2009 International Conference on Intelligent Networking and Collaborative Systems

Harnessing Pseudonyms with Implicit Attributes for Privacy-Respecting Mission Log Analysis

Stefan G. Weber
Telecooperation Group, Department of Computer Science, Technische Universität Darmstadt
Darmstadt, Germany
sweber@tk.informatik.tu-darmstadt.de

***Abstract*— Many applications in the area of collaborative work can be enhanced by tracking users regularly. Consider a future emergency management application, in which mobile first responders are continuously tracked in order to support a better coordination of the rescue missions and to create a mission log. However, continuous tracking of individuals and storing the data for later use is often in conflict with individual privacy preferences. Therefore, it is a challenge to deal with conflicting traceability and privacy protection requirements. A common way to implement some kind of privacy protection is to use pseudonyms instead of fixed IDs for each user. However, in order to build a multilateral secure and acceptable solution, a more complex system design w.r.t. to pseudonym linkability is required, that also allows third parties to analyze the logs for organizational and legal reasons.**

In this paper, we present our approach to deal with this issue: we propose to encode additional information into pseudonyms that are used in location tracking systems and stored in data logs. Our concept comprises both access rights for the user herself and implicit attributes that may be verified by third parties in a privacy-respecting manner. We introduce the cryptographic constructions, which employ cryptographically secure pseudo-random number generators, threshold cryptography and techniques for securely evaluating encrypted data. Moreover, in this paper, we sketch a practical application example in the area of emergency mission log analysis and discuss the main security properties of our concepts.

I. Introduction

Ubiquitous Computing (UbiComp) denotes "a powerful shift in computation, where people live, work, and play in a seamlessly interweaving computing environment" [1]. According to this vision, coined more than 15 years ago by Mark Weiser, people will work and carry out their personal actions continuously supported by computers. As forerunners of true UbiComp applications, a large range of applications has been proposed during the last years which can be enhanced by tracking users regularly. Imagine a future emergency management application, in which mobile first responders are continuously tracked in order to support a better coordination of the rescue missions and to create mission logs that can be analyzed afterwards. Since emergency work is characterized by extreme conditions like working under time pressure and with psychological burdens of having the responsibility of saving lives, we believe it is an adequate scenario for studying how technologies need to be designed to unobtrusively support collaborative work [2]. In this paper, we especially look at the issue of employing UbiComp technologies for collaborative work from a data privacy perspective. Security and especially privacy issues have been identified as one of the greatest barrier to the long term success of such upcoming UbiComp applications [3] that rely on continuous large scale data collection, like location tracking or employee tracking. Once data related to individuals has been collected, organizational and legal requirements may be in conflict with individual privacy preferences. Therefore, we believe that it is a major challenge to design data security mechanisms that allow for balancing between the conflicting goals of traceability for legal and organizational reasons and individual privacy protection [4]. In order to deal with the arising inherent tradeoffs of future UbiComp applications, this paper presents our work towards implementing balanced systems w.r.t. multilateral security. Especially, we consider the case of location tracking applications. Hereby, we focus on the ambiguous potential of data logs that emerge from continuous tracking. Instantiated in the context of collaborative emergency management, our concepts allow us to implement functionalities for privacy-respecting mission log analysis and also support the individual user in dispute cases.

Our contributions. We describe a new approach towards a pseudonym system that allows the users to exercise control on location data that has been collected along with pseudonyms, and protocols that allow 3rd parties to analyze the resulting location data logs in a privacy-respecting manner for traceability reasons. Basically, we enable users to be tracked pseudonymously, with the possibility to authenticate the own pseudonyms after use, thus making them personalized again. Therefore, we present our concept for generating pseudonyms on mobile devices, which is based on the use of cryptographically secure pseudo random number generators. Moreover, we provide protocols that allow for analyzing location log data in a privacy-respecting manner, based on threshold cryptography and techniques for securely evaluating encrypted data. Basically, our proposed approach exhibits the following main properties:

- Users are empowered to access the location data collected along with a pseudonym by encoding access rights into the pseudonyms.
- 3rd parties may selectively analyze log data by evaluating attributes that are implicitly encoded in the pseudonyms while preserving users' privacy.

978-0-7695-3858-7/09 $26.00 © 2009 IEEE
DOI 10.1109/INCOS.2009.40

- Verifiability of the log analysis functionality is given, i.e. non-compliance of 3rd parties w.r.t. to the allowed privacy-preserving actions can be detected.

Organization of the paper. In the following section, we introduce the basic idea of our approach. In section III, we describe necessary cryptographic background. Then, in section IV, we introduce the main concepts for pseudonym generation and privacy-respecting log analysis. This is followed by a detailed presentation of our approach in section V. An evaluation of our proposal with regard to the security requirements fulfilled can be found in section VI, followed by a discussion of related work in section VII. Finally, in section VIII, we conclude our work.

II. Basic Idea

In this section, we describe the basic idea of our approach. For this purpose, we sketch an illustrative application scenario first.

