Dynamically Semantic Reasoning for Privacy-Preserving Personalisation

Roghaiyeh(Ramisa) Gachpaz Hamed, Kaniz Fatema, Owen Conlan, and Declan O'Sullivan

ADAPT Centre
Trinity College Dublin - Ireland
{ramisa.hamed,kaniz.fatema,owen.conlan,declan.osullivan}@adaptcentre.ie

Abstract. Personalised systems are able to provide user-adaptive content and services for individual users based on their preferences and behaviour. Such systems are becoming increasingly popular. In order to function they need to collect comprehensive information about users, which amplify a variety of privacy implications. The tension between personalisation and privacy has become more prevalent due to both the availability of more sources of user data and intensive analysis facilities on one hand and increasingly tough restrictions imposed by privacy legislation on the other. The aim of this research is to propose a framework to enable semi-automated decision making by identifying sensitive information in a given context according to how a user has disclosed data in other circumstances.

Keywords: Privacy, Personalisation, Semantic Models, Ontology Mapping, OWL, Semantic Reasoning

1 Introduction

Personalized and user adaptive system have been developed in recent years in order to tailor the ever increasing number of web-based services to accommodate specific individual's needs and preferences. Using such system brings mutual benefits to both users and providers of such services. While users can gain appropriate products/services, providers are able to present appropriate services/products to their customers enabling better selling, an increase in customer satisfaction, an increase in benefits and an improvement in the quality of their services. To perform better adaptation, large amounts of data about users, typically their behaviour and interest, need to be collected which can often include sensitive information about users. Unauthorised access to such sensitive information may reveal the identifying information of individuals which may raise a variety of privacy challenges. Previous surveys and studies have highlighted the privacy concerns of personalisation system users, specifically in regard to the disclosure of their private and sensitive data [30]. Many surveys also reveal that if users know how their sensitive information is being used and have control over this usage they would be more willing to disclose their personal data

[17]. Therefore, potential risks of disclosing personal information conflict with its potential benefits, such as the value being assigned to personalisation, financial rewards and/or social adjustment benefits [1]. Consequently, privacy-related decisions, particularly in reference to personalisation, is not absolute and can be subjective, context-oriented and affected by many factors such as domain, usage, period of time, and stakeholder.

The aim of this early stage PhD research is to propose a framework for helping users in semi automated decision making by dynamically reasoning over the user profile, policies, context, and logs. To undertake the research we plan to use semantic technologies, not only to visualise context and user models but also to provide the capabilities to come to an automated or semi-automated decision for access to data and creation of interoperable and reusable architecture solutions. The proposed framework will combine existing standard ontology vocabularies in order to achieve the description of a domain independent user profile. Also, it utilises ontology alignment/mapping techniques to inter-operate and integrate heterogeneous resources of context ontologies. Meanwhile this framework deploys ontology reasoning to on-the-fly context-aware privacy policy reasoning in order to improve awareness and control of users over the usage of his/her information. In addition, as an illustration of the proposed approach, this paper describes how the framework applies to a use case from the health domain.

The rest of the paper is structured as follows: Related work of privacy preserving systems is outlined in Section 2. Then,we describe our proposed framework in Section 3. The use case for evaluating the framework and a sequence diagram representing the interactions between users and privacy preserving unit of the framework is presented in Section 4. The next section mentions the technical methods and tools that are planned to be deployed for applying the framework in various use cases. Finally, we conclude this paper in Section 6 and present some future work.

