Revisiting Utility Metrics for Location Privacy-Preserving Mechanisms

Virat Shejwalkar
Information and Computer Science
University of Massachusetts, Amherst
vshejwalkar@cs.umass.edu

Amir Houmansadr
Information and Computer Science
University of Massachusetts, Amherst
amir@cs.umass.edu

Hossein Pishro-Nik
Electrical and Computer Engineering
University of Massachusetts, Amherst
pishro@engin.umass.edu

Dennis Goeckel
Electrical and Computer Engineering
University of Massachusetts, Amherst
dgoeckel@engin.umass.edu

ABSTRACT

The literature has extensively studied various location privacy-preserving mechanisms (LPPMs) in order to improve the location privacy of the users of location-based services (LBSes). Such privacy, however, comes at the cost of degrading the utility of the underlying LBSes. The main body of previous work has used a generic distance-only based metric to quantify the quality loss incurred while employing LPPMs. In this paper, we argue that using such generic utility metrics misleads the design and evaluation of LPPMs, since generic utility metrics do not capture the actual utility perceived by the users. We demonstrate this for ride-hailing services, a popular class of LBS with complex utility behavior. Specifically, we design a privacy-preserving ride-hailing service, called PRide, and demonstrate the significant distinction between its generic and tailored metrics. Through various experiments we show the significant implications of using generic utility metrics in the design and evaluation of LPPMs. Our work concludes that LPPM design and evaluation should use utility metrics that are tailored to the individual LBSes.

CCS CONCEPTS

• **Security and privacy → Privacy-preserving protocols**; **Usability in security and privacy**; *Social network security and privacy*; Pseudonymity, anonymity and untraceability.

KEYWORDS

Location Privacy-Preserving Mechanisms, Location Based Services, Utility Metrics, Ride Hailing Services

ACM Reference Format:
Virat Shejwalkar, Amir Houmansadr, Hossein Pishro-Nik, and Dennis Goeckel. 2019. Revisiting Utility Metrics for Location Privacy-Preserving Mechanisms.

ACSAC '19, December 9–13, 2019, San Juan, PR, USA

ACM ISBN 978-1-4503-7628-0/19/12…$15.00
https://doi.org/10.1145/3359789.3359829

In *2019 Annual Computer Security Applications Conference (ACSAC '19), December 9–13, 2019, San Juan, PR, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3359789.3359829

1 INTRODUCTION

Various location-based services (LBS) require their users' location information to operate. Such LBSes range from ride-hailing services like Uber to fitness applications like FitBit to recommendation systems like Yelp. Unfortunately, the platforms hosting such LBSes, e.g., cellphone operating systems, share fine-grained locations of users, which, in many cases, are more accurate than what is required by the LBSes to operate. For instance, a typical weather app LBS gets access to users' accurate locations, although it can function properly even with less accurate location information. To address this, the research community has investigated various location-privacy preserving mechanisms (LPPMs) [1, 5, 18, 20, 21], which aim to constrain the location information revealed to LBSes. The common techniques deployed by existing LPPMs are *anonymization* and *obfuscation* techniques, for instance location truncation, cloaking, generalization, and additive noise [1, 6, 13, 20, 27].

LPPM techniques improve privacy by trading off the utility of the underlying LBSes. For instance, a point of interest (POI) search LBS is likely to produce less useful recommendations when provided with inaccurate user locations. Previous work has used different metrics to quantify the impact of LPPMs on the privacy and utility of LBS services. Specifically, privacy has been quantified with metrics such as adversarial inference error [27], geo-indistinguishability [1], k-anonymity [13], conditional entropy [20], mutual information [18], and plausible deniability [3]. On the other hand, previous work has mainly used a *generic, distance-based metrics* to quantify the impact of LPPMs on utility. Such generic metrics measure the distance, e.g., Euclidean or squared Euclidean distance, between the real and obfuscated locations. The use of this distance-based utility metric is motivated by the intuition that the performance of an LBS is highly correlated with obfuscation amplitudes, e.g., increasing the obfuscation distance in a POI search LBS will degrade its utility by decreasing the quality of its recommendations [10].

In this paper, *we challenge the community's common use of a generic distance-based utility metric in designing and evaluating LPPMs.* We argue that a generic distance-only-based utility metric

does not capture the *actual utility perceived by users* of the underlying LBSes; this leads to LPPM designs being suboptimal with respect to the perceived utility. Therefore, we argue for using *application-tailored utility metrics* as opposed to generic (distance-only-based) utility metrics. An application-tailored utility metric (simply called tailored metric) aims at capturing the in-the-wild utility perceived by LBS users, and therefore is defined differently for different LBS systems. For instance, the tailored utility metric of a ride-hailing service like Uber should capture the time it takes for a rider to complete a ride, and the tailored utility metric of a fitness application should capture the burned calories (among other things); for both of these applications, distance is one of the features contributing to the perceived utility, but is not the only feature (as in generic utility metrics).

We demonstrate the implications of using tailored versus generic utility metrics by focusing on ride-hailing services (RHS). We choose RHS due to the complexity of their perceived utility, which can better demonstrate the distinction between tailored and generic utility metrics. However, our conclusions apply to any LBS with a utility metric/s that depends on parameters more than just the distance between real and obfuscated locations. We design a privacy-preserving RHS, called PRide, and define tailored and generic utility metrics for it. Our tailored metric captures the total time for ride completion and accounts for factors including surge pricing, behaviour and distribution of drivers.

A challenge to evaluating tailored metrics is the lack of public real-world LBS data (e.g., ride-hailing services do not make riding traces available to the public due to privacy and IP reasons). To overcome this challenge, we build an *RHS emulator* RHSE, which synthesizes RHS data and can be adjusted to various RHS environments and emulate different types of RHSes. We use our RHSE emulator to study multiple RHS environments for PRide, and compare the implications of tailored versus generic utility losses using state-of-the-art LPPM techniques [1, 5]. Our evaluations demonstrate that the utilities quantified using tailored and generic metrics are significantly different; therefore, using different utility metrics to deign LPPMs (i.e., optimizing privacy for a target utility or vice-versa) will result in substantially different LPPM parameters. We also show that using generic versus tailored metrics significantly impacts the outputs of the state-of-the-art utility improvement techniques, and that different utility losses associated with a given LBS should be combined according to user preferences for the utility improvement to be more effective and user-centric. To summarize, we make the following contributions:

- We demonstrate that the generic distance-only based metric, commonly used by the community to evaluate and design LPPMs, offers an incorrect perception of the actual utility perceived by users in practice. We therefore argue for the need to derive and use tailored utility metrics in the design, evaluation, and comparison of LPPM techniques for LBS services.
- We choose ride-hailing services to demonstrate the implications of generic versus tailored utility metrics, due to the complex nature of utility in such services. Towards this, we design a privacy-preserving ride-hailing protocol called PRide, discuss its privacy guarantees, and define a tailored utility metric for it.
- To overcome the lack of public real-world RHS data, we implement an RHS emulator, RHSE, that can emulate different RHS systems and environments.
- We perform extensive evaluations using our emulation of PRide. We demonstrate that the generic utility metric does not capture various important parameters of PRide that contribute to its in-the-wild utility; this motivates the need for tailored utility metrics in the design and evaluation of LPPMs.

Organization. The rest of this paper is organized as follows: § 2 reviews privacy and utility loss metrics from the previous literature and § 3 describes preliminaries. § 4 details the privacy preserving RHS, PRide, used to demonstrate our claims. § 5 details the RHS emulator, RHSE, built for data synthesis. § 6 details experiments on PRide data synthesized using RHSE along with results and their implications. § 6.6 introduces comprehensive utility loss and details its effects on the utility improvement techniques. We conclude the work in § 7.

2 RELATED WORK AND MOTIVATION

We briefly review the privacy and quality loss (QL) metrics proposed in the literature. We use the term *quality loss* to quantify utility degradation.

Various metrics have been proposed to quantify privacy improvements of LPPMs. Gruteser et al. [13] propose k-anonymity which provides privacy by adding a user's location to a set of other $(k-1)$ users' locations. Shokri et al. [27] argue that the privacy of a user is the inference error of the adversary, and propose adversarial inference error as the privacy metric. Oya et al. in [20] propose conditional entropy as a complementary metric to adversarial inference error to narrow the spectrum of optimal mechanisms for given QL expectation. Andres et al. [1] propose Geo-indistinguishability based on differential privacy [5, 10, 20, 21]. Similar to the differential privacy, it abstracts from the prior of the adversary and is robust with respect to composition. Due to the simplicity and theoretical guarantees of attaining privacy by adding Laplacian noise, Geo-indistinguishability is widely adopted by many tools, viz. Location Guard [1], LP-Guarding [12], and LP-Doctor [11].

Unlike privacy metrics, only a few QL metrics are proposed, which are variants of the Euclidean distance metric. Andres et al. [1] quantify QL of POI search LBS using (C, rad_I)-accuracy but without capturing the user preferences which can significantly affect QL improvement techniques [5] as we show in § 6.6. Chatzikokolakis et al. [5] propose a QL improvement by remapping obfuscated locations and evaluate using Euclidean distance as a QL metric. We provide empirical evidence that remapping using one QL metric can prove suboptimal towards other viable QLes in an LBS(§ A.2.2). Given the Euclidean distance-based QL constraints, Shokri et al. [28] construct the optimal LPPM against an optimal inference attack adversary. Oya et al. [20] consider average and worst case Euclidean distance as QL metrics.

Only a few works explicitly consider factors affecting the QL of specific applications to evaluate LPPMs. Micinski et al. [17] study POI search LBSes using three metrics namely, edit distance between the lists, overlap between retrieved results and the additional distance required to reach the closest entry on the list. Bilogrevic

Table 1: Privacy and quality loss metrics used in previous location privacy research

Previous work	Application	LPPM	Privacy metric	QL metric
Shokri [27]	None	Cloaking, precision reduction	Adversarial inference	None
Andres [1]	POI retrieval	Laplace noise	Differential privacy	Squared Euclidean
Chatzikokolakis [5]	Check-ins	Laplace noise	Differential privacy	Euclidean
Bilogrevic [2]	Check-ins	Semantic/geographical obfuscation	None	Check-in motivations
Shokri [28]	None	Cloaking, precision reduction	Adversarial inference	Euclidean
Fawaz [12]	POI, Healthcare	Laplace noise, precision reduction	Differential privacy	User survey
Micinski [17]	POI retrieval	Cloaking	None	Retrieved sets' overlap
Pham [22, 23]	RHS	Cloaking	N/A	N/A
Oya[20]	None	Laplace noise	Differential privacy, conditional entropy	Average Euclidean, worst case Euclidean

et al. [2] introduce *perceived utility* metric for check-in services based on different categories of motivations. However, the QL metrics considered will fail in the case of continuous LBS such as RHS because applying semantic obfuscation has no effect on the QLes encountered commonly in continuous LBS. Hence, evaluating LPPMs for continuous LBS vs for static/one-time LBS, such as POI search, is different from the point of view of QL metrics. The work in [2] is based on a user survey which is a gruesome task; therefore, such works need to devise QLes tailored to applications and release them to the community for further research. Oya et al. [21] reconsider privacy vs QL trade-offs of LPPMs that guarantee Geo-indistinguishability to devise an alternate privacy metric to improve the trade-offs; evaluation of all metrics is using the generic QL metrics. Pham et al. [22, 23] consider privacy of different players in ride hailing services and propose PrivateRide - a new protocol that provides security and privacy guarantees based on established cryptographic paradigms. However, the work neither evaluates PrivateRide in terms of QL due to cloaking LPPM nor analyzes trade-offs of privacy-QL.

