IIAI International Journal of Service and Knowledge Management
International Institute of Applied Informatics
2019, Vol. 3, No. 2, 1 – 14

Privacy Protection for Multi-Option Problem of Participatory Sensing Using Random Noise Addition

Tomomichi Hayakawa [*], Teruhisa Hochin [†], Tokuro Matsuo [‡]

Abstract

The rapid proliferation of smartphone-mounted multiple sensors has been accompanied by the increasing utilization of participatory sensing, which is a type of crowdsourcing by which many users effectively share sensing data. Privacy protection is important for this purpose because the sensing data often contain private information about the users. Existing privacy protection methods do not enable effective and precise data restoration in this application when there is many choices and few data. In this study, we developed a method for addressing this issue. The randomized response method and negative survey method are used to conceal private information contained in individual data by the addition of random noise to the data. Moreover, the proposed method utilizes a novel procedure whereby the transmission is repeated multiple times when selecting one option from multiple options. The proposed method is evaluated by simulation and is found to be more effective than existing methods.

Keywords: Participatory Sensing, Privacy Protection, Randomized Response, Negative Surveys

1 Introduction

Participatory sensing [1] has attracted much attention in recent years due to the widespread use of smartphones equipped with many sensors. Participatory sensing is a type of crowdsourcing that enables effective sharing of information obtained by sensors mounted on the smartphones of multiple users. In conventional sensing, dedicated equipment is required for the acquisition of the sensing information, for which the process can be both time-consuming and costly. Participatory sensing affords significant cost savings and utilizes the sensor-equipped smartphones of ordinary users. The users acquire sensing information through sensors such as GPS receivers, microphones, and acceleration sensors installed on their smartphones and report the information to participatory sensing administrators. A participatory sensing operator collects all the data and aggregates and analyzes them.

[*] National Institute of Technology, Ichinoseki College, Japan
[†] Kyoto Institute of Technology, Kyoto, Japan
[‡] Advanced Institute for Industrial Technology, Tokyo, Japan

However, there are some privacy problems associated with participatory sensing. The sensing data collected by participants include private information such as their action history and a malicious attacker may eavesdrop on data received by the administrator. Some participants may thus be concerned about the compromise of their privacy and may refrain from sharing sensing information with administrators. For participants to safely and confidently supply participatory sensing information, it is necessary for administrators to apply a privacy protection method that ensures that only necessary information is collected.

In this study, we developed a privacy protection method for participatory sensing involving many options. The proposed method represents an extension of the randomized response method, which is a current privacy protection method for participatory sensing. The randomized response method and negative survey method (another existing privacy protection method) are used to conceal private information contained in individual data by the addition of random noise to the data. These methods are suitable when there is a large amount of sensing data and a small number of options. However, when applied to a case with a small amount of data and many options, the restoration accuracy of the data distribution in the server would be low, which is unacceptable in practice. In this paper, the restoration accuracy is the accuracy of the difference between the sensing data distribution transmitted by all the participants and the data distribution restored based on the data received by the administrator using the particular privacy protection method. In the proposed method, we utilize a technique that enables effective restoration of the data distribution in the server, even when the number of choices observed by the sensor is large and the sensing data sent to the server is small. Experimental simulations were performed to evaluate and compare the restoration accuracy of the proposed method with those of existing methods. The results confirmed the effectiveness of the proposed method.

The rest of this paper is organized as follows. Section 2 describes two existing privacy protection methods, the problems of which are further discussed in Section 3. Section 3 also presents the proposed privacy protection method, which is effective for cases with a small amount of sensing data and many options. In Section 4, we compare the proposed method with existing methods based on the results of simulation experiments. Section 5 finally summaries the study and its findings.

2 Related Work

2.1 Participatory Sensing

Many studies have been conducted on developing a privacy protection method for participatory sensing that is optimally applicable to various cases. In this section, to properly identify the nature of the proposed method, we will first describe and classify different existing methods based on the process of the participatory sensing and the subjects.

Cryptography technology [7] [8] has been used to protect private information from eavesdropping by a malicious third party during the transmission of data between a participant an administrator of participatory sensing. Two encryption methods have been employed for this purpose, namely, public key encryption and common key encryption. In both methods, the participant on the data transmission side encrypts the data and the administrator on the data reception side restores the encrypted data. The restored data may be different in some ways from that before encryption.

A method referred to as the k-anonymity is one of those used to protect the private information of users in data contained in a database. This protection method makes any k-1 user indistinguishable from other users [9] [10]. Differential privacy protection is another method used to protect not only individual data but also statistical data outputted from private data [11] [12].

Random noise addition has also been used to protect private information contained in sensing information sent by a participant to an administrator in participatory sensing. This method involves adding a value other than the true value to the acquired data during the transmission to the administrator. Examples of this type of protection method are randomized response [13] [14] and negative surveys [17] [18]. Multidimensional negative surveys [19] [20] is an extension of randomized response and is particularly applied to multidimensional data containing multiple fields of information. The method is, however, only effective when the number of options observed by the sensor is small and the amount of acquired data transmitted is large. However, when there are many options and a small amount of acquired data, the restoration accuracy achieved by this method is insufficient. In the use of the random noise addition method, which is the basis of the method proposed in this paper, even when the number of choices observed by the sensor is large and the number of acquired data is small, the administrator is able to obtain a restoration accuracy equivalent to that of existing methods in the converse situation. The proposed method utilizes a novel procedure whereby the transmission is repeated k times when selecting one option from multiple options. Even for a small number of sensing data and many options, the administrator is able to obtain the distribution of the entire sensing data collected by each sensing device, and may also achieve the same restoration accuracy as existing methods.

