Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Computer Science and Statistics

Privacy Aware Travel Assistant

Manasi Rane

June 14, 2022

A dissertation submitted in partial fulfilment
of the requirements for the degree of
MSc (Computer Science - Data Science)

Declaration

I hereby declare that this dissertation is entirely my own work and that it has not been submitted as an exercise for a degree at this or any other university.

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at `http://www.tcd.ie/calendar`.

I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at `http://tcd-ie.libguides.com/plagiarism/ready-steady-write`.

Signed: _____ Date: _____

Abstract

In the era of digital assistants, Travel assistants are widely used in everyday travel to easily access the route updates. On-route travel assistants requires continuous user location updates from the start of the journey, during the journey and till the end. Currently there are various approaches to protect the static user location such as location obfuscation, anonymity and perturbation but the constant updating makes location correlation protection more challenging.

To address this issue, we propose using the Differential Privacy with RAPPOR mechanism to construct a privacy protection mechanism that is based on user locations anticipated using a Markov model rather than actual user locations. This aids in ensuring the user's location privacy while also maintaining the linkage between them, hence increasing the effectiveness of the system.

Acknowledgements

Contents

List of Figures

List of Tables

Nomenclature

1 Introduction

Humans have been superseded by digital assistants in almost every industry, and travel is one of them. Travel assistants are smarter versions of travel planners, offering real-time updates on the routes chosen. Passengers or users can get real-time information about the route, such as the estimated time to the next stop on the route or the estimated time to the traveller's destination, via travel assistants. Giannopoulos in his paper (7) has listed multiple benefits of on-route planning such as the possibility to induce people to travel at non-peak hours, use more 'traffic friendly' destinations, share vehicles with other travellers, reducing journey-time uncertainty, making real-time route changes to avoid congestion, choosing the most appropriate connections/interchanges thus maximising use of spare capacity, minimising waiting times, and increasing 'Inter-modality' in passenger transport services. Thus, with so many advantages, a Travel Assistant is a necessity to keep up with today's fast-paced world.

While location-based services (LBS) have shown to benefit both individuals and society, the increasing exposure of users' location data creates serious privacy concerns. The user must provide his current location information to the LBS server in order to receive the required service. Users may face major threats if this information enters into the hands of malevolent opponents.

When considering a Travel planner the privacy protection mechanisms are only limited to the obfuscation or perturbation of the location data used only when the service is queried but when it comes to a Travel assistant, the location data is updated continuously to the LBS server and must be protected from the beginning of the journey, during the journey till the end of the journey. Also the continuous update of the location enhances the correlation between the various location data making defense against location correlation more challenging.

There are various approaches in which LBS servers protect the static location queries including encryption, anonymity, etc. These methods allow the data collectors to collect the original and accurate data and implement location obfuscation, perturbation and encryption techniques helping them to draw conclusions and overall statistics thus protecting the individual data from being stored. But even the LBS server is also considered to be an

adversary since users nowadays are not willing to provide the accurate location even for statistics collection.

Differential privacy resolves the trust issues with LBS servers perturbing the data before sending it to the data collectors. But adversaries are able to infer the perturbed data as well if the same data is queried multiple times from the database. Google has developed the RAPPOR mechanism which perturbs the location data at two different layers making it more difficult for the adversaries to infer the original value. However the increasing correlation amongst the location values allows the adversary to infer background knowledge more than expected.

In this paper, we suggest using Markov model to predict the locations the user might travel slowing the correlation of the original location data thus boosting the RAPPOR mechanism. The Markov model predicts data based on the current state and thus reduces the correlation on the previous states protecting the location correlation.

The main contributions of this paper are as following:

1. Implementing first order Markov model as the location prediction algorithm which predicts the locations of the users based only on the current state and not on the previous states protecting the auto correlation of the continuously updated location data.

2. Implementing Local Differential Privacy with RAPPOR mechanism as the location protection algorithm which provides both one-time and longitudinal privacy framework.

3. Visualizing and understanding the trade off between utility and privacy along with management of the privacy budget parameter ϵ value.

2 State of the Art

2.1 Travel Planners

Currently there are various web-mapping services like Google Maps which provide the shortest routes with the real time updates. **Map Quest** (12) is one such free web-mapping service which uses current location based on user's IP address (if using the web) or phone GPS (if using mobile). Along with route mapping, it assists travellers with various details such as the estimated fuel costs, hotels, shops and gas stations.

Google Maps introduced in 2005 by Google Inc. is being used widely for travelling from one point to another anywhere throughout the whole world. Google Maps representations are considered to be very accurate as compared to many other routing applications such as Map Quest. Google Maps collects and stores data about the users such as the search terms entered, IP address and the latitude and longitude co-ordinates. Google uses encryption for data privacy when in transit and has secured access. It sets a cookie NID in user's browser to optimise services and to provide user personalized services (8). Users can protect the data shared with the Google Maps in different ways such as by opening the Google Maps with incognito mode in browser or using Google Maps without signing in. Also, the privacy settings can be modified to control what data can be shared with the user. Sharing less data, on the other hand, results in fewer tailored updates.