Table 1 summarizes the privacy and QL metrics proposed in the past. We see that the majority of QL metrics are distance-based only, presumably to facilitate problem formulation and make the analysis of optimal algorithms tractable [20].

3 PRELIMINARIES

We start by introducing preliminaries on the main concepts used across the paper.

3.1 Ride-hailing Services

A ride-hailing service (RHS) has three main players: *riders*, *drivers*, and a *service provider* (*server*). The server collects and manages all of the data of riders and drivers, and is responsible for matching riders to nearby drivers, calculating ride fares to charge riders and fascilitating payments to drivers. Using RHS mobile application, a rider sends a requests to the server which contains her location and reputation. Then the server searches for drivers nearby the rider's location and sends a request to the nearest available driver using the RHS driver application. This request contains the rider's location, reputation, and surge around the rider's location. If a driver rejects the request, it is sent to the next available driver. Once a request is accepted, driver details are sent to the rider and she chooses to accept or reject the driver. If accepted, the rider and driver collaborate on pick up. At any time, if the driver rejects a ride request, a penalty in terms of money or reputation is imposed. Note that the details of such interactions may differ across in-the-wild RHSes.

3.2 Utility (Quality Loss) Metrics

A quality loss (QL) metric quantifies the *utility* degradation of an LBS due to the use of LPPMs. The expected QL is formulated as [1, 5, 28]:

$$\hat{QL}(\text{LPPM}, \pi, d_Q) = \sum_{\substack{l_r \in \mathcal{X} \\ l_o \in \mathcal{Z}}} \pi(l_r) \cdot \text{LPPM}(l_r)(l_o) \cdot d_Q(l_r, l_o) \quad (1)$$

where $\pi(l_r)$ is locations' prior distribution, $\text{LPPM} : \mathcal{X} \rightarrow \mathcal{Z}$ is the LPPM and, l_r and l_o are the real and obfuscated locations, respectively. $\text{LPPM}(l_r)(l_o)$ denotes the probability of obfuscating l_r to l_o.

$d_Q(l_r, l_o)$ quantifies the QL metric for a single LBS access, and $\hat{QL}$ quantifies the expected loss over $\mathcal{X}, \mathcal{Z}$. Prior works mainly use distance-based measures to define $d_Q(l_r, l_o)$, i.e., generic metrics. In this work, we argue to tailor the definition of $d_Q(l_r, l_o)$ to specific applications.

3.2.1 Generic QL Metric. The generic QL metric models $d_Q(l_r, l_o)$ as a function only of the obfuscation distance, i.e., Euclidean distance between l_r and l_o: $d_Q : \mathcal{X}, \mathcal{Z} \rightarrow \mathbb{R}$. The majority of previous works use this metric to simplify formulation and analysis. The average Euclidean and average squared Euclidean distances are the most commonly used metrics [1, 5, 6, 26, 28] for $d_Q(l_r, l_o)$. We denote Euclidean distance as QL_g in the rest of the paper; in case of geolocation coordinates, one can replace Euclidean with Haversine distance [29].

3.2.2 Tailored QL Metrics. A Tailored QL metric aims to capture the QL perceived by LBS users in the wild. Such metrics are not well studied in the literature due to their complexity, highly subjective nature, and consequent non-tractability. For instance, for a fitness LBS the calories burnt during an activity is one of the tailored QL metrics. Calories burnt captures speed, elevation, heart rate, and

basal metabolic rate etc. [14] of the users. Hence, the tailored QL is a function of all the attributes listed above, instead of obfuscation distance only. Similarly, for RHSes the ride completion time and/or ride fares can be the tailored QL metrics. We note that, *it is hard to formalize these metrics*; nevertheless, as we systematically demonstrate, tailored QL metrics are important for LPPM design, and for better user-experience LBSes can devise them using the comprehensive QL notion we propose.

3.3 Geo-indistinguishability and Location Privacy Preserving Mechanisms

We use geo-indistinguishability [1] in our experiments as the privacy metric. Geo-indistinguishability, derived from differential privacy, formalizes privacy guarantees for location sensitive data. Suppose $\mathcal{X}, \mathcal{Z} \subseteq \mathbb{R}^2$ are the domains of real and obfuscated locations, and $P(\mathcal{Z})$ is the set of probability distributions over $\mathcal{Z}$. $\mathsf{LPPM} : \mathcal{X} \rightarrow P(\mathcal{Z})$ provides Geo-indistinguishability if:

$$\ln\left|\frac{\mathsf{LPPM}(l_r)(l_o)}{\mathsf{LPPM}(l'_r)(l_o)}\right| \leq \epsilon \mathbf{d}(l_r, l'_r) \tag{2}$$

In (2), we assume that $\ln(\frac{x}{y})$ is 0 if both x,y are 0 and ∞ if one of them is 0. $\mathbf{d}(l_r, l'_r)$ is an arbitrary distance function, and ϵ is the privacy budget. An LPPM provides (ℓ, r)-*geo-indistinguishability*, if it provides an ℓ level of privacy within radius r of the actual location; this is achieved by setting $\epsilon = \ell/r$ in (2). Geo-indistinguishability ensures that an LPPM obfuscates both the locations l_r and l'_r to l_o with near equal probabilities, making it difficult for an adversary to reverse-engineer the real location among l_r and l'_r after observing l_o.

In our evaluations, we use three state-of-the-art LPPMs with geo-indistinguishability guarantees: Planar Laplace (LPPM_p), Geometric (LPPM_g), and Exponential (LPPM_e) mechanisms [1, 5]. We consider a 10Km × 10Km region for all of our experiments. As the considered region is finite, we employ truncated versions of the planar Laplace and geometric mechanisms. In all our experiments, $d(.)$ is Euclidean distance; LPPM_p incurs average Euclidean loss of $2/\epsilon$, but due to the truncation, the loss incurred is less than $2/\epsilon$ in our experiments for small ϵ values. For the probability mass function of the three LPPMs and implementation details, we refer the reader to [5].

4 PRIDE: A PRIVACY-PRESERVING RHS

To demonstrate the impact of tailored and generic utility metrics on design and evaluation of LPPMs, we choose ride-hailing services due to the complex nature of their perceived utility. We introduce a privacy-preserving RHS instance PRide which uses the state-of-the-art privacy mechanisms (§ 3.3) to preserve privacy of RHS drivers and riders.

4.1 Threat Model and Privacy Guarantees

We consider the PRide server to be adversarial, who tries to learn locations of riders. The server has some prior information about the distribution of rider's true locations and uses posterior information from each ride to infer the rider's true locations. The posterior information includes the rider's hailing location and the locations of drivers around it, the location of the matched driver, and the ride

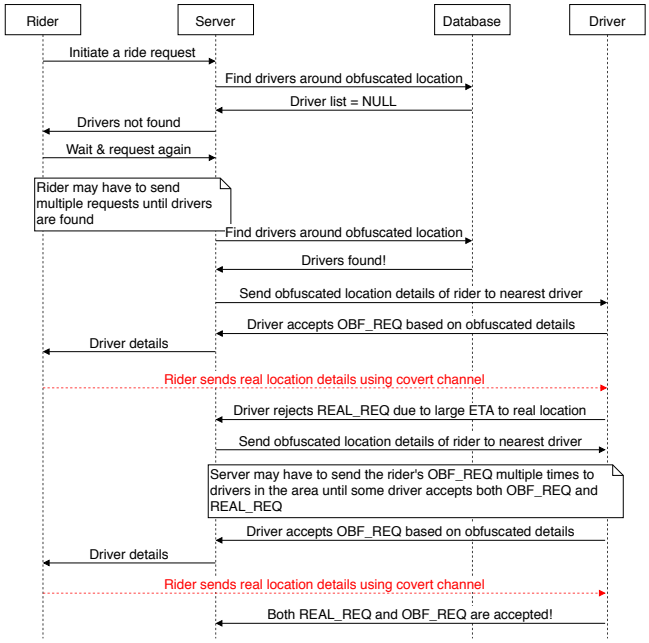

Figure 1: A common PRide scenario described in Section 4.3.

destination. We assume all of the locations in the posterior information are obfuscated using geo-indistinguishability mechanisms, and below, analyze the corresponding privacy protection to *riders*.

The specifics of obfuscations in PRide are as follows: 1) drivers' locations used for ride matching are obfuscated using a constant privacy budget ϵ_1, 2) drivers report only the pick-up and destination locations of a ride to the server; pick-up location is the rider's obfuscated location l_o which spends ϵ_r budget and destinations are obfuscated by the rider using budget ϵ_2. In essence, all the posterior information is obfuscated and prevents the server from inferring true locations of riders. Therefore, by obfuscating locations in this way, riders achieve $(\epsilon_r + \epsilon_1 + \epsilon_2)$-geo-indistinguishability due to composability [9]. Hence, using ϵ_r privacy budget, riders can achieve privacy linear in ϵ_r.

Note that our work aims to thoroughly investigate the effects of tailored versus generic QL metrics. Therefore, to abstract from PRide's privacy analysis, we make the following assumptions. First, *drivers' locations can be obfuscated to achieve any desired distribution of drivers*. For instance, we assume that uniform distribution of drivers can be achieved after obfuscation. This assumption allows us to obfuscate only the riders' locations in the rest of the paper. Second, along with drivers' location obfuscation, out-of-band secure channels between riders and drivers, and anonymous payments are used by PRide [22, 23] for riders to achieve the above-specified geo-indistinguishability. This assumption allows us to abstract from the detailed privacy analysis of PRide protocol and focus on QL quantifications when only riders obfuscate their locations.

4.2 Notations

We clarify some notations below. The expected time of arrival ETA is calculated for a combination of the *hailing location* of rider wrt

Table 2: Notations used in describing PRide and RHSE. These are explained in detail in § 4.2

Symbol	Significance
l_r, l_o	Real location of rider, obfuscated location obtained by applying LPPM on l_r
$\mathsf{ETA_t}$	ETA tolerance of drivers that contributes to their choice of serving a ride
$\mathsf{ETA_r^r}$	ETA *to real* location, l_r, when hailing location is *real*, l_r
$\mathsf{ETA_o^o}$	ETA *to obfuscated* location, l_o, when hailing location is *obfuscated*, l_o
$\mathsf{ETA_r^o}$	ETA *to real* location, l_r, when hailing location is *obfuscated*, l_o
OBF_REQ,REAL_REQ	Ride requests that, respectively, reveal obfuscated location, l_o, and real location, l_r
(ℓ, r)	Geo-indistinguishability guarantee: privacy level ℓ within radius r
$\mathsf{LPPM_p}$, $\mathsf{LPPM_g}$, $\mathsf{LPPM_e}$	Continuous planar Laplace, Geometric Laplace, Exponential LPPMs
S	Surge factor calculated as ratio of number of active ride requests and active drivers in area
(H, M, L)	Levels of strictness of drivers for different surge factors - High, Medium, Low respectively
$\mathsf{M_d}$	Driver's acceptance model: tuple of levels of strictness for three different surge factor ranges
$\mathsf{QL_t}$, $\hat{QL}_t$	Tailored QL for a single LBS access, empirical expected tailored QL for a user
$\mathsf{QL_g}$, $\hat{QL}_g$	Generic QL for a single LBS access, empirical expected generic QL for a user

(with respect to) server and the pickup location of the rider with respect to driver. Either l_r or l_o can be used to *hail a ride* and is called a *hailing location*. If a ride request sends l_r to a driver, it is called REAL_REQ and if it sends l_o, it is called OBF_REQ. Specifically, if l_r is used to hail a ride, the only pickup location for the driver is l_r; this is denoted as $\mathsf{ETA_r^r}$ i.e., hailing and pickup locations are both l_r (row 1 in Table 3). However, to preserve privacy, PRide hails a ride using l_o i.e., the rider first releases only l_o and all drivers see ETA to l_o (row 2 in Table 3) which is denoted as $\mathsf{ETA_o^o}$ i.e., the hailing location is l_o and the pickup location is l_o.[1] If some driver accepts this request, then, only to that driver, the rider releases her l_r i.e., the hailing location is l_o but the pickup location is l_r hence corresponding ETA is $\mathsf{ETA_r^o}$ (row 3 in Table 3). All of the notations are summarized in Table 2 and 3.

