2011 Fifth IEEE International Conference on Semantic Computing

Preserving Privacy in Context-Aware Systems

Pramod Jagtap, Anupam Joshi, Tim Finin, Laura Zavala,
Computer Science and Electrical Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250 USA
pramod1@umbc.edu, joshi@cs.umbc.edu, finin@cs.umbc.edu, rzavala@umbc.edu

***Abstract*—Recent years have seen a confluence of two major trends – the increase of mobile devices such as smart phones as the primary access point to networked information and the rise of social media platforms that connect people. Their convergence supports the emergence of a new class of context-aware geosocial networking applications. While existing systems focus mostly on location, our work centers on models for representing and reasoning about a more inclusive and higher-level notion of context, including the user's location and surroundings, the presence of other people and devices, and the inferred activities in which they are engaged. A key element of our work is the use of collaborative information sharing where devices share and integrate knowledge about their context. This introduces the need for privacy and security mechanisms. We present a framework to provide users with appropriate levels of privacy to protect the personal information their mobile devices are collecting, including the inferences that can be drawn from the information. We use Semantic Web technologies to specify high-level, declarative policies that describe user information sharing preferences. We have built a prototype system that aggregates information from a variety of sensors on the phone, online sources, and sources internal to the campus intranet, and infers the dynamic user context. We show how our policy framework can be effectively used to devise better privacy control mechanisms to control information flow between users in such dynamic mobile systems.**

***Index Terms*—Privacy, social networking, mobile computing, policy, generalization**

I. Introduction

Phones, especially smartphones, are increasingly the most common gateway for people to access the information infrastructure and services available on the web. Smartphones are programmable devices that come with variety of powerful embedded sensors such as GPS, accelerometers, microphones, cameras, gyroscopes and others. These sensors can be used to collect information about users and their surroundings in terms of location, motion, temperature, other people in the vicinity and so on. This information allows us to infer a user's context, and has the potential to change the way people interact with the information infrastructure. For instance, many new phone based applications enhance the social networking experience with additional social dynamics that emerge from allowing users to interact relative to location and time. Location awareness is one important aspect of a context-aware system. However, context encompasses more than just the user's location, because other things of interest are also mobile and changing [23]. Other important aspects include the ambiance, resources and people nearby, and the activities in which they are engaged. The rise of online social networking systems along with recent improvements in mobile technology, smartphones, and sensor networks present a unique opportunity for context-aware systems.

A very important issue in such applications is that of privacy. Unfortunately, this is often overlooked, or only superficially discussed. While context aware systems and applications face security threats similar to other distributed and mobile applications, privacy and security aspects are more prominent due to the sensitive nature of context information. The existing controls in context-aware systems are based on the static information (particular users or groups), and predetermined. In fact, on many smartphone systems, the user is asked to make the decision to share sensor information such as location at install time of the application. These controls are not adequate for context-aware systems, since context is dynamic, and itself determinative of what data can be shared. For instance, it might be acceptable to share location and accelerometer information generally, but not when one is exceeding the posted speed limit while driving! This environment calls for better access controls with finer control over the context data to preserve user privacy. The user needs to be in control of the release of her personal information at different levels of granularity, from raw sensed data to high level inferred context information.

What is required is a privacy system that allows a user to specify policies to control the information flow from sensors based on the changing context of the users. None of the existing models allows users to specify information sharing policies based on such information. In this paper, we present a semantically rich, policy-based framework to constrain the information flow in a context-aware system. It uses an OWL ontology to represent dynamic aspects of context-aware system, and and a combination of OWL-DL and Jena rules specifying the policy to perform reasoning. The dynamic elements such as user context, requester context, temporal restrictions and context restrictions are taken into consideration along with static information like profile information and social relationships before making access control decisions to sensed and inferred context data. The framework also allows for automatic generalization of context attributes in order to protect user privacy. This allows users to share context information on different levels of accuracy along with different types of information, from raw sensed data to inferred context. The framework can be extended and incorporated in existing social networks including location-based mobile social networks. We have validated our architecture in an on-campus context-aware prototype system that aggregates

978-0-7695-4492-2/11 $26.00 © 2011 IEEE
DOI 10.1109/ICSC.2011.87

information from a variety of sensors on the phone, online sources, and sources internal to the campus intranet, and infers the dynamic user context. We show how our policy framework can be effectively used to devise better privacy control mechanisms to control information flow between users in such dynamic mobile systems.

