Enhancing Privacy in Ride-Sharing Applications Through POIs Selection

Francesca Martelli
Institute of Informatics and Telematics (IIT)
Italian National Research Council (CNR)
Pisa, Italy
Email: francesca.martelli@iit.cnr.it

Maria Elena Renda
IIT-CNR, Pisa, Italy
and DUSP - MIT
Cambridge, MA
Email: elena.renda@iit.cnr.it, erenda@mit.edu

Abstract—**The problem of data privacy preservation is of central importance in ride-sharing applications, because in order to efficiently match passengers with vehicles, these services rely on exact location information. Yet, transportation and location data can reveal personal habits, preferences and behaviors, and users may prefer not to share their exact location. Masking location data in order to avoid the identification of users in case of data leakage, and/or misusage would help protect user privacy, but could also lead to poorer system performance, in terms of efficiency and quality of service as perceived by users.**

In this paper, we compare classic data masking techniques, namely *obfuscation*, k*-anonymity*, and l*-diversity*, applied to users' location data, before sending it to a carpooling system. While the first two techniques use randomly generated points to mask the actual location, l-diversity uses actual points of interest, having the additional benefit of ensuring that the disclosed location is always an accessible and safe pickup or drop-off location. Given that users in a real ride-sharing system could choose to protect or not protect their location data when using the system, we also evaluate the effect of privacy preservation penetration rate, by varying the percentage of users choosing to have their location data protected. The results show that l-diversity performance is better than the others' even when the privacy penetration rate is high, suggesting that this technique has the potential to meet both users' and system's needs, and thus being a better option to provide privacy within carpooling systems.

I. Introduction

Ride-sharing systems hold promise for improving the efficiency of transportation and reducing congestion and pollution by decreasing vehicle miles traveled (VMT). In order to accurately match drivers and riders on systems like UberPool, WazeCarpool, and Lyft Share, at least riders' origin, destination, and departure time must be communicated to the application. Ride-sharing applications usually require and/or collect data from users, and store private information, such as where and when users are going, how frequently and so on. Mobile apps collect data to improve the services offered, for location-based marketing, information or alerts, by means of mechanisms like geofencing and geo-targeting [1], [2]. In some countries, such data must be collected and retained (in some cases, for extended periods) and reported to governmental or other authorities (e.g., [3]). Sometimes, companies might share general use statistics, or may send user-identifying data to third parties [4], [5]. To help users in protecting their privacy and avoid their potential identification, masking location data is in order. However, the loss of information could decrease data utility (a metric usually related to the performance based on the effectiveness of the underlying data) and lead to poor quality or low efficiency of the location-based system [6], also because the resulting masked location could be not accessible for passengers' pickup or drop off. Furthermore, even if the resulting location point is accessible or if it has been adjusted to be accessible, curbside passenger pickups/drop offs (PUDO) could still endanger or congest traffic. For this reason, besides adopting two well established methodologies to mask location data, namely *obfuscation* [7], [8] and k*-anonymity* [9], [10], we also implement l*-diversity* [11], [12]. In fact, by using real Points of Interest (POIs) close to the real location instead of random "close" locations, l-diversity could also be seen as an option to alleviate curbside congestion, and an alternative to create dedicated urban PUDO zones [13], as it is happening in cities like Boston, San Francisco, and Washington DC.

The evaluation framework we propose here allows to compare the number of shared miles that would be achieved by optimally matching trips in a carpooling system using users' exact location information with those achieved through anonymized data, and to quantify the average increment of users' waiting and riding time. This way, we are able to understand the effects of location data-masking on ride-sharing applications both in terms of system efficiency (VMT) *and* Quality of Service (QoS) (users' riding time and waiting time). To understand the impact of users opting in or out from location privacy within the same system, we also test the sensitivity of its performance to the percentage of riders requesting location data privacy preservation. In this study, we specifically analyzed the case of carpooling between home and work, which is the largest contributor to traffic congestion and air pollution, but also because these regular trips have the potential to reveal users' repetitive route patterns and their recurrent time schedules.

