# Protecting Location Privacy from Untrusted Wireless Service Providers

Keen Sung
University of Massachusetts Amherst
ksung@cs.umass.edu

Brian Levine
University of Massachusetts Amherst
brian@cs.umass.edu

Mariya Zheleva
University at Albany, SUNY
mzheleva@albany.edu

## ABSTRACT

Access to mobile wireless networks has become critical for day-to-day life. However, it also inherently requires that a user's geographic location is continuously tracked by the service provider. It is challenging to maintain location privacy, especially from the provider itself. To do so, a user can switch through a series of identifiers, and even go offline between each one, though it sacrifices utility. This strategy can make it difficult for an adversary to perform location profiling and trajectory linking attacks that match observed behavior to a known user.

In this paper, we model and quantify the trade-off between utility and location privacy. We quantify the privacy available to a community of users that are provided wireless service by an untrusted provider. We first formalize two important user traits that derive from their geographic behavior: *predictability* and *mixing*, which underpin the attainable privacy and utility against both *profiling* and *trajectory linking* attacks. Second, we study the prevalence of these traits in two real-world datasets with user mobility. Finally, we simulate and evaluate the efficacy of a model protocol, which we call *Zipphone*, in a real-world community of hundreds of users protecting themselves from their ISP. We demonstrate that users can improve their privacy by up to 45% by abstaining minimally (e.g., by sacrificing at most 5% of their uptime). We discuss how a privacy-preserving protocol similar to our model can be deployed in a modern cellular network.

## CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**; *Pseudonymity, anonymity and untraceability*; *Mobile and wireless security*.

## KEYWORDS

Location privacy, trajectory privacy, mobile privacy.

**ACM Reference Format:**
Keen Sung, Brian Levine, and Mariya Zheleva. 2020. Protecting Location Privacy from Untrusted Wireless Service Providers. In *13th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '20), July 8–10, 2020, Linz, Austria.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3395351.3399369


*WiSec '20, July 8–10, 2020, Linz, Austria (Virtual Event)*

ACM ISBN 978-1-4503-8006-5/20/07…$15.00
https://doi.org/10.1145/3395351.3399369

## 1 INTRODUCTION

When mobile users connect to the Internet, they authenticate to a cell tower, allowing service providers such as Verizon and AT&T to store a log of the time, radio tower, and user identity [69]. As providers have advanced towards the current fifth generation of cellular networks, the density of towers has grown, allowing these logs to capture users' location with increasing precision. Many users are persistently connected, apprising providers of their location all day. Connecting to a large private Wi-Fi network provides similar information to its administrators. And some ISPs offer cable, cellular, and Wi-Fi hotspots as a unified package.

While fixed user identifiers are useful in supporting backend services such as postpaid billing, wireless providers' misuse of identifier data is increasingly leading to privacy concerns [14]. Users concerned about their location privacy [10] may use existing tools that allow protection only at the network and application levels. For example, VPNs and Tor [21] mask the IP address of a user from a *remote* server, and hide the remote server location from the service provider. Additionally, access control features allow users to hide or reduce location information sent to location-based services. No such tools exist for *protection of geographic locations from local service providers* — but that does not mean that users are complacent about their ISPs having knowledge of their locations. A recent class action lawsuit demonstrates that mobile users do not want cellular service providers to sell their historic movement records to third parties, such as *location aggregators* [14].

To gain privacy, a user $u$ may attempt to anonymously use a wireless service by obtaining a mobile identity $i_1$ without revealing personal information. The service would provide data connection, while phone calls would be signalled over a VPN using Voice over IP (VoIP). The user may switch to a new pseudonymous identity, $i_2$, before the first is compromised, eventually going through a series of identities over time [12]. However, two primary attacks prevent the user from having location privacy, as illustrated in Figure 1.

(1) In *location profiling*, an attacker identifies one or more of the identities $i_1, i_2, \ldots$ as user $u$ by exploiting the uniqueness of the locations the user is known to regularly visit.
(2) In *trajectory linking*, an attacker infers that activity by $i_1$ is linked to activity by $i_2$ despite the change in identifier. The union of locations can enhance the success of location profiling.

There is a fundamental location privacy cost to connecting to a mobile service. To reduce the success of these attacks without modifying their behaviors, users can *(i)* switch identifies frequently, and *(ii)* remain offline for a period of time between connection sessions, which both reduce user utility. In this paper, we model and quantify this trade-off between utility and location privacy. We

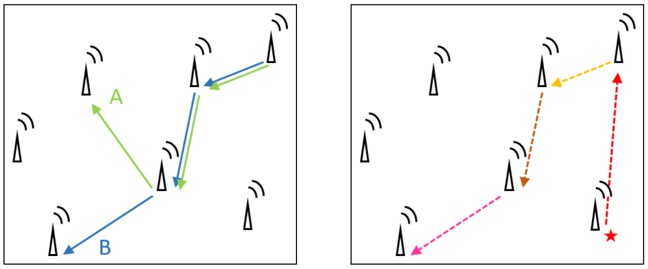


**Figure 1: Left: Diverging paths that are regularly taken by two users. During training, an attacker would encode each labelled transition into a transition matrix for *location profiling*. Right: Separate, unlabelled activity where an unknown user reconnects using a new pseudonym at every tower. If the anonymous user does not successfully mix at these towers (i.e. does not remain offline long enough), the attacker can *link* the trajectories together and match the concatenated trace to User B's profile.**

define utility as the proportion of time the user may stay connected throughout the day while behaving in a privacy preserving manner.

Our work complements existing research in location privacy. Location profiling has been long known to be a problem [19]; attacks typically classify either the set of locations cells visited by an unlabelled user during a time period, or the list of transitions between locations [51]. Trajectory privacy studies, including a body of work in VANETs [35, 48], generally link disconnected traces using Euclidean information. Defenses against these attacks generally utilize a mixing strategy or, more recently, differential privacy. While the latter can separately protect against either location profiling or trajectory linking [24, 65, 66], it requires the cooperation of ISPs. In contrast, our work assumes the ISP is an adversary, and we evaluate robustness against attackers using both profiling and linking.

For our analysis, we model defensive strategies as a protocol we call *Zipphone*, and we define specific ISP-based attacker algorithms as well. We assume a set of users employ Zipphone, using ephemeral identifiers and go offline to prevent trajectory linking. Notably, users do not need to coordinate mixing; naturally occurring mix zones are enough to significantly reduce linking success. Our attacker model looks to historical transition probabilities to model linking, rather than Euclidean distance. Using two real-world datasets [23, 52], we quantify the *path predictability* and *mixing degree* of user activity. With the same data, we demonstrate how a small community can reduce an attacker's re-identification accuracy substantially while sacrificing a limited amount of utility.

**Contributions.** We make the following contributions.

- We formalize two important user traits that derive from their geographic behavior: predictability and mixing, which underpin the attainable privacy and utility against both profiling and linking. To our knowledge, prior work has not analyzed the combination of the profiling and trajectory linking attacks.
- We analyze two real-world datasets [23, 52] and quantify the predictability and mixing behavior of mobile users. While these datasets are relatively small (100–150 active users), they provide a realistic look at the behavioral properties of a set of users.
- We use the same two datasets to quantify attacker accuracy in the re-identification of a community of users running Zipphone. Predictable, mixing users are identifiable only 24% of the time if they renew their identifiers every ten minutes. At the same time, users with permanent identifiers are susceptible to attacks in 69%. We quantify the trade off between the frequency of identifier renewals and user utility. We find that renewals as often as even one hour offer little protection.
- Finally, we discuss how our model Zipphone protocol can be employed in emerging mobile cellular networks without explicit cooperation of the provider.
  We additionally estimate the incurred user-side overhead from Zipphone in terms of time and battery consumption for 3G and 4G networks. Specifically, we measured power consumption during network association and disassociation, and we demonstrate that a user may incur at most 1% battery overhead per day regardless of network technology or desired privacy if Zipphone were used. We detail the challenges that such deployment would face.

In what follows, we first summarize related work in Section 2. We then present our attacker model and corresponding attacker-defender dynamics in Section 3. We evaluate Zipphone's privacy preserving performance in Section 4. We then discuss avenues for employing Zipphone in emerging mobile cellular networks and quantify the user overhead in Section 5. We discuss limitations and ethical implications in Section 6 and conclude in Section 7.