2.2 Randomized Response

Randomized response [13] is basically the alternatives method proposed by S. L. Warner. It enables respondents to answer without revealing private information, as well as determination of the statistical distribution of all the respondents.

Agrawal et al. extended the application of the randomized response method to cases in which a single answer is given to the multiple options of a category [14]. In this application, the private information of the user is protected by transmitting a true value different from the true acquired from the sensing device, using a certain probability. Randomized response is used in various fields that utilize position information and participatory sensing [15] [16].

The procedure when using randomized response for privacy protection in the transmission of participatory sensing data is as follows.

(1) The sensor acquires sensing data and the sensing device observes one sensing data option among the total number of options, α.

(2) The privacy protection process is implemented on the sensing data, which is then transmitted to the server. The sensing device transmits is sensing data to the server with a preset probability p. Alternatively, the other α-1 options among the obtained sensing data are selected with a probability $\frac{1-p}{\alpha-1}$ and transmitted to the server. Let A be the data distribution of the entire data collected by the sensing device. After the implementation of the privacy protection process, the distribution Y of the entire sensing data transmitted to the server is given by Equation (1). The privacy protection process can be expressed by a square matrix M of size $\alpha \times \alpha$, as in Equation (2).

(3) The server reconstructs the statistical information of the sensing data based on the received data, using Equation (3) and the distribution A of the entire sensing data. This is done without knowing the true value collected by each sensing device, thus ensuring privacy protection for all users.

$$Y = AM \tag{1}$$

$$M = \begin{pmatrix} p & \frac{1-p}{\alpha-1} & \cdots & \frac{1-p}{\alpha-1} \\ \frac{1-p}{\alpha-1} & p & \frac{1-p}{\alpha-1} & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1-p}{\alpha-1} & \cdots & \cdots & p \end{pmatrix} \tag{2}$$

$$A = YM^{-1} \tag{3}$$

2.3 Negative Surveys

The negative surveys privacy protection method is an extension of randomized response developed by Esponda et al. [17] [18]. In Step (2) of the randomized response method, the probability of transmitting sensing data to the server in its true form is only set to $p=0$. The unique feature of the method is that only the options different from the true value are transmitted to the server, without the need to transmit a true value.

However, information about each sensing device (that is, which device observed which value) cannot be obtained when only the alternatives different from the true value are sent to the server. Under this situation, the server can obtain only the distribution of the entire sensing data received after their aggregation. Therefore, compared with randomized response, when using negative surveys, it is more difficult for the server to make an inference from the sensing data based on individual sensing devices.

The procedure when using negative surveys for data protection in the transmission of participatory sensing data is as follows.

(1) The sensing device acquires the sensing data and observes one option among the total number of options, α.

(2) The sensing device implements the privacy protection process on the sensing data and transmits the data to the server. The sensing device transmits α-1 options different from the obtained sensing data to the server with a probability of $\frac{1}{\alpha-1}$. Let A be the distribution of the entire sensing data collected by the sensor device, and Y the distribution of the entire sensing data received by the server after the implementation of the privacy protection process. The privacy protection process can be expressed as in Equations (4) and (5).

(3) The server reconstructs the statistical information of the sensing data based on the data received from the sensing device. The server then obtains the data distribution A of the sensing data acquired by each sensing device using Equation (6). In addition, for a number of acquired data N, the data distribution A can also be obtained by Equation 7.

$$Y = AM \tag{4}$$

$$M = \begin{pmatrix} 0 & \frac{1-p}{\alpha-1} & \cdots & \frac{1-p}{\alpha-1} \\ \frac{1-p}{\alpha-1} & 0 & \frac{1-p}{\alpha-1} & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1-p}{\alpha-1} & \cdots & \cdots & 0 \end{pmatrix} \quad (5)$$

$$A = YM^{-1} \quad (6)$$

$$\forall_i \mid A(i) = N - (\alpha - 1) \cdot Y(i) \quad (7)$$

3 Proposed Privacy Protection Method

3.1 Issues Requiring Solution

Figure 1 compares cases of using randomized response, negative surveys, and a data transmission method without privacy protection. In negative surveys, the probability of transmitting a true value is $p = 0$, and data other than the true data are randomly transmitted. In randomized response, the probability of transmitting a true value is $0 < p < 1$, and $p \neq \frac{1}{\propto}$. When the probability of transmitting a true value is $p = \frac{1}{\propto}$, the transmission probability of all the values is also $\frac{1}{\propto}$. This is because perfect random transmission will be achieved, and it will be impossible to restore the transmitted value to a true value. Furthermore, when the probability of transmitting a true value is $p = 1$, only the true value will be transmitted, and a privacy protection method would not be applied.