A. Application Scenario and Setting

We return to the introductory example of the future emergency management system. In this system, mobile first responders are continuously tracked for the duration of their missions out in the field. Their real-time position and tracking information can be displayed on a digital map in an emergency management headquarter, so that the data can by analyzed by a decision maker. While the *current* tracking information supports the coordination of rescue missions through the decision maker in the headquarter [5], collected *historical* position information of the first responders also creates some kind of log which documents the rescue missions. This data log is what we call a *mission log*. It can be analysed for several purposes in the postprocessing phase of an emergency. In the following, we assume a mission log to obey to a simple structure: it contains several entries in the form *entity ID - time - location*.

B. Security Requirements

Having introduced the concept of the mission log, we next outline basic data security requirements that have to be met to implement a *multilateral secure mission log*. Thereby, multilateral security [6] is considered in the sense that we want to take into account security and privacy requirements of both the individual first responder and of the emergency management organization.

Basically, the individual first responder wants his privacy to be respected. Especially, despite the need for fine-grain location tracking to complement the decision making work, she wants to remain anonymous during the course of the mission. In contrast, the organization which accounts for the emergency management wants to be able to analyze processes after a mission. Also, organizations tend to verify compliance. In the first response application example, this relates to misuse detection in the mission log, e.g. by following questions like ”Did first responder X behave properly and did not exploit the current situation for malicious purpose?”. This is a highly critical question, since, in real world rescue missions, first responders need to break existing laws in some cases, to save lives. Therefore, psychological burdens of being traceable must be diminished so as not to hinder the rescue work. Moreover, a first responder might be accused of non-assistance of a person in danger, which could be judged upon according to the mission log. The challenging question is how to realize a technical solution that allows to balance between individual privacy preferences and organizational and legal traceability needs, and still allows to provide tracking information to support mission control.

C. Envisioned Functionality

A very basic kind of privacy protection can be implemented by allowing the user to be tracked und a pseudonym. However, in order to achieve a multilateral secure and acceptable solution, a more complex system design is required. As Flegel [7] lines out, two important facts have to be dealt with in the context of creating privacy-respecting log analysis functionalities:

1) ”The controlled disclosure of pseudonyms is the controlled ability to make pseudonymized objects accountable again. This ability is controlled by controlling who can use the pseudonymity mapping.”
2) ”The disclosure of pseudonyms should be bound to a priori specified purposes.”

Therefore, a central point is to implement a selective control functionality regarding the pseudonym linkability for both users and organizations. Moreover, mechanisms that allow for detection of crucial events inside a log are required. According to the discussion in the last two sections, we devise the following set of functionalities we want to implement.

First, regarding the first responders’/users’ perspective:

- The user should be tracked pseudonymously to provide some kind of basic privacy privacy protection.
- The user should be able to authenticate a pseudonym, that he has been tracked under. In dispute cases, this functionality allows him to repudiate accuses like ”you did not provide assistance during the mission at time X place Y” by proving evidence of having been ”at place Z at time X”.

Second, regarding the perspective of third parties, e.g. emergency management organizations:

- It should be possible to selectively analyze the entries that are recorded in a mission log in a privacy-respecting manner. Therefore, the logs should contain only pseudonymized entries that hide the real world identities of tracked persons.
- It should be possible to link sample entries, i.e. to verify if they belong to the same user.
- Moreover, it is desirable to check if one entry relates common organizational function, e.g. if it belongs to ”fire department Darmstadt”.
- If legally convincing pieces of evidence for misbehavior have been identified in the course of this analysis, e.g. the

case of non-assistance of a person in danger, it should be possible to reveal the true identity of a tracked person.
- The whole process of mission log analysis should be audited, i.e. it should be detectable if the parties that are responsible for it do not comply with the rules set up for the privacy protection of the individual user.

III. Cryptographic Building Blocks

In this section, we briefly describe the cryptographic tools that we will employ in our constructions.

A. Cryptographically secure PRNGs

A pseudo-random number generator (PRNGs) is a tool for generating sequences of numbers which seem to exhibit the properties of random numbers, even though they are created by a deterministic algorithm. PRNGs use an internal source of entropy called seed to derive the output values. The stronger notion of cryptographically secure PRNGs relates to PRNGs that produce sequences of numbers with stronger security requirements, i.e. it is actually impossible for an attacker to guess or derive any future or previous numbers by analyzing the output of a cryptographically secure PRNG. In the following, we simply use the term PRNG, even though we actually employ cryptographically secure ones.

B. Threshold ElGamal Cryptosystem

In our approach, we employ the ElGamal cryptosystem [8], over subgroups G_q of order q of the multiplicative group Z_p^*, for large primes $p = 2q + 1$. We treat the primes p, q and a primitive element g of G_q as common system parameters. More specifically, we utilize a threshold variant of it, according to Cramer et al. [9], offering robustness and distributed trust. Herein, the private key $s \in_R Z_q$ is generated via the distributed key generation protocol of Pedersen [10], and consequently it is secret shared [11] among all n participating authorities. Thus, the power to decrypt is distributed among all participating authorities, and a minimal number of t out n authorities are necessary to perform the threshold decryption operation. The authorities common public key is $h = g^s \bmod p$. A message $m \in G_q$ is non-deterministically encrypted by choosing $r \in_R Z_q$ and by computing $(g^r, h^r m)$.

