Forging Trust and Privacy with User Modeling Frameworks: An Ontological Analysis

Federica Cena
Department of Computer Science
University of Torino
Torino, Italy
cena@di.unito.it

Nima Dokoohaki
Software and Computer Systems (SCS)
Royal Institute of Technology (KTH)
Stockholm, Sweden
nimad@kth.se

Mihhail Matskin
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
misha@idi.ntnu.no

Abstract—With the ever increasing importance of social networking sites and services, socially intelligent agents who are responsible for gathering, managing and maintaining knowledge surrounding individual users are of increasing interest to both computing research communities as well as industries. For these agents to be able to fully capture and manage the knowledge about a user’s interaction with these social sites and services, a social user model needs to be introduced. A social user model is defined as a generic user model (model capable of capturing generic information related to a user), plus social dimensions of users (models capturing social aspects of user such as activities and social contexts). While existing models capture a proportion of such information, they fail to model and present trust and privacy. To this end, in this paper, we introduce an ontological model of social user, composed by a generic user model component, which imports existing well-known user model structures and captures the basic concepts regarding the user; and a social model, which contains social dimensions. In this model, trust, reputation and privacy become the pivotal concepts gluing the whole ontological knowledge models together.

Keywords—trust and reputation; privacy; user modeling; ontologies; semantic adaptive social web

I. INTRODUCTION

Social intelligence according to the original definition of Edward Thorndike is “the ability to understand and manage men and women, [..], to act wisely in human relations” [1]. Some authors have restricted the definition to deal only with knowledge of social situations, where social intelligence is an aggregated measure of social awareness, social progressive-ness and interests for new experiences. With the advent of social web, users have the possibility to exploit their social intelligence also in virtual environment, by using available social networking sites and services to maintain contact with other people as well as for sharing contents and experiences. We define “socially intelligent agents” as software agents which are responsible for gathering, managing and maintaining knowledge surrounding individual users in the social web. With the ever increasing significance of social networking sites and services, socially intelligent agents are of increasing value to both computing research communities as well as industries. For these agents to be able to fully capture and manage the knowledge relating to a user’s interaction with these social sites and services, a social user model needs to be defined and introduced.

A "social user model” is defined as a generic user model (with generic information about a user [2]), plus social dimensions of users (with social aspects of user, such as social activities, relationships with other users, groups they belong to and social contexts). While existing models (see Section II) capture a portion of such information, they fail to model ones of the most important dimensions of social connectivity: privacy, trust and reputation. To this end, in this paper, we introduce an ontological model of social user, composed by a generic user model component, which imports existing well-known user model structures and captures the basic concepts regarding the user; and a social model, which contains social dimensions. In this model, trust, reputation and privacy become the pivotal concepts gluing the whole ontological knowledge models together. We adopt the definition of ”privacy” as defined by Westin [3] as “the right of an individual to determine the amount of information available to other”. Privacy is particularly relevant in adaptive systems, since they gather a lot of personal information to provide adaptive services, and in social web, where users share a lot of data to other people. With respect to trust, we consider Golbeck’s definition [4]. According to this point of view, trust between two individual exists if the truster executes an action upon understanding that trustee’s actions in future will lead to a good outcome or utility for truster. Since our view of trust is reputation-based, it allows us to profile behavior of two individuals in a single relationship. To be able to port such profile across several applications, we have proposed to model reputation separately to profile the behavior or performance of an individual in several contexts.

Our ultimate goal is to propose a model that i) can be used as a reference to model users in the social web context, ii) can be used directly by the adaptive applications, for example, using some mechanism that, given a user, is able to populate such model on the fly according to the user information available on the Web. In this paper, however, we focused on modeling privacy and reputation/trust in social context, and thus in particular we described our models for such concepts. The paper is structured as follows. In Section II, we present existing approaches for modeling users in social web systems, focusing on how they deal with privacy, trust and reputation. Then, in Section III, we describe our application scenario. Sec-
tion IV briefly describes our framework and all the components involved: user data, domain, context, actions, privacy and trust. Section V focuses on the privacy model, while Section VI focuses on trust and reputation model. Finally, Section VII concludes the paper presenting possible future directions from this work.

