Security and Privacy

Topics · https://mobidatalab.eu/knowledge-base/security-and-privacy/

Deliverable D2.3, “State of the Art on Transport and Mobility Data Protection Technologies,” offers an overview of the privacy risks related to mobility and transportation data and provides a description to several techniques in the literature aiming at limiting or eliminating such risks.



What is Mobility Data?

Mobility data in its simpler form are data about individuals that include their locations at specific times. Sources of real-time raw individual location data include, but are not limited to, cell towers, Wi-Fi access points, RFID tag readers, location-based services, or credit card payments. Historical location data, in form of data sets in which each of the records corresponds to an individual and includes her location data for some time periods are referred to as trajectory microdata sets.

Such trajectory microdata sets are often of interest to transport authorities, operators, and other stakeholders to evaluate and improve their services, the state of the traffic, etc. and thus are often publicly released or shared. Sharing of mobility data is occasionally shared as aggregates (*e.g.,* heat maps) instead of at an individual level. For example, recently, due to the COVID-19 pandemic, the health authorities have also become interested in mobility data to predict the spread of infectious diseases.

What are the risks?

Whichever the form of these mobility data, they all share some statistical characteristics that make their sharing a potential privacy risk. Mobility data are highly *unique* and *regular.*

Unicity refers to the data of different individuals to be easily differentiable, particularly at some specific locations. The starting and ending locations of users' trajectories are often their home and work locations which, again are highly unique and can lead to reidentification. Studies show that user full trajectories can be uniquely recovered with the knowledge of only two locations.

The *regularity* of trajectories means that for single individuals, their data follows periodic patterns. Namely, individuals tend to follow the same trajectories during workdays—home to work and back to home.

Trajectory microdata is prone to privacy attacks on individual users because of these two defining characteristics. It is hard to anonymize. Experiments show that 50% of mobile subscribers in a database with 25 million different users can be uniquely detected with the knowledge of only 3 frequent locations. Additionally, experiments have shown that only very coarse generalization (e.g., at the city level) produces reliable anonymized trajectory microdata sets.

The difficulty in assessing the privacy risk of one's own data and the lack of appropriate anonymization mechanisms are two of the main barriers to mobility data sharing.

How to ensure the protection of mobility data? Data anonymization.

General Data Protection Regulation (GDPR) states that data controllers and processors must implement appropriate technological and organizational measures to secure personal data. Data encryption is the best way to secure stored personal data. However, it is not feasible for releasing trajectory datasets to research organizations or for making them available to public. In all these cases, it is important that when such data contain information about individuals some measures are taken to prevent the release of sensitive information about such individuals. These measures are collectively known as anonymization measures. Anonymous data is not personal data and hence falls outside the scope of the GDPR.

Anonymization measures aim to produce modified versions of the original datasets which are as close as possible to the original datasets but that do not allow attackers to identify individuals or to infer confidential attributes about concrete individuals. An effective anonymization solution should prevent all parties from singling out an individual in a dataset, from linking two records within a dataset (or between two separate datasets) and from inferring any information in such dataset. Therefore, removing directly identifying elements (elements that unambiguously identify individuals, for example, a passport number) is not enough to ensure that identification of the data subject is no longer possible. This is even more important to be taken into account for mobility data.

In the next sections, we give an overview of how anonymization mechanisms work with all kinds of datasets (not just mobility datasets). All of these concepts can then be applied to mobility datasets.

Attacks to datasets

Before introducing the types of attacks that can be performed on a dataset, we should provide some useful background.

A dataset can be seen as a matrix with n rows that correspond to single individuals and k columns, which correspond to attributes about the individuals. Attributes in a dataset can be classified as follows:

- **Identifiers** are attributes that unambiguously identify individuals, for example, a passport number.
- **Quasi-identifiers** are attributes that, while not being identifiers themselves, can identify single persons when combined. Examples of this are date of birth, city of birth or, in location data,

position of home and position of workplace.

- **Confidential attributes** are those which contain sensitive information about an individual, for example the salary, religion, or political affiliation.
- **Non-confidential attributes** are the rest of attributes about an individual that do not impact on their privacy.