II. RELATED WORK

While context-aware systems have been studied for a long time, the focus has been mainly on the location and activity inference. Recently research about privacy controls in these systems has received significant attention. AnonySense [24], a privacy-aware architecture for collaborative pervasive applications that use mobile sensing. Mobile sensor data is anonymized before its use by any of the applications. Project Aware Home [20] captures, processes and stores data (collected by sensors) about home residents and their activities. It uses access control mechanism based on Role-based Access Control (RBAC) by defining environment roles similar to subject roles of RBAC and it is used to capture security-relevant aspects of the environment in which an application executes. Context Privacy Service (CoPS) [22] describes the design and implementation of a privacy service which control how, when and to whom you could disclose a user's context information. Using the end-user survey and results of other research groups, it has identified requirements for flexible and efficient privacy service. This system is most closely related to our work. However, it does not handle context-dependent privacy policies, which can be specified by users on dynamic context data. There has been a lot of work done to develop access control frameworks. Rei [18] is a policy language designed for pervasive computing applications. It has been used to built a security framework [17] that addresses the issues of security for web resources, agents and services in the Semantic Web. Rein (Rei and N3) [16] is a distributed framework for describing and reasoning over policies in the Semantic Web. It supports N3 rules [4] for representing interconnections between policies and resources.

III. SYSTEM ARCHITECTURE

The major components of this system are client devices, server side modules and the Internet services that provide social media. The client devices are location aware smartphones. These client devices as well as the server side modules contain a user profiles repository, a privacy control module and content preferences. The server side also contains a content aggregator, a learn and share module and a privacy control module. The content aggregator combines social media like event updates, photos, and videos from Internet services like YouTube, Flickr, Facebook or university information portals. The learn and share module infers the user's dynamic context using sensor data collected by a variety of sensors on the phone, the information from the content aggregator and online sources such as user's calendar. The inferred context is shared with corresponding client device so that the device along with the server can handle further context sharing queries from other clients. The requester queries are passed through the privacy control module to constrain the information flow and hence to protect the user privacy. The privacy control module provides the access control mechanisms and aids in controlling the information flow within system. On the client device, it enables privacy sensitive and resource sensitive reasoning over sensed data along with privacy enforcement between peer devices sharing contextual information.

The information sharing occurs in three different ways as (i) context information sharing between the client devices, (ii) sensor data sharing between a client device and the server, and (iii) context and sensor information sharing between a client device and the server. The information sharing is controlled by the privacy control module in order to preserve user privacy. Our system uses the Jena Semantic Web software tools [6] on Android devices [21] to perform reasoning and constraining sensed data flow according to user-defined privacy policies. We will focus our discussion on our privacy mechanisms and the relevant system components which have most direct influence on the information flow in the system.

A. Privacy Related Components

The privacy control module aims to protect user privacy by performing reasoning over the context. It deals with the resource to be protected, the owner of a resource and the requester who wants to access it. More abstractly, it accepts an RDF triple (*U, C, Q*), where *U* is the identity of the requester, *C* is the requester's context (expressed as RDF triples in our ontology), and *Q* is the query pertaining to context information. The module has access to owner's profile information and the group information along with specified privacy policies. It enforces owner's privacy policies using static information about the owner as well as dynamic information observed and inferred from her context. It consists of, (i) a set of ontologies for describing activities/context, policies and access requests, (ii) the knowledge about the owner, (ii) the privacy preferences, and (iv) a reasoning engine that accepts requests and performs the reasoning.

