Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 6): pp. 4208-4216
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.06.402

Location Privacy-Preserving Mobile Crowd Sensing Using Pseudonyms Approach

Dhirendra Kumar Tripathi Research Scholar, Department of Computer Science, Sri Satya Sai University of Technology & Medical Sciences, Sehore, M.P., India.

Dr. Jitendra Sheethlani Research Guide, Department of Computer Science, Sri Satya Sai University of Technology & Medical Sciences, Sehore, M.P., India.

ABSTRACT

As mobile technology has advanced, the issue of privacy leakage has emerged as a primary area of study within the subject of mobile crowdsourcing. Therefore, this study presents two methods for protecting users' location privacy, each of which relies on anonymizing factors such pseudonyms and separate data collectors/senders. As a first step, we provide a well-known current method that involves transmitting data to a centralised server while using pseudonyms. After that, we provide a rather straightforward method called DDCS, which employs a distributed network of independent data collectors and transmitters (i.e., one user collects information and swaps with other user that transmits collected information to the central server or platform). Finally, we provide a hybrid method that combines pseudonyms with DDCS to improve privacy. The simulation results demonstrated that the hybrid strategy provides the highest level of privacy protection compared to the pseudonym-based approach and the DDCS.

Keywords: Crowdsensing, Mobile users, Hybrid, Simulation, Location

I. INTRODUCTION

Consumer-focused mobile sensing and computing devices, such as cellphones, music players, and in-vehicle sensors, are an up-and-coming category of devices near the Internet's periphery. These gadgets are crucial to the development of the IoT because of the massive amounts of sensor data they will contribute to the Internet. The concept of mobile crowd sensing (MCS) has been gaining traction recently, with several systems and apps developed to utilize users' mobile devices to measure ambient context. An MCS consists of the governing body that hosts the platform on which the application runs and the users who make use of the programme to gather data. Detecting earthquakes, as well as monitoring city noise, climate, people density, emergency behavior, traffic abnormalities, and other conditions, are only some of the many uses for MCS.

Because there are so many people already using cell phones, MCS applications may collect information in ways never before possible, in locations previously out of reach, and at low cost. In the case of traffic congestion applications, for instance, MCS has the ability to gather real-time data from secondary and even tertiary highways, which is now prohibitively expensive with conventional technology but would be greatly beneficial. The funding, installation, and maintenance expenses associated with deploying static sensors along all highways are high.

Typically, mobile users' geolocation are included in the MCS data. An attacker might potentially pinpoint a mobile user's position and then act accordingly. Pseudonymous pseudonyms can help protect mobile users from prying eyes, but an enemy can still use their whereabouts to deduce sensitive details about the users, such as their political leanings, their hobbies, and even whether or not they have any health issues.

Therefore, securing people's right to privacy in their locations is crucial. Large amounts of research have been devoted to protecting users' privacy in Location Based Service (LBS) and crowd sensing systems since users frequently do not trust the underlying servers or platforms. Several strategies were offered to protect users' right to anonymity in their locations by preventing servers and platforms from determining their precise whereabouts.

II. PRIVACY PRESERVING APPROACHES

Here, we introduce an existing system based on pseudonym that has been widely utilized to offer privacy and security in VANET and other new technologies. Then, we describe two proposed methods of protecting users' location privacy when communicating with a centralized server, such that the centralized server is unable to determine users' whereabouts or the paths they took when moving.

These methods take into account the fact that a large number of people spread out throughout a specific area and at a variety of places are tasked with gathering information over a period of time. They gather data about a given environment, such the ambient noise level, mobile signal strength, Wi-Fi fingerprint, and temperature. Also, once a device has started taking this kind of take or collecting this kind of data, it should be able to do so without being connected to the platform for a while.

Existing Pseudonym-based Approach

Here, the node ID is used as a seed to construct a string of pseudonyms, which are then distributed around the network. The node always includes the pseudonym as the data collector ID whenever it collects information from a specific place inside a specified data collection region. As a result, there are several aliases spread throughout each collection site. All data collected by a node will be sent to a central server or platform, but the server will be unable to determine the node's location or identity because to the use of a series of pseudonyms. As a result, nodes and devices may maintain their locational anonymity.

