HIPPOCRATIC MULTI-AGENT SYSTEMS

Ludivine Crépin
LHC - LIG - ENSMSE, 46, avenue Félix Viallet, 38031 Grenoble, France

Laurent Vercouter, Olivier Boissier
ENSMSE - Centre G2I, 158 cours Fauriel, 42023 Saint-Étienne, France

Yves Demazeau
CNRS - LIG, 46, avenue Félix Viallet, 38031 Grenoble, France

François Jacquenet
Université Jean Monnet - LHC - CNRS, 18 rue Benoit Lauras, 42000 Saint-Étienne, France

Keywords: Multi-agent system, privacy preservation, sentive information protection and management.

Abstract: The current evolution of Information Technology leads to the increase of automatic data processing over multiple information systems. In this context, the lack of user's control on their personal data leads to the crucial question of their privacy preservation. A typical example concerns the disclosure of confidential identity information, without the owner's agreement. This problem is stressed in multi-agent systems (MAS) where users delegate their control on their personal data to autonomous agents. Interaction being one of the main mechanism in MAS, sensitive information exchange and processing are a key issue with respect to privacy. In this article, we propose a model, "Hippocratic Multi-Agent System" (HiMAS), to tackle this problem. This model defines a set of principles bearing on an agency to preserve privacy. In order to illustrate our model, we have chosen the concrete application of decentralized calendars management.

1 INTRODUCTION

One of the main characteristics of multi-agent systems (Jennings and Wooldridge, 1999) is the interaction. This feature implies information communication and many sensitive information can often spread throughout the system without taking into account this sensitiveness. Spam is certainly a typical example: spammers get a user's email address without the user knowing how his mail has been disclosed. In this paper we focus on privacy in multi-agent systems.

(Deswarte and Melchor, 2006) define five kind of privacy enhancing technologies: IP and localization protection, anonymous access, authorization needed sensitive information and sensitive information protection. This article present a new model for sensitive data management and protection in multi-agent systems. Sensitive information can be in relation with the user (e.g. his identity) or with an agent (e.g. its strategy of the negotiation). In our work, we do not differentiate this two kind of information.

In this article, we propose a model we call "Hippocratic Multi-Agent System" (HiMAS) to tackle the privacy preservation problem in relation with sensitive information by using artificial agents.

2 SOME APPROACHES ON PRIVACY

This section focuses on various data-processing technologies in order to present the main privacy aspects.

2.1 Platform for Privacy Preferences

The Platform for Privacy Preferences (P3P) (W3C, 2002; Cranor, 2002) is an initiative of the W3C consortium that aims to develop a standard to make sensitive information management possible on both client and server sides. A user specifies his **preferences** to

Crépin L., Vercouter L., Boissier O., Demazeau Y. and Jacquenet F. (2008).
HIPPOCRATIC MULTI-AGENT SYSTEMS.
In *Proceedings of the Tenth International Conference on Enterprise Information Systems - AIDSS*, pages 301-307
DOI: 10.5220/0001702803010307

define the constraints that he wishes to impose on his personal data. The server which has to manage this data specifies a **policy** on the collected information.

This standard thus makes it possible to specify constraints on sensitive data management. Several critics of the P3P have focused on the impossibility for users to check if a server respects its engagement. Other standards are under development at that time in order to try to solve some of the drawbacks of the P3P.

2.2 Role-Based Access Control

Role-Based Access Control (RBAC) (Sandhu et al., 1996) has been designed in order to allow management and **dynamic data access control** in dynamic organizations and complex information systems.

A role is defined here as a set of access permissions and a set of users. To ensure a flexible and dynamic management of the data access, the RBAC uses sessions. Each session represents a mapping between a user and a subset of roles. Such a system allows to dynamically assign permissions to a user via a role.

This technology is only dedicated to accessing to the sensitive information after its collection. Even if the RBAC imposes more constraints on the use of sensitive data than the P3P, we can regret a lack of control on what happens to the data after it has been accessed.

2.3 Hippocratic Databases

The Hippocratic Databases model (Agrawal et al., 2002), including some principles of the P3P and RBAC, defines ten principles for privacy preservation.