We consider attacks in which an attacker, with some background knowledge about some individual (e.g., knows some quasi-identifiers of the victim, such as their age or their place of birth) can identify such individual in a released dataset and, thus, learning some additional information about the individual. Note that the presence of the individual in a dataset might itself be a violation of the individual's privacy. For example, if an individual is found in a dataset containing cancer treatments, one can infer that such individual is being treated for cancer, and thus that the individual has cancer. The risks here is that the identity of individuals or some attributes of such individuals are disclosed.

In **record linkage attacks**, the attacker knows some quasi-identifiers about some victim and checks for combinations of quasi-identifier attribute values in the released data base. If the attacker can find a small group of individuals with the same combination of quasi-identifier attribute values as her victim, then the attacker is able to link a record in the dataset to her victim with high confidence. For example, if we release a dataset with the salary of all our employees along with their gender, age and marital status, it can be easy to find a combination of these attributes that is shared by a very small number of employees or worse, just one employee (e.g. a young widow in a development company).

Attribute linkage attacks are those in which an attacker learns some information on a confidential attribute from her target individual. Again, the attacker knows some quasi-identifier attribute values of her victim and proceeds in the same way as in the record linkage attacks. In case all individuals that share the same quasi-identifiers also share some confidential attribute, the attacker learns that confidential attribute even if he cannot pinpoint exactly who the individual is. In the example above, if we found a combination shared by 10 employees but they all have very similar salaries, their privacy is still broken.

Utility and privacy

Anonymization mechanisms modify the original data to prevent disclosure of personal information, both of identities and their related confidential attributes. This modification may hurt the utility of the

data. Anonymization methods aim to protect the privacy of the respondents while producing as little as possible effects on the utility of the data. As utility, we may refer to the differences between the original and the modified data or to the differences in the results obtained from some processing on the original and the modified data (e.g., a heatmap computed on the original mobility dataset versus a heatmap computed on the anonymized mobility dataset).

Generic measures for data utility compare different statistics between the original dataset and the anonymized dataset. Some of these statistics include the means and covariances of some attribute subsets, and covariance and correlation matrices. Other measures include the mean absolute or mean squared errors between attributes in the original and anonymized datasets, or distance metrics between the empirical distributions of the attributes.

When the analyses to be performed on data are known before any anonymization measure is applied, one can measure the utility of the anonymization process by conducting such analyses on the original and the anonymized datasets and compare the results using some distance metric.

When trying to balance the utility of the anonymized datasets and the privacy guarantees offered by the anonymization measures taken, one can follow two approaches:

- In utility-first anonymization, the objective is to cause the minimum disruption to the dataset while minimizing the disclosure risks. First, the data is transformed using some anonymization method with some parameters and then the identity or attribute disclosure risks are calculated. If the risk is considered too high, the anonymization methods are re-run using more strict parameters. This process is repeated until the disclosure risks are considered low enough.
- In privacy-first anonymization, some privacy model is enforced, such as k-anonymity or ϵ-differential privacy, that ensure some boundaries on the risks of re-identification or attribute disclosure. Anonymization methods in this case are model dependent, and the parameters are derived from the model.

Privacy Models

k-Anonymity and extensions

A well-known privacy model is k-anonymity , which requires that each tuple of quasi-identifier attribute values be shared by at least k records in the database. This condition may be achieved

through generalisation and suppression mechanisms, and through microaggregation.

Unfortunately, while this privacy model prevents identity disclosure, it may fail to protect against attribute disclosure. The definition of this privacy model establishes that complete re-identification is unfeasible within a group of records sharing the same tuple of perturbed quasi-identifier attribute values. However, if the records in the group have the same value (or very similar values) for a confidential attribute, the confidential attribute value of an individual linkable to the group is leaked.

To fix this problem, some extensions of k-anonymity have been proposed, the most popular being l-diversity and t-closeness . The property of l-diversity is satisfied if there are at least *l* 'well-represented' values for each confidential attribute in all groups sharing the values of the quasi-identifiers. The property of t-closeness is satisfied when the distance between the distribution of each confidential attribute within each group and the whole dataset is no more than a threshold *t*.