## 2 RELATED WORK

Our study is related to a broader category of prior work on location privacy. Most prior work assumes the service provider is trusted and in fact responsible for user privacy. Prior approaches have a variety of goals, including: *(i)* properly anonymizing mobility datasets before public release; *(ii)* adding privacy for users of locations based services; and *(iii)* increasing location privacy for mobile device users from third-party attackers but *not* the service provider itself. In contrast to these works, our goal is to provide mobile users location privacy from the wireless provider itself. This presents a unique challenge: the user is responsible for her own privacy, and the only control she has over this is whether to remain connected to the service at any moment in time.

In our preliminary work [59], we examined the efficacy of ephemeral IMSIs. This paper significantly expands upon that work by: including trajectory linking as an attack; including user utility, off time, and cool down in the renewal algorithm, which is more practical and also thwarts trajectory linking; quantifying predictability and mixing of users; using a new data set; and quantifying overhead.

**Location privacy with provider cooperation.** Many studies focus on enlisting a trusted carrier to protect against a third party attacker [29, 32, 33, 46]. Reed et al. [56] propose privacy from the carrier using onion routing, but does not consider the direct connection that must be made to a tower. Federrath et al. [28] propose a similar scheme that prevents linkability of calls between two parties but omit critical details regarding authentication to the carrier. Fatemi et al. [27] propose an anonymous scheme for UMTS using identity-based encryption, but unlike our approach, their scheme involves the carrier in the cryptographic exchange; they enumerate the vulnerabilities of similar works [41, 54, 67, 70]. Kesdogan et al. [42] proposes using a trusted third party to create pseudonyms for GSM users, but also routes all calls through that provider, which allows it to characterize the calling pattern and infer the caller.

**User-driven trajectory privacy.** Mix zones [12, 30] can be employed by a user against a provider attacker when the network service provider is non-cooperative. While the concept of mix zones is fairly old, it remains the only available option for users who want to hide their own location privacy from a service provider. Work in VANETs also uses mix zones to protect vehicle trajectory [25, 35, 48]. Given that their focus is on trajectory, these studies do not consider location profiling. Other work involves the introduction of false information [44, 58]. Few studies use this concept to protect the user from an omnipresent network attacker. Chan [15] focuses on call metadata privacy, rather than location privacy.

**User-driven profiling privacy.** Work that increases the privacy of location-based services (LBS) [38, 53, 62, 63] generally add noise to location queries. These works are not viable or applicable against an untrusted service provider: a user cannot manipulate which tower they connect to, and the provider knows the physical locations of the towers serving users.

**Dataset protection.** Works that aim to prevent leaks in personally identifiable information in shared or publicly released datasets [68] primarily rely on obfuscation. They also strive to prevent trajectory recovery [34, 60]. Older work on deanonymization of mobile users' traces assumes the user's pseudonym is unchanged throughout the trace. But a small amount of external information, such as the person's home or work address [40], can deanonymize an obfuscated trace [11, 12, 31, 45, 49, 51] given a consistent identifier. Zang and Bolot [69] show that suitably anonymizing a trace of 25 million cellular users across 50 states (30 billion records total) requires only that users have the same pseudonym for no longer than a day. A day's duration is unsuitable for Zang and Bolot's goal of supporting researchers that wish to characterize the behaviour of users over time (while maintaining their privacy). On the other hand, the result is promising for users seeking privacy, who might be able to change their pseudonyms more frequently than once per day.

**Differential privacy.** More recently, differential privacy approaches [22, 50] are used to add noise to datasets while preserving its aggregate characteristics. Palamidessi et al. [9] introduce geo-indistinguishability, and ElSalmouny & Gambs [24] further discuss $(D, \epsilon)$-location privacy. Xiong et al. [65, 66] formalize situations where location queries can be temporally correlated and linked. These methods all assume the service provider is trusted and are, thus, not applicable to our problem setting.

**Outside threats.** Several studies protect against third party attackers and vulnerabilities in 3GPP implementations [36, 39]. Khan et al. [43] provide a cryptographic mechanism to generate LTE pseudonyms and prevent third-party attackers or IMSI catchers from linking users.

In comparison to related work, we differ in that we do not trust the wireless service to ensure the user's privacy, and we assume in our analysis that the adversary is attempting to link together traces. Our evaluations are based on traces of real users [23, 52], which allows us to quantify the periodicity of identifier changes in the context of modern cellular infrastructure.

# 3 ATTACKER AND DEFENDER ALGORITHMS

Our primary goal is to quantify the privacy-utility trade-offs present in systems that provide geographic anonymity from mobile ISPs. To do so, first we instantiate a specific protocol for users and provide well-defined attacker algorithms. The protocol, Zipphone, is based on mechanisms available to the user only; i.e., the ISP is not cooperative, an assumption not shared by many location privacy systems. In short, users can control only their active identity (i.e. pseudonym) and whether or not they are connected; providers attempt to link the activities of identities to existing user profiles.

## 3.1 Problem Statement

Zipphone users seek to use the network, but not have their real identities associated with mobility recorded in traces. Upon joining the network, the user $u$ is assigned a pseudonym $i$. The pseudonym lets the user maintain a connection session for some period of time. The user attaches to a sequence of towers as it moves according to signal strength and the corresponding handoff procedures. By registering as identity $i$ and then moving, the user provides to the ISP a *trace*: $(i, (s_1, s_2, \ldots))$, where each value of $s$ indicates a specific wireless transceiver and a timestamp. The provider knows the locations of the transceivers and can, thus, trace a user's mobility. It is not the goal of the user to hide that they are using Zipphone.

The goal of the attacker is to infer and label their identities from the traces. The attacker is a wireless provider such as a Mobile Network Operator (MNO) that already has a history of traces for each Zipphone user. The attacker then tries to determine which user from a set $u_1, u_2, u_3, \ldots$ is the one that created the trace $(i, (s_1, s_2, \ldots))$ based on a classifier trained from the known history, where $i$ represents an IMSI. Since longer traces are easier to classify, users must regularly renew their identity; programmable solutions such as an eSIM could facilitate this process. Section 5 provides a discussion on how this may be implemented in a modern cellular infrastructure.

In Section 4.3, we demonstrate that longer traces are easier to identify and link with other traces; users should regularly renew their identifier in order to keep these traces short. We assume the user does not perturb their own movement patterns. Therefore important parameters are *(i)* the identity renewal frequency, and *(ii)* the user's offline duration. When the renewal frequency is higher, privacy also increases; but each identity renewal incurs an offline period and increases power usage. Longer offline durations improve privacy but reduce utility. We assume all such parameters are public and known to the attacker.

## 3.2 Attacker Model

The attacker's goal is to determine the identify $u$ of a trace $(i, (s_1, s_2, \ldots))$ of consecutive tower connections. We assume the attacker *(i)* has all traces of all Zipphone devices, and *(ii)* has labelled/identified traces of historic movement for all Zipphone users, for training a classifier; in other words, the attacker is a service provider such as a mobile network operator. The attacker performs *trajectory linking*, which patches together separate traces if a classifier predicts they are from the same user.

**Algorithm 1** User identifier renewal strategy (*Zipphone*)

```
utility ← Minimum utility between 0.0 and 1.0
max_off_time ← Maximum time offline during renewal
while device is online do
    WAIT(until device moves outside range of tower)
    DISCONNECT
    off_time ← UNIFORM(0,max_off_time)
    WAIT(off_time)
    CONNECT                        ▷ connect with new identifier
    cooldown_time ← utility × off_time
    WAIT(cooldown_time)
```

We assume that all Zipphone users are of equal interest to the attacker, and that it uses only normal cellular infrastructure to attack. For example, we assume that the attacker does not install cameras on towers to identify users via facial recognition, nor would they follow a particular user by car. It does not make sense for the attacker to set up an IMSI catcher[17] since they already own the entire real infrastructure. We assume that location accuracy is on the level of cell tower; while features such as RSSI or TDOA could locate wireless devices with more precision, devices could in turn artificially slightly reduce performance as a defense, effects of which are outside the scope of this paper.

We assume that the attacker gains no other information from the users; in mobile phones, information such as IMEI, device model, or OS signatures, are easily turned off via OS settings. In practice, such features would assist the attacker (see [16]), but are not the focus of this paper as they are more easily obfuscated or falsified than real geographical movement. For example, IMEIs, which are akin to a MAC address, can be modified by the user since she controls the handset hardware (e.g., SilentCircle's blackphone [6]). Since users are likely identifiable by the unique set of outgoing calls they make, they should make calls via VoIP through an anonymizing proxy or circuit instead of using a conventional phone connection. Encryption of the VoIP stream can thwart carrier eavesdropping. Stronger protection is available by using VoIP over Tor [8].

A user tries to maximize their utility (i.e. uptime) while remaining private; thus, their reidentifiability depends on their predictability and mixing behaviour. A user who visits vastly different location than her peers could not mix easily; her activity could be easily linked and profiled. A user who is not predictable could not be easily identified regardless of mixing behaviour.

