Transition-Entropy: A Novel Metric for Privacy Preservation in Location-Based Services

Sina Shaham*, Ming Ding**, Bo Liu^{+}, Zihuai Lin*, Jun Li^{++}

*School of Electrical and Information Engineering, The University of Sydney, Australia
**Data61, CSIRO, Sydney, Australia
$^{+}$Department of Engineering, La Trobe University, Australia
$^{++}$Nanjing University of Science and Technology, Nanjing, China
Email: {sina.shaham, zihuai.lin}@sydney.edu.au, ming.ding@data61.csiro.au, b.liu2@latrobe.edu.au, jun.li@njust.edu.cn

***Abstract*—The advent of location-based services has created the need for preserving the location privacy of users. An adversary such as an untrusted location-based server can monitor the queried locations by a user to infer sensitive information such as the user's home address, health conditions, shopping habits, etc. To address this issue, dummy-based algorithms have been developed to increase the anonymity of users, and thus, protecting their privacy. Unfortunately, the existing algorithms only consider a limited amount of side information known by the adversary whereas they may face more serious challenges in practice. In this paper, we consider a new type of side information based on consecutive location changes of users, and propose a new metric called transition-entropy to investigate the location privacy preservation. Furthermore, we develop a greedy algorithm to significantly improve the transition-entropy performance for a given dummy generation algorithm. Via experiments conducted on a real-life dataset, we evaluate the performance of the proposed metric and algorithm.**

I. Introduction

With the ubiquitous use of smartphones and social networks, location-based services (LBSs) have become an essential part of the contemporary society. The users of smart devices can simply download location-based applications and query the information from the LBS provider. For example, LBSs offered by companies like Alibaba, Apple, and Google can be used to find nearby restaurants, track the parcels, and provide personalized weather notifications. The annual market for LBSs is expected to reach USD 77.84 Billion by 2021, with an annual growth rate of 38.9% [1].

In spite of countless advantages of LBSs, the privacy issues associated with the user locations have raised many concerns in our society. An untrusted server can collect the location data of users and analyze it to learn sensitive information such as the type of queries submitted, shopping habits of users, and the address of users' properties or workplaces. Such information can be easily abused by the server or disclosed to other parties. Therefore, it is of great importance to devise new ways to preserve the location privacy of users defined as "the ability to prevent other parties from learning one's current or past locations" [2].

The techniques to address the threats to the location privacy of users have attracted much attention among researchers [2]–[8]. Most of the literature is based on an approach called k-anonymity [9]. Using this criterion, the release of a location is said to provide k-anonymity, if the real location of any user is not distinguishable from at least $k-1$ other locations. Initially, the approach to hide the location of the user was conducted using a trusted anonymization server [10], but later on, due to the shortcomings of this approach such as the anonymizer becoming the bottleneck itself, the use of dummy locations to achieve the k-anonymity was proposed in [11]. Since then, the researchers have strived to develop dummy generation algorithms to preserve the k-anonymity for users.

The principal idea behind the dummy generation algorithms is to generate $k-1$ dummy locations aside from the real location of the user and submitting them all together to the LBS server while asking for a query from the LBS provider. Thus, it becomes difficult for an untrusted LBS provider, or so-called the adversary, to identify the real location of the user. The groundwork in this field was laid by the authors in [11]. They generated the dummies randomly throughout the map and evolved them as users move. Followed by this work, the authors in [12] and [13] proposed to choose the candidate dummies from a virtual circle or grid constructed around the current location of the user. Unfortunately, in all of the mentioned works, the fact that the adversary might have some side information which can rule out the dummies or reveal the real location of the user was overlooked.