ϵ-Differential Privacy and relaxations

Another important privacy model is differential privacy. This model was originally defined for queryable databases and consists in perturbing the original query result of a database before outputting it. This may be viewed as equivalent to perturbing the original data and then computing the queries over the modified data. Thus, differential privacy can also be seen as a privacy model for microdata sets.

Methods

In Statistical Disclosure Control, masking refers to the process of obtaining an anonymised dataset *X'* by modifying the original *X*. Masking can be perturbative or non-perturbative. In the former approach, the data values of *X* are perturbed to obtain *X'*. In contrast, in nonperturbative masking *X'* is obtained by removing some values and/or by making them more general; yet the information in *X'* is still true, although less detailed; as an example, a value might be replaced by a range containing the original value.

Perturbative masking

Perturbative masking generates a modified version of the microdata set such that the privacy of the respondents is protected to a certain extent while simultaneously some statistical properties of the data are preserved. Well-known perturbative masking methods include:

- **Noise addition.** This is the most popular method, which consists in adding a noise vector to each record in the dataset. The utility preservation depends on the amount and the distribution of the noise.
- **Data swapping.** This technique exchanges the values of the attributes randomly among individual records. Clearly, univariate distributions are preserved, but multivariate distributions may be substantially harmed unless swaps of very different values are ruled out.
- **Microaggregation.** This groups similar records together and releases the average record of each group. The more similar the records in a group, the more data utility is preserved.

Non-perturbative masking

Common non-perturbative methods include:

- **Sampling.** Instead of publishing the whole dataset, only a sample of it is released.
- **Generalisation.** The values of the different attributes are recoded in new, more general categories such that the information remains the same, albeit less specific.
- **Top/bottom coding**. In line with the previous method, values above (resp. below) a certain threshold are grouped together into a single category.
- **Local suppression.** If a combination of quasi-identifier values is shared by too few records, it may lead to re-identification. This method relies on replacing certain individual attribute values with missing values, so that the number of records sharing a particular combination of quasi-identifier values becomes larger.

Synthetic microdata generation

An anonymisation approach alternative to masking is synthetic data generation. That is, instead of modifying the original data set, a simulated dataset is generated such that it preserves some properties of the original data set. The main advantage of synthetic data is that no respondent re-identification seems possible since the data are artificial. However, if, by chance, a synthetic record is very close to an original one, the respondent of the latter record will not feel safe when the former record is released. In addition, the utility of synthetic data sets is limited to preserving the statistical properties selected at the time of data synthesis.

Some examples of synthetic generation include methods based on multiple imputation and methods that preserve means and co-variances. An effective alternative to the drawbacks of purely synthetic

data are hybrid data, which mix original and synthetic data and are therefore richer. Yet another alternative is partially synthetic data, whereby only the most sensitive original data values are replaced by synthetic values.

Anonymization of mobility datasets

A trajectory microdata set is a microdata set that contains trajectory data. Thus, each row in the dataset represents an individual and each column represents an attribute of the individual. At least one of the attributes must correspond to a trajectory, that is, a list of spaciotemporal points. Additionally, other columns may correspond to other attributes of the individual.

Trajectory microdata sets are special because the location information included in them can be considered both as quasi-identifiers and sensitive information. Let us illustrate this using an example. The location information of a certain individual contains several recurring positions, some of them occur at night-time hours and some of them at daytime hours. Most likely, those positions correspond to the place of residence and the workplace of the individual, respectively. It is also very likely that the combination of place of residence and workplace is unique (although it may not be the case in all situations), and so the individual can be uniquely identified. This would allow an attacker to perform a record linkage attack if, for example, the attacker has background information on the workplace of their victim. If the location information also includes recurring visits to a medical institution, then we also have attribute disclosure, since the attacker can be confident that the individual has some chronic medical condition (if the medical institution is specialized in some kinds of conditions, then the information the attacker gets is more precise).

