Method of differential privacy protection for web-based communities based on adding noise to the centroid of positions

Chun Guan

School of Information Engineering,
Nanchang University,
Nanchang 330031, Jiangxi, China
Email: guanchun@yeah.net

Jun Hu* and Yu Zhou

School of Software,
Nanchang University,
Nanchang 330047, Jiangxi, China
Email: hujun@ncu.edu.cn
Email: 1761517051@qq.com
*Corresponding author

Alexander Shatalov

School of Information Engineering,
Nanchang University,
Nanchang 330031, Jiangxi, China
Email: 1850109810@qq.com

Abstract: The research of location privacy protection has a very important significance for the emergence and sustenance of web-based communities in the big data era. The method of differential privacy for position is a privacy model strictly for the measurement of privacy. However, under the privacy protection for multi-positions, it will bring much bigger errors by adopting the method of adding noise to protect the privacy of multi-positions. In this paper, the author puts forward a new method of differential privacy for multi-positions based on mechanism of adding noise to the centroid of positions and compares with the independent mechanism of adding noise. The experiment shows that the centroid mechanism of adding noise is better than the mechanism of independently adding noise.

Keywords: privacy; privacy protection; web-based communities; centroid; location-based services; multi-positions.

Reference to this paper should be made as follows: Guan, C., Hu, J., Zhou, Y. and Shatalov, A. (2021) 'Method of differential privacy protection for web-based communities based on adding noise to the centroid of positions', *Int. J. Web-Based Communities*, Vol. 17, No. 1, pp.53–64.

Biographical notes: Chun Guan received her BE degree from the Department of Electronic Engineering from Jiangxi University in China in 1993. She obtained her ME degree in Applied Mathematics from Nanchang University in China in 1998, and obtained her PhD degree in Management Science and Engineering from Nanchang University in China in 2005. She is currently an Associate Professor at the School of Information Engineering, Nanchang University. Her research interests include artificial intelligence, privacy protection and web-based communities.

Jun Hu obtained his BE degree from the Department of Computer Science from Jiangxi University in China in 1992, his ME degree in Applied Mathematics from Nanchang University in China in 1997, and his PhD degree in Computer Application from BIT (Beijing Institute of Technology) in China in 2003. He is currently a Professor at School of Software, Nanchang University. He visited the School of Computer Science, Nanyang Technological University, Singapore from October to December 2012, and visited the Clemson University School of Computer Science for one year from 2017 to 2018. He is mainly engaged in research works on artificial intelligence, computational intelligence, trust computing, deep learning and other related fields.

Yu Zhou received his BE degree in Software Engineering from Nanchang University, China, in 2016. He is currently pursuing his Master's degree at Nanchang University, China with a research subject regarding privacy protection in web-based communities.

Alexander Shatalov is an undergraduate student at School of Information Engineering, Nanchang University in China with a research subject regarding privacy protection in web-based communities.

1 Introduction

With the mature development of positioning technology, such as mobile communication technology and sensing device, the geographic location information between human beings and objects is effectively associated. Placing a sensor chip in the mobile object can locate the position information of the mobile object. For example, the positioning devices, such as GPS and Wi-Fi, are placed in the mobile devices widely. The smart phones and vehicle navigation can directly achieve the exact location information of the mobile object, for example, most of current social websites application can collect and publish the location information of the users (Jabeur et al., 2013). Moreover, the rise of wearable devices bring new ways for location information, it can locate the location of the users by processing the data of the collective acceleration and optical image (Anguelov et al., 2010; Ugolotti et al., 2013).

Members of web communities often use the GPS information in daily life. The personal location service brings community members a lot of convenience. By collecting the places which the user often goes, it can be presumed the user's living habits and trip rules, then individual service can be recommended to both the user and his/her social network members. For example, in the document (Abowd et al., 1997; Cui et al., 2019), by analysing a large number of historical trajectory data, reasonable routes can be recommended to everyone; and according to the current traffic congestion, more fast and convenient route can be recommended to the user and so on. The wireless data science and technology (Jana) collected the mobile data from more than 100 countries which supplied by over 200 wireless operators (Piao et al., 2019). Those data that cover about the information of 3.5 billion community members among Latin America, Europe and Africa, can try to explain the prosperity history of the urban and the spread of a disease.

