Towards A Comprehensive Requirements Architecture For Privacy-Aware Social Recommender Systems

Shan Chen **Mary-Anne Williams**

Innovation and Enterprise Research Laboratory
Centre for Quantum Computation and Intelligent Systems
Faculty of Engineering and Information Technology
University of Technology, Sydney
Email: shanc|mary-anne@it.uts.edu.au

Abstract

Social recommendations have been rapidly adopted as important components in social network sites. However, they assume a cooperative relationship between parties involved. This assumption can lead to the creation of privacy issues and new opportunities for privacy infringements. Traditional recommendation techniques fail to address these issues, and as a consequence the development of privacy-aware cooperative social recommender systems give rise to an important research gap. In this paper we identify key problems that arise from the privacy dimension of social recommendations and propose a comprehensive requirements architecture for building privacy-aware cooperative social recommender systems.

1 Introduction

Content-based filtering (Billsus & Pazzani 2007) vs. collaborative filtering (Goldberg et al 1992) have been dominating the traditional recommender systems. Typically, users are classified by their interests and/or preferences based on some similarity measures. Grouping using these approaches connects users to each other, implicitly or explicitly. Such connections can create new and reveal social contexts for users, and can make the system prone to privacy breaches. On the other hand, the development of online social networks has addressed the need to support social recommendations, i.e., provide users the ability to introduce people (i.e., friends) to others or to offer social referrals between users to facilitate network consolidation and expansion. However, such recommendations are either based on existing connections or created using social factors in individual's personal space. For example, many *social network sites* (SNS) - e.g., Facebook (2009), LinkedIn (2009) and Pulse (2009) - provide a list of "People you may know" for users to build and/or expand their networks, or to invite people who were previously unknown. People being introduced in this way are either socially connected explicitly, or identified based on the similarity of some attributes (e.g., interests, geography location, occupation, etc.) This kind of recommendation can inject many privacy issues for parties involved, and as a consequence increase the opportunities for privacy infringements.

In the problem domain of social recommendation, we argue that the privacy issues arise due to the lack of *choice* offered to users as to whether they want to be introduced or referred - i.e., *consent* - and the ability to *control* "who knows what about me". The problem of *choice* concerns users' rights to choose their preferences and give *consent*. The *control* problem has two important aspects: the *who* and the *what* dimensions - the latter concerns "things about me", while the former concerns "things about others". As a consequence, concerns about users' rights involve both "me" and "others". Privacy issues related to the *control* problem highlight the balance of rights between users. By identifying key problems that arise from the underlying *rights* (i.e., choice, consent and control), this paper studies privacy requirements and proposes a requirements architecture for building privacy-aware social recommender systems.

The rest of this paper is organized as follows. Section 2 studies the problem domain, section 3 idenfifies architectural requirements, section 4 investigates privacy issue in social connections, section 5 establishes social connection privacy preserving requirements, section 6 describes system requriements, and section 7 presents a discussion and future work.

2 Problem Domain

This section describes a motivating example that reflects a social recommendation service - i.e., *People you may know* (PYMK) - which has been offered by many popular social network sites such as Facebook, LinkedIn and Pulse. It then discusses problems arise from the motivating example to uncover the fundamental privacy problem in social recommendations.

2.1 Motivating Example

Mary joined a social network on MySN site and she received a PYMK list, on which those who went to the same school as her and those who share the same type of profession as her are listed, allowing her to send a message to them. Mary was surprised to find many school friends she had lost contact with on the PYMK list. However, she also felt compromised to be on the MySN because she wanted to keep her professional information disjoint from her personal social network, and to keep her away from those professionals she did not want to network with. She reasoned that if she saw other's information they could also see hers.

2.2 The Right Problem

From the motivating example, above, we can see that if Mary was not asked if she wanted to be on the PYMK list that appeared to others (e.g., to the public or to specific targeted groups), then she had no way of choosing preferences and giving permissions to control her information privacy. We refer to this problem the user's *right of choice.* If choice is offered, the ability to consent on the usage of information is then required

Copyright ©2010, Australian Computer Society, Inc. This paper appeared at the Seventh Asia-Pacific Conference on Conceptual Modelling (APCCM 2010), Brisbane, Australia, January 2010. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 110, Sebastian Link and Aditya K. Ghose, Ed. Reproduction for academic, not-for profit purposes permitted provided this text is included.

to control the information. In other words, the user's *right to consent* and *right to control* their information are essential to fulfill privacy requirements. In this light, the privacy problem in social recommendations mainly involves users' rights of *choice*, *consent* and *control*. We refer to these rights the *3C Rights* (3CR) framework and describe them as follows:

- *choice* - the ability to choose to-be or not-to-be introduced or referred;
- *consent* - the freedom to give permissions of personal information usage to others; and
- *control* - the power to control personal information and ways of sharing it.