2 Related Work

In recent years, many works have been done on the topic of privacy preserving for users data in personalized systems[31]. In particular, the access control model is the most frequently used method for defining privacy policies to let users control the access to their information. Role-Based Access Control (RBAC) [27, 2, 16] presents a model to manage roles that control the access to the users information. RBAC cannot provide dynamic access control because context-aware elements are not included [18]. On the other hand, Contextaware RBAC (C-RBAC) [29] does not ensure privacy protection and integrity because it does not consider the level of security between objects. Authors of [24] have analysed a large number of solutions that support C-RBAC. Context aware-task role based access control (CA-TRBAC) provides security services of user authentication and access control by context-aware security manager, and provides context-aware security services [10]. There are different approaches for specifying privacy policies using semantic modeling based on ontologies [11, 15, 25]. Ontology-based Access

Control model (Onto-ACM) [7] is a model of intelligent context-aware access for proactively applying the access level of resource access based on ontology reasoning and semantic analysis method. Also, Context-Aware PRIvacy preserving system Supervised by users (CAPRIS) [6], provides users with groups of policies that form profiles to protect their privacy in the environment in which they are located. Although the last two models address the need for dynamic access control, the privacy policies they provides are not capable to cope with uncertain and constantly changing context.

3 The Framework

As mentioned in the introduction, privacy-related decisions in personalised systems can be diverse with regards to different contexts and circumstances. For instance, a user who intensively shares his/her data such as diseases name and hospitalisation in an on-line health support community, may not wish to reveal any information about his/her hospitalisations to a health insurance application because of the fear of price discrimination. S/he is even more circumspect to share his/her data about either hospitalisation or diseases name in general purpose social networks of his/her friends due to anxiety of reputation undermining.

This section describes the framework proposed to dynamically help a user make a decision on disclosing his/her personal information to different stakeholders (simply assumed as a context) who requests to access the data. Fig.1 depicts the design and components of the proposed framework. The components of the framework can be categorised in three different groups: ***actors***, ***data resources*** and ***functional components***.

3.1 Actors

In this framework, we have defined two kinds of actors; different stakeholders who request to access to personal data is considered a ***Requester*** and the user of the personalised system who possess his/her data and has full authority to share them, named as a ***Data Owner***. The terms used for these actors are drawn from the Policy Based Management research domain [33].

3.2 Data Resources

Data resources are the places, which store data and knowledge. Since ontology is a knowledge model that not only represents a set of concepts and their relations within a domain but also facilitates the concepts instantiation, so we propose to use two separate data resources for each ontology, one for instances of concepts and other one for the knowledge model of the concepts. Accordingly,the user model is stored in the framework as ***User Ontology*** resource but user data is kept in ***Personal Data*** resource, which is out of the framework, to be secured from any unauthorised access in malicious way. Since the tendency of the data owner to share his/her information with any requester can be changed,

the framework assumes and models a requester as a partial context which can affect the system behaviour and stores the model in the ***Context Ontology*** resource and the data related with the same ontology in ***Context KB*** resource, respectively. There is another resource component for storage of primitive privacy policy set defined by Data Owner named as ***User-defined preference***. All of the data resource components are placed in and will be referred to as the ***Knowledge Base (KB)*** component of the framework. Meanwhile, all events and activities of the framework are recorded in a ***Logs*** resource.