While the location-based service (LBSs) and the publish of location data bring much convenience to our life, the exposure of real users' location information causes huge security risks to users, especially in the prevailing big data era, the attacker can easily presume the user's behaviour habits and living status by analysing big data technology of the user's location information, which leaks privacy seriously. Therefore, location privacy protection is an urgent problem to resolve, studying the location privacy protection is a very important field for the sake of community-integrity and privacy protection, the location privacy protection can help both the individual and the social circle to enjoy the location service and doesn't need to worry about the risk of leaking personal privacy. The research of location privacy protection has a very important significance for the survival of web-connected communities.

2 Related work

The traditional location privacy protection methods include two categories (Friedman et al., 2016). One is to conceal one personal or unit's identification information, that is, to conceal its real identification information, but to supply the accurate location information to the server so that to get the high-quality service. This method has anonymity technology, disguising name technology, mix zone and so on. The other method is to expose the user's identification information, but to supply the server with the user's location information after changing with a way. This method uses wrong or false location information, sign object or regional location information (Yu and Cao, 2018; Tian et al., 2018; Soria-Comas et al., 2017; Yang et al., 2018; Wang and Sinnott, 2017; Jan, 2019; Kutlu et al., 2018; Lyu and Lim, 2018) to conceal its real location information. Gong et al. (2019) initially adopt obfuscation mechanism to do the obfuscation on the location, and put forward a query processing algorithm for obfuscation location information, and ensure a higher service quality. The main thought of this technology is to reach the goal of protecting location privacy by reducing the accuracy of the user's location.

Therefore, the traditional location privacy protection assumes that the attacker has no background knowledge, and couldn't do the quantification on the privacy protection level. Differential privacy protection (Kim et al., 2018; Wang et al., 2017; Shin et al., 2018; Ni et al., 2018; Md and Vijayakumar, 2019; Taamallah et al., 2019) does quantification on the privacy protection level, which is a strict privacy protection model, and can protect the attack of the biggest background knowledge. This model was initially applied in the statistics database. They applied this thought and put forward the concept of geo-indistinguishability, and supplied a quantisable privacy protection model for location privacy protection. This model is expended from the definition of differential privacy protection and expends the one-dimensional Laplace mechanism implementation method to two-dimensional position coordinates, because the location differential privacy reaches the goal of protecting the privacy by adding little random noise to every position. When the number of the position is small, to add little random noise to every position can get good effect. But when the number of positions is large, to add much noise can also reduce the degree of privacy protection. In the position-based service, adding noise to the position for many times will reduce the degree of privacy protection, and cause too much error, so Ul Hassan et al. (2019) put forward a predictive mechanism to reduce the volume of addition of the noise and improve the privacy protection efficiency. The thought of the predictive mechanism is that if the distance between the current position and one position of the historical positions is within a certain threshold, the result of historical position after adding noise will be adopted to be the search result, to reduce the volume of addition of the noise. Aiming at the privacy protection of multi-positions of the concept of geo-indistinguishability, Wang et al. (2018) determine the optimal mechanism with adding the least amount of noise in the finite positions. The concept of geo-indistinguishability is applied to protect users' location privacy in social applications (Hua et al., 2018; Xiong et al., 2017; Ma et al., 2018). It is also the research focus of this paper to utilise this model to protect the location privacy and try its best to reduce the error.

3 Method based on mechanism of adding noise to the centroid of positions

3.1 Differential privacy for multi-positions

Differential privacy for positions applies the method of differential privacy protection into the protection for positions privacy, the principle thought is to add a concern random noise on the original positions to create new positions as output, so that it can ensure the function of the real positions.

Andres et al. (2013) applies differential privacy to put forward the concept of geo-indistinguishability privacy protection, defines every single position's privacy protection. However, many applications often involve many positions; it is of great value to study the protection for multi-positions. In this paper, there are different definitions for privacy protection under multi-positions and from differential privacy. In the differential privacy, it requires two groups of datasets are adjacent datasets, that is, among two datasets, at most, there is a different element. In this paper, the authors still adopt the definition of differential privacy to expend to the definition of positions privacy protection under multi-positions data.