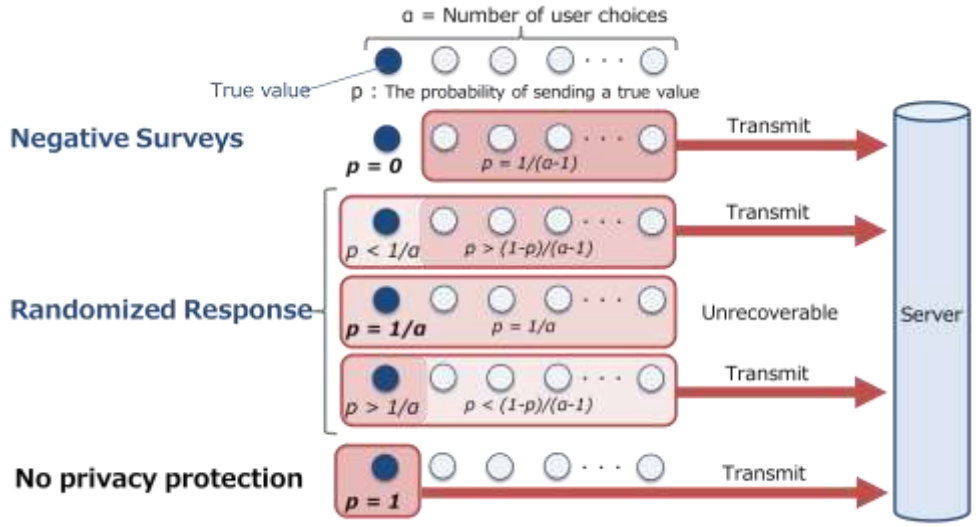

Figure 1: Outline of the randomized response and negative surveys methods

Existing privacy protection methods, namely, randomized response and negative surveys, are suitable for many sensing data with a small number of options. They are, however, characterized by the following problems.

- When the sensing data transmitted to the server is small, there is a low restoration accuracy in the restoration of the true value data distribution from the sensing data subjected to privacy protection processing.

- As the number of options observed by the sensing device increases, the restoration accuracy decreases.

In the use of random noise addition, a total number of data D is required to maintain the restoration accuracy when the number of options α increases. This is because of the transmission of values other than the true value as noise. The relationship between the total number of data D and the number of choices α can be expressed as in Equation (8) for a probability p of transmitting a true value, where $p\alpha$ (the product of the probability and the number of choices α) is constant. Here, $R(p\alpha)$ represents an equal distribution of data with respect to the number of choices α; the larger the value of $R(p\alpha)$, the more uniform is the data distribution, indicating higher restoration accuracy.

$$R(p\alpha) = \frac{\log D}{\log \alpha} \tag{8}$$

To obtain a good restoration accuracy, it is necessary for the data to be uniformly distributed with respect to the number of choices α. Even when the number of choices α increases, the achievement of the same restoration accuracy requires the total number of data D to be equal to $\alpha^{R(p\alpha)}$. When existing privacy protection methods are applied to a case with a small number of sensing data and many options, accurate restoration of the data distribution is impractical. This necessitates the development of a privacy protection method applicable to such cases.

3.2 Privacy Protection Method Using Multiple Transmissions

In the proposed privacy protection method, we utilize a technique for obtaining the data distribution of the entire sensing data in the server even when both the number of options observed by the sensing device and the sensing data are small.

In the proposed method, when one option is selected among multiple options and transmitted to the server, it represents k repeated transmissions. The distribution of the entire data collected by each sensing device can be obtained on the side of the server, even for a small number of data and many options. The procedure when using the proposed method for privacy protection in the transmission of participatory sensing data is as follows.

(1) The sensing device acquires sensing data and observes one option among the total number of options, α.

(2) The sensing device implements the privacy protection process on the sensing data and transmits the data to the server. The device selects true sensing data obtained with a probability p, or alternatively selects α-1 options different from the true data with a probability $\frac{1-p}{\alpha-1}$. The above selection is repeated k times and notifications of k options are sent to the server. Let A be the distribution of the entire sensing data collected by the sensing device. Let k be the number of repeated transmissions to the server. The privacy protection process is implemented by the sensing device and the distribution of the entire sensing data transmitted to the server is represented by Y, given by Equation (9). The privacy protection process can be represented by the matrix M, which is a square matrix of size $\alpha \times \alpha$ given by Equation (10).

(3) The server reconstructs the sensing data based on the data received from the sensing device. The server then obtains the distribution A of the entire sensing data collected by the device, according to Equation (11).

$$Y = kAM \tag{9}$$

$$M = \begin{pmatrix} p & \frac{1-p}{\alpha-1} & \cdots & \frac{1-p}{\alpha-1} \\ \frac{1-p}{\alpha-1} & p & \frac{1-p}{\alpha-1} & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1-p}{\alpha-1} & \cdots & \cdots & p \end{pmatrix} \tag{10}$$

$$A = YM^{-1} \tag{11}$$

We investigated the difference between the randomized response method and the presently proposed privacy protection method. Figure 2 is a schematic illustration of the proposed method as an extension of the randomized response method. In the latter, when the sensing device acquires the sensing data, it transmits the true value of the data to the server with a probability p. The device also transmits α-1 options different from the true value with a probability $\frac{1-p}{\alpha-1}$. In the existing method, the sensing device transmits one option to the server for sensing data acquisition. In the proposed method, the transmission is repeated k times, where k is an integer ≥2.

In the proposed method, because the data is repeatedly transmitted to the server, the total number of data transmitted to the server is $D' = Dk$. The relationship between the total number of data D and the number of choices α is expressed by Equation (12), where p is the probability of transmitting a true value, and $p\alpha$ (the product of the probability and the number of choices α) is constant.

$$R(p\alpha) = \frac{\log D\prime}{\log \alpha} \qquad D' = Dk \tag{12}$$

By repeating the data transmission k times and increasing the total number of data D, the proposed method can achieve the same effect as when a large amount of sensing data is transmitted.