One important piece of side information which can be exploited by the adversary is the query probability of the locations across the map. The adversary can utilize the recorded data and infer the number of times that the users have queried over various locations on the map. Using this information, the adversary can calculate the query probability of each location, and then, identify the dummy locations according to the history of interests in locations. For instance, if a dummy has been chosen on a lake, where the query probability is basically close to zero, the adversary will then know with a high likelihood that such queried location is a dummy. And therefore, such naive selection of dummy locations compromises the location privacy of the user. To solve this issue, an enhanced algorithm was proposed by [14], referred to as the dummy-location selection (DLS) algorithm. Basically speaking, the authors used an entropy metric [15] to evaluate the queries submitted

in different locations and generated the dummies in a way to maximize the entropy.

Although the DLS algorithm is promising for a stationary set of the queried locations including the real location and its associated dummies, the algorithm fails to address the privacy issues caused by the consecutive queries made to the LBS provider. In more detail, the authors have limited the side information to queries submitted in different locations but overlooked the fact the adversary has also access to the trajectories, and consequently, the number of times the paths between locations have been traveled. Having access to such extra side information, the adversary can expose the dummies and compromise the k-anonymity of the users. For further explanation, a toy example has been provided in Fig. 1, where

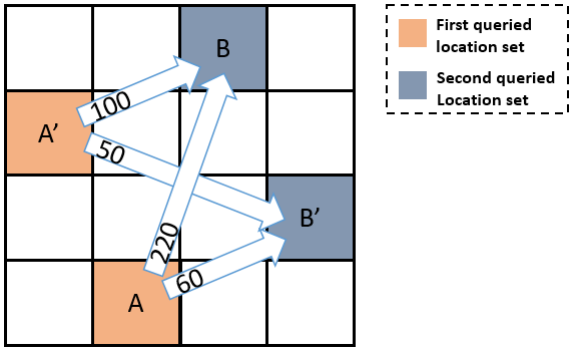

Fig. 1: An example of location privacy of the user being compromised by considering the introduced side information.

we show a user moving from location A to location B with k set to two. The associated dummies of the real locations A and B are denoted by A' and B', respectively. The dummies in each location set are generated using the DLS algorithm, hence, they have a similar probability of being selected. The numbers on the directed edges indicate the number of times that users have queried the end location of the edge right after asking about the starting point of the edge. For instance, the users have queried location B for 100 times immediately after location A'. According to the DLS algorithm, the k-anonymity requirement has been preserved for each location. However, let us look at the four paths connecting the two sets of locations together and consider the number of times that each path has been inquired. It can be seen from Fig. 1 that location B has been inquired for 320 times after locations A and A' whereas location B' has only received 110 times of inquiries. Therefore, the adversary can infer with a high likelihood that the real location is possibly location B, and thus, compromising the location privacy of the user.

To address the existing issues, we incorporate a new type of side information which can be utilized by the adversary to expose the generated dummies, or even real location of users. We propose a metric called transition-entropy which enables us to evaluate the performance of dummy-based algorithms considering the introduced side information. Furthermore, we develop a greedy algorithm which can significantly improve the performance of the proposed metric for a given dummy generation algorithm while maintaining the robust performance in terms of the traditional metric formulated in our work. Finally, we analyze the performance of the metric and algorithm on a real-life dataset.

The rest of the paper is organized as follows. Section II describes the system model used throughout the paper including the system architecture, the adversary model, and the side information used by the adversary. In section III, we introduce our proposed metric followed by proposing a novel algorithm to improve the user's location privacy. Finally, the analysis of the proposed metric and algorithm is provided in section IV, and we conclude our work in section V.

II. System Model

For this paper, we adopt a non-cooperative system architecture [16], as shown in Fig. 2. In this architecture, the LBS users are directly in contact with the LBS provider with no middle-man or a third party service provider.

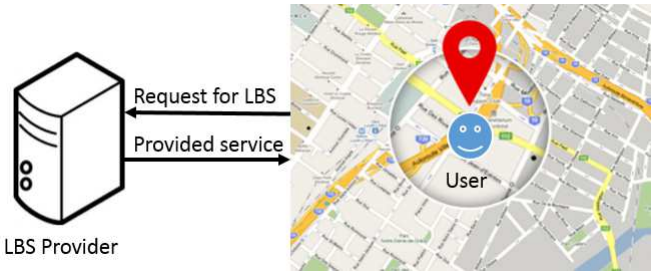

Fig. 2: Non-cooperative system architecture for LBSs.