## 3.3 Attacker-defender dynamics

*3.3.1 User strategy.* Algorithm 1 defines the Zipphone user algorithm. As described in the previous section, Zipphone users renew their identifiers only when three conditions are met: *(i)* they are in the process of switching towers, and *(ii)* the renewal *cool down period* (in seconds) has expired; *(iii)* they are not actively using the phone. To renew, users first detach, then stay offline, and then reattach with a new profile. The offline time is selected uniformly at random from a *maximum offline period.* It must be random, otherwise linking traces would be trivial. The cool down period ensures that the loss of utility remains at a minimum for the user. This aggressive renewal strategy is frequent enough to allow the natural formation of mix zones, and does not require users to coordinate times or places to mix.

**Algorithm 2** Location profiling algorithm

**function** PROFILE_USER($u$) ▷ $u$ is the user index

  $T^u_{0,q} \leftarrow \frac{Count(q)}{\sum_{q' \in \mathbb{C}} Count(q')}$ ▷ The prior for user's initial location

  **for all** $p \rightarrow q \in$ TRANSITIONS($u$) **do** ▷ $p \rightarrow q$ denotes a transition

    $T^u_{p,q} \leftarrow \frac{Count(p \rightarrow q)}{\sum_{q' \in \mathbb{C}} Count(p \rightarrow q')}$ ▷ This transition matrix may be sparse

  **return** $T^u$

**function** CLASSIFY_USER($s$) ▷ $s = (s_0, s_1 \dots), s \in \mathbb{C}$ is a sequence of tower IDs

  **return** $\arg\max_u T^u_{0,s_0} \prod_{i=0}^{n-2} T^u_{s_i,s_{i+1}}$

*3.3.2 Attacker strategies.* The attacker's goal is to take a timestamped sequence of visited towers and infer the user, given a training set. We first describe a *location profiling classifier* that could be employed by the attacker. We then define a *trajectory linking classifier* to aid the attacker in trajectory linking.

*Location profiling algorithm.* Our classifier (Algorithm 2) is a Markov model that chooses the most likely user for a sequence of tower attachments; the classifier is adapted from Mulder et al. [51]. This algorithm is well suited to identify users of a device that has its location constantly logged throughout the day. With this classifier, the attacker labels a sequence of locations with the most likely user, based on all possible users' transition histories. In our model, vector $s$ is a sequence of locations in the location set $\mathbb{C}$: $s = (s_0, s_1, s_2 \dots), s \in \mathbb{C}$. In the steps below, the attacker identifies the most probable user given each candidate user's history, $\hat{u} = \arg\max_u p(u|s)$.

We determine the most likely user, given a sequence of locations.

$$\Pr(u|s) = \Pr(u|s_0, s_1, s_2, \dots)$$

We apply Bayes' rule, and consider the likelihood of a sequence given a user.

$$\Pr(u|s) = \frac{\Pr(s_0, s_1, s_2, \dots |u) \Pr(u)}{\Pr(s_0, s_1, s_2, \dots)}$$

We assume that each user is equally likely.

$$\begin{aligned} \Pr(u|s) &\propto \Pr(s_0, s_1, s_2, \dots |u) \\ &= \Pr(s_0|u) \cdot \Pr(s_1|u, s_0) \cdot \Pr(s_2|u, s_0, s_1) \cdot \\ &\quad \Pr(s_3|u, s_0, s_1, s_2) \dots \end{aligned}$$

Each transition is independent per the Markov assumption.

$$= \Pr(s_0|u) \prod_{i=0}^{n} \Pr(s_{i+1}|s_i, u)$$

We determine the most likely user $\hat{u}$.

$$\hat{u} = \arg\max_u \Pr(s_0|u) \prod_{i=0}^{n} \Pr(s_{i+1}|s_i, u)$$

The attacker computes a transition matrix $T$ for each user in the training data by counting the historical transitions. The probability of the first location in the sequence $\Pr(s_0|u)$ is computed from the overall number of a user's occurrence at a location. The attacker does not consider the probability of a trace ending at a certain location, since a sequence can end for arbitrary reasons.

**Algorithm 3** Linking algorithm

max_t ← Maximum time offline during renewal
**function** TRAIN_LINK_TRANSITIONS
  **for all** $p \xrightarrow{max\_t} q$ **do** ▷ all locations $q$ seen within max_t of $p$
    $T^{l}_{p,q} \leftarrow \frac{Count(p \xrightarrow{max\_t} q)}{\sum_{q' \in C} Count(p \xrightarrow{max\_t} q')}$ ▷ transition matrix used for linking
  **return** $T^{l}$
**function** CLASSIFY_USER_WITH_TRAJECTORY($s$)
  **while** link_count<max_links **do**
    candidates ←FIND_CANDIDATES($s$)▷ traces ≤ max_off_time after $s$ ends
    **if** EMPTY(candidates) **then**
      break
    $\hat{s}' \leftarrow \arg\max_{s'} T^{l}_{s_n, s'_0}$ ▷ $\forall s' \in$ candidates
    $s \leftarrow$ CONCATENATE($s$,$s'$)
  **return** CLASSIFY_USER($s$)

The success of such an attack depends on two factors: the number of users in the anonymous community, and the similarity of the user's location transitions to the other users. If there is one registered cell phone user on the network, then linking the user to location is trivial; however, if there are many users who behave similarly, it would be difficult for the attacker to tell the user apart.

We also designed and tested a classifier that exploited diurnal features of user mobility, however, it did not perform significantly better than the above outlined algorithm. Thus, in the remainder of the paper, our attacker model does not employ diurnal features.

*Trajectory linking algorithm.* In Algorithm 3, we extend Algorithm 2 to model the attacker's ability to do trajectory linking. The attacker uses the transitions of all users and builds a semi-Markov *linking transition matrix*. This matrix is similar to the one described in Algorithm 3, except that it is built by considering all subsequent locations within a given offline time, rather than only the next immediate location. This strategy ensures that unreasonable transitions do not confuse the classifier, and any unseen transitions occurring within that time frame are accounted for.

Our trajectory linking first searches for candidate traces that *start* within the maximum offline time. If a number of traces start within the offline time, the targets have a chance to mix, and the attacker must infer which trace comes next by using the semi-Markov transition matrix. This process is repeated until the trace is of sufficient length for classification, or there are no more candidates.

# 4 EVALUATION

In this section, we determine the parameters in our model and evaluate the algorithms using two real-world datasets that contain geotagged user data coupled with tower attachment logs: PhoneLab [52] and RealityMining [23]. First, we characterize the amount of *predictability* and *mixing* behaviour exhibited by users in these datasets. We demonstrate that both characteristics are related to the success of the attacker's accuracy. Next, we simulate a deployment of Zipphone amongst a community of users, and determine their reidentifiability with respect to sacrificed utility.

## 4.1 Datasets

Both datasets were collected by university affiliates who carried phones instrumented to log network attachment and user activity.

| Type | Trait Predictable | Trait Mixing | Privacy hypothesis | PhoneLab | Reality Mining |
|---|---|---|---|---|---|
| P/M | Yes | Yes | Moderate-Low | 18% | 18% |
| P/nM | Yes | No | Low | 26% | 30% |
| nP/M | No | Yes | High | 30% | 24% |
| nP/nM | No | No | Moderate | 26% | 29% |

**Table 1: User typology and their proportions in our target datasets, with a hypothesis about the amount of privacy a user could attain from Zipphone.**

(1) PhoneLab [52] is an Android testbed comprising 593 phones distributed to students at the University of Buffalo campus. As a part of this testbed, users contributed geotagged traces of their cellular network associations. We use 24 months from January 2015 to January 2017 of cellular network association traces from PhoneLab to assess the privacy preservation potential of Zipphone.
(2) RealityMining [23] is a dataset released by MIT that tracks a group of 100 mobile phone users across various contexts. Similar to PhoneLab, RealityMining contains geotagged network association information. For our analysis, we leverage 12 months of RealityMining data from July 2004 to July 2005.

We are unaware of other public datasets that could be used to analyze our algorithms. Larger datasets [13, 61] do not contain sufficient information about users' association with towers and, thus, do not cater to our analyses. (We filed IRB protocol 2017-3900 as part of this project, and it was approved as exempt.)

