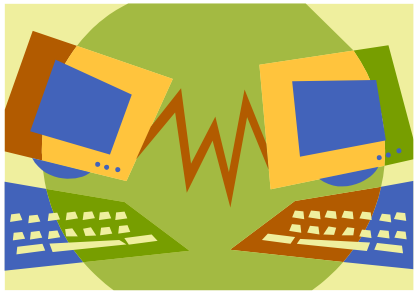
Privacy-Preserving Ride Sharing Scheme for Autonomous Vehicles in Big Data Era



Ahmed B. T. Sherif, Khaled Rabieh, Mohamed M. E. A. Mahmoud, Xiaohui Liang,

IEEE Internet of Things Journal, April 2017

Chun-Ting Po

Outline

- Introduction
- Privacy-Preserving Scheme
  - Similarity Measurement Over Encrypted Data
- Evaluation
- Conclusion

Introduction

- Autonomous vehicles (AVs) are equipped with advanced sensing and communication capabilities, etc., to enable the vehicles to autonomously drive themselves.
- Ride-sharing (or carpooling) allows AVs to be shared by users, e.g., to share the cost of on-demand cab service.
- However, the organization of ride sharing requires the users to disclose sensitive information, e.g., the pick-up/drop-off locations.

Introduction (cont.)

- In this paper, the authors propose a privacy-preserving scheme to organize ride sharing.
- The authors use a group signature scheme to ensure users anonymity, and they also use a similarity measurement technique over encrypted data to enable a server to measure the similarity of the users' trip data without knowing the data.

Privacy-Preserving Scheme

- The primary user creates a group of secondary users who can share rides with him, and he calculates a group signature credentials and distributes them to the group members.
- The primary user should also compute the secret key and send it to the group members.
- If a primary user wants to share a trip with secondary users, he sends a packet to the server having the indices for his trip data along with a signature and other information.

Privacy-Preserving Scheme (cont.)

- Similarly, the secondary users submit packets to the server, and the server uses the indices to find the secondary users who can share the ride with the primary user.

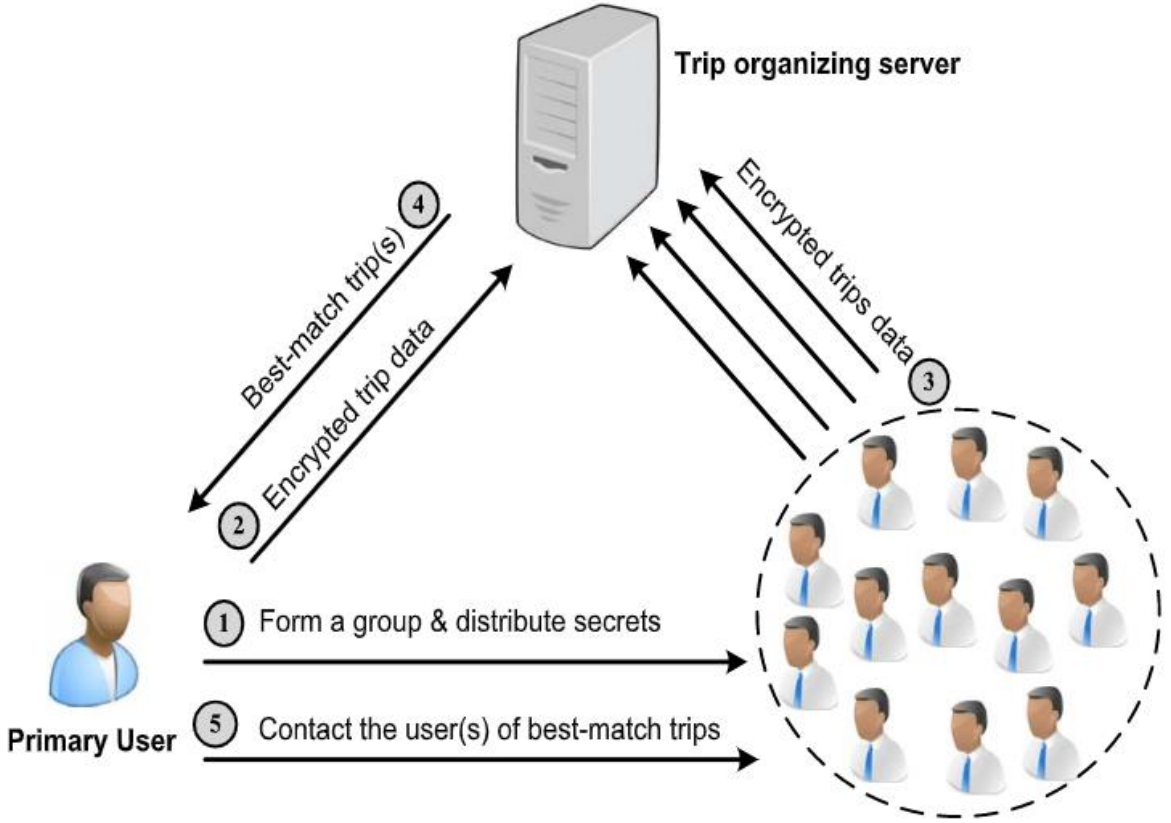
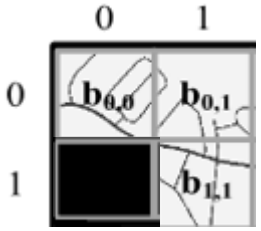


Fig. Exchanged messages in the scheme.