Assume that the location map is divided into an $n \times n$ grid and a user communicates with an LBS server for a service. At time t^q, the user intends to make his/her q-th query from the service provider, preserving k^q-anonymity. Here, k^q quantifies the privacy protection requirement of the user. This metric implies that the adversary is not able to identify the real location of the user with a probability higher than $\frac{1}{k^q}$. Hence, such user needs to transmit $k^q - 1$ dummy location to hide its true location from the observer. Note that by the term location we refer to the cell in which the user is located. We denote the set of locations transmitted to the LBS provider at q-th query by

$$LS^q = \{l_1^q, l_2^q, ..., l_{k^q}^q\}. \tag{1}$$

Also, the real location is shown by r^q where $r^q \in LS^q$. The probability of location l_x^q being the real value is shown by

$$Pr(l_x^q = r^q) \text{ for } x = 1, ..., k^q. \tag{2}$$

In the next query, the user requires k^{q+1}-anonymity and queries the location set $LS^{q+1} = \{l_1^{q+1}, l_2^{q+1}, ..., l_{k^{q+1}}^{q+1}\}$ from the LBS provider. The probability of location $l_y^{q+1} \in LS^{q+1}$ being queried consecutively after $l_x^q \in LS^q$ is shown by

$$Pr(l_x^q \Rightarrow l_y^{q+1}) \text{ for } x = 1, ..., k^q. \tag{3}$$

Furthermore, two types of adversary models are considered in our work: an active adversary, and a passive adversary. The

passive adversary can eavesdrop the communication between the users and the LBS provider. An active adversary, on the other hand, compromises the LBS provider and has access to the information stored on the server. In our work, the active adversary is assumed to be the LBS provider itself.

The adversary is assumed to possess the location map of the area where the users are distributed. He has access to the queries made by the users and can record them over time to obtain the history of the locations where the users have queried from. Moreover, the adversary can calculate the query probability of different locations in the map, which is defined as the number of times a particular location has been queried. The adversary can exploit the query probability to infer the probability of a location being genuine or fake in the future queries. Furthermore, we assume that the adversary has access to the number of times each path has been traveled on the map. Therefore, the adversary not only has the data on the number of queries made on each location, but it is well-aware of the number of times that a location has been queried consecutively after the other locations.

III. Privacy Metrics and the Corresponding Dummy Location Generation Algorithm

In this section, we briefly explain a metric which was partially developed in [14]. Then, we propose a metric called transition-entropy for consecutive queries from the LBS provider followed by proposing a greedy algorithm which can be applied to any dummy generation method increasing the transition-entropy while maintaining the high performance in terms of cell-entropy.

A. Cell-Entropy Metric

Although not mentioned as a metric in [14], cell-entropy was implicitly proposed as part of the DLS algorithm. We have named this metric cell-entropy to distinguish it from the transition-entropy metric proposed in this paper. For a given location set $LS^q = \{l_1^q, l_2^q, ..., l_{k^q}^q\}$ which includes the real location of a user and $k^q - 1$ dummies chosen to preserve k^q-anonymity, the set of query probabilities are shown by $B^q = \{b_1^q, b_2^q, ..., b_{k^q}^q\}$ where b_j^q is the query probability of location (cell) l_j^q for $j = 1...k^q$. The query of probability of cell l_j^q is calculated by

$$b_j^q = \frac{\text{number of queries in } l_j^q}{\text{number of queries in whole map}}. \tag{4}$$

The cell-entropy borrows the concept of entropy from information theory to quantify the uncertainty in query probability of the locations in LS^q. The cell-entropy metric for location set LS^q can be defined as [14]

$$h = -\sum_{j=1}^{k^q} b_j^q \log_2 b_j^q. \tag{5}$$

Intuitively speaking, the larger the cell-entropy, the better the privacy preservation.