Table 3: PRide request types. In no-privacy case (row-1), only the REAL_REQ is sent. In privacy preserving (row-2+3) case, OBF_REQ is sent followed by the REAL_REQ. In row-2, the rider does not send anything to any specific driver but in row-3, the rider sends l_r, through a out-of-band secure channel, to the driver who accepts the OBF_REQ. R,D,S are rider, driver and server respectively.

Request	R→S	R→D	S	D	ETA
REAL_REQ	l_r	NA	l_r	l_r	$\mathsf{ETA_r^r}$
OBF_REQ	l_o	NA	l_o	l_o	$\mathsf{ETA_o^o}$
OBF_REQ → REAL_REQ	l_o	l_r	l_o	l_r	$\mathsf{ETA_r^o}$

4.3 The PRide Protocol

In PRide, a rider first uses obfuscated location, l_o, to send *obfuscated* ride request, OBF_REQ, and only when some driver accepts the OBF_REQ, the rider reveals her real location, l_r using REAL_REQ through some secure channel. RHS applications allow drivers to see only the ETA to rider's hailing location. Therefore, drivers first see ETA to obfuscated location, $\mathsf{ETA_o^o}$, and then see ETA to real location, $\mathsf{ETA_r^o}$. Drivers may accept OBF_REQ due to low ETA, $\mathsf{ETA_o^o}$, to l_o but ultimately cancel it due to high ETA, $\mathsf{ETA_r^o}$ to l_r. However, the minimum acceptance policies [8, 15] in RHSes do not allow drivers to cancel the accepted rides very often. Sequence diagram in Fig. 1 demonstrates the following scenario common in PRide (and in any RHS):

- Rider obfuscates l_r to l_o and sends OBF_REQ; she may have to resend the request due to unavailability of drivers around l_o (the hailing location).
- Server finds drivers within search radius and forwards the OBF_REQ to nearest driver.
- A driver accepts OBF_REQ based on $\mathsf{ETA_o^o}$ but rejects REAL_REQ due to high $\mathsf{ETA_r^o}$; server may forward the request to drivers multiple times.
- Subsequently, some driver accepts both OBF_REQ and REAL_REQ based on her $\mathsf{M_d}$ (§ 5.1.2).

4.4 Tailored QL Metric for PRide

In this section, we formalize the tailored QL, $\mathsf{QL_t}(l_r, l_o)$[2], used in our evaluations: *Difference in time to complete a ride when hailing location is real versus obfuscated.* We ignore the non-quantifiable factors such as behavior of drivers as they do not change with obfuscation and cannot affect design and/or evaluation of LPPMs. Modeling $\mathsf{QL_t}$ analytically is difficult; however, RHSE (§ 5) can synthesize data of the time required to complete rides which can be used as $\mathsf{QL_t}(l_r, l_o)$.

The three stages of a ride (with corresponding times in parentheses) are *driver allocation* ($\mathsf{T_{DA}}$), *rider pick up* ($\mathsf{ETA_r^r}$ or $\mathsf{ETA_r^o}$), and *ride to destination* ($\mathsf{T_{SD}}$). Note that, $\mathsf{T_{SD}}$ is the same for l_r and l_o, hence does not contribute to $\mathsf{QL_t}$. However, based on which location (l_r versus l_o) is used to hail a ride, the driver allotted to a rider changes due to the probabilistic nature of drivers' models (§ 5.1.2). This changes the ETA as different drivers can have different ETAs. Therefore, with change in the hailing location, time to allocate a driver, $\mathsf{T_{DA}}$, also changes and hence contributes to $\mathsf{QL_t}$.

An example is shown in Fig. 2. Here, a ride hailed using the real location, l_r, gets accepted by $\mathsf{D_2}$, because $\mathsf{D_2}$ has higher $\mathsf{ETA_t}$ than her ETA to l_r i.e., $\mathsf{ETA_r^r}(\mathsf{D_2}) < \mathsf{ETA_t}(\mathsf{D_2})$. However, when the ride is hailed using obfuscated location, l_o, the server first sends obfuscated request, OBF_REQ, to $\mathsf{D_5}$ and she accepts it, but $\mathsf{D_5}$

[1] Superscript of ETA denotes location from where the ride is hailed and subscript denotes the pick-up location with respect to the driver.

[2] We drop (l_r, l_o) when it is clear from the context.

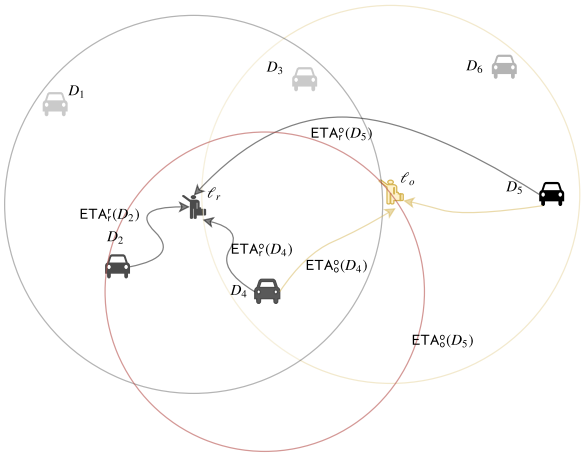

Figure 2: Schematic explaining tailored QL quantification in PRide. Black or yellow circles denote the search radius with respect to l_r or l_o, respectively, and the red circle denotes ETA_t of driver D_4. Various notations are in Table 2.

rejects the REAL_REQ because D_5's ETA to l_r is higher than her ETA_t i.e., $\text{ETA}^o_r(D_5) > \text{ETA}_t(D_5)$ (for demonstration we assume (HHH) model, which we define in § 5.1.2, for D_5). The next nearest driver, D_4, accepts both OBF_REQ and REAL_REQ, because both $\text{ETA}^o_o(D_4)$ and $\text{ETA}^o_r(D_4)$ are less than $\text{ETA}_t(D_4)$ as shown by the red circle. Hence, for the rider in Fig. 2,

$$\text{QL}_t = (\text{T}_{\text{DA}}(D_4) + \text{ETA}^o_r(D_4)) - (\text{T}_{\text{DA}}(D_2) + \text{ETA}^r_r(D_2))$$

Therefore, for a rider who is allotted drivers D_o, D_r when hailing locations are l_o, l_r respectively, QL_t can be written as:

$$\text{QL}_t(l_r, l_o) = \overbrace{(\text{T}_{\text{DA}}(D_o) + \text{ETA}^o_r(D_o))}^{\text{ride completion time with } l_o} - \overbrace{(\text{T}_{\text{DA}}(D_r) + \text{ETA}^r_r(D_r))}^{\text{ride completion time with } l_r} \tag{3}$$

Finally, assuming all of the players behave similarly in a single PRide run, we can calculate the expected QL_t using the following equation:

$$\hat{QL}_t(\text{LPPM}, \pi, d_T) = \sum_{\substack{l_r \in \mathcal{X} \\ l_o \in \mathcal{Z}}} \pi(l_r).\text{LPPM}(l_r)(l_o).\text{QL}_T(l_r, l_o) \tag{4}$$

In (4), $\text{QL}_t(l_r, l_o)$ quantifies the tailored QL of PRide as defined by (3), rider prior, $\pi(l_r)$, is assumed uniform, i.e., $(\frac{1}{A})$ where A is the area under consideration. Previous works calculate expected Euclidean loss in (1) as the numerical average of the Euclidean loss over multiple runs of LPPMs. Similarly, we calculate $\hat{QL}_t$ numerically using the data of all the rides, $\mathcal{R}$ in one PRide run as follows, where QL_t is for one access of LPPM:

$$\hat{QL}_t(\text{LPPM}, \pi, d_T) = \frac{1}{|\mathcal{R}|} \cdot \sum_{\mathcal{R}} \text{QL}_t(l_r, l_o) \tag{5}$$

5 RHSE: OUR RHS EMULATOR

As motivated in §1, we design RHSE to synthesize relevant PRide data. We detail the RHSE emulator below.

5.1 RHSE Players and Their Behaviors

5.1.1 Riders. In PRide, riders initiate a ride request by generating l_r, (optionally) $l_o = \text{LPPM}(l_r)$, and a destination, and wait until some driver accepts the request. The hailing location of riders can be either l_r or l_o; our tailored QL metric is based on the difference in the times to complete rides in these two cases (§4.4).

If there are no drivers around the hailing location or none of the available drivers accepts the ride request, the rider waits or sends the request again to the server. Once the request is accepted, the rider waits for the serving driver to pick up and then the ride completes as usual. Note that request acceptance includes accepting both REAL_REQ and OBF_REQ if the hailing location is l_o, i.e., the privacy preserving case. Without privacy, i.e., if hailing location is l_r, accepting just the REAL_REQ is sufficient. At the end of the ride, the rider is relocated to another location. We assume that the riders never give up on a ride search until the end of an RHSE run.

5.1.2 Drivers. Drivers take important decision in RHSE which affects the tailored QL of riders the most: *whether to accept a ride request or not.* Drivers, when active, are either waiting for a ride request or serving someone's request. Driver locations used for ride-matching are assumed obfuscated and their destinations of rides are the same as that of the riders they serve. Drivers get requests from riders within a search radius and based on their acceptance model, M_d, and ETA_t, they chose to accept/reject the ride requests. If multiple drivers are around the hailing location of the rider, they are arranged in a queue based on their ETA to the *hailing* location of the rider; note that the server need not know the real location for this task. If a driver rejects the request, the next driver in the queue is sent the request. Following the RHS policies [8, 15, 16], in RHSE, a driver with acceptance rate less than 80%, after rejecting a ride, is blocked for 5 RHSE time units; however, if a driver accepts a request and acceptance rate is below 80% no penalty is imposed. Below, we describe the three main attributes of drivers' acceptance model, M_d.

ETA tolerance: Drivers prefer rides with ETAs below a particular threshold ETA [4] called ETA tolerance, ETA_t, and is an attribute of drivers in RHSE. In reality ETA_t can vary with drivers and times of the day but for the ease of interpretation of results we keep ETA_t constant for all of the drivers at the *beginning* of RHSE. However, beyond a particular number of consecutively rejected rides, called *threshold rejection count*, ETA_t is increased linearly. ETA_t is reset to its initial value if a driver accepts a request. This realizes an intuitive scenario where drivers accept higher ETA rides if they do not get a suitable ride for a long time.