Similarity Measurement Over Encrypted Data

- The primary user creates bit vector p for the trip data.
- The secondary users create bit vector s for the trip data.



(0, 0, 1, 0)

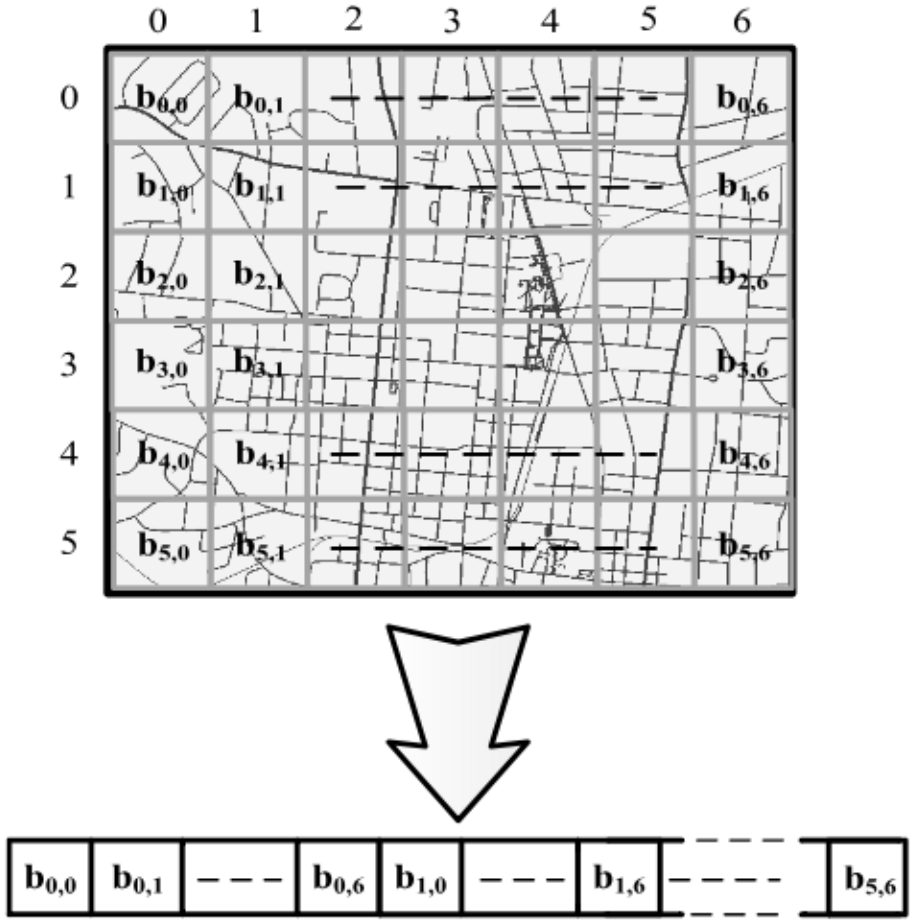


Fig. Representing a ride sharing area as a binary vector.

Similarity Measurement Over Encrypted Data (cont.)

- The secret keys used to create the indices include two matrices and a row vector.
- The matrices M_1 and M_2 and a vector S are used by the primary user, while the matrices $(M_1^{-1})^T$ and $(M_2^{-1})^T$ with the same row vector S are used by secondary users.
- Then, p is split into two random vectors p' and p".
- Finally, the trip data's index is $(p' \cdot M_1, p'' \cdot M_2)$.

Similarity Measurement Over Encrypted Data (cont.)

