

Classification of Privacy Management Techniques in Pervasive Computing

Stan Kurkovsky, Oscar Rivera, Jay Bhalodi

Central Connecticut State University
1615 Stanley Street, New Britain, CT 060505, USA
kurkovskysta@ccsu.edu

Abstract. Context awareness, the ability to adapt to the needs of each user, is a fundamental property of pervasive computing systems. Context information is created by tracking the actions and collecting real-time user data, such as location or body temperature. As the system collects more information about its users, the quality of its context-aware services increases, and so does a potential threat to the users' privacy if the context information is compromised. This paper surveys a wide range of existing pervasive computing systems that address this complex problem, and provides a generalized classification of privacy management techniques used by these systems. A comparative analysis of the surveyed systems and their privacy features is also presented.

Keywords: Privacy, security, management, pervasive computing.

1 Introduction

Since the earliest days of research in pervasive and ubiquitous computing, preservation of user privacy has been the subject of an active discussion [2,9,51]. On the one hand, a pervasive computing environment (PCE) needs to collect a substantial amount of information about its users to be able to adapt and respond to their demands without requiring much explicit user interaction. On the other hand, the more information a system has about its users, the greater is the potential threat to the users' privacy.

It is a core design feature of most PCEs to preserve the anonymity of the clients (users, services, or smart devices) interacting with service providers. However, an inherent contradiction lies in the fact that the service provider cannot fully trust the client that does not reveal its true identity. Without such a trust, the service provider cannot ascertain that the client has a rightful access to the requested resource or service. As a result, there is a perceived conflict between privacy and the technologies that enable the features of a PCE dealing with sensing and storing of user-related information. Personal privacy, in the aspects related to the information about the user's identity, past and current activities, and location, may be challenged because these kinds of information are typically among the data collected and stored by the sensors and services employed by most PCEs [30].

Large amounts of data potentially containing personal or personally identifiable information may become a serious temptation for misuse by a PCE service provider

or a third party intruder. Protection of such information as personal identity, location and past activities becomes more difficult due to the inherently distributed nature of PCEs and a continuous influx of newly collected data.

The main objective of this paper is to review and classify a broad range of privacy management techniques, methods and tools implemented by various existing and proposed pervasive computing systems and architectures. This paper is organized as follows. The background and motivation of this study is discussed in Section 2. Section 3 describes the identified privacy management techniques and illustrates them with the implementation details of one or more existing system. Section 4 concludes the paper with a summary.

2 Background and Motivation

Research literature provides a wide range of definitions of privacy in PCE: “control over information disclosure” [17], “privilege of users to determine for themselves when, how, and to what extent information about them is communicated to others” [29], and “ability of an individual to control the terms under which their personal information is acquired and used” [21]. A detailed discussion of many broad aspects of privacy in various contexts can be found in [5,9,18,38,39].

In PCEs, users may find an increasing amount of obstacles to maintain their privacy because of the growing volume of potentially identifiable data collected by the systems [1]. Data collected by perceptual interfaces that are capable of recognizing users, their gestures and facial expressions, may become a potential threat to the user privacy if it is somehow made accessible to a third party. Finally, some users may be fearful of compromised privacy and not be willing to accept and use a system that tracks and collects information about them, even though this data by itself may pose no privacy threats [1].

While there is a significant amount of research in pervasive computing aimed at design and implementation of privacy management tools and techniques, their practical usability and acceptance remains an important challenge [14]. In 2003, Computing Research Association identified the ability to “give computer end-users security they can understand and privacy they can control” in the “dynamic, pervasive computing environments of the future” as one of the four major research challenges of trustworthy computing [20].

Recently, a number of studies explored social implications of pervasive technologies and their impact on the changing perceptions of privacy [3,24,35]. As the technologies enabling PCE become ever more miniaturized and further blend with the surroundings, the users may not recognize or feel their presence and lose the awareness of the fact that some or all of their actions may be monitored and recorded by a PCE. Continuous collection of information about the users will inevitably expose the details of their personal traits: patterns of behavior, driving and walking routes, current location, shopping preferences, likes and dislikes, social associations, etc. Most studies suggest that some users may be willing to expose certain information about their behavior and actions in order to enable the adaptive nature of a PCE; at the same time, they may prefer to establish some boundaries and let other aspects of their

activities remain private. Users also may prefer to have the tools to limit the amount or granularity of the collected information, or to stop the monitoring altogether if they decide to do so [3]. Although the very notion of pervasive computing suggests that the technology should become invisible and blend into the environment, some users prefer to know that the technology actually is there [24].