Behavior based on $(\text{ETA} - \text{ETA}_t)$**:** Different probability distributions over the difference[3] $(\text{ETA} - \text{ETA}_t)$ also represent different behaviors of drivers. That is, while some drivers may strictly reject ride requests with ETA greater their ETA_t, some may be more tolerant. Hence, based on the strictness with which drivers follow their

[3] As explained in Section 4.2, ETA can be either ETA^r_r, ETA^o_r, or ETA^o_o depending on the hailing and pick up locations.

Table 4: Tunable parameters of RHSE with usability of each in the Usability column. The ENV_b column specifies the baseline environment.

Parameter	Meaning	Usability	ENV_b
M_d	Driver model (§ 5.1.2)	Drivers behavior for different surges → high/medium/low	HHH
ETA_t	ETA tolerance of drivers	Drivers behavior towards ETA value → high/low	400
R, D	Number of riders and drivers	Density of players → NYC vs Luxembourg	(200,120)
ℓ	Geo-indistinguishability privacy level	Privacy-QL awareness → high vs low	ln(1.4)
r	Obfuscation radius	Privacy-QL awareness → high vs low	1km
$(Lat_i, Lon_i)_{i\in[0,3]}$	Geographical region considered	Route conditions and typical ETAs → Cincinnati vs Manila	Paris
$DELAY_{req}$	Delay in subsequent requests	Rider's urgency → high vs low	0
LPPM	Location privacy preserving mechanism	LPPM under test	$LPPM_p$
$GRID_{res}$	Grid resolution for discrete LPPMs	Privacy-QL requirement → high vs low	NA
$\pi(R)$, $\pi(D)$	Initial distribution of players	Measuring effect of surge → uniform vs non-uniform	Uniform

ETA_t, there are three types of drivers – high (H), medium (M), low (L). This behavior is configured using the probability of accepting a ride as a function of $(ETA - ETA_t)$, where the step, exponential and slow-exponential functions are used for high, medium, and low strictnesses, respectively. This is demonstrated in Fig. 3: The drivers with high strictness follow a step function with threshold 0 for $(ETA - ETA_t)$, i.e., if $(ETA - ETA_t) > 0$ they accept the ride otherwise reject it. Fig. 7 shows the combined effect of varying ETA_t for three different driver models, namely, HHH, MMM, LLL while keeping $\hat{QL}_g$ constant.

Behavior based on surge factor: In Uber/Lyft, due to monetary advantages, drivers prefer to serve requests in the region of high price surge [25]. Similarly, drivers in RHSE choose to serve a ride request based on *surge factor* S in the area around the *hailing* location. Therefore, a driver adapt her strictness of following ETA_t with S, i.e., changes her $P(ETA - ETA_t)$. S is calculated as the ratio of the number of active requests and the number of idle drivers in an area:

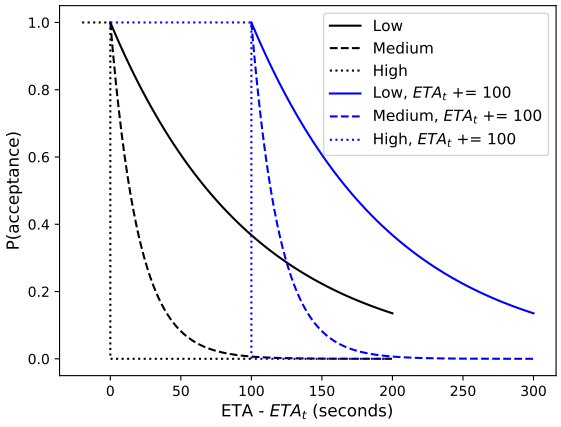

Figure 3: Strictness of ETA_t: For the same tolerance, different drivers behave differently. High (dotted), medium (dashed) and low (solid) strictnesses are modeled using step, exponential and, slow exponential functions, respectively.

$$S = \frac{\text{Number of active requests}}{\text{Number of idle drivers}} \tag{6}$$

In (6), the denominator and numerator are both positive integers. Ride requests with $S > 1$ (more riders than drivers, S^+) are favored over those with $S < 1$ (more drivers than riders, S^{0-1}) by drivers. An event of no drivers around the hailing location of a rider, S^-, is assumed high surge and given the highest priority. For the three surge factor ranges, (S^-, S^{0-1}, S^+), drivers behave differently according to different $P(ETA - ETA_t)$ as shown in the Fig. 3.

We can now define the drivers' model, M_d, as a tuple of probability distributions for three surge ranges: $(P_{S^-}, P_{S^{0-1}}, P_{S^+})$. Hence, (L,H,M) (or simply LHM) would imply a drive model that follows low, high, and medium strictness in (S^-, S^{0-1}, S^+) ranges respectively. Note that M_d defined in the above specified manner effectively captures all the three attributes of a driver.

5.1.3 Server. The RHSE server acts as a medium between riders and drivers. The server does not actively change the distribution of drivers or riders and therefore does not affect the privacy-utility trade-offs for riders. Therefore, the service provider (server) behavior is not particularly modeled.

5.2 Configuring Scenarios

By tuning the parameters of RHSE, various commonly occurring scenarios of a typical RHS can be synthesized. The parameters are summarized in Table 4 along with their significance in devising various environments. For example, scenario with high demand and low supply, a common occurrence during peak office hours, can be realized by increasing the number of riders, R and decreasing number of drivers, D. For the same number of riders and drivers, changing geographical area, e.g., Cincinnati versus Manila, significantly changes the pick up ETAs; daily average time to travel 4-5 Km in Manilla is 31-32 minutes vs that in Cincinnati it is 9-10 minutes [19]. In RHSE, a region of consideration can be specified using a tuple of region's geo-coordinates. Different LPPMs can be plugged into RHSE (§ 6.3). Any combination of values of these parameters is termed *environment*, ENV, that realizes a unique scenario of RHS.

6 COMPARING TAILORED VERSUS GENERIC QL METRICS

In this section, we show the effects of using generic versus tailored QL metrics on the design and evaluation of LPPMs using PRide as an example LBS. First, we analyze the planar Laplace mechanism, $\mathsf{LPPM_p}$, under different PRide environments and show the effects of PRide parameters on *evaluation* of $\mathsf{LPPM_p}$. In the later sections, we show how the use of generic versus tailored QL metric affects results of LPPM comparison (§6.3), state-of-the-art utility improvement techniques (§6.5 and §6.6) and choice of privacy budget (§6.4).

6.1 Experimental Setup

In all of our experiments, we use the expected generic and tailored QLes as formulated in (1) and (5), respectively. We consider a square region of 10Km×10Km bounded from left and right by longitudes 2.275873 and 2.421079, respectively, and bounded above and below by latitudes 48.810519 and 48.901606, respectively. $\mathsf{LPPM_p}$ and $\mathsf{LPPM_g}$ are designed for infinite domains; therefore, we use their truncated versions. Note that truncation preserves privacy due to its deterministic nature. We perform experiments for both uniform and nonuniform driver distributions, $\pi(\mathsf{D})$. To realize nonuniform $\pi(\mathsf{D})$, drivers are relocated to the left half of the considered region with probability 0.9 instead of a uniform relocation at the end of the rides. *We use such extreme drivers' distribution as the nonuniform case to show the PRide parameters' effect vividly.* To get ETAs and distances between locations, we query the OSRM Table Service API [24]. In the figures with boxplots, solid brown boxplots denote $\hat{QL}_g$ while empty blue ones denote $\hat{QL}_t$.

6.2 Impact on LPPM Evaluation

In this section, we compare $\hat{QL}_t$ of $\mathsf{LPPM_p}$ while varying different PRide parameters and keeping $\hat{QL}_g$ constant. We keep all parameters of PRide constant as in $\mathsf{ENV_b}$ given in Table 4 and observe the effect of modifying a single PRide parameter, e.g. drivers' acceptance model, on $\hat{QL}_g$ and $\hat{QL}_t$. We show that although $\hat{QL}_t$ changes with change in these parameters, $\hat{QL}_g$ remains constant. This demonstrates that the generic metric does not account for the contribution of PRide parameters to the tailored metric; we will show in § 6.3 how this impacts the design aspects of LPPMs.

6.2.1 Drivers' acceptance model. We change the strictness of drivers' acceptance models, $\mathsf{M_d}$, from high to low, to understand how it affects $\hat{QL}_t$. To check consistency across ϵ, we vary $r \in \{0.5, 1\}$Km; higher r implies higher $\hat{QL}g$ and vice-versa. As seen in Fig. 4 (left), drivers' tendency to accept rides with the same average ETA increases with decreased strictness of $\mathsf{M_d}$ which reduces driver allocation time, $\mathsf{T_{DA}}$, and hence, $\hat{QL}_t$ (denoted by empty blue boxes). This is consistent across different r, because for all r's $\hat{QL}_t$ decreases with decreased strictness of $\mathsf{M_d}$. The drivers' distribution here is uniform, however, we observe the similar reduction in $\hat{QL}_t$ as $\mathsf{M_d}$ becomes more tolerant with nonuniform $\pi(\mathsf{D})$, as shown in Fig. 5 (left). These experiments validate our hypothesis that the tailored QL metrics account for the changes in parameters of the LBS, which the generic QL metrics miss.

6.2.2 Drivers' ETA tolerance. With increase in $\mathsf{ETA_t}$ of drivers from 400 to 2000 seconds, we observe a reduction in $\hat{QL}_t$, as shown in Fig. 4 (middle). With higher $\mathsf{ETA_t}$, drivers' probability to accept rides with higher ETAs increases; hence, $\mathsf{T_{DA}}$ (§ 4.4) decreases and so does $\hat{QL}_t$. However, $\hat{QL}_g$, having no such relation to $\mathsf{ETA_t}$, remains almost constant; this is observed for different obfuscation radii r's and non-uniform driver distributions (Fig. 5 (right)). In addition, we observe that at higher $\mathsf{ETA_t}$, $\mathsf{M_d}$ does not contribute to $\hat{QL}_t$, because at higher $\mathsf{ETA_t}$ even the strictest driver model HHH accepts all ride requests. Hence, QL_t remains almost constant and equal to the time to cover constant $\hat{QL}_g$ as shown in Fig. 7. *Therefore, $\hat{QL}_t$ is a function of $\mathsf{ETA_t}$, $\mathsf{M_d}$ and $\hat{QL}_g$.*

6.2.3 Number of players. The number of active riders and drivers can affect $\hat{QL}_t$ of riders. For instance, high demand can be realized by increasing active riders and their request and/or decreasing the number of active drivers; and vice versa for high supply. Therefore, without loss of generality, we realize high supply and/or low demand by increasing the number of active drivers from 120 to 400 in steps of 40 while keeping $\mathsf{QL_g}$ constant. The resulting $\hat{QL}_t$ is shown in Fig. 4 (right), where we observe the decrease in $\hat{QL}_t$ with an increase in the number of drivers. The **above evaluations demonstrate that the independence (dependence) of generic (tailored) metrics from LBS parameters can affect evaluation of LPPMs**.

