Influence of tracking duration on the privacy of individual mobility graphs

Conference Paper

Author(s):
Martin, Henry; Wiedemann, Nina; Suel, Esra; Hong, Ye; Xin, Yanan

Publication date:
2022

Permanent link:
https://doi.org/10.3929/ethz-b-000572753

Influence of tracking duration on the privacy of individual mobility graphs

Henry Martin[1,2,*], Nina Wiedemann[1,*], Esra Suel[1,3], Ye Hong[1], Yanan Xin[1]

[1] Institute of Cartography and Geoinformation, ETH Zurich, Switzerland
[2] Institute of Advanced Research in Artificial Intelligence (IARAI), Austria
[3] Center for Advanced Spatial Analysis, University College London, United Kingdom
[*] Authors contributed equally

Keywords: privacy, mobility graphs, time dependence

***Summary**: Location graphs are a compact representation of individual mobility that can be used as a mobility profile to personalize location-based services. While location graphs are more privacy-preserving than raw tracking data, it was shown that there is still a considerable risk for users to be re-identified by their mobility graph topology. However, it is unclear how this risk depends on the tracking duration. Here, we consider a scenario where the attacker wants to match new tracking data of a user to a pool of previously recorded mobility profiles, and we analyze the dependence of the re-identification performance on the tracking duration. For our experiment, we use a one-year long tracking dataset of 137 users divided into subsets of varying durations (4, 8, 16, 20, 24, and 28 weeks). We find that the re-identification performance increases with growing pool- and test-user tracking duration, and even the smallest tested duration allows to match users significantly better than random. The provided evidence for a tracking duration dependency of user privacy has clear implications for the data collection and storage strategies. It is advised for data collectors to limit the tracking duration or to reset user IDs regularly when storing long-term tracking data.*

1 Introduction and Background

Companies are increasingly gathering and using spatio-temporal location data from personal mobile devices. User location data have substantially improved location-based services (LBS) and personalized offers [11]. However, detailed mobility traces collected from individuals may contain sensitive personal data that are associated with high privacy risks [18]. A particular concern is the increasing integration of user data from different sources [23], enabling companies to build more detailed and complete user profiles [16]. Therefore, identifiability (and matching) of individuals from different datasets is a critical dimension of data privacy risk [11].

Previous studies showed that the removal of basic identity information from mobility traces is insufficient in this context, as users can be re-identified using the information on frequently visited locations [2, 3, 7, 8, 22, 28]. One solution proposed in the literature is to obscure the geographic coordinates in order to guarantee ε-differential privacy [4, 10, 25]. Another promising possibility for privacy-preserving storage and processing of individual tracking data is given with so-called *location graphs* or *mobility networks* [19, 21]. In these graphs, nodes usually represent visited locations and edge weights correspond to the number of observed movements between these locations. Graph representations offer several benefits: 1) they can be enriched with node and edge features based on the application needs, 2) they are compact and grow sub-linearly in size with increasing tracking duration, and 3) they can be analyzed efficiently with graph neural networks for various applications such as activity purpose imputation [13].

The privacy and unique identifiability properties of individual mobility graphs, however, are not well understood. Recently Manousakas et al. [12] showed that the graph topology of personalized mobility graphs, even when all coordinate and time stamp information is removed from its nodes, is often uniquely identifying. In this paper, we build upon their work and aim to understand the dependency of privacy-preservation on the tracking duration. Intuitively, location graphs over short time periods are more diverse for one user and may reduce the risk of deanonymization. To investigate this possibility, we divide a tracking dataset of 137 users into distinct time periods of different durations and analyze attack scenarios where a new location graph is matched to a pool of location graphs of known users. Our experiments indeed show that matching performance depends on the tracking duration of both pool data and new data; however, there is a considerable re-identification risk even with just four weeks of tracking duration.

2 Materials and Methods

2.1 Data and preprocessing

We analyze the time dependency of topology privacy on a high-quality tracking dataset, collected through the SBB Green Class tracking study [14]. The study was conducted by the Swiss Federal Railways (SBB) to evaluate the impact of a mobility-as-a-service offer on individuals' travel behavior. Study participants are predominantly male with above average income. All study participants were tracked over a full year using an application installed on their phone[1] that segments tracking data into stationary periods called *staypoints* and movement behavior called *triplegs*. We summarize *staypoints* recorded at the same place to *locations* and aggregate all movement between two significant stays at locations with a duration larger than 25 minutes to *trips*. All preprocessing is done in Python and PostgreSQL using the Trackintel movement data processing library [15]. A detailed overview of the preprocessing steps and parameters is given in Appendix A.