1) Context ontology: Our context ontology [13] captures the user location and surroundings, the presence of other people and devices, and the inferred activities in which they are engaged. We adopt description logics (DL), specifically OWL (Web Ontology Language), and associated inferencing mechanisms to develop a model of context. The context ontology captures the semantic notion of context in a mobile context-aware system. Using the ontology, each device contains a declarative knowledge base with semantically rich information about user's information, activities, inferences, and further contextual information. The knowledge base aligns with the context ontology which defines the key context concepts used for making access control decisions. The ontology supports the generalization of context infomation by creating hierarchical models for different aspects of context viz. activity and location. It aims to protect user privacy and helps users to have finer control over their contextual information and hence allows sharing of context information on different levels of

granularity. The following section describes the generalization in detail.

2) Generalization: Generalization involves replacing a value with a less specific but semantically consistent value in order to protect user data privacy [25]. Our system uses context-data generalization to allow information sharing on different levels of granularity.

Location Generalization

The location information is sensitive and hence it should be shared with legitimate set of people as decided by user. A privacy policy such as "*Share my location with teachers on weekdays from 9am-5pm*" allows a group of people defined by the user as "teachers" to access user's GPS location between the specified hours. In many cases the user might be interested to share the location but not exactly. For instance, the user can have privacy policy like "*Share my building-wide location with teachers on weekdays from 9am-5pm*" which allows location sharing but at the same time it does not reveal the exact location. The system will share the building names with "teachers" rather than exact GPS position of user. In order to support the location generalization, our ontology uses hierarchical model for location. Location is a super class of Point, Room, Building, City and State classes. The Point class is used for denoting the GPS coordinates whereas Room and other subclasses are used to denote different levels of abstractions for the location. The transitive "Part_Of" property creates a location hierarchy based on some simple axioms like "Room is a part of Building". The reasoning engine will use this ontology to infer the different relations existing between instances of these subclasses.

Activity Generalization

Along the lines of location generalization, we present activity generalization for allowing users to share different descriptions of their current activity to different set of requesters. Consider a policy like "*Share my activity with friends on weekends*"; this will share user's current activity to the people belonging to a "Friend" group. In many cases, the user is willing to share more generalized activity rather than precise one. For instance, if a user is attending a confidential "project meeting" then she might want to share it in a more generalized way as "working" or simply as a "meeting". In such cases, the user clearly needs to obfuscate certain pieces of activity information. We permit the user to differentiate between the set of activities by attaching a confidentiality parameter e.g. visibility option. The visibility option specifies the sensitivity level of activity from the user perspective. Our ontology supports different visibility options such as Public, SemiPublic, Private, SuperPrivate. The Public option implies that the corresponding activity is least sensitive whereas SuperPrivate option indicates that the activity is at most sensitive. The SemiPublic and Private are listed in increasing order of sensitivity. These visibility options can be used to share more generalized/less sensitive/public activities instead of specific/sensitive/private ones. The activity generalization is supported by using a hierarchical model of activities.

3) Reasoning Architecture: The reasoning engine handles the requester queries and performs reasoning for access control decisions. Our system uses the Jena Semantic Web framework[6] for performing the reasoning over context data. Jena inference system allows the support of various inference engines or reasoners. In our system, the reasoning engine uses the context ontology, users context information and group information along with the user-specified privacy rules to generate an inference model. This inference model is used for responding to the requester queries.

4) Knowledge about the user: A user can create her personal profile and put in information like name, email address, hobbies and interests and can manage different groups of her friends. Apart from that, the system has dynamic knowledge information about user's context including her current activity and location. Our context ontology defines the entities required to represent a user information in addition to the FOAF [11] vocabulary. This knowledge is specified using N3 [4] in our system. All the attributes in a user's personal profile as well as data sensed by mobile devices are considered as resources to be protected.

5) Privacy preferences: Privacy preferences are access control rules that describe how a user wants to share which information, with whom, and under what conditions. All the privacy preferences are represented as N3 rules in the system. The user can specify privacy preferences to share personal information based on her (i) profile and context information, (ii) requester's context information, (iii) temporal and spatial restrictions and (iv) generalization of context data. For example, the user can have privacy policy like *"Share my activity with friends all the time except when I am attending a lecture"*. which emphasizes on user's context and group information. A privacy policy like *"Share my context information with anyone who is attending same class as me"*. considers user's context and requester's context before making information sharing decision is shown in Table I whereas a privacy policy like *"Do not share my context with anyone during super-private activities"* utilizes activity generalization and ensures activities with "super-private" property are not shared with anyone.