6.3 Impact on LPPM Comparison

The correct choice of LPPM for a given LBS can affect the utilty of its users; therefore, in this part we consider the effect of QL metrics on LPPM comparison. In this part, we compare $\mathsf{LPPM_e}$ and $\mathsf{LPPM_p}$ using $\hat{QL}_g$ and $\hat{QL}_t$ for the same PRide ENV. We run PRide with baseline ENV for these mechanisms and different values of ϵ. The motivation here is to understand *whether comparing performances of two LPPMs using generic and tailored metrics leads to the same conclusions.* We will show that there is a difference in conclusions when LPPMs' performances are quantified with the two metrics. Therefore, tailored QL metrics are required to compare and choose LPPMs for a particular LBS scenario.

Fig. 6 (left) shows the comparison of $\mathsf{LPPM_p}$ and $\mathsf{LPPM_e}$ for $\mathsf{ENV_b}$ with uniform $\pi(\mathsf{R})$ and $\pi(\mathsf{D})$, and Fig. 6 (middle) shows the comparison for nonuniform $\pi(\mathsf{R})$ and $\pi(\mathsf{D})$. For uniform π's, comparing the LPPMs using the two metrics points to unambiguous superiority of $\mathsf{LPPM_p}$, as $\mathsf{LPPM_p}$ performs better on both the metrics for all r's. However, for nonuniform π's, the two LPPMs perform similarly on $\hat{QL}_g$ as with uniform π's, but the performance of $\mathsf{LPPM_e}$ improves significantly when measured with $\hat{QL}_t$. The reason for this is, as follows. Note that $\mathsf{LPPM_e}$ obfuscates locations farther than $\mathsf{LPPM_p}$ (compare the lines with brown boxplots in the middle figure). High obfuscations due to $\mathsf{LPPM_e}$ increases the probability of finding drivers with the considered nonuniform distributions and improves the performance of $\mathsf{LPPM_e}$. This performance improvement on $\hat{QL}_t$ may allow the use of $\mathsf{LPPM_e}$ due to the additional advantage of its easy adaptivity to any distance metric over $\mathsf{LPPM_p}$ [5, 6].

Similarly, we compare $\mathsf{LPPM_p}$ with $\mathsf{LPPM_g}$ which are very close in performance when measured with $\hat{QL}_g$; the results for uniform

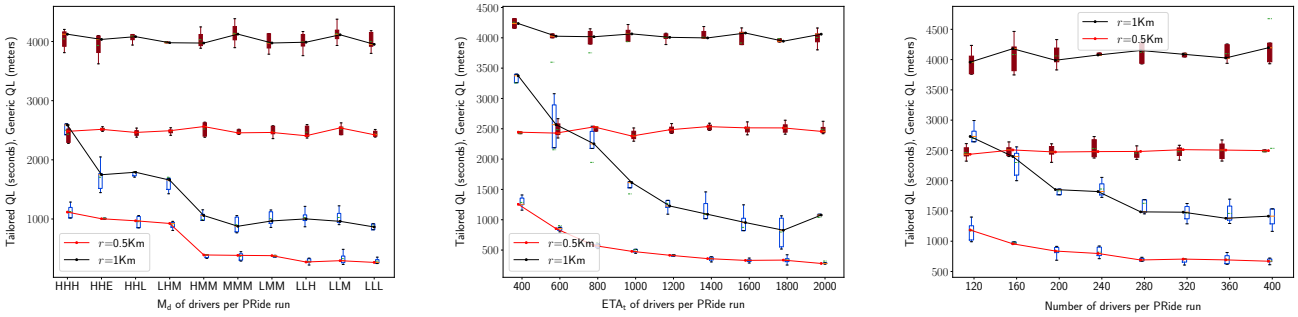

Figure 4: (left): As the drivers' acceptance model, M_d, becomes more tolerant, $\hat{QL}_t$ reduces but $\hat{QL}_g$ stays almost constant; this is consistent for different $\hat{QL}_g$'s. (middle): Similarly, with increase in ETA_t of drivers, $\hat{QL}_t$ reduces but $\hat{QL}_g$ does not, even for different $\hat{QL}_g$'s. (right): In PRide, with increase in number of drivers, $\hat{QL}_t$ decreases but $\hat{QL}_g$ does not, for different $\hat{QL}_g$. This confirms the dependence of $\hat{QL}_t$ on another important PRide parameter.

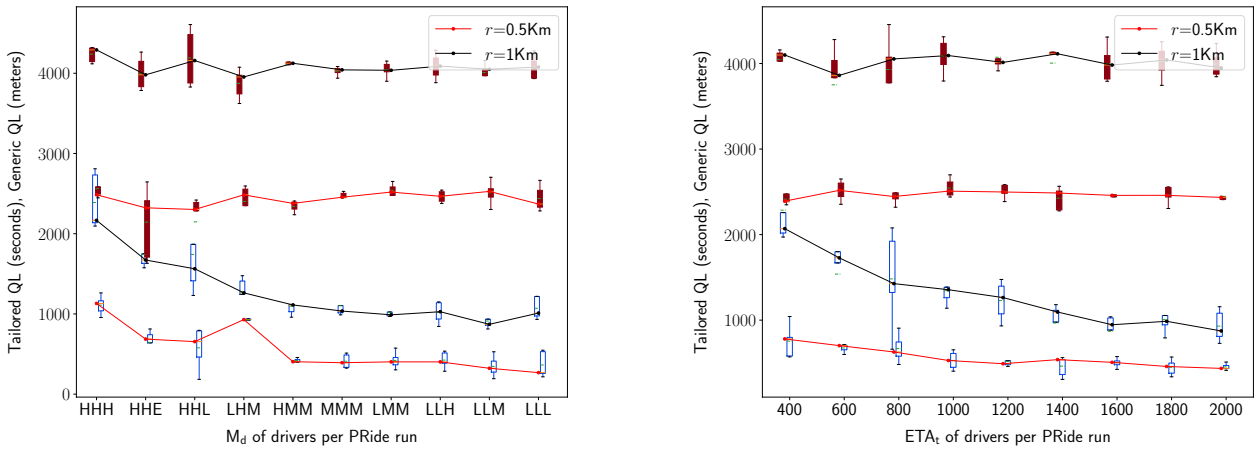

Figure 5: LPPMs' evaluations under a nonuniform distribution of drivers. (left): $\hat{QL}_t$ decreases with increase in the tolerance of drivers' acceptance models. (right): $\hat{QL}_t$ decreases with increase in ETA_t while $\hat{QL}_g$ does not.

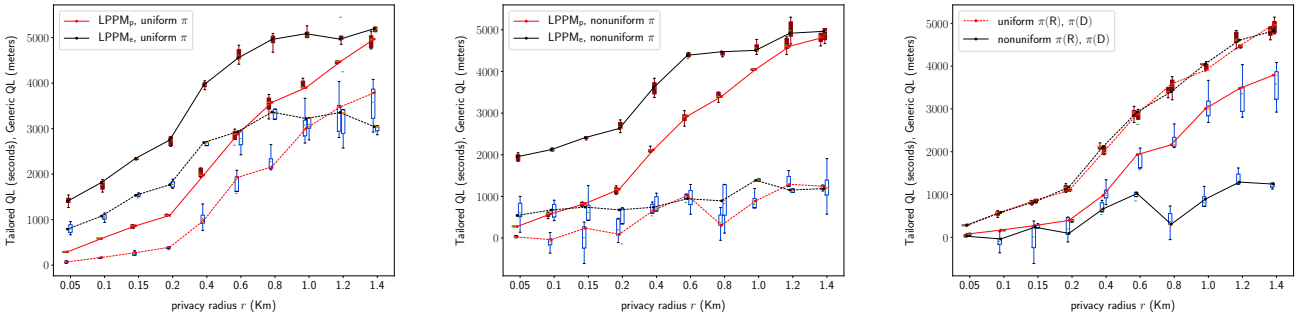

Figure 6: (left): Comparing $LPPM_p$ and $LPPM_e$ for ENV_b: For uniform π's, $\hat{QL}_g$ and $\hat{QL}_t$ both conclude $LPPM_p$ to be better. (right): Comparing the LPPMs for ENV_b with nonuniform π(D) and π(R): For some r ranges, $\hat{QL}_g$ prefers $LPPM_p$ but $\hat{QL}_t$ does not, e.g., $r = 0.4$Km, 0.6Km. Note also that, with nonuniform π's, $\hat{QL}_t$ of $LPPM_e$ is almost constant for the entire range of r considered, which is a strong evidence of the nonlinearity of $\hat{QL}_g$ and $\hat{QL}_t$.

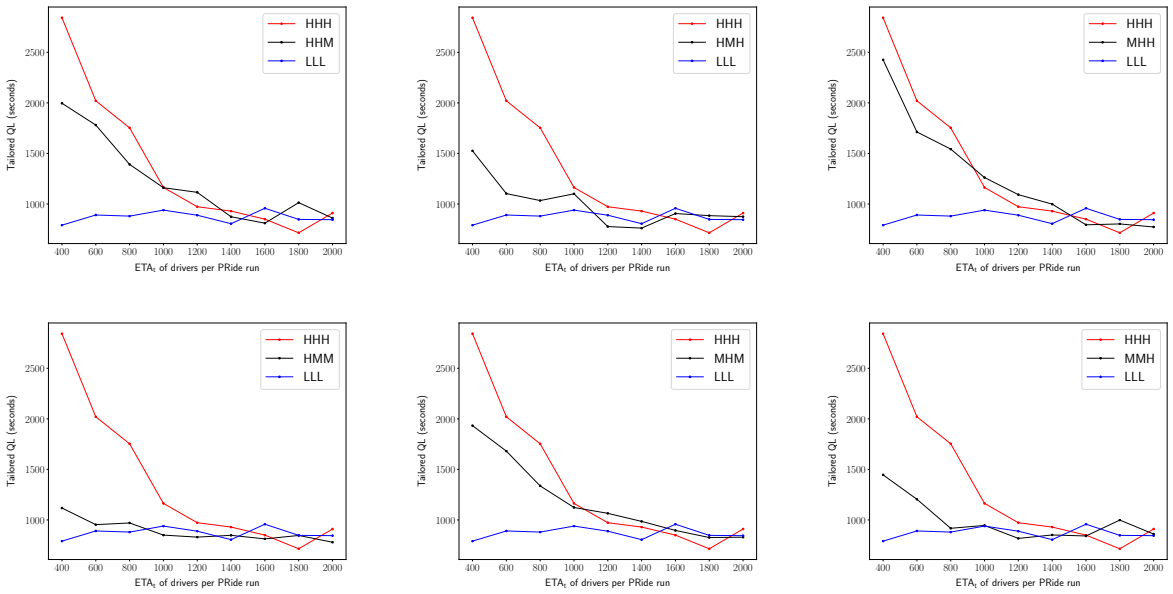

Figure 7: Effect of different ETA tolerances, ETA$_t$ (x-axis), for different driver models M$_d$ per PRide run. We experiment with six driver models while keeping r = 1Km, i.e., $\hat{QL}_g$ constant. For different strictnesses, $\hat{QL}_t$ (y-axis) of the driver models varies significantly and is shown by the black line. At high ETA$_t$ drivers, even the strictest ones with HHH model, accept all of the rides, therefore, M$_d$ will not contribute to $\hat{QL}_t$ and all models have almost the same $\hat{QL}_t$.

and nonuniform distributions for π(D) and π(R) are shown in Fig. 9. We find that, unlike in the uniform case, generic and tailored metrics *completely* disagree on the performance of the two mechanisms when π(D) and π(R) are nonuniform and grid resolution for LPPM$_g$ is 0.2Km. The reason for better the performance of LPPM$_p$ on $\hat{QL}_t$ is the same as explained above.