This paper is organized as follows: in the next Section we review the related literature; in Section III we present the evaluation framework and the techniques we adopt; in Section IV we report the performance evaluation of the considered approaches; Section V concludes the paper.

II. Related Work

How mobility companies use data collected from users is becoming a compelling issue for citizens, authorities, and companies themselves. As data masking allows to avoid the identification of users in case of data leakage and/or misusage, these techniques are becoming the answer to the increased request for privacy protection. The simplest way to mask a location position is by obfuscating it [7], randomly selecting a point in a circled area centered on it. Similarly, *cloaking* [14] replaces the real position with the location of the center (or centroid) of the census block to which it belongs. More efficient methods to mask location data have been inherited from the database realm and reformulated for location-based services, such as k-anonymity [9], [10], and differential privacy [15]. L-diversity has been firstly proposed in [16] as a more efficient technique than k-anonymity to protect location data privacy against data breaches. The advantage of this technique is that the masked location is chosen from a pool of actual POI positions, instead of randomly generated ones. In [12], l-diversity is used to protect activity data databases, while [11] analyzes its efficiency in protecting location data when published on social networks. When combined with other techniques, l-diversity could be used to preserve trajectory data privacy from semantic attacks [17], [18].

As for understanding the different impacts location data privacy could have on mobility sharing applications, in [19] the authors evaluated computational performance and communication overhead for ride matching algorithms; [20] analyzed the impact of location data privacy on cybersecurity within a ride sharing system, in case of external attacks or malicious drivers' behaviors; in [21], an extensive evaluation of the effects of location privacy control on the mobility sharing system performance, shows a potential tradeoff between data location privacy and the "price" users, society, and the system would pay to achieve it.

Some ride-sharing systems have studied methodologies to mask users' physical coordinates before they are sent to the mobility sharing application [22], [23], while others provide options for users to meet at a pickup point instead of providing their real address [24], [25]. Recent approaches take advantage of dedicated urban areas created for managing mobility sharing curbside congestion, called PUDO zones [13].

In this paper, we show how the use of POIs as pickup/drop off points could jointly meet users' needs for privacy protection, safety and acceptable levels of QoS, and those of an efficient carpooling system.

III. Model and Algorithms

In this Section we describe the framework we designed to evaluate privacy preserving techniques within location-based carpooling applications, providing details on the process of matching trips, masking the location data, and evaluating the performance of the system in terms of efficiency and QoS.

A. Trip Matching

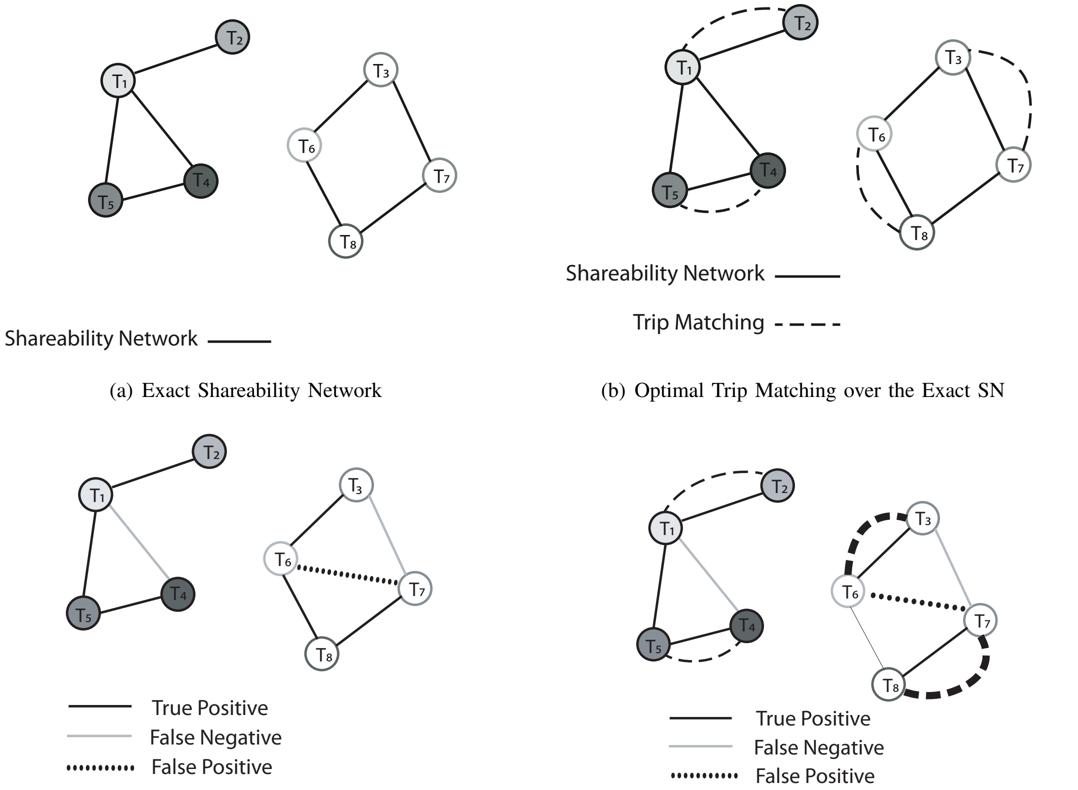
Our evaluation model requires trip matching algorithms efficient enough to be run several times with different parameter settings for each chosen data-masking algorithm. For this reason, to find matching opportunities given riders' and drivers' needs, we use the *Shareability Network* (SN) model, presented for the first time in [26] for taxi rides, and refined in [27] for carpooling applications.