This method is helpful for avoiding disclosure of private information because the user is not using their real name when interacting with the service. This implies that the platforms are able to communicate with any service or user without requiring any identifying information [3]. When a user is in close proximity to a retail establishment, for instance, they may receive a notification about the current pricing structure. Although the system cannot identify the user, it does know that they are interested in this service. Pseudonyms are used to protect real identities from prying eyes. Pseudonymous information (such as an individual's residence and place of employment) can be utilised to create a detailed profile and ultimately reveal their true identities. An effective countermeasure is to use pseudonyms that are constantly being updated. Therefore, in the proposed method, users utilise both a consistent set of pseudonyms across all services and a unique set of pseudonyms for each service.

In addition, the node that collects the data is also the node that transmits the data to the central server; in this setup, nodes do not communicate with their counterparts in the crowd. Therefore, the central server may be able to determine the node's ID by doing calculations on the pseudonyms of several data packets and producing a mathematical equation to tie pseudonym to actual ID or vice versa (a backwards procedure to utilise a set of pseudonyms to produce the node's ID). Further, it may not be energy efficient for each device to gather data and send it to a centralised server. This method also does not take advantage of the fact that devices may interact with one another to conceal their true identities by exchanging data upon meeting.

Proposed Approaches

At first, we present a basic method we call Different Data Collector and Sender (DDCS), which is suitable for lightweight networked applications. Then, we offer a hybrid technique that combines DDCS with the current pseudonym-based approach.

- **Different Data Collector and Sender (DDCS) –** When users relocate to a new area, they start collecting data there as well (i.e., temperature, humidity and pressure information). If users, for instance, record temperatures at ten different sites in a given region and over the course of a given time period, this might provide useful information. Here, we presume that all parties involved in a transaction can be trusted except for the platform or server on which the transaction takes place. In other words, the server shouldn't be privy to users' whereabouts or private data. Every time a user crosses paths with another, they exchange their info to keep their whereabouts secret. The more users share information with one another, the less likely it is that the platform or central will be able to determine a user's precise position when that user finally sends information to the central at the conclusion of the time. That is, the platform or central server will be unable to determine which user collected which data and where it came from at the conclusion of any given time period, regardless of how often the user reconnects. Keep in mind that GPS coordinates might be part of the data that has been gathered. This is due to the fact that the person or entity collecting the data is not necessarily the same person or entity sending the data to the centralized server. The data collector may have to upload all obtained information to a server if he or she lives in a less densely populated location, such as a rural area. Therefore, the collector's position may be determined by the server. However, this security is broken if an unauthorized third party compromises the network by impersonating legitimate users.

- **Hybrid Approach –** This method combines the data-gathering and anonymous techniques into a single whole. Pseudonyms are created at the node level in this method. The device or node always appends a pseudonym to the data it gathers at a given place, according to the data collecting need. In addition, the node communicates with other nodes anytime it comes across another device or data in the crowd, on the assumption that the data are brought to the central server by other nodes. In this setup, the data collector is different from the data transmitter. As a result, it is challenging for the platform or any intruder to ascertain the position or ID of the gadget. Data packets in the proposed hybrid solution (shown in Figure 1) contain both the pseudonym and the real data. In Figure 1, each data packet contains the pseudonym of the node that is exchanging it with a neighbouring node.