Definition 1: Given two positions datasets $x = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, $x' = \{x_1, x_2, \ldots, x_i',$ $\ldots, x_n\}$, inquiry function f, for any x, x', there is $d(x_i, x_i') \leq r$, Z is the result set of the inquiry function. If one function or mechanism meets the formula requirement for all $S \subseteq Z$ as follows:

$$\frac{P(f(x) = S|x)}{P\left(f(x') = S|x'\right)} \leq e^{\varepsilon r}, \qquad \forall r > 0, \forall x, x' : d\left(x, x'\right) \leq r. \tag{1}$$

Then, this mechanism meets the requirement of ε-geo-indistinguishability. Where x, x' are the adjacent location datasets, at most, there is one different position? The inquiry function f is the action function for position data, for example: $f(x) = x$, $f(x) = centroid(x)$ and so on.

As showed in Definition 1, it is very similar to the definition of differential privacy protection in the statistical database, the difference is the datasets here are two groups of position data, and it abstracts the inquiry function f of position data. By analysing the definition 1, the definition 1 also meets the requirement of the definition of every single position's differential privacy protection which put forward by Andres et al. (2013). However, among these two datasets, there is separately one position and two positions are different, and when its inquiry function returns the position itself, it is equivalent to be $f(x) = x$, therefore, Definition 1 includes the differential privacy protection model of single position and multi-positions, which is more versatile. In this paper, it adopts this definition to do the further study.

3.2 The mind of adding noise to the centroid

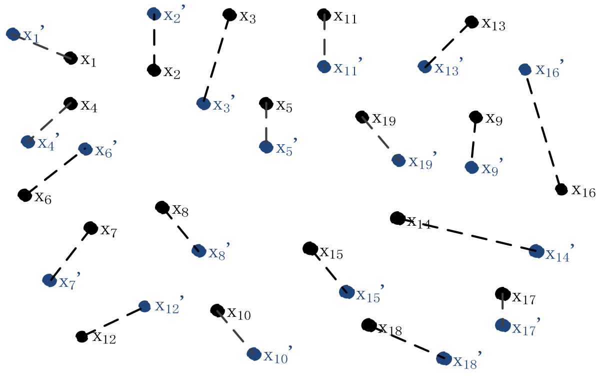
The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph). In real life, the users often pay attention to several positions, the most visual thought is that the position can be protected by adding Laplace noise to every position; it is called 'independent adding noise mechanism'. This method has one problem, that is, when the number of positions increases, the consumption of privacy protection budget also increases; it also means the protection degree reduces. For example, the user makes multiple location service requests at the same position, the point generated by adding noise with the independent adding noise mechanism will surround by the real positions; thus, the attack can guess the real-time position of the user according to the position after adding noise.

In addition to the independent adding noise mechanism, in this paper, the author puts forward algorithm of differential privacy for multi-positions based on the mechanism of adding noise to the centroid (hereinafter referred to as the mechanism of adding noise to the centroid)., aiming at similar multi-positions, the author adopts the method of adding noise to the centroid of multi- positions to resolve the problem of too much error caused by adding noise many times with the independent adding noise mechanism. Certainly, adopting the method of the centroid reduces the amount of the addition of noise compare to the independent adding noise mechanism; however, it increases the reconstruction error at the same time, because there is a little distance between every position and the centroid. Nevertheless, under the condition of close distance between the positions, adopting adding noise mechanism to the centroid is better than the independent adding noise mechanism. Take an extreme example, when many positions are the same, adopting the method of the centroid to add noise has no reconstruction error, and only needs to add the noise for one time, but adopting the independent adding noise mechanism requires adding the noise for many times. Therefore, when the distance between positions is very close, adopting adding noise mechanism to the centroid is better than the independent adding noise mechanism.

The mechanism of independently adding noise, which is to independently add Laplace noise to every position for location datasets, meets the requirement of the differential privacy protection for positions. The privacy protection degree is the accumulation of the privacy protection budget for all positions in the set. When the number of the positions is small, and the distance between positions is farther, the effect is ideal.