- To calculate the similarity of the primary user and the secondary users, the two indices $(p' \cdot M_1, p'' \cdot M_2)$ and $(s' \cdot (M_1^{-1})^T, s'' \cdot (M_2^{-1})^T)$ are multiplied using inner product.

$$(p' \cdot M_1, p'' \cdot M_2) \cdot (s' \cdot (M_1^{-1})^T, s'' \cdot (M_2^{-1})^T)$$

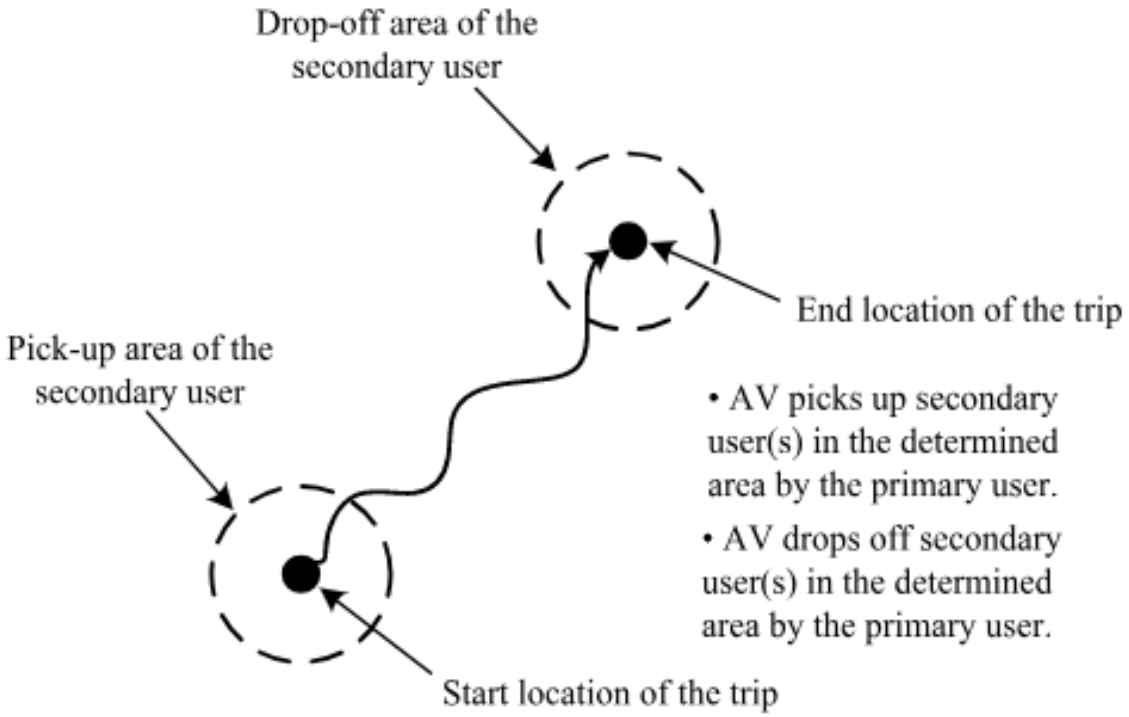
$$= p' \cdot s' + p'' \cdot s''$$

$$= p \cdot s$$

DAP-DAD Shared Rides

- In determined area for pick up and determined area for drop off (DAP-DAD) , the primary user decides the pick-up and drop-off areas of the secondary users.
- The primary user's indices { $I_p^{(p)}$, $I_p^{(d)}$, $I_p^{(t)}$ } and the secondary user's indices { $I_s^{(p)}$, $I_s^{(d)}$, $I_s^{(t)}$ } are for the pick-up location, drop-off location, and pick-up time, respectively.
- In the simulations, when the pick-up area is greater than the trip's start cell, we call this case "with flexibility."

DAP-DAD Shared Rides (cont.)



Determined Area for Pick up and Determined Area for Drop off (DAP-DAD)