In Fig. 6 (right), we compare uniform and nonuniform distribution cases for LPPM$_p$. We note that with uniform π(R) and π(D), both $\hat{QL}_t$ and $\hat{QL}_g$ increase monotonically with the obfuscation radii. But, with nonuniform distributions the increase is not monotonic: $\hat{QL}_t$ remains constant in some intervals of obfuscation radii. This experiment demonstrates that there are *can* be RHS scenarios in which $\hat{QL}_g$ would change significantly with obfuscation while $\hat{QL}_t$ may not, and vice-versa. This emphasizes the need for the use of $\hat{QL}_t$ in comparisons of LPPMs.

Furthermore, in §6.4, we show how the use of generic versus tailored QL affects the choice of privacy budget ϵ, an important parameter LBS users tune according to their privacy needs. Specifically, we show that the use of tailored QL allows for an LBS-aware choice of ϵ: When ETA$_t$ of drives in a region is 400 versus 900 seconds, the use of $\hat{QL}_g$ always leads to the same ϵ, while $\hat{QL}_t$ adapts to the variations ETA$_t$ to give appropriate ϵ. Note that here we make a reasonable assumption that the tailored QL is a more appropriate QL metric than the generic QL metric.

To summarize, **the performance evaluation using LBS- tailored metrics can disagree completely with that using generic metrics and mislead the choice of an LPPM for a particular LBS.** This disagreement may not be significant in all scenarios.

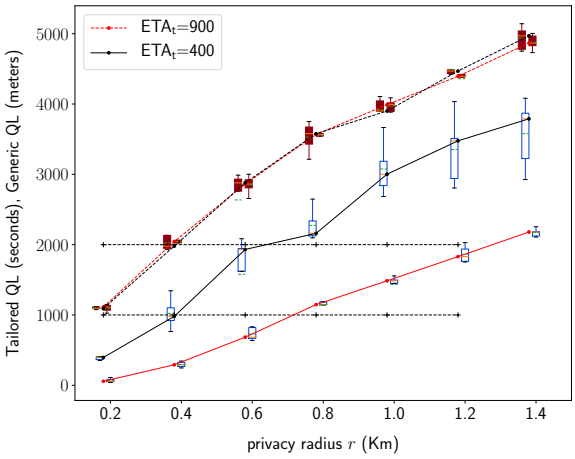

Figure 8: LPPM$_p$ with different ETA$_t$ for different obfuscation radii r, hence, also for different ϵ. $\hat{QL}_g$ in both ETA$_t$ cases remains the same while $\hat{QL}_t$ reduces with ETA$_t$. This allows to choose desired ENV-tailored ϵ without violating QL requirements.

However, it is important to find corner cases and tailor LPPMs to the LBS for reliable designs. Apart from designing, **while choosing from the existing LPPMs, generic metrics may discard a**

more suitable LPPM due to its independence from the LBS parameters. For instance, in Pride, an LPPM that considers players' distribution and employs LPPM_p for uniform while LPPM_e for nonuniform distribution is better than any of the two LPPMs separately.

6.4 Implications to choice of ϵ:

In geo-indistinguishable mechanisms, privacy budget ϵ is an important parameter, because it directly affects LPPM outputs and also specifies generic QLes for some LPPMs. Hence, the choice of ϵ is an important aspect of LPPM design. In this section, we evaluate the effect of the use of $\hat{QL}_g$ vs $\hat{QL}_t$ on the choice of ϵ. Two commonly used QLes for a given metric are average loss, $\hat{QL}$ [1, 5], and worse case loss, QL^+ [1, 20, 28]. While optimizing mechanisms for the best privacy, QL^+ is commonly used as a constraint [1, 20, 28]. However, it is hard to incorporate QL^+ in the design process of geo-indistinguishability mechanisms [20]. Instead, we solve the objective in (7) to find the least ϵ that upper bounds $\hat{QL}$ by QL^+.

$$\epsilon^* = \underset{\epsilon \in \mathbb{R}^+}{\text{argmin}} \ \hat{QL}(\text{LPPM}, \pi, d_Q) \quad \text{s.t. } \hat{QL} < QL^+ \tag{7}$$

For LPPM_p, this objective can be easily solved, because $\hat{QL}_g$ has a closed form expression and is given by $\frac{2}{\epsilon}$. While we do not have such a closed form expression for tailored QL, $\hat{QL}_t$, varying PRide parameters (§ 6.2), we can find the minimum ϵ. To understand this, consider $\text{QL}_\text{t}^\text{max}$ and $\text{QL}_\text{g}^\text{max}$ to be the maximum tolerable tailored and generic QLes, respectively. We execute PRide with two different ETA_t of drivers, namely 400 and 900, for ϵ ranging from 0.2 to 1.4. The corresponding results are plotted in Fig. 8. Here due to truncation at low ϵ (high r), $\hat{QL}_g$ does not scale linearly with ϵ. We set $\text{QL}_\text{t}^\text{max}$ to 1000 seconds and $\text{QL}_\text{g}^\text{max}$ to 2Km as show in Fig 8 by horizontal lines.

Using $\hat{QL}_g$ and $\text{QL}_\text{g}^\text{max}$, the optimal ϵ is approximately $\frac{\ln(1.4)}{0.4}$ while that using $\hat{QL}_t$ and $\text{QL}_\text{t}^\text{max}$, the optimal ϵ is $\frac{\ln(1.4)}{0.75}$ or $\frac{\ln(1.4)}{0.4}$ *depending on the value of* ETA_t. Unlike $\text{QL}_\text{t}^\text{max}$, solving (7) using $\text{QL}_\text{g}^\text{max}$ ignores the variations in LBS parameters, ETA_t in this case. Note that, for the nonuniform $\pi(\text{D})$ case as well, the result will be similar as in Fig. 8 for different ETA_t. Hence, the choice of ϵ will be affected undesirably in the nonuniform $\pi(\text{D})$ case as well. Therefore, **using a tailored metric and considering LBS parameters, LBS-specific privacy budget can be chosen, but generic metrics will always choose the same optimal ϵ due to the independence from LBS parameters**.

6.5 Implications on Utility Improvement Techniques

Utility improvement techniques are commonly used along with LPPMs due to the degradation of utility that LPPMs cause. Therefore, we study the impact of the use of QL metrics on the outputs of such techniques. In this part, we consider ENV_b with nonuniform $\pi(\text{D})$ and call the area with high density of drivers the *special zone*; check §6.1 for the details of a nonuniform distribution. We introduce a *greedy remapping* strategy where the LPPM remaps each l_o to the zone via truncation, we call it l_o^t. We run PRide with this setting for different obfuscation radii r's and plot the results in Fig. 10. Note that ϵ and r are inversely proportional.

From Fig. 10, it can be seen that *the greedy remap improves upon $\hat{QL}_t$ for all ϵ values, but, does not always improve $\hat{QL}_g$*. The reasons for these are: At low r, $\hat{QL}_g$ is low, hence both l_r, l_o will be out of the special zone with high probability. But l_o^t will be inside the zone, hence, will have 9 fold more chances of finding a driver. Note that, due to this remap, l_o^t will be farther away, i.e., $\text{QL}_\text{g}(l_r, l_o) < \text{QL}_\text{g}(l_r, l_o^t)$; this can be seen in Fig. 10 where QL_g increases for low r values. On the other hand, $\hat{QL}_g$ is high at high r, hence, *if l_r is in the special zone, l_o will be out of the zone with high probability* and in this case remapping will again increase chances of finding a driver. But, with this remap, l_o^t will be closer than l_o, i.e., $\text{QL}_\text{g}(l_r, l_o) > \text{QL}_\text{g}(l_r, l_o^t)$; this can be seen in Fig. 10 where QL_g decreases for high r values.

The greedy remapping experiment implies that considering LBS parameters can improve $\hat{QL}_t$ without improving $\hat{QL}_g$; in Fig. 10, $\hat{QL}_g$ with remapping increases while $\hat{QL}_t$ decreases. More importantly, for some ENVs, **remapping to improve $\hat{QL}_g$ may not improve $\hat{QL}_t$**. This can be understood as: at low r, such remap will reduce $\hat{QL}_g$ for l_r's out of the special zone and increase the distance of l_o^t from the special zone, hence reduce the chances of finding drivers and will increase $\hat{QL}_t$. Therefore, **remapping using a generic metric will not necessarily improve tailored QLes and vice versa**. This emphasizes the need to consider LBS parameters to devise QL metrics for the utility improvement techniques to be effective.

6.6 Effect of Multiple QLes on Remapping in LPPMs

In this part, we investigate how multiple viable QLes, called QL-dimensions, associated with LBSes and different preferences of users towards them affect remapping when these QLes are considered separately and in combination. Specifically, we investigate: 1) effect of tailored vs generic metrics on remapping proposed in [5], 2) effect of combination of QLes, considering user preferences for each QL-dimension. We define the notion of *comprehensive QL*, $\widetilde{QL}$, that combines different QL-dimensions according to preferences input by users. Remapping [5] requires a closed form expression for the QL metric used which is not available for RHS LBSes. Therefore, for demonstration, we consider POI search LBS with two QLes: Euclidean distance as generic QL, $\hat{QL}_g$, and count of POIs returned by the LBS as a tailored QL, $\hat{QL}_t$. $\hat{QL}_g$ represents the average distance the user needs to travel to the finally selected POI while $\hat{QL}_t$ represents the number of choices the user has to choose the final POI from. Due to space constraints, we defer the details of the framework and its application to POI search LBS to Appendix A and only discuss representative results here.

Key insights: 1) Remapping using one QL metric need not generalize to other viable QL metrics. Fig. 11 shows $\hat{QL}_g$ for LPPM outputs —without remap, with remap using $\hat{QL}_g$, with remap using $\hat{QL}_t$, and with remap using $\widetilde{QL}$. It can be seen that if remapping is performed using only the Euclidean distance as the QL

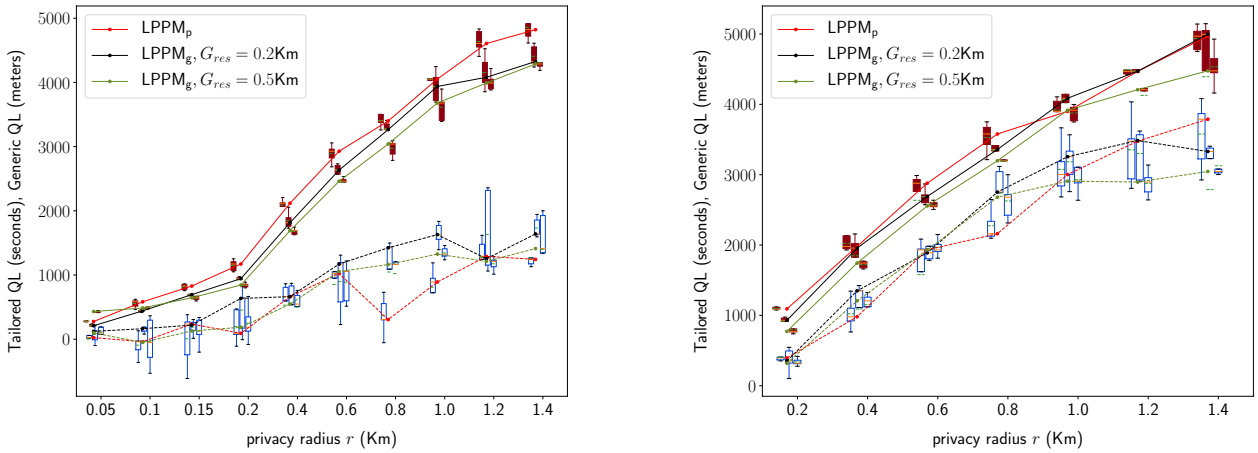

Figure 9: Comparison of LPPM$_p$ and LPPM$_g$ with uniform (left) and nonuniform (right) π(D) and π(R). Unlike in the uniform distribution case, $\hat{QL}_g$ and $\hat{QL}_t$ completely disagree on the performance of the two LPPMs when G_{res} is 0.2Km for LPPM$_g$ when the distributions π(D) and π(R) are nonuniform.