The donor must know the **purpose specification** for each sensitive data stored in the database in relation with him. He had to give also his **consent** for each data collected. The **limited collection** implies a minimal data collection for the realization of the purposes. The **limited use** principle imposes on the database to use the collected information only for the specified purpose. Due to the **limited disclosure** principle, such a database had to communicate the stored information only for the purpose and only with the donor's consent. Information shall be only stored until the purpose realization (**limited retention**). The information **accuracy** had to be enforce. Such a database had to guarantee the **safety** of the stored information. A donor shall have access to all information about him with the **openness** principle. The last principle, **compliance**, implies that the donor can verify that all the above principle is respected.

These principles allow to preserve privacy by focusing on safety, storage and communication of sensitive data and also on the database operation.

2.4 Privacy and Peer-to-peer

(Damiani et al., 2004) proposed a decentralized privacy preserving approach for spam filtering with a structured peer-to-peer (P2P) architecture. E-mail servers share knowledge by a P2P network in order to reduce spam. It allows to detect more spam messages with collaborative and filtering techniques.

(Belenkiy et al., 2007) presented an other vision of privacy preserving P2P. This work focuses on the security (with cryptographic technologies) and the anonymity respect by using trusted entities and e-cash to make interaction tracing not possible.

2.5 Multi-agent Systems and Privacy

Privacy preservation is becoming an important field in the area of multi-agent systems. We propose here some different visions of privacy in this domain.

In distributed constraint problems, privacy is related to data protection by decreasing the sharing thus increasing the secret within the agency. Privacy preservation focuses here on hiding the agent current state. The main problem is that privacy preservation makes algorithms less efficient (Freuder et al., 2001). A first approach focuses on cryptology (Yokoo et al., 2005) but is too expensive. There are also many algorithms based on a random permutation for privacy preservation that aren't so much expensive (Nzouonta et al., 2004; Greenstadt et al., 2006).

In multi-agent systems, many works propose to preserve privacy using a guarantor agent in addition to a high level of security. This agent guarantees communicated sensitive information between two agents with respect to their desires (Bergenti, 2005), by using filter entity and profile (Cissée and Albayrak, 2007; Rezgui et al., 2002). The main advantage of these works is the use of only one trusted entity: the guarantor agent.

These different approaches allows us to define three step for the information management in order to preserve privacy. The first one is about the **storage** of sensitive information: security is required. The second one is the information **communication** which must be safe. Moreover the users must know what and how he gives his information. The last one implies that a guarantee about the **behavior** of this entity is required. This entity must describe the information manipulations and makes a commitment to respect the constraints fixed by the donor on the information.

3 FOUNDATIONS OF HiMAS

The previous section makes us focus on problems raised by the privacy preservation. Following this rapid study we propose a model we call HiMAS, that is Hippocratic Multi-Agent Systems. It defines the private sphere concept in order to model privacy. It is based on nine principles for privacy preservation inside multi-agent systems.

In order to illustrate the HiMAS model, we consider a decentralized calendars management application (Demazeau et al., 2006): each user is represented by an agent in charge of the scheduling of events (tasks or meetings). Timetables can be shared with other agents. When the sharing is not possible, a negotiation system is necessary to fix the meetings.

3.1 Private Sphere

From many researches, we define the dimensions of a private sphere as follows. The private sphere concerns information that an agent considers as sensitive. The **ownership rights** of the sensitive information are only assigned to the agent concerned by this information (Thomson, 1975). The private sphere is also **personal** (Demeulenaere, 2002; Baase, 2003), **personalizable** (the agent chooses what its private sphere contains) (Westin, 1967) and **context-dependent** (Palen and Dourish, 2003).

In order to introduce the private sphere inside multi-agent systems we need to specify two tasks. The first one is the **private sphere management** considering only one agent. The second task concerns the **private sphere protection** that required the agency.

3.2 Nine Principles for Hippocratic MAS

Our model, HiMAS, is inspired by the model proposed by (Agrawal et al., 2002). Indeed it defines all the fundamental principles for privacy preservation: storage, communication and information becoming.

To represent the possible positions of an agent, we define three roles (see figure 1). The **consumer** characterizes the agent which asks for sensitive information and which will use it. The **provider** characterizes the agent which discloses a sensitive information. The **subject** describes the agent subject of the sensitive information.

According to the HiMAS model an hippocratic MAS must respect the following nine principles.

Purpose Specification. The *provider* must know what are the objectives of the sensitive data transaction. In this way it can evaluate the communication consequences. For example, the *consumer* asks the *provider*'s plans in order to fix a meeting.