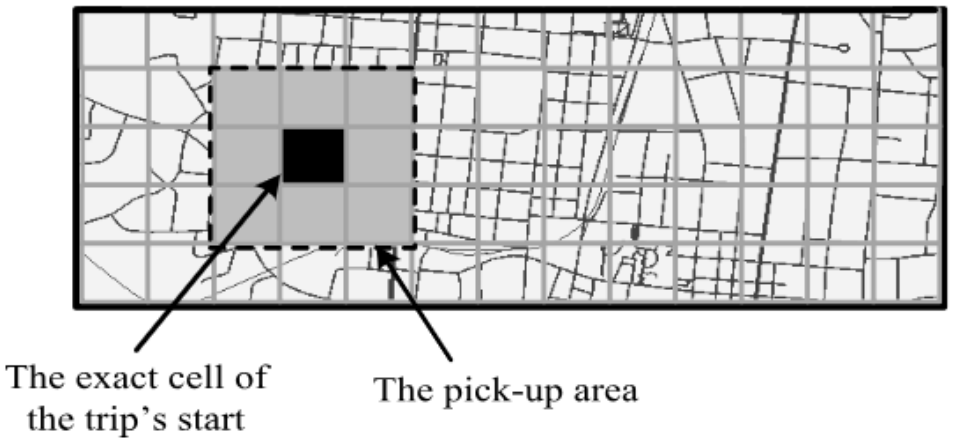
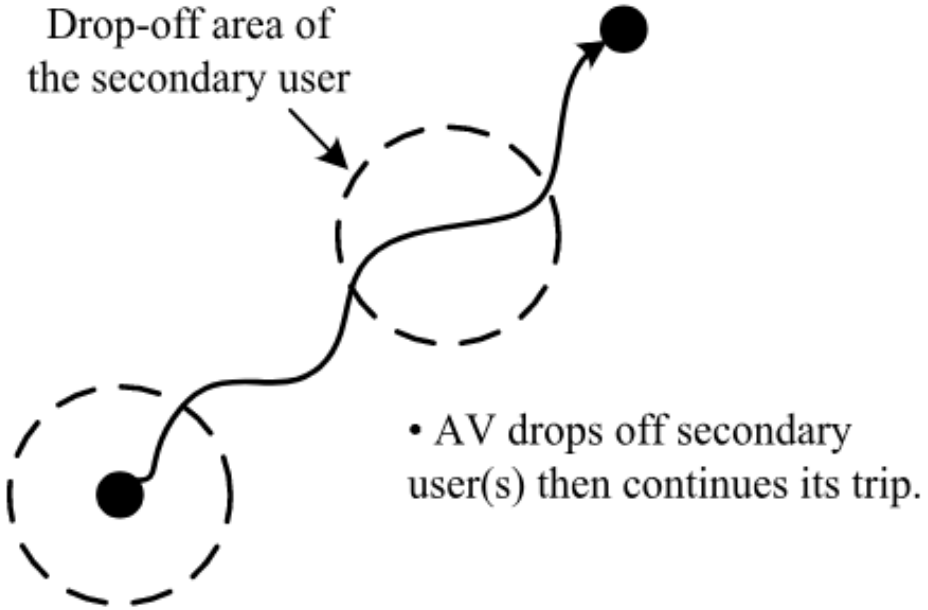


Fig. Primary user's pick-up vector.

DAP-ORD Shared Rides

- The second case of ride sharing is called determined area for pick-up and on route drop-off (DAP-ORD).
- In this case, the secondary user's drop-off location is in the surrounding area of the trip route.
- The primary user should submit indices for the pick-up area, pick-up time, and route, and the secondary user should submit indices for the pick-up location, pick-up time, and drop-off location.

DAP-ORD Shared Rides (cont.)



Determined Area for Pick up and On-route Drop off (DAP-ORD)