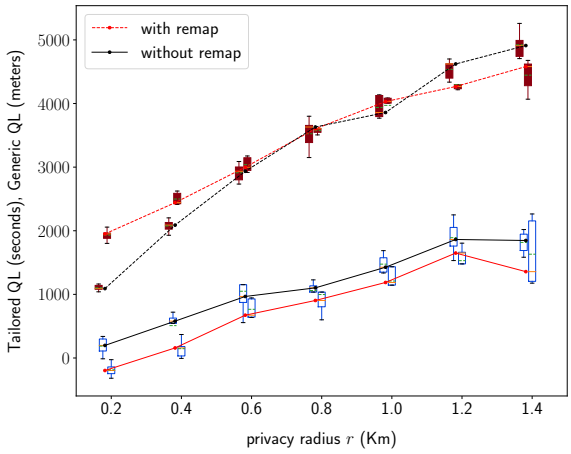

Figure 10: PRide with nonuniform π(D) with and without greedy remapping. Greedy remap helps low r by increasing obfuscation distance when l_r is not in special zone and helps high r by reducing obfuscation distance when l_r is inside the special zone.

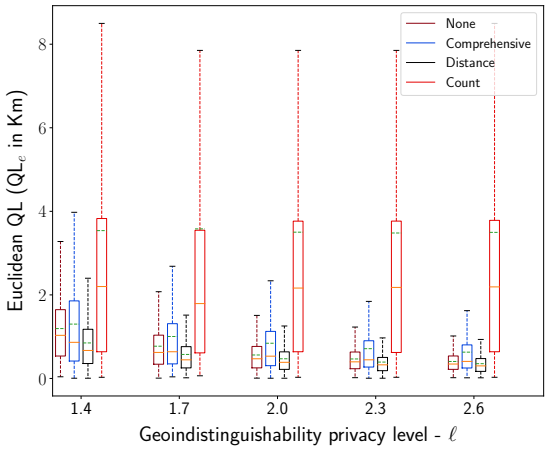

Figure 11: Comparing QL$_e$, i.e., $\hat{QL}_g$ of three remaps: locations using $\hat{QL}_g$-remap are closest while that using $\hat{QL}_g$-remap (red) are farther from l_r; locations remapped using comprehensive QL (blue) are balanced according to user preference. l_r here is at 0 of the y-axis.

metric, the remapped locations have low POI count (Fig. 13). On the other hand, if remapping is performed using only the POI-count as QL, the remapped locations are far from real locations (Fig. 11). This restates (§ 6.5) the importance of tailored QLes for remapping based utility improvements. **2) All the possible QL-dimensions associated with an LBS combined according to user preferences for each dimension is a more effective (Fig. 14) and user centric way to employ remapping.** We show that, with change in preferences of QL-dimensions, remapping outputs change significantly as shown in Fig. 12. Note that as the preference for $\hat{QL}_g$ dimension increases along the x-axis, locations remapped using $\widetilde{QL}$ shift towards that using $\hat{QL}_g$. Further details of the experiments are presented in Appendix A.

7 CONCLUSION

We challenge the community's common use of a generic distance-based utility loss metric and argue for the need of LBS-tailored utility loss metrics for the LPPM design and evaluation. Motivated by the lack of real-world data, we build an extensible RHS emulator

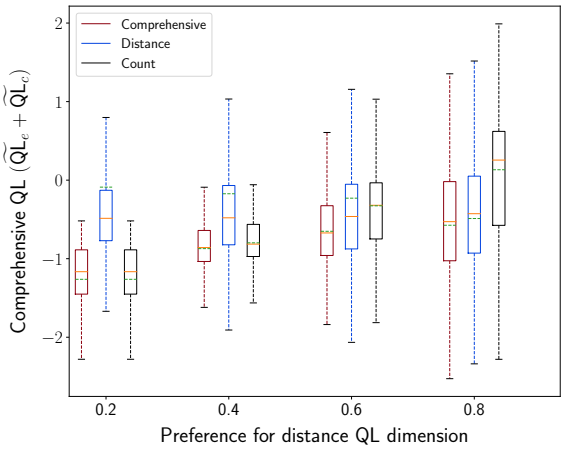

Figure 12: Comparing $\widetilde{QL}$ for different preferences for QL_e, p_e: With p_e, difference between remap using $\widetilde{QL}$ and that using QL_e reduces because the earlier remap considers the user preferences.

for RHS data synthesis. We thoroughly evaluate the established LPPMs using the data synthesized by the emulator for the specific privacy preserving instance of RHS that we define We demonstrate the implications of using generic versus tailored utility loss metrics on different aspects of LPPM design process namely, choice of parameters, comparison, design of utility improvement techniques and, evaluation. We also demonstrate the need to consider user-centric combination of utility metrics while employing the state-of-the-art utility improvement techniques. Our work brings to notice the inadequacy of the generic metrics and its effects on the LPPM designs.

8 ACKNOWLEDGEMENT

The work was supported by the NSF grant CPS-1739462.

REFERENCES

[1] 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*. ACM, 901–914.

[2] Igor Bilogrevic, Kévin Huguenin, Stefan Mihaila, Reza Shokri, and Jean-Pierre Hubaux. 2015. Predicting users' motivations behind location check-ins and utility implications of privacy protection mechanisms. In *22nd Network and Distributed System Security Symposium (NDSS)*.

[3] Vincent Bindschaedler, Reza Shokri, and Carl A Gunter. 2017. Plausible deniability for privacy-preserving data synthesis. *Proceedings of the VLDB Endowment* 10, 5 (2017), 481–492.

[4] H Campbell. 2017. *WhatâĂŹs the Farthest You Should Drive To Pick Up A Passenger?* https://maximumridesharingprofits.com/whats-the-furthest-you-should-drive-to-pick-up-a-passenger/.

[5] Konstantinos Chatzikokolakis, Ehab Elsalamouny, and Catuscia Palamidessi. 2017. Efficient utility improvement for location privacy. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 308–328.

[6] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Marco Stronati. 2015. Constructing elastic distinguishability metrics for location privacy. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 156–170.

[7] E Cho, SA Myers, and J Leskovec. [n.d.]. Friendship and mobility: Friendship and mobility: User movement in location-based social networks. *Proc. ACM SIGKDD 2011* ([n. d.]).

[8] Ridesharing Driver. 2018. *Fired from Uber: Why drivers get deactivated, and how to get reactivated.* https://www.ridesharingdriver.com/fired-uber-drivers-get-deactivated-and-reactivated/.

[9] Cynthia Dwork, Aaron Roth, et al. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science* 9, 3–4 (2014), 211–407.

[10] Wisam Eltarjaman, Rinku Dewri, and Ramakrishna Thurimella. 2017. Location privacy for rank-based geo-query systems. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 77–96.

[11] Kassem Fawaz, Huan Feng, and Kang G Shin. 2015. Anatomization and protection of mobile appsâĂŹ location privacy threats. In *24th {USENIX} Security Symposium ({USENIX} Security 15)*. 753–768.

[12] Kassem Fawaz and Kang G Shin. 2014. Location privacy protection for smartphone users. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 239–250.

[13] Marco Gruteser and Dirk Grunwald. 2003. Anonymous usage of location-based services through spatial and temporal cloaking. In *Proceedings of the 1st international conference on Mobile systems, applications and services*. ACM, 31–42.

[14] FitBit Help. 2019. *How does Fitbit estimate how many calories I've burned?* https://help.fitbit.com/articles/en_US/Help_article/1381.

[15] Uber Help. 2018. *What are acceptance rates?* https://help.uber.com/h/b6da86a4-2938-497c-a4fd-fd6f386aeefa.

[16] Business Insider. 2015. *Uber's internal charts show how its driver-rating system actually works.* https://www.businessinsider.com/leaked-charts-show-how-ubers-driver-rating-system-works-2015-2?r=UK&IR=T.

[17] Kristopher Micinski, Philip Phelps, and Jeffrey S Foster. 2013. An empirical study of location truncation on android. *Weather* 2 (2013), 21.

[18] Zarrin Montazeri, Amir Houmansadr, and Hossein Pishro-Nik. 2017. Achieving perfect location privacy in wireless devices using anonymization. *IEEE Transactions on Information Forensics and Security* 12, 11 (2017), 2683–2698.

[19] Uber Movement. 2017. *Uber Movements.* https://movement.uber.com/cities?lang=en-US.

[20] Simon Oya, Carmela Troncoso, and Fernando Pérez-González. 2017. Back to the drawing board: Revisiting the design of optimal location privacy-preserving mechanisms. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 1959–1972.

[21] Simon Oya, Carmela Troncoso, and Fernando Pérez-González. 2017. Is Geo-Indistinguishability What You Are Looking for?. In *Proceedings of the 2017 on Workshop on Privacy in the Electronic Society*. ACM, 137–140.

[22] Anh Pham, Italo Dacosta, Guillaume Endignoux, Juan Ramon Troncoso Pastoriza, Kévin Huguenin, and Jean-Pierre Hubaux. 2017. Oride: A privacy-preserving yet accountable ride-hailing service. In *26th {USENIX} Security Symposium ({USENIX} Security 17)*. 1235–1252.

[23] Anh Pham, Italo Dacosta, Bastien Jacot-Guillarmod, Kévin Huguenin, Taha Hajar, Florian Tramèr, Virgil Gligor, and Jean-Pierre Hubaux. 2017. Privateride: A privacy-enhanced ride-hailing service. *Proceedings on Privacy Enhancing Technologies* 2017, 2 (2017), 38–56.

[24] OSRM Project. 2019. *Table Service API.* http://project-osrm.org.

[25] N Scheiber. 2017. *How Uber Uses Psychological Tricks to Push Its DriversâĂŹ Buttons.* https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html.

[26] Reza Shokri. 2015. Privacy games: Optimal user-centric data obfuscation. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 299–315.

[27] Reza Shokri, George Theodorakopoulos, Jean-Yves Le Boudec, and Jean-Pierre Hubaux. 2011. Quantifying location privacy. In *2011 IEEE symposium on security and privacy*. IEEE, 247–262.

[28] Reza Shokri, George Theodorakopoulos, Carmela Troncoso, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. 2012. Protecting location privacy: optimal strategy against localization attacks. In *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 617–627.

[29] Wikipedia. 2018. *Haversine Formula.* https://en.wikipedia.org/wiki/Haversine_formula.

A COMPREHENSIVE QL FRAMEWORK

Here, we detail the framework for comprehensive QL, $\widetilde{QL}$, and demonstrate its effects on remapping for POI search application. The reason we use POI search LBS is the availability of a closed form QL expression required for the remapping proposed in [5].