The 3CR framework can provide a specification of higher-level privacy requirements to the recommender provider. To fulfill these requirements, lower-level detailed requirements for recommender system implementation are required. To discover fundamental problems behind the 3CR framework, we analyze the interplay between each right of the 3CR using scenarios educed from the motivating example above.

2.2.1 Choice

Scenario MySN adds Mary to the PYMK list with people that share the same type of profession as her and then presents the PYMK list to her boss who is on the MySN. Mary's boss sends a friending request to her and she adds her boss upon his request because she does not want to be impolite. However, after her boss joins her social network, Mary fails to maintain the disjointedness of her personal social network and professional social network.

Problem The user's choice of being introduced or referred will have an impact on his/her information privacy. In the light of 3CR, a shortage of the right to *choice* naturally leads to a deficiency of rights to *consent* and *control* information.

2.2.2 Consent

Scenario MySN offers recommendation choice options: to-be-recommended or not-to-be-recommended. Mary wants to expend her social networks but does not want her boss to be in her social network on MySN. She knows that her boss is on MySN. If she chooses the option to-be-recommended then her boss will know of her existence and might request friending. But if she chooses not-to-be-recommended then she will loose the opportunity to be known to potential social contacts.

Problem Having binary choice options is insufficient for supporting consent and leads to failures of information control because social relationships are not binary: friend or not. Users should have the freedom to give different permissions to different contacts.

2.2.3 Control

Scenario The new version of MySN allows Mary to specify who she will not be recommended to. Mary believes on MySN she can now stay away from her boss because she has specified the name of her boss not to receive recommendation about her. However, two days later she receives a friending request from her boss. She does not know that her MySN new friend Phoebe is her boss's little daughter who shares her online experience with her father.

Problem Even though Mary has sufficient *rights of choice and consent*, she does not have sufficient power to control what information is made available to her contacts - i.e., ways of sharing in the network.

It can be seen that, the magnitude and dimension of 3CR as well as the interplay between the three rights have a major impact on the privacy. We describe the magnitude of 3CR in terms of *3CR values*:

- the range of available choice options,
- the type and detail of consent the user can set, and
- the level of power the users have in controlling his/her information.

These 3CR values and the ways they interoperate can lead to different impacts of the information usage and in turn the privacy. Since the problems reflected in these values are closely related to social problems in social networks, to gain an insight into their privacy implications, we study the related social problems in the next sub-section.

2.3 The Social Problem

Given that social interactions are fundamental activities in social networks and interactions are based on social connections, the context of a social network is framed by social entities (e.g., users) and relationships connecting them. In this light, fundamental to the privacy problem in social networks is philosophy of social relationships - i.e., the relationship privacy is attributed as the primary privacy problem in social networks. Processed in social networks, social recommendations inherit the philosophical privacy problem - i.e., relationship privacy as social recommendation privacy.

Relationship privacy involves several problems that needs to be addressed. In light of the 3CR, these problems typically are:

- the selection of potential parties - who can be considered as appropriate candidate(s) to network with; and
- the selection of specific information to share - who can share some certain information with and in what way.

It has been evidenced that different networks tolerate different connections, reflected in properties such as types, degrees, directions and multiplex (Chen & Williams 2009). The way these properties cohere to balance users' rights in the 3CR space provides a key to the preservation of users' information privacy in social recommendations. However, the dynamic of social networks gives users no way of knowing the status of these properties. Consequently the 3CR problem in social networks leads to several operational issues described in the next sub-section.

2.4 The Operational Problem

Information privacy requires users be aware of the current status of the social network in which they interact with others. One way to address this problem would be to allow users to query the network about self and others. However, to promise a balance of rights between users, queries cannot return comprehensive information to the user that violates others' privacy. On the other hand, queries can potentially reveal the user's privacy because they reflect the querier's intentions. Accordingly, privacy-aware queries need to be constructed with consideration to the following issues.

- content - i.e., *what information can be retrieved* such that the maximum information can be obtained without violating privacy; and
- behavior - i.e., *how to query* such that the querier's intentions that can reveal or be used to infer privacy are not disclosed while at the same time necessary information can be obtained as complete as possible.

While the social problem concerns current status of a social network, the operational problem gives consideration to the dynamic aspect of the network - i.e., the evolution of the network. This can be reflected in both the content issue and the behavior issue taking privacy implications into account upon each operation. The key to uncover privacy implications in evolving social networks is to learn potential ways the user connects to others - i.e., possible relationships that can be established. Since each candidate is a social entity playing specific social roles in the network, they can have different impacts on their social connections and in turn impact the privacy of those connected to them.

