An Authorised Pseudonym System for Privacy Preserving Location Proof Architectures

Janaka Seneviratne\textsuperscript{1} Udaya Parampalli\textsuperscript{1} Lars Kulik\textsuperscript{1}

\textsuperscript{1} Computing and Information Systems University of Melbourne

Email: janakaw@student.unimelb.edu.au Email: udaya@unimelb.edu.au Email: 1kulik@unimelb.edu.au

Abstract

An emerging class of Location Based Services (LBSs) needs veriﬁed mobile device locations for service provision. For example, an automated car park billing system requires veriﬁed locations of cars to conﬁrm the place and the duration of parked cars. Location Proof Architectures (LPAs) allow a user (or a device on behalf of its user) to obtain a proof of its presence at a location from a trusted third party. A major concern in LPAs is to preserve user location privacy. To achieve this a user’s identity and location data should be maintained separately with additional measures that prevent leaking sensitive identity and location data. In this paper, we present a privacy preserving LPA in which users appear under pseudonyms. Our main contribution is a third party free pseudonym registering protocol based on blind signature schemes. We show that our protocol allows to build a pseudonym system with a guaranteed degree of privacy agreed at the time of pseudonym registration. We also demonstrate that a pseudonym can be authenticated across different organizations in an LPA. Our system ensures that (i) only authenticated users can register their pseudonyms, (ii) the pseudonyms have a consistent degree of privacy at the point of registration and (iii) a user cannot take another user’s pseudonym.

Keywords: Privacy, Location Proof Architecture, Pseudonym

1 Introduction

In Location Proof Architectures (LPA) (1; 2; 3), users provide their veriﬁed locations to obtain a Location Based Service (LBS). In the previously proposed LPAs, a trusted third party uses a Location Proof Issuer (LPI) that veriﬁes a user location and provides a Location Proof (LP). A user obtains a service from an LBS by submitting an LP. In such an architecture, users need to be able to protect their location privacy (4; 5) from a third party that veriﬁes the user location (i.e. LPI) and possibly even from the LBS. We distinguish LBSs that require the user identity from those that only need a pseudonym together with veriﬁed location data. We will describe such LBSs later in this section.

In LBSs which require veriﬁed user location data, it is most important that user location privacy is preserved in an LPA. A greater degree of location privacy can be achieved by keeping a user anonymous and using imprecise instead of exact location data (4; 6). In this paper we propose a pseudonym system that can be used in LPAs to improve location privacy by maintaining user anonymity.

We illustrate in Figure 1, how we improve privacy of LPAs by eliminating the need for trusted organizations through the use of our pseudonym system. A user registers his pseudonym with a Pseudonym Issuing Organization (PIO) and obtains a certificate of the pseudonym. A key difference of our system is that a user is able to obtain a pseudonym without a trusted organization who can link a user’s true identity to the user pseudonym. When the user needs an LP he submits his certiﬁed pseudonym together with his location to an LPI and proves the knowledge of a secret associated with the pseudonym. The LPI veriﬁes the location and issues the user with an LP. The user submits the LP together with his certiﬁed pseudonym to an LBS to obtain a service. In contrast to existing LPAs which use two independent trusted organizations for authentication and location veriﬁcation, we only use a single organization for issuing LPs. Our LPA can serve a sub-group of LBSs which need veriﬁed user

This research was supported under Australian Research Councils Discovery Projects funding scheme (project number DP110100757). We acknowledge Peter Lee for proof reading.

Copyright ©2014, Australian Computer Society, Inc. This paper appeared at the Australasian Information Security Conference (AISC 2014), Auckland, New Zealand, January 2014. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 149, Udaya Parampalli and Ian Welch, Ed. Reproduction for academic, not-for-profit purposes permitted provided this text is included.

Figure 1: An overview of our pseudonym based Location Proof Architecture (LPA). PID: Pseudonym. A user obtains a certiﬁed pseudonym from PIO. Then obtains an LP from an LPI by submitting PID. In the end, obtains a service from an LBS by submitting the LP and PID. User remains anonymous throughout all the steps.
location but does not require true user identity. The following scenarios are examples for such LBSs.

Scenario 1: Assume a car park that is equipped with sensors. A user parking his car obtains location proofs when entering and leaving the car park. Such location proofs could be used for automated billing purposes. The car park administrators need a reliable way to track the usage of car parks and charge users accordingly. On the other hand a user might not want to reveal all the car parks he visited and the duration spent at each place. These requirements call for a private and authorised pseudonym registering system which does not rely on a trusted third party.

Scenario 2: Currently, location based games have gained popularity. For example, in the Geo-Cache game (7) users hide valuable items at certain positions and others should find the items. At the moment, these applications rely on user honesty to obtain correct locations. To maintain the credibility of the game, an administrator can ask for LPs from its users. The users in the game might want to keep their location information private.

In both scenarios a user’s true identity is not compulsory to obtain the service. It is important that a user provides verified location data. The organizations (i.e., LPIs and LBSSs) that use pseudonyms in place of true user identities need to ensure that only authenticated users are able to obtain pseudonyms. The users need to be able to stay anonymous with a known degree of privacy.