In a carpooling SN, each node represents a trip, and an edge between two nodes represents the fact that the trips associated to these nodes are shareable. The trips in the SN are shareable if and only if they satisfy three spatial and temporal "shareability conditions", related to users' physical location and destination, and users' departure time and potential flexibility. The first condition states that the required *detour time* for the driver to pick the passenger up and drop them off does not exceed a given threshold value. The second condition ensures that the starting time windows of the trips are properly overlapped, so that the driver could pickup and then drop off the passenger. While this condition establishes the temporal compatibility of the trips, it is not enough to ensure that carpooling actually reduces VMT. In fact, it is possible that the driver performs a very long detour to pickup or drop off the passenger, thus actually increasing VMT. The third shareability condition ensures that the length of the shared trip is no longer than the sum of the lengths of the individual trips, so to avoid the above described possible negative side effects of shared mobility (see [27] for details). Depending on the metric used and the objective to achieve (reduce kilometers, time, number of vehicles, etc.), a specific weight on each edge of the SN could be set; in this study, focused on VMT, the weight on each edge between two nodes has been set to be equal to the amount of kilometers that could be saved if the correspondent trips are selected for the matching. Once the SN has been generated, we compute the optimal trip matching over it to maximize the overall saved VMT, by applying the well known Edmond's maximum weight matching algorithm.

B. Location Privacy

To anonymize location data and preserve users' privacy, we first implemented two classic data masking techniques, namely obfuscation [7], [8] and k-anonymity [9], [10]. The former consists in replacing the location data point with a randomly selected one within a given radius from the original one, while the latter selects the masked location from a set of k positions, $k - 1$ randomly generated within a given radius from the original position, plus the real location. One drawback of these techniques is that the masked location could be not accessible for pickup/drop off, and even if we overcome this problem by providing the closest accessible location, this could still result to be a problematic pickups/drop-off point and could somehow endanger and/or congest traffic.

L–diversity [11], [12] is another well known data masking technique, pretty similar to k-anonymity, with the difference that instead of a random "generic" location point selected to replace the original one, a random POI location is selected.