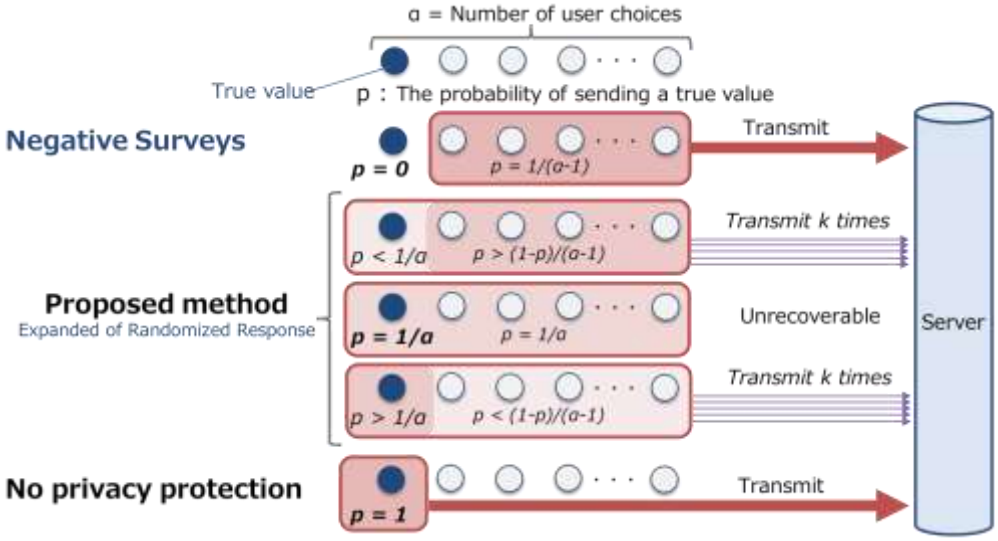

Figure 2: Outline of the proposed method

4 Evaluation Experiments

4.1 Evaluation Method

Evaluation experiments were performed to compare the randomized response method and the proposed method with respect to their accuracy for restoring the true data value. Jensen-Shannon divergence [22] [23] is used as an evaluation scale for restoration accuracy, being a measure of the difference between two probability distributions. It is an extension of Kullback-Leibler divergence [21].

The Kullback-Leibler divergence between two equal probability distributions is zero, and it increases with increasing difference between the probability distributions. For two probability distributions P and Q, the value $D_{KL}(P/\!/Q)$ of the Kullback-Leibler divergence of Q as seen from P can be obtained by Equation (13). However, because it does not satisfy Equation (14), it is asymmetric.

$$D_{KL}(P||Q) = \sum_i P(i) \log_2 \frac{P(i)}{Q(i)} \quad (13)$$

$$D(P||Q) = D(Q||P) \quad (14)$$

Jensen-Shannon divergence represents an extension of the Kullback-Leibler divergence because it enables the determination of the difference between two probability distributions that are considered symmetric under Kullback-Leibler divergence. For two probability distributions P and Q, the Jensen-Shannon divergence value $D_{JS}(P/\!/Q)$ can be determined using Equations (15) and (16). The Jensen-Shannon divergence is symmetric because it satisfies Equation (14). When $D_{JS}(P||Q)$ is small, it indicates that the accuracy of the restoration to the true data value is high.

$$D_{JS}(P||Q) = \frac{1}{2}(D_{KL}(P||M) + D_{KL}(Q||M)) \quad (15)$$

$$M(x) = \frac{1}{2}(P(x) + Q(x)) \quad (16)$$

In using the Jensen-Shannon divergence to evaluate the accuracy of restoration to the true data value in this study, we performed one hundred experiments per data set for a given probability p and number of options α, and determined the average of the $D_{JS}(P/\!/Q)$ values. The number of data was linearly increase. For example, when the number of options α was 10, the numbers of data for the respective options were {1, 2, 3... 10} and the total number of data D was 55. Similarly, when the number of options α was 20, the numbers of data for the respective options were {2, 4, 6... 40} and the total number of data D was 210. Furthermore, to compare the results for the different values of α and D, the division data was normalized by D and the restoration accuracy was evaluated.

4.2 Simulation of Randomized Response Method

First, for the existing methods (randomized response and negative surveys), we performed simulation experiments to examine how the accuracy of the restoration of the true data value varies with increasing number of options α and total number of data D. In the simulations, the probability p of transmitting the true value was varies as 0, $\frac{1}{5\alpha}$, $\frac{1}{4\alpha}$, $\frac{1}{3\alpha}$, $\frac{1}{2\alpha}$, $\frac{2}{\alpha}$, $\frac{3}{\alpha}$, $\frac{4}{\alpha}$, and $\frac{5}{\alpha}$, relative to the number of options α. The probability p of *0* was only applied to negative surveys.

Figure 3 shows the simulation results for the existing methods with respect to D and α. From the

upper left corner, the values of α are respectively 5, 10, and 20, while they are 100 and 200 from the bottom left. The y-axis of each graph represents the Jensen-Shannon divergence, with a small value indicating high accuracy. The x-axis represents the probability p of transmitting a true value. Where the middle line is missing, the probability p becomes $\frac{1}{\alpha}$, indicating that recovery is impossible. In the graphs, when α=5, D is 15, 30, 75, and 150; when α = *10*, D is 55, 110, 275, and 550; when α = 20, D is 210, 420, 1,050, and 2,100; when α = 100, 5,050, 10,100, 25,250, and 50,500; and when α=200, D is 20,100, 40,200, 100,500, and 201,000.