II. USER MODELING ON SOCIAL WEB: STATE OF THE ART

In the user modeling field, there were several attempts to define a generic user model which contains the definition of user features and of his/her physical and social context, expressed with semantic web language and made available for all user-adaptive systems via Internet. In fact, a commonly accepted top level ontology for user and context models is of great importance for the user modeling and context research community. The major advantage is the simplification of using and exchanging user model and context data between different user-adaptive systems. The most known (and adopted) models are the General User Model Ontology (GUMO) [5], the Unified User Context Model (UUCM) [6], and Friend of A Friends (FOAF) [7]. GUMO includes basic user dimensions, such as demographic data, user knowledge, emotional state and personality aspects, user skills, capabilities, user interests, preferences, user goals and plans, etc. Moreover, GUMO also models the environment by representing data like location, time, device, etc. However, the current version lacks of modeling of social data, even if the authors started to work on it [8]. UUCM models several features of the user and his/her situation: cognitive characteristics (area of interest, competence, preference), usage data (current task, task role, task history), social data (relationships the user is involved in), environment data (device, current time, language, location). FOAF focuses more on social data than on user and usage data, since it mainly aims at describing the links between people and the things they create and do over the web. FOAF is weak in defining other user features, such as interests and preferences, knowledge and expertise. Only interests are represented, by means of the "interest" property, which represents an interest of a user through indicating a document whose topic(s) is of interest for him/her. Describing interests in full is a complex undertaking: FOAF provides no support for characterizing levels of interest.

A recent attempt to model users in the social web has been done by the Grapple project [9]. Within this project, the Grapple User Model Framework (GUMF) is defined for storing, retrieving and sharing information about users between components of the framework. In the framework, the Grapple User Modeling Ontology[10] is proposed, in order to describe all the possible statements about a user, and concepts like creator of the statement, rating of the statement, temporal and spatial dimensions. Most of such existing UM frameworks fail to capture and present privacy policies as well as user’s trust statements. GUMO simply has the attribute gumo:privacy which defines the default privacy status for each class of user dimensions. UUCM and FOAF do not explicitly model privacy. In GUMF privacy is modeled only with a property (hasPrivacyPreference) which expresses the level of privacy concerns of the users. However, privacy in user modeling is a crucial, multidimensional and complex aspect, that cannot be expressed by means of a single property. Personalized interaction and user modeling bear significant implications on privacy, due to the fact that personal information about user needs to be collected to perform personalization [11]. Moreover, Social Web context is particularly challenging for privacy, since social applications gather a lot of data about the user and his/her activities. Thus, the concept of privacy should be decomposed in several dimensions. A first theoretical attempt to define all the privacy dimensions involved in the user modeling process has been made by the Unified Model for Privacy Preferences [12], a formal model which defines the main categories of information in social web context. However, to the authors’ knowledge, there are no attempt to integrate such privacy model in a global user model.

At the same time, little attention has been paid to effective incorporation of trust and reputation into user models. Among adaptive Web applications, recommender systems have been quite successful in utilizing and leveraging social trust and reputation. Golbeck first introduced the notion of ontological modeling of trust in semantic social Web [13], [14]. Later on, Golbeck and Ziegler [15] pointed out the importance of profile similarity as a metric to infer reputation-based trust values in a social network and they utilized resulting trust values for improving word-of-mouth style recommendations. Following the Golbeck’s ontology, functional models of social trust are proposed. Dokoohaki and Matskin introduce a functional, yet very light-weighted ontology of trust [16]. The semantic model captures the semantics of relationship concept, where topic and metric of trust is documented under MainProperties of relationship concept, while the context of relationship (e.g. date of relation initiation, goal of it, etc.) is kept under AuxiliaryProperties concept. This trust ontology was used later on by Zarghami and Fazeli [17], as the main knowledge model of a trust-based recommendation system. Ontologies of reputation have been proposed as well. Casare and Sichman [18] have introduced a functional ontology of reputation to model reputation of intelligent agents. Since they utilize legal norms, they model social control mechanisms for software agents. As a result, such model becomes suitable for utilization in Social Web as well. Chang et al. [19] propose a basic reputation ontology and an advanced reputation ontology. They also distinguish between the entities towards which reputation is modeled for. Since the major focus is on e-commerce agencies, this model is not entirely suitable for modeling reputation of social users. Main argument for both previous models is lack of quantifiable semantics leading to lack interoperability in between them. Reputation interoperability can be enabled through utilizing semantic technologies [20]. Alnemr et al.[20] propose a functional reputation ontology that can serve as a vocabulary to be utilized in several applications. In this work, reputation is modeled as a complex object Reputation Object (RO). While RO captures the semantics
of reputation assertions, \textit{ReputationValues} represent the metric for reputation object instances, while the context of reputation is described using the \textit{Criteria} concept, that documents the provenance of the facts surrounding these assertions, such as algorithms used for gathering and computing the values.