Also, it is important that a user cannot appear under another user’s pseudonym in an LPA. For instance, in the car park example, an adversary might want to pretend as another user to avoid the fee. Considering these requirements, we provide a precise description of properties of our authorised pseudonym system in Section 2.

In this paper, we present a pseudonym registering protocol based on blind signatures schemes without requiring a trusted organization that can link a user with his or her pseudonym. In addition, our system provides a consistent degree of privacy (i.e., similar to privacy guarantee offered by $k - \text{anonymity}$ (8)) for users joining the system at different times at the time of registration. We explain our protocol in Section 3.

A problem associated with pseudonyms is that over time the degree of privacy decreases (6). This is due to the fact that an adversary’s background knowledge increases over time. We discuss measures to ensure that the degree of privacy stays over a threshold in Section 7.

Furthermore, we show that pseudonyms can be authenticated using Zero Knowledge (ZK) proofs. Using ZK proofs in contrast to sharing credentials with an organization, protects users from untrustworthy organizations. For example, although a corrupt LPI can issue a false LP on behalf of a user, the LPI cannot prove the pseudonym used in the LP to the LBS as it does not have secret for the pseudonym. In addition, the ZK based authentication protocol allows a user to use the same pseudonym across different organization without needing a “single-sign-on" infrastructure.

We make following contributions in the paper. We provide a pseudonym system based on blind signature schemes that does not need a trusted third party to register pseudonyms. In addition, we show that pseudonyms can be authenticated with the organizations in an LPA while maintaining user anonymity. We also provide a consistent degree of privacy associated with the pseudonyms at the time of registration.

This paper is organized in the following manner. In Section 2, we provide a precise description of our pseudonym system. We explain our pseudonym registering protocol in Section 3. In Section 4, we show how to use pseudonyms within an existing zero-knowledge based authentication protocol. In Section 5, we explain how our pseudonym registering protocol achieves the properties described in Section 2. In Section 6 we provide an analysis of privacy of our system. In Section 7, we explain how to maintain a degree of privacy. In Section 8, we present related work. We conclude in Section 9.

2 Authorised Pseudonyms

In the following, we provide a precise description of the properties of our third party free authorised pseudonym registering protocol.

- Authorisation: An authority called Pseudonym Issuing Organization (PIO) allows only authenticated users to register pseudonyms in the system. The PIO selects a user sub-set $U_m$ from the user population $U$ to register pseudonyms in each protocol round. Each user $u_i$ in $U_m$ is authenticated and issued receipts to take part in the pseudonym registering protocol.

The PIO may not allow users who have a history of billing rule violations to enter the system. Because of this property, LBSSs can trust the users who use pseudonyms in the system, which in turn allows LBSSs to be more flexible in their service offerings. For example, an LBS may allow longer billing cycles to its users.

The PIO can control the number of pseudonyms issued per user through the number of receipts issued per user. In some LPAs, a user may require multiple pseudonyms (9) to maintain a higher degree of privacy.

- $m$-Unlinkability: Consider a user sub-set $U_m = \{u_i\}$ where $|U_m| = m$. The corresponding pseudonym set is $P_m = \{p_i\}$. If, at the time of pseudonym registration, the probability with which the PIO is able to link a user $u_i \in U_m$ with the pseudonym $p_i \in P_m$ is $1/m$, we say that a pseudonym system is $m$-unlinkable at the time of pseudonym registering.

This guarantees a degree of privacy at the time of pseudonym registration. Inherently, pseudonyms lose the degree of privacy over time (6). We discuss how to address this problem within our system in Section 8.

- Unforgeability: We identify two forms of unforgeability. Online-Unforgeability refers to a user’s ability to prove the ownership of a pseudonym, remaining anonymous while Offline-Unforgeability refers to a user’s ability to link the user to a pseudonym.

  - a) Online-Unforgeability (authentication): An adversary user $u_a$ cannot appear under a pseudonym of another user $u_b$ to an organization (i.e. LPI or LBS) unless the second user $u_b$ willingly shares the secret $x_a$ associated with the user pseudonym $p_a$.

  - b) Offline-Unforgeability: In case of a dispute, a user $u_a$ is able to reveal the pseudonym $p_a$ at the user’s own risk of losing privacy. A user $u_a$ cannot take the ownership of any other user’s $u_b$ pseudonym $p_b$, without the consent of the second user $u_b$. In a case where a user $u_b$ believes to be treated unfairly by an LBS, the user $u_a$ has the ability to defend his claims by appearing under the true identity. On the other hand, an adversary user cannot mis-use this privilege.

Our pseudonym registering protocol allows users to register pseudonyms without a trusted third party which is not available in the existing pseudonym systems (10; 11;...
In our system, the PIO cannot identify an actual user $u_i$, based on his pseudonym $p_i$, from within a group of $m$ users without colluding with another of the users at the time of pseudonym registration. We show that pseudonyms can be authenticated across organizations (i.e., LPI and LBSs) in an LPA using an existing Zero Knowledge Proof (ZKP) based authentication protocol.