B. Transition-entropy metric for consecutive queries

The main purpose of the metric we propose here is to provide a benchmark for the comparison between dummy-based algorithms taking into account the comprehensive side information we consider in this paper. The metric indicates the susceptibility of the existing algorithms to attacks on location privacy of the users as the k-anonymity requirement of the users can easily be compromised in trajectories. Hence, we need to develop new algorithms for preserving the location privacy of the users. We start by illustrating the metric for two consecutive queries and then generalizing it for trajectories.

Assume that at time t^q a user makes its q-th query and has an anonymity constraint of k^q, and requests the service for the location set of $LS^q = \{l_1^q, l_2^q, ..., l_{k^q}^q\}$. The set LS^q includes $k^q - 1$ dummies and the real location of user. Then, at time t^{q+1} the user moves to a new location with the anonymity constraint of k^{q+1} and makes his $(q+1)$-th query providing the server with the location set of $LS^{q+1} = \{l_1^{q+1}, l_2^{q+1}, ..., l_{k^{q+1}}^{q+1}\}$ consisting of the real location of the user and the associated dummies. The dummies can be generated using any of the existing algorithms.

Using the sets LS^q and LS^{q+1}, we generate a bipartite graph shown in Fig. 3, where each set forms the vertices at a side of the graph. We denote the number of times the location

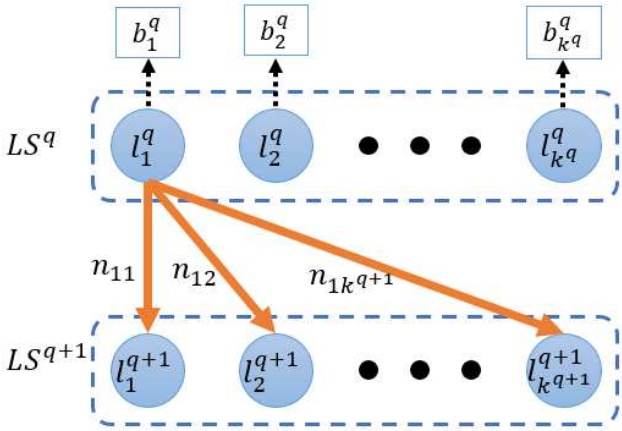

Fig. 3: The bipartite graph generated by the consecutive queries of a user.

$l_y^{q+1} \in LS^{q+1}$ follows the location $l_x^q \in LS^q$ by n_{xy} and assign it to the directed edge connecting l_x^q to l_y^{q+1}. Also, for every location $l_x^q \in LS^q$, we denote query probability of the location l_x^q by b_x^q. The query probability of a cell is calculated by dividing the number of times that cell has been called over the whole number of queries of the map. This data is calculated from the history of data LBS provider holds.

We would like to find out how probable it is for each member of the location set LS^{q+1} to be the real location of the user (r^{q+1}) given the location set LS^q in the previous query from the LBS provider. In other words, the aim is to calculate the posterior probability of the members in LS^{q+1} with respect

to LS^q. This probability for each member of LS^{q+1} can be calculated based on the LS^q as

$$\forall\ l_y^{q+1} \in LS^{q+1}:$$

$$\Pr(l_y^{q+1} = r^{q+1} | LS^q) \tag{6}$$

$$= \sum_{s=1}^{k^q} \Pr((l_s^q \Rightarrow l_y^{q+1}), (l_s^q = r^q)) \tag{7}$$

$$= \sum_{s=1}^{k^q} \Pr(l_s^q \Rightarrow l_y^{q+1} | l_s^q = r^q) \Pr(l_s^q = r^q), \tag{8}$$

where the equation (7) is the joint probability of l_s^q being the real location of LS^q and moving to the location l_y^{q+1} after l_s^q. The former probability in equation (8) can be calculated as

$$\forall\ l_y^{q+1} \in LS^{q+1},\ \forall\ l_x^q \in LS^q:$$

$$\Pr(l_x^q \Rightarrow l_y^{q+1} | l_x^q = r^q) = \frac{n_{xy}}{\sum_{y=1}^{k^{q+1}} n_{xy}}, \tag{9}$$

and the latter probability which indicates the normalized query probability as

$$\forall\ l_x^q \in LS^q:\ \Pr(l_x^q = r^q) = \frac{b_x^q}{\sum_{j=1}^{k^q} b_j^q}. \tag{10}$$