But why do these services store user's details? Google mentions in (8) that it might share user's personal information with other companies, organizations or individuals outside of Google for different reasons. Some free services share location to generate revenue with location-based advertisements (6).

2.2 Travel Assistants

Personalized Travel assistant (PTA) differs from typical map-based direction finders or trip planners by offering a 'virtual assistant' to access real-time traffic and transportation information, choice of transport and user-specific travel guidance such as allowing the user to decide whether time or expense is more important. PTA continuously collects current

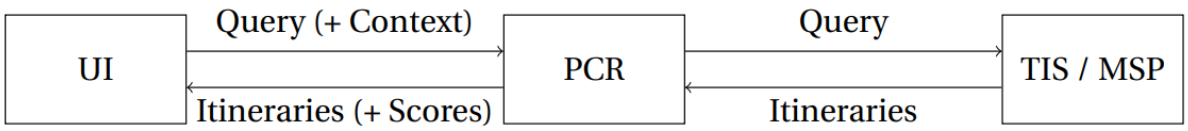

Figure 2.1: Travel information architecture with PCR.

(13)

location data of the user and updates the traffic details based on the location.

Seoul PTA (3) incorporates "virtual assistant" features that provide transit guidance based on user preferences and trip context (for instance, whether time is more critical than expense for a particular trip), employing real-time traffic and public transportation information. It consists of 4 key features: Transport Information Service (Public transportation details), Personal Travel Planner (considering user preferences), Carbon Calculator (to manage carbon footprint and travel modes) and Real-Time Router (reroute during trip in case of disruptions).

In (13), Christian Samuel has proved with the evaluations conducted that personalized travel planning is preferred over a traditional departure time-based travel itinerary sorting. The paper explains architectural changes to include integrated inter modal travel information services i.e., new components and interfaces. It mentions the issues faced when multiple interfaces are combined to create an integrated itinerary at one place. Here inter modal includes public transport, personal cars, walking, car-sharing as well.

Personalized Travel Planning is implemented here in multiple ways by considering the user preferences such as car-sharing, user's weather specific requirements, user's age, etc with the help of various recommender systems. In 2.1, it shows how the privacy is secured by only sharing the query with the recommender system and not the context to external Mobility service provider (MSP) to preserve user data privacy but adversary nowadays are able to infer background knowledge from the query data as well.

2.2.1 Privacy and Security threats in Christian Samuel's design

The model architecture of this paper comprises a lot of third-party service providers in architecture which includes sharing of the user data within all of them which increases the risk of misuse of private data. Even though the user management module acts as a supervisor providing tokens to authenticate users and unique identifiers for desired itineraries, it is not difficult for an outsider to access the tokens with proxy-based or brute force attacks. Absence of distributed approach shows the dependency on the centralized server for all the data related tasks which makes the server high at risk for adversary attacks. Also the design requires the accurate location sharing with the server/data collector which in

cases of adversary attacks can fall in malicious hands. The paper fails to explain any location privacy protection algorithms if used.

2.2.2 Privacy Risks in Travel Assistants

In spite of such a great progress in Travel Assistants development, there are many issues which are unaddressed yet because of the impact to the utility of the services. Privacy is one such issue which is very much in conversation right now. Privacy can mean different for different people at different times. The data at risk can have different sensitive attributes like personal, commercial and research. It is constantly changing and thus the different characteristics of privacy needs to be addressed in the privacy protection mechanisms. In (21), Martina Ziefle et al, have concluded that user's find location data and lifestyle habits as too personal to be shared. When using the travel assistants, the major risk is of sharing the location data of the users.

Existing travel assistants have taken precautions to protect personal information, but there is always a trade-off between privacy and performance, quality, or cost. Privacy preserving mechanisms restrain various Location-Based services (LBS) preventing the misuse of user's private data, but they are rarely used since users are unaware of the privacy threats as they do not intensively share their location. Even if communications are pseudonymous, the spatio-temporal correlation of location traces may serve as a quasi-identifier 2.2. For e.g., work and home locations uniquely identify most of the population of US (6). Studies and evaluations performed in (6) shows the privacy erosion problem arises when even a small amount of information is shared with such services.

Privacy policy at Whim (15) explains the data they collect from the users, how they use the data and what data is shared with third party users. It mentions that along with the travel and trips data, they may collect data such as the IP address and many more which is not even specified clearly. If the non-personal data is stored with personal data, ultimately making both non-personal and personal data identifiable to the individual. Also, to improve the usability of the service, personal data is shared with third party users which is under the Data protection law but if this data is lost by any of them to the hands of malicious attackers, then can be harmful. There have been various ways (16) by which users have tried to protect privacy while using such applications like searching destination with fake source location or by modifying privacy levels to zip codes or city names. However, all these measures do not provide accurate results and require manual efforts of entering source locations.