C. Non-interactive Zero Knowledge Proofs

Zero knowledge proofs (ZKPs) are used to guarantee correctness of private actions in order to implement verifiability. Especially, we use non-interactive zero knowledge proofs (NIZKPs) by applying the Fiat-Shamir heuristic [12]. See e.g. [13] for discussions of these techniques. Since NIZKPs may be stored, they may be checked for verifiability purposes after a protocol run.

D. Broadcast Channels with Memory

Our constructions require broadcast channels with memory to store and exchange information in protocols involving distributed computations. These channels can be seen analogous the concept of so called bulletin boards [14]. In our setting, a broadcast channel is append-only, i.e. once the information is sent over it, it is stored and cannot be changed or deleted afterwards. Authorities of the threshold cryptosystem use such channels to provide their partial computations to the other authorities. Also, they store NIZKPs during the computations, which allows to later verify their correctness.

E. Plaintext Equality Tests

A plaintext equality tests (PET) [15] is a primitive for pairwise blind comparisons of ciphertexts of e.g. the ElGamal cryptosystem. A PET allows to test whether two randomized encrypted ciphertexts represent the same plaintext by performing algebraic operations on the ciphertext, but without revealing the plaintext.

F. Reencryption Mixnet

A mixnet, originally introduced by Chaum [16], is a cryptographic tool to anonymize ciphertexts. In our following construction, we employ ElGamal based reencryption mixnets [17], which basically reencrypt and permute ciphertexts to afford anonymity. Moreover, we require the mixnets to be verifiable, i.e. to provide proofs of correctness of their operations. We do not go into details here, but point to the proposal of Furukawa et al. [18], that implements verifiability for reencryption mixnets.

IV. Basic Model

We now introduce the basic cryptographic constructions of our approach. These constructions implement functionalities for pseudonym mapping and linking, for both considered parties, i.e. users and organizations.

A. Pseudonym Registration and Generation

In this section, we introduce the concepts for pseudonym generation. Basically, we propose to encode a static attribute inside malleable pseudonyms by generating the pseudonym as encryption of an attribute under the public key of a threshold ElGamal cryptosystem. The resulting construction is what we call a *pseudonym with implicit attribute*.

First, each user must participate in a registration phase. In the registration phase, each user receives a trusted personal device which includes a cryptographically secure PRNG. Therefore, the user interacts with a trusted registration authority. The registration phase consists of the following main steps:

1) Each user receives a trusted device. The registration authority registers a unique seed in the PRNG of the device to enable it for pseudonym generation.
2) Each user receives a base pseudonym from the registration authority. The base pseudonym contains a distinct implicit attribute which associates the user with the issuing organization. For example, a user may receive a pseudonym with implicit attribute *Firefighter* $\#1$.
3) The user's real world ID is encrypted and stored together with the implicit attribute on a registration list.

In the registration process, all encryptions are done under the public key belonging to a set of so called verification authorities inside the organization, e.g. in the emergency

management organization. To generate the base pseudonym for a user, the registration authority encrypts the chosen attribute A under the public key of the verification authorities: $P = E_{PK_{VA}}(A) = (g^r, h^r A)$. Moreover, the random value r is transferred to the user and stored on the device. The user now derives *tracking pseudonyms* from the base pseudonym in the following way:

1) The seeded PRNG is used to generate a sequence of random numbers.
2) Each random number r_i is used to compute randomization factors $F_{r_i} = (g^{r_i}, h^{r_i})$.
3) F_{r_1} is used to construct the *first* tracking pseudonym by multiplying it with the base pseudonym: $P_1 = (g^r, h^r A) * (g^{r_1}, h^{r_1}) = (g^{r+r_1}, h^{r+r_1} A)$.
4) Further tracking pseudonyms are created by consecutive multiplication: $P_{i+1} = P_i * F_{r_{i+1}}$.

By this procedure, a user creates a set of different tracking pseudonyms that all contain the same implicit attribute. These pseudonyms are used instead of fixed IDs, during the location tracking. Moreover, due to the construction, users are enabled to authenticate a tracking pseudonym that is stored in a mission log. Therefore, he needs to show that he is in possession of the base pseudonym and the correct aggregated random factor, which allows to reproduce a recorded pseudonym, thereby authenticating it. This functionality is what we call *attestation*.

B. Method of Privacy-Respecting Mission Log Analysis

In this section, we introduce the concepts for privacy-respecting mission log analysis. Basically, we harness the possibility to do algebraic operations on the randomized-encrypted ciphertexts that represent pseudonyms.

Remember the structure of a mission log, as introduced in section II-A: *entity ID - time - location*. Assuming that the *entity ID* values are actually values of *pseudonyms with implicit attributes*, we allow the verification authorities of the responsible organization to execute two basic operations for privacy-respecting log analysis:

1) check if two log entries relate to the same entity but without revealing the ID of the entity;
2) check if one entry relates to a group of entities with a common organizational function.