## 4.2 Behaviour that affects attacker accuracy

We begin by characterizing user behaviour. Intuitively, there are two behavioural traits that affect mobile users' privacy: *(i) Predictability*, or to what extent users travel over fixed routes; and *(ii) Mixing* behaviour, or how likely are users to visit popular locations that see a large volume of other Zipphone users. To highlight the effect of user behaviour on privacy, we categorized PhoneLab and RealityMining users *post hoc* into four groups:

- predictable (P) or unpredictable (nP); and
- mixing (M) or not mixing (nM).

The four resulting user types are described in Table 1, where we also set forth a hypothesis of how user behaviour would affect privacy. We verify and confirm these hypothesis in our evaluation (Section 4).

**Predictability** We calculate the user predictability in terms of the similarity of the set of cellphone towers they visited during the testing and training period. For each user, let $C_{pre}$ be the set of towers visited during the training phase and $C_{post}$ be the set of towers visited in the testing phase. We express the predictability in terms of a user's Jaccard similarity score between $C_{pre}$ and $C_{post}$, defined as

$$J_C = \frac{C_{pre} \cap C_{post}}{C_{pre} \cup C_{post}}, \tag{1}$$

where $0 \le J_C \le 1$. $J_C = 0$ when the sets of visited towers in testing and training are completely disjoint, while $J_C = 1$ means that the sets of visited towers in testing and training are the same. Intuitively, a higher $J_C$ means a more predictable trajectory.

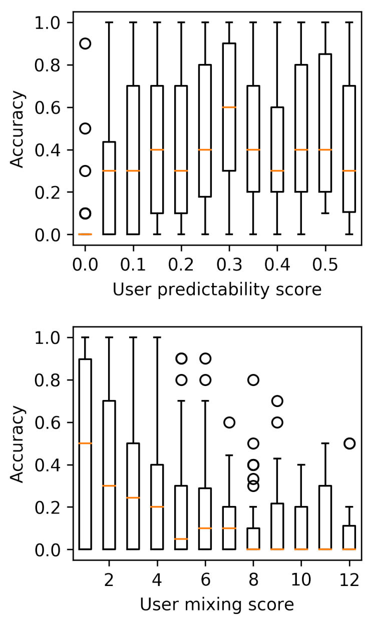


**Figure 2: Top: User predictability versus attacker accuracy, showing that attacker accuracy is near zero with low predictability. Bottom: User mixing versus attacker accuracy, showing that median accuracy falls to zero as user mixing increases. The plots were computed from the** PhoneLab **dataset. The presented results are for a maximum offline time period of 30 seconds and a set utility of 95%. Utility and accuracy metrics are discussed in detail in Section 4.3.**

Figure 2 (top) presents the attacker's accuracy (i.e., the probability that a user would be identified) as a function of the users' Jaccard score in the PhoneLab dataset. We note that the trends and respective thresholds are similar for the RealityMining dataset and omit these results due to space limitations. For this setup, 91% of users fall within the 0.0–0.4 Jaccard score range. Users with a Jaccard score below 0.1 are less identifiable. Using this analysis of our test dataset, we set the Jaccard score to 0.1 as a cut off to differentiate between predictable users (such with $J_C > 0.1$) and unpredictable users (such with $J_C \leq 0.1$).

**Mixing behaviour** We establish a mixing score $\mathcal{M}_C$ as a metric that evaluates a user's likelihood to mix with other Zipphone users. Intuitively, the higher the mixing score, the more efficient ID switching will be and the harder it will be for the adversary to evade a user's privacy. We calculate $\mathcal{M}_C$ for each individual user. Let $t_i^k$ be the duration of time a user $i \in (1, N)$ spends at tower $k \in (1, K)$. During the period $t_i^k$, other users $j \in (1, N'), j \neq i, N' \subset N$, may arrive and depart from tower $k$. Let $\tau_{ij}^k$ be the time of user $j$'s arrival or departure. Intuitively, $t_i^k$ and $\tau_{ij}^k$ define the temporal granularity of tower mobility and Zipphone user encounter events, respectively, from the perspective of a single user $i$. Let $C(\tau_{ij}^k)$ be the number of users in user $i$'s vicinity at time $\tau_{ij}^k$. We define the mixing score as:

$$\mathcal{M}_C = \sum_{k=1}^{K} \sum_{j=1}^{N'} \frac{C(\tau_{ij}^k)}{\tau_{ij}^k - \tau_{i(j-1)}^k} \tag{2}$$

Figure 2 (bottom) presents the attacker's accuracy as a function of the users' mixing score in the PhoneLab dataset. The trends and respective thresholds are similar for the RealityMining dataset. The attacker's accuracy deteriorates as the users' mixing score increases. Based on this analysis, we set a mixing score of 4 as the cutoff to determine whether a user is mixing or not mixing, i.e. users with $\mathcal{M}_C \leq 4$ are not mixing and these with $\mathcal{M}_C > 4$ are mixing.

**User typology in our datasets.** As detailed earlier, we differentiate between four types of users based on their predictability and mixing behaviour. Using the presented analysis in Figure 2, we set a Jaccard similarity threshold of 0.1 and mixing score threshold of 4. We note that these thresholds are solely used to establish the user topology in the following evaluation and do not play a role in the profile classification carried out by the attacker. Figure 1 presents the amount of users that fall in each user type category. We see a relatively even user representation across all categories. We use these user types and the corresponding user populations in all results presented in the evaluation of Zipphone (Section 4.3).

## 4.3 Results

To determine the affect of Zipphone on the utility and privacy of users, we simulated the protocol using the PhoneLab and RealityMining datasets. In these simulations, the attacker uses the inference algorithms outlined in Section 3.3.2 to develop a location profile for each user. We split the data up into several sets of three months; training was done on the first two months, and testing was done on the third month.

*4.3.1 Utility-privacy trade-off.* We evaluated the utility-privacy tension with regard to the four user types. We quantify privacy gained in terms of reduced attacker accuracy. We measured loss of utility in terms of time spent offline during the testing period. Figure 3 displays the privacy gained by each user group during the one-month testing periods.

Users gained significant privacy from sacrificing 5% utility, on average remaining online for 9.5 minutes, and going offline for 30 seconds. In particular, Type P/M (predictable but mixing users) gained 45% in the PhoneLab dataset, and 49% in the RealityMining dataset. Interestingly, Types nP/M and P/nM also show a similar trend: Type nP/M benefits from having the divided traces be less predictable, and for Type P/nM any small amount of predictability is reduced to none. Type nP/nM does not mix, and enjoys uniformly high privacy because they are unpredictable. Users were more private in general in the PhoneLab, since it represented a larger community of users, making mixing easier for the user, and user inference more difficult for the attacker.

*4.3.2 Trace length and location profiling.* The main driver of attacker accuracy is trace length. Longer traces contain more information, allowing more accurate reidentification. In these experiments, the attacker tries to identify an independent trace of varying length, increasing from one second to four weeks. Figure 4 shows the result.

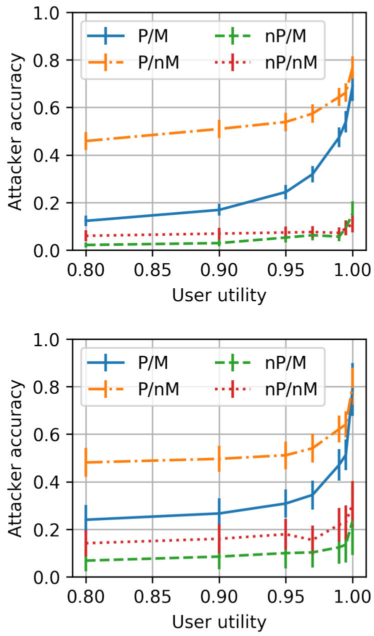


**Figure 3: Top:** PhoneLab**. Bottom:** RealityMining**. In both datasets, predictable but mixing users (Type P/M) gain the most from using Zipphone. Ten test traces were evaluated per user, and accuracy is represented as a mean of the proportion of successful reidentifications per user. Error bars represent a 95% confidence interval.**

The longer the trace, the more identifiable (and thus less private) an individual is. Users who exhibit more predictable behaviour have less privacy; generally, they benefit from traces that are at most one hour long. In other words, predictable users should change their identifier at least once per hour while in motion. Those who travel to unique locations as compared to others benefit significantly less from the shorter trace. This result highlights the benefit of Zipphone. Users should change their identifiers more than once per hour, and this system obviates the need to physically change an identifier, and handles this process automatically. While a temporary SIM device may grant some measure of privacy, a system that renews a user's identifier a lot more quickly can be a lot more effective.

*4.3.3 Compromises in utility.* While users may renew identifiers by prearranging mixing strategies with other users, such coordination is impractical. A frequent enough renewal strategy and long enough renewal times allow mix-zones to naturally form, which enables users to mix without any coordination. In Figure 5 (top), we examine the amount of time a user should remain offline. The frequency of renewal is informed by the utility desired, which we set at 95%.