As we said before, the uniqueness of trajectory data makes it very difficult to anonymize while preserving the utility of the data.

Attacks to mobility datasets

One of the main risks in the sharing or release of a microdata set, be it a statistical database or a database with mobility data is that of record linkage. In a record linkage attack, an attacker attempts to uniquely match record in a published dataset to a record previously (partially) known by the attacker. The previous knowledge of the attacker might determine the kind of protection necessary to avoid such attacks.

Some sections above, we introduced the concepts of identifiers, quasi-identifiers, and confidential attributes in microdata sets. In the case of mobility datasets, we should include one more kind of attribute:

- **Location-based quasi-identifier**, a certain combination of spatiotemporal points that uniquely identify a record in a database or an individual

A straightforward **record linkage attack** is possible when an attacker knows a subset of locations visited by her victim, or a subtrajectory of a full trajectory of a user. Due to the huge unicity of trajectory data, a small subset of positions can be enough to single out individuals. Methods to protect based on k-anonymity or similar try to make any subset of positions (any subtrajectory) be shared by at least *k* individuals, reducing the unicity of the data. It is also possible that the attacker has information on the trajectory of her victims, potentially with the same granularity as in the target database, but with different sampling times. For example, an attacker might obtain mobility information from geo-tagged photos in her victim's social media and intend to find her complete record in a target database (possibly to learn some confidential attribute of her victim). The information the attacker has might have the same granularity, but the sampling times are different.

Instead of exploiting the unicity of mobility data, an attacker might exploit its regularity. The attacker could know a Markovian model of the pathing of her victim, that is, it knows a model that indicates the probabilities of moving from some position to another, irrespective of all previous or future positions. After accessing the released mobility dataset, the attacker can build a similar Markovian model for each of the individual records comparing them to his own. The authors who presented this attack reported an 80% success rate on a dataset of 100 individuals.

While record linkage attacks rely on the unicity of mobility microdata to single out individuals, **attribute linkage attacks** exploit the homogeneity of the data to discover sensitive information about the target individual. As an example, consider a mobility dataset that has been anonymized using some method based on k-anonymity. In this example, an adversary is not capable of linking her victim to an individual record in the released dataset since the background information of the attacker matches *k* or more different individuals. However, if the confidential attributes of the *k* individuals (which may also be some location) are highly homogeneous, then the attacker might infer the value for this attribute with high confidence.

Privacy methods for anonymizing mobility datasets

In this section we briefly introduce some anonymization methods for mobility data presented in the literature.

Disclosure risk mitigation

Mitigation strategies follow the utility-first anonymization approach. These mechanisms do not provide any formal privacy guarantees but aim to reduce reidentification risks by applying different techniques, such as noise addition, generalization and coarsening with heuristic parameter choice. After the application of such techniques, the disclosure risk is calculated (for some objective disclosure prevention, i.e., identity disclosure or attribute disclosure). If the obtained risk is still too high, the techniques are applied with more strict parameters. Several of the following techniques can be applied both in the context of location-based services and in static trajectory microdata sets:

- One of the first mechanisms introduced within the general anonymization literature, and within the location privacy literature is noise addition or obfuscation. Agrawal and Srikant[1] introduced a method for privacy-preserving data mining in which users introduce noise drawn from a uniform or Gaussian distribution to their sensitive attributes. Later, the data aggregator can reconstruct the original distribution from the noisy data to train a classifier. Note that techniques to enforce ϵ-differential privacy use noise addition, but differential privacy calibrates the added noise to achieve strict privacy guarantees.
- Cloaking techniques aim at reducing the granularity of the location data, both temporal and spatial. Hoh et al.[2] reduce the temporal granularity to prevent home identification. First, they mount an attack to identify the home location from 239 mobility traces spanning one week and with sample frequency of 1 location per second while the vehicles are switched on. The authors report 85% of the homes are identified in this scenario. Then, 75% of the points in the trajectories are dropped (sample frequency is reduced to 1 location per 4 seconds), resulting in a home identification rate of 40%.
- Song et al.[3] proposed the segmentation of trajectories. The authors measure the risk of reidentification by a uniqueness metric, that is the fraction of individuals in the dataset which are uniquely identifiable by a set of spatiotemporal points. To reduce uniqueness, the trajectories of users are split and assigned to new different identifiers, mimicking partially consistent protection mechanisms for location-based services. While the method, reduces the uniqueness of trajectories, but anonymized trajectories remain highly unique. These methods do not perturb any of the spaciotemporal information but may reduce the utility of the data sets when the intended

analysis requires the mobility traces for users over long periods of time.