Based on the sequence of locations and trips of a user, we construct the individual location graph (or mobility network) as described by Manousakas et al. [12]: In the graph $G(V,E)$, each location is one node, and each trip between two locations increases the weight of the directed edge by 1. The edge weight $w(e)$ thus corresponds to the number of transitions during the observation period. To analyze the impact of different tracking periods, we build the graphs on subsets of the dataset that are created by binning the dataset into non-overlapping time periods of 4, 8, 16, 20, 24, and 28 weeks (see Figure 1).

2.2 Feature based graph matching

Graph matching describes the problem of either identifying if two graphs are isomorphic (exact graph matching) or identifying the best match from a set of candidate graphs (inexact graph matching problem) [20]. The exact solutions for both problems are computationally intractable, and we therefore rely on heuristics to accomplish inexact graph matching. Related works have proposed so-called R-convolution graph kernels [9] that measure the difference between two graphs in terms of the counts of certain substructures, such as paths. Similarly, we compare the distributions of selected graph features to approximate the graph similarity. We represent each graph in a fixed-size vector $v(G)$ that expresses graph characteristics, e.g.,the distribution of node in-degrees. Two graphs G_1 and G_2 are compared in terms of the distance between their vector representations, $d(v(G_1),v(G_2))$. As distance metrics d, we test a simple Mean Squared Error (MSE), Kullback-Leibler divergence, and Wasserstein distance.

We experiment with five vector-based graph representations $v(G)$:

- $v_{indegree}$: Distribution of (unweighted) node in-degrees, i.e., the number of connections of one location to other locations. The distribution of in-degrees over the 20 most popular locations is used.

- $v_{outdegree}$: Similar to the in-degree, the distribution of out-degrees over the 20 locations with the highest out-degree is computed.

[1] https://play.google.com/store/apps/details?id=ch.sbb.myway

- $v_{transition}$: The distribution of transition weights over the 20 most popular trips. Intuitively, some users commute between very few locations more frequently than other locations, whereas some users transit more evenly among locations [17].

- $v_{shortest\_path}$: The distribution of shortest-path lengths in the graph. All-pairs shortest paths were computed with the Floyd-Warshall algorithm [5, 26] and the ratio of shortest paths with length x is reported in $v_{shortest\_path}$ for $x \leq 10$.

- $v_{centrality}$: The betweenness centrality [6] of a node denotes its centrality (in terms of network hops) with respect to other nodes, which is bounded between 0 and 1. Since many nodes have low centrality in mobility graphs, we construct 10 bins from 0 to 1 in log space and report the number of nodes per centrality bin.

Finally, we concatenate all five graph descriptors into one combined vector v_{comb}.

2.3 *Experiment design*

We analyze the following privacy attack scenario: The adversary is a data broker who has access to a pool of users and their tracking data. The attacker then gets access to additional tracking data of a test user which she wants to match to the correct user in the pool to create a combined user profile. All tracking data are represented in the form of weighted and directed individual location graphs without any node or edge features such as coordinates. In the following, we define $u_i^{pool}, i \in [1..n]$ as the i-th user in a pool of n users, and $u_j^{test}, j \in [1..m]$ as a user of the test dataset, $D_{test} = \{u_j^{test}\}$. Let G_i^{pool} and G_j^{test} further denote the corresponding location graphs.

The adversary now aims to find the best match out of the pool users for each test user u_j^{test}. This is accomplished by computing the distance of the graph descriptors presented in Section 2.2. The pairwise distances from the test user to all users of the pool are computed as $d\big(v(G_j^{test}), v(G_i^{pool})\big)$ and the pool users are ranked according to their distance. As a result, we obtain the rank that was assigned to the true match of a user in the pool. In other words, we are only interested in the rank that was assigned to the user in the pool that corresponds to the test user ($u_i^{pool} = u_j^{test}$) and the assigned rank $r_j = r(u_j^{test})$ means that this user had the r_j-highest similarity to herself compared to all other users in the pool.