The user can specify privacy policies to protect the sensed data. Before the data is collected from sensors or whenever there is a request for sensed data, the privacy control module evaluates the user-defined privacy policies and decides which sensor data can be collected. Only allowed sensors' data is collected and sent to the server for further context inferring. For instance, a user can have policy like *"share GPS coordinates on weekdays from 9am-5pm only if he is in office"*. Table II shows it's corresponding Jena rule.

In our system, negative permission "prohibited" always takes preference over positive "permitted" permission. It means, if the output inference model has both "prohibited" and "permitted" values for predicate such as "contextAccess" then system assumes information sharing permission is "prohibited". Along with user-level privacy policies, our system has provision for system-level policies. The context-aware systems are used by individuals to organization and from

TABLE I
POLICY TO SHARE CONTEXT INFORMATION BASED ON OWNER'S CONTEXT AND REQUESTER'S CONTEXT.

```
[ShareContextRule:
    (?requester ex:requester "True")
    (?requesterActivity platys:is_performed_by ?requester)
    (?requesterActivity platys:occurs_at ?requesterPlace)
    (?requesterPlace platys:has_location ?requesterLocation)
    (?requesterLocation platys:part_of ?requesterRoom)
    (?requesterRoom rdf:type platys:Room)
    (?user ex:systemUser "True")
    (?userActivity platys:is_performed_by ?user)
    (?userActivity platys:occurs_at ?userPlace)
    (?userPlace platys:has_location ?userLocation)
    (?userLocation platys:part_of ?userRoom)
    (?userRoom rdf:type platys:Room)
    equal(?requesterRoom, ?userRoom)
    equal(?requesterActivity, ?userActivity)
    equal(?userActivity, platys:Listening_To_Lecture)
->
    (?requester ex:contextAccess ex:userPermitted)
]
```

TABLE II
POLICY TO SHARE GPS COORDINATES. IT STATES THAT GPS DATA CAN BE SHARED ON WEEKDAYS FROM 9AM-5PM ONLY IF USER IS IN OFFICE.

```
[ShareGPSRule:
    (?request ex:hasRequester ?requester)
    (?request ex:requestTime ?localTime)
    (?requester ex:systemUser "True")
    (?localTime time:dayOfWeek ?day)
    ge(?day, 1) le(?day, 6)
    (?localTime time:hour ?hour)
    ge(?hour, 9) le(?hour, 17)
    (?user ex:latitude ?latitude)
    (?user ex:longitude ?longitude)
    Equal(?latitude, ?officeLat)
    Equal(?longitude, ?officeLong)
->
    (?requester ex:canAccessGPSCoordinates "True")
]
```

social-networking application to military domains. In case of military domains or organizations, the user may not be the sole owner of client device and there is a strong need of robust security mechanisms. Such organizations can have system-level privacy policies which should always override user-specified policies. The system-level policies should be defined by the system-administrator to ensure that the sensitive resources are always protected from illegitimate access. For instance, system-administrator can define a system-level policy like *"Do not share the user's context with anyone if she is inside a NSA building 2"* which ensures that user's context is not shared with anyone if she is inside NSA building numbered as 2. Our system ensures that user-level permissions are always overridden by system-level permissions.

IV. SYSTEM IMPLEMENTATION AND EVALUATION

Our prototype implementation uses smartphones such as iPhone or Android phone as client devices. For the detailed description of server side modules please refer to [8] and [14]. The client device can send different types of access requests to another client device or the server. The main distinction between the access requests made by a client device to a peer device and to a server is that the latter request contains a specific userId. This userId is used to retrieve specific user's information. These requests can be detailed context access requests or specific resource requests like location or activity access requests. The access request is processed by the policy framework present on client device or server.The access query is processed by the policy framework and its result is shown to the requester with valid accuracy level. The system considers contextual information and sensor information as the resources that changes dynamically for the user, and has provided mechanisms to specify more expressive policies to control its sharing. The users can create policies by using Policy Editor interface. It is a Web interface which can be used by users from client device to specify and edit privacy preferences. They can specify access control rule as - 'who' by selecting friends or groups of friends, 'what' by selecting resources such as location or activity, 'conditions' by selecting allowed days of the week or specifying the allowed time range during day or by specifying region on the map as sensitive. Users can also specify allowable type of activity like sleeping, eating, working, chilling. The policies are created and stored in N3 format on both server and client sides in persistent memory and reloaded when required by reasoning engine.