While (9) suggested the mixing of a user's location with other users' locations in query so that it's hard to identify the exact or real location of the user, it is not the good approach since there will be more damage if all the user's locations are exposed. (16) also explains

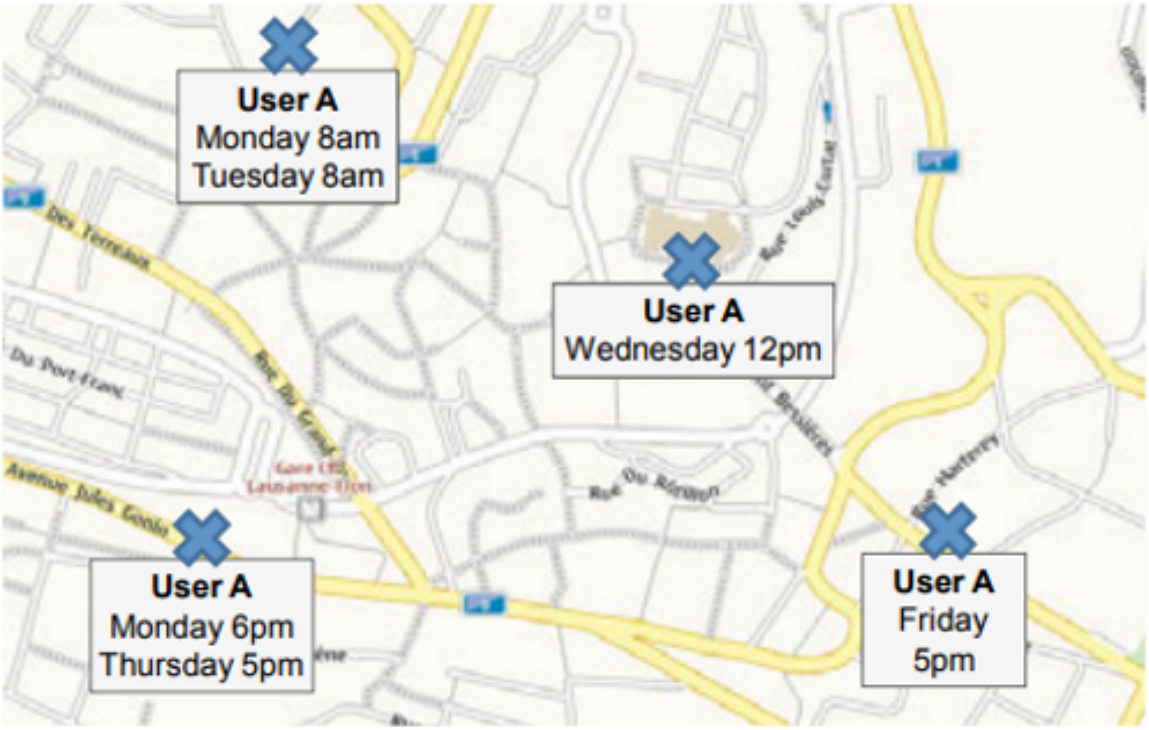

Figure 2.2: Quasi-identification of a user A whose identity is not revealed (6)

that the measures where either identity of the user or the location is hidden, it is not difficult nowadays to find out one from the other where every application asks to sign in before the service is provided.

2.3 Existing privacy protection mechanisms (PPM)

There are various approaches which protect the privacy for such location based services using location encryption, obfuscation, confusion techniques.

Zhang *et al.* in (20) have proposed a system which uses dual k-anonymity mechanism 2.3 combined with dynamic pseudonyms to overcome majority of the privacy concerns mentioned in the above studies. Here, when sending a query request user is assigned with dynamic pseudonym and sent to the different anonymizers along with K-1 additional locations. These query requests are sent to the third-party location service provider for another query, and the results are returned to the user via multiple anonymizers. Thus, it's hard for an attacker to identify the user due to dynamic pseudonym. The advantage of the DKM scheme is that user trajectory cannot be acquired from the Location service provider or a single anonymizer. In addition to offering more protection to the user's trajectory privacy, it provides an effective approach for the anonymizer's performance bottleneck and

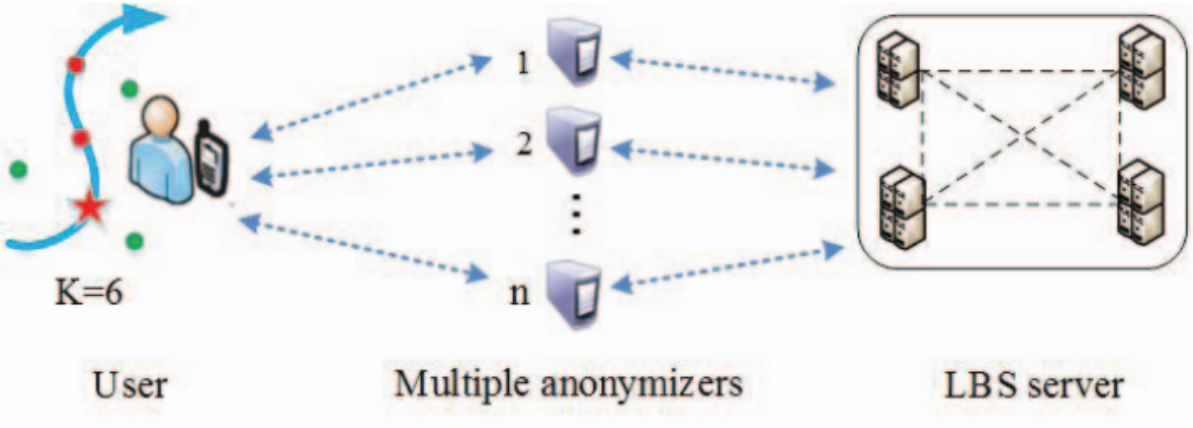

Figure 2.3: Dual-k anonymizer

(20)

the single point of failure problem.