- SwapMob[4] can also be used to protect static trajectory microdata sets. In SwapMob, partial trajectories are swapped among users after they meet at some point.

Indistinguishability-based methods

This strategy aims at reducing or eliminating the uniqueness in the quasi-identifiers (in this case, the location-based quasi-identifiers) and typically are based on k-anonymity and its extensions. By applying techniques such as noise addition, generalization, location coarsening, and microaggregation, these strategies try to produce groups of individuals for which their quasi-identifiers are the same, so that an adversary with some background information cannot distinguish between individuals in such groups.

- **GLOVE:** Gramaglia and Fiore[5] study the difficulty of enforcing k-anonymity in trajectory databases and propose the GLOVE algorithm, based on a specialized generalization technique. The proposed algorithm aims to protect respondents against record linkage attacks and does not make any assumptions on the attackers' knowledge (attackers might know full trajectories of their victims). The authors' criterion towards indistinguishability of records is k-anonymity, and aim at anonymizing full trajectories of respondents, using generalization and (possibly) suppression of spaciotemporal points.
- **KAM:** Monreale *et al.*[6] propose an anonymization method based on k-anonymity through the generalization and suppression of spatial points (this approach does not consider time). In this case, the authors assume that an attacker may know a subtrajectory (a trajectory contained in some other trajectory or trajectories) of her victim.
- **NWA/W4M:** Abul et al. introduces a trajectory database anonymization method based on a relaxation of k-anonymity named (k,δ)-anonymity, whereby the indistinguishability requirement among k entries in the database required for k-anonymity is relaxed by some uncertainty δ.
- **SwapLocations:** Domingo-Ferrer *et al.*[7] propose trajectory database anonymization mechanism, called SwapLocations, based on ***k***-anonymity that does not introduce fake, perturbed, or generalized trajectories. The method is based on microaggregation of trajectories, but instead of substituting the trajectories of k-anonymous sets by the cluster representatives, locations are swapped among individuals in the *k*-anonymous sets.

Generation of Synthetic Trajectory Microdata

Deep learning (DL)-based methods aim to generate synthetic trajectories that can realistically reproduce the patterns of individuals' mobility. The intuition is that the generated synthetic data come from the same distribution of real trajectories (thereby preserving utility). At the same time, they do not correspond to real trajectories (thereby preserving privacy).

Existing DL methods leverage sequence natural language processing (NLP) models, such as RNNs, or generative models, such as generative adversarial networks (GANs), to approximate the distribution of the real trajectory data and then sample synthetic trajectories from that distribution. Following you can find some examples of methods using NLP models:

- Kulkarni and Garbinato[8] exploit the ability of RNNs to model problems over sequential data having long-term temporal dependencies. Like training a next-word prediction model, they train a next location prediction model using the real trajectory data as training data. Then, they construct a synthetic trajectory by starting at some arbitrary location and iteratively feeding the current output trajectory sequence as input to the next step in the trained model.
- In a work carried out within the **MOBIDATALAB project**, Blanco et al.[9] show preliminary work on the generation of synthetic trajectory microdata using machine learning models typically used for natural language processing and time series. In this case, they use Bidirectional LSTM (BiLSTM) models.

GANs set up a game between two neural networks: the generator G and the discriminator D. G's goal is to generate "synthetic" data classified as "real" by D, whereas D's goal is to correctly distinguish between real and synthetic data and provide feedback to G to improve the realism of the generated data.