Note that equation (10) indicates that the posterior probabilities of the cells in LS^q are set to the normalized query probability of the locations. Calculating equation (8) for every member of the location set LS^{q+1}, the posterior probabilities of the locations in LS^{q+1} are derived based on the LS^q. Having these probabilities, we exploit the concept of entropy to infer the uncertainty in identifying the dummies or the real location of the users calculated by

$$h = -\sum_{y=1}^{k^{q+1}} \Pr(l_y^{q+1} = r^{q+1} | LS^q) \log_2 \Pr(l_y^{q+1} = r^{q+1} | LS^q). \tag{11}$$

We call h, the transition-entropy of the location set LS^{q+1} with respect to LS^q. The transition-entropy metric represents the uncertainty of identifying the real location by the adversary in consecutive queries from the LBS provider. Having a larger value for the transition-entropy indicates that for each member of LS^{q+1}, the probability of the paths originating from the LS^q to the destination of that member is similar to the other members of LS^{q+1}. Hence, it would be more difficult for the adversary to compromise the k^{q+1}-anonymity of the users based on the transitions made from their previous query. The formal algorithm for calculating the transition-entropy of the location set LS^{q+1} with respect to LS^q is presented in algorithm 1. The main advantages of the metric are as follows: (i) it considers the performance of the dummy-based algorithms in trajectories and not just a stationary set of locations; (ii) it is able to investigate the performance of the dummy-based algorithms for users with varying k-anonymity requirements in their trajectory; (iii) it entails many other factors such as time reachability or direction similarity considered in other works.

Algorithm 1: Calculation of transition-entropy for the location set LS^{q+1} with respect to LS^q.

Input: The location sets LS^q and LS^{q+1}.
Output: The transition-entropy of LS^{q+1} with respect to LS^q.
Initialization: $CellSum = 0$, $h = 0$.
for $1 \le x \le k^q$ **do**
  $EdgeSum = 0$
  for $1 \le y \le k^{q+1}$ **do**
    $EdgeSum = EdgeSum + n_{xy}$
  end
  for $1 \le y \le k^{q+1}$ **do**
    $\Pr(l_x^q \Rightarrow l_y^{q+1} | l_x^q = r^q) = n_{xy}/EdgeSum$
  end
end
for $1 \le x \le k^q$ **do**
  $CellSum = CellSum + b_x^q$
end
for $1 \le x \le k^q$ **do**
  $\Pr(l_x^q = r^q) = b_x^q/CellSum$
end
for $1 \le y \le k^{q+1}$ **do**
  $\Pr(l_y^{q+1} = r^{q+1} | LS^q) = 0$
  for $1 \le x \le k^q$ **do**
    $\Pr(l_y^{q+1} = r^{q+1} | LS^q) = \Pr(l_y^{q+1} = r^{q+1} | LS^q)$
      $+\Pr(l_y^{q+1} = r^{q+1} | l_x^q = r^q) \Pr(l_x^q = r^q)$
  end
  $h = h -$
    $\Pr(l_y^{q+1} = r^{q+1} | LS^q) \log_2(\Pr(l_y^{q+1} = r^{q+1} | LS^q))$
end
return h