For users to gain privacy during identifier renewal, they must remain offline long enough to mix with other users. Additionally, users must not have a fixed offline time, since this would be susceptible to a timing attack. Users must choose an offline time that is not so long to be disruptive, but not so short as to offer little privacy. The Zipphone population's policy should fix a chosen utility, and employ a cool down time between each user's identifier renewal

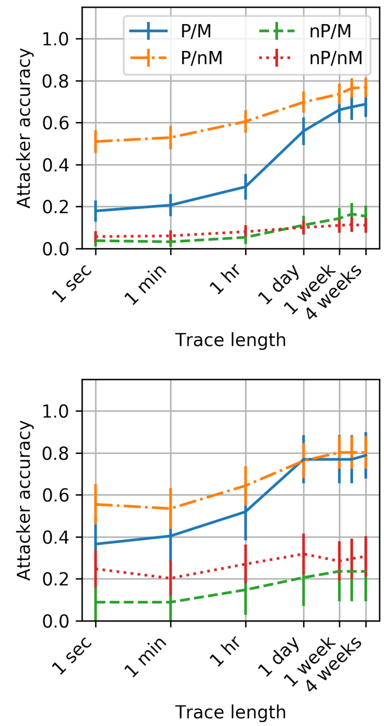


**Figure 4: Top:** PhoneLab**. Bottom:** RealityMining**. Users lose a significant amount of privacy when traces are on the order of one day long. The accuracy at one month is equivalent to the accuracy in Figure 3 at 100% utility.**

based on that desired utility. For example, if users' offline-times are 30 seconds, and are aiming to maintain 95% utility, they will keep every identity for at least 30 seconds $\div$ $(1 - 0.95)$ =10 minutes.

Because going offline for 30 seconds can be fairly disruptive, we analyzed scenarios where reconnections are disallowed if *(i)* the user is in the middle of a phone call, or *(ii)* the device screen is active. This data was available in only the PhoneLab dataset. Since phone calls were intermittent, active calls could be kept online without sacrificing privacy. However, within the offline periods, users would on average miss 4 calls out of 24 per month while maintaining 95% utility. Looking at screen usage, we show in Figure 5 (bottom) that users could preserve active usage of phone undisturbed, but in doing so would sacrifice additional privacy by a small amount (i.e. about 2% across all utility levels).

## 5 INTEGRATING ZIPPHONE WITH EMERGING MOBILE NETWORKS

In this section we discuss how Zipphone could be integrated in emerging mobile cellular networks towards improved user privacy. We first present necessary background on user authentication in emerging cellular networks. We then detail how Zipphone can utilize these networks for privacy-preserving services without requiring network modifications. Finally, we present empirical results for user-side energy overhead.

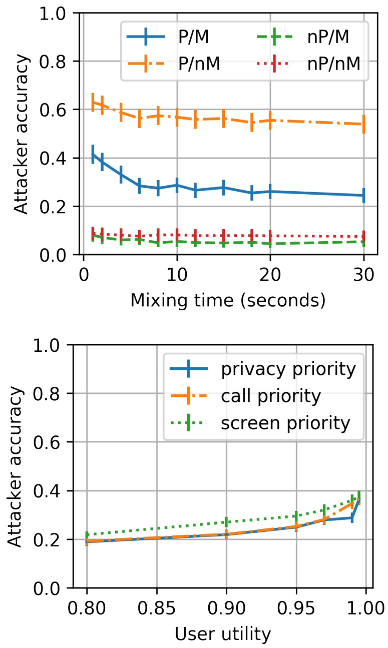


**Figure 5: Top: the effect of mixing-time on privacy while maintaining a 95% utility for the** PhoneLab **dataset. Bottom: privacy/utility of all users depending on whether their priority is privacy, phone calls, or screen use. Calls can be prioritized without sacrificing privacy. However, remaining online while the screen is on significantly reduces privacy.**

## 5.1 Background

Traditionally, hardware SIM cards installed in mobile devices provide the basis for user provisioning in Mobile Network Operators (MNO). Each SIM has a unique International Mobile Subscriber Identity (IMSI), which is pre-programmed by the vendor prior to being sold to a mobile subscriber. At the point of sale, when a user purchases the SIM card, an entry is created in the MNO's Home Location Registry (HLR) connecting the IMSI with a Mobile Station International Subscriber Directory Number (MSISDN; i.e., a phone number). In addition, the IMSI is paired with a Ki value at the MNO's Authentication Center (AuC) and used for user equipment (UE) authentication. We note that this procedure requires a mapping between IMSIs and devices, not IMSIs and users and, thus, it supports both pre-paid and post-paid services.

This hardware SIM approach to user provisioning is plagued with high overhead, wasted IMSI allocations, and manual processes. To address these limitations, the eSIMs standard [1] has been developed, which allows programmatic and on-the-fly provisioning of a user's identity on a network. With eSIMs, mobile users can maintain multiple simultaneous mobile network identities and use heterogeneous services from one or multiple MNOs. Three out of the four major carriers in the US currently support eSIM, with one major carrier supporting eSIM in 42 other countries worldwide [2].

eSIMs introduce new components to user management that are useful for Zipphone. Similar to traditional SIMs, the eSIM functional *profile* [5] carries phone identification information and is jointly maintained in the MNO's HLR and the AuC. The Subscription Manager Data Preparation (SM-DP+), is responsible for provisioning a user's profile onto the eSIM. Thus, the SM-DP+ is the first point of contact between an aspiring subscriber and the MNO, from which the subscriber obtains their functional profile. There is no upper limit on the amount of *profiles* an eSIM can maintain; it depends on *(i)* the size of a single *profile*, *(ii)* the eSIM integrated memory and, *(iii)* the operator's preferences. As an example, T-Mobile currently supports up to 10 concurrent eSIM *Profiles* [4]. Responding to the eSIM revolution, both major mobile operating systems, Android[1] and iOS[2], integrate APIs that allow the development of carrier apps for programmatic user subscription management.

## 5.2 Proposed Zipphone Architecture

*5.2.1 Overview.* Zipphone can be realized as a smartphone application. Upon installation and then periodically, the Zipphone app will anonymously acquire multiple functional profiles and associated *service quants* from the MNO's SM-DP+. We define a service quant as a set of mobile services (i.e. data, SMS and voice calls) that the subscriber will use while active with the particular profile and note that these quants can be obtained in the form of an anonymous prepaid service [3, 7]. Zipphone then programmatically swaps these profiles as discussed in Section 3.3 and uses the corresponding service quant for the duration in which a profile is active. This functionality can be achieved without explicit cooperation from the network provider or any modifications in the network as long as the provider is eSIM-capable and offers anonymous prepaid plans.

*5.2.2 Purchasing Credentials.* Zipphone requires that users anonymously purchase *profiles* without linking to a consistent financial or network identifier. This purchase would be a significant challenge to deploying Zipphone as it must also not be used to profile the user. Here we offer a sketch of how it could be done.

Purchase can be made through traditional means, such as a credit card, to a third-party Mobile Virtual Network Operator. The MVNO can issue *Privacy Pass* tokens [18]. These cryptographic tokens cannot be forged by the client and cannot be spent twice, and yet they are unlinkable to the purchase. The advantage of this approach is that the MVNO has the option of keeping track of who its customers are while not knowing where they are geographically. In contrast, the MNO would know clients have paid the MVNO, but not know who they are. The use of Privacy Pass makes it hard for the MVNO and MNO to share knowledge. If the tokens are sold by an MVNO, then signaling is required to the MNO to cancel the IMSI a period of time after they are first used (e.g., 15–30 minutes). To purchase the Privacy Pass tokens anonymously from an MNO or MVNO is more challenging. Cash can be used in person. To pay online, anonymous currencies such as *Zcash* [37, 57] can be used. Protocols such as Dandelion++ [26] allow transactions to be issued to Zcash with network anonymity. It's also possible that an MNO could accept Zcash payments, issue Privacy Pass tokens, and accept the anonymized tokens later. It's worth noting that Zipphone offers benefits even when anonymous purchases cannot be made. For example, law enforcement, activists, or journalists and other large

[1] https://source.android.com/devices/tech/connect/esim-overview
[2] https://developer.apple.com/documentation/coretelephony/ctcellularplanprovisioning

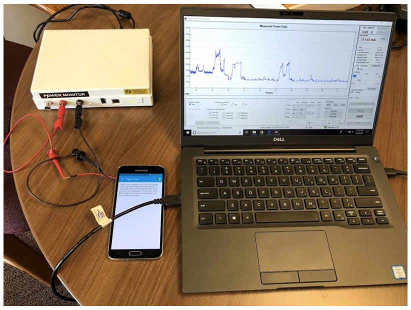

**Figure 6: Experimental setup for power measurements on 3G and 4G networks.**

organizations for whom security is crucial can create their own trusted MVNO and maintain location privacy from an untrusted MNO.