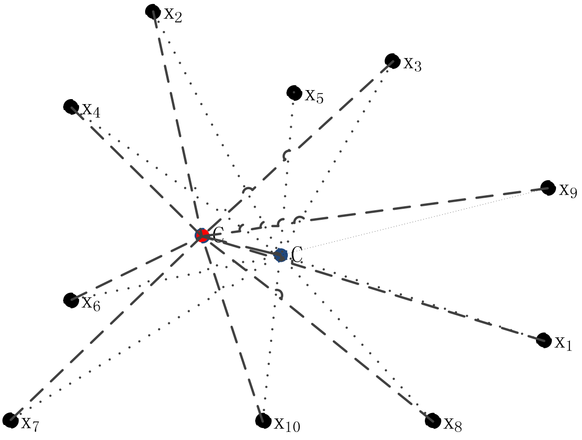
When the number of positions is less, there is a better effect to apply the independent adding noise mechanism. When the number of positions is more, the mechanism effect is lower, because the privacy budget is getting linear increase with the number of positions, and the amount of the addition of noise is getting linear increase too, which reduces the accuracy of positions. The independent adding noise mechanism is as showed in Figure 1, every position independently add random Laplace noise.

Figure 1 The mechanism of independently adding noise (see online version for colours)



On the existing problem of the mechanism of independently adding noise, in this paper, it adopts the method of adding noise to the centroid of multi-positions instead of adding noise to every position to reduce the amount of the addition of noise. By this method, it gets the centroid for multi-positions, then adopting Laplace mechanism to add noise to the centroid, which equals to that every position after adding noise of the position is the position after adding noise to the centroid.

Figure 2 The mechanism of adding noise to the centroid (see online version for colours)



The advantage of this method, which compares to the independently adding noise mechanism, is to reduce the amount of the addition of noise but it will increase the reconstruction error, because there is error between the centroid position and the real position, it requires to find out the reasonable centroid to reduce the comprehensive error, this mechanism applies to the cite when the number of positions are bigger and the distance between positions is closer. The centroid adding noise mechanism is as showed in Figure 2, it gets C for all positions, and then gets C' by adding random Laplace noise to the centroid, that is , all x_i', that got by random algorithm for all positions, are C'. This mechanism also meets the requirement of differential privacy protection for positions.

3.3 The method of adding noise to the centroid

Method of differential privacy protection for multi-positions based on mechanism of adding noise to the centroid of positions is shown as follows. Under many positions, using adding noise method of expending every single position completes the privacy protection of multi-positions.

Inference 1: Supposing one position set is $x_0 = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, $d(x_i, x_i') \leq r$, inquiry function is f, if the position X after adding noise meets the requirement of the following probability distribution $D(x_0)(x) = \frac{\varepsilon}{2\pi}\exp(-\varepsilon d(f(x_0), x))$, then it satisfies $\varepsilon \cdot d(f(x_0), f(x'))/r$ – geo-indistinguishability.

Proof: According to the definition of differential privacy and the triangle inequality, it gets:

$$D_\varepsilon(x_0)(x) \leq \exp\left(\varepsilon d\left(f(x_0), x_0'\right)\right) D_\varepsilon(x_0')(s) \tag{2}$$

To do integral on both sides of inequality on the query result S, it can get:

$$\int_S D_\varepsilon(x_0)(x)dx \leq \int_S \exp\left(\varepsilon d\left(f(x_0), x_0'\right)\right) D_\varepsilon(x_0')(x)ds \tag{3}$$

According to the definition of the probability density function, it can get to know:

$$P\left(f(x_0) = S|x_0\right) = \int_S D_\varepsilon(x_0)(x)ds \tag{4}$$

Thus,

$$P\left(f(x_0) = S|x_0\right) \leq \exp\left(\varepsilon d\left(f(x_0), f(x_0')\right)\right) P\left(f(x_0') = S|x_0'\right) \tag{5}$$

$$P\left(f(x_0) = S|x_0\right) \leq \exp\left(\frac{\varepsilon d\left(f(x_0), f(x_0')\right)}{r} \cdot r\right) P\left(f(x_0') = S|x_0'\right) \tag{6}$$

That is, it satisfies $\varepsilon \cdot d(f(x_0), f(x'))/r$-geo-indistinguishability. Q.E.D.

From the above inference, under multi-positions, adopting different positions to inquire function, there will be different degree of privacy protection. According to Laplace adding noise mechanism, it can be deduced that adding multi-positions centroid needs adding the amount of the noise.

Inference 2: Supposing one position set is $x_0 = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, $d(x_i, x_i') \leq r$, inquiry function is $f(x) = centroid(x)$, if the position x after adding noise meets the requirement of the probability distribution $D(x_0)(x)=\frac{\varepsilon^2}{2\pi}\exp(-\varepsilon d(f(x_0),x))$, then it meets the requirement ε/n-geo-indistinguishability.