(a) Exact Shareability Network

(b) Optimal Trip Matching over the Exact SN

(c) Privacy-preserving SN

(d) Optimal Trip Matching over the Privacy-preserving SN

Fig. 1: Example of the optimal matching over the exact Shareability Network (SN) and of a good matching over the privacy-preserving SN: (a) Network of feasible shareable trips based on exact locations. (b) Optimized set of matched trips based on exact locations (optimal trip matchings indicated by curved, dashed lines). (c) The SN obtained by masking trip locations. In this case, trips T3 and T7 are determined to be not shareable (false negative) while T6 and T7 are determined to be shareable (false positive), when compared to the exact network (as shown in (b)). (d) Trip matching over the privacy-preserving SN. Good matches (bold, dashed lines) are trips selected for the matching over the privacy-preserving SN but not included in the one over the exact SN because not optimal. By selecting the good matches T_3-T_6 and T_7-T_8, we only measure a lower efficiency in the privacy-preserving matching, since the good matches would not optimize the system goal.

More specifically, a close POI is selected among l randomly selected POIs, located within a given radius from the exact location. This technique not only provides location anonymity, but has the additional benefit, unlike other methodologies, of ensuring that the disclosed location is always an accessible and safe pickup or drop-off location.

C. Key Performance Indicators

To evaluate system efficiency and QoS, we need to know how the system performs in terms of saved VMT, users' riding time, and users' waiting time, both with and without applying any masking technique to users' location data. We compute the optimal trip matching solution for the *exact* SN, derived from the real location data, and, for each masking methodology, the optimal trip matching solution for the *privacy-preserving* SN, generated using the masked data. Our evaluation methodology consists in comparing both the exact and privacy-preserving SNs, and the corresponding optimal trip matchings, by identifying specific Key Performance Indicators (KPI). In particular, by comparing the two SNs we are able to identify those trips incorrectly selected as shareable (*false positive*), and those incorrectly labeled as not shareable (*false negative*) in the privacy-preserving SN, due to the inexact location information. By comparing the optimal matching over the exact SN and the optimal matching over the privacy-preserving SN, we are able to quantify how many true positive and false positive shareable trips are actually selected when matching trips in the latter (called *Good Matches* and *Bad Matches*, respectively), and how the presence of good and bad matches could degrade the performance in terms of VMT and users' waiting and riding time. In fact, by selecting good matches in the privacy-preserving matching we could measure a lower efficiency with respect to the optimal matching over the exact SN, since the good matches would not optimize the system goal of reducing total VMT, while by selecting bad matches in the privacy-preserving matching, i.e. trips that are not shareable in the exact SN but selected and matched in the privacy-preserving SN because of the anonymized locations, could degrade both system efficiency (VMT) and QoS (users' waiting and riding time). In Figures 1 and 2 we report an example of an exact and a privacy-preserving SN and their corresponding optimal trip matchings, in case only good matches are selected (Figure 1),

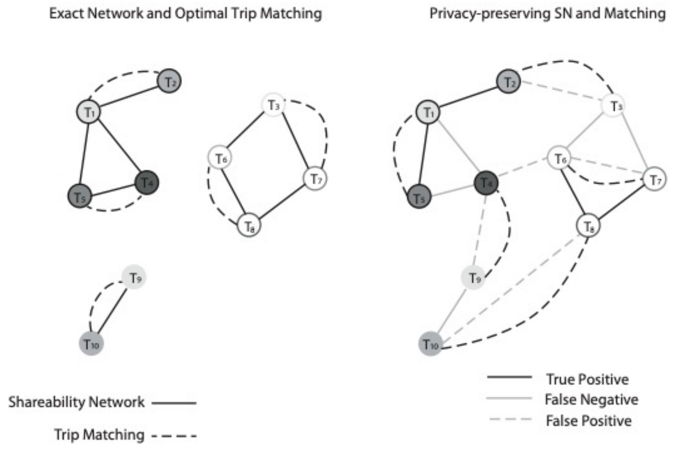


Fig. 2: Example of a "bad matching" (on the right) over the privacy-preserving SN compared to the exact SN (on the left). The SN (straight lines) and matching (dashed lines) for the exact data on the left, and for the privacy-preserving data on the right. Some of the selected links for the matching in the latter are false positive (light dotted lines), i.e. not present in the exact SN. Bad matches (namely $T_2 - T_3$, $T_4 - T_9$, $T_6 - T_7$, and $T_8 - T_{10}$) are trips that are not shareable in the exact SN but selected in the privacy-preserving SN because of the anonymized locations, degrading system efficiency QoS in the matching.

and in case both good and bad matches are selected (Figure 2).