A. System Evaluation

The goals of evaluation were (i) to see if the system satisfies a basic criteria by allowing access from privileged user and restricting illegal user, (ii) to test whether the actual computing time of reasoning over mobile devices is acceptable and (iii) to determine how it scales with different size of user information like number of users in group list. The system behaved as expected by allowing information access to privileged users and denying access to illegal users as per user-defined privacy rules. Here, we define a privileged user as a requester who is allowed to access user's context as per user-specified privacy rules whereas other's are modeled as illegal users.

We have evaluated the system performance in terms of reasoning time taken for the requester query. It is measured when the access requests are made to server PC and to the android client device. To evaluate scalability of the system, we varied the number of users in group list and noted the time taken (reasoning time) by the system to provide access levels for the requester. Figure 1 shows the results of evaluation where the obtained values are average of several computations. It describes the growth of reasoning time (in milliseconds) against number of users in the group list. It clearly shows that reasoning on mobile devices can be done without any scalability issues and it can be efficiently used to enforce privacy over sensed and contextual data.

V. CONCLUSION AND FUTURE WORK

Our mobile devices are becoming the dominant way we communicate with people, access information, and consume

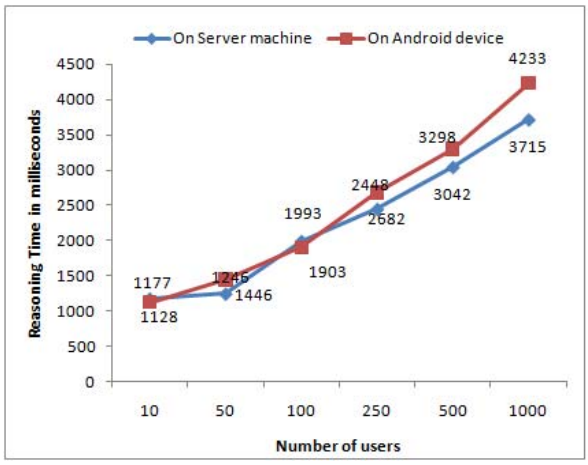

Fig. 1. Reasoning time (in milliseconds) for different number of users in owners group list.

services. As they become more intelligent, they can and will model our interests, activities and behavior in order to understand our current context and using it, better serve our needs. When appropriate, aspects of this learned context may be shared with other devices in order to collaborate and provide enhanced service. This development introduces a strong need to allow users greater control of what information is shared with who and with what level of detail.

We described a policy based framework to control information flow in collaborative context aware application. It allows users to specify a rich suite of privacy preferences that consider the static and dynamic knowledge about user, along with generalization rules to regulate the accuracy of results. Protected resources can be activities, location information, or media such as photos, videos posted by participants of the social network. We showed some example policies that state of the art systems do not support. Our privacy mechanisms constitute a baseline that can be extended and incorporated by any of the existing social networks including location based mobile social networks. We plan to extend the prototype implementation to address the engineering challenge of scalability. We plan to carry out user studies to evaluate the utility of the proposed privacy control mechanisms. We also plan to address the issues of incorporating incentives to allow for even more flexibility in the definition of policies for context-dependent release of information.

ACKNOWLEDGMENT

The research described in this paper was partially supported by the National Science Foundation (award 0910838) and the Air Force Office of Scientific Research (MURI Grant FA9550-08-0265)

REFERENCES

[1] A. Acquisti and R. Gross. Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. In *Proceedings of 6th Workshop on Privacy Enhancing Technologies*, 2006.

[2] D. Beckett. Turtle - Terse RDF Triple Language. Technical report, 2007.