To obtain statistically robust results, we evaluate the scenario on all possible tracking period combinations for the pool and the test user. Figure 1 shows an overview of the experimental setup and also shows that the tracking period combinations are not unique. For example, for our total tracking time of 56 weeks there are 14 distinct 4-week periods and 7 distinct 8-week periods. We do not evaluate all 98 possible combinations but restrict ourselves to combinations where the test user is matched to the closest, directly preceding tracking period in the pool. This choice of valid pool and test user pairs is exemplified by the black arrows in Figure 1.

For every valid time bin combination for a given combination of tracking periods, we match every available test user to the users from the pool and evaluate the matching success using the metrics introduced below. All code for the experiments is publicly available[2], however we can not publish the tracking dataset to protect the privacy of the study participants.

2.4 *Metrics for re-identification performance*

To evaluate the success of the matching attack, we employ two metrics: the top-k matching performance and the mean reciprocal rank (MRR) [24]. Both rely on the rank that was assigned to the true match of a test user in the pool as introduced above, $r(u_j^{test})$.

We then report the top-k matching performance in one set of test users D_{test} as

$$Acc(D_{test}, k) = \frac{1}{|D_{test}|} \sum_{u_j \in D_{test}} \mathbb{1}\{r(u_j^{test}) \leq k\}.$$

[2] https://github.com/mie-lab/topology_privacy

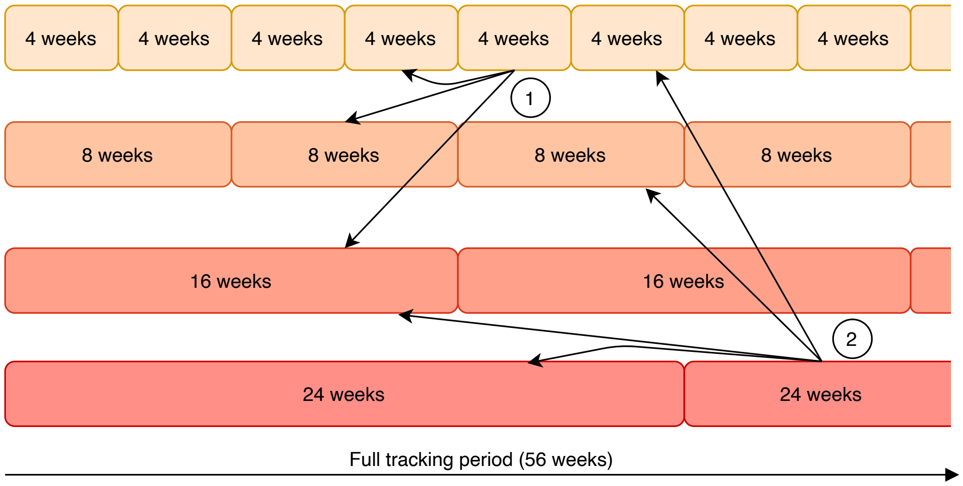


Figure 1: Experimental setup: The tracking data, comprising 56 weeks, are split into non-overlapping bins of varying duration. In the attack scenario, new tracking data from one period is matched to a pool of users at a previous time period. In example 1) the test data of four weeks length can be compared to the pool in the preceding 4 weeks, 8 weeks, or the preceding 16 weeks. In the second example (marked as 2) a test user with tracking data from the second 24-weeks-period is matched to users from the preceding 4-, 8-, 16- or 24-weeks pool.

This considers a match as successful if the true match of the test user is among the top-k closest users in the pool.

Furthermore, we use the MRR as a second evaluation metric, which is defined as the average of the inverse of the ranks in a test dataset. It is a common metric used in information retrieval and re-identification tasks [1]. The MRR of a test set is

$$MRR(D_{test}) = \frac{1}{m} \sum_{u_j \in D_{test}} \frac{1}{r(u_j^{test})}.$$

The MRR can be interpreted as the harmonic mean of the ranks, with the property that good matches (high rank) have much higher influence than bad matches (low rank).