A.1 Comprehensive QL

In LBSes, QLes incurred and their preferences are highly user dependent; therefore, we consider an LBS with D QL dimensions,

$\{\text{QL}_\text{i}\}_{i\in D}$ and each user has preference p_i and threshold QL_i^{th} for each QL_i. Threshold values for QLes are introduced in previous works [20, 28]; we only use them here along with p_i to formalize comprehensive QL, $\widetilde{QL}$. The significance of p_i and QL_i^{th} specific to POI search LBS is described in § A.2. We assume l_r is obfuscated to l_o using LPPM, $\text{LPPM} : \mathcal{X} \rightarrow \mathcal{X}$. Then, given the actual QL value for a QL-dimension, $\text{QL}_\text{i}(l_r, l_o)$, or simply QL_i, and the preference tuple, (p_i, QL_i^{th}), user's *perceived QL* can be defined as: $f_i(\text{QL}_\text{i}, p_i, \text{QL}_i^{th})$. Finally, total $\widetilde{QL}$ incurred can be written as in (8), and the remapping objective in [5] can be tweaked to include $\widetilde{QL}$ as in (9).

$$\widetilde{QL}(l_r, l_o) = \sum_{i\in D} \sum_{l_r, l_o \in \mathcal{X}} f_i(\text{QL}_i, p_i, \text{QL}_i^{th}) \tag{8}$$

$$\mathbf{R}(l_o) = \underset{l_o^* \in \mathcal{X}}{\text{argmin}} \sum_{l_x \in \mathcal{X}} \sigma_{l_r|l_o} \widetilde{QL}(l_r, l_o) \tag{9}$$

A.2 Example scenario: POI search LBS

In this section, we define $\widetilde{QL}$ for a POI search application. Euclidean distance as a generic QL, $\hat{QL}_g$ (denoted as QL_e in rest of this section), and count of POIs returned by the LBS as a tailored QL, $\hat{QL}_t$ (denoted as QL_c in rest of this section). QL_e captures the extra distance one needs to travel, on average, to the final POI while, QL_c captures the difference in the number of POIs returned. QL_e is the conventional Euclidean distance between l_r and l_o while $\text{QL}_\text{c} = \frac{\text{COUNT}(l_o, s)}{\text{COUNT}(l_r, s)}$. Here, $\text{COUNT}(l_r, s)$ gives the count of POIs around location l_r within search radius of s. Corresponding to the two dimensions, preference tuples are denoted as $\{(p_e, \text{QL}_e^{th}), (p_c, \text{QL}_c^{th})\}$ with $p_c = 1 - p_e$. For QL_e, QL_e^{th} signifies the maximum tolerable distance from the real location l_r, while for QL_c, QL_c^{th} signifies[4] the minimum ratio of number of POIs at l_o to that at l_r i.e. $(\frac{\text{COUNT}(l_o, s)}{\text{COUNT}(l_r, s)})_{\text{min}}$. The corresponding perceived QLes, $\widetilde{\text{QL}}_e$ and $\widetilde{\text{QL}}_c$, and remapping objective are:

$$\widetilde{\text{QL}}_e = p_e \times \ln\Big(\frac{\text{QL}_\text{e}(l_r, l_o)}{\text{QL}_e^{max}}\Big) \quad \widetilde{\text{QL}}_c = p_c \times \ln\Big(\frac{\text{QL}_c^{th}}{\text{QL}_\text{c}(l_r, l_o)}\Big) \tag{10}$$

$$\mathbf{R}(l_o) = \underset{l_o^* \in \mathcal{X}}{\text{argmin}} \sum_{l_o \in \mathcal{X}} \sigma_{l_r|l_o} (\widetilde{\text{QL}}_e + \widetilde{\text{QL}}_c) \tag{11}$$

The logarithmic formulation in (10) is to ensure negative QL if results returned are better than expected: When a user finds POIs within $r < \mathbf{d}_e^{max}$, $\widetilde{\text{QL}}_e$ is negative. Similarly, $\widetilde{\text{QL}}_c$ is negative if the expected ratio of POI count at l_o versus l_r is more than $\mathbf{d}_c^{th}$ i.e. if there are more POIs around the final obfuscated location than what the user expected.

A.2.1 Experimental setup. For experiments, we use the Gowalla dataset [7] which contains 6,442,890 check-ins of 196,591 users. We implement the COUNT, with search radius $s = 400m$, by performing geolocal queries to OpenStreetMap database. We consider the two QL dimensions (10), separately, as baselines and compare performances of three remaps: remap using $\widetilde{\text{QL}}_e$, R_e, remap using $\widetilde{\text{QL}}_c$, R_c, and remap using comprehensive QL, $\tilde{R}$. We fix, the obfuscation radius, r, at 0.2Km, $p_e = 1$ for R_e, $p_c = 1$ for R_c, $p_e = p_c = 0.5$ for $\tilde{R}$, $\mathbf{d}_e^{th}$ is 1000m and $\mathbf{d}_c^{th}$, is 0.7. For remapping, we use the exact same setup as in [5] for data splitting and global prior generation.

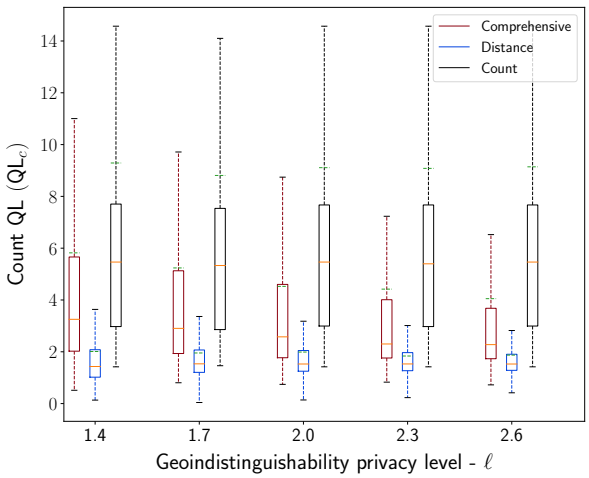

Figure 13: Comparing QL_c of three remaps. QL_c is a ratio without any unit. POI count at R_e-locations is lowest while POI count at R_c-locations are highest with $\tilde{R}$ striking a balance. Note that, QL_c denotes ratio of counts of POIs at l_o and at l_r therefore, high QL_c values are desired.

A.2.2 Results and discussion. The optimal remapped locations in Fig. 11, 13 and 14 are the same for respective remaps but, Fig. 11 and 13 show QL_e and QL_c, respectively while Fig. 14 shows $(\widetilde{\text{QL}}_e + \widetilde{\text{QL}}_c)$[5]. In Fig. 11 the four boxes per privacy level denote QL_e of locations for four cases: no remap, $\tilde{R}$, R_e and R_c; Fig. 13 and 14, show the same for the count and comprehensive QLes, respectively.

Fig 11 shows that locations remapped using just QL_c dimension (red boxes), perform poorly along QL_e dimension for all privacy levels, ℓ. For $\ell = 1.4$, the median of the distance between l_r and remapped locations is 2.4km for R_c, 0.6km for R_e, 1km for no-remap and 0.8km for $\tilde{R}$. Similarly, Fig 13 shows that if remapping is performed using just QL_e dimension (blue boxes), the corresponding locations perform poorly along QL_c dimension for all ℓs. For $\ell = 1.4$, the ratio of the number of POIs at the remapped location and at l_r is 2.32 for $\tilde{R}$, 1.28 for R_e and 4.23 for R_c. In both cases, the lowest value of the QL dimension is for the remap that uses the corresponding QL, but that remapped location's QL value along the other dimension is undesirably high. However, $\tilde{R}$ strikes a balance between two QL dimensions and remaps to locations that are optimal for given user preferences. Fig. 14 shows the sum of two QL dimension values; we note that $\tilde{R}$ finds locations with the lowest comprehensive QL and therefore, improves utility of LPPM in a user-centric way. Finally, we compare the three remaps for different user preferences for the two dimensions. Fig 15 demonstrates that

[4] the significance of these thresholds are for demonstration and can change as required

[5] Note Fig 11 and 13 show actual QL values while Fig 14 shows perceived QL values

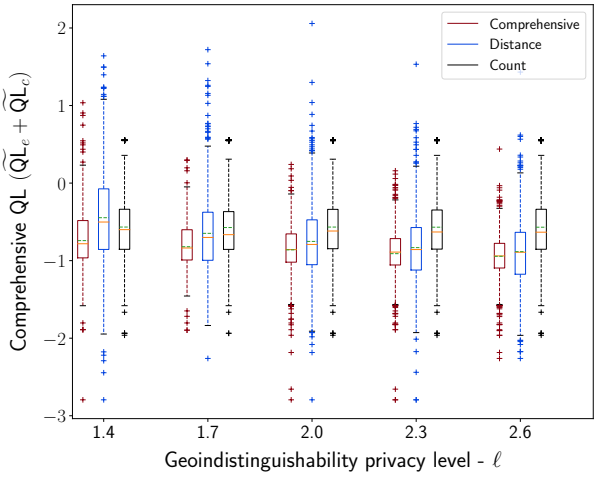

Figure 14: Comparing $\widetilde{QL}$ for R_e, R_e and $\tilde{R}$. When user preferences are considered, $\tilde{R}$ outperforms other remaps.

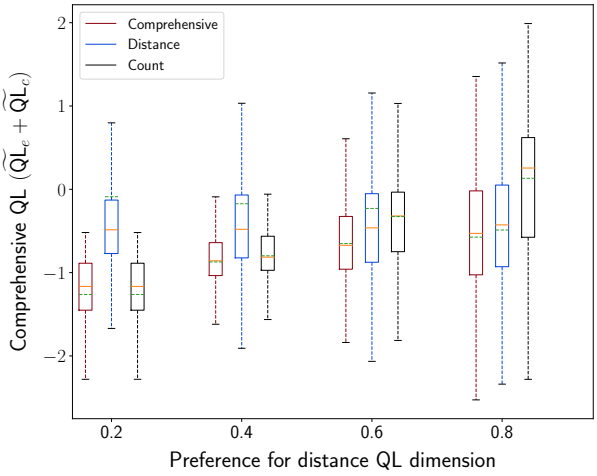

Figure 15: Comparing $\widetilde{QL}$ for different p_e values: With p_e, difference between $\tilde{R}$ and R_e reduces as $\tilde{R}$ considers user preferences.

increasing preference for QL_e from 0.2 to 0.8 decreases the difference between both means and medians of $\tilde{R}$ and R_e while increases that for the R_c. Note that, the increased variance of comprehensive QL values of R_c in Fig 15 is due to the uncontrolled nature of the experiment, the outcomes of which can depend on the density of POIs in the area around the locations involved.

Results from § 6 and § 6.6 emphasize that un-tailored and incomplete QLes can lead to an incorrect perception of the privacy-utility trade-off. That is, though the privacy guarantees of an LPPM are as expected, QL_g claims utility that is not necessarily guaranteed. Further, if provable QL improvement techniques such as remapping are used, the locations with optimal QL vary significantly with the QL metric used (Fig 11). Finally, using the comprehensive QL for the remapping can greatly improve QL of LBS in a user-centric manner and so the user experience (Fig 14,15). We note that devising tailored and robust QL metrics for complex LBSes like RHS is not a trivial task. However, using advanced machine learning techniques on data synthesized using emulators such as RHSE can lead to QL models that are more representative of perceived QL incurred in the wild; we leave this to the future work.