The first operation is implemented by executing a *plaintext equality test* on the pseudonym values of two entries. Suppose that $P_a = (g^{r_a}, h^{r_a} A_a)$ and $P_b = (g^{r_b}, h^{r_b} A_b)$ represent two entries of that kind. If they relate to the same entity, they contain the same implicit attribute. In order to verify this, the pseudonyms are divided: $P_c = P_a / P_b = (g^{r_a - r_b}, h^{r_a - r_b} A_a / A_b)$, which is an encryption of the attribute "1", if A_a equals A_b. By performing a threshold decryption, the verification authorities yield an explicit attribute which is either "1" or a meaningless different value.

The second operation can be seen as an extension of the test above to a global instead of pairwise comparison of pseudonyms. For example, it allows to test if one log entry relates to the organizational unit "Firedepartment Darmstadt", but without disclosing which of the involved firefighters is the actual originator.

The method is based on operations of an ElGamal threshold cryptosystem. Basically, it works as following:

1) First, all participating verification authorities jointly generate a shared key z.
2) Then, the authorities select all attribute values that are relevant to the organizational function and create base pseudonyms for each relevant attribute.
3) Next, all base pseudonyms are processed by a reencryption mixnet. This creates an anonymized list of base pseudonyms, w.r.t. to the position of the individual attribute in the list.
4) The verification authorities cooperatively apply their shares of z each anonymized base pseudonym. This process achieves blinding of the attribute inside the pseudonym.
5) After that, each blinded pseudonym is decrypted. This yields a blinded attribute, which is used as a deterministic yet blind fingerprint of the original attribute related to the base pseudonym.
6) After processing all base pseudonyms that need to be considered, they can be compared without leaking information about the implicit attribute by comparing only the blind fingerprints.
7) In order to do so, the authorities also derive a blind attribute fingerprint for the pseudonym value of the one entry of the mission log that shall be verified.

Having outlined the method step by step, we go more into details. The whole approach makes use of secret sharing techniques according to Shamir [11] and of distributed key generation according to Pedersen [10]. To jointly generate the required secret shared key z, the verification authorities employ the distributed key generation protocol due to Pedersen. In this protocol, each authority j $(j = 1 \cdots n)$, receives a share z_j of the key z.

In the following, we describe the complete protocol for *distributed blinding*, which is analog to the distributed decryption protocol [9] of the ElGamal threshold cryptosystem. This protocol can be used to blind an arbitrary element $x \in G_q$ using the shared key z. This method is used to apply z cooperatively to the pseudonyms, which are effectively encoded as two elements of G_q.

1) Each authority computes $b_j = x^{z_j}$, a partial blinding of x, by applying its secret z_j. Also, each authority publishes publicly b_j together with a NIZKP that

$$\log_g \rho_{z_j} = \log_x b_j$$

. The latter is realized using a proof of knowledge for equality of discrete logs [19]. The proof assures that the authority indeed utilized its correct share to produce the partial blinding.
2) For any subset Λ of n authorities with valid zero-knowledge proofs, the blinded value x^z is reconstructed

using the discrete Lagrange interpolation

$$x^z = \prod_{j \in \Lambda} b_j^{\lambda_{j,\Lambda}} \bmod p$$

where

$$\lambda_{j,\Lambda} = \prod_{l \in \Lambda \setminus \{j\}} \frac{l}{l-j} \bmod q$$

are the appropriate Lagrange coefficients.

Let $A_i \in G_q$ be an attribute plaintext, and $(g^r, h^r A_i)$ an pseudonym with implicit attribute A_i with $r \in_R Z_q$, the authorities produce the deterministic fingerprint through the following steps:

1) To each component of $(g^r, h^r A_i)$ the distributed blinding protocol is applied, blinding it to a fix secret shared exponent $z \in Z_q$: $((g^r)^z, (h^r A_i)^z) = (g^{rz}, h^{rz} A_i^z)$.
2) The blinded pseudonym is jointly decrypted to the blinded attribute A_i^z using the distributed decryption protocol of the threshold ElGamal cryptosystem.

Here, A_i^z represents a deterministic fingerprint produced with a key z. It is used to blindly compare attributes encoded in the pseudonyms.

C. Disclosure of Pseudonyms

As stated earlier, "the disclosure of pseudonyms should be bound to a priori specified purposes" [7]. In the last sections, we provided methods for analyzing entries of a mission log with regards to pseudonym linkability. However, it is beyond the scope of this paper to define conditions and rules for misuse detection. However, in order to complement the two provided operations for log analysis, we . Simply, the operation for complete disclosure of pseudonyms works as following:

- Upon misuse detection, the verification authorities decide to completely disclose the identity behind a pseudonym.
- First, they cooperatively decrypt the pseudonym part of the relevant entry of the mission log. This yields the plaintext of the distinct attribute encoded in the pseudonym.
- Next, the authorities select the corresponding entry on the registration list. Then, they cooperatively decrypt the deposited ciphertext to the real world ID.