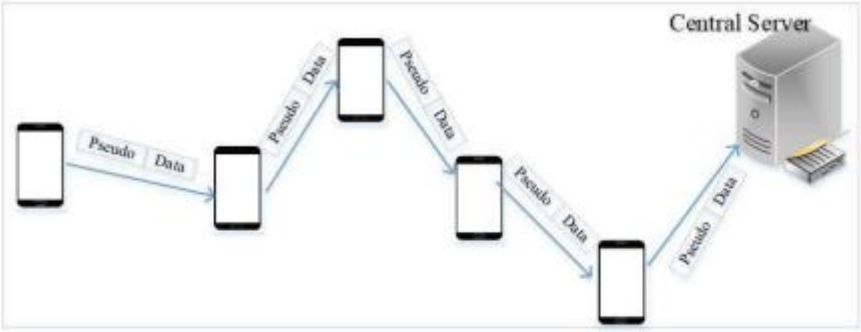

Figure 1: Proposed Hybrid Approach

Most methods don't think about paying people or gadgets to take part in the crowdsensing process. To increase privacy, this method incorporates applying incentive to users/devices to swap information with other users to successfully send data to the central server. Each node's reward will be proportional to the total number of swaps they participate in. During the swap procedure, priority is given to nodes based on their total reward points. If two nodes A and B meet or pass each other, the node with the greater reward points can trade data with the other. Devices will be prompted to join in this form of connection as a result of this procedure.

III. SIMULATION SETUP AND RESULTS

We simulate the performance of the various methods to compare them to the current pseudonym-based approach and to determine which method is most effective at protecting users' location privacy during crowdsensing.

Setup

As can be seen in Figure 2, the simulated network region is assumed to be of a square form with a tile layout (i.e. split into a number of little squares). Each tile in the network is assumed to be 10 metres by 10 metres in size. To be regarded "not near enough," two roaming users must have non-overlapping tiles in their immediate vicinity. In Figure 2, adjacent tiles for Users #1 and #2 do not overlap, hence these Users are not "close enough" to share data.

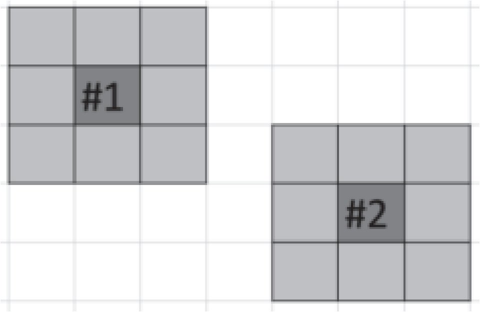

Figure 2: Network Model with "Not Near Enough" Users

Figure 3 show that if the tiles around any two users overlap, the users are judged to be "near enough."

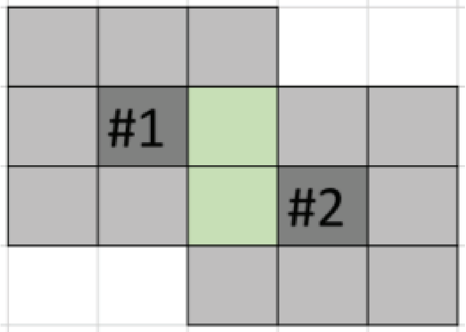

Figure 3: Network Model with "near enough" Users

We simulate the network for some time with a fixed number of users moving at a fixed speed across a fixed region. If two users X and Y exchange data, X will not immediately get Y's copy of his data if they remain in close proximity forever; instead, a cooldown will be triggered.

We next evaluate how each user's journey matched up with the information they uploaded to the server or platform at the conclusion of the simulation. The proportion of successfully traced data increases if the given location and time match the actual data. If the platform doesn't know where the user has been, it can't follow their movements. Therefore, confidentiality is maintained. In Table 1 you'll see the values used for each of the simulation's parameters.

