Towards HiMAS: A model for privacy preserving multi-agent systems

Ludivine Crépin
LHC - LIG - ENSMSE
46, avenue Félix Viallet
F-38031 Grenoble Cédex
Ludivine.Crepin@imag.fr

Laurent Vercouter
ENSMSE - Centre G2I
158 cours Fauriel
F-42023 Saint-Étienne Cédex 2
Laurent.Vercouter@emse.fr

Yves Demazeau
CNRS - LIG
46, avenue Félix Viallet
F-38031 Grenoble Cédex
Yves.Demazeau@imag.fr

François Jacquenet
Université Jean Monnet - LHC (UMR CNRS 5516)
18 rue Benoît Lauras
F-42000 Saint-Étienne
Francois.Jacquenet@univ-st-etienne.fr

Olivier Boissier
ENSMSE - Centre G2I
158 cours Fauriel
F-42023 Saint-Étienne Cédex 2
Olivier.Boissier@emse.fr

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 personal data control 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 the users’ privacy but also the agents’ privacy.

The next section briefly presents various visions of the privacy concept. In section 3 we present the fundamental principles of the HiMAS model. Section 4 allows us to specify the mechanisms of privacy management inside HiMAS while section 5 focuses on its protection. To conclude we propose some hints on the work that has to be done in order to implement such a model.

II. SOME APPROACHES ON PRIVACY

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

A. Platform for Privacy Preferences

The Platform for Privacy Preferences (P3P) [23], [8] 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 define the constraints that he wishes to impose on his personal data. The server which has to manage this data specifies a policy (objectives, collection, use and retention).

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 commitment [21]. Other standards are under development at that time in order to try to solve some of the drawbacks of the P3P.

B. Role-Based Access Control

Role-Based Access Control (RBAC) [20] 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.

### C. Hippocratic Databases

The Hippocratic Databases model [1], including some principles of the P3P and RBAC, strengthens RBAC in the field of databases. Citing Agrawal et al. the ten principles of Hippocratic Databases are listed below.

**Purpose Specification:** For personal information stored in the database, the purposes for which the information has been collected shall be associated with that information.

**Consent:** The purposes associated with personal information shall have consent of the donor of the personal information.

**Limited Collection:** The personal information collected shall be limited to the minimum necessary for accomplishing the specified purposes.

**Limited Use:** The database shall run only those queries that are consistent with the purposes for which the information has been collected.

**Limited Disclosure:** The personal information stored in the database shall not be communicated outside the database for purposes other than those for which there is consent from the donor of the information.

**Limited Retention:** Personal information shall be retained only as long as necessary for the fulfillment of the purposes for which it has been collected.

**Accuracy:** Personal information stored in the database shall be accurate and up-to-date.

**Safety:** Personal information shall be protected by security safeguards against theft and other misappropriations.

**Openness:** A donor shall be able to access all information about the donor stored in the database.

**Compliance:** A donor shall be able to verify compliance with the above principles. Similarly, the database shall be able to address a challenge concerning compliance.

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

The Hippocratic Databases model give a vision more complete than the two previously presented approaches. Such a database manages the storage, the communication and the becoming of the sensitive information. This work provides many aspects of privacy preservation but we can notice that one facet isn’t handled: the prevention of the users against the malicious entities.

Some limits of this model are presented in [16]. This article proposed an integration of the Hippocratic Databases in secure Tropos in order to complete the privacy preservation in the field of databases. The main contribution of this approach is the introduction of trust for the relation between the different databases. Trust is used to increase the safety level between the databases. However the user/database relation is not studied in this work while Agrawal et al. define it as essential for the privacy preservation.

### D. Privacy and peer-to-peer

[9] 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 the level of spam. It allows to detect more spam messages with collaborative and filtering techniques.

[3] presented another vision of privacy preserving P2P. This work focuses on the security (with cryptographic technologies) and the anonymity respect. They use two trusted entities and e-cash to ensure anonymity in order to make interaction tracing not possible.

This two approaches give two complementary visions of privacy preservation. [9] focus on the prevention of the users whereas [3] studies the behavior of the entity which receive the sensitive information with cryptographic techniques. We can notice here that this two contributions must be applied together for a global privacy preservation.

### E. 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. A typical example is the meeting negotiation. In this situation, privacy preservation focuses on hiding the agent current state. The main problem is that privacy preservation makes algorithms less efficient [12]. A first approach focuses on cryptography [26] but is too expensive. There are also many algorithms based on a random permutation for privacy preservation (keeping the current agent’s state secret) that aren’t so much expensive [17], [13].

These approaches focus on the discovery of sensitive information by the manipulations of other data. This aspect is essential for the privacy preservation but it is not enough to assure the privacy. Indeed, the information collection or the future use of this information for example is not handled here.