2.5 Summary

The problems described above suggest the privacy problem domain in social recommendations can be divided into *choice*, *consent* and *control* issues, with the core in *relationships*. Consequently it requires an adaptable and extendable choice space, rich relationship semantics, and privacy-aware queries. In the subsequent sections we identify low-level challenges and fine-grained requirements to address these problems.

3 Architectural Requirements

3.1 Choice, Consent & Control

Our architecture is based on three core pillars: choice, consent and control (Williams 2009). Users are given *choice* to *consent* and to *control* their information privacy. To create consent for recommendations users need to be able to accurately express their needs. In addition, preferences for wishes and interests are preferable because they help to determine users' intentions and in turn privacy management decisions. Since people's desires, wishes and interests are highly situated and can be multiplex, to accurately capture preferences adaptability and extendability of choice are essential. To this end, a capable privacy-aware recommender will provide users with the following:

- *choice options* that allow them to
  - express their needs and preferences accurately, and
  - adjust and change their needs or preferences to new conditions.
- *consent mechanisms* that enable them to
  - learn the context of their own networks in relation to privacy implications, and
  - specify permissions for using their data and obligations attached to the usage.
- *control devices* that provide them ability to
  - control ways of sharing personal information, and
  - verify expected controls.

It can be seen that, the ability to preserve privacy largely depends on the power to control information usage. To achieve this, comprehensive guidelines for privacy protection are essential. The set of principles for privacy protection identified by the Organization for Economic Cooperation and Development (OECD) (OECD 2009) is a good candidate because they represent as far as possible a global consensus.

3.2 OECD Privacy Principles (OECD_PP)

The eight principles for privacy protection identified by OECD are as follows:

1. Collection Limitation (CL) limits the collection of personal data.
2. Data Quality (DQ) ensures personal data is relevant to the purposes of used.
3. Purpose Specification (PS) restricts the collected data to the purposes of collection.
4. Use Limitation (UL) restricts data to be used within the permission of the purpose specification.
5. Security Safeguards (SS) ensures data is protected by safeguards.
6. Openness (OP) ensures policies with respect to personal data are open to the user.
7. Individual Participation (IP) ensures individual rights of actions related to own personal data.
8. Accountability (AC) ensures the principles above are complied.

Based on the notion of the 3CR, this set of eight principles are categorized into the 3CR groups and serve as a higher-level guideline for specifications of layer requirements. Fig. 1 shows a two-layer requirement architecture.

3.3 OECD_PP in Recommendations

The increasing number of privacy breaches reported in the media almost everyday has demonstrated that, from a users' perspective, there is insufficient support for the principles of Purpose Specification and Use Limitation in existing SNS. One might argue that these SNS do provide limited access control support. However, users are not made aware of nor do they have the ability to specify the purpose and usage of their information being collected. We argue that these two principles dominate social recommendations because:

- *The Purpose Specification Principle* restricts the use of the collected data to the purposes of collection - it concerns the consistency of data usage and the purpose for which they were collected. For example, relationship information was collected for the purpose of sending social recommendations - i.e., determination of recommendation target/candidates, or personal information was collected for sending social recommendations and not for other types of recommendations like buying or selling.
- *The Use Limitation Principle* restricts data to be used within the permission of the purpose specification - it concerns deviations from specified purposes. For example, the problem of inconsistency where a social relationship exists for the purpose of social interactions, e.g., a religious relationship is for religious interactions and not for trading interactions.

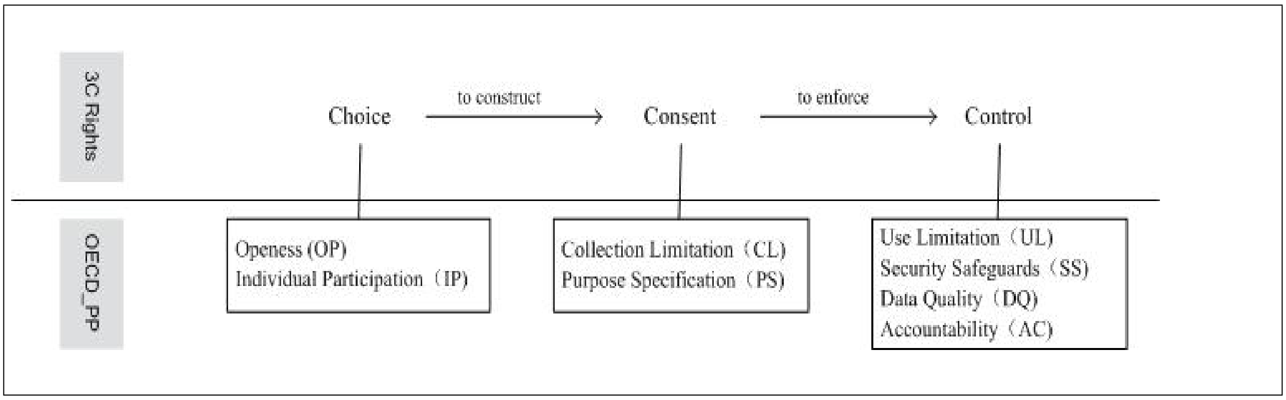

Figure 1: Two-Layer Requirements Architecture

3.4 Layered Requirements

For the above reasons, we aim to address the problem of users' awareness with respect to the principles of PS&UL in the problem domain of 3CR - in this light, Users' awareness is about:

- self information and ability (i.e., power to control their information on the SNS); and
- others' information relevant to his privacy - i.e., users' knowledge about the social context in which he lives, i.e., users' own social networks.

To this end, it is necessary to provide users mechanisms to query the network to alleviate their concerns and to increase their awareness, i.e., *query-answering* as an essential component of the recommender. With the focus on the principles of PS&UL, query-answering addresses the problem of user awareness with respect to privacy. In this regard, the following requirements are established:

- at the 3CR layer, users are provided *choice* options and *consent* to important aspects of *control.*
- at the OECD_PP layer, users are provided conceptual guidelines on what needs to be taken care of in terms of constructing consent and enforcing control.
  - choice, the recommender allows user interactions (IP) for query-answering based on the provided options (OP);
  - consent, the recommender allows user consents set on the collection (CL) and the usage (PS); and
  - control, the recommender allows user information control in usage (UL).
- at the Codes of Practice (CoP) layer, users learn to situate themselves to better establish controls using queries that are based on the set of CoP - in this paper, learning is mainly used to address the problem of social connections, i.e., relationship privacy.

It can be seen that, the CoP layer is a new layer added to the two layered requriements. The purpose of layering requirements is to allow externality (in opposition to "pure" system requirements) input such as law and social norms as higher-level conceptual guidelines for system design situated at a lower-level where CoP applies. The role that CoP plays requires fine-grained requirements of *social connection* and *relationship privacy* be understood and specified.

4 Privacy in Social Connections

4.1 Social Connection and Relationship

To learn the meaning of relationship privacy, we begin from intrinsic properties of a relationship. As we have identified in (Chen & Williams 2009) relationships are multifaceted and can be symmetric or asymmetric. A relationship is symmetric if both ends of the relationship share the same attitude of the relationship, i.e., recognize the relationship under the same conditions; otherwise, the relationship is asymmetric. For example, the relationship between A and B is symmetric if both set the relationship to the same type under the same conditions. The relationship is asymmetric if A sees B as a friend but B sees A as a colleague. The asymmetric property implies the existence of *direction* in a relationship. On the other hand, the same pair of social entities can hold more than one type of relationship. For example, A and B are siblings, classmates and both are members of club A&B. The connection between A and B is therefore described as "sibling, classmate and A&B member".

To better understand the privacy issues that can arise between two social entities, we distinguish the concept of relationship and the concept of social connection. A *social connection* indicates two social entities are connected by some reason. Each involve the connection of two social entities in a *relationship*, which has a type and a direction. A social connection is *multiplex* if there is more than one relationship held between the two entities on the connection.

Social referrals naturally introduce indirect social connections. When one entity connects to another via a referral, the connection is indirect. The concept of *connection degree* is used to indicate the distance between two entities. For example, if A connects to B, and B connects to C, then A is said to be 2 degrees away from C, i.e., the connection degree between A and C is 2. If there are multiple paths connecting A to B, technically the degree of A and B is the length of the shortest path between A and B.

4.2 Privacy Concerns

In the context of social recommendations, we study what can have an impact on a relationship in a specifc context. In the following we begin with a set of scenarios elicited from a social referral.

Assume that on MySN everyone can refer their social contacts to each other. In this example M refers S to X. Scenarios where privacy concerns arise might include the following:

- If M introduces S to X explicitly, then possible privacy costs from such a referral include:

  - X knows "*M knows S (the latter may not know the former)*" - the existence of the one-way relationship is disclosed;
  - X knows "*M and S know each other (i.e., there is a relationship between them)*" - the existence of the two-way relationship is disclosed; or
  - X knows "*M and S are connected by R reason (e.g., working in the same department)*" - the type of relationships is disclosed.

- If M refers S to X implicitly by i) anonymity, ii) using a different ID, or iii) via a third-party (i.e., the system or another social entity), if these features available, then privacy costs from such a referral could be:
  - X knows "*someone that knows S does not want to be known to X*" (by anonymity);
  - X knows "*the owner of the ID knows S*" (by a different ID), or
  - X knows "*the third-party knows M does not want to reveal his relationship with S to X*" (by a third-party).