C. Greedy Algorithm

Suppose that at time t^q the user has made its q-th query for the location set of $LS^q = \{l_1^q, l_2^q, ..., l_{k^q}^q\}$ which includes the real location and its associated dummies. As the user changes its location and makes his $(q+1)$-th query at time t^{q+1}, assuming k^{q+1}-anonymity for the user, we wish to generate the location set $LS^{q+1} = \{l_1^{q+1}, l_2^{q+1}, ..., l_{k^{q+1}}^{q+1}\}$ to maximize the transition-entropy metric. The principal idea is to choose the members one by one from a pool of dummies such that they maximize the transition-entropy. The greedy algorithm is formally presented in Algorithm 2. The algorithm starts by generating a pool of dummies using the DLS algorithm. The DLS algorithm has been chosen due to its robust performance in terms of cell-entropy. The proposed algorithm here is applicable for other dummy generation methods as well. Initially, the location set LS^{q+1} only includes the real location of the user at $(q+1)$-th query. The next member is added by trying out all the members in D and calculating the transition-entropy of LS^{q+1} including that member and choosing the one

Algorithm 2: The proposed greedy algorithm for location privacy preservation of the users.

Input: k^{q+1}, the location set $LS^q = \{l_1^q, l_2^q, ..., l_{k^q}^q\}$ in q-th query, and the location set $LS^{q+1} = \{l_1^{q+1}\}$ which only includes the real location of the user at $(q+1)$-th query.
Output: the location set LS^{q+1} which includes the real location and $(k^{q+1}-1)$ dummies.
Initialization: .
$D \leftarrow$ generate a pool of $4k^{q+1}$ dummies using the DLS algorithm
for $1 \leq member \leq k^{q+1}-1$ **do**
  $entropy = zeros(1 \times |D|)$
  for $1 \leq d \leq |D|$ **do**
    $LS^{q+1} = LS^{q+1} \cup \{D_d\}$
    $entropy[d] \leftarrow$ transition entropy of LS^{q+1} w.r.t. LS^q
    $LS^{q+1} = LS^{q+1} - \{D_d\}$
  end
  $NewMember \leftarrow$ {member of D which maximize $entropy$}
  $LS^{q+1} = LS^{q+1} \cup \{NewMember\}$
  $D = D - \{NewMember\}$
end
return LS^{q+1}

which maximizes the transition-entropy. Then, we proceed to the selection of the third member and the same procedure is repeated until all the $k^{q+1}-1$ dummies are chosen.

IV. Performance Evaluation

A. Experiment Setup

In our experiment, we use the data collected by Geolife project [17]–[19], which includes the GPS trajectories of 182 users from April 2007 to August 2012 in Beijing, China. We have conducted our experiments on 1km $\times$ 1km central part of the Beijing map with the resolution of 0.01km $\times$ 0.01km for each grid cell. The location privacy requirement (k) of the users are investigated for the values 2 to 30. For each value of k, the algorithms are repeated 3000 times to ensure the reliability of the results. Although the proposed algorithm and metric can be used for the users who have varying location privacy requirements in consecutive queries of the LBS, for the sake of comparison, we have assumed that the k value stays the same in consecutive calls for the LBS. Additionally, the experiments are performed on a PC with a 3.40GHz core-i7 Intel processor, 64-bit Windows 7 operating system, and 8.00GB of RAM. Moreover, Python program is used to implement the algorithms.

B. Cell-Entropy Performance

In order to calculate the cell-entropy metric, the adversary records the number of times each cell has been queried over time, and using this information calculates the query probability of each cell. Once the dataset including the real location and dummies are submitted to the server, the adversary can calculate the cell-entropy of the user. A higher value for the cell-entropy indicates more uncertainty in finding the real location or recognizing the dummies. Therefore, maximum cell-entropy is desirable to maintain the k-anonymity of the users.