IV. Performance evaluation

A. Data

To evaluate efficiency and QoS, we needed a dataset containing not only users' location data and departure times, but also their flexibility. To this end, we use the data collected through the MobilitandoPisa survey issued in 2016 to people living and/or working in Pisa, Italy, and its surroundings [28]. This survey data contains detailed information on origins and destinations of daily car commuters and their departure and arrival times from home to work and vice versa. It also has the unique feature of containing a quantification of commuter's flexibility in departure and arrival time, allowing us to define a departure time window for each trip, and evaluate the effects of data privacy on users' waiting time. We checked and cleaned the survey entries to remove invalid or incomplete ones, ending up with a total of 1966 trips.

B. Simulation settings

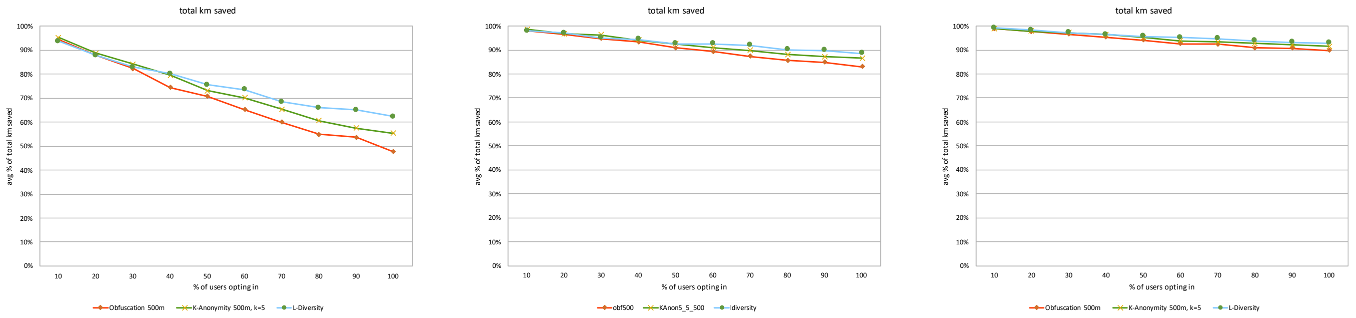
To build the SNs given trips' origin and destination, both with and without data-masking, we need to know trips' travel times; on this regard, we used OpenStreetMap [29] travel times for free-flow traffic. A different SN is computed for each value of detour time within the set $[0, 1, ..., 15]$ minutes. As for the masking techniques, after analyzing the data and discovering that there are a limited number of POIs within short distance of many trip locations, we chose a 500 meters radius for all the techniques, and set $l = 5$ POIs for l-diversity and $k = 5$ for k-anonymity. As for the penetration rate, we varied the percentage of users opting in for privacy preservation from 10% to 100%, in steps of 10%.

For each system configuration, namely by setting the data-masking algorithm (when used) and the percentage of users opting in, and the allowed detour time, we ran 10 instances of the matching algorithm, and averaged the results.

C. Results

Without using any privacy preservation technique, the maximum kilometers saved with the trip matching over the exact SN is 8,658 km with an allowed detour time of 5 minutes, 10,590 km with an allowed detour time of 10 minutes, and 11,135 km with an allowed detour time of 15 minutes. These are the savings achievable by optimally matching the trips over the exact SN, compared to having all the trips driven separately, i.e. without trip matching, which accounts for a total of 28,708 traveled km. In Figure 3 we report the average saved distance achievable with the trip matching computed over the privacy-preserving SN, expressed as a percentage of the saved distance achievable with the optimal trip matching over the exact SN, built on the real location data. Each sub-figure reports the results for the three privacy-preserving methodologies we have analyzed, while varying the percentage of users opting in from 10% to 100%, allowing three different values for the detour time: 5, 10 and 15 minutes. The experimental results confirm the intuition that more privacy implies lower efficiency. They clearly show that whenever users have a low tolerance for detour time (see Figure 3 (a)), the overall saved distance of the system is much lower compared to that of the optimal matching computed on the real data, and the efficiency degrades even more when more users opt into privacy. In fact, due to privacy preservation, users' locations have been anonymized, and it could happen that the shareability condition designed to reduce total travel distance is not actually optimized (in presence of good matches) and/or satisfied (in presence of bad matches). Increasing the allowable detour time up to 15 minutes (Figure 3 (c)) results in much less sensitivity to people opting into privacy, and higher savings. In comparing the efficiency degradation of the methodologies we adopted, l-diversity performs slightly better than k-anonymity and obfuscation, showing to have the best VMT savings, even with 100% of users opting into privacy.

