Privacy Management for Context Transponders

Michael Fahrmair, Wassiou Sitou, Bernd Spanfelner
Technische Universität München, Department of Informatics,
Boltzmannstr.3, 85748 Garching (Munich), Germany
{fahrmair, sitou, spanfeln}@in.tum.de

Abstract

While by now feasible solutions to protect privacy for complex ubiquitous applications are available, very small devices, called context transponders (CTP) still lack resources to run sophisticated full scale approaches like DRM and strong encryption. These devices however, will play an important role in the success of ubiquitous computing as they support the idea of pervasiveness (small, mobile etc.). We propose a set of techniques outlined as the CTP-classmark that aims to introduce data avoidance and obfuscation strategies to enable information security in context aware applications. Using this classmark, the users' privacy can be protected even on low resource devices and hence the users' trust in ubicomp applications will rise.

1. Introduction

Several context aware service provisioning systems have been developed in our research group during the last seven years. Among them were a mobile community based search engine [6], an in-house navigation and information portal for our campus and a long term study about a smart home [5]. By deploying these systems and making them available to a non involved audience (e.g. when asking for permission to use the campus wide WLAN for limited tracking experiments) we learned several important lessons.

At first and above all that privacy is a major concern for ubiquitous computing, especially if it relies on monitoring large amounts of context information. Second we learned that larger ubicomp applications often depend on crucial smaller subsystems that are adaptive and context aware themselves. Such context transponders (CTPs), are both the smallest hardware and software entities within the domain of context adaptation. CTPs at least manage their own context space (a semantically encapsulated context storage and retrieval service with separate access control either virtually inside a context server implementation or within the CTP's own local context server).

These devices will play a vital role in the success of ubicomp since they are very small and mobile and hence enable everywhere distribution. Nevertheless compared to simple sensors they are more multi purpose, because they have the capability of limited interaction and intelligent information preprocessing. Location based services, matchmaking (see PCP example in Section 2), or unified notification scenarios often make use of such devices. For example the PDAs in the campus example above act as semi intelligent sensors with their own context management to adapt to different positioning technologies depending on available infrastructure (WLAN, GSM, GPS) while being part of several ubiquitous applications at the same time. So it is important that privacy protection in a specific application needs to enclose all information paths including security in any lower sub systems.

Unlike larger ubicomp applications, which can rely on sufficient resources for this task to run approaches like [7], for CTPs such strategies are not always feasible. With the CTP-classmark (Section 3) we therefore introduce a set of resource saving privacy techniques based on avoidance and obfuscation instead of access control. This enables CTPs of maintaining user's privacy without major collapses of performance for most local service provisioning or data matching scenarios. The classmark consists of:

- A privacy capability descriptor to negotiate avoidance and obfuscation of context exchange where reasonable (i.e. depending on application and situation).
- Several avoidance techniques to ensure that only exact matches are able to know the compared context.
- An extension proposal to obfuscate data even to the matching partner while maintaining the matching functionality

To validate our approach we have implemented a test system based on the CAWAR-Framework (see [13]). We tested the performance (Section 4) and verified that no major collapse in performance is introduced by the CTP-classmark. In addition we compared the approach to a fully

featured DRM based privacy protection system for ubiquitous computing described in [7].

2. CTP privacy requirements

To better expose the relevant problems with privacy protection for context aware systems in resource constrained situations we have developed a new dedicated prototype scenario called the Peer City Ping (PCP).

PCPs are simple wireless transponder devices (comparable to active badges [14]) that can hold both a simple personal profile (age, sex, hobbies etc.) and a search profile to look out for. Whenever two PCPs with mutual matching profiles are within radio range, a match is signaled to both users. The name Peer City Ping stems from the fact that it is difficult to signal such a match between two devices in a crowd of people (like in a city), and therefore a manual ping is used to localize matches.

This ping is manual because most of the PCP-Devices lack the power to form cross-layer and end-to-end adaptation synchronization control structures [3] to negotiate different sounds for different matches. Instead, each PCP is required to have at least one button to trigger a ping signal which causes the partner to send out an audio echo. Changing the rhythm of his button presses and listening for the answers allows the user to locate his right match. In addition this has the advantage of options for plausible deniability (anonymous decline) without anyone loosing her face.