Yuwen *et al.* in () uses efficient and reliable cryptography approach to ensure privacy of the transmitted location data but still the delay due to the cryptography approach reduces the utility of the system and is thus not preferred over other mechanisms.

In the big data era, PPMs such as cryptography, k-anonymity provides privacy protection to a certain limit. Issues such as slowed systems, delayed responses, quasi identification attacks were increased which demanded new mechanisms for privacy protection.

2.4 Differential Privacy

To balance the trade off between utility and privacy, Cynthia Dwork *et al.* have proposed Differential privacy which resolves most of the former mentioned issues. Instead of sharing the exact data, DP shares the aggregated statistics and increases uncertainty of the data by adding noise to it when sharing with third party service providers.

Differential Privacy(DP) as explained in (4) is a privacy preserving mechanism which adds statistically controlled noise to the user data thus restricting the identification of the user data by the adversary or any third party. It is mainly classified into two types: Local and Global. Local differential privacy (LDP) is used by Google and Apple to protect their users data since LDP works best with large amount of data. This distributed variant of DP does not demand a trusted third party, which avoids the data collector stealing user information. To protect privacy, each user chooses the level of privacy of LDP he expects to perturb their data before sending the data to the aggregator, that is, the data collector. In this way, only obfuscated data is gathered and manipulated by the data collector. When the amount of data is large enough, the utility becomes acceptable, so the data collector can still find out

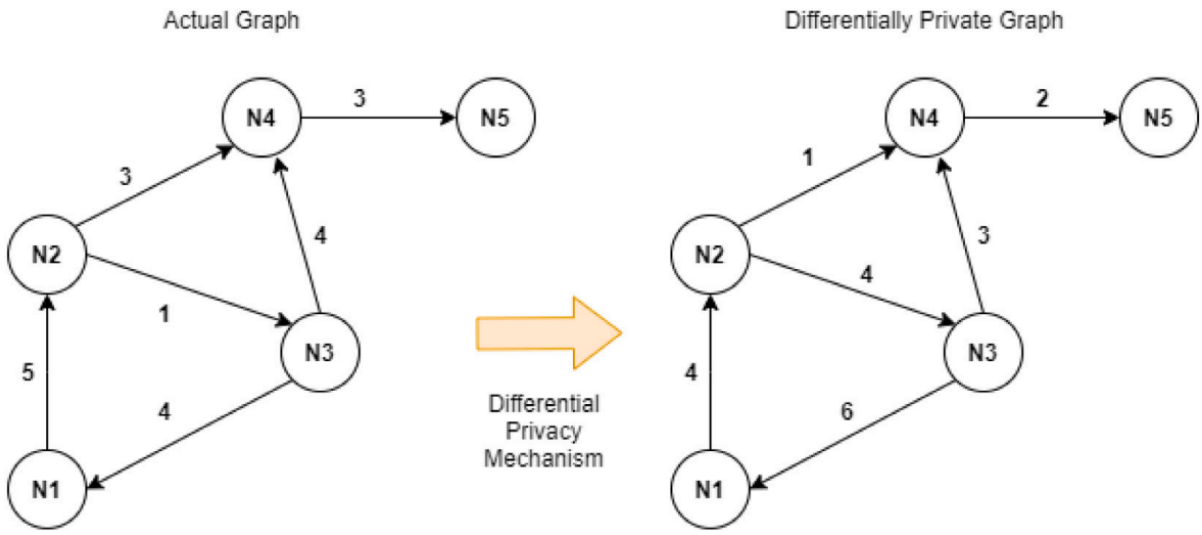

Figure 2.4: DP applied in graph topology preserving edge weight functions (2)

the accurate frequency of specified attributes.

Global differential privacy (GDP) provides more accuracy as compared to the LDP but is used only when we can guarantee the privacy with data collector. However LDP allows the user to add noise before passing the data to the data collector and thus is more privacy-preserved and preferred for the designs with distributed systems.

Table 2.6 shows the existing applications of Differential Privacy in variety of areas wherein sensitive user information is at risk.

2.5 Proposed Methodology

Because it has a big domain size, a user's location data can be safeguarded via a local differential privacy technique. Location data, as described in (19), can be generalized over a greater region and regarded as a categorical feature rather than a numerical property (19) (1).

Privacy budget ϵ can be helpful in knowing the degree of privacy protection thus letting us to study the trade-off between utility and privacy. Location data is a great example for LDP because for a user it does not matter if the city name is exposed but will definitely matter if the apartment number is exposed which can be managed by changing the ϵ value.