Among many real world application domains where pervasive computing systems have been implemented, healthcare is one area in which privacy plays the most prominent role [23,44,53]. Hospitals, clinics and emergency rooms are a perfect testbed to implement a PCE: information services provide real-time data about patients, their medical history and treatment records; this information must be treated with confidentiality, yet available anytime to authorized users. Healthcare environment requires a high availability of information services, constant coordination among colleagues, and rapid response to emergencies [53]. A healthcare-oriented PCE must support a high degree of mobility for its collaborating users, who must have real-time access to the patient data protected by strong privacy safeguards. Depending on the specialization of a medical professional and their role in the treatment of a given patient, they may have access to different areas of records of that patient. In this application domain, privacy support may be further enforced by various legislative acts; in the US, protection of private medical records is mandated by the Health Insurance Portability and Accountability Act (HIPAA). Implantable medical devices used in a healthcare-oriented PCE present an even wider range of privacy-related challenges that are unique to this domain [23].

3 Privacy Management Techniques in PCE

In this section, we describe a number of privacy management techniques used in the design and implementation of existing pervasive computing systems and architectures. Due to space limitations, we omit a detailed discussion of each individual system; instead in each section describing a particular privacy management technique, we discuss the implementation of the corresponding technique by one or more system or PCE architecture.

3.1 Access Rights and Policy Management

It is often a challenging problem to control access to confidential information in a PCE [16,25,32]: users requesting access could lack knowledge about which access rights may be required, access control should enable easily adjustable and context-sensitive access rights, there may be a multitude of diverse information services making the management of access rights essential. Users may interact with many different smart devices and providers in order to obtain their services. Since these devices and services could be malicious or genuine, the user privacy may be compromised.

RAVE is a media space designed to support people who are distributed geographically, but work together on common tasks. Participants use video cameras, monitors, microphones and speakers, and can choose from a number of ways to

distribute their audio-video content or receiving the same from others in the media space [9,40]. The users of RAVE can control who can connect to them and what kind of connection is allowed. Feedback provided by this system alerts the users of the type of data being sent and who is allowed to access such data. In addition, one of the main benefits of RAVE is that it provides a privacy control at all levels of specificity, allowing users to make early decisions about granting permissions for specific kind of service to specific users.

The Trusted Platform Module (TPM) is a hardware solution for mobile computing devices that can provide common security architecture for pervasive environments [7]. TPM implements access rights without revealing user identities to external parties. A TPM-enabled system can authenticate a mobile user by issuing a challenge to the device, which is then signed by the access requestor. This technique provides robust security in the possible event of a third party stealing non-shareable information. The outside user will not be able to authorize the operation signature and thus not be able to enter a secure PCE.

The Privacy Awareness System (PawS) offers a set of tools that allows users protect their privacy while helping others respect it [37]. A user in PawS can specify the capabilities disclosed in the privacy policy such as who can update or delete the data versus who can only view the data. Thus, when a user enters any environment, a privacy proxy checks these policies against user's predefined privacy preferences. Services can collect information and users can utilize these services if the policy agrees, otherwise the user may not choose to use the service.

Privacy Violation Avoider (PriVA) is a privacy-aware PCE model [4], aimed at avoiding information leaks while sharing resources and information among the users. PriVA has built-in generalized policies for sharing resources. It has a default policy for certain resources and these policies cannot be changed to reduce the workload of the model. If a user does not want to share the resource they can tag for resource as non-sharable, rather than applying complex customized policies. Although PriVA provides a default policy for each resource, users might want to modify this policy. User can choose additional policies from the list of the customized policies provided in the model, which provides some policy flexibility. Different documents can have different policies, for example in teacher's handheld; the document containing grades have different level of importance versus the document with a list of reading materials.

Context-aware access policy management has been successfully implemented in the MOSQUITO project [36], which supports security requirements policies. Policy rules are expressed in XML and typically capture such context information as location and proximity. MOSQUITO's policies are aimed at evaluating trust between a consumer and a provider of a service in a PCE. The framework provides the flexibility to choose the implementation details of location-based trust evaluation policies, which can be customized to meet the application-specific requirements.