For a given number of options α, both randomized response and negative surveys exhibit higher restoration accuracy with increasing D, and vice versa. In addition, the restoration accuracy deteriorates with increasing α. It can therefore be concluded that, for both randomized response and negative surveys, the restoration accuracy deteriorates with decreasing D and increasing α.

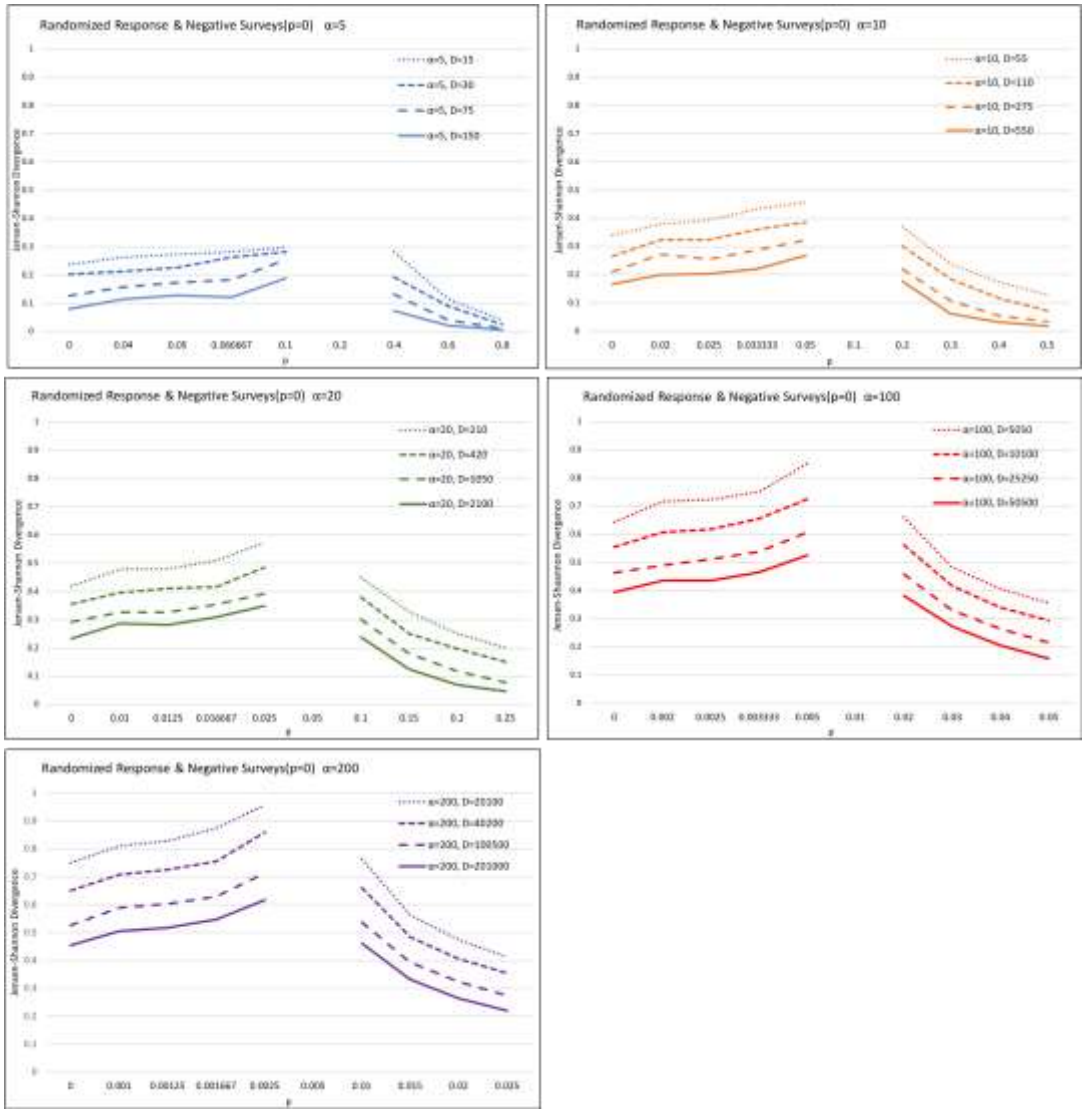

Figure 3: Restoration accuracies of randomized response and negative surveys for number of options α values of 5,10,20,100, and 200)

4.3 Simulation of Proposed Method

In the simulation of the proposed method, if the product of the total number of data D and the number of iterations k was the same as the total number of data to be transmitted in the simulation of the existing methods, the total number of data D' received by the server would be the same. This was an indication of a restoration accuracy equivalent to those of the existing methods. As in the simulation of the existing methods, the probability p of transmitting the true value in the simulation of the proposed method was varied as 0, $\frac{1}{5\alpha}$, $\frac{1}{4\alpha}$, $\frac{1}{3\alpha}$, $\frac{1}{2\alpha}$, $\frac{2}{\alpha}$, $\frac{3}{\alpha}$, $\frac{4}{\alpha}$, and $\frac{5}{\alpha}$, relative to the number of options α. The probability $p = 0$ was only applied to negative surveys. The number of iterations k in the implementation of the proposed method was also varied as 5, 10, 20, 100, 200, and 500.