D. Application Example

Having introduced the allowed operations for mission log analysis, we next provide an example that shall illustrate how these operations can actually be employed in practice for analyzing log data. Again, the example takes place in an emergency management context. We assume the scenario, that a large airport is affected by a large scale emergency. Several airplanes have caught fire due to an accident. The burning is spreading over to the terminal buildings. Since the fire brigade of the airport is unable to handle the situation on its own, additional forces from nearby fire departments are requested. Also, the *fire department Darmstadt* sends a group of 50 mobile first responders to support the rescue missions. Arriving at the airport, the first responders register with the tracking system, receiving base pseudonyms with implicit attributes *Firefighter* #247 to *Firefighter* #296. During the course of the successful rescue mission, their movements and actions are continuously tracked and stored in the mission log. In the preprocessing phase of the emergency, a group of airport officers is appointed to analyze a specific incident which has been reported to them by an anonymous eye witness: it is mentioned that a group of 4 firefighters suddenly disappeared from an important task to extinguish fire in an office wing, located in the vicinity of a jeweller in the shopping area. The store reports that expensive items have disappeared, and several offices have been destroyed due to the lack of man power. Analyzing the mission logs, the officers manage to identify traces of 5 pseudonyms that move away from the office wing in question (*by using the first operation*). Moreover, the 5 pseudonyms can be identified to belong to *fire department Darmstadt* (*by using the second operation*). On request, the commander from Darmstadt asserts, that a group of his first responders decided to change the mission task, due to the observation of strange knocking sounds nearby. While, in that case, the commander's information suffices to resolve the incident in question, the individual firefighter could have also used the *attestation* functionality, to document scenes of the mission.

V. Complete Approach

This section presents our complete approach. First, we sum up the phases of the protocol.

A. Phases

- *Setup phase:* In this phase, system parameters for the threshold cryptosystem and private keys of the verification authorities are created.
- *Registration phase:* Each user receives a personal device and a unique seed from the registration authority. Moreover, a base pseudonym with an implicit attribute is registered to the use. The base pseudonym allows to derive tracking pseudonyms in the next phase.
- *Pseudonym Generation:* A user creates tracking pseudonyms locally on the personal device. Basically, each tracking pseudonym is a randomized threshold encryptions of the same attribute. The devices PRNG is used for providing random factors in the pseudonym creation.
- *Tracking phase:* After that, in the tracking phase, the user is regularly tracked under one of the tracking pseudonyms. Thereby, the pseudonym is regularly changed, according to the preference of the user.
- *Attestation phase:* After the tracking phase, a user may authenticate an entry of the mission log by providing the correct keying material that allows to reconstruct the pseudonym of the entry.
- *Log Analysis phase:* In this phase, entries of the log are processed by verification authorities for organizational and legal reasons. Upon convincing detection of misuse, pseudonyms can be disclosed.

- *Authority Auditing phase:* In this phase, the correctness of the actions of the verification authorities in the log analysis phase is verified. This done by checking the NIZKP stored on the broadcast channels with memory.

B. Protocol Description

We next describe our complete approach in detail. Thereby, we take into account the building blocks from section III, and the operations described in section IV.

In the following protocol description, we denote the participants as: U_i a user of the location tracking application; VAs the n verification authorities that share the private key SK_{VA} with the corresponding public key PK_{VA}; RA a trustworthy registration authority.