Proof: It is given that in $x = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, $x' = \{x_1, x_2, \ldots, x_i', \ldots, x_n\}$, only the position x_i and x_i' are different, $d(x_i, x_i') \le r$, every position uses $x_i = (s_i, t_i)$, $C(x)$, $C(x')$ show the centroid of two groups of positions. Then:

$$C(x)=\left(\frac{s_1+s_2+\ldots+s_n}{n},\frac{t_1+t_2+\ldots+t_n}{n}\right) \tag{7}$$

$$\begin{aligned} C(x') &= \left(\frac{s_1'+s_2'+\cdots+s_n'}{n},\frac{t_1'+t_2'+\cdots+t_n'}{n}\right) \\ &= \left(\frac{s_1+s_2+\cdots+s_n+r\cos\theta}{n},\frac{t_1+t_2+\cdots+t_n+r\sin\theta}{n}\right) \end{aligned} \tag{8}$$

$$C(x)-C(x')=\left(\left(\frac{r\cos\theta}{n}\right)^2+\left(\frac{r\sin\theta}{n}\right)^2\right)^{1/2}=\frac{r}{n} \tag{9}$$

That is: $d(f(x), f(x')) = r/n$.

Therefore, it meets the requirement ε/n-geo-indistinguishability. Q.E.D.

In the case of that multi-positions need protection, independently adopting Laplace mechanism to add noise will make the adding amount of noise too much, the more reasonable way is to use a typical position to replace the similar position. Multi-positions centroid is a much more reasonable choice. Supposing there is n positions' set $x = \{x_1, x_2, \ldots, x_i, \ldots x_n\}$, $d(x_i, x_i') \le r$ every position needs protection in the set, if adopting independent Laplace mechanism to add noise, by default using $f(x) = x$, assuming the privacy protection budget of the consumption of every position is ε, in which, n is the number of positions, therefore, the consumption of privacy protection budget is ε/n, it is known that adopting the way of multi-positions centroid can increase the degree of privacy protection degree. On the contrary, in the case of that the independent adding noise method and the centroid adding noise method have common privacy protection budget, to reduce the privacy protection degree under the centroid adding noise mechanism can reduce the addition of noise.

Known from Laplace adding noise mechanism, adding noise r is related to privacy protection parameter ε, assuming the privacy protection budget of each position is ε, the privacy protection degree got by adopting independent adding noise mechanism is n; the privacy protection degree got by adopting adding noise to the centroid mechanism is ε/n, therefore, in order to get ε-geo-indistinguishability under the mechanism of adding noise to the centroid that is, $n\varepsilon = \varepsilon'/n$, it gets $\varepsilon' = n^2\varepsilon$. Therefore, it can be deduced the method of adding noise to the centroid as follows:

Given a position set $x = \{x_1, x_2, \ldots, x_i, x_n\}$, $d(x_i, x_i') \le r$, inquiry function $f(x) = centroid(x)$, to add the noise of (r, θ) to the centroid, then it meets the requirement of ε-geo-indisthinguishability, in which, $r = C_{g,n}^{-1}(r)$, θ is the random angle for $[0, 2\pi)$.

$$C_{g,n}(r)=1-\left(1+n^2\varepsilon r\right)\exp\left(-n^2\varepsilon r\right) \tag{10}$$

3.4 Error analysis

It only exists agitation error under the independently adding noise mechanism, that is, by the random noise added under Laplace mechanism, but Laplace distribution is a continuous random distribution, that is, every noise added every time is different, perhaps, the random noise added for the first time is smaller, the random noise added for the second time is louder, to get the adding average error of the mechanism, it can get the expectation of the error after addition.

Supposing one position set is $x = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, $d(x_i, x_i') \le r$, adopting the formula (10) to independently add noise $error_i = C_{\varepsilon,n}^{-1}(r)_i$ for every position, since there is no reconstruction error, the total error is the sum of noise after adding of every position. $Error = \sum_{i=1}^{n} error_i$. Since the noise caused by the random algorithm is different, it can be estimated the theoretical error by the noise probability density function. Supposing the expectation after adding noise every time is $E(r)_i$, then:

$$\begin{aligned} E(r)_i &= \int_0^{+\infty} r \cdot p(r) dr \\ &= \int_0^{+\infty} r \cdot \varepsilon^2 r e^{-\varepsilon \cdot r} dr \\ &= \varepsilon \cdot \int_0^{+\infty} \varepsilon \cdot r^2 e^{-\varepsilon \cdot r} dr \\ &= \varepsilon \cdot \frac{2}{\varepsilon^2} \\ &= \frac{2}{\varepsilon} \end{aligned} \tag{11}$$

It can be known from the noise expectation that the adding noise is inversely proportional to the privacy protection budget, while the privacy protection degree is higher, the noise expectation which need adding is larger, vice versa.