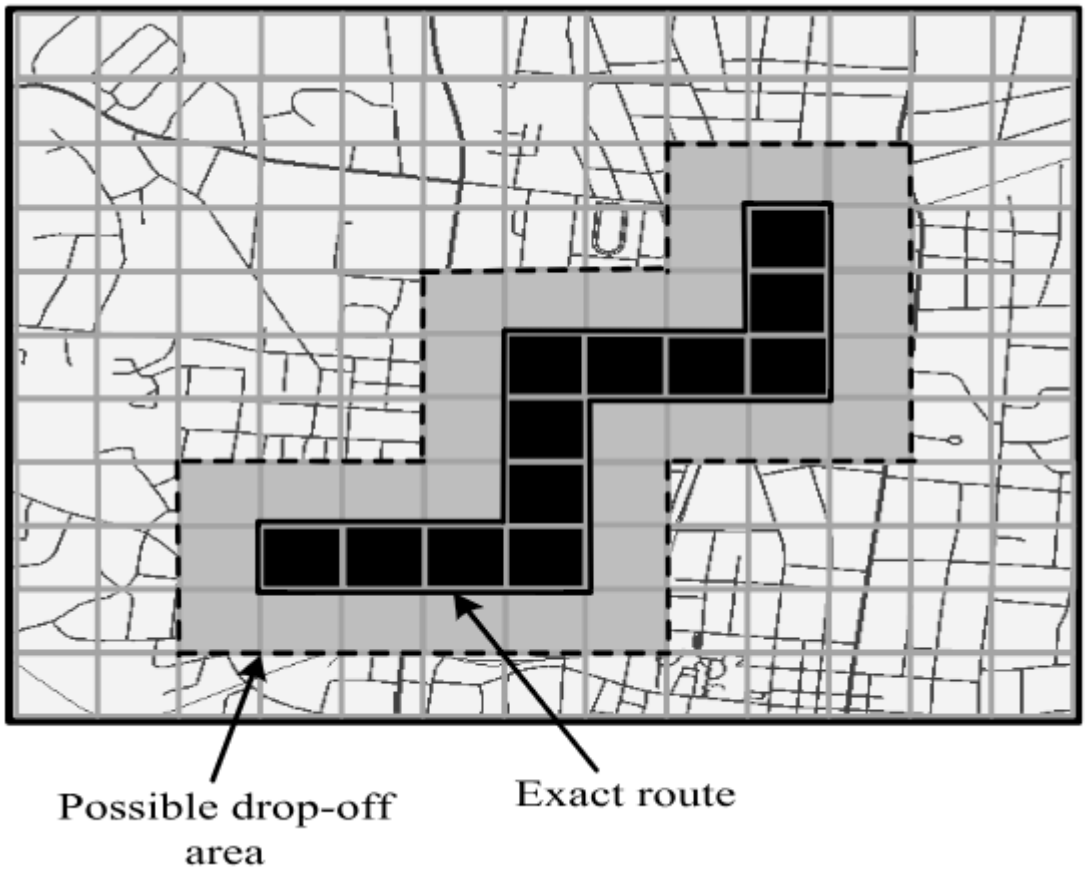
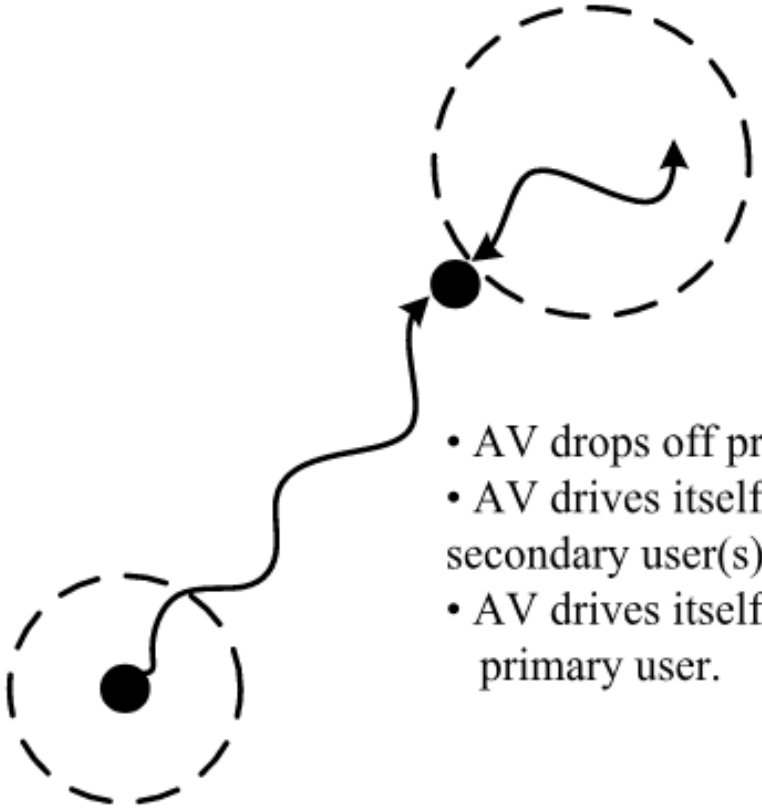


Fig. Primary user's drop-off area in DAP-ORD.

DAP-EAD Shared Rides

- The third case of ride sharing is called determined area for pick-up and extended area for drop-off (DAP-EAD).
- In this case, the primary user should submit indices for the pick-up area, pick-up time, drop-off location, and route, and the secondary user should submit indices for the pick-up location, pick-up time, and route.

DAP-EAD Shared Rides (cont.)



- AV drops off primary user.
- AV drives itself to drop off the secondary user(s).
- AV drives itself back to the primary user.

Determined Area for Pick-up and Extended Area for Drop-off (DAP-EAD)

Search Time Enhancements

- In the scheme, location information is represented by cells, and users have the same location if they share the same cell.
- If the cell size is large, the distances between users can be large but the scheme considers them in the same location.
- If the cell size is small, this setting will definitely increase the search time.

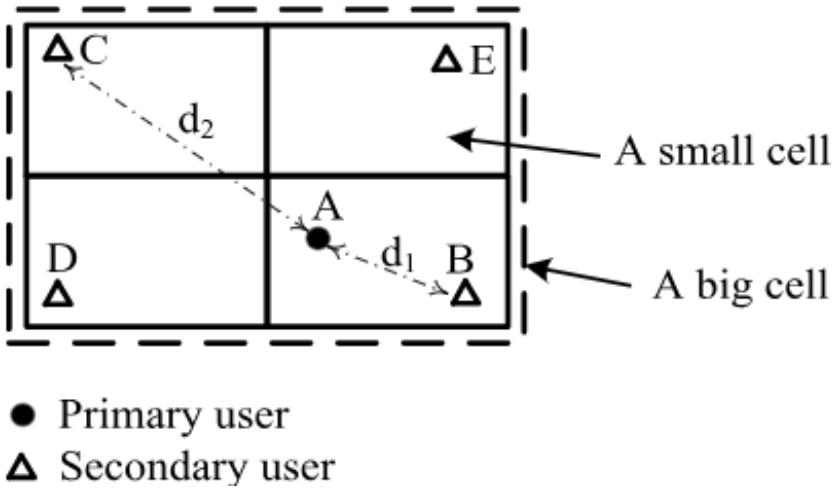


Fig. Effect of cell size on precision.