- *Setup:* The participating verification authorities (VAs) jointly agree on the appropriate system parameters p, q, g and cooperatively generate PK_{VA}, and SK_{VA} by using the threshold ElGamal cryptosystem. In addition, they produce a shared key z via the distributed key generation protocol of the threshold ElGamal cryptosystem. PK_{VA} as well as the ElGamal system parameters are published.
- *Registration:* Each user U_i is registered by a trustworthy registration authority RA. After handing out a personal mobile device, the RA generates an unique seed s_{U_i} and transmits it to the mobile device of U_i. Additionally, each user receives a base pseudonym with an implicit attribute: $P_{U_i,B} = E_{PK_{VA}}(A_{U_i})$. The random factor $r_{i,B}$ used in the encryption process is also stored on the user's device. At end of this phase the RA publishes a reference list C. It contains entries of all registered users (encrypted real world ID) and their assigned attributes: $C_{U_i} = E_{PK_{VA}}(U_i) - A_{U_i}$.
- *Pseudonym Generation:* Each user U_i derives from his base pseudonym $P_{U_i,B}$ a set of tracking pseudonyms $\{P_{U_i,j}\}$. In order to do so, first, the seeded PRNG is used to produce a set of random factors $\{r_{i,j}\}$. Then, the random factor inside the base pseudonym is updated with a random factor from the set: $P_{U_i,1} = P_{U_i,B} * (g^{r_{i,1}}, h^{r_{i,1}})$, $P_{U_{i,j+1}} = P_{U_i,j} * (g^{r_{i,j+1}}, h^{r_{i,j+1}})$.
- *Tracking:* When using the tracking system, each user is tracked under tracking pseudonyms from the set $\{P_{U_i,j}\}$. According to the preferences of each user, the pseudonym $P_{U_i,j}$ is be changed to $P_{U_i,j+1}$ in a specified interval of time. Note that changing a tracking pseudonym does not change the implicit attribute an user is tracked under. In the tracking phase, a mission log is created. Its entries are in the form *entity ID - time - location*. The *entity ID* field records the value of a pseudonym with implicit attribute, i.e. a tracking pseudonym.
- *Attestation:* In dispute cases after the tracking phase, a user U_i may access data recorded in the mission log. In order to do, he selects an entry k of the mission log by handing out attestation information in form of a random factor $r_{i,A} = \sum_{j=1}^{j=k} r_{i,j}$ and the base pseudonym $P_{U_i,B}$ to the verification authorities. They verify whether the pseudonym of the mission log entry k matches the reconstructed tracking pseudonym $P_{U_i,k}$. In that case, the user receives a tuple *time - location*, which is additionally certified by the verification authorities. The user can use the tuple to invalidate location- and time-dependent accusations against him.
- *Log Analysis:* According to an organizationally and legally defined set of conditions and rules, the verification authorities use the provided operations to detect evidences of misuse in the mission log. If convincing evidences have been identified, they introduce a pseudonym disclosure. Therefore, the authorities cooperatively decrypt the pseudonym part of the relevant entry of the mission log. The resulting attribute plaintext is used to select the corresponding entry on the registration list. Then, the authorities cooperatively decrypt the deposited ciphertext to the real world ID.
- *Authority Auditing:* External verifiers may verify the correctness of actions of the verification authorities in the log analysis phase. In order to do so, they can access the content of the broadcast channels with memory. These channels provide a log of the committed actions. Actions that do no comply with the allowed operations can be identified, as well as attempts to corrupt the cooperative operations, since stored NIZKPs cannot be verified correctly in that case.

VI. Analysis

In this section we sketch the analysis of the presented approach with respects to main security requirements. We also outline practical aspects that have to be considered in the course of implementing a system that actually employs our concepts, afterwards.

A. Security Requirements

- *User Privacy:* The privacy of users of location tracking applications is basically protected due to the use of pseudonyms. Since a user is able to adjust the frequency of pseudonym changes, he is empowered to adjust the level of anonymity provided by the system.
- *Pseudonym Linkability:* Even though the basic pseudonym construction stems form an encryption operation, no single verification authority is able to decrypt a pseudonym and thus to link a pseudonym to a user. Both registration list and pseudonym creation are performed as encryption operations of a threshold cryptosystem. Therefore, only a quorum of at least t verification authorities may directly link a pseudonym to a user. The allowed operations for log analysis also require the cooperation of a quorum of authorities.
- *Pseudonym Authentication:* In order to implement an attestation functionality, which could actually entail legal effects, the pseudonym authentication needs to assure a uniqueness property. Especially, it must not be possible for an attacker to provide authentication information for a pseudonym without being the real originator. This is

achieved in our construction due to a two-factor authentication process: The users needs to provide both the base pseudonym and the aggregated random factor. An attacker cannot intercept a base pseudonym, since the registration is assumed to be trustworthy. Moreover, being able to reproduce matching random factors contradicts the inherent assumptions of the employed cryptographically secure PRNGs.

- *Trust and Robustness:* The proposed construction allows to represent a large range of inter-organizational distributions of duties and powers. This is due to the use of a (t, n) threshold ElGamal cryptosystem as well as the distributed computation of fingerprints. Moreover, this approach can also tolerate the failure of at maximum $n-t$ authorities, which may be due to unavailability or due to corruption. Additionally, the robustness of scheme stems from its ability to tolerate attacks against verification authorities or failures of them, without corrupting the whole system. However, we require the registration phase of the scheme to be trustworthy, since it depends on a single registration authority, which is contrary to the distributed design of the rest of the scheme.
- *Verifiability:* The protocol is verifiable by anyone who can read the broadcast channels. Thus, the correctness of the scheme can be made transparent to external verifiers. Each distributed computation step can be verified by either checking the broadcast NIZKPs stored on broadcast channel in case secret inputs are involved, e.g. private keys or secret permutations in the mixing, or by re-performing the deterministic steps of the computation.

B. Practical Aspects

While a prototype implementation of our proposed concepts is not finished, yet, we next list practical aspects that have to be considered:

- *Handling changing IDs:* In order to implement tracking pseudonyms, in our approach we assume that the personal device of the user is able to handle changing IDs. Especially, the tracking may not proceed under some kind of static MAC address. However, the technical feasibility to implement changing IDs is a standard assumption. We share e.g. with [20].
- *Pseudonym Bit size:* In the proposed concept, pseudonyms are encoded as *two* elements of an algebraic group. Therefore, the bit size of a pseudonym depends on the characteristics and chosen parameters of the group. In order to minimize the communication and storage overhead, we propose to employ elliptic curve groups in order to implement the cryptographic operations of the threshold ElGamal cryptosystem.
- *Generation of Randomness:* To support users in accessing tracking pseudonyms recorded in a mission log, we propose to use random factors produced by a cryptographically secure PRNG as authentication information. Therefore, it is required to implement the generation of randomness on mobile devices. In order to do so, we are currently following the approach of [21].

