VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

Project report of

A Privacy Preserving Online Raid Hailing System

Submitted to

Prof. Dr. Deebak B D

Submitted by

Kiran Galla (19BCE2583)
Alokam Nikhitha (19BCE2555)
Ayush Tiwary (19BCE2049)

in partial fulfilment for the award of the degree of

Bachelor of Technology
in
Computer Science & Engineering

School of Computer Science and Engineering

Table of Contents

1. Objective

The growing popularity of online ride-hailing (ORH) services has made our daily travel much easier. It enables a rider to quickly and easily request the nearest driver via mobile devices. Existing ORH systems, on the other hand, necessitate the collection of users' location data, raising serious privacy concerns. Although several ORH service solutions have been proposed for privacy protection, most of the existing systems are based on a third trusted party to compute the distance between a rider and a driver. For practical deployment, such a security assumption cannot fully address privacy concerns. We present a new ride-matching scheme for ORH systems that allows for privacy-preserving and effective distance calculation without the use of a third-party server in this paper.

2. Domain Introduction

Many ORH services have started which help the riders to easily book a ride on their phone. It was great until people started to realize the privacy concerns, It is then that the use of a third trusted server came into play but this couldn’t fully address the privacy concerns. So techniques like Road Network Embedding (RNE) are used for distance calculation and this is uniquely bridged with cryptographic primitives like Property-preserving Hash (PPH) to safeguard the riders privacy.

3. Background and Motivation

Our proposed scheme allows ORH systems to securely compute the user distance while protecting both riders' and drivers' location privacy. In particular, we employ cutting-edge distance calculation techniques based on Road Network Embedding (RNE) and demonstrate how to uniquely bridge cryptographic primitives such as Property-preserving Hash (PPH) with RNE in depth to support privacy-preserving ride-matching services. Furthermore, we propose an optimised design to improve matching efficiency. We conduct a formal analysis of the security strengths and put the system prototype into action. The evaluation results show that our design is safe and efficient for ORH systems.

4. Literature Review

S.No:	Title	Citation	Abstract
1.	Quantifying the Tradeoff Between Cyber-security and Location Privacy	Dajiang Suo, M. Elena Renda, and Jinhua Zhao	Increasing location privacy in Location Based Services could protect users' data in case of breaches, it could also lead to security issues for users. In this paper they examined the impact of location data privacy-preservation on the performance of the anomaly detectors. A Density-based Spatial Clustering of Applications with Noise (DBSCAN) and a Recurrent Neural Network (RNN) framework are used to match the trip and two dimensional Laplace noise is used to preserve the location privacy of the users. The investigation results clearly show that while the level of privacy increases, by increasing the perturbation noise, the capacity for the system to identify anomalies could decrease quite sensibly, especially when using the clustering.
2.	pShare: Privacy-Preserving Ride-Sharing System	Junxin Huang , Yuchuan Luo , Ming Xu, Bowen Hu and Jian Long	They primarily investigate the privacy and utility of the ride-sharing system, which allows multiple riders to share one driver,

	with Minimum-Detouring Route		in this paper. They proposed pShare, a privacy-preserving ride-sharing system, to solve the privacy problem and reduce ride-sharing detouring waste. They used a zone-based travel time estimation approach to privately compute over sensitive data while cloaking each rider's location in a zone area to hide users' precise locations from the service provider. To compute the matching results, as well as the least-detouring route, the service provider first computes the shortest path for each eligible rider combination, then compares the additional travelling time (ATT) of all combinations, and finally chooses the combination with the lowest ATT.
3.	SRide: A Privacy-Preserving Ridesharing System	Ulrich Matchi Aïvodji,Kévin Huguenin , Marie-José Huguet, Marc-Olivier Killijian	Modern ridesharing systems, which have been enhanced with location-based features, have improved user experience by allowing drivers and riders to plan trips in near real time. The fine-grained nature of location data collected by service providers and exchanged between users, on the other hand, raises privacy concerns that could stymie the adoption of such systems. They presented