In any of these cases, the referrer (i.e., M, or the third-party) knows that S is known to X.

4.3 Relationship Privacy

Relationship privacy concerns illustrated in this example can be encapsulated in terms of *existence* and *relationship type*:

- *Existence* concerns whether one has a relationship with another. For example, M refers S to X implicitly because of this concern. On the other hand, did M know if S had the same concern?
- *Relationship type* concerns the kind of relationship between two social entities. For example, sibling, friend and colleague are different relationship types.

On the other hand, social entities in a social network are interconnected. Social entities can connect directly or indirectly. A consequence of relationship privacy between the two social entities who hold the connection and their social contacts, existing or potential, is extended to *access control* and *distribution control*:

- *Access control* concerns the kind of social contacts that can be granted access to certain information, or the condition of building a relationship with others that can grant access to the information (recourse). In the context of social recommendation, relationship privacy has three main relationships:
  - the relationship between the recommender agent and the recommended agent,
  - the relationship between the recommended agent and the recommendation recipient, and
  - the relationship between the recommender agent and the recommendation recipient.

  In this paper, we give consideration to the recommended agent - i.e., we consider the recommended agent as a privacy priority agent such that *access control* is the recommended agent's power to control his information being accessed by the recommender agent (i.e., to disclose information about the recommended agent) and the recommendation recipient (i.e., to know information about the recommended agent). For example, does M know if S is willing to make a connection to social entities like X (e.g., X is a member of group GA and if S is concerned about his privacy when having a relationship to GA).
- *Distribution control* concerns the kind of social contacts knowing (i.e., the *existence concern* and the *relationship type concern*) or having access control to some information can also grant permission to distribute information under certain constraints - e.g., before or after accepting the recommendation (i.e., accept S as a social contact for some purpose), can X refer S to others? Can X disclose the relationship between M and S (e.g., when referring S to other social entities)? On the other hand, X also got to know something about M (if not before) - can X distribute M's information (e.g., refer M to others - i.e., reveal the existence of M - and the relationship between them)?

When each of these concerns on each dimension of a social connection are evident, the privacy issue tends to be multi-layered. Consider the referral example above, If M makes the referral explicitly, the *existence* concern on each dimension extends to

- *direction* - can X know if S and M holds a symmetric or asymmetric relationship? In the case of an asymmetric relationship, who is the dominant partner in the relationship?
- *multiplex* - can X know if S connected to M in various ways and how are they connected?
- *connection degree* - can X know if S is a direct contact of M? If not, how many degrees away?

It can be seen that these concerns on each dimension can take the problem to a level where more sensitive and negative implications can be discovered. This suggests multi-layer relationship privacy requirements (RPR) for connection-driven social recommenders. The multi-layer RPR serves as a guideline for establishing the codes of practice.

5 Social Connection Privacy Preserving Requirements

Given that our aim is to preserve relationship privacy in social recommenders, the *codes of practice* (CoP) focus on the semantics of relationship privacy on four dimensions categorized as *disclosure* and *control.*

Disclosure

This category concerns the properties of *existence (EX)* and *relationship type (RT)*. Let P denote a property of a relationship privacy in Disclosure, such that $P = EX$ for property "existence" and $P = RT$ for property "relationship type". Then, the CoP for Disclosure are as follows:

- Direction (P:D)
  - The disclosure of property P of a symmetric relationship should only be made with the explicit consent of both parties of the relationship.
  - The disclosure of property P of an asymmetric relationship should be made with the consent of the dominant party. If $P = EX$, the consent includes the existence of the relationship and its direction.

- Multiplex (P:M)
  - The disclose of property $P = EX$ of each relationship in a multiplex connection should only be made with the consent of the dominant party of each relationship[1].
- Connection Degree (P:CD)
  - The disclose of property P of an indirect relationship should only be made with the consent of both parties of the relationship. Such consent includes the value of connection degree and the social entities connected on the path between two ends of the relationship.[2] If $P = RT$, the consent involves parties on the relationship that RT refers to.

Control
This category concerns the properties of *access (AC)* and *distribution (DT)*. Let P denote a property of a relationship privacy in Control, such that $P = AC$ for property "access" and $P = DT$ for property "distribution". Permissions for control of property P on certain information of an agent can only be granted upon the consent of the agent. Within the scope of a social recommendation, if

- $P = AC$
  Permission for agent A to introduce agent B to agent C should only be granted upon B's consent that allows his (B) certain types of social contacts (i.e., A) to introduce him (B) to certain types of social entities (i.e., C). Permission for agent A to establish a relationship R to agent B should only be granted upon B's consent that allows certain types of social contacts (i.e., A) to connect to him (B) on a relationship type of R.
- $P = DT$
  Permission for agent A to distribute certain information of agent B to agent C (e.g., when making recommendation) should only be granted upon B's consent that allows his (B) certain types of social contacts (i.e., A) to introduce him (B) to certain types of social entities (i.e., C). The kind of the information of interest must be considered for this property in the CoP below.