*5.2.3 Communication without Leaking Identity or Location.* For an additional layer of privacy, Zipphone users should ignore the MSISDN (phone numbers) provided by a profile. In other words, users should not use MSISDN-based services such as text and voice calls and instead should rely on IP based services over the data plan. If a Zipphone user initiated or received overt LTE or unencrypted VoIP calls, they risk being identified via a profile of call records held by the carrier. Incoming calls are spam or attacks and should be ignored. Note that the E911 service, which is tied to a handset and not a user or SIM, would be still available if needed.

Some protection would be gained from using an encrypted VoIP service, since it would not reveal to the carrier the identity of the user's contact, whom she calls, or from whom she receives calls. However, if the IP address of the VoIP service is unique, then connecting to it would help the MNO link a collection of profiles together. An anonymous VoIP service, such as Torfone can be used; note that anonymous VoIP has a performance penalty [47].

In general, an anonymous communication system, i.e., Tor, must be used for all Zipphone communication (voice or data). However, there is one change required. Tor chooses a consistent, single *guard* relay to start all three-relay circuits through the Tor network. If Zipphone users send all traffic to a single guard relay, it would be a consistent identifier despite changing IMSIs. Instead of a guard at the start of the circuit, Zipphone users should use a consistent relay as the exit. This switching of roles allows Zipphone users to receive all protections against the Predecessor Attack [64] that Tor normally provides via guard nodes at the entry.

## 5.3 Zipphone Overhead

Zipphone triggers periodic disassociation/association from the mobile carrier, which together incur additional battery draw and connect/disconnect delays on the mobile device. Thus, in this section, we quantify the overhead in terms of battery drain and latency, incurred by Zipphone on 3G and 4G networks.

**Experimental setup.** In order to evaluate the power consumption of mobile network association/disassociation, we used a Samsung Galaxy S5 Duos phone with a bypassed battery and a Google Fi SIM card, and a Monsoon Power Meter. We connected the phone to the main channel of the power meter, as illustrated in Figure 6,

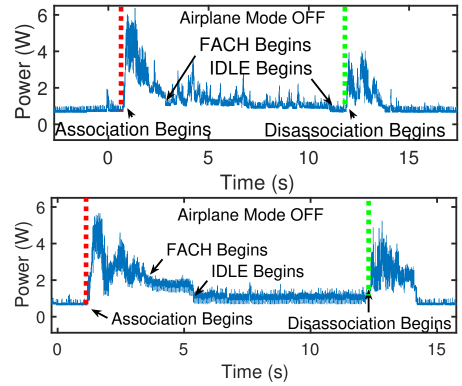


**Figure 7: Power trace for 3G (top) and 4G (bottom) association and disassociation.**

which allowed us to both power up the phone and measure its energy consumption. In order to measure the power draw at 3G and 4G networks, we forced the phone to the respective technology and sampled the power draw at a granularty of 200$\mu$s. We used the phone's Settings screen to toggle between Airplane Mode OFF and Airplane Mode ON every 10 seconds for 4G and every 20 seconds for 3G. We disabled all background services on the phone. This ensured that we are only measuring the power draw from association/disassociation, plus a baseline of about 700mW used by the display for the Airplane Settings page. For each of 3G and 4G we completed 10 full association/disassociation cycles. The average experienced time and power to connect inform our simulation.

Figure 7 presents a zoomed version of a single associate/disassociate cycle for 3G (top) and 4G (bottom)[3]. There are several important points to note on each trace. First, the red vertical line indicates the phone's transition from Airplane Mode ON to OFF state, which immediately triggers a network association. After the association procedure completes, the phone enters FACH (Forward Access CHannel) state in anticipation for the user to begin accessing the Internet. Since this does not happen in our controlled activity, the phone futher transitions into IDLE state. At the instant designated with a green vertical line, we toggle Airplane Mode ON, which immediately triggers a disassociation procedure.

A Zipphone user would experience two types of overhead: *(i)* offline time, and *(ii)* power draw. We measure the offline time as the time between the beginning of network association and the beginning of the FACH state. We measure the power overhead as the sum of power to associate and power to disassociate, whereby the power to associate is incurred from the begining of the network association to the beginning of the FACH state, while the power to disassociate is measured from the beginning till the end of the disassociation procedure.

Figure 2 presents the average incurred overhead for our measurement campaign. We see that the offline time incurred by 3G is nearly double that of 4G. The power consumption, on another hand, is comparable across the two technologies. We use these results to quantify the battery usage per day for users in our datasets. To this end, we convert the measured power consumption for a

[3]Note that the timescale (i.e. the $x$-axis range) for 3G is longer than that for 4G. On 3G, the phone takes significantly longer to transition to IDLE mode compared to 4G.

| | | mean | (std dev) |
|---|---|---|---|
| 3G | Power to connect (mW) | 2,098 | (435) |
| | Power to disconnect (mW) | 1,282 | (157) |
| | Time to connect (s) | 5.0 | (0.8) |
| | Time to disconnect (s) | 4.0 | (1.0) |
| 4G | Power to connect (mW) | 2,006 | (171) |
| | Power to disconnect (mW) | 1,120 | (295) |
| | Time to connect (s) | 2.6 | (0.2) |
| | Time to disconnect (s) | 3.0 | (1.2) |

**Table 2: Time and power overhead incurred by a single association/disassociation procedure on 3G and 4G in our experiments. Results are averaged over 10 runs.**

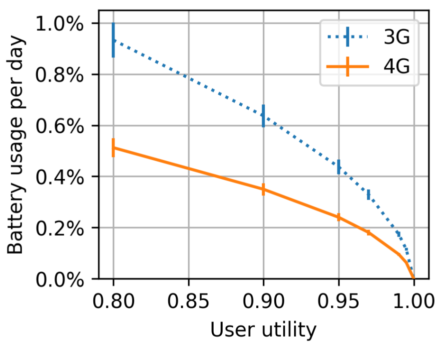


**Figure 8: Battery usage does not exceed 1% per day, regardless of desired privacy or network type.**

single connect/disconnect from mW to mWh using the values in Figure 2. We assume a 3.85V battery with a capacity of 2800mAh, which is typical. On the $x$-axis we control the desired user utility from 0.8 to 1, which effectively controls the amount of network disconnect/connect cycles a user will incur for the duration of a day. We multiply that number by the energy consumption (in mWh) and then divide by the battery's capacity to determine what fraction of the battery is consumed due to Zipphone. Table 8 presents our results, which indicate that the battery usage is at most 1% per day regardless of technology (3G or 4G) or desired privacy.

**Network control overhead.** Finally, although we do not explicitly quantify it, we do not expect that Zipphone users would incur significantly higher signalling overhead on the cellular network compared to non-Zipphone users. In order to release network resources and optimize clients' battery life, network providers forcefully disassociate users from the network after a network-defined timeout [55], typically in the order of a few seconds as illustrated by our measurements in Figure 7. Since Zipphone only operates when a user is inactive, the control overhead incurred by the network will be comparable with that from non-Zipphone users.

# 6 DISCUSSION

## 6.1 Limitations

Our technique has limitations. Privacy from the MVNO, and not just the MNO, requires that users make purchases anonymously. As such, our approach requires deliberate action from the user. And we require devices that accept software SIMs. Skyroam is one provider of devices based on a software SIM that operates in tens of countries around the world. Another limitation is that users would never be able to quantify their privacy gains as there is no way to determine the number of other Zipphone users. In addition, we do not address other privacy risks, which include physical attacks (e.g., radio frequency fingerprinting [20]), software vulnerabilities, use of location-based services, advertising fingerprints, browser cookies, and malware.

Our evaluations are limited as well. For example, we do not explicitly consider users mixing when they are stationary; if they do, attackers could also consider these additional mixes when linking. Attackers may also use more advanced classifiers that account for yet additional features (e.g., time of day or favourite locations [69]) to increase accuracy. Conversely, users could develop more efficacious methods to prevent linking.

Finally, our results are tied to our datasets, which are relatively small and limited to university populations. Obtaining a usable large-scale dataset is difficult, as MNOs are generally unwilling to anonymize and share such data. Furthermore, collecting user mobility data first-hand requires a fairly involved longitudinal effort.