			SRide: a privacy-preserving ridesharing protocol that addresses the matching problem for dynamic ridesharing systems in this paper. They designed and build an SRide prototype that works in four steps. Firstly, it generalises user spatiotemporal data. Then, to compute feasible matches, it employs a privacy preserving protocol. . Then, for each feasible pair, it computes a ridesharing score using an improved version of Priv-2SP-SP, a privacy-preserving protocol for computing meeting points for ridesharing. Finally, based on their ridesharing scores, it computes the optimal assignment of drivers and riders. To demonstrate the proposed scheme's practical feasibility, we conduct an experimental trace-driven evaluation.
4.	HERMES: Scalable, Secure, and Privacy-Enhancing Vehicular Sharing-Access System	Iraklis Symeonidis , Dragos Rotaru, Mustafa A. Mustafa , Bart Mennink, Bart Preneel, and Panos Papadimitratos	In this paper, They propose HERMES, a scalable, secure, and privacy-enhancing vehicle sharing and access system. HERMES securely outsources Vehicle Access Token (AT) generation operations to a group of untrusted servers. It extends the system design of an earlier proposal, SePCAR, for improved

			efficiency and scalability. HERMES efficiently employs and combines several cryptographic primitives with secure multiparty computation (MPC) to meet system and user needs for secure and private computations. It hides vehicle secret keys and transaction details from servers, such as vehicle booking details, AT information, and user and vehicle identities. It also ensures user accountability in the event of a dispute. Furthermore, they provide semantic security analysis and demonstrate that HERMES satisfies its security and privacy requirements.
5.	Location privacy-preserving in online taxi-hailing services	Xiaoying Shen1 · Licheng Wang1 · Qingqi Pei2 · Yuan Liu1 · Miaomiao Li	Because of its convenience and low cost, online cab hailing has become the most common means of transportation. However, because online taxi-hailing service providers can track their specific movement trajectories, it creates a privacy risk to consumers and drivers. Furthermore, with the existing online taxi-hailing system, there is a time delay between the time a passenger makes a request and the time the driver arrives at the passenger's boarding point.

			They provided a new and efficient location privacy protection approach based on the MinHash algorithm to address these two issues (LPPM). The LPPM converts the precise positions of passengers and drivers into a set of points of interest surrounding them, and the distance between them into the similarity between the two sets. Using the MinHash technique, a service provider can efficiently match passengers and drivers without giving their particular location information. To overcome the second obstacle, they deploy mobile edge computing technologies in an online taxi-hailing system in this work. It can speed up data processing, allow drivers to plan ahead of time, and lessen the likelihood of traffic gridlock. LPPM has a high level of security, according to the security study, and the final experimental findings indicated that LPPM is effective.

5. Techniques & its Related Challenges

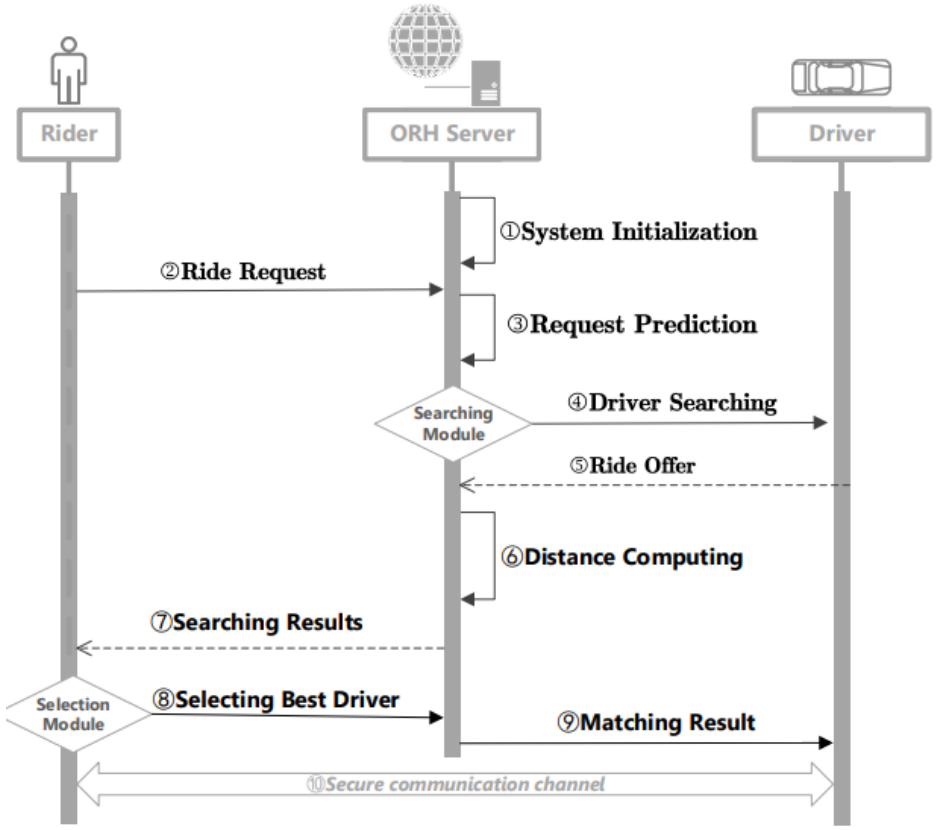
The main issue with the ORH services is regarding privacy of both Rider and Driver. Techniques like Homomorphic Encryption can perform computations on the data which is encrypted.But this technique can only give the Euclidean distance between the rider and driver which is not viable for ORH in road networks as vehicles are obliged to go along the streets.where in Property-Preserving Hash (PPH) can calculate the distance by using Road Network Embedding which estimate the distance with Road networks and do the ride-matching, but the matching performance may get affected as it doesn't have advanced road network to support dynamic road weights and some malicious riders/Drivers intentionally leak the private keys which leads to privacy issues.