[3] Tim Berners-Lee. Cwm - a general purpose data processor for the semantic web.

[4] Tim Berners-Lee and Dan Connolly. Notation3 (N3): A readable RDF syntax. Technical report, 2008.

[5] Tim Berners-Lee, Dan Connolly, Eric Prud'hommeaux, and Yosi Scharf. Experience with n3 rules. In *Rule Languages for Interoperability*, 2005.

[6] Jeremy J. Carroll, Ian Dickinson, Chris Dollin, Dave Reynolds, Andy Seaborne, and Kevin Wilkinson. Jena: implementing the semantic web recommendations. pages 74–83, New York, NY, USA, 2004. ACM.

[7] Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday, and Christos Efstratiou. Developing a context-aware electronic tourist guide: some issues and experiences. In *CHI*, pages 17–24, 2000.

[8] Audumbar Chormale. Constraining information flow in social networks with privacy policies. Master's thesis, University of Maryland, Baltimore County, 2009.

[9] Jon Doyle. Truth maintenance systems for problem solving. Technical report, Cambridge, MA, USA, 1978.

[10] Catherine Dwyer, Starr R. Hiltz, and Katia Passerini. Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. In *Proceedings of the Thirteenth Americas Conference on Information Systems (AMCIS)*, 2007.

[11] Mike Graves. FOAF: Connecting People on the Semantic Web.

[12] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, WPES '05, pages 71–80, New York, NY, USA, 2005. ACM.

[13] Pramod Jagtap. Context ontology, 2011. http://ebiquity.umbc.edu/_file_directory_/resources/317.txt.

[14] Pramod Jagtap. Privacy preservation in context-aware systems. Master's thesis, University of Maryland, Baltimore County, 2011.

[15] H. Jones and J.H. Soltren. Facebook: Threats to privacy. ethics and the law on the electronic frontier course, 2005.

[16] Lalana Kagal and Tim Berners-lee. Rein : Where policies meet rules in the semantic web. Technical report, Laboratory, Massachusetts Institute of Technology, 2005.

[17] Lalana Kagal, Tim Finin, and Anupam Joshi. A policy based approach to security for the semantic web. In *2nd International Semantic Web Conference (ISWC2003)*, September 2003.

[18] Lalana Kagal, Tim Finin, and Anupam Joshi. A policy language for a pervasive computing environment. In *IEEE 4th International Workshop on Policies for Distributed Systems and Networks*. June 2003.

[19] Lalana Kagal, Chris Hanson, and Daniel Weitzner. Using dependency tracking to provide explanations for policy management. In *Proc. IEEE Workshop on Policies for Distributed Systems and Networks*, pages 54–61, Washington, DC, 2008. IEEE Computer Society.

[20] Cory Kidd, Robert Orr, Gregory Abowd, Christopher Atkeson, Irfan Essa, Blair MacIntyre, Elizabeth Mynatt, Thad Starner, and Wendy Newstetter. The Aware Home: A Living Laboratory for Ubiquitous Computing Research. volume 1670, pages 191–198. 1999.

[21] Lorecarra. Androjena : Jena android porting, 2009.

[22] Vagner Sacramento, Markus Endler, and Fernando Ney Nascimento. A privacy service for context-aware mobile computing. In *Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks*, pages 182–193, 2005.

[23] Bill Schilit, Norman Adams, and Roy Want. Context-aware computing applications. In *In Proceedings of the Workshop on Mobile Computing Systems and Applications*, pages 85–90, 1994.

[24] Minho Shin, Cory Cornelius, Dan Peebles, Apu Kapadia, David Kotz, and Nikos Triandopoulos. AnonySense: A system for anonymous opportunistic sensing. *Journal of Pervasive and Mobile Computing*, 2010.

[25] L. Sweeney and Latanya Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10:2002, 2002.

[26] Roy Want, Andy Hopper, Veronica Falcão, and Jonathan Gibbons. The active badge location system. *ACM Trans. Inf. Syst.*, 10:91–102, January 1992.

[27] D. A. Waterman and F. Hayes-Roth, editors. *Pattern-Directed Inference Systems*. 1978.