LDP is generally implemented with Laplace and Exponential mechanisms, but Laplace mechanism cannot be applied since the location data is treated as a categorical variable and Laplace only applies to numerical data and Exponential mechanism is applied for Global/Central Differential Privacy since it requires the design to follow a centralized approach unlike ours where the users are distributed and there is no trusted third party

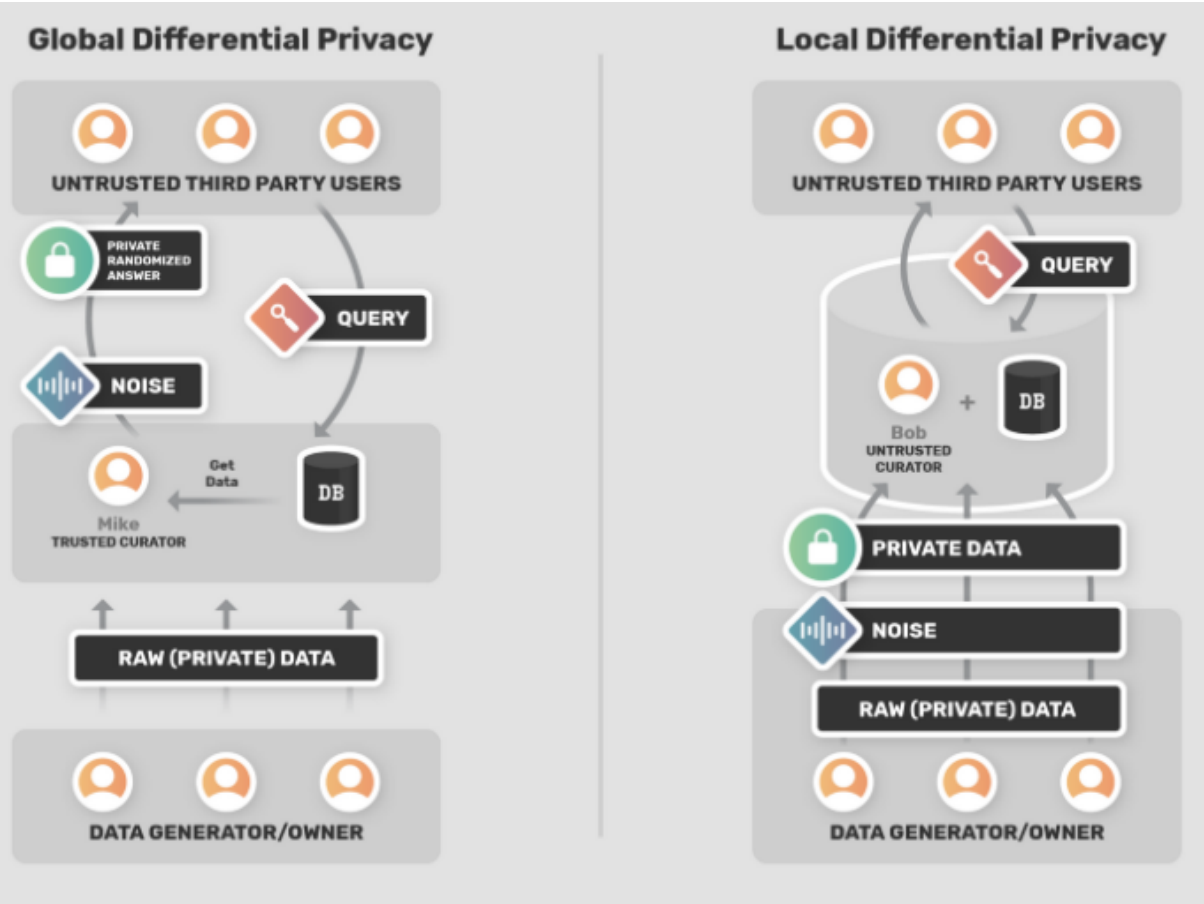

Figure 2.5: Global Differential Privacy vs Local Differential Privacy (5)

Differential privacy techniques	Differential privacy in Internet of vehicles
	Differential Privacy in social network analysis
	Local Differential Privacy on Metric spaces

Figure 2.6: Differential Privacy in real world applications

server with access to all the records.

LDP also restricts the performance of the algorithm if the same queries are repeated over time that is, an adversary can infer the real value from the database if the same data is sent over and over again to the database using the same perturbed value. This allows LDP to use memoization to answer the same queries with the same data. However, memoization also can let the adversary to find the real raw value after a point. As stated in (22) traditional randomized responses algorithms fail to provide privacy in case of memoizations. Multiple queries on same perturbed data or user are also not protected by the traditional randomized responses. Thus, to overcome this problem, we will be using Randomized Aggregatable Privacy-Preserving Ordinal Response(RAPPOR) mechanism to perturb the data which uses two perturbation levels, one is the permanent randomized parameter and the other is instantaneous randomized parameter. The permanent randomized parameter can be used to maintain the longitudinal privacy of the location data whereas the instantaneous randomized parameter will be helpful in one time attacks.

Amongst very few RAPPOR implementation papers, in (14), (17) Differential Privacy with RAPPOR mechanism is implemented along with Map density segmentation mechanisms like Modified Hilbert Curve or Dynamic Hilbert Curve which helps to map the 2D space to 1D space to ensure spatial correlation. These region segmentation is updated with the map changes at the server side regularly to sync the changes at the user side. The participant's location is then perturbed at the user side before sending it to the server.