6. Methodology

The conventional methodology was to use a third party server to address the customers privacy but this did not completely solve the issues as the key was still in the third persons hand, Hence the methodology followed in this paper is as follows:

- Homomorphic Encryption to perform computations on encrypted data. So that the data can be computed without the access to the secret key.But in this method we can only calculate the euclidean distance i.e: the displacement between two points which is not useful to the online ride hailing system.
- Property preserving hash which can use Road network Embedding to calculate the distance between the driver and the rider with the streets taken into account, but the matching performance may get affected as it doesn't have advanced road network to support dynamic road weights.
- For secure distance calculation, trusted hardware technologies (e.g., Intel SGX). However, the security of their architecture is contingent on the presence of trusted server hardware.
- Bit-block encryption and a Property-preserving Hash (PPH)-based difference evaluation algorithm can be used. Here the main idea is to use

bilinear mapping to encrypt RNE vectors into bit-blocks, then incorporate the bit weight with random masks in each block cypher.



7. Research Findings

Conventionally an attacker who tries to steal the sensitive information from the ORH server can monitor the encrypted RNE vectors and the ride-matching data but the results obtained through this paper are pertaining to maintain the privacy of the customer in the online ride-hailing system, without involving a third trusted party and match the nearest driver for each rider while preserving both driver privacy and rider privacy.

8. SECURITY ASPECTS

CONFIDENTIALITY

Confidentiality refers to the protection of information from access by unauthorized parties. Homomorphic encryption is being used for encrypting data. The data includes the location coordinates of the user and the driver. Even when the transmission or storage medium has been compromised, the encrypted the information is virtually useless to unauthorized persons without the proper keys for decryption. If intercepted, the interceptor will not be able to crack unless they know the key.

ACCESSCONTROL

Users and drivers are provided with passcodes and usernames to grant access for the information that is actually useful for them. It ensures that only the people eligible to be the part of a particular resource are granted permission and able to log in to the chat room created.

NON-REPUDIATION

Non repudiation ensures that the user cannot deny sending something. The location coordinates once posted cannot be deleted and users cannot deny sending their location to the server and it will be stored once the session between the user and the driver in a group is terminated/ended. In the online ride hailing application non-repudiation is ensured so that the user cannot delete his location or change it abruptly.

AUTHENTICATION

User ID and password can be used for authenticating that the user belongs to a registered set of people who can access the online ride hailing application. All the passwords are encrypted, So there won't be any compromise with the users safety and only the authorized users can access the app.

9. Tools Description

9.1 Features

The various features of the Online Raid Hailing System without a third party are to use Encrypted Locations for Ride-Matching which is safe, the secure value comparison which is based on PPH. Building a Safe Ride-Matching Scheme so that the original location is not being leaked. And also for the improvement of ride-matching original performance.

9.2 Specifications

- We will be using Homomorphic Encryption to perform computations on encrypted data.
- As we need the distance and not the displacement we will use road network embedding to effectively calculate the distance between the rider and the driver.
- We will then bridge it with the cryptographic primitives like property preserving hash to safeguard the users privacy.

10. Comparison with existing system

In the existing system of an online ride hailing system, the location coordination of the rider and the user are sent to a third party server and here the coordinates are computed and compared inorder to match a rider with raised ride request from the users end. In this existing system the problem is that the coordinate information that is meant to be sensitive could be compromised. so in the updated system we are trying to study about a system that can take the coordinates of the location that are encrypted and then match the rider to a user in a safe way without involving a 3rd party server.

11. Contributions

The authors have propose an Encryption method called as bit-block encryption method , with which they can be computing difference of given encrypted inputs.

The three entities that have been considered in the proposed system are ORH server, Riders and Drivers.

ORH server: It helps to for calculate the distance between Rider and Driver . It stores the location of the both rider and Driver in the form of encrypted data and will calculate the distance on getting request from the Rider.

Rider and Driver: The location of both the Rider and Driver will have to be secured, for this both the driver's and Rider's original location will be converted in to set of road vectors with RNE and encrypt those vectors with the cryptographic primitives that they have proposed.