PerGym is a pervasive system for a gym that provides personalized services based on privacy-sensitive context information [46,47]. PerGym aggregates context information from distributed heterogeneous sources. Using a set of distributed policies, this information is aggregated by a context-aware privacy module, whose main responsibility is to enforce the system-wide compliance with a set of privacy policies set by each individual user.

Daidalos (Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimized personal Services) is an infrastructure for enabling privacy in a pervasive environment [13,45,49]. Daidalos provides a Privacy Negotiation Manager, which is involved in privacy policy definition, structures and relations to personal information of the users. It also creates a negotiation protocol within the agent software where statements of privacy policy are exchanged between peers.

3.2 Classification of Resources

Mobile devices carried by the users of a PCE may contain a wide range of information, not all of which may be shareable. It could include personal names, phone numbers, addresses, appointment schedules, to do lists, bank accounts, as well as official or business information regarding customers, account holders, etc. Some elements of this information may be confidential and sensitive; others may be open for access. In an environment that supports resource sharing, unwanted parties could retrieve the confidential information causing information leakage and leading to the violation of user privacy. One method to assure the users that resource sharing will prevent such confidential information leakage is to provide a mechanism to classify all information resources by their degree of “share-ability.”

PawS [37] provides the encoding of privacy policies into machine-readable XML to classify data. This structure indicates the users with whom an item of information may be shared.

PriVA [4] uses a resource manager that maintains a list of all resources available on each device. PriVA’s TagR module tags these resources as shareable or non-shareable. This model provides the users the flexibility to share resources as they want, whenever they want. PriVA’s ClsR module can be used to categorize all resources into different groups, such as images, videos, or text, adding to the simplicity of this model.

3.3 Data Persistence Control

In pervasive environments, information can be automatically collected and stored over a long period of time. Personal data is intrinsically valuable; it must be protected indefinitely against misuse and unauthorized disclosure. This task can be complicated by the fact that storage may be distributed among many systems and their components. Similarly to the ability to control the dissemination of information, the ability to control the persistence of data is also a fundamental property of privacy. Pervasive environments must provide tools to control the disclosure of collected and stored data. This can be done by placing time constraints on the data as defined by the owner or a service provider. A time to live attribute used in the Capability-Based Privacy-Preserving Scheme [34] or a restricted lifetime of access rights implemented in Hierarchical Identity-Based Encryption [27] could ensure that data does not exist beyond an established period of time. By limiting the data storage period, the risk of data leakage can be minimized.

PawS [37] uses a lifetime attribute for data that can be set by the owner, but it is the responsibility of a service provider or a user to observe that attribute. However, this framework stops short of providing the capability to virtually shred the data to satisfy the owner's privacy preferences. Such a feature can be the best tool to ensure confidentiality and privacy in pervasive environments since the capability to control information also implies the capability to extinguish private information. Therefore, PCE users should be in control of the lifespan of the information related to all private and personal activities.

3.4 Granularity Awareness

Information accessible in a PCE, such as time and location, can be grained down to details with different levels of privacy requirements. At the same time, there is a clear need to maintain the balance between the reasonable resolution of information the users are willing to provide and the amount of information needed by a particular service in a PCE. For example, some location-based systems may not function properly if they receive updates on a scale of kilometers and hours, as opposed to meters and seconds [10]. Similarly, privacy granularity restrictions used in SPARCLE [14] provide the ability to control database access down to a single field in the database, which may significantly hamper the functionality of the application on a large scale.

In Hierarchical Identity-Based Encryption [25,27], context information may be transformed along the way from the sender to receiver by decreasing its granularity level and therefore removing the details of information, for which the receiver may not have sufficient access rights. Additionally, information access requests may be evaluated and denied based on the granularity of the requested information. It is up to the owners of information to define the rules specifying the levels of granularity of the information they own and the associated levels of privacy.

The Publish/Subscribe Substrate model [43] is event-based; it uses three kinds of granularity levels to provide enhanced security and privacy guarantees to the users. System granularity assigns system events into categories that may or may not be received by third parties. Event type granularity determines whether subscribers of the system can receive events of a given type. Matching set granularity determines which set of interested and authorized subscribers can receive access to a particular set of events.

In PawS [37,38], users are allowed to adjust the granularity of their location data that is made available to the system; the resolution of the user location can be indicated in meters, or fractions thereof. PawS service providers can also adjust the level of location granularity they require by requesting only the necessary elements from the full set of location and/or information, e.g., only asking for the building name instead of the building and a room number.