However, RAPPOR also works best only when the user data does not change too often and when there is less correlation amongst the different user data. But in a Travel assistant, with the continuous update of location data to the server, there are high chances of correlation amongst two consecutive locations. If two different location values are correlated then an adversary can infer knowledge from the location correlation problem. In other words, the permanent randomized response parameter ensures longitudinal privacy only when the user data is not changing rapidly and if changing then the values should not be correlated.

To solve this correlation protection problem, we have proposed using a Markov Model to predict the user location rather than using the accurate/original location data. Markov model and its variants are used in multiple existing designs (10) (11) to predict the location data to protect the location correlation problems when implementing Differential privacy.

Thus our proposed methodology comprises of a location prediction algorithm based on Markov model to protect the correlation amongst different location data and a location protection algorithm based on Differential Privacy with RAPPOR to protect the privacy of the location data.

2.6 Trade off between utility and privacy

Author of (21) proposed a trade-off framework by using Pareto efficiency. Pareto efficiency, or Pareto optimality, is a state that no one can get better when no one else gets worse. In LDP, if the privacy enhances, the loss of privacy is bound to increase, which will lead to reduced data utility. Conversely, if the utility of data increases, the utility loss will decrease, which will inevitably lead to lower privacy protection. By definition, the whole system state is always Pareto optimality. The privacy metric and utility metric is measured by privacy loss and utility loss respectively.

When privacy budget ϵ is changed as an independent variable, it will cause changes in the privacy measure and utility measure of the dependent variable. The larger the value of the privacy budget ϵ, the weaker the privacy protection, the greater the privacy loss, the higher the utility, the smaller the utility loss. In other words, ϵ-LDP, the privacy is determined by epsilon. If ϵ=0, which means exp(0) = 1, namely the two tuples are closely enough. Therefore, the privacy is ensured perfectly. However, when the ϵ=∞ there is no privacy guarantee. Thus, choice of ϵ is critical in practice as the increase in privacy risks is in direct proportion to exp(ϵ).

2.7 Research Statement

When using privacy preserving mechanisms there will always be trade-off in privacy and utility. Though there have been many studies showing this trade-off in various areas, to the best of our knowledge, this trade-off is not yet studied fully dedicating to the travel assistant system. Travel assistant provides frequent real-time reports on the user's location and behavior, which can reveal a great deal of personal information. It is not clear to the customers how the privacy impacts the performance of the system. Thus, our research topic is to build a prototype addressing the trade-off between utility and privacy. This paper will compare and show how the increased privacy while applying differential privacy directly restricts the performance of the travel assistant.

The main focus of this paper will be on applying a privacy protection mechanism such as RAPPOR to perturb the predicted location generated by the Markov Model protecting location correlation problem. Proposed design can help in overcoming the location correlation limitation in the existing RAPPOR mechanism which is not addressed yet in any studies conducted so far.

The design also aims to allow users control over data sharing so that they can decide how much data to share for the system's required performance. LDP is chosen over GDP since the data is anonymized before sharing it with the data collector and thus reducing the risk of

misuse of user data by adversary, third party or any doubted data collector. Also GDP requires the third party centralized collect data which is not suitable for Distributed versions of this model.

To address this trade-off between utility and privacy, we can use the implementation of the framework proposed in (21) by using the utility loss as utility metric and privacy loss as privacy metric. The framework will be implemented to design a payoff function which identifies the best system state where in the privacy protection and utility of sensitive information is desirable.

3 Design and Implementation

3.1 Project Overview

3.1.1 Functional Overview

To ensure strong privacy and support the local differential privacy approach, this design will be comprised of distributed computing where the user location data will be perturbed at the user devices itself and will be sent to the server only after perturbation.

To ensure on route updates even during the journey, the application needs to continuously send current location data to the server for perturbation. But the location data sent for perturbation will not be the actual location data but the predictions of the first order Markov model results which will generate the locations which the user is most likely to visit. First order Markov model uses only the value of the current location for prediction of future location and no dependency on the previous/older values.

Location prediction algorithm:

Markov model is a memory less random process uses actions and transition matrix which will be the movement and the matrix of probabilities of result of every movement.

$P(Xn+1 = x|X1 = x1,\cdots, Xn = xn) = P(Xn+1 = x|Xn = xn)$

As stated in (18), in a finite state space Xn: n= 0, 1, 2,.. then $Xn=i$ indicates that the object is in state I at time n. If for any n where n >= 0, the above equation is true then the process is a Markov chain.