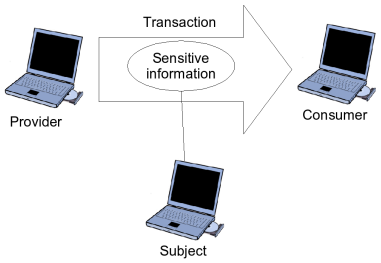

Figure 1: Agents roles in a privacy preserving environment.

Consent. Each sensitive data transaction requires the *provider*'s consent. For example, when a *consumer* asks a *provider* for its planning at a precise date, the *provider* has to give its consent. If the *provider* and the *subject* aren't the same agent, the *subject*'s consent is also needed.

Limited Collection. The *consumer* commits to cutting down to a minimal number of data for realizing its objectives. For example, when a *consumer* asks a *provider* for its planning in order to fix a new meeting, the *consumer* needs only to know its free slots and occupied slots. It must not try to obtain more information like meetings subject.

Limited Use. The *consumer* commits to using sensitive *provider*'s information only to satisfy its objectives. In the previous example, the *consumer* must only use the required planning to fix a new meeting between the *provider* and itself. The *consumer* can't transmit this sensitive information to another agent if it isn't defined in its objectives.

Limited Disclosure. The *consumer* commits to disclosing a sensitive information only to reach its objectives. Moreover it must disclose it the least time as possible and to the least agents as possible. To fix a meeting, for example, the *consumer* doesn't need to disclose the whole *provider*'s planning.

Limited Retention. The *consumer* commits to retain a sensitive information only during the minimal amount of time for the realization of its objectives. For example, while deciding a new meeting, the *consumer* commits to deleting *consumer*'s planning once the appointment has been set or after the meeting date.

Safety. The system must guarantee sensitive information safety during storage and transactions.

Openness. The transmitted sensitive information must remain accessible to the *subject* and/or the *provider* during the retention time. For example, if the *provider*'s plans change, it must have the choice to update the planning known by the *consumer* so that the appointment check is based on true information.

Compliance. Each *provider* shall be able to check the respect of previous principles.

Notice that the accuracy principle proposed for Hippocratic Databases isn't kept for HiMAS. Indeed, we consider that an agent may lie to protect its private sphere. For example, the act of denying access to information at a malicious agent can often reveal sensitive information. When a *provider* marks a *consumer* as malicious, there are two possibilities. The first one is that the *provider* doesn't reply to it. The second one is that the *provider* lies about the sensitive information in order to protect it. Using a lying allows the *provider* not to warn the malicious *consumer*. This solution also allows to discredit this *consumer* by the agency when it will disclose the false information.

4 PRIVATE SPHERE MANAGEMENT IN HiMAS

Given the foundations of the HiMAS model presented above, let's turn now to the description of requirements for integrating these principles in the private sphere management inside multi-agent systems. We describe first the private sphere representation.

We define a private sphere *PS* as a quadruplet:

$$PS = < Elements, Authorizations, Rules, Norms >$$

where $Elements$ is a set of elements, *Authorizations* is a set of authorizations, *Rules* is a set of rules, and *Norms* is a set of norms.

4.1 Private Sphere Elements

A private sphere element, *element*, is a sextuplet:

$$element = < id, information, Owners, context, Subjects, References >$$

where *id* is the element identifier, $information$ is the sensitive information to protect, *Owners* is a finite set of owners known by the agent, *context* is the information context, $Subjects$ is a finite set of subjects, and $References$ is a finite set of references on elements concerning sensitive information which can be found using $information$.

Given the identifier of the element $e128$ concerning a sensitive information *meeting* representing the meeting in agent *alice*'s calendar.

$$< e128, meeting, \{alice, charlie\}, professional, \{alice, charlie\}, \{monday - 10AM, \{alice, charlie\}\} >$$

This meeting takes place at a precise date, $monday - 10AM$. The agents *alice* and *charlie* are the participants and only these agents are aware of this sensitive information. These agents are also concerned by this information, so they are also the subjects.

An information can refer to other information, e.g. *meeting* refers to $monday - 10AM$ and $\{alice, charlie\}$.

In order for an agent to reason about sensitive information disclosure, an element is associated with a set of owners, e.g. $\{alice, charlie\}$ for $e128$.

An element is also in relation with a given context, e.g. the context of $e128$ is $professional$. This context allows the agent to reason about sensitive information management with the help of rules.