Road Network Embedding

This will allow to represent a particular location in the form of a Road vector by the system, and can help in calculating distance between the locations.
The Road network will be represented as RN = (V, E) .
Where V is set of road nodes and E be the set of Road segments.

Bilinear Map

It's a mapping represented as $e : G1 \times G2 \rightarrow GT$. G1,G2.. GT are multiplicative cyclic groups, they are provided with the same prime order p,
The g1 will be generator for G1, similarly g2 for G2.
The properties :

$$e(g1^a , g2^b) = e(g1, g2)^{ab}, \text{ for } \forall a, b \in Z.$$
$$\text{and } e(g1, g2) != 1, \text{ for } g1 \in G1 \text{ and } g2 \in G2.$$

Property-Preserving Hash

It is searchable encryption scheme, which uses the bilinear mapping functions to reveal the orders of PPH.

Property P defined as:

$$P(x, x^*) = \begin{cases} 1 & x = x^* + 1 \\ 0 & otherwise \end{cases}$$

PPH ciphertext of x can be defined as
$PPH(x) = (g1^{r1} , g1^{r1\times H(k,x)} , g2^{r2} , g2^{r2\times H(k,x+1)})$

Where H is a Hash function.
Where r1,r2 are random values.

PPH(x) can be denoted as PPH(x) = (P1, Q1, P2, Q2),
where $P1 = g1^{r1}$, $Q1 = g1^{r1\times H(k,x)}$ 1 , $P2 = g2^{r2}$ and $Q2 = g2^{r2\times H(k,x+1)}$

For the cypher text of x we can take the tuple (P1,Q1) and tuple (P2,Q2) for cypher text of x+1
PPH(x)= (P1, Q1, P2, Q2), and PPH(x*)=(P1*,Q1*,P2*,Q2*).

We can check if P(x,x*)=1 by checking
e(P1,Q2*)=e(Q1,P2*), if this equation holds it is considered as matching, i.e., x matches $x* + 1$, this property will help determine the order of 2 1-bit values in easier way.

Let v and $v*$ be the vectors. They considered v as a element of the location vector corresponding to the the rider, and $v*$ be the element of the vector for the driver. The foremost concept is to divide v and $v*$ into binary bit-blocks in parallel and encrypt every block with weighted variations with the aid of using the usage of PPH schemes. Later, they will match the block-wise ciphertext of $v*$ with that of v, and achieve the block-wise differences of $v*$ and v in ciphertext. Finally, we compute the difference among values $v*$ and v.

The two ciphertexts will be given $v_{|i}$^ and $v\hat{}_{|i}$ *, the server finds z_{j*} which $z_{j*} = v_{*|i}$ by testing whether the following equation holds with all $v\hat{}_{|i}$, $z_j \in \hat{}v|i$ and $v\hat{}* |i$:

$$e(g_1^{r1} , g_2^{r2\times H(k1,v_* |i)}) = e(g_1^{r1\times H(k1,zj)} , g_2^{r2})$$

where g_1^{r1} and $g_1^{r1 \times H(k1,zj)}$ are from $v\hat{}_{|i}$, z_j , and g_2^{r2} and $g_2^{r2 \times H(k1,v_* |i)}$ are from $v\hat{}* |i$. The difference between z_{j*} and $v_{|i}$ in z_{j*} 's ciphertext will be the difference between $v_{|i}$ *and $v_{|i}$ if z_j^* matches $v_{|i*}$. According to the above discussion, the hash token in driver's ciphertext H(k2, v∗ |i||i) is equal to H(k2,zj ∗ ||i) in the mask of zj ∗ . Therefore, the weighted difference of the (i + 1)-th pair of blocks can be revealed by XORing the random masks F(H(k2, v∗ |i||i), γj ∗), where H(k2, v∗ |i||i) is from driver's token and γj ∗ is from rider's ciphertext. Finally, by summing all the weighted differences the difference between the values v∗ and v is found . it is represented as :

$$v^* - v = \sum_{i=0}^{n-1} d_{j^*,i},$$

where d_j ∗,i is the weighted difference of z_{j*} which z_{j*} is the matching value in the (i + 1)-th block.

The computed difference will allow to calculate the difference in distance between rider and Drivers and find a match on which driver is near to the rider and thus riding match will be done with out loosing the privacy of Rider and Driver.

They have used various real-world datasets inorder to find how efficient the proposed method is working. And they found that their proposed design was able to give better performance.

12. Future Research

We intend to extend our research findings to implement them in other domains like food delivery management services, logistic services etc. We would also like to explore more upon the conventional and the traditional anatomy of the online ride hailing systems, logistic services, food delivery management apps etc and know more about how they actually work.