The transition probability matrix is generated with the help of above equation for consecutive states. One-step transition probability can be expressed as the below equation:

$Pi,j = P(Xn+1 = xj|Xn = xi)$

When the above equation arranged in a matrix form we get,

$$P = \begin{bmatrix} P_{1,1} & \cdots & P_{1,j} & \cdots & P_{1,m} \\ \vdots & \ddots & \vdots & & \vdots \\ P_{i,1} & \cdots & P_{i,j} & \cdots & P_{i,m} \\ \vdots & & \vdots & \ddots & \vdots \\ P_{m,1} & \cdots & P_{m,j} & \cdots & P_{m,m} \end{bmatrix}$$

Location protection algorithm: RAPPOR is used to perturb the predicted location as follows:

Let A = a1, a2, ...an is the area ID after the map is divided by the server, where n represents the total number of areas after the map is divided. After the participant gets the area ID from the server, find which ID the current location belongs to. Let *ai* (1 <= i < = n) be the ID with the current location. Then, a n-bit array, L (which denotes the current location of a specific user) is defined as,

$$L_j = \begin{cases} 1, & j = i \\ 0, & otherwise \end{cases} \quad (17)$$

Permanent randomized response: For each client's value v and bit i, 0 <= i < k in B, create a binary reporting value B' which equals to

$$B'_i = \begin{cases} 1, & \text{with probability } \frac{1}{2}f \\ 0, & \text{with probability } \frac{1}{2}f \\ B_i, & \text{with probability } 1-f \end{cases} \quad (22)$$

This B' is memoized and is always used in multiple queries of same data which helps in maintaining longitudinal privacy. Here, it is very important that every report on B should return B' as the perturbed value which helps to not allow adversary extract the averaging information from multiple queries thus allowing multiple queries.

Instantaneous Randomized Response: Allocate a bit array S of size k and initialize to 0. Set each bit i in S with probabilities

$$P(S_i = 1) = \begin{cases} q, & \text{if } B'_i = 1. \\ p, & \text{if } B'_i = 0. \end{cases} \quad (22)$$

Finally, perturbed S is sent to the server t query the database. The above encoding method as mentioned in (22) (17) satisfies ϵ-differential privacy.

3.1.2 Application flow

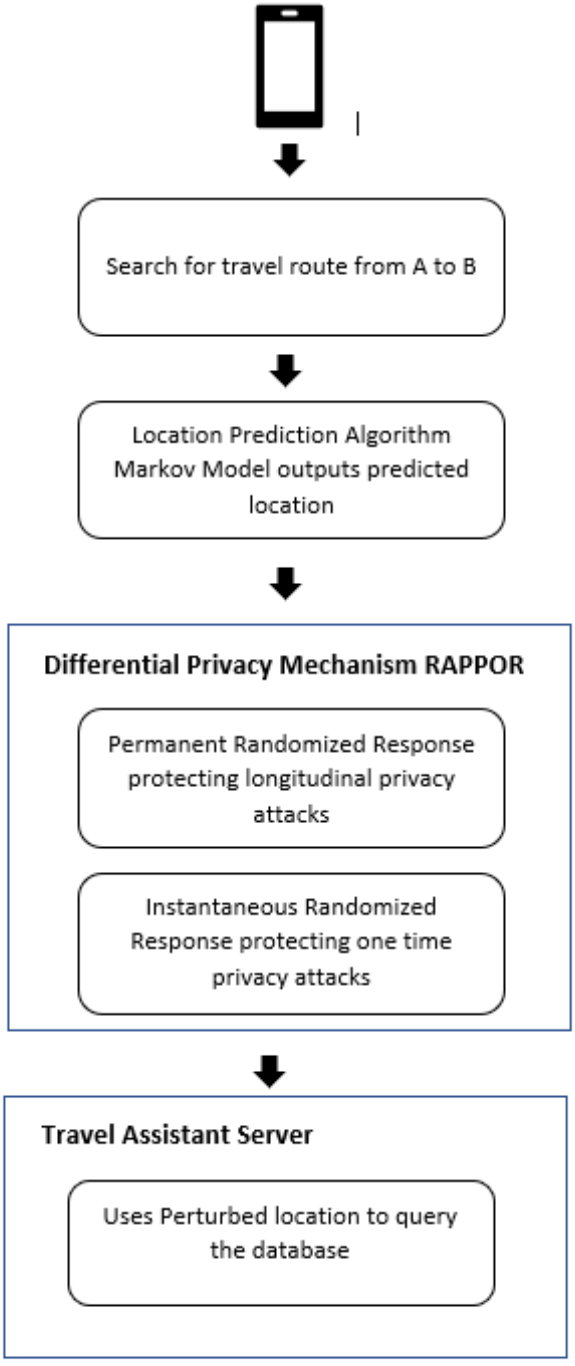

Figure 3.1: Architecture Flow of the Privacy Aware Travel Assistant

3.1.3 Technology choices

Mobile Application: The Web application handles the front end and manages the general flow of the application. User can interact with the application through this module to select the routes and get real time updates about the journey. The programming language chosen for the application is Python 3.8 version. This choice was made as it is a flexible programming language that allows for quick prototyping. The downside of Python's flexibility manifests with the slower computation times. However, the slower times are still satisfactory in regards to the scope of this dissertation.

Kivy is the open-source python library used to build the mobile application to ease the integration on the Android platform. Kivy does not support all the Android features but can be sufficient for the prototype version we are planning to build with this dissertation.

Location prediction algorithm is implemented with Markov model since its an memory less random process which predicts the future values only based on the current value and not depends on previous values. The reason for choosing first order Markov process is to reduce the correlation amongst the predicted values when compared to the original values.