3.5 Constraints and Permissions

Role Based Access Control (RBAC) [50] is one of the most widely used methods to control access to resources and services. In RBAC, users are associated with roles, which, in turn, are associated with a set of permissions. In pervasive systems, a straightforward implementation of RBAC may be difficult due to a large number of diverse service providers and transient users, which may lead to privacy violations and information leaks [26]. In a PCE, constraints on access rights often provide means to balance the tradeoff between the amount of privacy a user is willing to concede and the value of service that will be provided to the user. For example, users of an application may give control over choosing their location that will be available to others based on the identity of the person receiving the information. In order for PCE to be successful, owners of information should have a convenient way to stipulate the conditions under which their information can be accessed.

RAVE [9, 40] provides users with control over connections: which users can connect and what kind of connections are allowed. The users are made aware of what data is being sent via the feedback provided by the RAVE system. Therefore, RAVE makes its users aware of the technologies used in the PCE and their potential implications to their privacy.

Information Space Model [31] uses information boundaries and allows its users to define a set of permissions for access to information, resources, and services. The information space boundary acts as a trigger for privacy control to enforce permissions defined by the owners of each information space. A boundary, whether physical, social, or activity-based, delimits an information space. For example, a school principal can create an information space model that contains all the information in his office.

Context Models for Privacy [28] uses fact-based and situation-based approaches to manage the sets of constraints and permissions. Fact-based model provides a tool that application developers can use to explore and specify context requirements; it defines entities about which context information is required and types of information that interests the users. Developers may also choose to identify an appropriate source for each fact type using this fact-based model. Situation-based model provides a way to describe contexts at a higher level than the fact-based model. Situations are defined using predicate logic and can be combined easily with logical connectives to form rich descriptions. The situations help yield a truth value by expressing conditions on the context that can be evaluated against a set of variable bindings and a context to allow or deny access to shared information.

3.6 Ownership of Context Information

Features that allow or enable ownership of contextual information mark information in a manner that allows an association with an individual, behavior or action. This association gives the system the ability to protect data by implementing rules based on context, which can be used as an additional detail to be considered when evaluating and interpreting privacy rules. Caveats based on contextual information and its

ownership association with an individual user provides additional restrictions to data access rights.

Privacy Protecting Middleware [17] provides several mechanisms to control the conscious disclosure of contextual data. The rules for disclosure are negotiated and established at the time data is requested. Only the data that meets the privacy criteria set by the rules during the negotiation phase can be released. It is also capable of modifying the contextual data for a specific use or situation. These modifications remove privacy specific details that are not relevant to the nature of the data exchange. These mechanisms require a conscious awareness of contextual ownership in order to negotiate and publish data that is consistent with the privacy concerns of the owner. The state or fact of being owned must be a prerequisite to any viable negotiation between an entity that is requesting access to information and one that is safeguarding information.

RAVE [9, 40] provides a mechanism to establish the user context ownership and control. Users have complete control over the audio-video content in their personal space by controlling the state (on/off) and placement (location) of all audio-video equipment. In the RAVE virtual office space, workers choose to contribute information or opt out of the environment at any time. They can also choose who will receive their multimedia streams and what data to share. In multimedia spaces, context information is related to images in the video and sounds from the audio stream. Users have full authority over the contextual information in their personal office space but not in the community areas.

Participants in the PawS environment [37] can also assign ownership to context. PawS uses the W3C Platform for Privacy Preferences (P3P) to express ownership and control over user content [1, 48, 55]. P3P is both machine and human readable and can be used to mark content and communicate privacy policies. PawS users can also assign a privacy statement to context information using the P3P syntax. Although P3P can communicate user's privacy preference, it cannot enforce the preference: it relies on trust for enforcement.

Trusted Platform Module [7] can establish ownership by creating a non-migratable key to lock the information. This technique assumes that the user who holds the keys to unlock the information is also the owner of the information. TPM can ensure privacy if the distribution of keys to decrypt the data is controlled. Otherwise, an unauthorized party can use the keys to access the information.

Context Models for Privacy [28] addresses the challenges of assigning ownership to context information and enabling users to express privacy preferences for their own information. This approach offers a direct link between an information source and the entities that should be entitled to control the corresponding context information for privacy purposes. This ownership relationship describes the connection between an entity and the context attributes in which they have an interest in terms of privacy. With context models, ownership can be assigned to facts or objects. Ownership can also be assigned to one user or a group of users. Finally, ownership can be based on situation by applying rules associated with the fact or object types referenced by the situation.