Table I reports the effects of privacy-preservation on k-anonymity and l-diversity performance, with different allowable detour times. This table reports the overall percentage of bad matches and the relative percentage of those violating each particular constraint, namely exceeding: the allowable detour time, the allowable waiting time, and the distance that would be required for performing the matched trips as "solo" trips. To show the efficiency degradation we report the number of kilometers saved by matching the trips with no privacy preservation and those saved by applying the specific location privacy preservation technique, while to show QoS degradation we report the average extra detour and waiting time. Obfuscation results, being pretty similar to the k-anonimity ones, are not reported here. The percentage of total bad matches is quite high when the allowable detour time is low, with failure of the detour time constraint contributing to the majority of the bad



(a) Detour time allowed 5'. Values are expressed as a percentage relative to the maximum savings of 8,658 km without privacy.

(b) Detour time allowed 10'. Values are expressed as a percentage of the maximum savings of 10,590 km without privacy.

(c) Detour time allowed 15'. Values are expressed as a percentage of the maximum savings of 11,135 km without privacy.

Fig. 3: Saved distance achievable with the trip matching computed over the privacy-preserving SN, expressed as a percentage of the saved distance of the optimal matching over the exact SN. Here we report the saved distance varying the percentage of users opting-in, and different values of allowed detour time of (a) 5, (b) 10, and (c) 15 minutes. For each detour time value, we report the amount of saved kilometers of the optimal matching over the exact SN compared to having all the trips driven separately, which accounts for a total of 28,708 traveled km.

Methodology	**Metric**	**Allowable Detour Time**		
		5'	**10'**	**15'**
l-diversity	**Bad Matches (%)**			
	Total	26.1	7.8	4.6
	Due to detour time constraint	92.3	55.7	27.4
	Due to time windows constraint	3.5	10.4	18.2
	Due to saved distance constraint	13.0	38.8	59.3
	Saved kilometers (km)			
	With no privacy	8,658	10,590	11,135
	With 100% privacy	5,396	9,374	10,345
	Average extra time (minutes)			
	Detour time	2.3	2.7	2.4
	Waiting time	1.0	1.0	1.0
k-anonymity	**Bad Matches (%)**			
	Total	28.4	9.2	6.7
	Due to detour time constraint	89.9	42.2	15.8
	Due to time windows constraint	4.1	11.3	16.2
	Due to saved distance constraint	20.9	55.8	73.2
	Saved kilometers (km)			
	With no privacy	8,658	10,590	11,135
	With 100% privacy	4,762	9,108	10,133
	Average extra time (minutes)			
	Detour time	2.5	3.0	4.4
	Waiting time	1.0	1.2	1.1

TABLE I: L-diversity and k-anonymity: percentage of matched trips on the privacy-preserving SN violating at least one shareability constraint compared to the matching on the exact SN. The results reported here are for 100% of users opting-in. Note that because multiple constraints may be violated simultaneously, the sum of the three percentages (detour time, time windows, and saved distance) may exceed 100%.

matches. With low values of allowed detour time, the fraction of bad matches is quite high primarily because matched trips cannot comply with the acceptable detour time constraint. This not only has an impact on saved VMT, but also on users' travel time (Detour time on Table I). In increasing the allowed detour, l-diversity maintains the average extra detour time pretty stable and a low waiting time.