VII. Related Work

Quite some work has been done in the area of pseudonymity, location privacy and selective traceability. In this section, we present a selective discussion on representative work that relates to our research. Classically, Chaum [16] introduced pseudonyms as basic tool for privacy protection in distributed systems. In the following years, several types of pseudonyms and applications have been identified. Kesdogan et al. [22] proposed to use changing pseudonyms in mobile GSM networks. Concerning the provided degree of linkability, pseudonyms can be classified into *transaction pseudonyms*, *role pseudonyms*, *relationship pseudonyms*, *role-relationship pseudonyms* and *person pseudonyms*. While a discussion of this issue can be found in [23], we see our approach as a conceptual combination of both transaction and person pseudonyms. The ability to use changing tracking pseudonyms in a freely determined frequency relates to the first property, while the implicit attribute can be interpreted as an implicit person pseudonym. In the context of Ubiquitous Computing applications, pseudonyms have been proposed as a basic mean for location privacy protection. In [24], Henrici et al., propose a hash-based construction of pseudonyms which allows for implementing changing pseudonyms for RFID applications. Notably, Heinemann [20] proposed to implement privacy-respecting opportunistic networks based on changing pseudonyms, whereby he proclaims that IDs must be changeable within every communication layer. Moreover, Beresford et al. [25] combine the use of changing pseudonyms with an abstraction of mix-nets to geographic regions, called mix-zones. For users of location-based service, these mix-zones are service-free zones, in which the users' actual change of pseudonyms is done. Delakouridis et al. [26] apply pseudonyms to the problem of storing and accessing location information in a privacy-preserving, decentralized manner, in their Share The Secret (STS) architecture. They propose to split the information to be protected according to Shamir secret sharing [11], and to distribute those shares on several servers, addressable via pseudonyms. Differently, the work of Biskup et al. [27] focuses on pseudonyms as means to implement a multilateral secure solution which allows to deal with conflicting privacy and traceability requirements. Especially, they propose a method for the recovery of transaction pseudonyms. Herein, real world identities can be recovered if a threshold of pseudonymous actions is detected. Our concept also employ a threshold construction. Differently, we require the participating verification authorities to exceed a threshold, in order to execute certain operations. Hereby, we build on several useful properties of threshold-homomorphic cryptosystems [28], which have a long tradition of research in the area of cryptographic protocols.

VIII. Conclusion and Outlook

In this paper we introduced a novel approach for generating and analyzing pseudonyms with implicit attributes. One goal of our research is to enable users that are continuously location-tracked to later access the resulting data logs for attestation purposes. However, we stress that only giving the users access powers is not enough to create a balanced system in the sense of multilateral security. Moreover, parties with supervisory functions, which we call verification authorities, should be able to harness the logs, to verify compliance to organizational and legal traceability needs. In this paper, we argued that both goals can conjointly be achieved. The basic idea to achieve this is to encode implicit attributes into the pseudonyms used, both for the user and for third parties. Therefore, we employ cryptographically secure PRNGs and threshold encryption techniques to generate pseudonyms for users of a location tracking system. To analyze collected data, we devise protocols that allow to exercise analysis functions in a fine-grained and selective manner, while preserving user privacy. Next, due to the use of non-interactive zero knowledge techniques, misuse of the analysis functionalities is also detectable.

Addressing inherent tradeoffs already in the construction of technologies that are pervading everyday life more and more is essential for their acceptance. We believe that the concepts proposed in this paper are very useful to realize multilateral secure UbiComp applications. Currently we are working on an implementation of a system that demonstrates the feasibility of our approach on a large scale. One practical aspect that we consider is the use of elliptic curves as underlying algebraic groups for the cryptographic operations. This will allow us to implement pseudonyms with a small size. While our current threshold construction allows to represent a large range of inter-organizational distributions of duties and powers, we will also consider to integrate a global attorney functionality into our security design in the future.

Acknowledgment

This work was supported by the German Federal Ministry of Education and Research in the context of the project SoKNOS and by CASED (www.cased.de). The authors are responsible for the content of this publication.

References

[1] M. Weiser, "The Computer for the 21st Century," *Scientific American*, vol. 265, no. 3, pp. 94–104, 1991.

[2] F. Flentge, S. G. Weber, A. Behring, and T. Ziegert, "Designing Context-Aware HCI for Collaborative Emergency Management," in *Int'l Workshop on HCI for Emergencies in conjunction with CHI 2008*, 2008.

[3] M. Satyanarayanan, "Privacy: The Achilles Heel of Pervasive Computing?" *IEEE Pervasive Computing*, vol. 2, no. 1, pp. 2–3, 2003.

[4] S. G. Weber, A. Heinemann, and M. Mühlhäuser, "Towards an Architecture for Balancing Privacy and Traceability in Ubiquitous Computing Environments," in *Int. Workshop on Privacy and Assurance (WPA-2008) at 3rd Int. Conference on Availability, Reliability and Security (ARES 2008)*. IEEE Computer Society, 2008, pp. 958–964.