To choose an algorithm for the LDP, Laplace mechanism cannot be applied since the location data is treated as a categorical variable and Laplace only applies to numerical data. Exponential mechanism is applied for Global Differential Privacy since it requires the design to follow a centralized approach unlike ours where the users are distributed and there is no trusted third party server with access to all the records. We have also reviewed the traditional randomized responses algorithms used and as stated in (22), they fail to provide privacy in case memoizations. Multiple queries on same perturbed data or user are also not protected by the traditional randomized responses. Thus, Local Differential Privacy algorithm is implemented using the Randomized Aggregatable Privacy-Preserving Ordinal Response (RAPPOR) technique for preserving the privacy of location data as implemented in (17).

The trade-off between the utility and the privacy can be visualized with the help of DPComp, a web based tool used to assess the accuracy of the differentially private algorithms.

4 Evaluation

5 Conclusion

People living in busy cities always are in a hurry and to travel in such a haste can be both dangerous and infuriating. Before leaving home, it is vital to understand the route. Humans become accustomed to a particular route for their workplace commute, but there are occasions when it is necessary to take a different route to work or to visit somewhere else before arriving at work. Additionally, owing to construction or an accident, the road to work may be closed. Even if we have routing tools like Google Maps that can provide us with instructions, there are always a few additional chores to be accomplished, such as booking tickets, check-in timings, and so on. At times like these, everyone feels like they need an assistant to help them with all of their responsibilities and get to work on time.

Bibliography

[1] Alvim, Mário and Chatzikokolakis, Konstantinos and Palamidessi, Catuscia and Pazii, Anna. Invited paper: Local differential privacy on metric spaces: Optimizing the trade-off with utility. In *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*, pages 262–267, 2018.

[2] Atmaca, Ugur and Maple, Carsten and Epiphaniou, Gregory and Dianati, Mehrdad. A privacy-preserving route planning scheme for the Internet of Vehicles. *Ad Hoc Networks*, 123:102680, 09 2021.

[3] Cisco. Personalized Travel Assistant [PTA], Seoul, 2022.

[4] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3–4):211–407, aug 2014.

[5] Shaistha Fathima. Global vs local differential privacy, Oct 2020.

[6] Julien Freudiger, Reza Shokri, and Jean-Pierre Hubaux. Evaluating the privacy risk of location-based services. In George Danezis, editor, *Financial Cryptography and Data Security*, pages 31–46, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[7] Giannopoulos, George. The application of information and communication technologies in transport. *European Journal of Operational Research*, 152:302–320, 02 2004.

[8] Google, 2022.

[9] Lee, Ken and Lee, Wang-Chien and Leong, Hong and Zheng, Baihua. Navigational path privacy protection: navigational path privacy protection. pages 691–700, 01 2009.

[10] Hongtao Li, Yue Wang, Feng Guo, Jie Wang, Bo Wang, and Chuankun Wu. Differential privacy location protection method based on the markov model. *Wireless Communications and Mobile Computing*, 2021:1–12, 2021.

[11] Lu Ou, Zheng Qin, Yonghe Liu, Hui Yin, Yupeng Hu, and Hao Chen. Multi-user location correlation protection with differential privacy. In *2016 IEEE 22nd International Conference on Parallel and Distributed Systems (ICPADS)*, pages 422–429, 2016.

[12] Dave Roos. How mapquest works, Dec 2005.

[13] Samsel, Christian. *Ubiquitous Intermodal Mobility Assistance*. PhD thesis, 03 2019.

[14] Jian Wang, Yanli Wang, Guosheng Zhao, and Zhongnan Zhao. Location protection method for mobile crowd sensing based on local differential privacy preference. *Peer-to-Peer Networking and Applications*, 12(5):1097–1109, 2019.

[15] Whim. Whim users privacy policy, 2022.

[16] Wikipedia contributors. Mobility-as-a-service — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Mobility-as-a-Service&oldid=1072224325`, 2022. [Online; accessed 21-February-2022].

[17] Wang Xiongjian and Yang Weidong. Protection method of continuous location uploading based on local differential privacy. In *2020 International Conference on Networking and Network Applications (NaNA)*, pages 157–161, 2020.

[18] Ming Yan, Shuijing Li, Chien Aun Chan, Yinghua Shen, and Ying Yu. Mobility prediction using a weighted markov model based on mobile user classification. *Sensors*, 21(5), 2021.

[19] Mengmeng Yang, Lingjuan Lyu, Jun Zhao, Tianqing Zhu, and Kwok-Yan Lam. Local differential privacy and its applications: A comprehensive survey. *CoRR*, abs/2008.03686, 2020.

[20] Zhang, Shaobo and Wang, Guojun and Liu, Qin and Wen, Xi and Liao, Junguo. A Trajectory Privacy-Preserving Scheme Based on Dual-K Mechanism for Continuous Location-Based Services. In *2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*, pages 1004–1010, 2017.

[21] Ziefle, Martina and Halbey, julian. Users' Willingness to Share Data on the Internet: Perceived Benefits and Caveats. 04 2016.

[22] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 21st ACM Conference on Computer and Communications Security*, Scottsdale, Arizona, 2014.