Increasing the allowable detour time reduces the number of badly matched trips while having a negligible impact on the average extra detour time required to share these trips. Even if the relative percentage of bad matches due to non-overlapping time windows for the shared trips slightly increases, the average extra waiting time is quite stable for all the techniques. These results show that the majority of performance loss comes at the expense of VMT and users' travel time.

V. Conclusion

In ride-sharing applications, the problem of privacy preservation is of central importance since to match drivers and riders, at least their origin, destination, and departure time must be communicated to the system. On the one hand, anonymizing location data to avoid the identification of users in case of data leakage, misuse and/or breaches increases user privacy; on the other hand, the loss of accurate location information could lead to poor quality or lower efficiency of the mobility system, and raise safety issues as well if the masked location is not (easily) accessible.

In this paper, we studied the impact of location masking techniques on the ride sharing system performance by varying the percentage of users opting in, and not only in terms of efficiency (saved VMT), but also in terms of QoS, as perceived by the user (riding and waiting time, accessible PUDO locations). Our results clearly show that the higher is the privacy protection penetration rate, i.e. more users opting in for privacy preservation, the more the efficiency degrades, with lower effects if users allow for higher detour times. On the other hand, the more the user is willing to spend longer time in the vehicle (longer allowable detour time), the less the efficiency is compromised, and vice versa. This could be used by the (private or public) stakeholder in the decision-making process in order to effectively pursue the overall system goal.

Moreover, the results reported here show that l-diversity performs better than the other methodologies, even when the privacy penetration rate is high, with lower saved VMT losses and users' riding and waiting extra time. The results, combined with the fact that l-diversity provides accessible and safe PUDO locations by definition, suggest that this technique has the potential to meet both users' and system's needs, being a better option to provide privacy within carpooling systems.

References

[1] S. Lian, T. Cha, and Y. Xu, "Enhancing geotargeting with temporal targeting, behavioral targeting and promotion for comprehensive contextual targeting," *Decision Support Systems*, vol. 117, pp. 28 – 37, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167923618302008

[2] M. Maiouak and T. Taleb, "Dynamic maps for automated driving and uav geofencing," *IEEE Wireless Communications*, vol. 26, no. 4, pp. 54–59, August 2019.

[3] L. Feiler, "The legality of the data retention directive in light of the fundamental rights to privacy and data protection," *European Journal of Law and Technology*, vol. 1, no. 3, 2010. [Online]. Available: http://ejlt.org/article/view/29

[4] G. Cleary, "Mobile privacy: What do your apps know about you?" 2018, last accessed: March 2022. [Online]. Available: https://www.symantec.com/blogs/threat-intelligence/mobile-privacy-apps

[5] M. H. K. Jennifer Valentino-DeVries, Natasha Singer and A. Krolik, "Our apps know where you were last night, and they're not keeping it secret," 2018, last accessed: March 2020. [Online]. Available: https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

[6] D. Calacci, A. Berke, K. Larson, and A. Pentland, "The tradeoff between the utility and risk of location data and implications for public good," *arXiv*, vol. 1905.09350, 2019. [Online]. Available: http://arxiv.org/abs/1905.09350

[7] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *Proceedings of the Third International Conference on Pervasive Computing*, ser. PERVASIVE'05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 152–170. [Online]. Available: http://dx.doi.org/10.1007/11428572_10

[8] R. Dewri, "Local differential perturbations: Location privacy under approximate knowledge attackers," *IEEE Transactions on Mobile Computing*, vol. 12, no. 12, pp. 2360–2372, Dec 2013.

[9] B. Gedik and L. Liu, "Location privacy in mobile systems: A personalized anonymization model," in *25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, June 2005, pp. 620–629.

[10] P. Shankar, V. Ganapathy, and L. Iftode, "Privately querying location-based services with sybilquery," in *Proceedings of the 11th International Conference on Ubiquitous Computing*, ser. UbiComp '09. New York, NY, USA: ACM, 2009, pp. 31–40. [Online]. Available: http://doi.acm.org/10.1145/1620545.1620550

[11] L. N. and D. S.K., "Applications of k-anonymity and l-diversity in publishing online social networks," in *Security and Privacy in Social Networks*, AY, EY, CA, AN, and PA, Eds. Oxford: Springer, 2013.