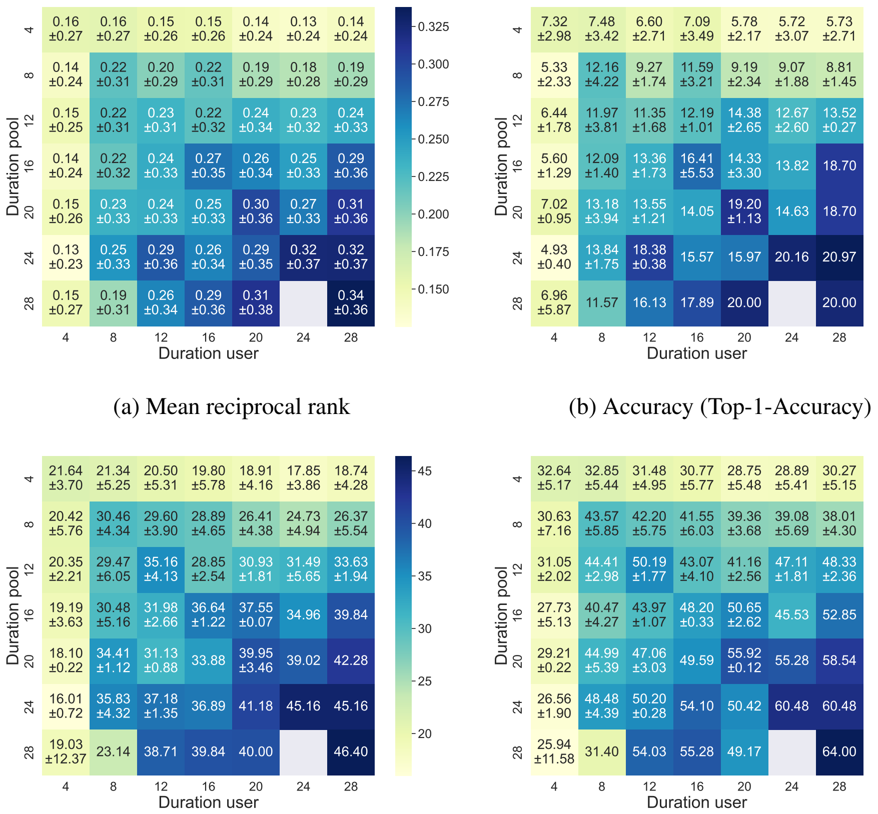
3 Results and Discussion

We run the experiment described in Section 2.3 for all combinations of tracking periods and consecutive start times, adding up to 167 combinations. For each of these combinations, we attempt a matching for every user available in the dataset, which results in over 2'500'000 user-to-user comparisons. We find that the best matching performance is achieved with the combined graph descriptor v_{comb} and the mean squared error (MSE) as the similarity metric d. See Table 2 and Section 3.2 for more details on this choice.

In the following, we report the MRR and top-k matching accuracy for each combination of pool- and test-user tracking duration. If several accuracy results for a tracking period combination are obtained (due to multiple time bin combinations), we report the average result and the standard deviation.

3.1 Effect of tracking period on re-identification performance

Figure 2 shows the average matching performance and the standard deviation for all duration combinations of the pool and the test users. All metrics show a significant dependency on both the duration of the pool and the duration of the test user data. This result shows that privacy-friendly applications should be designed such that their tracking duration is as short



(a) Mean reciprocal rank

(b) Accuracy (Top-1-Accuracy)

(c) Top-5-Accuracy

(d) Top-10-Accuracy

Figure 2: Dependency of matching performance on tracking duration. Top-k accuracy and MRR increase with both the tracking duration of the pool-users as well as the one of the test user.

as possible. This is especially true because a privacy concerned person does not have control over the duration of the pool in our scenario, as the pool represents data that were already collected by a third party.

Furthermore, Figure 2 shows that even for the shortest tracking duration that was tested (i.e., four weeks combined with four weeks), the re-identification capability of our simple matching strategy is significantly better than random. A random rank assignment would result in a top-10 accuracy of 7.6%, compared to the accuracy of 32.6% from the shortest tracking duration. This result shows that the graph representation, even without any additional context or coordinate information, is not anonymous, which is in line with the conclusion reported from [12].

We further analyzed the importance of the pool duration, the test user duration, and the difference of their durations, using linear regression with the duration as the independent variable and the average performance as the dependent variable. The resulting coefficients in Table 1 show that while both duration-variables have a positive impact on the performance, the influence of pool duration is slightly higher. For every additional week of pool tracking duration, the top-10 identification accuracy increases by 0.72%. As the pool is not under the control of the user, a potential solution to minimize the privacy risk is to require data brokers to reset user IDs after a specific tracking period. Notably, Table 1 also shows a major effect from the absolute difference between pool and test tracking duration, corresponding to the strong performance on the diagonals in Figure 2. This can be explained by the higher similarity of graphs that are constructed from the same tracking duration, making it easier to match the correct user.