We assume that the goal of users in our system is to preserve their location privacy by using a pseudonym. A user $u_i$ can share his pseudonym $p_i$ with another user $u_j$ at the risk of revealing his locations to $u_j$ and losing the credibility associated with $p_i$. In such a case, the user $u_j$ will have to share the credential (i.e. a secret $s_j$) associated with the pseudonym $p_j$ with the user $u_j$. Also, the organizations (i.e. LPIs and LBSs) in our LPA provide their services solely based on a user location but do not need the true identity of the user, unless there is a dispute with a user.

3 Pseudonym Registration

To achieve a guaranteed degree of privacy the pseudonym registration protocol is executed in two phases. In phase 1, $m$ users $u_i \in U_m$ obtain receipts for registering pseudonyms from the PIO. In phase 2, all $m$ users $u_i$ together anonymously register their pseudonyms with the PIO.

In the presentation of our protocol, we use an RSA based blind signature scheme without loss of generality. This can be replaced by any secure blind signature scheme, which (i.e., a blind signature) on a number $n_i$ from the PIO, without letting the PIO know the number $n_i$. This can be replaced by any secure blind signature scheme without loss of generality.

We note that to make RSA signatures secure against RSA multiplicative property attack, the pseudonyms should be formatted using a secure padding scheme such as Optimal Asymmetric Encryption Padding (OAEP).

3.1 Obtaining Receipts from the PIO

In phase 1 of our protocol, a user $u_i$ authenticates to the PIO and obtains two types of receipts called a regular receipt $r_i$ and a collision resolving receipt $cr_i$. See Table 1.

![Figure 2: The protocol steps in phase 1. A user obtains a regular receipt $r_i$ and collision resolving receipt $cr_i$ from the PIO. The user is identified by the PIO in this phase.](image)

The PIO’s RSA public key and private key are $(n, e)$ and $d$ respectively. The PIO maintains two sets of numbers $R$ and $T$ where for $r \in R$ and $t \in T$, $0 \leq r + t \leq n - 1$.

<table>
<thead>
<tr>
<th>1</th>
<th>User Authenticates</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$r_i$</td>
</tr>
<tr>
<td>3</td>
<td>$n_i k^e$</td>
</tr>
<tr>
<td>4</td>
<td>$(n_i k^e)^d$</td>
</tr>
<tr>
<td>5</td>
<td>$(r_i + n_i) k^e$</td>
</tr>
<tr>
<td>6</td>
<td>$(r_i + n_i)^d k^e$</td>
</tr>
</tbody>
</table>

Table 1: A user obtains two receipts a regular receipt and a collision resolving receipt from PIO. The user is not anonymous in this stage.

Next, PIO displays a number panel containing the set $T$ to each user $u_i$. A user $u_i$ selects a number $n_i$ from the panel without acknowledging the PIO. Then the user $u_i$ selects another number $k$ in random such that $0 \leq k \leq n - 1$ and $gcd(n,k) = 1$.

In the next step, the user gets the numbers $n_i$ and $n_i + r_i$ blindly signed by the PIO. We call the blindly signed $n_i$, a regular receipt $r_i$ and blindly signed $n_i + r_i$ a collision resolving receipt $cr_i$. This concludes the phase 1. See Figure 2.

If the LPA requires multiple pseudonyms per user, the user may be issued with multiple receipts in this phase.
Registering Pseudonym by Users

In phase 2, the $m$ users in $U_m$ register their pseudonyms using the regular receipt $rr_i$, in random order. In this step, the users remain anonymous. See Table 2.

Firstly, a user $u_i$ selects a pseudonym $p_i$ of his choice. Then submits $p_i$ and the regular receipt $rr_i^d$ to the PIO. In this phase, the $m$ users submit their pseudonyms $p_i$ and regular receipts $rr_i$ in random order. In addition a user $u_i$ can deposit a bond together with his pseudonym $p_i$ to increase the trustworthiness of the pseudonym to organizations.

In order to be a valid protocol round, in phase 2 the following four conditions must be satisfied. If this is the case each user’s $u_i$ pseudonym is registered in the system and the user is given a signed pseudonym (i.e. In this case, PIO uses a regular signature scheme as opposed to blind signatures used for issuing receipts). The four conditions are, i) the signature on a regular receipt $rr_i$ is valid, ii) the $p_i$ is not already registered, iii) the same number $n_i$ in the regular receipt is not submitted by two or more users and iv) all $m$ users in the group submit the $rr_i$. We explain these conditions in detail in Section 5. See Figure 3.

Next, we will consider how to handle if any of the conditions cannot be satisfied. Firstly, if any of the first two conditions do not satisfy, the protocol round is invalidated. The numbers submitted by the users are removed from $T$ and protocol starts again from phase 1. If any of the last two conditions do not satisfy, the PIO will generate new parameters for the blind signature and the protocol will be executed again from phase 1.