4.2 Authorizations Attached to a Private Sphere Element

Authorizations of private sphere element allow an agent to define operations that it authorizes on sensitive information. These authorizations concern the use, the deletion, the disclosure, the modification and the alteration of the information. Given the element $e128$ previously defined for example, we may define:

$use(e128)$: The sensitive information contained in $e128$ can be used by the agent.

$delete(e128)$: This authorization allows an agent to delete element $e128$ from its private sphere.

$disclose(e128)$: The agent knowing element $e128$ can disclose *meeting*.

$change(e128)$: This authorization allows an agent to modify $e128$.

$lie(e128)$: The agent can lie about the $e128$'s sensitive information in order to protect it.

4.3 Private Sphere Rules

Because the private sphere is defined in a certain context, it dynamically evolves over time and because it is intrinsically personal, we attach a set of rules to it, allowing to specify the activation conditions on the authorizations described above.

We define a private sphere rules as:

$$authorization \leftarrow condition$$

condition represents the activation condition of the authorization. It depends on application context and refers to agent's belief.

For example given the current context $currentcontext$ and the context of the element $e128$ $professional$, we can specify the rule:

$$use(e128) \leftarrow (currentcontext \in professional)$$

This rule allows an agent to use $e128$ if the $currentcontext$ belongs to in the context of $e128$.

Private sphere rules allow an agent to define the internal dynamic of the sphere according to its desires. This dynamic is unique to each agent because of the private sphere personalization.

These rules are dynamic: they are influenced by the various produced events. For example if an information of its private sphere is known by all the other agents, an agent can decide to remove it of its sphere.

4.4 Private Sphere Norms

We define norms like private sphere rules. However norms are known by the agency as opposed to rules and must be respected by each agent.

$$norm \leftarrow condition$$

Private sphere delimitation can be influenced by the *society's rules*, even if everyone chooses his behavior with respect to these rules (Demeulenaere, 2002). Some dynamic norms can be imposed to agency on private sphere element but an agent can violate this. The consequences deserve some studies in order to define the various impacts on the agency.

4.5 Global Organization of the Private Sphere

An agent personalizes its private sphere by defining the set of its elements (so the set of sensitive information) and the set of rules which is in relation with authorizations about private sphere element.

At the agent reasoning level, norms may infer new private sphere rules. Afterward these rules infer new authorizations for elements concerned by norms.

5 PRIVATE SPHERE PROTECTION IN HiMAS

Let's pursue our investigations on requirements imposed by the HiMAS model on the private sphere protection in multi-agent systems.

This step needs to focus on sensitive information communication between a *provider* and a *consumer*. We define this kind of communication as a **data transaction**.

In an agency, private sphere protection must be provided by the following means:

1. *before* the data transaction: an agent must determine risks to disclose a sensitive information,
2. *during* the data transaction: an agreement must be stated between the *provider* and the *consumer* on their behavior with respect to the sensitive data,
3. *after* the data transaction: the agency must guarantee the *consumer* behavior.

5.1 Protection before a Data Transaction

The first principle which has to be guaranteed before data transaction is the safety of sensitive information during its storage. Indeed, a HiMAS must impose a non intrusion rule into the agents' private sphere.

An agent must not disclose sensitive information without evaluating the possible impacts. An agent must have a representation of the context in addition to the private sphere representation.

HiMAS agents must also be able to pass judgment on other agents in order to determine the risk incurred by disclosing an element of their private sphere. Such risk-taking can be evaluated using for instance trust and the social trust network of the agent that received the data, like in (Damiani et al., 2004).

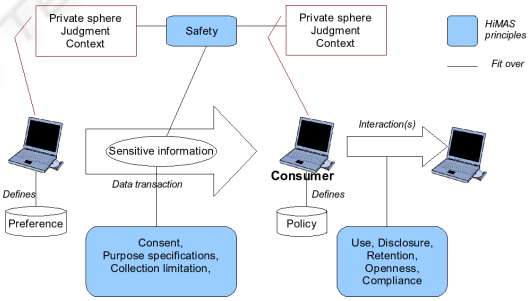

Figure 2: Global view of the HiMAS model.

5.2 Protection during a Data Transaction

The first principle to guarantee during a data transaction concerns communication safety. A data transaction needs a secure medium of communication, stopping from every intrusion in the transaction.