For the interpretation of the results, it is important to note that the results with small bins are statistically more robust than the results with large bin combinations, because more bins are available. For several combinations of large bins, only one trial was available and therefore no standard deviation was reported, and no distinct time bins are available for the combination of 28 weeks pool duration and 24 weeks test tracking duration.

	pool duration	test duration	absolute difference between pool and test duration	Intercept
1-Accuracy (in %)	0.40	0.29	-0.32	4.36

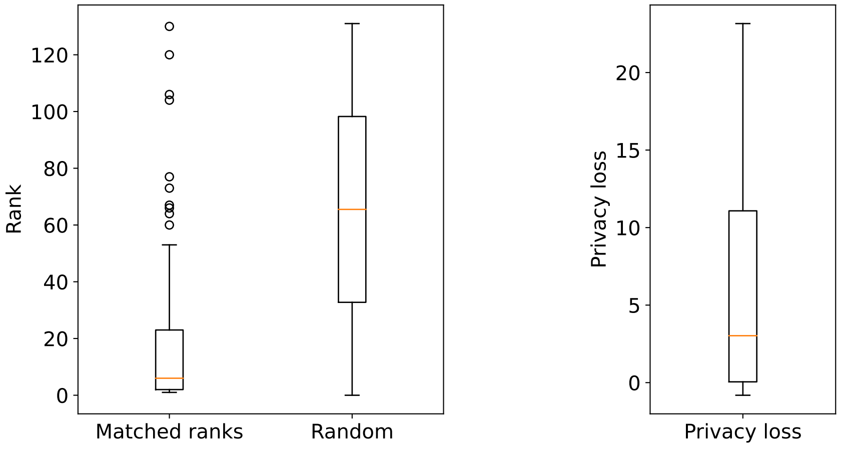
Table 1: Regression analysis of the effect of pool- and test-user tracking duration on the matching performance. Both have a positive effect on the re-identification performance (=negative impact on privacy); however, the effect of the pool duration is slightly higher. The matching performance is higher if the absolute difference between pool and test user duration is low. All results are significant (p-values « 0.01).

3.2 *Features that enable re-identification*

In Section 2.2, we proposed several graph descriptors to calculate a distance between graphs. Table 2 shows the matching performance of different graph features and distance functions. We note that the distance function does not have a strong effect on the matching performance. In contrast, the features result in very different performances in re-identifying graphs. The transition weight and in-degree-distribution are the most useful features, whereas node centrality obtains low matching capability. Based on the results in Table 2, we chose the MSE of all features combined, as this performs best on average according to three out of four error metrics. While our focus is on the time-dependency of privacy preservation, future work could analyze the limits of re-identification of location graphs by using more complex matching methods such as deep graph kernels [27].

		Recip. rank		1-Accuracy		5-Accuracy		10-Accuracy	
Distance metric d	$v(G)$	Mean	Max	Mean	Max	Mean	Max	Mean	Max
KL-divergence	transition	0.13	0.19	5.49	9.60	17.48	28.46	28.18	40.80
	in degree	0.12	0.20	4.47	10.40	15.65	24.80	26.37	41.60
	out degree	0.12	0.17	4.11	7.32	15.14	24.80	26.76	36.29
	shortest path	0.08	0.11	2.36	4.13	9.30	14.52	16.40	26.61
	centrality	0.05	0.06	0.86	2.02	4.34	8.06	9.10	15.32
	combined	0.22	0.35	12.00	23.33	30.34	51.20	42.71	62.40
MSE	transition	0.13	0.19	5.35	9.24	17.00	28.00	28.07	41.60
	in degree	0.12	0.18	4.70	9.76	15.44	22.40	25.47	35.48
	out degree	0.11	0.16	3.62	6.61	14.62	25.60	25.76	39.20
	shortest path	0.08	0.11	2.33	4.04	9.40	15.32	16.83	29.03
	centrality	0.05	0.07	1.39	3.25	6.06	11.16	10.67	16.74
	combined	**0.22**	0.34	**12.31**	20.97	30.61	46.40	**43.25**	**64.00**
Wasserstein distance	transition	0.13	0.19	5.35	9.21	17.73	30.08	27.75	41.60
	in degree	0.12	0.19	4.53	10.40	15.59	25.60	25.84	36.00
	out degree	0.11	0.17	3.93	8.13	14.91	24.00	26.18	40.32
	shortest path	0.07	0.11	2.10	4.20	7.72	16.00	13.82	24.00
	centrality	0.05	0.09	1.38	4.13	5.71	11.38	10.88	16.53
	combined	0.20	**0.36**	10.57	**24.00**	27.62	52.80	38.86	61.60
Sum all metrics	combined	0.22	0.36	12.01	22.50	**30.65**	**52.80**	42.88	62.40