- trajGANs[10] consists of a generator G which generates a dense representation of synthetic trajectories from a random input vector z and a discriminator D, which classifies input trajectory samples as "real" or "fake".
- SVAE[11] builds its generator G based on an LSTM and a Variational Autoencoder (VAE) to combine the ability of LSTMs to process sequential data with the ability of VAEs to construct a latent space that captures key features of the training data.
- MoveSim[12] uses a self-attention-based sequential model as a generator to capture the temporal transitions in human mobility.

The anonymization module of the MobiDataLab Transport Cloud prototype

The CRISES research group at URV has developed an anonymization module within the Mobidatalab project that allows users to anonymize a mobility dataset, to perform an analysis in a private-way and to compute some utility and privacy measures over both the original and the anonymized datasets in a straightforward way.

The final version of the anonymization module includes 6 anonymization methods, 1 privacy-preserving analysis method and 5 methods to compute different utility and privacy metrics. It also provides a command line interface (CLI) that allows users to use all the module functionalities in a straightforward way. The module is also ready to be deployed in a server and to process requests through an API.

The anonymization module has been designed with a focus on modularity, where pseudonymization or anonymization methods can be built using different components dedicated to preprocessing, clustering, distance computation, aggregation, etc. We have focused on making it easy to add new methods and components, in order to encourage contributions from other researchers.

The module is available at https://github.com/MobiDataLab/mdl-anonymizeralong with a detailed documentation.

[1] Agrawal, R. and Srikant, R., 2000, May. Privacy-preserving data mining. In Proceedings of the 2000 ACM SIGMOD international conference on Management of data (pp. 439-450).

[2] Hoh, B., Gruteser, M., Xiong, H. and Alrabady, A., 2006. Enhancing security and privacy in traffic-monitoring systems. IEEE Pervasive Computing, 5(4), pp.38-46.

[3] Song, Y., Dahlmeier, D. and Bressan, S., 2014, January. Not so unique in the crowd: a simple and effective algorithm for anonymizing location data. In PIR@ SIGIR.

[4] Salas, J., Megías, D. and Torra, V., 2018, September. SwapMob: Swapping trajectories for mobility anonymization. In International Conference on Privacy in Statistical Databases (pp. 331 -346). Springer, Cham.

[5] Gramaglia, M. and Fiore, M., 2015, December. Hiding mobile traffic fingerprints with glove. In Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies (pp. 1–13).

[6] Monreale, A., Andrienko, G.L., Andrienko, N.V., Giannotti, F., Pedreschi, D., Rinzivillo, S. and Wrobel, S., 2010. Movement data anonymity through generalization. Trans. Data Priv., 3(2), pp.91–121.

[7] Domingo-Ferrer, J. and Trujillo-Rasua, R., 2012. Microaggregation-and permutation-based anonymization of movement data. Information Sciences, 208, pp.55-80.

[8] Kulkarni V, Garbinato B. “Generating synthetic mobility traffic using RNNs.” In Proceedings of the 1st Workshop on Artificial Intelligence and Deep Learning for Geographic Knowledge Discovery 2017 Nov 7 (pp. 1-4).

[9] Blanco-Justicia A., Jebreel N., Manjón J.A and Domingo-Ferrer J., “Generation of Synthetic Trajectory Microdata from Language Models”, Privacy in Statistical Databases-PSD 2022, Paris, France, In Lecture Notes in Computer Science vol. 13463, pp. 172-187, ISBN: 0302-9743, Sep 2022.

[10] Xi L, Hanzhou C, Clio A. “trajGANs : Using generative adversarial networks for geo-privacy protection of trajectory data.” Vision paper 2018.

[11] Huang D, Song X, Fan Z, Jiang R, Shibasaki R, Zhang Y, Wang H, Kato Y. “A variational autoencoder based generative model of urban human mobility.” In 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) 2019 Mar 28 (pp. 425-430). IEEE.

[12] Feng J, Yang Z, Xu F, Yu H,Wang M, Li Y. “Learning to simulate human mobility.” In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining 2020 Aug 23 (pp. 3426-3433).