Figure 2 represents a data transaction between a *provider* and a *consumer*. When a *consumer* asks for information to a *provider*, they have already evaluated risk-taking for the data transaction and have taken

context into account in order to estimate if this transaction is possible or not.

In this part we describe all the elements needed to protect the private sphere during a data transaction.

Policy & Preference. We define a policy and a preference, *policy* and *preference*, as a quadruplet:

$$policy = < Objectives, date, Agents, format >$$

$$preference = < Objectives, date, Agents, format >$$

where *Objectives* is a non empty finite set of objectives[1], *date* a retention date, *Agents* a finite set of agents which represent the possible disclosure list and *format* the information format (in order to clarify all the information details).

Let's for example consider two agents *bob* and *alice*. *bob* requires *alice*'s sensitive information *meeting*. So *bob* is the *consumer* and needs this information in order to be present at this meeting and will not disclose it. In fact, it needs all the details (date, place, subject, participants...). $policy_{bob}$ is therefore:

$$< \{bepresent\},\ date_{meeting},$$

$$\emptyset,\ \{date_{meeting}, Participants, place\} >$$

$preference_{alice}$ agrees with *bob*'policy because this policy does not contradict its private sphere rules and the society norms:

$$< \{bepresent, discloseTeam\},\ date_{meeting},$$

$$Coworkers,\ \{date_{meeting}, Participants, place\} >$$

A *consumer* defines its policy with respect to sensitive information which is required by a *provider*. This special information defines the *consumer*'s behavior with respect to the sensitive information.

On the provider's side, a preference is defined in the same way that a policy in order to allow the *consumer* and the *provider* to look for an agreement about their behavior with respect to the required sensitive information. A preference is defined using authorizations bearing on private sphere elements, which refer to the sensitive information required. A preference is also based on the different representations that agents build about the agency and on the different reasonings that it evaluates before the data transaction.

A first advantage can be put forward with this model: during the data transaction the agreement between a policy and a preference allows to represent the *provider* consent or disagreement.

[1]The objectives are close to the concept of goal, like for example in BDI model (Bratman, 1987).

Data Transaction. We define a data transaction as:

$$transaction =$$

$$< information, policy, preference, consent >$$

where *information* is the sensitive information, *policy* the *consumer* policy, *preference* the *provider* preference and *consent* a boolean representing the agreement (or not) between the *consumer* and the *provider*.

Let's considerer $policy_{bob}$ and $preference_{alice}$ as defined previously. The data transaction between *bob* and *alice* concerning *meeting* is also:

$$< meeting, policy_{bob}, preference_{alice}, true >$$

The *consent* is true because the policy and the preference match together.

Once the information is received, the *consumer* inserts a new element about this information into its private sphere. Moreover it deduces from its policy a set of authorizations in order to manage this element.

For example, once *bob* has received *meeting*, it inserts into its private sphere a new element $e578$ about it. The agent *alice* modifies also the element $e128$ in its private sphere by adding the agent *bob* to the participants and to the owners of this sensitive information.

This formalization of data transaction allows to check agents behavior on the following principles: the ***provider*'s consent** is checked by using the agreement between the preference and the policy, the **purpose specifications** by using the *consumer*'s policy and the **collection limitation** from the *consumer* using the *provider*'s knowledge about its policy.

5.3 Protection after a Data Transaction

After a data transaction, several mechanisms must be introduced in order to preserve privacy. These mechanisms concern five HiMAS principles: the **limited use, disclosure** and **retention** of sensitive information by the *consumer*, the sensitive information **transparency** by the *consumer* and the **compliance** about the respect of all the HiMAS principles.

These principles allow the detection of malicious agents behavior in relation with the private sphere. An agent is malicious if it infringes at least one of these five principles according to its preference or its policy.

Figure 2 gives a global view of the HiMAS model with a representation of the different required protection levels for privacy preservation. Each HiMAS principle is attached to a step of a data transaction.

6 CONCLUSIONS AND PERSPECTIVES

In this paper we have proposed a model we called hippocratic multi-agent systems, HiMAS. Such a system has to respect nine principles to preserve privacy.

HiMAS agents must be able to represent their private sphere by storing its characteristics and by managing it by itself. After a data transaction, the agency must play a role in privacy preservation.