Table 2: Matching performance of different combinations of features, distance functions, and evaluation metrics. The highest matching accuracy is achieved with an R-convolution kernel that computes the MSE between all graph-features distributions combined.



(a) Rank distribution

(b) Privacy loss [12]

Figure 3: Evaluation of rank distribution and privacy loss as proposed by Manousakas et al. [12].

3.3 Validation of matching methodology based on related work

We validate our method by comparing the results to the ones of Manousakas et al. [12]. In their longitudinal study, Manousakas et al. [12] split the tracking data user-wise into two parts at a random point in time, sampled uniformly between 30% and 70% of the whole period (around one year). The most comparable experiment from our study is the one where both the pool and the test duration is 28 weeks. In accordance with the evaluation by Manousakas et al. [12], we show the distribution of ranks and the "privacy loss" in Figure 3. Although the absolute ranks are not informative due to the different pool size (132 users[3] for our dataset versus 1500), the re-identification ability seems comparable. Specifically, the mean of the true rank is shifted from 66 (random) to 17.1 (informed adversary) for our dataset, and from 750 to 140 in their experiment [12, p. 13]. This corresponds to a median privacy loss of 3.03 with our method, while [12] report a "considerable privacy loss with a median of 2.52 [which] means that the informed adversary can achieve a median deanonymization probability 3.52 times higher than an uninformed adversary" [12, p. 14]. Overall, we reproduced the results successfully and extended their results with additional analysis of the impact of the tracking duration.

3.4 Intra user vs. inter user variability of re-identification performance

The main results of this study (Figure 2) are reported as average matching performance. We now further analyze the sources of variance of the matching performance by analyzing the variance of the rank assigned to users during the matching. In particular, we aim to answer the following question: Is the variance due to strong differences between users (e.g., easy-to-match vs. hard-to-match users), or do users have a strongly varying re-identification ability over time? To answer this question, we calculate the standard deviation between different users in the same timesteps (inter-user) and for the same user over several timesteps (intra-user).

Figure 4 shows that the inter-user standard deviation is consistently higher than the intra-user standard deviation. This indicates the existence of user groups that are consistently hard or easy to match. Moreover, the intra-user standard deviation decreases as the tracking duration increases, which can be explained by the higher stability of long-term location graphs. Future work could analyze the factors that characterize hard-to-match or easy-to-match users,

[3] For long time bin durations, not all users matched the criteria set for tracking coverage. Details on the filtering of users are given in Appendix A.

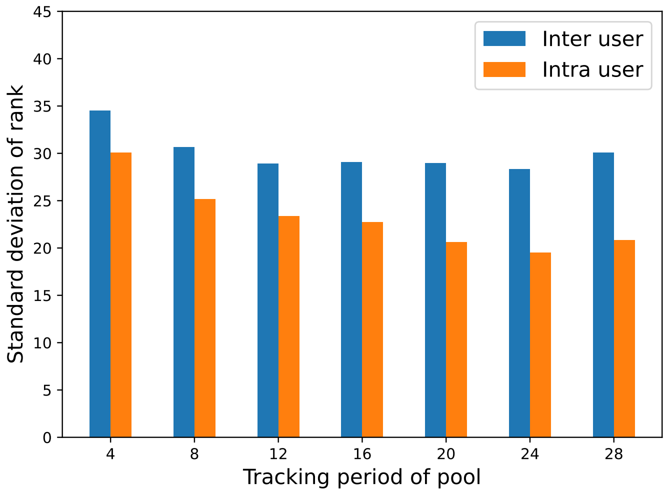


Figure 4: Inter vs. intra person variability of matching performance. The variance over users is higher than the variance over time bins. Intra-user variance decreases with growing tracking duration.

leading to interesting advice for individuals that would like to keep their mobility data hard-to-match.