Examples of adoption of reputation and trust in user models as pointed out earlier have been limited. Grapple project [9] investigates capturing and utilization of reputation to model the trust between users, by allowing the users to rate each other’s opinions and statements, following the eBay model [21]. Adoption of such a plain model of reputation is not successful, nor sufficient in generic and unified models of users, due to several reasons. First of all, rating is an implicit model of reputation, and representing it as a simple form of property-rating or a vector of ratings strips it from its original notion and postulation, according to Alnemr [20]. On the other hand, many systems are already using explicit trust statements to evaluate users, such as Epinions [22]. Second, since trust and reputation convey different semantics on Social Web, then frameworks for modeling users should be capable of describing trust and reputation separately. This difference is pointed out when you introduce a trust model capable of describing trusted peers of a user on a social network, e.g. Facebook or LinkedIn, as well as a reputation model capable of storing and presenting the reputation of user across different communities on-line, such as reputation of a user as a reviewer on Amazon, or reputation of a user as blogger in a blogging community such as Twitter.

III. UNDERSTANDING IMPORTANCE OF SOCIAL USER MODELS IN CROSS-SYSTEMS PERSONALIZATION

The aim of this section is to better address the advantages of bringing trust and privacy together to improve system’s adaptation. We present a brief use case where we describe how our social user model can work in a social web environment. Tom has a strong interest in art and he loves dancing tango. He lives in Turin and he joins iCITY [23], a social community dealing with events and attractions in the city, in order to get suggestion about what to do in the city. Tom use many of the most popular social site, like del.icio.us, Flickr, Facebook and LinkedIn. All these social applications collect a lot of data about his current interests, preferences, activities, which make available to other users and other applications. Thus, Tom wants to control the release of such information to other people: for example, he wants that only friends who share the passion for tango with him can see the news about tango he posted on Facebook wall. Furthermore, among such people, he wants that only the people he trusts more can see his score in the latest tango competition, like her friend Jill. Tom is planning a weekend in Florence, and he would like to visit the Institute and Museum of the History of Science in Florence. Smartmuseum application [24] is available for such museum. Smartmuseum is able to collect all the information about Tom the social web applications he interacted with made available and, using them, to build a user model of Tom on the fly. This information can be used to initialize the adaptation process.

This model also considers the preferences Tom declared about the release of information to other applications: his personal information can be delivered only to trusted applications which are forbidden to use them for commercial purpose. In particular, the information iCITY maintains about the events Tom has seen, the tags he inserted and the topics he is interested in could be very useful for the museum system to quickly identify his focus of interest and offer him a personalized visit to the museum. Since iCITY agreed on that privacy policy, after the interaction, Smartmuseum will send to iCITY some novel information about Tom that can be used to update the current user model of the application. This scenario can serve as a guideline for re-use of user interaction data generated by one application into another across similar domains. In this way, we illustrated how three user modeling problems can be solved, i.e. (1) cold-start problem in Smartmuseum, that can initialize the user model and start the recommendation from a point closer to user’s interests, (2) maintaining an integrated user profile, which reflects larger scope of user interests and activities, (3) the release of information (to other applications and to other people) take the user’s preferences for privacy and trust into account. In this paper, we focused on this third advantage. In the current situation, this scenario is far to happen, due to lack of integration among social applications and user data, and due to the lack of policies which integrate trust and privacy.

IV. OUR FRAMEWORK FOR USER MODELING IN THE SOCIAL WEB

Since modeling the users on the Social Web is a very complex task, an investment is needed for putting these separate pieces together. At the same time, we also aim at bridging the space left by the previous work by considering privacy, reputation and trust, the most crucial concepts within Social Web as the key missing concepts and dimensions surrounding the notion of user on the Social Web. To this end, we have proposed for a user modeling framework within which any user model can be imported and extended with social dimensions and enriched with privacy preferences, reputation and trust assertions. Our model of social web users will contain the following models:

- \textit{User model}, the description of user features according to existing de-facto standards such as GUMO [25], UUCM [6].
- \textit{Domain models} specific for the domain, such as standard domain vocabularies as AAT [26], ULAN [27] for artworks, etc.
- \textit{Context model}, which describes both the physical context (e.g., place, time, etc) and the social context (e.g., relations with other users and roles).
- \textit{User Activities model}, which describes the actions of the users (such as ATOM model [28]).
- \textit{Social data model}, which describes the social data: service data, disclosed data, incidental data, behavioral data, derived data (following Schneier model [29]).
• Privacy model, which describes the main privacy concepts for a user to be able to specify his/her own privacy preferences and policies.
• Trust and Reputation model, which describes main trust concepts between two individuals as well as expressing reputation towards a single or a group of individuals.