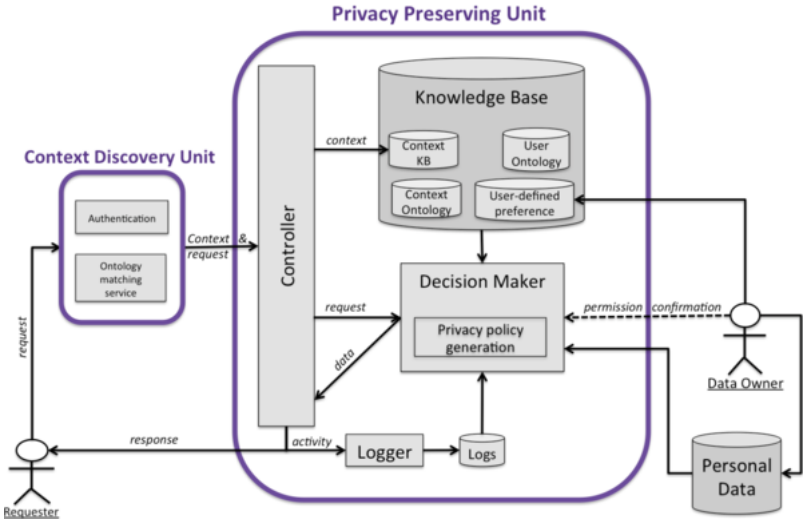

Fig. 1. The Framework

3.3 Functional components

The framework involves two general units named ***Context Discovery Unit (CDU)*** and ***Privacy Preserving Unit (PPU)***, which together are responsible for the main functionality of the system and each unit consists of different functional components. The ***Context Discovery Unit*** authenticates the requester in ***Authentication*** component and converts the native information of authorised requester to match with ontological model of context. If there is any heterogeneity in requesters information derived from other ontologies, the ***Ontology matching service*** would be deployed to align them with context ontology of the framework. ***Controller*** of PPU is the enforcement point of the framework, which intercepts the access request to information from CDU and makes a decision request as a SPARQL query to the ***Decision Maker*** component to obtain the access decision. Also, it is in charge of receiving the ontological

model of requester and stores it in appropriate data resource component of ***KB***. Another function of ***Controller*** is receipt of data from ***Decision Maker*** component in case of access permission and sharing with ***Requester***. ***Decision Maker*** evaluates access request against privacy policies before issuing access decision. This evaluation is done initially against predefined privacy policies by user. If the access request violates any of the policies, the request will be denied directly. Otherwise, ***Decision Maker*** will apply reasoning over the ontological models stored in ***KB*** to come to an access decision. ***Decision Maker*** reveals requested information to ***Controller*** if access decision is allowed. ***Logger*** component records all events and activities of PPU in ***Logs*** resource for learning the typical behaviour of framework in complicated reasoning methods later.

Inspired by [12], example vocabularies which may be used to implement the framework are FOAF, WebIDs and the Web Access Control (WAC) vocabulary in integration with existing infrastructure. The FOAF vocabulary [5] allows the description of domain independent user profiles and provides a container for other information from different domains. WebIDs [28] securely connect a user identity to the information in a user profile and can be used for authenticating a user. The Web Access Control (WAC) vocabulary [12] allows the user to authorise third party services for accessing different parts of his profile information by using an Access Control List (ACL).

4 Use Case

In order to illustrate how the proposed framework could assists user to make a decision on access request to his/her information by different stakeholders, in this section we outline a use case from Personally Controlled Health Record (PCHR) platforms. In PCHR, patients are intensively willing to engage with their health situation and disclose access to their private data include biographical, condition/disease, treatment, symptom and genetic information by third party experts/services in order to gain added value. This added value/benefit can be experienced through exchange with other patients having the same diseases or recommended clinical trials matching with their condition from domain experts. However, they would potentially be concerned to disclose data to other unauthorised requesters.

To help the user to assess and make a decision on access request in inherently uncertain contexts, our proposed framework can be applied to use run time ontology reasoning methods to reason over privacy policies, the context, the user's preferred general policies and the user model in order to propose a decision. In this use case the user profile would include: 1.A FOAF profile such as name, gender and birthday . 2. His/her private personal health record which contains his/her health condition, medications, laboratory results, hospitalisation and so on that can be described by using Clinical Document Architecture [19]. All this information is hosted by the user and along with his/her predefined policy preferences can be emended at any time by him/herself. Figure 2 explains communication and interactions between different components of the framework through

a sequence diagram when reasoning undertaken by the Decision Maker informs the patient, Data Owner, with regard to recommendations about permission to access request of his/her general practitioner (GP) to his/her hospitalisation record. Main steps of the diagram are commented as following:

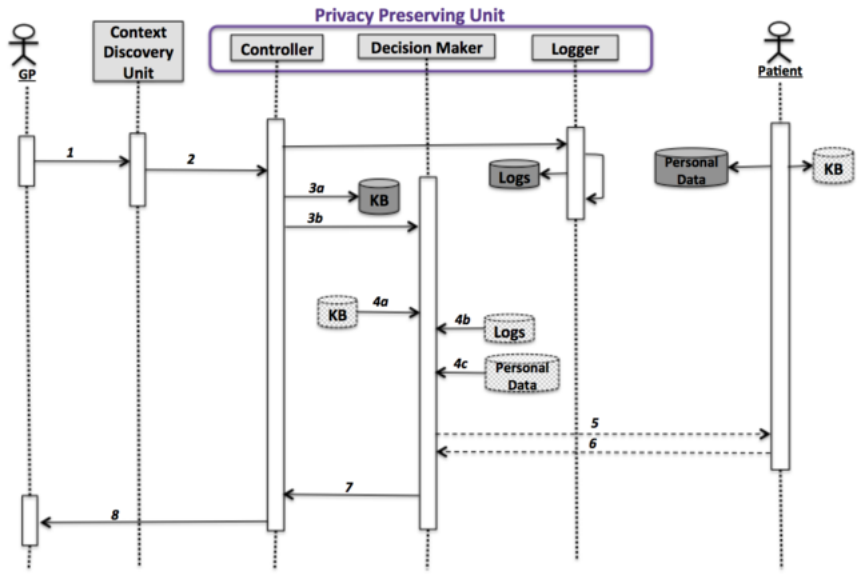

Fig. 2. The Sequence Diagram for the Use case

1. The GP sends his request of hospitalisation record to ***CDU***.
2. ***CDU*** authenticates GP as an allowed user of the system and aligned his former ontology with his ontological model according to FOAP+WebID vocabularies (if applicable) and sends them to ***Controller*** along with his request of hospitalisation record
3. Firstly ***Controller*** stores the ontological model of GP in appropriate resources at ***KB***. Then it creates a SPARQL query based on GPs data and his request of hospitalisation and hands over to ***Reasoner***.
4. Based on patients ontologies and his/her predefined preferences with regards to context ontology of GP, ***Decision Maker*** infers new privacy policies and evaluates them against his request of hospitalisation record. Previous activities of system, obtained from provenance log can also affect inference process of ***Decision Maker*** in some complicated cases.
5. The results of evaluation are sent to patient for obtaining his/her confirmation.
6. The final decision of patient on disclosing hospitalisation record sends back to ***Decision Maker***.

7. ***Decision Maker*** reveals hospitalisation record and access permission to ***Controller***.
8. ***Controller*** makes hospitalisation record accessible to GP.

5 Technical Approach

A typical policy-based management system includes a policy refinement process which is responsible to transform an abstract policy specification into low-level, concrete policies that can be enforced on the managed system. The main tasks of the refinement process are the following: Determine the resources that are needed to satisfy the requirements of the policy, Translate high-level goals into operational policies that the system can enforce and verify that the low-level policies actually meet the requirements specified by the high-level goal[26]. Difference over traditional policy refinement process of privacy preserving systems, the proposed framework primarily acts as an Attribute Based Access Control system (ABAC) [13, 14], where attributes associated with a given user is also engaged into the access decision. Meanwhile, it offers an enforcement point where proposes an access decision through dynamically reasoning over these user models along with the context, the users preferred general policies. To implement the framework, initially it is planned to build suitable ontologies according to the vocabularies mentioned in Section 3. In order to benefit from its consistency with the lower-level web standards and rich representational formalisms such as frame theory [21] and description logic [3], the research tends to express these ontologies in Web Ontology Language (OWL) [20] using Protege, the popular tool for ontology management developed at the Stanford University School of Medicine [23]. Ontology development exploits reusability of domain knowledge, although providing a unified model, that can effectively serve as a basis for semantic sharing and data integration, is not straightforward. To overcome such heterogeneity issues, various ontology alignment and mapping techniques will be deployed during this research. Lexical labels given to the elements of the ontologies [8], the resembling structural patterns in the ontology graphs [22] can signify as a most common source of evidence for ontology mapping which will be targeted through this research. Deploying ontology reasoning to improve awareness and control of users over the usage of his/her information specifically in personalisation contexts is one of the main contributions of this research. Explicit context is acquired from context sources directly, while implicit context is the additional information deduced from explicit context [32]. Multiple surveys [4, 32, 34, 35] provide overviews of ontology-based context reasoning approaches. Ideally, this research tends to apply two categories of reasoning tasks: 1) Ontology reasoning using description logic, which will use a restricted set of first order formulas for specifying a terminological hierarchy and 2) User-defined reasoning, which will infer a wide range of higher-level, conceptual context from relevant low-level context [9].