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 [5]. [7] and [19] also use the same technology, the first one using a filter entity and the later one using users profile in Web Services. The main advantage of these works is the use of only one trusted entity: the guarantor agent.

With these two approaches we are always confronted with the same problem encountered in P3P: no verification is made after the communication of sensitive information. Moreover any prevention is made in order to detect a malicious agent.
F. Discussion of these privacy approaches

These different visions on the privacy 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 [20], [1].

The second one is the information communication which must be safe [20], [1], [3], [5]. Moreover the users must know what and how he gives his information [23], [1].

The last one is about the entity who receives the sensitive information. This entity must describe the information manipulations and makes a commitment to respect the constraints fixed by the donor on the information [23], [1]. This step implies that a guarantee about the behavior of this entity is required.

III. 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 [10]: each user is represented by an agent in charge of the scheduling of events, either tasks or meetings. Timetables can be shared with other agents. When the agents do not share their timetables, a negotiation system is necessary to fix the meetings.

A. Private sphere

In multi-agent systems, we consider that a private sphere refers to three different types of entities. It can be related to a user, an agent or to an agent representing a user. To simplify our HiMAS model, we call private sphere any sphere of those three entities.

From many researches in social science, 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 [22]. The private sphere is also personal [11], [2], personalizable (the agent chooses what its private sphere contains) [25], [24], [15] and context-dependent [4], [18].

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 is required by the agency. How can agents verify that their private sphere isn’t violated after a transaction of sensitive information?

B. Nine principles for Hippocratic MAS

Our model, Hippocratic Multi-Agent Systems, is inspired by the Hippocratic Databases model proposed in [1]. Indeed it defines all the fundamental principles for privacy preservation: safety and constraints on communication, storage and information becoming.

To represent the possible positions of an agent with respect to the private sphere, we define three roles (see figure 1). The consumer role characterizes the agent which asks for sensitive information and uses it. The provider characterizes the agent which discloses a sensitive information. The last role, 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 respect the objectives of the sensitive data transaction. In this way it can evaluate the communication consequences, according to its desires. For example, the consumer asks the provider’s plans in order to fix a meeting.

- **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 motivating the collection, previously specified. 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 or participants.

- **Limited use**: The consumer commits to using sensitive provider’s information only to satisfy the objectives that it has specified and nothing more. 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 agent 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 lie allows the provider not to warn the consumer about the fact of being judged as malicious. This solution also allows to discredit this consumer by the other agents when it will disclose the false information.

IV. 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:

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

- **Elements**: a set of elements.
- **Authorizations**: a set of authorizations.
- **Rules**: a set of rules.
- **Norms**: a set of norms.

We define these sets farther in this section.

A. Private sphere elements

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

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

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

For example, let’s consider the following sextuplet representing the private sphere of an element whose identifier is $e379$. It concerns the sensitive information $meeting$ representing the meeting in agent alice’s calendar.

$$< e379, meeting, \{alice, charlie\}, professional, \{alice, charlie\}, \{monday – 10AM, \{alice, charlie\}\} >$$

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: an information can give some details, e.g. meeting refers to the details $monday – 10AM$ and $\{alice, charlie\}$. When an agent give a sensitive datum, there are two possibilities. The first one is that an agent can give all the details of the information and the References set is not used. The second one is that an agent would give the sensitive information but not completely. For example, alice can give the date of meeting but not the participants $\{alice, charlie\}$. In this case, the References set is used to know what are the other informations given with the complete information.

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 $e379$.

An element is also in relation with a given context, e.g. the context of $e379$ is professional. This context allows the agent to reason about sensitive information management with the help of rules.

B. 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. We have defined these five kind of authorizations in relation with all the possibilities of data manipulations concerning privacy in our work: to use, to delete, to disclose, to change (in order to update the information for example) and to lie about a sensitive data.

Given the element $e379$ previously defined for example, we may define:

- **use($e379$)**: The sensitive information contained in $e379$ can be used by the agent.
- **delete($e379$)**: This authorization allows an agent to delete element $e379$ from its private sphere.
- **disclose($e379$)**: The agent knowing element $e379$ can disclose meeting.
- **change($e379$)**: This authorization allows an agent to modify $e379$.
- **lie($e379$)**: The agent can lie about the $e379$’s sensitive information in order to protect it.
C. 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 rule as:

\[ \text{authorization} \leftarrow \text{condition} \]

The condition \( \text{condition} \) depends on application context and refers to the agent’s belief.

For example given the current context \( \text{currentcontext} \) and the context of the element \( e_{379} \) \( \text{professional} \), we can specify the rule:

\[ \text{use}(e_{379}) \leftarrow (\text{currentcontext} \in \text{professional}) \]

This rule allows an agent to use \( e_{379} \) if the \( \text{currentcontext} \) belongs to the context of \( e_{379} \).