All such models have been represented as OWL ontologies. In the following, we will describe in more details the Privacy model (see Section V) and the Trust and Reputation model (see Section VI), since they are the main contributions of the paper.

V. PRIVACY MODEL

According to Kim et al. [30] the most important piece of a privacy-respecting Semantic Web is a privacy ontology that enables agents to exchange privacy-related information using a common language. The privacy ontology should be able to clearly define the various dimensions of privacy (e.g. privacy of personal behavior vs. privacy of communications), and contain enough parameters and index terms to enable specification of a privacy policy in a standard machine-understandable format. It should be descriptive enough to specify the highest known standards of data protection and privacy. Following former suggestions, we have defined a light-weight privacy ontology in OWL-2 which describes the main concepts of privacy in a social context, and the relations among such concepts. We took inspiration from the Unified Model for Privacy Preferences [12] (see Section II).

We also use some of the concepts OWL-S privacy ontology [31], a simple and easy-to-use ontology for expressing privacy policies as well as a protocol to support matching of such policies among Web Services. However, we developed our own ontology, since our point of view is the user in a social context, and not the provided services, as in that case. Our goal was to have a model that is platform independent and can be used in different contexts, able to cross the borders of social platforms (the so called Walled Garden of the Social Web [12]), and expressed by the means of semantic web language to promote interoperability among applications. As we will see, some portion of the ontology has been imported from OWL-s ontology, for re-usability purposes. We have defined the following main concepts:

• **Who** (the recipient of data): individuals (friends, family members, colleagues, companions, etc); agents; organizations*, business*, government agency*
• **What** (the data that are the objects): user model**, context model, domain model (link to some domain ontology), social model***
• **When** (retention time): week day (working days, week end), day hours (morning, afternoon, evening)
• **Where** (place the data are physically stored): address, location information (link to some geo ontology).

Notice that the dimensions signaled with: * means that they are imported from the OWL-s ontology; ** imported from the GUMO ontology; *** imported from the Grapple model.

• Why (purpose why the data are collected): to be processed (for adaptation purpose, for marketing purpose, for inference purpose, for data mining purpose), to be sold, to be transmitted.
• How (process made to the data): data protection techniques; privacy actions*, privacy policies*.

Figure 1 visualizes the privacy ontology representing the taxonomy of the involved concepts.

This ontology model allows then to define privacy policies according to such information. A set of SWRL rules for describing privacy policies can be defined for each specific users; in particular, "what" and "who" associations have been chosen as a first domain. An example of rule is the following (we omitted prefixes to enhance readability): it can express the fact that a user can let his/her colleagues see where she is or access her calendar activities only between 8am and 5pm on the weekdays but not over the weekend.

\[
\text{Location}(?x) \land \text{Tasks}(?y) \land \\
\text{Day}(?x, \text{Working\_days}) \land \\
\quad \rightarrow \text{can\_be\_disclosed}(c, \text{Colleague})
\]

The choices about privacy policies are largely subjective, and cannot be defined at priory, but it depends of course on the users preferences and situations conditions. Therefore, privacy policies are not a priority at this stage of the project, and they need further investigation.

VI. TRUST AND REPUTATION MODEL

Artz and Gil [32] categorize the notion of trust in computer science domain into three main categories: policy-based trust, reputation-based trust and general models of trust. While Semantic Web has benefited from research of all three subcategories, it is well-accepted that a Social Web model of trust is reputation-based. Golbeck first referred to such model as a Web of Trust [13]. A Web of Trust is a directed-edge network between a group of entities (or resources), within which each link carries a trust value and, assuming a transitivity of
trust, reputation can be collected and inferred for each single individual across such network. Within the context of Web of Trust [13], reputation can be defined as a measure of trust, within which individuals can gather and maintain reputation of other individuals across the network. To express trust and reputation information we have used ontologies allowing for expression and quantification of trust for use in algorithms to make a trust decision about any two entities [16], e.g. Tom trusts Jill highly with respect to dancing. We propose for a combined model of social trust and reputation, bearing in mind the details described previously. To model the trust we adopt the concepts of Alnemr’s ontology [20]. We fuse two sub-ontologies together using a new concept, called Context for modeling both trust relations and reputation concepts, through which contextual details of trust and reputation can be captured and stored. While ontologies of trust have allowed for expressing trust between two individuals, it is important to be able to express collective knowledge of trusted opinions about an entity as well. This form of reputation demands a model capable of documenting reputation assertions on its own without pointing to provenance of the assertion of trust [20], e.g. Jill is well-known for her skill in dancing. While trust ontology enables us to model a trust network of social inter-relations, extended ontology of reputation enables us to model assertions of reputation separately as well. This way we can fully capture the semantics of reputation-based trust on social web. Following previous discussion, we model trust and reputation using concepts below:

- **Trust** (Main concept of trust): Abstract trust (relationship).
- **Relationship** (Connection between two trusting peers): Relationship is the most important concept of our trust model. Relationship always has a sink and a source, which we have described here as *truster* and *trustee* entities. We have used two exact cardinalities on *hasTrustee* and *hasTruster*, in order to state having exactly one truster and one trustee per each relation.
- **Entity(Truster)** (Source of trusted relation): We distinguish between source and target of trust as a trust network is always a directed graph [13]. We distinguish between source and target of trust as a trust network is always a directed graph [13].
- **Entity(Trustee)** (Sink of trusted relation): Same as Truster, the target or sink of trust relationship. We need both entities to be able to determine the credibility of statements issued.
- **Trust Topic and Value** (Main properties of trust): Every trust relation is established surrounding a topic and is quantified using a metric. Following this assumption, we use main properties concept to model the subject and value of trust. Restrictions allow us to assign a single value and subject for each single relation subject to trustworthiness modeling.
- **Context** (Context of trust): Contextual properties of trust is realized using this concept. Defining context for trust relations allows us to specify functional or non-functional auxiliary properties of trust in our model. In the case of functional properties, for instance the algorithms used to gather and compute the trust values can be presented. For instance, we might use spreading activation [33] for gathering trust values, or T-index[17] for computing the trust values. Having context allows us to record the time, date or location that such relationship was established or the type of social network this relation was created, such as business in the case of LinkedIn. We use this concept to merge Trust model to Reputation model by defining Context as superclass of Criterion (see figure 2).

- **Reputation** (Reputation assertion): A Reputation assertion about an entity. Using this concept we can assert and define reputation for any entity (person, organization, group). The model adopted here allows us to define completely mention the trust statements used to.
- **ReputationValue** (reputation metric): Reputation of an entity (truster) is quantified and stored using instances of this concept. We can use the current value to represent the current reputation score while collection of reputation values asserted can be stored in history list. This allows us to gather and store all explicit (trust) or implicit (votes) statements towards an entity. Gathering provenance about an entity’s reputation history allows us to later assess the credibility of statement issuers. concept of PossibleValues allows us to define different ranges and values for reputation and store them together.
- **Criterion** (Context of reputation): Contextual properties of reputation is realized using this concept. Defining context for reputation assertions allows us to specify functional or non-functional auxiliary properties of reputation in our model. In the case of functional properties, the algorithms used to gather and compute the trust values can be presented. For instance, we might use a simple web crawler for gathering trust values, or we might utilize Sum or Bayesian functions for computing reputation scores [34]. Similar to trust, having criterion allows us to record the time, date or location that reputation was asserted.

Figure 2 visualizes the trust and reputation ontologies, representing the taxonomy of the involved concepts.

We aimed at proposing an interoperable model for embedding trust and reputation into any user-centric adaptive system, as well as sharing statements and assertions of trust and reputation across multiple systems. Thus any model of trust and reputation modeled for social context, should be capable of being aligned with our model. Taking this into account we avoid making choice between metrics for either trust or reputation. This should also be mentioned that choice of metric is also heavily dependent on application, user behaviour as well as data at hand. As a result choices of metric or algorithms are not a priority at this stage, and we will investigate further.
which metrics or mechanisms suite best for similar scenarios.

VII. CONCLUSION

In this paper, we have presented an approach for modeling the user in Social Web. The goal of our research work is to study how to put together all the standards and initiatives separately made by different entities in order to provide a complete model of a user which interacts with social web context. In detail, the main contribution of our work is to propose a model of user in a social context:

- that can be used as a reference to model users in social web context;
- which contains explicit modeling of privacy and trust dimensions, that usually existing models do not consider all together;
- that can be directly used by socially intelligent agents and by adaptive systems, populating and consuming it using real user data.

In our future work, we are planning to exploit the model in a existing social recommender systems, and evaluate the impact in recommendations and the final user satisfaction.

REFERENCES