6 Conclusion

This research plans to introduce new ways of using semantic reasoning for privacy-preserving personalisation that take greater advantage of the capabilities of semantic models. To semi-automatically identify potential conflicts of usage data across different contexts, this approach models different contexts to know how data is used and provides reasoning that reasons across all of these models. Also, we have to reason through across ontology mapping because different data can be stored in different ways. This would enhance the state of the art approaches by empowering personalisation technologies to be more privacy sensitive yet it is still capable of operating of across different applications and contexts.

Acknowledgments. The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

References

1. Adjerid, I., Peer, E., Acquisti, A.: Beyond the privacy paradox: Objective versus relative risk in privacy decision making
2. Ahn, G.J., Sandhu, R.: Role-based authorization constraints specification. ACM Transactions on Information and System Security (TISSEC) 3(4), 207–226 (2000)
3. Baader, F.: The description logic handbook: Theory, implementation and applications. Cambridge university press (2003)
4. Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., Riboni, D.: A survey of context modelling and reasoning techniques. Pervasive and Mobile Computing 6(2), 161–180 (2010)
5. Brickley, D., Miller, L.: Foaf vocabulary specification 0.98. Namespace document 9 (2012)
6. Celdrán, A.H., Pérez, M.G., Clemente, F.J.G., Pérez, G.M.: What private information are you disclosing? a privacy-preserving system supervised by yourself. In: High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC, CSS, ICESS), 2014 IEEE Intl Conf on. pp. 1221–1228. IEEE (2014)
7. Choi, C., Choi, J., Kim, P.: Ontology-based access control model for security policy reasoning in cloud computing. The Journal of Supercomputing 67(3), 711–722 (2014)
8. Doan, A., Madhavan, J., Domingos, P., Halevy, A.: Learning to map between ontologies on the semantic web. In: Proceedings of the 11th international conference on World Wide Web. pp. 662–673. ACM (2002)
9. Ejigu, D., Scuturici, M., Brunie, L.: An ontology-based approach to context modeling and reasoning in pervasive computing. In: Pervasive Computing and Communications Workshops, 2007. PerCom Workshops' 07. Fifth Annual IEEE International Conference On. pp. 14–19. IEEE (2007)