Using Pseudonyms together with existing Zero-Knowledge based Identification protocols

In this section we show that by defining a pseudonym as a public key in Fiat-Shamir zero knowledge proof authentication protocol, the pseudonym can be authenticated to different organizations (i.e., LPI and LBSs) in an LPA by using a secret held by a user. Because, a user pseudonym is authenticated, his true identity remains hidden from the organizations preserving privacy. Only location information of the user is known to the organizations.

An LPI or an LBS can associate a user identity with his pseudonym, if they know the user’s home location. In LPA design, necessary precautions are taken to obfuscate a user’s home location. In this paper, we do not consider privacy attacks possible by using such location patterns.

Table 2: All $m$-users register their pseudonyms in parallel using the regular receipt. A users receive a certified pseudonym from PIO if the user does not collide and his pseudonym is unique. The users remain anonymous in this phase.

We show that the pseudonyms in our system can be defined as a public key in Fiat-Shamir identification protocol as presented in (14) (i.e., we use a basic version for demonstration). See Table 3. In the full version of Fiat-Shamir identification protocol (14), a user’s public key (i.e. his registered pseudonym) $v$ is replaced by $(v_1, ..., v_k)$.

On-time setup phase, a trusted center selects and publishes the modulus $n = pq$ ( $n$ is different to that used in Section 3) with large prime factors $p$ and $q$ and keep the factors $p$ and $q$ secret, We point out that unlinkability is not affected by using such a one-time trusted center.

In phase 2 of our protocol, a user $u_i$ selects his
5 Analysis against the requirements

In this section, we show how our protocol satisfies the properties of an authorised pseudonym registering protocol given in Section 2. First we explain a number of properties of regular receipts \(rr_i\) and how we satisfy these properties in a protocol round. Then we show how the regular receipts allow us to obtain the properties in Section 2.

5.1 Properties of Regular Receipts

To obtain Authorisation and \(m\)-Unlinkability properties we need the following properties from a regular receipt \(rr_i\) issued in phase 1. In Section 5.2, we show how these properties are used to obtain Authorisation and \(m\)-Unlinkability properties.

- Authorised: A receipt \(rr_i\) is issued to a user \(u_i\) authenticated by the PIO.
- Unforgeable: An adversary user cannot generate a forged receipt \(rr_i\) without the PIO.
- One-timeness: A user \(u_i\) can use \(rr_i\) only once. In other words, a user cannot repeatedly submit the same \(rr_i\) to register different pseudonyms.
- Phase-bound: A receipt \(rr_i\) can be used only in the \(m\)-user protocol round within which it was issued.
- Unlinkable: The PIO cannot identify a user \(u_i\) from the receipt \(rr_i\).

Table 3: We show that pseudonyms in our system can be defined according to the Fiat-Shamir identification protocol, which allows a user to authenticate the same pseudonym across different organizations.

We show that by executing the protocol steps in the following manner we are able to achieve the properties of a regular receipt \(rr_i\) listed above.

5.1.1 Step1: Obtaining a Regular Receipt

In our protocol, an authenticated user \(u_i\) obtains a regular receipt \(rr_i\) in following steps. (1) A user \(u_i\) selects a number \(n_i\) randomly from a set \(T\). The PIO maintains \(T\) so that the PIO has not signed any number in \(T\) before.

(2) The user \(u_i\) applies a blinding function \(c(n_i)\) to \(n_i\) only known to the user \(u_i\). (3) The user submits \(c(n_i)\) to the PIO and the PIO signs and return \(\hat{s}(c(n_i))\). where \(\hat{s}\) is signature function known only to the PIO. (4) The user applies the inverse function \(c^{-1}\) to obtain the regular receipt \(rr_i = \hat{s}(n_i) = c^{-1}(\hat{s}(c(n_i)))\).

5.1.2 Step2: Verifying a Regular Receipt

A user \(u_i\) anonymously submits the \(rr_i\) together with \(m\) other users so that the PIO cannot identify \(u_i\) based on time of submission. The PIO obtains \(n_i = s(\hat{s}(n_i)) = s(rr_i)\) by
applying inverse signature s to rr. The PIO accepts rr given that (1) \( n_i \in T \), (2) no two or more users submit the same \( n_i \), which we call a collision instance and (3) in addition to be a valid protocol round the PIO must receive \( m \) distinct rr. If a collision occurs or number of receipts \( rr \) received by the PIO is less than \( m \) the protocol round is re-executed from the start with new signing functions \( s \) and \( s' \).

**Theorem 1:** If regular receipts \( rr \) are obtained and validated according to Step1 and Step2, they satisfy the properties of regular receipts listed in Section 5.1.

**Proof:** In the following we explain how we achieve the above properties. The Authorised property follows simply from the fact that only authenticated users are issued \( rr \). Unforgeable property follows due to the unforgeability of underlying signature function \( s' \) (i.e., conservation of signature property of blind signatures) (13). One-time ness is guaranteed in the following manner. A user \( u_i \) selects a number \( n_i \in T \) that has not been signed before. When a user submits \( rr, n_i \) within is removed from the set \( T \). This ensures that \( rr_i \) can be used only once.