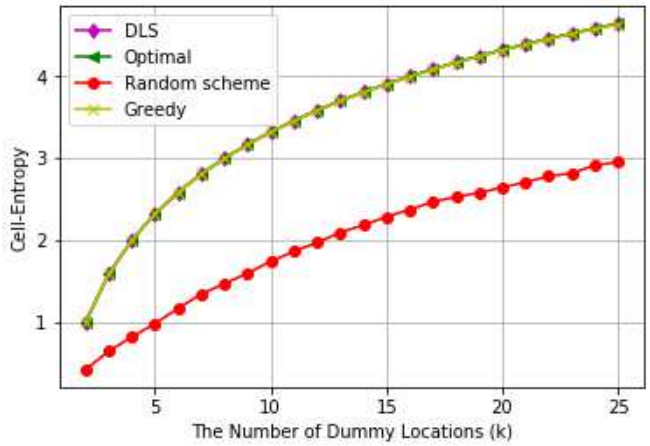

Fig. 4: Comparison of algorithms in terms of cell-entropy for different values of k.

Fig. 4 represents the comparison of different algorithms in terms of cell-entropy. The optimal value is achieved when the k locations queried form the LBS provider all have the same probability of $\frac{1}{k}$, or equivalently, the location set has the cell-entropy of $h = \log_2 k$. The optimal value is the target for all the algorithms since it is the maximum entropy that a location set can achieve. In the random scheme [11], the dummies are generated randomly which expectedly results in a lower cell-entropy compared to the other algorithms. As it can be seen in the figure, the DLS algorithm achieves near-optimal performance in terms of the cell-entropy. Therefore, the adversary is unable to compromise the k-anonymity of the user from the stationary set of locations submitted to the server using the available query probabilities. The greedy algorithm can also achieve near-optimal performance. It must be noted that the proposed greedy algorithm is adaptable to any dummy generation algorithm, therefore, the reason for a high cell-entropy performance of the greedy algorithm is that we have chosen DLS as our baseline algorithm. Hence, if other algorithms are chosen, the cell-entropy performance must be evaluated for them as well to ensure the robust performance in terms of the cell-entropy.

C. Transition-Entropy Performance

The currently established cell-entropy metric only considers the location privacy for the stationary set of queried locations submitted to the LBS server, but overlooks the fact that the adversary has access to the trajectories traveled by the users as

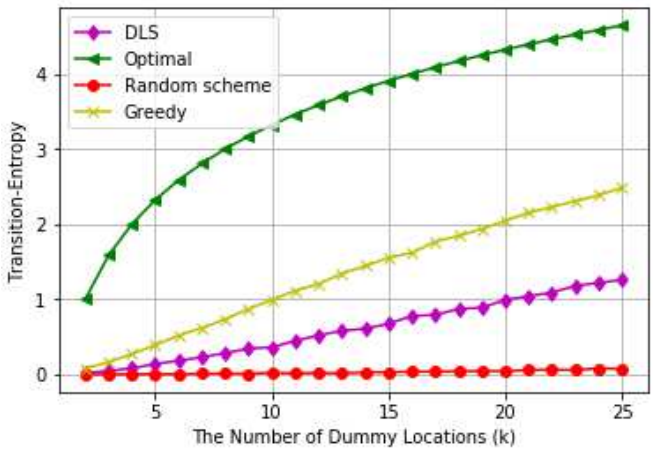

Fig. 5: Comparison of algorithms in terms of transition-entropy for different values of k.

well. Fig. 5 compares the performance of different algorithms in terms of the transition-entropy for a path length of two. For all the algorithms, based on the value of k, two consecutive location sets are generated, each including the real location and its associated dummies.

The optimal value in Fig. 5 corresponds to a scenario in which all the members in the second location set have the same likelihood to be queried immediately after the members in the first location set. The optimal values can be calculated in a similar way as the optimal number for the cell-entropy for different values of k. As it can be seen from the figure, the random scheme has a very poor performance which means that the adversary can easily recognize most of the dummies from the transition-entropy even for the two consecutive location sets queried by the user. It is important to note that although DLS algorithm achieved a near-optimal performance in terms of cell-entropy, its transition-entropy performance indicates that the adversary can violate the location privacy of the users by calculating the posterior probabilities of location transitions. The transition-entropy of the proposed greedy algorithm in this paper is shown to significantly improve the transition-entropy performance, almost doubling the performance compared with the DLS algorithm. In other words, the k-anonymity requirement is is more likely to be achieved by the proposed algorithm which leads to a higher location privacy for the users of LBSs.