Adopting the independently adding noise mechanism to add noise to every position of the position set, then the total expectation error is

$$Error = \sum_{i=1}^{n} error_i = n \cdot E_i(r) = \frac{2n}{\varepsilon}.$$

There are two parts of errors by adopting the mechanism of adding noise to the centroid, one part is the noise error added to the centroid, the other part is the error between the real position and the centroid.

The advantage of this method, which compares to the independently adding noise mechanism, is to reduce the amount of the addition of noise but it will increase the reconstruction error, because there is error between the centroid position and the real position, it requires to find out the reasonable centroid to reduce the comprehensive error, this mechanism applies to the cite when the number of positions are bigger and the distance between positions is closer. The centroid adding noise mechanism also meets the requirement of differential privacy protection for positions.

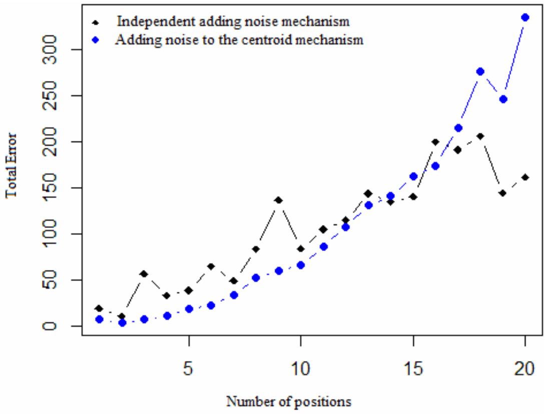
4 Experiments

Besides analysing the error of differential privacy for multi-positions based on the centroid from the position evacuation degree, the number of positions in the position dataset is also a very important factor to influence error. In this paper, with the experiment, the author analyses the error size of the differential privacy for multi-positions based on the centroid and the independently adding noise mechanism. In the experiment, setting $\varepsilon = 0.2$, the position evacuation degree is 2, that is, to generate randomly n positions under the area of $2n \times 2n$, the experiment result is as shown in Figure 3. It can be known from analysing the result of the experiment, when the position number is small, the total error caused by differential privacy for multi-positions based on the mechanism of adding noise to the centroid is smaller than the error caused by the independent adding noise mechanism. In this experiment, if the number of positions is less than 15, the algorithm in this paper is much better than the independent adding noise mechanism; but when the number of positions is bigger than 15, the total error of centroid method is bigger than the independent adding noise mechanism, because in the case of that n positions in the area of $2n \times 2n$ is uniformly distributed, the expectation of reconstruction error is

$$E = \sum_{i=1}^{n} \sqrt{\left(\frac{2i}{2}\right)^2 + \left(\frac{2i}{2}\right)^2} = \sum_{i=1}^{n} \sqrt{2i} = \sqrt{2} \sum_{i=1}^{n} i,$$

when $n > 15$, the reconstruction error is bigger than $\frac{2b}{\varepsilon}$, which is beyond the theory demands, therefore, the error of independent adding noise mechanism is smaller instead.

Figure 3 Total error under different number of positions (see online version for colours)



With the experiment, authors analyse the error introduced by differential privacy protection method for multi-positions based on the mechanism of adding noise to the centroid, and compares with the independently adding noise mechanism based on the same privacy protection conditions. The experiment result shows that under the same privacy protection level, the differential privacy protection algorithm for multi-positions based on the mechanism of adding noise to the centroid is better than the independent adding noise mechanism under certain conditions, the algorithm applies in the case when the distance between positions is close.

