TrPF: Location Aware Participation System
For Mobile Users

Anju Augustine
Mtech Scholar
Dept. Of Computer Science
ICET Muvattupuzha

Abstract—The ubiquity of the various cheap embedded sensors on mobile devices, such as cameras, microphones, accelerometers, and so on, is enabling the emergence of participatory sensing applications. During participatory sensing individuals can collect and analyse the participant’s location as well as trajectory privacy which break the participators privacy. In the proposed system we improve a trajectory privacy preserving framework. For that we improve a the critical mix zone model with considering the time factor from the perspective of graph theory. It does not reveal the participants information. Here considers the dummy locations also.

Keywords—Mix Zones; Weighted Graph; Trajectory’s;

I. INTRODUCTION
Participatory sensing [1] which is the process that enables individuals to collect, analyze and share local knowledge with their own mobile devices, emerges as required under these well conditions. The development of wireless communication technologies, such as WLAN, 3G/4G, WiMAX, Bluetooth, Zigbee, and so on, mobile devices are equipped with a variety of embedded sensors that has powerful sensing, storage and processing capabilities which enables participatory sensing to emerge as a new and powerful technology. Now participatory sensing depends on a wide geographic area[2]. The participant may upload the data and tag with the spatial-temporal information when the reading where recorded[7]. In typical applications, the uploaded data reports may reveal participants’ spatial-temporal information. Analysts could obtain some valuable results from the published trajectories for decision making, for example, merchants may decide where to build a super market that can produce maximum profit by analyzing trajectories of customers in a certain area.

In this paper, we propose a trajectory privacy-preserving framework named TrPF, for participatory sensing. We observe that the locations on or nearby participants trajectories may not all be sensitive, and with this thought, our proposal only deals with the sensitive trajectory segments. For the privacy we have to construct the mixzone. Mix-zones are regions [3] where no applications can track participants’ movements. Then calculate the weight construct to calculate the weight.

Some work focused on road network[4] mixzones which are not useful for participatory sensing. They all build mix zones at road intersection which may restrict the random data collection time and the number of ingress/egress locations; for another thing, the trajectory segments at the road intersection may not be sensitive, while the others would be. So here using theoretical mixzone model for protecting sensitive trajectory segment from the perspective of graph theory.

II. RELATED WORK
In the survey shows the privacy and security implications in three types of application scenarios. In [2], [7], they analyzed the privacy challenges in participatory sensing applications in detail. Several methods are surveyed for the privacy protection in terms location privacy protection and trajectory privacy protection in location-based services.

A. Location Privacy Protection
Location privacy is defined as the ability to prevent other unauthorized parties from learning one’s current or past location. There are several methods surveyed for privacy protection. They are described as follows.

1) Dummy Locations: This method mainly employs the idea of dummy locations [8] to protect a user’s location privacy. A location—dependent query is abstracted as Q = (pos; P), where parameter pos is the mobile user location, P denotes userspecified predicates and Q the original query. With this approach, the original query is typically converted into a query Q0 = (pos1; pos2; : : : ; posk; P), where the posi include the user’s real location and k - 1 dummy locations, and P is the original query predicate that applies to all k-locations. We call query Q0 a location privacy query, since it hides the user location.

2) Obfuscation: Obfuscation[6] protect user’s location privacy by deliberately degrading the accuracy of his/her spatial-temporal information.

3) K-Anonymity: It[5] is a general wide spread concept which is not restricted to location privacy.

4) Mix zones: Pseudonyms [3] is used to break the linkage between a user’s identity and his/her events. The process of its change is usually performed in some pre-determined areas called mix-zones, [4] which are the regions where no applications can track participants’ movements.

B. Trajectory Privacy Protection
By identifying users trajectories anyone can easily track his/her location. Most simple method for Trajectory protection [11] are dummy trajectory and suppression method.

1) Dummy trajectory confusion: This method propose to generate dummy trajectories in order to confuse the adversaries. Dummy trajectory’s are produced based on two principles. i) the movement patterns of dummies should be similar to real users; ii), the intersections of trajectories should be as more as possible.
1) Suppression-based method: This is based on the assumption that different adversaries may have different and disjoint part of users’ trajectories. This method [11] reduces the probability of disclosing whole trajectories. Here Trajectory pieces should be suppressed and the publication of these pieces may increase the whole trajectory’s until reaches the probability above a certain threshold.

III. ARCHITECHTURE

![Architecture of TrPF](image)

In this fig 1 shows the architecture that contains the following:

Data collectors: Wireless Nodes are apparatus with the capabilities of feeling, computation, memory and wireless connection, which act as data collectors in participatory sensing. They can collect data and share with each participant. The data collector who wants to upload data can get a certificate from Trusted third party.

Trusted Third Party Server (TTPs): TTPs shops participators’ applicable information such as certificates and pseudonyms information. Certificates are used for verifying participators’ validity so as to discard malicious attacker. The revelation of the spatial-temporal information may also break the participants’ privacy.

Report Server: The report server mainly has two functions. (a) Interact with TTPs to verify the validity of the each participators’ identity by the certifcates contained in the data reports; (b) Simplify the uploaded data reports such as data aggregation, and then gives the data reports to Application Server.

Application Server: Application Server acts as a data center. It gives all data services for end users and play the following functions:

(a) Data Storage: store the processed data received from data report server; (b) Data Sharing: any end users can access data services; (c) Data Publish: release the data accounts for the end users to query.

Queriers: Queriers are the end users that request sensor accounts in a granted participatory sensing submission, which can be individual users or community users. They can access the data uploaded by the data collectors. Only the registered end users can get access to the distributed data accounts.