V. Conclusions

In this work, we considered new side information which can be exploited by the adversary to compromise the location privacy of the users. We proposed a metric called transition-entropy to evaluate the performance of the dummy-based algorithms. The metric is based on the transitions between two consecutive locations in the map and considers the deplorable effect of new side information on location privacy of the users. Furthermore, we developed a greedy algorithm which can significantly improve the transition-entropy metric for a give dummy generation algorithm. Performing numerous experiments on a real-life dataset, we demonstrated enhanced performance of the proposed algorithm in terms of the transition-entropy while also maintaining the high performance in terms of the traditional cell-entropy metric.

References

[1] "Location-based services (lbs) and real time location systems (rtls) market by location (indoor and outdoor), technology (context aware, uwb, bt/ble, beacons, a-gps), software, hardware, service and application area - global forecast to 2021." [Online]. Available: https://www.marketsandmarkets.com/Market-Reports/location-based-service-market-96994431.html

[2] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive computing*, vol. 2, no. 1, pp. 46–55, 2003.

[3] T. Jiang, H. J. Wang, and Y.-C. Hu, "Preserving location privacy in wireless lans," in *Proceedings of the 5th international conference on Mobile systems, applications and services*. ACM, 2007, pp. 246–257.

[4] C.-Y. Chow, M. F. Mokbel, and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location-based service," in *Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems*. ACM, 2006, pp. 171–178.

[5] B. Liu, W. Zhou, T. Zhu, L. Gao, and Y. Xiang, "Location privacy and its applications: a systematic study," *IEEE access*, vol. 6, pp. 17 606–17 624, 2018.

[6] S. Shaham, M. Ding, B. Liu, Z. Lin, and J. Li, "Privacy preservation in location-based services: A novel metric and attack model," *arXiv preprint arXiv:1805.06104*, 2018.

[7] ——, "Machine learning aided anonymization of spatiotemporal trajectory datasets," *arXiv preprint arXiv:1902.08934*, 2019.

[8] S. Shaham, M. Kokshoorn, Z. Lin, M. Ding, and Y. Wu, "Raf: Robust adaptive multi-feedback channel estimation for millimeter wave mimo systems," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–6.

[9] L. Sweeney, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 557–570, 2002.

[10] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st international conference on Mobile systems, applications and services*. ACM, 2003, pp. 31–42.

[11] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in *Pervasive Services, 2005. ICPS'05. Proceedings. International Conference on*. IEEE, 2005, pp. 88–97.

[12] B. Niu, Z. Zhang, X. Li, and H. Li, "Privacy-area aware dummy generation algorithms for location-based services," in *Communications (ICC), 2014 IEEE International Conference on*. IEEE, 2014, pp. 957–962.

[13] H. Lu, C. S. Jensen, and M. L. Yiu, "Pad: privacy-area aware, dummy-based location privacy in mobile services," in *Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*. ACM, 2008, pp. 16–23.

[14] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving k-anonymity in privacy-aware location-based services," in *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014, pp. 754–762.

[15] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *International Workshop on Privacy Enhancing Technologies*. Springer, 2002, pp. 41–53.

[16] R. Cheng, Y. Zhang, E. Bertino, and S. Prabhakar, "Preserving user location privacy in mobile data management infrastructures," in *International Workshop on Privacy Enhancing Technologies*. Springer, 2006, pp. 393–412.

[17] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from gps trajectories," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 791–800.

[18] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on gps data," in *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 2008, pp. 312–321.

[19] Y. Zheng, X. Xie, and W.-Y. Ma, "Geolife: A collaborative social networking service among user, location and trajectory." *IEEE Data Eng. Bull.*, vol. 33, no. 2, pp. 32–39, 2010.