By adapting nine of the principles of (Agrawal et al., 2002) to multi-agent systems, a HiMAS can enable to guarantee the sensitive data communication and give a vision of data becoming, contrary to classic agent models or in (W3C, 2002). Our model also takes advantage of the multi-agent systems characteristics like decentralization, autonomy and openness in an application context such as the Web.

The HiMAS model opens a lot of research and development perspectives. On a theoretical standpoint the formalization of many features of a HiMAS can be studied with interest. On a more practical level, the design of various components of a HiMAS is also an interesting issue. In fact, we hope this model will be a useful basic block for the research community.

ACKNOWLEDGEMENTS

This work was supported by Web Intelligence project, financed by the ISLE cluster of Rhône-Alpes region.

REFERENCES

Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y. (2002). Hippocratic databases. In *Proceedings of the International Conference Very Large Data Bases*.

Baase, S. (2003). *A Gift of Fire: Social, Legal, and Ethical Issues in Computing*. Prentice-Hall.

Belenkiy, M., Chase, M., Erway, C., Jannotti, J., Kupcu, A., Lysyanskaya, A., and Rachlin, E. (2007). Making P2P accountable without losing privacy. In *Proceedings of Workshop on Privacy in the Electronic Society*.

Bergenti, F. (2005). Secure, trusted and privacy-aware interactions in large-scale multiagent systems. In *Proceedings of the Workshop From Objects to Agents*.

Bratman, M. E. (1987). Intention, plans, and practical reason. O'Reilly, Harvard University Press: Cambridge,MA.

Cissée, R. and Albayrak, S. (2007). Experimental analysis of privacy loss in dcop algorithms. In *Proceedings of 6th International Joint Conference on Autonomous Agents and Multiagent Systems*.

Cranor, L. F. (2002). *Web Privacy with P3P*. O'Reilly.

Damiani, E., di Vimercati, S. D. C., Paraboschi, S., and Samarati, P. (2004). P2P-based collaborative spam detection and filtering. In *Proceedings of 4th International Conference on Peer-to-Peer Computing*.

Demazeau, Y., Melaye, D., and Verrons, M.-H. (2006). A decentralized calendar system featuring sharing, trusting and negotiating. In *Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*.

Demeulenaere, P. (2002). Difficulties of private life characterization from a sociologic point of view. In *Privacy in Information Society*, volume 11.

Deswarte, Y. and Melchor, C. A. (2006). Current and future privacy enhancing technologies for the internet. *Annales des Télécommunications*, 61:399–417.

Freuder, E. C., Minca, M., and Wallace, R. J. (2001). Privacy/efficiency tradeoffs in distributed meeting scheduling by constraint-based agents. In *Proceedings of 7th International Joint Conference on Artificial Intelligence Workshop on Distributed Constraint Reasoning*.

Greenstadt, R., Pearce, J. P., Bowring, E., and Tambe, M. (2006). Experimental analysis of privacy loss in dcop algorithms. In *Proceedings of 5th International Joint Conference on Autonomous Agents and Multiagent Systems*. ACM.

Jennings, N. R. and Wooldridge, M. (1999). Agent technology: Foundations, applications and markets. *Journal of Artificial Societies and Social Simulation*, 2(4).

Nzouonta, J., Silaghi, M.-C., and Yokoo, M. (2004). Secure computation for combinatorial auctions and market exchanges. In *Proceedings of 3rd International Joint Conference on Autonomous Agents and Multiagent Systems*.

Palen, L. and Dourish, P. (2003). Unpacking "privacy" for a networked world. In *Proceedings of the 2003 Conference on Human Factors in Computing Systems*. ACM.

Rezgui, A., Ouzzani, M., Bouguettaya, A., and Medjahed, B. (2002). Preserving privacy in web services. In *Proceedings of the Workshop on Web Information and Data Management*.

Sandhu, R. S., Coyne, E. J., Feinstein, H. L., and Youman, C. E. (1996). Role-based access control models. *IEEE Computer*, 29(2).

Thomson, J. J. (1975). The right of privacy. Philosophy and Public Affairs 4: 295-314.

W3C (2002). Plateform for privacy preferences, http://www.w3.org/p3p/.

Westin, A. F. (1967). Special report: legal safeguards to insure privacy in a computer society. *Commun. ACM*.

Yokoo, M., Suzuki, K., and Hirayama, K. (2005). Secure distributed constraint satisfaction: reaching agreement without revealing private information. *Artificial Intelligence*, 161(1-2).