Let KI be "the kind of the information of interest", the CoP for Control are as follows:

- Direction (P:D) The type of A is determined by the relationship between A and B; and KI, if applicable. The type of C is determined by the relationship between A and C, and the potential relationship between B and C; and KI, if applicable. Direction should be concerned if the relationship is asymmetric.
- Multiplex (P:M) The type of A is determined by the set of relationships on the connection between A and B; and KI, if applicable. The type of C is determined by the set of relationships hold on the connection between A and C, and the potential connection between B and C.
- Connection Degree (P:CD) The type of A is determined by the set of relationships held on the set of connections between A and B; and KI, if applicable. The type of C is determined by the set of relationships held on the set of connections between A and C, and the potential connection between B and C.

By establishing the CoP above, the layered requirements can be shown in Fig. 2.

It can be seen that, these CoP provide a basis to support consent's construction. Central to the CoP is the problem of "who can see/use what?" (WCS/UW). In the social recommendation problem domain, social entities are identified by their social connections. Subsequently the problem of WCS/UW requires the user not only be able to identify the kind of social entities, but also their connections to the user. This requirement suggests the development of concepts of *abstraction* and *granularity*, where the former reflects the level of detail on social entities - e.g., the abstraction levels of individuals, groups, communities, organizations and networks are from the lowest to the highest, the latter refers to the fineness with which relationship types are categorized on a certain abstraction level - e.g., friends, business partners, family members, classmates, co-workers, etc. In this light, the AC and DT of CoP necessary consider abstraction levels, i.e., groups - i.e., for AC:D, AC:M, AC:CD, DT:D, DT:M and DT:CD, when determining B's relationship to A and C, criteria should include groups (if any) to which they belong.

6 System Requirements

6.1 Component Requirements

Towards the privacy-aware social recommender system that we propose, components provide functionality to fulfill the requirements established above are: *rights* components, *relationship registry* component and *query-answering* components:

- Rights components: Choice, Consent and Control three components to provide mechanisms for users to express their 3CR;
- Relationship Registry component to store all the relationships and policies;
- Query-Answering components:
  - Query Library to store queries and permissions attached; and
  - Obligation and Permission Reasoner to reason about obligations and permissions.

6.1.1 Choice

The Choice component provides a space for users to establish a basis of recommendation consents. To construct consents for privacy purposes, *choice* requires the following key abilities: *extensibility*, *expressivity* and *adaptability*.

Extensibility: It is essential for the users to accurately express their needs. In the choice space, users are given higher-level options as initial suggestions. Upon selections of these options, users can further detail their needs and preferences if necessary.

Expressivity: Users' requirements that can include short-term and/or long-term needs, and preferences for their wishes and interests, require rich semantics options. From a system design perspective, this requirement shows the need of expressive representations to capture comprehensive requirements.

[1] This code is not applicable to the property RT since a relationship is simplex. Thus, $P : M = EX : M$.

[2] This code is made on the assumption that an indirect relationship is symmetric since otherwise the combination of symmetric and asymmetric leading to the complex consent is out of the scope of this paper.