[12] P. Parameshwarappa, Z. Chen, and G. Koru, "Anonymization of daily activity data by using l-diversity privacy model," *ACM Trans. Manage. Inf. Syst.*, vol. 12, no. 3, jun 2021. [Online]. Available: https://doi.org/10.1145/3456876

[13] A. Fielbaum, X. Bai, and J. Alonso-Mora, "On-demand ridesharing with optimized pick-up and drop-off walking locations," *Transportation Research Part C: Emerging Technologies*, vol. 126, p. 103061, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X21000887

[14] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, ser. MobiSys '03. New York, NY, USA: ACM, 2003, pp. 31–42. [Online]. Available: http://doi.acm.org/10.1145/1066116.1189037

[15] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 901–914. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516735

[16] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, "L-diversity: privacy beyond k-anonymity," in *22nd International Conference on Data Engineering (ICDE'06)*, April 2006, pp. 24–24.

[17] Z. Tu, K. Zhao, F. Xu, Y. Li, L. Su, and D. Jin, "Protecting trajectory from semantic attack considering k-anonymity, l-diversity, and t-closeness," *IEEE Transactions on Network and Service Management*, vol. 16, no. 1, pp. 264–278, March 2019.

[18] L. Yao, Z. Chen, H. Hu, G. Wu, and B. Wu, "Sensitive attribute privacy preservation of trajectory data publishing based on l-diversity," *Distributed and Parallel Databases*, vol. 39, no. 3, pp. 785–811, 2021.

[19] M. Li, L. Zhu, and X. Lin, "Efficient and privacy-preserving carpooling using blockchain-assisted vehicular fog computing," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4573–4584, June 2019.

[20] D. Suo, M. Renda, and J. Zhao, "Quantifying the tradeoff between cybersecurity and location privacy," 2021, tech. Rep. n. 2021-TR-07, Istituto di Informatica e Telematica - CNR. Submitted and revised also on arXiv. [Online]. Available: https://arxiv.org/abs/2105.01262v2

[21] F. Martelli, M. E. Renda, and J. Zhao, "The price of privacy control in mobility sharing," *Journal of Urban Technology*, vol. 28, no. 1-2, pp. 237–262, 2021.

[22] U. M. Aïvodji, K. Huguenin, M.-J. Huguet, and M.-O. Killijian, "Sride: A privacy-preserving ridesharing system," in *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, ser. WiSec '18. New York, NY, USA: ACM, 2018, pp. 40–50. [Online]. Available: http://doi.acm.org/10.1145/3212480.3212483

[23] Y. Luo, X. Jia, S. Fu, and M. Xu, "pRide: Privacy-preserving ride matching over road networks for online ride-hailing service," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 7, pp. 1791–1802, July 2019.

[24] U. M. Aïvodji, S. Gambs, M.-J. Huguet, and M.-O. Killijian, "Meeting points in ridesharing: A privacy-preserving approach," *Transportation Research Part C: Emerging Technologies*, vol. 72, pp. 239–253, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X1630184X

[25] M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar, "The benefits of meeting points in ride-sharing systems," *Transportation Research Part B: Methodological*, vol. 82, pp. 36–53, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261515002088

[26] P. Santi, G. Resta, M. Szell, S. Sobolevsky, S. H. Strogatz, and C. Ratti, "Quantifying the benefits of vehicle pooling with shareability networks," *Proceedings of the National Academy of Sciences*, vol. 111, no. 37, pp. 13 290–13 294, 2014.

[27] F. Librino, M. E. Renda, P. Santi, F. Martelli, G. Resta, F. Duarte, C. Ratti, and J. Zhao, "Home-work carpooling for social mixing," *Transportation*, vol. 47, no. 5, pp. 2671–2701, 2020.

[28] Pisa Municipality, "Mobilitando Pisa," 2016, last accessed: March 2022. [Online]. Available: https://www.pisamo.it/wp/2016/02/mobilitando-pisa

[29] [Online]. Available: https://www.openstreetmap.org