4 Conclusion

In this work, we present a set of experiments to analyze how tracking duration influences the re-identification ability of individual mobility graphs. The tracking data of a one-year study is binned into time periods of varying size and transformed into location graphs to represent individual mobility behavior. We run a re-identification experiment based on a graph matching task for all pair-wise combinations of time periods.

We can confirm results from Manousakas et al. [12] that location graphs without additional context information are sufficient to re-identify users with a success rate that is significantly higher than random. Furthermore, we show that the re-identification ability increases with increased tracking duration of the pool of candidate users as well as with increased tracking duration of the test user. Therefore, privacy friendly applications should only require tracking data over periods that are as short as possible, and data brokers should be required to reset the user IDs of their data regularly to limit the pool duration.

Furthermore, we analyze the matching result of users for different time steps and showed that users vary in their exposure to be re-identified. Characterization of these user groups should be explored in future work. Another extension of this work would be to further investigate the influence of other dataset properties, such as the tracking intensity (e.g., the number of location points within a unit time), and to collect evidence from more diverse datasets. The latter is straightforward with this method as the individual location graphs have very few requirements (e.g., no specific features or labels needed). Finally, it is important to mention that we only employed a simplistic matching strategy and a more sophisticated matching approach could lead to even higher success rates for matching. The results should therefore be considered as a lower bound of possible matching success. The presented analysis however augments the understanding of the privacy risk of tracking data - even if it is reduced to topology - and can improve the regulation of anonymization practices.

5 Acknowledgement

This work was supported by the ETH Zurich Foundation [MI-01-19] and by the Hasler Foundation through grant 1-008062;

6 CRediT author statement

Martin, Henry: Conceptualization, Methodology, Software, Writing - Original Draft, Writing - Review & Editing; **Wiedemann, Nina**: Conceptualization, Methodology, Software, Visualization, Writing - Original Draft, Writing - Review & Editing; **Suel, Esra**: Conceptualization, Writing - Original Draft, Writing - Review & Editing; **Hong, Ye**: Conceptualization, Writing - Review & Editing; **Xin, Yanan**: Conceptualization, Writing - Review & Editing;

References

[1] Nick Craswell. Mean reciprocal rank. *Encyclopedia of database systems*, 1703, 2009.

[2] Yves-Alexandre De Montjoye, César A Hidalgo, Michel Verleysen, and Vincent D Blondel. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, 3(1):1–5, 2013.

[3] Yoni De Mulder, George Danezis, Lejla Batina, and Bart Preneel. Identification via location-profiling in gsm networks. In *Proceedings of the 7th ACM workshop on Privacy in the electronic society*, pages 23–32, 2008.

[4] Matt Duckham and Lars Kulik. A Formal Model of Obfuscation and Negotiation for Location Privacy. In *Pervasive Computing*, volume 3468, pages 152–170. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005. doi: 10.1007/11428572_10.

[5] Robert W Floyd. Algorithm 97: shortest path. *Communications of the ACM*, 5(6):345, 1962.

[6] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[7] Sébastien Gambs, Marc-Olivier Killijian, and Miguel Núñez del Prado Cortez. De-anonymization attack on geolocated data. *Journal of Computer and System Sciences*, 80(8):1597–1614, 2014.

[8] Philippe Golle and Kurt Partridge. On the anonymity of home/work location pairs. In *International Conference on Pervasive Computing*, pages 390–397. Springer, 2009.

[9] David Haussler. Convolution kernels on discrete structures. Technical report, Department of Computer Science, University of California, 1999.

[10] Ammar Haydari, Michael Zhang, Chen-Nee Chuah, Jane Macfarlane, and Sean Peisert. Adaptive Differential Privacy Mechanism for Aggregated Mobility Dataset. *arXiv:2112.08487 [cs]*, December 2021.

[11] Carsten Keßler and Grant McKenzie. A geoprivacy manifesto. *Transactions in GIS*, 22 (1):3–19, 2018.

[12] Dionysis Manousakas, Cecilia Mascolo, Alastair R Beresford, Dennis Chan, and Nikhil Sharma. Quantifying privacy loss of human mobility graph topology. *Proceedings on Privacy Enhancing Technologies*, 2018(3):5–21, 2018.