Despite the limitations, this paper introduces an effective method for mobile network users to take charge of their own location privacy, and provides a detailed look at the efficacy of such a service.

## 6.2 Ethical implications

Mobile devices are an essential part of most people's daily routine. Accordingly, there is a tension between the right to location privacy and the need to investigate crimes and threats to public safety. The techniques we introduce and evaluate are effective to protecting privacy, but unfortunately would thwart a common method of investigation as well. Any deployment of Zipphone would have to take into account this difficult, zero-sum game ethical dilemma.

# 7 CONCLUSION

Our work demonstrates that, fundamentally, users do not need to trust wireless service providers with their location information. We evaluated a deanonymization attack that uses a combination of location profiling and trajectory linking, and showed that it is effective in identifying long-term pseudonyms. Using two separate datasets of call detail records, we then demonstrated that a Zipphone user can defend against such attacks by renewing her identifier regularly. We also evaluated the utility cost in terms of time offline and battery life, and showed it to be minimal. Users who do not use any anonymization scheme are always identifiable. In our trace-driven evaluations, a non-Zipphone user who is habitual and conventional (predictable and mixing) who renews her pseudonym monthly is identifiable 69% of the time, and one who uses Zipphone is identifiable 24% of the time if she sacrifices 5% of her utility and 1% of battery life, towards a lower bound of 19% if she sacrifices more. In other words, users can significantly reduce their identifiability by up to 45% by renewing their pseudonym after offline periods consuming less than 5% of their uptime.


# ACKNOWLEDGMENTS

This work was funded in part by NSF award CMMI-1831547.

# REFERENCES


[1] 2018. eSIM Whitepaper: The what and how of Remote SIM Provisioning. https://www.gsma.com/esim/wp-content/uploads/2018/12/esim-whitepaper.pdf.

[2] 2019. Find wireless carriers that offer eSIM service. https://support.apple.com/en-us/HT209096.

[3] 2020. Boost Mobile Prepaid Plans. https://www.boostmobile.com/plans.

[4] 2020. eSIM settings: Apple iPhone on iOS 12. https://support.t-mobile.com/docs/DOC-39253.

[5] 2020. eUICC Technical Releases. https://simalliance.org/euicc/euicc-technical-releases/.

[6] 2020. Silent Circle blackphone. http://silentcircle.com.

[7] 2020. T-Mobile SimplyPrepaid Plans. https://prepaid.t-mobile.com/prepaid-plans.

[8] 2020. Torfone. http://torfone.org.

[9] Miguel E Andrés, Nicolás E Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2013. Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. 901–914.

[10] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

[11] Alastair R. Beresford and Frank Stajano. 2003. Location Privacy in Pervasive Computing. *IEEE Pervasive Computing* 2, 1 (Jan. 2003), 46–55. https://doi.org/10.1109/MPRV.2003.1186725

[12] Alastair R. Beresford and Frank Stajano. 2004. Mix zones: user privacy in location-aware services. In *Proc. Pervasive Computing and Communications Wrkshps*. 127–131. https://doi.org/10.1109/PERCOMW.2004.1276918

[13] Lorenzo Bracciale, Marco Bonola, Pierpaolo Loreti, Giuseppe Bianchi, Raul Amici, and Antonello Rabuffi. 2014. CRAWDAD dataset roma/taxi (v. 2014-07-17). Downloaded from https://crawdad.org/roma/taxi/20140717. https://doi.org/10.15783/C7QC7M

[14] Case No. 19-cv-4063. 2019. Scott, Jewel, And Pontis, et al. v. AT&T Inc.; AT&T Services, Inc.; AT&T Mobility, LLC; Technocom Corp.; and Zumigo, Inc. https://www.courthousenews.com/wp-content/uploads/2019/07/ATTlocationservices-COMPLAINT.pdf.

[15] Eric Chan-Tin. 2015. AnonCall: Making Anonymous Cellular Phone Calls. In *2015 10th International Conference on Availability, Reliability and Security*. IEEE, 626–631.

[16] Mark D. Corner, Brian Neil Levine, Omar Ismail, and Angela Upreti. 2017. Advertising-based Measurement: A Platform of 7 Billion Mobile Devices. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*.

[17] Adrian Dabrowski, Nicola Pianta, Thomas Klepp, Martin Mulazzani, and Edgar Weippl. 2014. IMSI-Catch Me If You Can: IMSI-Catcher-Catchers. In *Proc. ACM ACSAC*.

[18] Alex Davidson, Ian Goldberg, Nick Sullivan, George Tankersley, and Filippo Valsorda. 2018. Privacy Pass: Bypassing Internet Challenges Anonymously. *Proceedings on Privacy Enhancing Technologies* 3 (2018), 164–180.

[19] Yves-Alexandre De Montjoye, César A Hidalgo, Michel Verleysen, and Vincent D Blondel. 2013. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports* 3 (2013), 1376.

[20] Shouyun Deng, Zhitao Huang, Xiang Wang, and Guangquan Huang. 2017. Radio frequency fingerprint extraction based on multidimension permutation entropy. *International Journal of Antennas and Propagation* 2017 (2017).

[21] Roger Dingledine, Nick Mathewson, and Paul Syverson. 2004. Tor: The Second-generation Onion Router. In *USENIX Security*. https://www.usenix.org/conference/13th-usenix-security-symposium/tor-second-generation-onion-router

[22] Cynthia Dwork. 2011. Differential privacy. *Encyclopedia of Cryptography and Security* (2011), 338–340.

[23] Nathan Eagle and Alex Sandy Pentland. 2006. Reality Mining: Sensing Complex Social Systems. *Personal and Ubiquitous Computing* 10, 4 (2006), 255–268.

[24] Ehab ElSalamouny and Sebastien Gambs. 2016. Differential Privacy Models for Location-Based Services. *Trans. Data Privacy* 9, 1 (April 2016), 15–48.

[25] Karim Emara, Wolfgang Woerndl, and Johann Schlichter. 2015. CAPS: Context-aware privacy scheme for VANET safety applications. In *Proceedings of the 8th ACM conference on security & privacy in wireless and mobile networks*. ACM, 21.

[26] Giulia Fanti, Shaileshh Bojja Venkatakrishnan, Surya Bakshi, Bradley Denby, Shruti Bhargava, Andrew Miller, and Pramod Viswanath. 2018. Dandelion++: Lightweight Cryptocurrency Networking with Formal Anonymity Guarantees. *Proc. ACM Meas. Anal. Comput. Syst.* 2, 2 (June 2018), 29:1–29:35.

[27] Mitra Fatemi, Somayeh Salimi, and Ahmad Salahi. 2010. Anonymous roaming in universal mobile telecommunication system mobile networks. *IET Information Security Journal* 4, 2 (2010), 93–103. https://doi.org/10.1049/iet-ifs.2009.0154

[28] Hannes Federrath, Anja Jerichow, Dogan Kesdogan, and Andreas Pfitzmann. 1995. Security in Public Mobile Communication Networks. In *Proc. IFIP/TC6 Personal Wireless Communications*. 105–116.

[29] Hannes Federrath, Anja Jerichow, and Andreas Pfitzmann. 1996. MIXes in Mobile Communication Systems: Location Management with Privacy. In *Proc. Intl. Wrkshp on Information Hiding*. 121–135.

[30] Julien Freudiger, Reza Shokri, and Jean-Pierre Hubaux. 2009. On the Optimal Placement of Mix Zones. In *Proc. PETS*. 216–234.

[31] Philippe Golle and Kurt Partridge. 2009. On the Anonymity of Home/Work Location Pairs. In *Proc. Intl. Conf. on Pervasive Computing*. 390–397. https://doi.org/10.1007/978-3-642-01516-8_26

[32] Maria Gorlatova, Roberto Aiello, and Stefan Mangold. 2011. Managing base station location privacy. In *Proc. MILCOM*. 1201–1206. https://doi.org/10.1109/MILCOM.2011.6127464

[33] Maria Gorlatova, Roberto Aiello, and Stefan Mangold. 2011. Managing location privacy in cellular networks with femtocell deployments. In *Proc. WiOpt Symposium*. 418–422. https://doi.org/10.1109/WIOPT.2011.5930055

[34] Marco Gramaglia, Marco Fiore, Alberto Tarable, and Albert Banchs. 2017. Preserving mobile subscriber privacy in open datasets of spatiotemporal trajectories. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 1–9.

[35] Nan Guo, Linya Ma, and Tianhan Gao. 2018. Independent mix zone for location privacy in vehicular networks. *IEEE Access* 6 (2018), 16842–16850.