Another problem besides identification and anonymous decline for context aware PCPs like data or service matching application scenarios is privacy protection. In this paper the following related questions are discussed:

- Is matching possible without sending sensitive data around that could fall into wrong hands?
- Is it possible to prevent positive matches to know my sensitive data just from the fact they searched for it and became a match?
- Is it possible to prevent a positive match to know what my search was?

As can be seen from the example what is making the overall process of mobile and specifically ubiquitous matchmaking so difficult, is the usually very limited interaction or computation resources in CTP devices. Especially potential limitations in encryption, authentication, processing and communication capabilities prevent the adoption of more sophisticated privacy rights management (PRM) concepts like described in [7, 9, 10]. This makes PCP an ideal example for protecting privacy in ubiquitous computing (i.e. available, operateable and useful in as many situations as possible [8]) in cases of very limited but flexible technical resources, i.e. small devices that can be temporarily part of several different ubiquitous applications (mobiles, weather station, RFID health card). In terms of the PCP scenario each PCP client is itself a ubiquitous application that can be adapted to temporarily incorporate other nearby PCPs (for the matchmaking), while at the same time being part of other applications (i.e. the other PCPs matchmakings). Therefore PCP clients qualify as CTP entities as previously defined.

Other CTP based privacy examples can be easily found when following the PCP matchmaking and signaling idea. Examples could be personalized advertisements or special offers on mobile phones when passing a shop, virtual graffiti, shopping and price comparing assistants in mobiles and so on. In principle also any ubiquitous service provisioning system could profit from a low resource privacy protection mechanism when comparing user preferences with service availability, load and pricing information that should not be revealed to competitors. Such non local scenarios with their centralized servers do not seem to suffer from hardware based resource restrictions in the first place though. However because of their wide area communication and high bandwidth they need to cope with an economically unpleasant asymmetry between the number of matching requests and actual brokered service usages. In such cases it is reasonable to let a large number of independent virtual CTPs (easier load balancing, failure recovery etc.) handle the initial service negotiations. In case of a positive service match it is then always possible to switch to a full scale privacy class mark later on in the provisioning protocol.

3. The CTP classmark

We have implemented the above described PCP scenario and evaluated its performance with full scale DRM using a rapid prototyping testbed for context adaptive middleware applications. This testbed has already been successfully used to evaluate predictive cross layer and end-to-end adaptation scenarios in a mobile telecommunication scenario [12]. As expected, the transmission of DRMed sensor information significantly reduced the scalability (see fig. 1). In addition CTP devices are usually cheap, heterogeneous and large in numbers and there is so far no reasonable economic way to reach an DRM-acceptable certification level for the involved devices.

Based on these experiments, we have developed a new classmark to amend the PRM classmark from [7]. The additional methods are especially tailored for CTP settings and ubicomp. The CTP classmark concentrates in obfuscation and avoidance of transferring sensitive information rather than protecting its access and controlling its usage. This means, instead of temporarily licensing and enforcing usage rights limitations (including copy protection) on sen-

sitive information, the CTP classmark avoids transferring sensitive information in the first place. Avoidance also has the advantage of fulfilling most of the privacy design guidelines as described in [11].

However in the PCP scenario for example, one side would have to send out some potentially privacy related information. Therefore the CTP classmark is based on the strategy to first make sure to *avoid* that any other than the one that really needs the information (positive match) will get it, second to *obfuscate* privacy related information towards positive matches as far as the service allows. The first part of the strategy is achieved by introducing at least one additional phase in the exchange protocol that is used as a reduced information match based authorization. This means the receiver first has to proof that she is most likely a match, without exposing too much valuable information. The second part of the strategy is usually achieved by adding some kind of fuzziness so that the data is still accurate enough to fulfill the service, not enough specific however to recreate the exact information. Of course what is 'valuable' and 'accurate enough' can sometimes depend on the exact service the data is used for.