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 from its sphere.

D. Private sphere norms

We define private sphere norms in the same way that private sphere rules. However norms are known by the agency as opposed to rules which are personal for each agent. Moreover each agent must respect these norms.

\[ \text{norm} \leftarrow \text{condition} \]

Private sphere delimitation can be influenced by the society’s rules of good behavior, even if everyone chooses his behavior with respect to these rules [11]. Indeed some norms of the society can be imposed to agents on what private sphere contain but an agent can violate one of these norms. Consequences of violation deserve some studies in order to define the various impacts on the agency.

Moreover these norms can evolve over time: some agents behavior can cancel a norm or establish a new norm. For example a norm which forbids professional meeting after 6 PM is canceled if every agent fixes this kind of meeting after 6 PM.

E. Global organization of the private sphere

An agent personalizes its private sphere by defining the set of its elements: the set of information which is sensitive according to it. It also personalizes the set of rules which is in relation with authorizations about private sphere elements (figure 2).

At the agent reasoning level, norms may infer new private sphere rules. For example, if a norm can forbid a professional meeting after 6 PM then every agent create a rule about this.

Afterward these rules infer new authorizations for elements concerned by norms. For example, if any agent has a professional meeting before 8 AM then a norm can be created for this.

V. 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.

Private sphere protection in a HiMAS 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 agent society, 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 agent society must guarantee the consumer behavior on the transmitted sensitive information.

The following section presents these three steps.

A. Protection before a data transaction

The first principle that 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 [9] or in [16].

B. Protection during a data transaction

The first principle that has to guarantee during a data transaction concerns communication safety. A data transaction needs a secure medium of communication, preventing from any intrusion in the transaction.

Figure 3 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 the 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. We start with the policy and the preference. These concepts are the first step in our model for privacy preservation because the agents reason from policies and preferences in order to protect their private sphere during and after a data transaction.

1) Policy & Preference: We define a policy and a preference, policy and preference, as a quadruplet:

\[
policy = \langle \text{Objectives}, date, \text{Agents}, format \rangle
\]

\[
\text{preference} = \langle \text{Objectives}, date, \text{Agents}, format \rangle
\]

- **Objectives**: a non empty finite set of objectives\(^1\).
- **date**: a retention date
- **Agents**: a finite set of agents which represent the possible disclosure list.
- **format**: the information format (in order to clarify the information).

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...). \(\text{policy}_{bob}\) is therefore:

\[
< \{\text{bepresent}\}, date_{\text{meeting}}, \emptyset, \{date_{\text{meeting}}, \text{Participants}, place\} >
\]

\(\text{preference}_{alice}\) agrees with bob’s policy because this policy does not contradict its private sphere rules and the society norms:

\[
< \{\text{bepresent}, \text{discloseTeam}\}, date_{\text{meeting}}, \text{Coworkers}, \{date_{\text{meeting}}, \text{Participants}, place\} >
\]

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 data 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.

2) Data transaction: We define a data transaction as:

\[
\text{transaction} = \langle \text{information}, \text{policy}, \text{preference}, \text{consent} >
\]

- **information**: a sensitive information.
- **policy**: a consumer policy.
- **preference**: a provider preference.
- **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:

\[
< \text{meeting}, \text{policy}_{bob}, \text{preference}_{alice}, \text{true} >
\]

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 e379 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 using the agreement between the preference and the policy,
- the purpose specifications using the consumer’s policy,
- the collection limitation from the consumer using the provider’s knowledge about its policy.

C. Protection after a data transaction

After a data transaction, several mechanisms must be introduced in order to ensure privacy preservation. 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 with respect to the provider.
- 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. A malicious agent is an agent which infringes at least one of these five principles according to its preference or its policy.

Figure 3 gives a global view of the HiMAS model with a short representation of the different required protection levels for privacy preservation. Each principle of a HiMAS is attached to one of the three steps of a data transaction.

VI. CONCLUSIONS AND PERSPECTIVES

In this paper we have proposed a model we called hippocratic multi-agent systems or HiMAS. A system based on our model 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. However, once a sensitive information is communicated, the agency must play a role in order to preserve privacy, based on the agents’ policies and preferences.

By adapting nine of the principles of [1] to multi-agent systems, the HiMAS model can enable to guarantee the sensitive data communication and provide also a vision of data becoming, contrary to classic agent models or the P3P [23]. Our model also takes advantage of the multi-agent systems characteristics like for example the decentralization, the autonomy and the 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.

ACKNOWLEDGMENT

This work is supported by Web Intelligence project, financed by the ISLE cluster of Rhône-Alpes region. We thank France Telecom R&D for supporting the research related to trust mentioned in this paper.

REFERENCES