Figure 4 shows the simulation results for both the proposed method and the existing methods (randomized response and negative surveys), namely, the restoration accuracies with respect to the number of options α. From the upper left corner, the α values are 5, 10, and 20, while they are

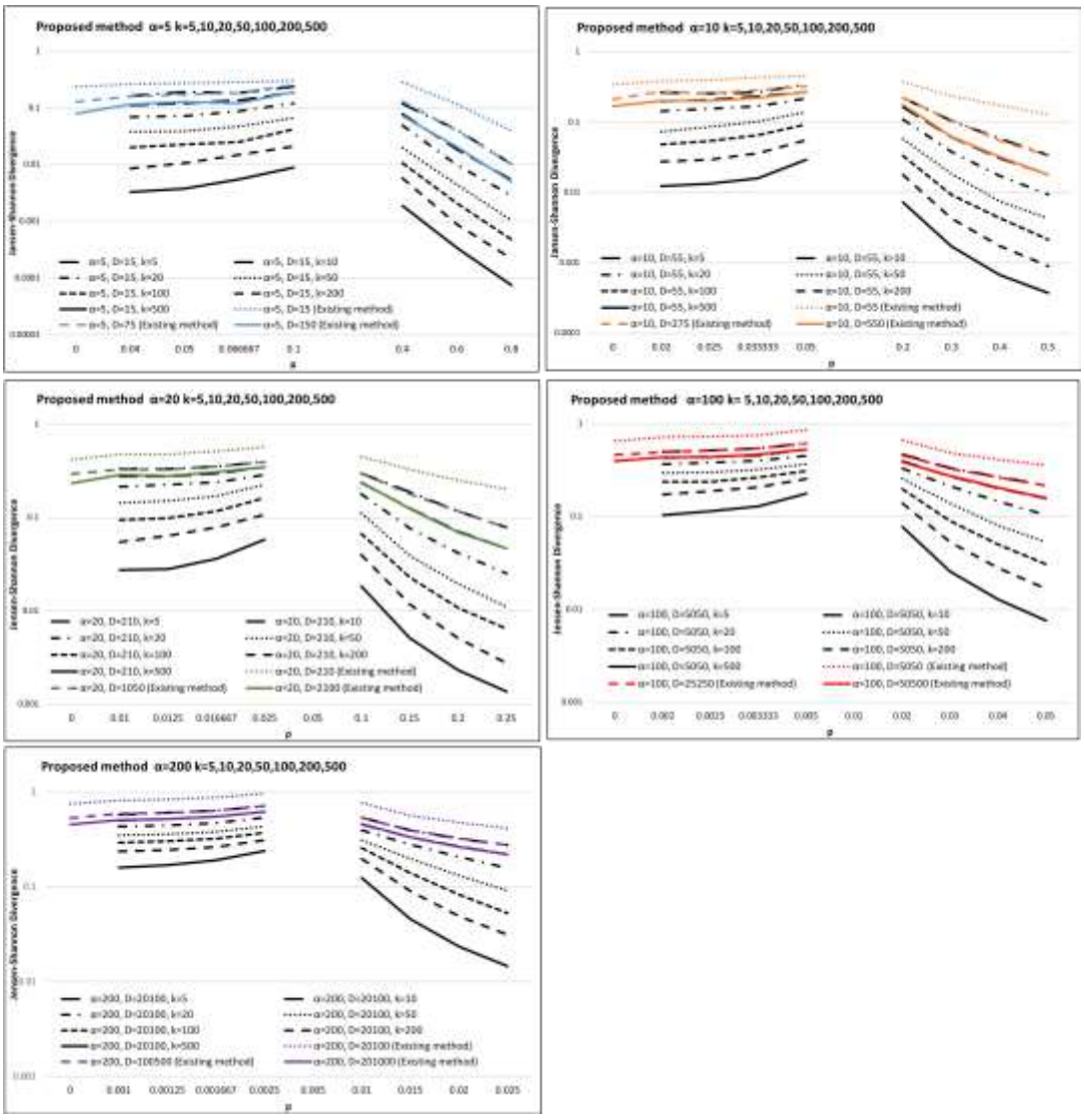

Figure 4: Comparison of the proposed method with the existing methods for α values of 5, 10, 20, 100, and 200, and k values of 5, 10, 20, 50, 100, 200, and 500

100 and 200 from the bottom left corner. In each graph, the number of iterations k is varied as 5, 10, 20, 100, 200, and 500, as indicated by the black lines. The colored lines correspond to the existing methods. The y-axis represents the Jensen-Shannon divergence, with a small value indicating a high restoration accuracy. To make each data more distinguishable, logarithmic display is employed. The x-axis represents the probability p of transmitting a true value. Where the middle line is missing, the probability p becomes $\frac{1}{\alpha}$, indicating that recovery is impossible. The restoration accuracies of the different methods were compared by superimposing the results of the proposed method on those of the existing methods in the previous section.

In Figure 4, for α=5 and D=75 in the existing methods, the result is the same as for D=15 and k=5 in the proposed method; and for D=150 in the existing method, the result is the same as for D=15 and k=10 in the proposed method. In addition, for α=10 and D=275 in the existing method, the result is the same as for D = 55 and k = 5 in the proposed method; and for D=550 in the existing method, the result is the same as for D=55 and k=10 in the proposed method. Furthermore, for α=20 and D=1,050 in the existing method, the result is the same as for D=210 and k=5 in the proposed method; and for D=2,100 in the existing method, the result is the same as for D=210 and k=10 in the proposed method. For α=100 and D=25,250 in the existing method, the result is the same as for D=5,050 and k=5 in the proposed method; and for D=50,500 in the existing method, the result is the same as for D=5,050 and k=10 in the proposed method. For α=200 and D=10,500 in the existing method, the result is the same as for D=20,100 and k=5 in the proposed method; and for D = 201,000 in the existing method, the result is the same as for D=20,100 and k=10 in the proposed method. These combinations reveal equal restoration accuracies of all the methods. It can also be seen that the restoration accuracy increases with increasing k for all α values.