Search Time Enhancements (cont.)

- The users should represent their location data with two indices: fine-grained index and coarse-grained index.
- The server first matches the coarse-grained indices and then matches the fine-grained indices of the users whose coarse-grained indices are matched.

Search Time Enhancements (cont.)

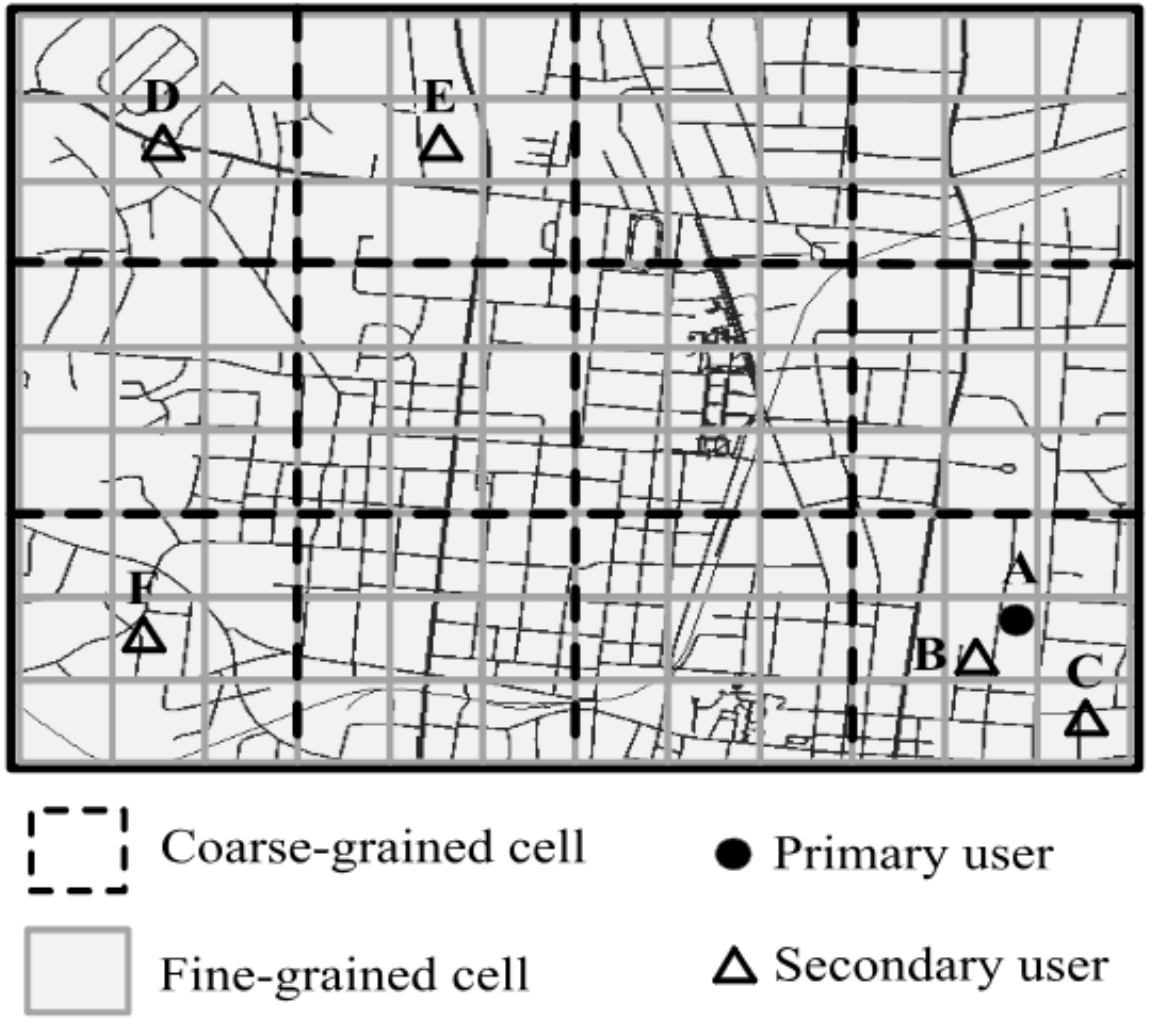


Fig. Enhancement technique.

Evaluation

- **Security/Privacy Analysis**
  - Trip Data Privacy: The server cannot infer the trip information of any user from the indices.
  - Primary-Secondary Users Unlinkability: The server cannot link the ride sharing requests (and offers) sent from the same user.
  - Identity Anonymity and Authentication: By using group signatures, each user can prove that he is a member in the group anonymously.
  - Access Control: The server can only organize shared rides for users who are members in the primary user's group.