A. Trajectory Mix-Zone Graph Model

In Trajectory mixzone graph model, we propose to anonymize the sensitive trajectory segment from the perspective of graph theory. The whole area is divided into several parts. According to the sensitive locations on or nearby the trajectories, we divide the whole trajectories into sensitive trajectory segments and nonsensitive trajectory segments. Here we only protect the sensitive trajectory based on mix-zones model and pseudonym technique. This reduces information loss and costs at a certain privacy-preserving level.

Any data collector who enters the Sensitive Area should select a pseudonym provided by TTPs to anonymize the linkability between his identity and his collected data reports. Meanwhile, they record their ingress and egress time. A participator’s information we describe as a tuple:

\[ I_i = (I_D p, R_i, S_i, t_{\text{ingress}}, \Delta t_{\text{egress}}) \]

where \( I_D p \) comprises the participator’s pseudonym supplied by TTPs, \( R_i \) is the mapping from participator’s identity to his pseudonym, \( S_i \) is the sensitive locality the participator passes by, \( t_{\text{ingress}} \) presents the set of participators’ enter time and \( \Delta t_{\text{egress}} \) is the participator’s egress time gap.

B. Trajectory Graph Construction

We propose to model the Trajectory Mix-zones as Directed Weighted Graph (DWG), which is formalized as \( G = (V,E) \) where \( V \) is the set of vertexes which are constructed by the pseudonyms provided by TTPs. \( E \) is the the set of edges that represent the participators’ trajectory mapping from the ingress to the egress in the sensitive area. A participator enters the sensitive area with a pseudonym and leaves it with another pseudonym. It can be depicted as

\[ V = \{(v11,v12,\ldots,v1n),(v21,v22,\ldots,v2n)\} \]

As a result of pseudonym technique, there may be some difficulties for adversary to link the ingress and egress participator with the same identity.

The process of the trajectory graph construction can be described by Algorithm 1

**Algorithm 1 Graph Construct**

Input : Trajectory \( Tr \) and Pseudonym set \( P \)
Output : Directed Weighted Graph (DWG)
Sensitive trajectory segment \( TF \leftarrow Tr \)
For each \( TF \), \( \epsilon \in TF \) then
Construct the sensitive area \( S \)
Select \( k \) pseudonyms as vertex to achieve \( k \)-anonymity
for each i < k then
select ingress pseudonym P_i ∈ P
select egress pseudonym P_j ∈ P
if P_i = P_j then
E_{ij} ← < P_i ≠ P_j >
else
Select another point P_0
End if
W_{ij} ← Weight Construct
G_i ← < V_i, E_i, W_{ij} >
End for
End for
Return graph G

Algorithm2 Weight Construct

Input t_ingress and t_egress = t_{j}, t_{j+1}
Output= Edge Weight W
For i=1 to k do
For j=1 to k do
§1=t_{j} - t_ingress (v_i), §2= t_{j+1} - t_ingress (v_i)
P(v_i, t) = integral[ f(§1, §2)]
P(v', t) = 0
P(v', t) = P(v', t) + P(v_i, t)
End both for loop
P(v_i[t_{j}, t_{j+1}]) ← (P(v_i, t)) / P(v, t)
W_{ij} ← P(v_i[t_{j}, t_{j+1}])
Return W

IV. METRIC

A. Privacy Level Metric

Privacy level metric is achieved based on Information entropy. Information entropy is a quantitative measure of information content and uncertainty over a probability distribution. In this paper, the probability distribution gives the chance that adversary can identify each participant. If the probability distribution is uniform, the information entropy is higher and is difficult to identify the actual participant. If there is significant difference in probability distribution it is easy to identify the real participant with low information entropy.

B. Privacy Loss Metric

Privacy loss is defined as the probability that an adversary will be able to gain sensitive trajectory segment about a participant. It could be calculated by combining the identity leakage and the pseudonym mapping index.

C. Information Loss Metric

Information loss is defined as the reduction in the probability with which people can accurately determine the position of an object. It is calculated by the sum of area size of anonymity regions.

IV. PERFORMANCE EVALUATION

TABLE 1: Two Groups of Statistical Parameters

<table>
<thead>
<tr>
<th>Ingress time Interval(T)</th>
<th>Arrival rate (λ)</th>
<th>Time interval Parameter(μ, σ)</th>
<th>The number of participants(k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>5</td>
<td>(2.5, 0.5)</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>(3.1)</td>
<td>20</td>
</tr>
</tbody>
</table>

Two groups of experiments with different statistical parameters are shown in Table I. As a result of participators’ different arrival rates λ = λ (5, 10) shows the number of participants that enters the mix-zones is different. The probability density function of time interval in mix-zones with (μ, σ) = [(2.5, 0.5), (3.1)].

According to the discussion, the maximum entropy achieves if and only if the mapping probabilities are equal. In this paper, we can improve the theoretical mix-zones model with considering the time factor. The maximum entropy and actual entropy can be computed according to the probability distributions depict correspondingly. When the mapping probabilities are equal, the maximum entropy achieves. Privacy level can be calculated depicted by Fig.2. It evaluates the privacy-preserving level. When the privacy level increases the trajectory privacy-preserving scheme becomes more stronger. Consequently, the privacy leak is lower.

Fig.2 Privacy Level

Fig.3 Privacy Loss
Fig. 3 the privacy loss decreases with the number of participants in the mix-zones increases. Discussion based on the number of participators that enters the mix-zones $k$ will changes with the arrival rate $\lambda$ and ingress time interval $T$ changing. When the number of participators increases the arrival rate and time interval also increases.

VI. CONCLUSION

In this paper proposes a trajectory protection framework by combining the dummy location method and theoretical mixzone model based on time factor is used.

ACKNOWLEDGMENT

The authors would like to thank each and every one reviewers for their insightful comments and suggestions, which have greatly improved this paper.

REFERENCES