[5] S. G. Weber, "Securing first response coordination with dynamic attribute-based encryption," in *Proceedings of 2009 World Congress on Privacy, Security, Trust and the Management of e-Business (CONGRESS 2009)*. IEEE Computer Society, 2009, pp. 58 – 69.

[6] K. Rannenberg, "Multilateral security a concept and examples for balanced security," in *NSPW '00: Proceedings of the 2000 workshop on New security paradigms*. New York, NY, USA: ACM, 2000, pp. 151–162.

[7] U. Flegel, *Privacy-Respecting Intrusion Detection*. Springer, 2007.

[8] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms." *IEEE Transactions on Information Theory*, vol. 31, no. 4, pp. 469–472, 1985.

[9] R. Cramer, R. Gennaro, and B. Schoenmakers, "A secure and optimally efficient multi-authority election scheme," in *Advances in Cryptology - EUROCRYPT 97*, ser. Lecture Notes in Computer Science, vol. 1233. Springer-Verlag, 1997, pp. 103–118.

[10] T. P. Pedersen, "A threshold cryptosystem without a trusted party (extended abstract)," in *Advances in Cryptology - EUROCRYPT 91*, ser. Lecture Notes in Computer Science, vol. 547. Springer-Verlag, April 1991, pp. 522–526.

[11] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[12] A. Fiat and A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems," in *Advances in Cryptology - CRYPTO 86*, ser. Lecture Notes in Computer Science, vol. 263. Springer-Verlag, August 1986, pp. 186–194.

[13] R. Cramer, I. Damgard, and B. Schoenmakers, "Proofs of partial knowledge and simplified design of witness hiding protocols," in *Advances in Cryptology - CRYPTO 94*, ser. Lecture Notes in Computer Science, vol. 839. Springer-Verlag, 1994, pp. 174–187.

[14] J. C. Benaloh, "Verifiable secret-ballot elections," Ph.D. dissertation, Yale University, Department of Computer Science, September 1987.

[15] M. Jakobsson and A. Juels, "Mix and match: Secure function evaluation via ciphertexts (extended abstract)," in *ASIACRYPT 00*, ser. Lecture Notes in Computer Science, vol. 1976. Springer-Verlag, 2000, pp. 162–177.

[16] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–88, 1981.

[17] C. Park, K. Itoh, and K. Kurosawa, "Efficient anonymous channel and all/nothing election scheme," in *Advances in Cryptology - EUROCRYPT 93*, ser. Lecture Notes in Computer Science, vol. 765. Springer-Verlag, 1993, pp. 248–259.

[18] J. Furukawa and K. Sako, "An efficient scheme for proving a shuffle," in *Advances in Cryptology - CRYPTO 01*, ser. Lecture Notes in Computer Science, vol. 2139. Springer-Verlag, 2001, pp. 368–387.

[19] D. Chaum and T. P. Pedersen, "Wallet databases with observers," in *Advances in Cryptology - CRYPTO 92*, ser. Lecture Notes in Computer Science, vol. 740. Springer-Verlag, 1993, pp. 89–105.

[20] A. Heinemann, "Collaboration in opportunistic networks," Ph.D. dissertation, Darmstadt University of Technology, 2007.

[21] J. Krhovjak, P. Svenda, V. Matyas, and L. Smolik, "The sources of randomness in smartphones with symbian os," in *Security and Protection of Information 2007*, 2007, pp. 87–98.

[22] D. Kesdogan, H. Federrath, A. Jerichow, and A. Pfitzmann, "Location management strategies increasing privacy in mobile communication," in *Proceedings of the 12th IFIP International Information Security Conference SEC'96*. Chapman Hall, 1996, pp. 39–48.

[23] K. Borcea-Pfitzmann, E. Franz, and A. Pfitzmann, "Usable presentation of secure pseudonyms," in *Digital Identity Management*, 2005, pp. 70–76.

[24] D. Henrici and P. Müller, "Hash-based enhancement of location privacy for radio-frequency identification devices using varying identifiers," in *PERCOMW '04: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops*. IEEE Computer Society, 2004.

[25] A. R. Beresford and F. Stajano, "Location Privacy in Pervasive Computing," *IEEE Pervasive Computing*, vol. 02, no. 1, pp. 46–55, 2003.

[26] C. Delakouridis, L. Kazatzopoulos, G. F. Marias, and P. Georgiadis, "Share The Secret: Enabling Location Privacy in Ubiquitous Environments." in *LoCA*, 2005, pp. 289–305.

[27] J. Biskup and U. Flegel, "Threshold-based identity recovery for privacy enhanced applications," in *ACM Conference on Computer and Communications Security*, 2000, pp. 71–79.

[28] B. Schoenmakers and P. Tuyls, "Client-Server Trade-Offs in Secure Computation," in *Security, Privacy and Trust in Modern Data Management*, M. Petkovic and W. Jonker, Eds. Springer, 2007, pp. 197–212.