Table 1: Simulation Parameter and Values

Parameter	**Values**
Network Shape	Square
Side length of each tile in the network	10 meters x 10 meters
Number of users per 1000 meter2	5 ~ 40
Simulation Time	900 seconds
Movement of the users	Random
Movement speed	2-3 km/Hour

Results

We evaluate how well the suggested methods function in comparison to a current pseudonym-based method by looking at the number of records that are tracked or matched by the centralised server. The results from the simulations are shown in Figures 4-6. If we hold the number of users in the network and the duration of the simulation constant at 30 and 900 seconds, respectively, then Figure 4 shows what proportion of the data is tracked by the central server as the network grows in size. Figure 4 shows that the DDCS method yields a substantially larger percentage of successfully traced data compared to the pseudonymous and hybrid methods. When compared to DDCS and pseudonyms, the hybrid strategy proves to be the most effective. In addition, for a given number of users, increasing the network capacity increases the proportion of data tracked. This is because it is more likely that the same user will be both the data collector and the data transmitter as the network size rises, causing the number of users per square meter to fall and resulting in a sparse network. Therefore, it is much simpler for the central server to track the device or user. However, if users attach pseudonyms to their data packets whenever they get access to new information or interact with new people, the central server would have a hard time keeping up.

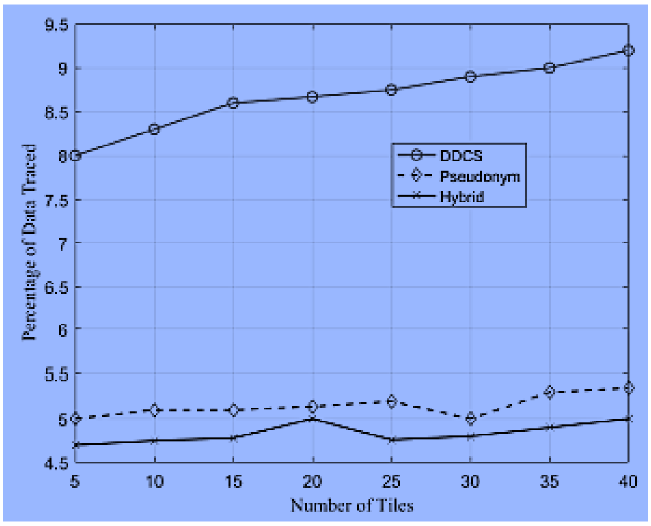

Figure 4: Percentage of Data Traced varying the side length of the network

The proportion of data traced as the number of users per square metre changes is shown in Figure 5. With a larger user base comes a greater chance that information may be shared between users, and that the person(s) sending information to a centralised server will be different from the people(s) collecting it in various locations. In this way, the amount of information that can be tracked will gradually decrease. Furthermore, the percentage of data tracked using pseudonyms and hybrid approaches (that employ pseudonyms and data swap) is substantially lower than with DDCS. Figure 6 depicts this trend clearly by showing how the fraction of traced data reduces as simulation duration increases using a variety of methods.

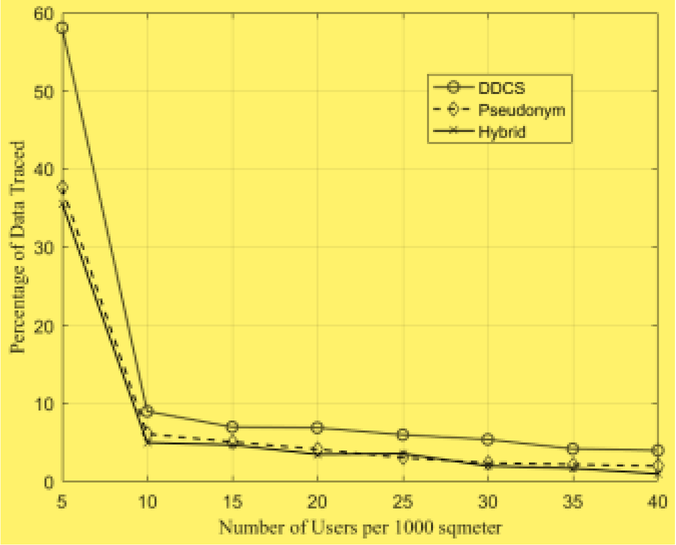

Figure 5: Percentage of Data Traced varying the number of users per 1000 square meters