**Phase-bound** property is obtained by requiring that all \( m \) users must submit the \( rr_i \). Otherwise an adversary user \( u_i \) may withhold his submission and use \( rr_i \) in another round. In addition, in a collision situation, the PIO cannot know whether the user choices genuinely coincided or a sub group of users colluded to share the same receipt \( rr_i \). If all users submitted \( rr_i \) containing distinct \( n_i \), the PIO is guaranteed that all users who were authenticated in phase 1 genuinely participated in the phase 2 of the protocol. If number of received \( rr_i \) is less than \( m \) or a collision is detected the protocol round is re-executed with new signing functions \( s \), \( s' \). This ensures that no previous \( rr_i \) are valid anymore.

Unlinkable property follows from the properties of blind signature scheme.

### 5.2 Properties of Authorised Pseudonym System

**Authorisation:** This property follows directly from those of a regular receipt \( rr_i \).

We note that, at anytime PIO can remove a pseudonym if the owner violates the rules of the system. The PIO maintains a revoked user list. An organization that uses a certified pseudonym will check the revoked list prior to using the pseudonym.

**m-Unlinkability:** Due to the unlinkable property of a regular receipt \( rr_i \) the PIO cannot identify the owner of \( rr_i \) from the content of \( rr_i \). Because in phase 2, \( m \) users submit \( rr_i \) in a random order the PIO cannot identify the owner of a \( rr_i \) from among \( m \) users, without colluding with \( m - 1 \) other users.

**Unforgeability:** a) **Online-Unforgeability (authentication):** We showed in Section 4 that by generating a pseudonym as the public key of Feige-Fiat-Shamir authentication protocol, it can be authenticated to different organizations in LPA using the secret kept with a user.

b) **Offline-Unforgeability:** At anytime, a user \( u_i \) can reveal his ownership of a pseudonym \( p_i \) by submitting both the regular receipt \( rr_i \) (i.e. which contains \( n_i \) and collision resolving receipt \( cr \) (i.e., which contains \( r_i + n_i \)) to the PIO.

By validating that the receipt number \( r_i \) issued in phase 1 to the user \( u_i \) is as same as that is in collision resolving receipt \( cr_i \), PIO knows \( u_i \) is a valid user. By using \( n_i \) in \( rr_i \), PIO is able to obtain the user’s pseudonym \( p_i \). Therefore, a user \( u_i \) is able to reveal his pseudonym at his own risk of losing privacy.

An adversary user \( u_j \) cannot pretend another user \( u_i \)’s pseudonym \( p_i \) because \( u_i \) cannot forge both regular receipt \( rr_i \) (i.e., \( u_i \)) and collision resolving receipt \( cr_i \). (i.e., \( r_i' + n_i' \)) of user \( u_i \). This follows from the conservation of signature property of blind signature scheme (13).

### 5.3 A known attack on blind signature schemes

We discuss a known attack on blind signature schemes and show that this attack does not affect our system.

A known attack on blind signature schemes is RSA blinding attack. In this attack a malicious user is able to learn a message encrypted with the PIO’s public key. The attack is performed in the following manner.

Assume the encrypted message of plain text \( m \) is \( m' = m^e \). The user calculates the blinding factor \( k' \) and send \( mk' \) to get blindly signed by the PIO. PIO returns the signed version as \( m'' = m'^d \cdot k' \). As \( k'd = k \), the user can obtain the plain text \( m = med = m'' / k \).

In order to prevent this attack in our protocol, PIO should never use its public key for encryption purposes.

### 6 A Consistent Degree of Privacy for Pseudonyms

We discuss the degree of privacy of pseudonyms at the time of registration provided in our protocol. We maintain a consistent degree of privacy for users joining at different times. Also we provide an expression for system parameter \( p (collision) \) in terms of user batch size \( m \) and size of set \( T, n \).

![Figure 5: A plot of collision probability versus size n of T. We maintain m = 100. We observe that collision probability ≤ 1% when n = 10^6.](image)

We provide a guaranteed degree of privacy at the time of registering pseudonyms. We can provide such a guarantee only at the point of registration due to the fact that, inherently pseudonyms lose the degree of privacy over time (6). We discuss how to maintain a degree of privacy over a threshold in Section 8.

Also, we empirically determine the optimal size \( n \) of set \( T \) so that we can maintain \( p (collision) \) below a threshold. It is important to maintain the set \( T \) with a constant size \( n \) in order to achieve a consistent degree of privacy. Otherwise user collision rate (i.e. two users selecting the same number) will fluctuate and as a result we cannot guarantee \( m \)-anonymity. We note that in our two phase pseudonym registering protocol, once a user has selected a number \( n_i \) from set \( T \), that number should be replaced for a new set of \( m \) numbers.