Experiment Setup

- The authors used the OpenStreetMap project, and SUMO program to create real routes for the users. The map is for the Cookeville city located in the Tennessee state in the USA.
- The size of the ride sharing area is 21×13 km. For the coarse-grained cells, the area is divided into 273 cells. For the fine-grained cells, the area is divided into 3010 cells.

Fig. Ride sharing region used in the experiment.

Experiment Results

- For the similarity measurement technique, the indices' sizes are 6.02 kB and 546 bytes, and the sizes of the secret keys are 149.3 kB and 18.2 MB in cases of fine-grained and coarse-grained cells.
- The computation of the secret key takes 846.5 ms and 18.8 mins, and the computation of an index takes 1.57 ms and 417 ms in cases of coarse-grained and fined-grained cells, respectively.
- For the group signature, the computation of signature needs 31.6 ms and signature verification needs 51.8 ms.

Experiment Results (cont.)

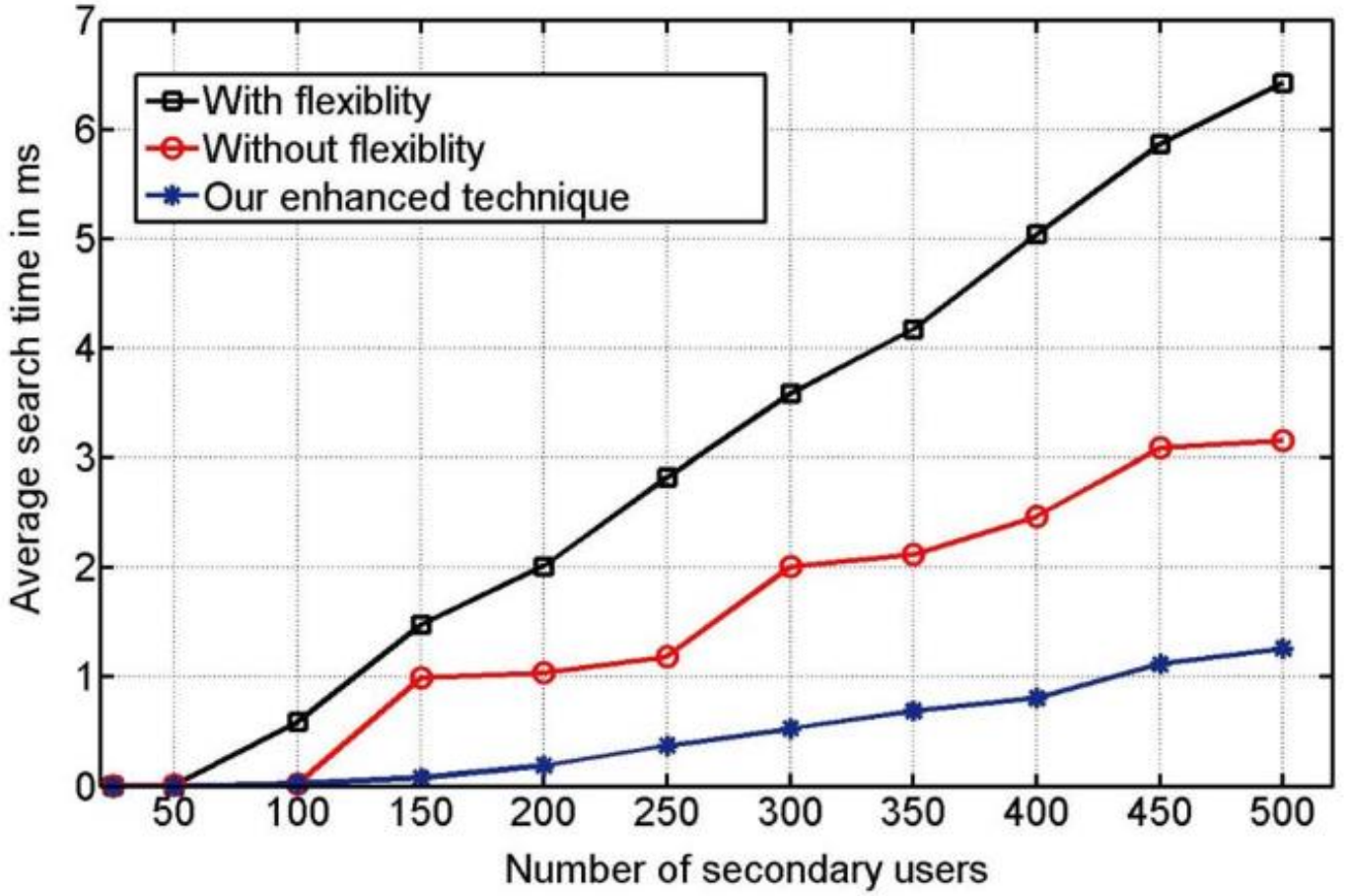


Fig. Search time versus the number of secondary users.