[13] Henry Martin, Dominik Bucher, Esra Suel, Pengxiang Zhao, Fernando Perez-Cruz, and Martin Raubal. Graph convolutional neural networks for human activity purpose imputation. In *NIPS spatiotemporal workshop at the 32nd Annual conference on neural information processing systems (NIPS 2018)*, 2018.

[14] Henry Martin, Henrik Becker, Dominik Bucher, David Jonietz, Martin Raubal, and Kay W Axhausen. Begleitstudie SBB Green Class-Abschlussbericht. *Arbeitsberichte Verkehrs-und Raumplanung*, 1439, 2019.

[15] Henry Martin, Ye Hong, Nina Wiedemann, Dominik Bucher, and Martin Raubal. Trackintel: An open-source python library for human mobility analysis, 2022. URL https://arxiv.org/abs/2206.03593.

[16] Steven Melendez and Alex Pasternack. Here are the data brokers quietly buying and selling your personal information. *The Fast Company*, 2019. URL https://www.fastcompany.com/90310803/here-are-the-data-brokers-quietly-buying-and-selling-your-personal-information.

[17] Luca Pappalardo, Filippo Simini, Salvatore Rinzivillo, Dino Pedreschi, Fosca Giannotti, and Albert-László Barabási. Returners and explorers dichotomy in human mobility. *Nature Communications*, 6(1):8166, November 2015. ISSN 2041-1723. doi: 10.1038/ncomms9166.

[18] Vincent Primault, Antoine Boutet, Sonia Ben Mokhtar, and Lionel Brunie. The long road to computational location privacy: A survey. *IEEE Communications Surveys & Tutorials*, 21(3):2772–2793, 2018.

[19] Martin Raubal, Dominik Bucher, and Henry Martin. Geosmartness for personalized and sustainable future urban mobility. In *Urban Informatics*, pages 59–83. Springer, 2021.

[20] Kaspar Riesen, Xiaoyi Jiang, and Horst Bunke. Exact and inexact graph matching: Methodology and applications. In *Managing and Mining Graph Data*, pages 217–247. Springer, 2010.

[21] Salvatore Rinzivillo, Lorenzo Gabrielli, Mirco Nanni, Luca Pappalardo, Dino Pedreschi, and Fosca Giannotti. The purpose of motion: Learning activities from individual mobility networks. In *2014 International Conference on Data Science and Advanced Analytics (DSAA)*, pages 312–318. IEEE, 2014.

[22] Luca Rossi, James Walker, and Mirco Musolesi. Spatio-temporal techniques for user identification by means of GPS mobility data. *EPJ Data Science*, 4(1):11, December 2015.

[23] Stuart A. Thompson and Charlie Warzel. The privacy project: Twelve million phones, one dataset, zero privacy. *The New York Times*, 2019. URL https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html.

[24] Ellen M Voorhees et al. The trec-8 question answering track report. *Trec*, 99:77–82, 1999.

[25] Leye Wang, Gehua Qin, Dingqi Yang, Xiao Han, and Xiaojuan Ma. Geographic Differential Privacy for Mobile Crowd Coverage Maximization. *arXiv:1710.10477 [cs]*, November 2017. URL http://arxiv.org/abs/1710.10477. arXiv: 1710.10477.

[26] Stephen Warshall. A theorem on boolean matrices. *Journal of the ACM (JACM)*, 9(1):11–12, 1962.

[27] Pinar Yanardag and SVN Vishwanathan. Deep graph kernels. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1365–1374, 2015.

[28] Hui Zang and Jean Bolot. Anonymization of location data does not work: A large-scale measurement study. In *Proceedings of the 17th annual international conference on Mobile computing and networking*, pages 145–156, 2011.

A Preprocessing

The tracking data is provided from the MyWay app as staypoints, labeled with activities, and triplegs, labeled with transport modes. The staypoints are clustered into locations with the DBSCAN algorithm with the parameter $\varepsilon = 30m$ and a minimum number of 1 point per cluster, i.e. each staypoint is assigned to a location. The Trackintel library merges consecutive staypoints and triplegs into trips as long as they are not interrupted by an activity (staypoints with duration >25 min or labeled with a purpose other than wait, unknown) or by a temporal gap (here 25 minutes). Finally, when constructing the graph, we filter out users with low tracking coverage in the selected time period. The users are required to have a tracking coverage of at least 70% in at least one third of the days. In our experiments, this leads to a varying number of 132-137 users depending on the time periods used.