Figure 5 compares the Jensen-Shannon divergence and $R(p\alpha)$ results for the simulation data in this section. The x-axis of the graph represents the Jensen-Shannon divergence, and the y-axis the $R(p\alpha)$ value. The colors of the lines correspond to different values of the probability p of transmitting a true value, namely, $\frac{1}{5\alpha}$, $\frac{1}{4\alpha}$, $\frac{1}{3\alpha}$, $\frac{1}{2\alpha}$, $\frac{2}{\alpha}$, $\frac{3}{\alpha}$, $\frac{4}{\alpha}$, and $\frac{5}{\alpha}$.

It can be seen from Figure 5 that there is a correlation between the Jensen-Shannon divergence and the $R(p\alpha)$ value. When the probability p of transmitting a true value is equal to α, which is the product of the number of options α, the restoration accuracy increases with increasing total number of transmitted data D.

4.4 Consideration

The simulation results reveal that the restoration accuracy for both randomized response and negative surveys deteriorate with decreasing total number of transmitted data D and increasing number of options α. However, the simulation results for the proposed method show that the same restoration accuracy as the foregoing methods can be achieved when the product of the number of data in the existing method, the total number of data D in the proposed method, and the number of iterations k in the proposed method are the same. In addition, the restoration accuracy of the proposed method increases with increasing k and α.

The existing privacy protection methods (randomized response and negative surveys) are suitable when there is many sensing data and a small number of options, but very unfavorable when there is many options and a small number of sensing data. However, the proposed method, which is an extension of randomized response, enables the restoration of effectively distributed data in the

server, even when there are many options and little sensing data.

Although the forgoing discussion of the proposed method for privacy protection in participatory sensing focused on restoration accuracy, it is necessary to also consider the method from the specific viewpoint of privacy protection. Normally, the privacy protection accuracy and the restoration accuracy are in a trade-off relationship. The privacy protection precision is the probability that the sensor data transmitted by the participant would be correctly estimated when the data distribution restored by the administrator is leaked by an attacker. The privacy protection accuracy decreases with increasing restoration accuracy. The feature of the proposed method that differentiates it from the existing methods is that the data from a given participant is transmitted k times. Because an attacker may know the repetition number k, there is the possibility that the transmission of the data k times may itself decrease the privacy protection. Although it is possible to maintain the restoration accuracy of the proposed method, it is not always possible to maintain the privacy protection precision of the existing methods.

In addition, when participants transmit sensing data using the proposed method, the amount of communication increases with the repetition of the transmission k times, and this may constitute a burden on the participants. It is therefore necessary to adjust not only the repetition number k, but also the probability p of transmitting a true value.

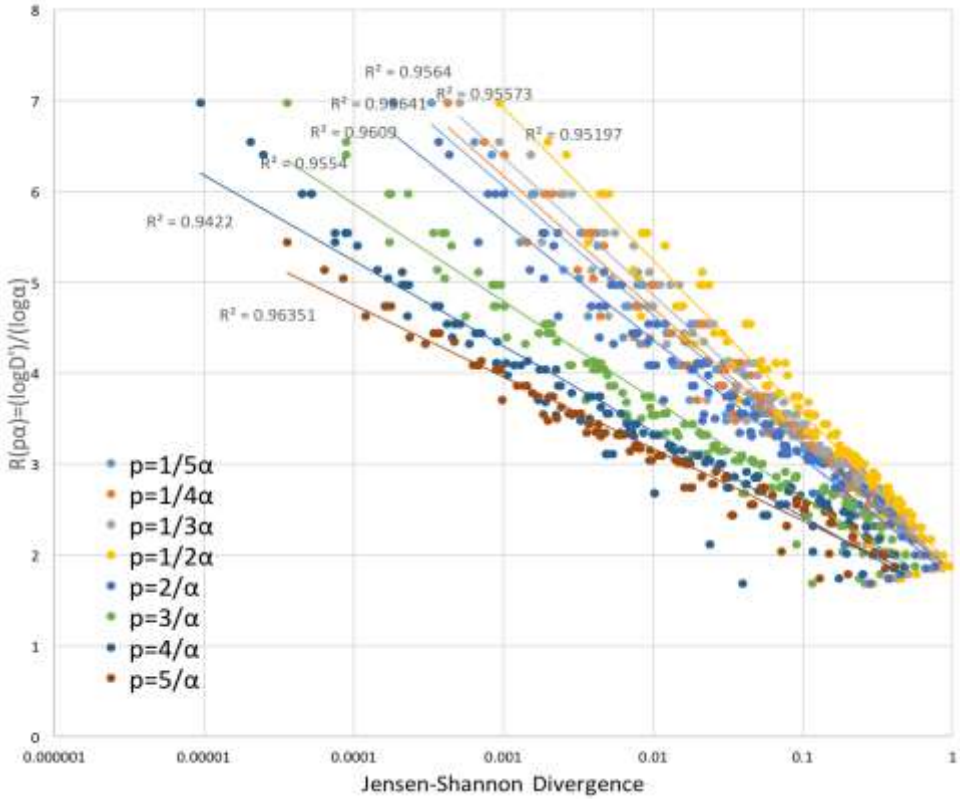

Figure 5: Comparison of the Jensen-Shannon divergences and R(p α) values for the different methods

5 Conclusion

In this study, we developed a privacy protection method for participatory sensing that addresses some of the problems of existing methods. Existing privacy protection methods, namely, randomized response and negative surveys, are only suitable for cases with many sensing data and a small number of options. However, the proposed method enables effective data distribution in the server even when the number of options observed by the sensor is large and the sensing data is small. Experimental simulations were used to evaluate and compare the restoration accuracies of the proposed and existing methods, with the observations verifying the effectiveness of the proposed method. Future studies are to take measures to the amount of traffic increases when participants transmit sensing data by the proposed method.