Experiment Results (cont.)

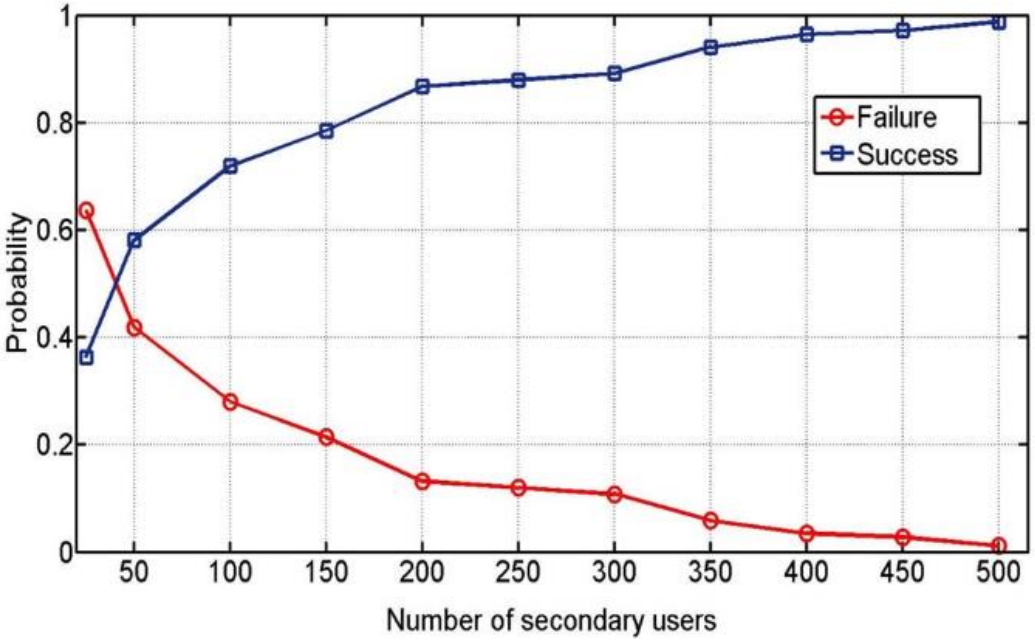


Fig. Success and failure probabilities in the flexibility case.

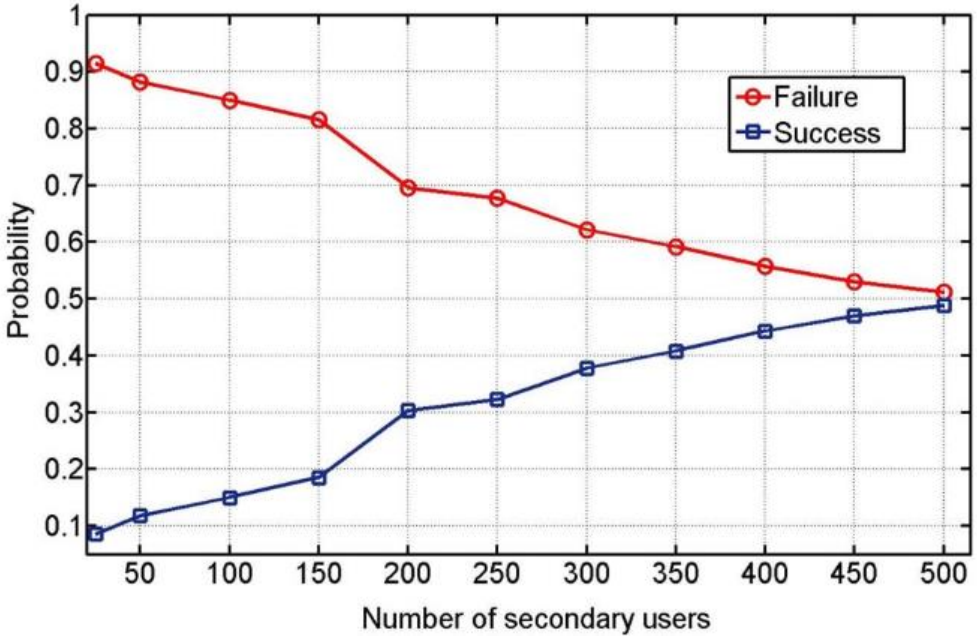


Fig. Success and failure probabilities in the nonflexibility case.

Conclusion

- In this paper, the authors have proposed a privacy-preserving ride sharing scheme for AVs.
- By using the enhanced technique of search time, better precision can be achieved with short search time.
- The experiment measurements have demonstrated that the search time is short and it can be used in the AV and big data era when ride sharing is very popular and a large number of rides should be organized in a short time.