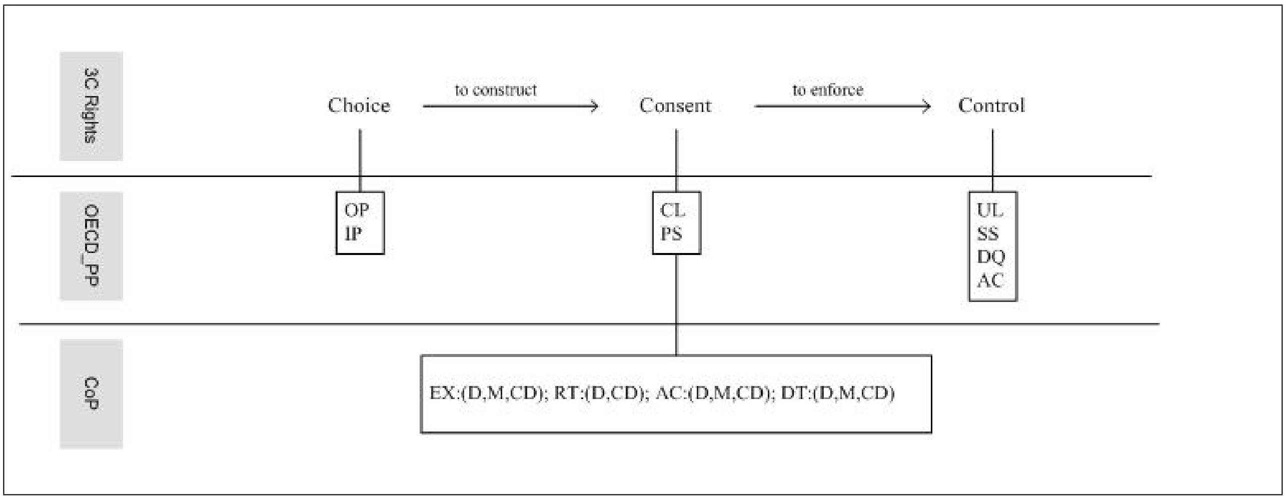

Figure 2: Three-Layer Requirements Architecture

Adaptability: Requirements can evolve over time. On the other hand, with the privacy issue in mind - i.e., relationship privacy (as identified above) - users' not only need to be able to express and control the evolution of their relationships, they may also need to identify their preferences at different level of detail for each relationship. As a result, *adaptability* to allow users to adjust their needs or preferences to new conditions is necessary.

6.1.2 Consent

The Consent component constructs permissions with privacy concerns and obligations in mind. As a prerequisite for *control* - i.e., the output of this component is provided to the Control component to generate controls - it concerns the problem of *fidelity* under the principle of Purpose Specification. As the successor to the Choice in 3C, it is transited with the problems of *extensibility*, *expressivity* and *adaptability.*

6.1.3 Control

The Control component manages users' intensions for the privacy of their information. It concerns the problem of *adequacy* under the principle of Use Limitation. As the top level of rights in the 3C, it inherits the problems of *extensibility*, *expressivity* and *adaptability* from its predecessor (i.e., the Consent component).

6.1.4 Relationship Registry

The Relationship Registry component stores all the relationships and associated policies. In this component, *granularity* and *abstraction* are the fundamental requirements and need to be highlighted. The finer granularity relationships tolerance, the more choice the user will have and the more accurate consents can be constructed, which in turn the more fine-graineded controls the user will have. This implies *adaptability* and *extensibility* requirements for fine-grained relationships. On the other hand, *abstraction levels* of individuals reflects their relationships to others within a certain scope. This implication not only stresses the need of *adaptability* and *extensibility*, but also requires *scalability* since abstractions can overlap in various degrees - i.e., when one's abstraction level changes, e.g., one can belong to multiple groups where some of these groups are nested and some are overlapped, when the relationship between groups are changed but one's group memberships remain, his relationship to other members in these group can change. The Relationship Registry is required to be able to adapt and scale to accommodate such changes.

The registry serves as a backbone to support the whole system with central information - i.e., relationships. Each time a new relationship is established, or an existing relationship is evolved, all the related information is required to register with the registry.

6.1.5 Query Library

The Query Library component stores all the queries available and conditions to restrict the kind of users using certain queries. Queries are constructed with the aim to assist users to learn more about their current positions or potential positions in the network in terms of privacy implications - i.e., identifying the choice options available to the user and the CoP for consents construction, queries look for information about the user's control power, and others information related to the user's privacy.

6.1.6 Obligation and Permission Reasoner

To answer queries with maximum information while committing to other users' information privacy preservation - i.e., it is required that the answers to be compliant with these users' privacy policies and not negative consequences. This reasoner serves to fulfill such requirements. It takes users' queries, checks privacy policies of other users that are involved in the answers, then reasons about

- whether the answers are consistent with those policies; and
- whether the querier's obligations and permissions in relation to the queries of interest are compliant with his/her policy and intentions to control own information privacy.

6.2 Overall Functional Requirements

Based on the requirements identified for each component in the previous subsection, the overall functional requirements for a privacy-aware recommender system are illustrated in the form of workflow described based on a social recommendation example as follows:

1. The user is offered two choice options for recommendation: to-be-recommended and not-to-be-recommended. If to-be-recommended is chosen, then the following options are provided.

2. The user is offered choice to query the existing network.

   The user has choice to learn about "who is around" - i.e., who is in the network and how they are connected. For social concerns, the user has choice to query about existence of groups and their members on their interests and habits - e.g., who is in his/her profession, who share the same interests, who has something in common (e.g., attending the same primary school), etc. For privacy concern, the user has choice to acquire knowledge about the social connections of those who are of interest, as well as their availabilities in terms of what they can do and what they cannot do in relation to the user's privacy concern. However, while the user is given choice for querying about the network, rights of social entities of interest to the user must also be taken care of - i.e., these entities' permissions for use of their information. If answers to the query will not respect to their privacy, i.e., answers conflict with other's permissions, then such answers should not be given. The Obligation and Permission Reasoner is responsible for checking such conflicts and provide advice in compliance with related entities' permissions. The CoP are applied to the related entities.