Therefore the overall CTP classmark is composed of a service GUID (128 bit), the ID (32 bit) of the avoidance method and the ID (24 bit) of the obfuscation method. The ID of the avoidance method is further on divided into an 8 bit method ID, and two values (8, 16 bit) indicating respectively the number of filter levels for match based authorization phases and the granularity of subelements that could be sequentially exposed. The ID of the obfuscation method is composed similar, indicating method and the worst case accuracy of an information that is transmitted to respectively a negative and a positive match. In technical terms there are several such possible methods to avoid sensitive data to be sent to negative matches or obfuscate data for positive matches in the PCP example:

n-phase negotiation avoids unnecessary sensitive information to be transferred by using a piece-by-piece exchange mechanism. This means the less sensitive piece of information is exposed (and compared) first. Only if the other PCP continues to qualify as a matching partner it is sent further information. While working pretty well for the PCP service type, this kind of solution is time and bandwidth consuming and hence not suitable for situations with fast moving clients. The n-phase negotiation method of course also has the prerequisite that there are actually less sensitive pieces of data.

2-phase index negotiation is an optimization of the n-phase method that works with one small fixed dataset of less privacy sensitivity (e.g. an index) and a second (extensible) larger dataset that is only compared if the index already matched. This helps to reduce bandwidth usage and negotiation time but on the other hand requires the data to be specifically structured.

Hashing obfuscation is a method already used in a similar form for storing passwords. Applied to PCP (e.g. using MD5 /SHA) this means, the searching client calculates a hash fingerprint of the expected result set. The queried client in turn calculates the hash from his profile and compares it with the search hash. Only in case of a positive match, both search and profile information are exposed. This method is simple and cheap (because hashes could be precalculated most of the time) but the disadvantage is that the search must be exact.

Range hashing is an extension of the normal hash obfuscation that allows searching for subsets of enumerations (e.g. like searching for [tennis, volleyball] in a profile that contains [tennis, volleyball and bicycling]). Hence, the querying PCP needs to send the expected result hash as well as the range length. The queried PCP then uses the range length to build all possible combinations of the given length from the enumeration and calculates (looks up) their hashes. This method however has drawbacks in case of profiles containing vectors of large enumerations. Therefore, it is only reasonable to be used in combination with n-phase avoidance, because this prevents unnecessary matching checks. On the other hand side n-phase information pieces should not be too small (like comparing each element of the range) because this would seriously reduce the obfuscation level.

Similarity hashes try to avoid this problem and moreover improve obfuscation towards positive matching partners by using locality sensitive hash functions (e.g. [2]). Similarity in this case means that similar profile data sets or queries will produce similar hash results (comparable to music fingerprinting mechanisms like [1]). This allows for "fuzzy" matching and therefore only exposes rough ideas instead of exact sensitive information even towards positive matching partners. However this comes at the cost of also making it easier for mismatching clients to guess the transmitted information so this needs to be well balanced.

4. Performance experiments

As already mentioned at the beginning of section 3 we used a rapid prototyping testbed for context adaptive middleware applications to evaluate the CTP classmark. So far the context information from sensor devices (e.g. the search profile of the user's terminal) had been transferred into the distributed system without any privacy protection. To better reflect the typical properties of a low power CTP, we reduced the profile to a set of 10 elements (10 byte each) and implemented a simple DRM system (encrypted container, per element access right).

To test the CTP classmark we have implemented the hashing obfuscation method for 100 byte sensor data package and compared the results with DRM-based and no-

privacy transmissions (see Figure 1). As expected the CTP classmark has a much better scalability and behaves almost exactly like the no-privacy approach in a high shared bandwidth environment (i.e. all sensor clients have to share a common bandwidth like it is common in most wireless scenarios).

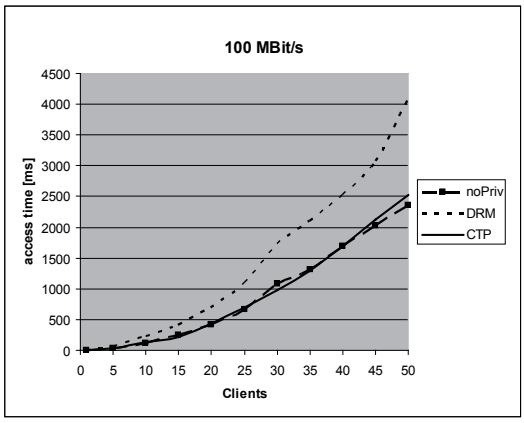

Figure 1. CTP classmark performance

Low bandwidth environments still lack advantages of hashing, because the used test-bed relies on XML based communication and therefore produces overhead that negates the positive effects of information avoidance there.