Acknowledgement

This work was supported by JSPS KAKENHI (Grant No. JP19K12130) and JST CREST (Grant No. JPMJCR15E1).

References

[1] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava: Participatory Sensing, World Sensor Web Workshop, ACM Sen-Sys 2006, Colorado, Oct 2006.

[2] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, and T. Choudhury: StressSense: Detecting Stress in Unconstrained Acoustic Environments using Smartphones, ACM International Conference on Ubiquitous Computing (UbiComp 2012), pp. 351–360, 2012.

[3] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong: MoodScope: Building a Mood Sensor From Smartphone Usage Patterns, The 11th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys 2013), pp. 389–402, 2013.

[4] M. Azizyan, I. Constandache, and R. Roy Choudhury: SurroundSense: Mobile Phone Localization via Ambience Fingerprinting, The 15th Annual International Conference on Mobile Computing and Networking (MobiCom 2009), pp. 261–272, 2009.

[5] N. Maisonneuve, et al.: NoiseTube: Measuring and mapping noise pollution with mobile phones, In: Information Technologies in Environmental Engineering, Springer Berlin Heidelberg, pp. 215–228, 2009.

[6] S. King and P. Brown: Fix My Street or Else: Using the Internet to Voice Local Public Service Concerns, The International Conference on Theory and Practice of Electronic Governance, Macau, pp. 72–80, 2007.

[7] K. L. Huang, S. S. Kanhere, and W. Hu: Preserving Privacy in Participatory Sensing System, Computer Communication, vol. 33, no. 11, pp. 1266–1280, 2010.

[8] G. Drosatos, P. S. Efraimidis, I. N. Athanasiadis, E. D'Hondt, and M. Stevens: A Privacy Preserving Cloud Computing System for Creating Participatory Noise Maps, Proceedings of

36th IEEE International Conference on Computer Software and Applications (COMPSAC2012), pp. 581–586, 2012.

[9] L. Sweeney: k-anonymity: A Model for Protecting Privacy, International Journal of Uncertainty, Fuzziness, and Knowledge-based Systems, vol. 10, no. 5, pp. 557–570, 2002.

[10] P. Samarati: Protecting Respondents Identities in Microdata Release, IEEE Transactions on Knowledge and Data Engineering, vol. 13, no. 6, pp. 1010–1027, 2001.

[11] C. Dwork, F. McSherry, K. Nissim, and A. Smith: Calibrating Noise to Sensitivity in Private Data Analysis, In: Theory of Cryptography, Springer Berlin Heidelberg, pp. 265–284, 2006.

[12] C. Dwork: Differential Privacy, In: Automata, Languages and Programming, Springer Berlin Heidelberg, pp. 1–12, 2006.

[13] S. L. Warner: Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias, Journal of the American Statistical Association, vol. 60, no. 309, pp. 63–69, 1965.

[14] R. Agrawal, R. Srikant, and D. Thomas: Privacy Preserving OLAP, Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data, pp. 251–262, 2005.

[15] S. Aoki and K. Sezaki: Negative Surveys with Randomized Response Techniques for Privacy-aware Participatory Sensing, IEICE Transactions on Communications, vol. E97-B, no. 4, 2014.

[16] D. Quercia, I. Leontiadis, L. McNamara, C. Mascolo, and J. Crowcroft: SpotME If You Can: Randomized Responses for Location Obfuscation on Mobile Phones, Proceedings of 31st International Conference on Distributed Computing Systems (ICDCS), pp. 363–372, 2011.

[17] F. Esponda: Negative Surveys, ArXiv Mathematics e-prints, 2006.

[18] F. Esponda and V. M. Guerrero: Surveys with Negative Questions for Sensitive Items, Statistics & Probability Letters, vol. 79, no. 24, pp. 2456–2461, 2009.

[19] M. M. Groat, B. Edwards, J. Horey, W. He, and S. Forrest: Enhancing Privacy in Participatory Sensing with Multidimensional Data, Proceedings of IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 144–152, 2012.

[20] S. Aoki and K. Sezaki: Privacy-Preserving Community Sensing for Medical Research with Duplicated Perturbation, Proceedings of the IEEE International Conference on Communications (IEEE ICC), pp. 4252–4257, 2014.

[21] S. Kullback and R. A. Leibler: On Information and Sufficiency, Annals of Mathematical Statistics, vol. 22, p. 79–86, 1951.

[22] J. C. Angulo, J. Antolin, S. Lopez-Rosa, and R. O. Esquivel: Jensen-Shannon Divergence in Conjugate Spaces: The Entropy Excess of Atomic Systems and Sets with Respect to their Constituents, Elsevier Physica A, vol. 389, pp. 899–907, 2010.

[23] B. Fuglede and F. Topsoe: Jensen-Shannon Divergence and Hilbert Space Embedding, IEEE International Symposium on Information Theory, 2004.