We can keep the set \( T \) with a constant size \( n \) by maintaining \( T \) as an indexed array where users pick \( n_i \) as an item in the array. When users submit the regular receipt
they must provide the index $i$ at which they found the number $r_u$. The PIO can refill the array indexes used in previous round with numbers from the universal set before starting the next round.

If the user batch size is $m$ and there are no collisions in a protocol round, we achieve a degree of privacy $m$-anonymity. This degree of privacy cannot be guaranteed if users collide (i.e. more than one user selects the same number from the set $T$).

If the size of set $T$ is $n$, the number of collision free instances out of all possibilities $n^m$ is $n(n - 1) . . . (n - (m - 1))$. This can be obtained by the fact that two users do not select the same number in the sequence of numbers selected by the users.

Assuming that users select numbers according to the uniformly distribution, we define collision probability as the probability that at least one collision occurs in a protocol round. In other words, this gives the probability that privacy is below $m$-anonymity. The collision probability is given by the following equation.

$$p(\text{collision}) = n^m - n(n - 1) . . . (n - (m - 1))/n^m \quad (1)$$

This can be re-written in the following form.

$$p(\text{collision}) = 1 - (1 - 1/n) . . . (1 - (m - 1)/n) \quad (2)$$

Assuming $1/n$ is relatively small and hence $(1/n)^2$ and higher order terms are negligible, we obtain the following equation (i.e. as an approximation).

$$p(\text{collision}) = 1 - 1 + (1 + . . . + (m - 1))/n \quad (3)$$

$$p(\text{collision}) = m(m - 1)/2n \quad (4)$$

$$n > 500 \times (m^2 - m) \quad (5)$$

In a practical implementation, we need to maintain collision probability below 0.1% so that $m$-anonymity is guaranteed to the users for most of the time. Therefore, $m$ and $n$ should be selected so that $0 \leq p(\text{collision}) < 0.1\%$.

The equation 4 implies that the collision probability increases when we have more users $m$ in a batch and decreases as the size $n$ of set $T$ is increased. In order to maintain a low collision rate we must select $m$ and $n$ according to the inequality given in Equation (5).

In the following, we show empirical results which demonstrates these effects.

We performed our experiments using MATLAB. To simulate a single round of our protocol, we draw $m$ numbers in random from discrete uniform distribution (i.e. of size $n$). In both the experiments presented, we perform 100 rounds.

In our first experiment, we observe that the percentage of collisions (i.e. all colliding rounds as a percentage of total number of rounds) for different $n$. We maintain $m = 100$. See Figure 5 and 6.

As we expect in equation (4) when the number of users are constant collision probability goes down with increasing $n$. Importantly, we observe that the collision probability becomes less than one percent only when $n$ is equal to million for a batch of 100 users.

In a second experiment, we change the batch size $m$ and observe minimal size $n$ (in terms of powers of 10) that makes collision probability $\leq 1\%$. See Figure 7.

We can observe that as the batch size $m$ increases, collision probability also increases. As a result, the threshold (in terms of powers of 10) at which collision probability becomes 1% increases with the batch size. This is as expected in equation (4).

## 7 Maintaining a Degree Privacy of Pseudonyms

In this section, identifying the information available to an adversary, we discuss ways that users can lose privacy of pseudonyms overtime. Also we discuss how to maintain a degree of privacy of pseudonyms at a higher level for a longer duration in the presence of an adversary who has such information.

We consider that such an analysis is necessary to present a complete pseudonym system that provides a consistent level of privacy for its users. We leave a comprehensive analysis as a future work.

### 7.1 Privacy Losses

We will examine the information available to colluding organizations and how they can exploit this information to identify a user using his pseudonym.

All location updates $(l_i, t_j)$ from a given pseudonym $p_i$ are available to LPs. Each time a user accesses a service $s_j$, the information $(p_i, s_j)$ is recorded by the service $s_j$. 

![Figure 6: A plot of collision probability versus size $n$ of $T$. We maintain $m = 1000$. We observe that collision probability $\leq 1\%$ when $n = 10^8$.](image)

![Figure 7: A plot of optimal size $(n)$ of $T$ versus size of user batches $m$.](image)
7.1.1 Service Access Matrix

We represent the information available to all colluding services as Service Access Matrix \( M \). Each time a pseudonym obtains a service, the service records the pseudonym. The services are able to establish the frequency of access for each service by a certain user during a period of time.

We show a Service Access Matrix in equation 6. We consider the total service count as \( v \) and total user count as \( w \). Also \( s_j \) is a service and \( p_i \) is a pseudonym. In addition, \( f_{(i,j)} \) is the frequency that the pseudonym \( p_i \) used the service \( s_j \).

\[
M = \begin{pmatrix}
    s_1 & s_2 & \ldots & s_v \\
p_1 \begin{pmatrix} f_{(1,1)} & f_{(1,2)} & \ldots & f_{(1,v)} \\ f_{(2,1)} & f_{(2,2)} & \ldots & f_{(2,v)} \\ \vdots & \vdots & \ddots & \vdots \\ f_{(w,1)} & f_{(w,2)} & \ldots & f_{(w,v)} \end{pmatrix}
\end{pmatrix}
\] (6)

For example, consider a health care service that is dedicated to heart disease patients. The frequency of use of the service is high for patients suffering from a heart disease but will be minimal for other users. If an adversary knows that the user \( A \) is suffering from a heart disease, the adversary is able to narrow down the subset of pseudonyms to which user \( A \) pseudonym belongs.