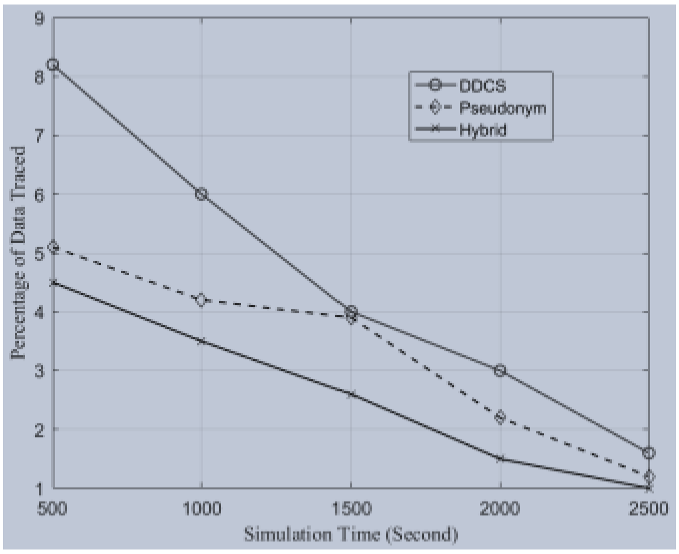

Figure 6: Percentage of Data Traced varying the Simulation time

IV. CONCLUSION

Utilizing the sensors already present in smart phones, "people-centric sensing" may be utilised for low-cost, large-scale sensing of the real environment. Although it offers several advantages, mobile people-centric sensing has two key challenges: (i) motivating the participants, and (ii) ensuring the accuracy of the sensed data. Unfortunately, current approaches to resolving these issues either need investment in infrastructure support or impose substantial additional burdens on users' mobile devices. Once the problems with data dependability are resolved, we expect mobile crowd sensing to become a common way for gathering sensing data in the real world. For complicated sensing tasks, mobile crowd sensing (MCS) makes use of the widespread availability of smartphones with many sensors. In order to provide a high-quality sensing service, mobile users must, on the one hand, reveal personal information (such as their names, where they are, what they're interested in, etc.) to other users. However, mobile users are less likely to volunteer for a sensing activity if they are not compensated for their time. As a result, it's crucial to think about how to protect users' personal information and design an appropriate incentive structure.

REFERENCES:

1. Wang, X., Liu, Z., Tian, X., Gan, X., Guan, Y., Wang, X.: Incentivizing crowdsensing with location-privacy preserving. IEEE Trans. Wirel. Commun. 16(10), 6940–6952 (2017)

2. Vergara-Laurens, I.J., Jaimes, L.G., Labrador, M.A.: Privacy-preserving mechanisms for crowdsensing: survey and research challenges. IEEE IoT J. 4(4), 855–869 (2017)

3. Christin, D.: Privacy in mobile participatory sensing: current trends and futurechallenges. J. Syst. Softw. 116, 57–68 (2016)

4. Bellavista, P., Corradi, A., Foschini, L., Ianniello, R.: Scalable and costeffectiveassignment of mobile crowdsensing tasks based on profiling trends and prediction: the participact living lab experience. Sensors 15(8), 18613–18640 (2015)

5. Ren, J., Zhang, Y., Zhang, K., Shen, X.S.: SACRM: social aware crowdsourcing with reputation management in mobile sensing. Comput. Commun. 65, 55–65 (2015)

6. Guo, B., Calabrese, F., Miluzzo, E., Musolesi, M.: Mobile crowd sensing: part 2. IEEE Commun. Mag. 52(10), 76–77 (2014).

7. To, H., Ghinita, G., Shahabi, C.: A framework for protecting worker location privacy in spatial crowdsourcing. In: Proceedings of VLDB 2014, pp. 919–930 (2014).

8. D. Yang, G. Xue, X. Fang, and J. Tang, “Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing,” in Proc. of ACM MobiCom, 2012, pp. 173–184.

9. Shina, M., Cornelius, C., Peebles, D., Kapadia, A., Kotz, D., Triandopoulos, N.: Anony Sense: a system for anonymous opportunistic sensing. Pervasive Mobile Comput. 7, 16–30 (2011).