   Example
   A query could be "are there anyone from my company?" or "anyone in Cooking group working in my company?". If a member of the Cooking group is working for the company and has specified that he does not want to reveal his professional information on the network, then the answer cannot include any information that can reveal his such information.

3. The user specifies his/her interests with consideration of the existing network - i.e., knowledge learned from Step 2. above.

   Example
   The use asks to be recommended to "people in the Cooking group".

4. The user specifies his/her needs. The CoP are applied to the user.

   Example
   After querying, the user is aware that his boss is a member of the Cooking group. He does not want his boss to know that he is in the Cooking group so he requires to "keep me away from my boss".

5. The user queries about obligations and permissions on the existing network and potential network in relation to his consent. The CoP are applied to both the user and the related entities.

   Example
   In the previous step, if the user was told that his boss is not in the Cooking group but somewhere in the network, then he might further query about "is there someone in the Cooking group will share information with my boss (i.e., someone permits or has obligations to share information with his boss).

6. The user is notified about potential implications with respect to his privacy concern. The CoP are applied to both the user and the related entities.

   Example
   After step 5, the user is informed that "Member A is connected to your boss B and A allows B to view all her connections".

7. Based on the information obtained, the user makes a decision regarding whether to change his/her choices, consents and/or expectations of what to control. Then, repeat previous steps if desired and applicable.

   Example
   Upon obtaining the information from step 6, the user requires "not to be recommended to the Cooking group" (i.e., repeat step 1), or "keep me away from Member A" (i.e., repeat step 4).

7 Discussion and Future Work

The need for privacy-aware social recommendations has been stimulated by the increasing prevalent privacy infringements in current online social networks and the failures of existing recommendation techniques to address such problems. The development of a privacy-aware system requires a privacy-by-design approach that necessarily begins from requirements development. In the area of requirement engineering, a number of contributions to the work on semantic privacy issues have been made. For example. Anton et al (2002) propose using goal taxonomies to structure privacy policies. Liu et al (2003) use a role-based approach to study trust relations and attacker-defender relations. Giorgini et al (2005) propose a goal-oriented secure tropos methodology to address security and trust issues. A common deficiency of these attempts is lack of considerations of the right problem at social level.

In the problem domain of social recommendations, we argue that the fundamental problem is philosophically the *right problem* and sociologically the *social problem*, of which technology solutions are necessary attributed to relationships and multiplex of relationship privacy must be addressed. To this end, this paper presents a preliminary work towards a comprehensive layered requirements architecture for building privacy-aware social recommender systems.

The proposed requirements architecture takes philosophical and sociological needs into account, allowing inputs from external regulations like legislation by using a layer approach for requirements development. Within the scope of social recommendations, requirements are developed for the central problem identified - i.e., social connection and relationship. As a result of lessons learned from existing social recommender business models such as Facebook (2009), LinkedIn (2009) and Pulse (2009) that are inadequate to meet the requirements, functional requirements towards a privacy-aware cooperative social recommender system are also developed. Future work will take steps towards methodologies and techniques for fulfilling the functional requirements to realize privacy-awareness in social recommendations.

References

Anton, A. I., Earp, J. B., Reese, A. (2002), Analyzing Website Privacy Requirements Using a Privacy Goal Taxonomy. IEEE Joint International Requirements Engineering Conference (RE'02). pp.23-31.

Billsus, D., Pazzani, M. J. (2007), Content-based recommendation systems. The Adaptive Web, LNCS.4321,325-341.

Chen, S., Williams, M-A. (2009), Privacy in social networks: A comparative study. *in* PACIS 2009 Proceedings. Paper 81.

Giorgini, P., Massacci, F., Zannone, N. (2005), Security and Trust Requirements Engineering, *in* A.

Aldini, R. Gorrieri, and F. Martinelli, eds, 'FOSAD 2004/2005', LNCS 3655, pp. 237-272.

Goldberg, D., Nichols, D., Oki, B. M., Terry, D. (1992), Using collaborative filtering to weave an information tapestry. Communications of the ACM. 35(12), 61-70.

Liu, L., Yu, E. S. K., Mylopoulos, J. (2003), Security and Privacy Requirements Analysis within a Social Setting. *in* Proc. of RE'03. pp.151-161.

Williams, M-A. (2009), Privacy, the law and global business strategies: A case for privacy driven design. *in* Proceedings of the AAAI 2009 Spring Symposium on Social Semantic Web: Where Web 2.0 meets Web 3.0.

OECD. (2009), http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html. (Accessed: 1 July 2009).

Facebook. (2009), www.facebook.com. (Accessed: 1 July 2009).

LinkedIn. (2009), www.linkedin.com. (Accessed: 1 July 2009).

Pulse. (2009), www.plaxo.com. (Accessed: 1 July 2009).