We obtain a normalized form of the above matrix \( M' \) by dividing the each frequency element by the sum of all frequencies. (Equation 7).

\[
\hat{f}_{(i,j)} = \frac{f_{(i,j)}}{\sum_{i,j} f_{(i,j)}}
\] (7)

Maximum privacy is guaranteed if the following condition is satisfied for any service \( j \) and pseudonym \( i \).

\[
\hat{f}_{(i,j)} = c_i \quad \forall i, j
\] (8)

\( \hat{f}_{(i,j)} \) (i.e. column-wise frequencies) gives information about users as to which user uses a given service more frequently and more sparsely. An adversary is able to correlate this information with background knowledge (i.e. in the previous example, background knowledge is that certain users belong to the super rich category) to narrow down the number of possible pseudonyms for a given user.

We identify the information that an adversary will gain by observing a row in the normalized matrix \( M' \) as the service usage profile. An adversary is able to use different type of background knowledge together with the service usage profile to categorise the pseudonyms. Such categorisation facilitates the task of an adversary in linking a pseudonym to its actual user.

We can measure the information revealed to an adversary (i.e. capturing the intuition that higher variability in \( M' \) reveals more information to an adversary) through \( M' \) as following. To maintain a high level of privacy \( H(\hat{f}) \) should be maintained closed to the maximum.

\[
H(\hat{f}) = -\sum_{i,j} \hat{f}_{(i,j)} \log 2(\hat{f}_{(i,j)})
\] (9)

7.1.2 Location Updates

An adversary will find that organizing location updates as a relation between a location update and different pseudonyms which send the same location update is useful to achieve the goal of linking a user to his true identity.

\[
L_i = (l_i, t_i) \rightarrow \{p_{i1}, p_{i2}, \ldots, p_{iq}\}
\] (10)

For example, consider the case where users obtain location proofs from a car park in the night. There will be only a few users who obtain location proofs around the same time. If an adversary has the background knowledge that a user regularly visits the same car park at that time, he is able to identify the user pseudonym using the above relation.

We observe that as the number of users \( (f'_i) \) who make similar location updates goes down, an adversary is able to link a user to his pseudonym more easily. The following relation expresses the frequencies of location updates.

\[
L_i \rightarrow f'_i
\] (11)

We can obtain a normalized relation by considering all location updates within a certain duration. See equation 10.

\[
\hat{f}'_{(i,j)} = \frac{f'_i}{\sum_i f'_i}
\] (12)

The information revealed to an adversary through the above relation will be minimum if \( f'_{i} \) is distributed evenly. We capture this property using the entropy of \( f'_{i} \) as shown in the following equation.

\[
H(\hat{f}') = -\sum_{i,j} \hat{f}'_{(i,j)} \log 2(\hat{f}'_{(i,j)})
\] (13)

7.2 Maintaining User Privacy

A major drawback of pseudonyms is that they are vulnerable to attacks by adversaries with background knowledge. Over time as the background knowledge of the adversary increases privacy provided by pseudonyms decreases (6). For example, consider the situation where an adversary observes that a user visits the same car park at night when there are only a few other cars parked. By colluding with the server, the adversary is able to identify a smaller set of pseudonyms in which the user’s pseudonym is included. It is expected that over time the adversary will be able to correctly identify the user’s pseudonym.

In order to overcome the degradation of privacy of pseudonyms over time (6), the pseudonyms need to be changed at regular intervals. We explain a procedure to achieve this in our system.

After a certain threshold duration (i.e. for the moment we will assume this duration is empirically determined), a user forms a batch of \( m \)-users with other users who are registered in the pseudonym system and requiring to change pseudonyms in a manner that maximizes \( H(\hat{f}) \) and \( H(\hat{f}') \).

The users in the batch execute phase 1 and 2 of pseudonym registration protocol (i.e. given in Section 4) to register new pseudonyms, using their existing pseudonyms for authentication purposes in place of the actual identity.

In (9), the authors use multiple pseudonyms per user to obtain a higher degree of user anonymity in the process of acquiring location proofs. In their approach, users change their pseudonyms according to Poisson distribution, when they update the server with a new LP.

Use of multiple pseudonyms per user can be enforced within our pseudonym registration system. In phase 1, a user obtains multiple regular receipts and in phase 2 registers multiple pseudonyms using those receipts. In order to maintain trustworthiness of pseudonyms to the organizations who use pseudonyms, a user may be required to place a bond together with each of his pseudonyms.