5. Conclusion and further works

In this paper we presented the CTP-Classmark for context transponders that are the smallest hard-/software based context adaptive systems. With these systems a fairly well suited technology is available to support ubiquitous computing. Nevertheless privacy will be a crucial factor for the success of such small devices. The context transponder classmark expands the variety of privacy enabling techniques for context aware systems to devices with very few resources. By using techniques like conventional and similarity hashing with precomputation we are able to reduce the effort for obfuscating context information compared with conventional DRM-based approaches. This approach assumes usage of context where a comparison of context information will result in additional information that will furthermore result in an adaptation decision (matchmaking). As a case study we used a PCP application, which matches the profiles of two users to find users with similar interests. A performance test showed that no major loss of performance is introduced by the hashing. Therefore the proposed CTP-Classmark performs better than a conventional DRM-System (tested as well) for matchmaking CTPs.

Further work has to be done exploring the use of locality sensitive hash functions for privacy protection and their performance optimization for matching multidimensional profiles using dynamic indexing concepts described in [4].

References

[1] P. Cano, E. Batlle, T. Kalker, and J. Haitsma. A Review of Algorithms for Audio Fingerprinting. In *International Workshop on Multimedia Signal Processing, US Virgin Islands*, December 2002.

[2] M. S. Charikar. Similarity estimation techniques from rounding algorithms. In *34th Annual ACM Symposium on Theory of Computing – STOC '02*, pages 380–388. ACM Press, New York, NY, 2002.

[3] M. Dillinger, E. Mohyeldin, J. Luo, M. Fahrmair, P. Dornbusch, and E. Schulz. Cross Layer and End to End Reconfiguration Management. In *WWRF11- Services and Applications Roadmaps- Invigorating the Visions*, 2004.

[4] M. Dillinger, E. Mohyeldin, J. Luo, P. Weckerle, P. Dornbusch, M. Fahrmair, and C. Salzmann. Dynamic Classmarks For SDR Equipment. In *13th IEEE International Symposium on Personal, Indoor and Mobile Communications – PIMRC*, 2002.

[5] M. Fahrmair. *Kalibrierbare Kontextadaption für Ubiquitous Computing*. PhD thesis, Technische Universität München, 2005.

[6] M. Fahrmair, C. Salzmann, and M. Schoenmakers. Verfahren zur Vorauswahl mobiler Dienste. Technical Report DE0010024368A1 [DE], Deutsches Patentamt, 2000.

[7] M. Fahrmair, W. Sitou, and B. Spanfelner. Security and Privacy Rights Management for Mobile and Ubiquitous Computing. In *Proc. of Privacy in Context – UbiComp Workshop*, 2005.

[8] M. Fahrmair, W. Sitou, and B. Spanfelner. An Engineering Approach to Adaptation and Calibration. In *Modeling and Retrieval of Context – MRC 2005*, volume 3946 of *LNCS*, pages 134 – 147, 2006.

[9] J. Hong and J. Landay. An Architecture for Privacy-Sensitive Ubiquitous Computing. In *2nd International Conference on Mobile Systems, Applications, and Services – MobiSys*, 2004.

[10] M. Langheinrich. A Privacy Awareness System for Ubiquitous Computing Environments. In *4th International Conference on Ubiquitous Computing – UbiComp*, 2002.

[11] S. Lederer, J. I. Hong, A. K. Dey, and J. A. Landay. Personal privacy through understanding and action: five pitfalls for designers. *Personal and Ubiquitous Computing*, 8, 2004.

[12] E. Mohyeldin, M. Dillinger, M. Fahrmair, P. Dornbusch, and W. Sitou. Interworking between Link Layer and Application Layer Adaptations in a Reconfigurable Wireless Middleware. In *15th IEEE International Symposium on Personal, Indoor and Mobile Communications – PIMRC*, 2004.

[13] E. Mohyeldin, M. Fahrmair, W. Sitou, and B. Spanfelner. A Generic Framework for Context Aware and Adaptation Behaviour of Reconfigurable Systems. In *16th IEEE International Symposium on Personal In-door and Mobile Radio Communications – PIMRC*, 2005.

[14] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The Active Badge Location System. *ACM Transactions on Information Systems*, vol. 10:pp. 91–102, 1992.