10. Eom, J.H., Park, S.H., Chung, T.M.: A study on architecture of access control system with enforced security control for ubiquitous computing environment. Journal of the Korea Institute of Information Security and Cryptology 18(5), 71–81 (2008)
11. Gao, F., He, J., Peng, S., Wu, X., Liu, X.: An approach for privacy protection based-on ontology. In: Networks Security Wireless Communications and Trusted Computing (NSWCTC), 2010 Second International Conference on. vol. 2, pp. 397–400. IEEE (2010)
12. Hollenbach, J., Presbrey, J., Berners-Lee, T.: Using rdf metadata to enable access control on the social semantic web. In: Proceedings of the Workshop on Collaborative Construction, Management and Linking of Structured Knowledge (CK2009). vol. 514 (2009)
13. Hu, V.C., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R., Scarfone, K.: Guide to attribute based access control (abac) definition and considerations. NIST Special Publication 800, 162 (2014)
14. Hu, V.C., Kuhn, D.R., Ferraiolo, D.F.: Attribute-based access control. IEEE Computer 48(2), 85–88 (2015)
15. Iqbal, Z., Noll, J., Alam, S., Chowdhury, M.M.: Toward user-centric privacy-aware user profile ontology for future services. In: Communication Theory, Reliability, and Quality of Service (CTRQ), 2010 Third International Conference on. pp. 249–254. IEEE (2010)
16. Joshi, J.B., Bertino, E., Latif, U., Ghafoor, A.: A generalized temporal role-based access control model. IEEE Transactions on Knowledge and Data Engineering 17(1), 4–23 (2005)
17. Kobsa, A., Teltzrow, M.: Contextualized communication of privacy practices and personalization benefits: Impacts on users data sharing and purchase behavior. In: Privacy Enhancing Technologies. pp. 329–343. Springer (2004)
18. Li, N., Tripunitara, M.V.: Security analysis in role-based access control. ACM Transactions on Information and System Security (TISSEC) 9(4), 391–420 (2006)
19. Liu, H., Hou, X., Hu, G., Li, J., Ding, Y.Q.: Development of an ehr system for sharing-a semantic perspective. In: MIE. pp. 113–117 (2009)
20. McGuinness, D.L., Van Harmelen, F., et al.: Owl web ontology language overview. W3C recommendation 10(10), 2004 (2004)
21. Minsky, M.: A framework for representing knowledge (1974)
22. Noy, N.F., Musen, M.A.: The prompt suite: interactive tools for ontology merging and mapping. International Journal of Human-Computer Studies 59(6), 983–1024 (2003)
23. Noy, N.F., Sintek, M., Decker, S., Crubézy, M., Fergerson, R.W., Musen, M.A.: Creating semantic web contents with protege-2000. IEEE intelligent systems (2), 60–71 (2001)
24. Perera, C., Zaslavsky, A., Christen, P., Georgakopoulos, D.: Context aware computing for the internet of things: A survey. IEEE Communications Surveys & Tutorials 16(1), 414–454 (2014)
25. Riaño, D., Real, F., López-Vallverdú, J.A., Campana, F., Ercolani, S., Mecocci, P., Annicchiarico, R., Caltagirone, C.: An ontology-based personalization of healthcare knowledge to support clinical decisions for chronically ill patients. Journal of biomedical informatics 45(3), 429–446 (2012)
26. Rubio-Loyola, J., Serrat, J., Charalambides, M., Flegkas, P., Pavlou, G.: A methodological approach toward the refinement problem in policy-based management systems. IEEE Communications Magazine 44(10), 60–68 (2006)
27. Sandhu, R.S., Coynek, E.J., Feinsteink, H.L., Youmank, C.E.: Role-based access control models yz. IEEE computer 29(2), 38–47 (1996)

28. Story, H., Harbulot, B., Jacobi, I., Jones, M.: Foaf+ ssl: Restful authentication for the social web. In: Proceedings of the First Workshop on Trust and Privacy on the Social and Semantic Web (SPOT2009) (2009)
29. Tahir, M.N.: C-rbac: Contextual role-based access control model. Ubiquitous Computing and Communication Journal 2(3), 67–74 (2007)
30. Teltzrow, M., Kobsa, A.: Impacts of user privacy preferences on personalized systems. In: Designing personalized user experiences in eCommerce, pp. 315–332. Springer (2004)
31. Toch, E., Wang, Y., Cranor, L.F.: Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. User Modeling and User-Adapted Interaction 22(1-2), 203–220 (2012)
32. Wang, X.H., Zhang, D.Q., Gu, T., Pung, H.K.: Ontology based context modeling and reasoning using owl. In: Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on. pp. 18–22. Ieee (2004)
33. Westerinen, A., Schnizlein, J., Strassner, J., Scherling, M., Quinn, B., Herzog, S., Huynh, A., Carlson, M., Perry, J., Waldbusser, S.: Terminology for policy-based management. Tech. rep. (2001)
34. Ye, J., Coyle, L., Dobson, S., Nixon, P.: Ontology-based models in pervasive computing systems. The Knowledge Engineering Review 22(04), 315–347 (2007)
35. Ye, J., Dobson, S., McKeever, S.: Situation identification techniques in pervasive computing: A review. Pervasive and mobile computing 8(1), 36–66 (2012)