If we allow \( p \) pseudonyms per user, the anonymity of a pseudonym at the point of initial pseudonym registration becomes \( m \times p \).
8 Related Work

In (9), authors provide a privacy-preserving LPA using multiple pseudonyms per user. In their LPA, a user selects a pseudonym according to Poisson distribution from the user’s set of pseudonyms when he updates next location onto the server. An adversary who has access to the server where location proofs for all users are stored cannot easily identify the locations visited by a user.

In this paper, we present a pseudonym system that can be used in LPAs, which allows users to register pseudonyms without needing a trusted third party to register pseudonyms. We ensure a level of privacy for pseudonyms at the time of registering. In addition a user can register multiple pseudonyms using our system.

The authors in (12) present a pseudonym management scheme for vehicular ad-hoc networks (VANETs). In their system a number of distributed authorities need to track vehicles using the pseudonyms. A key requirement of their system is trace-ability where $m$ number of colluding authorities should be able to reveal the identity of a vehicle. In contrast, in our system we eliminate the need for any trusted party that can link a user to his pseudonym.

The authors use blind signature schemes to register pseudonyms similar to our system. In our system we guarantee a level of privacy at the time of pseudonym registration by allowing a group of $m$ users to register their pseudonyms at the same time. Also, we present a pseudonym authentication mechanism based on zero-knowledge protocols which is different to their certificate based authentication mechanism.

A pseudonym system that is motivated by the need for a user to remain anonymous when obtaining health services from different organizations is given in (10). In their system authors need to avoid the linkability of user information by colluding organizations. To achieve this they use different pseudonyms for the same individual for different organizations.

In our system, the user information that is verified by an LPI and used by an LBS is the user location. Even if a user uses different pseudonyms for different organizations the colluding organizations are able to identify the set of pseudonyms belonging to the same user by correlating pseudonyms provided at the same location. In our system, we use the same single or multiple pseudonyms per user to identify the same user anonymously across different organizations.

The pseudonym system provided in (10) uses a trusted organization to register pseudonyms. Considering the application scenarios addressed in this paper, we believe that the main privacy goal of users in our system is to remain anonymous with a guaranteed level of anonymity. To achieve a guaranteed level of anonymity at the time of pseudonym registration we provide a pseudonym registration mechanism eliminating the need for a trusted organization who knows the mapping between true user identities and their pseudonyms within a group of $m$-users.

The authors describe a pseudonym scheme (13), where a mediator issues a pseudonym to the users. The goal of their scheme is to provide a credential mechanism using which a user is able to transfer personal information obtained from one organization to another in a privacy preserving manner. The organizations identify a user only by his pseudonym and cannot link personal information with the actual user. In addition, different organizations have different pseudonyms for the same user making it difficult to track a user’s information by colluding organizations.

We observe such a credential mechanism is not useful in an LPA because the organizations (i.e. LPIs and LBSs) can identify the pseudonyms used by the same user by correlating pseudonyms corresponding to the same location. In addition we need a pseudonym system that allows users to easily change their pseudonyms because adversaries with background knowledge can link true identity of a user with his pseudonym over time. Also we point out that having a centralised mediator which must be contacted for each transaction creates a bottle neck in an LPA.

In the pseudonym authorization system provided in (11), a user registers a pseudonym with a trusted third party. In addition, the users obtain a pseudonym per each service access. The authors use blind signature scheme to register users. In this paper, we use blind signature based pseudonym registering in a novel way so that users have a guaranteed degree of privacy when they register pseudonyms. We achieve this by allowing users to register pseudonyms in a group of $m$.

In (15) the authors explain how a vehicle can change its pseudonyms in consecutive location updates to preserve privacy. The authors study in (16) how to change pseudonyms effectively to safeguard a vehicle’s privacy.

The authors in (17) provide a pseudonym system with access control. They identify unforgeability and unlinkability as the main properties of their System. Also they provide certificate based credential mechanism for access control. In contrast, we provide a pseudonym authentication protocol using Zero Knowledge Proofs. In our system user does not have share secrets with a trusted third party thus allowing complete anonymity. In addition, zero knowledge proofs are less computation intensive making it possible to use such a scheme in mobile applications such as LPA. Table 4 gives a comparison of our system with existing systems.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Trusted Third Party Free</th>
<th>Authorised</th>
<th>m-Unlinkability</th>
<th>Access Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9)</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>(12)</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>(10)</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>TTPFAPS</td>
<td>T</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 4: A comparison of our pseudonym system against the existing systems. TTPFAPS: Trusted Third Party Free Authorised Pseudonym System (our system).

9 Conclusion

We identified the need for an authorised pseudonym registering protocol without a trusted third party that can be used in LPAs to improve user location privacy. We gave a precise description of properties of such a protocol. Then, we presented our two-phase protocol based on blind signature schemes. Also, we showed that if a pseudonym is defined according to an existing zero-knowledge proof authentication protocol, the same pseudonym can be authenticated across different organizations in an LPA.

We discussed the level of privacy of pseudonyms at the time of pseudonym registration. In addition, we discussed how to maintain a consistent degree of privacy of pseudonyms overtime in the presence of adversaries with background knowledge.

References