[36] Byeongdo Hong, Sangwook Bae, and Yongdae Kim. 2018. GUTI Reallocation Demystified: Cellular Location Tracking with Changing Temporary Identifier.. In *Proc. ISOC Network and Distributed Systems Security (NDSS) Symposium*.

[37] Daira Hopwood, Sean Bowe, Taylor Hornby, and Nathan Wilcox. 2019. Zcash Protocol Specification Version. https://github.com/zcash/zips/raw/master/protocol/protocol.pdf.

[38] Haosheng Huang, Georg Gartner, Jukka M Krisp, Martin Raubal, and Nico Van de Weghe. 2018. Location based services: ongoing evolution and research agenda. *Journal of Location Based Services* 12, 2 (2018), 63–93.

[39] Syed Rafiul Hussain, Mitziu Echeverria, Ankush Singla, Omar Chowdhury, and Elisa Bertino. 2019. Insecure connection bootstrapping in cellular networks: the root of all evil. In *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*. ACM, 1–11.

[40] Sibren Isaacman, Richard Becker, Ramón Cáceres, Stephen Kobourov, Margaret Martonosi, James Rowland, and Alexander Varshavsky. 2011. Identifying Important Places in People's Lives from Cellular Network Data. In *Proc. Intl. Conf. on Pervasive Computing*. 133–151.

[41] Yixin Jiang, Chuang Lin, Xuemin Shen, and Minghui Shi. 2006. Mutual Authentication and Key Exchange Protocols for Roaming Services in Wireless Mobile Networks. *IEEE Trans. on Wireless Communications* 5, 9 (2006), 2569–2577. https://doi.org/10.1109/TWC.2006.05063

[42] Dogan Kesdogan, Hannes Federrath, Anja Jerichow, and Andreas Pfitzmann. 1996. Location Management Strategies Increasing Privacy in Mobile Communication. In *Information Systems Security*. 39–48.

[43] Mohsin Khan, Philip Ginzboorg, Kimmo Järvinen, and Valtteri Niemi. 2018. Defeating the downgrade attack on identity privacy in 5G. In *International Conference on Research in Security Standardisation*. Springer, 95–119.

[44] Hidetoshi Kido, Yutaka Yanagisawa, and Tetsuji Satoh. 2005. An anonymous communication technique using dummies for location-based services. In *Proc. Intl. Conf. on Pervasive Services*. 88–97.

[45] John Krumm. 2007. Inference Attacks on Location Tracks. In *Proc. Intl. Conf. on Pervasive Computing*. 127–143.

[46] Denis Foo Kune, John Koelndorfer, Nicholas Hopper, and Yongdae Kim. 2012. Location leaks on the GSM Air Interface. In *Proc. ISOC Network and Distributed Systems Security (NDSS) Symposium*.

[47] Marc Liberatore, Bikas Gurung, Brian Neil Levine, and Matthew Wright. 2011. Empirical Tests of Anonymous Voice Over IP. *Elsevier Journal of Network and Computer Applications* 34, 1 (January 2011), 341–350.

[48] Rongxing Lu, Xiaodong Lin, Tom H Luan, Xiaohui Liang, and Xuemin Shen. 2011. Pseudonym changing at social spots: An effective strategy for location privacy in vanets. *IEEE transactions on vehicular technology* 61, 1 (2011), 86–96.

[49] Chris Y.T. Ma, David K.Y. Yau, Nung Kwan Yip, and Nageswara S.V. Rao. 2010. Privacy Vulnerability of Published Anonymous Mobility Traces. In *Proc. MobiCom*. 185–196.

[50] Darakhshan J Mir, Sibren Isaacman, Ramón Cáceres, Margaret Martonosi, and Rebecca N Wright. 2013. Dp-where: Differentially private modeling of human mobility. In *2013 IEEE International Conference on Big Data*. 580–588.

[51] Yoni De Mulder, George Danezis, Lejla Batina, and Bart Preneel. 2008. Identification via Location-profiling in GSM Networks. In *Proc. ACM Wrkshp on Privacy in the Electronic Society*. 23–32. https://doi.org/10.1145/1456403.1456409

[52] Anandatirtha Nandugudi, Anudipa Maiti, Taeyeon Ki, Fatih Bulut, Murat Demirbas, Tevfik Kosar, Chunming Qiao, Steven Y Ko, and Geoffrey Challen. 2013. Phonelab: A large programmable smartphone testbed. In *Proceedings of First International Workshop on Sensing and Big Data Mining*. ACM, 1–6.

[53] Ben Niu, Qinghua Li, Xiaoyan Zhu, Guohong Cao, and Hui Li. 2015. Enhancing privacy through caching in location-based services. In *2015 IEEE conference on computer communications (INFOCOM)*. IEEE, 1017–1025.

[54] Jaegwan Park, Jaeseung Go, and Kwangjo Kim. 2001. Wireless authentication protocol preserving user anonymity. In *Proc. International Symposium on Wireless Personal Multimedia Communications*. 159–164.

[55] Feng Qian, Zhaoguang Wang, Alexandre Gerber, Zhuoqing Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2010. Characterizing radio resource allocation for 3G networks. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. 137–150.

[56] Michael G. Reed, Paul F. Syverson, and David M. Goldschlag. 1998. Protocols using anonymous connections: Mobile applications. In *Security Protocols*. LNCS, Vol. 1361. 13–23.

[57] Eli Ben Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. 2014. Zerocash: Decentralized Anonymous Payments from Bitcoin. In *2014 IEEE Symposium on Security and Privacy*. 459–474.

[58] Reza Shokri, George Theodorakopoulos, George Danezis, Jean-Pierre Hubaux, and Jean-Yves Boudec. 2011. Quantifying Location Privacy: The Case of Sporadic Location Exposure. In *Proc. PETS*. 57–76.

[59] Keen Sung, Brian Neil Levine, and Marc Liberatore. 2014. Location Privacy without Carrier Cooperation. In *Proc. IEEE Workshop on Mobile System Technologies (MoST)*. http://forensics.umass.edu/pubs/Sung-MoST-2014.pdf.

[60] Zhen Tu, Fengli Xu, Yong Li, Pengyu Zhang, and Depeng Jin. 2018. A New Privacy Breach: User Trajectory Recovery From Aggregated Mobility Data. *IEEE/ACM Transactions on Networking* 26, 3 (2018), 1446–1459.

[61] Daniel T Wagner, Andrew Rice, and Alastair R Beresford. 2013. Device analyzer: Understanding smartphone usage. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*. Springer, 195–208.

[62] Jinbao Wang, Zhipeng Cai, Yingshu Li, Donghua Yang, Ji Li, and Hong Gao. 2018. Protecting query privacy with differentially private k-anonymity in location-based services. *Personal and Ubiquitous Computing* 22, 3 (2018), 453–469.

[63] Shengling Wang, Qin Hu, Yunchuan Sun, and Jianhui Huang. 2018. Privacy preservation in location-based services. *IEEE Communications Magazine* 56, 3 (2018), 134–140.

[64] Matthew Wright, Micah Adler, Brian Neil Levine, and Clay Shields. 2004. The Predecessor Attack: An Analysis of a Threat to Anonymous Communications Systems. *ACM Transactions on Information and System Security (TISSEC)* 4, 7 (November 2004), 489–522.

[65] Yonghui Xiao and Li Xiong. 2015. Protecting locations with differential privacy under temporal correlations. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. 1298–1309.

[66] Yonghui Xiao, Li Xiong, Si Zhang, and Yang Cao. 2017. Loclok: Location cloaking with differential privacy via hidden markov model. *Proceedings of the VLDB Endowment* 10, 12 (2017), 1901–1904.

[67] Guomin Yang, DuncanS. Wong, and Xiaotie Deng. 2005. Efficient Anonymous Roaming and Its Security Analysis. In *Applied Cryptography and Network Security*. LNCS, Vol. 3531. 334–349. https://doi.org/10.1007/11496137_23

[68] Ling Yin, Qian Wang, Shih-Lung Shaw, Zhixiang Fang, Jinxing Hu, Ye Tao, and Wei Wang. 2015. Re-identification risk versus data utility for aggregated mobility research using mobile phone location data. *PloS one* 10, 10 (2015), e0140589.

[69] Hui Zang and Jean Bolot. 2011. Anonymization of Location Data Does Not Work: A Large-scale Measurement Study. In *Proc. ACM MobiCom*. 145–156.

[70] Jianming Zhu and Jianfeng Ma. 2004. A new authentication scheme with anonymity for wireless environments. *IEEE Trans. on Consumer Electronics* 50, 1 (2004), 231–235. https://doi.org/10.1109/TCE.2004.1277867