5 Conclusions

Under the differential privacy protection for multi-positions, when the number of positions is less, adopting the independently adding noise mechanism can reach the goal of protecting the differential privacy, and it will not cause bigger error. But when the number of positions data is more, it will cause the adding noise too loud, which can reduce the effect of privacy protection. Aiming at this problem, in the paper, the author puts forward adopting adding noise to the centroid to all positions to reduce the addition of the error. The method applies to the situation when the distance between positions is close. In this paper, the author compares the difference and its advantage of these two methods and also compares the errors. The experiment shows that under the same privacy protection level, the differential privacy protection algorithm for multi-positions based on the mechanism of adding noise to the centroid is better than the independent adding noise mechanism under certain conditions.

Acknowledgements

This work was supported by the Natural Science Foundation of China (Grant No. 61662046) and the Special Foundation for Visiting Scholars of Young and Middle-Aged Teachers' Development Plan in Jiangxi Ordinary Universities (Grant No. 2018109).

References

Abowd, G.D., Atkeson, C.G. et al. (1997) 'Cyberguide: a mobile context-aware tour guide', *Wireless Networks*, Vol. 3, No. 5, pp.421–433.

Andres, M.E., Bordenabe, N.E. et al. (2013) 'Geo-indistinguishability: differential privacy for location-based systems', Paper Presented at *2013 ACM SIGSAC Conference on Computer and Communications Security*, Association for Computing Machinery. 4–8 November, Berlin, Germany.

Anguelov, D., Dulong, C. et al. (2010) 'Google Street View: capturing the world at street level', *Computer*, Vol. 43, No. 6, pp.32–38.

Cui, L., Qu, Y., Reza, N.M. et al. (2019) 'Improving data utility through game theory in personalized differential privacy', *Journal of Computer Science and Technology*, Vol. 34, No. 2, pp.272–286.

Friedman, A., Berkovsky, S. and Kaafar, M.A. (2016) 'A differential privacy framework for matrix factorization recommender systems', *User Modeling and User-Adapted Interaction*, Vol. 26, No. 5, pp.425–458.

Gong, M., Pan, K. and Xie, Y. (2019) 'Differential privacy preservation in regression analysis based on relevance', *Knowledge-Based Systems*, 1 June, Vol. 173, pp.140–149.

Hua, J., Tong, W., Xu, F. and Zhong, S. (2018) 'A geo-indistinguishable location perturbation mechanism for location-based services supporting frequent queries', *IEEE Transactions on Information Forensics and Security*, Vol. 13, No. 5, pp.1155–1168.

Jabeur, N., Zeadally, S. and Sayed, B. (2013) 'Mobile social networking applications', *Communications of the ACM*, Vol. 56, No. 3, pp.71–79.

Jan, S.K. (2019) 'Investigating virtual communities of practice with social network analysis: guidelines from a systematic review of research', *International Journal of Web Based Communities*, Vol. 15, No. 1, pp.25–43.

Kim, J.W., Kim, D-H. and Jang, B. (2018) 'Application of local differential privacy to collection of indoor positioning data', *IEEE Access*, 9 January, Vol. 6, pp.4276–4286.

Kutlu, B., Bozanta, A. and Coskun, M. (2018) 'Usage factors of location-based social applications: the case of Foursquare', *International Journal of Web-Based Communities*, Vol. 14, No. 2, pp.128–148.

Lyu, J. and Lim, H. (2018) 'The role of sense of community in brand online social networking', *International Journal of Web-Based Communities*, Vol. 14, No. 2, pp.149–171.

Ma, X., Ma, J. and Li, H. (2018) 'AGENT: an adaptive geo-indistinguishable mechanism for continuous location-based service', *Peer-to-Peer Networking and Applications*, Vol. 11, No. 3, pp.473–485.

Md, A.Q. and Vijayakumar, V. (2019) 'Dynamic ranking of cloud services for web-based cloud communities: efficient algorithm for rating-based discovery and multi-level ranking of cloud services', *International Journal of Web-Based Communities*, Vol. 15, No. 3, pp.248–270.

Ni, L., Li, C., Wang, X. et al. (2018) 'DP-MCDBSCAN: Differential privacy preserving multi-core DBSCAN clustering for network us er data', *IEEE Access*, 7 April, Vol. 6, pp.21053–21063.

Piao, C., Shi, Y., Yan, J. et al. (2019) 'Privacy-preserving governmental data publishing: a fog-computing-based differential privacy approach', *Future Generation Computer Systems*, January, Vol. 90, pp.158–174.

Shin, H., Kim, S., Shin, J. and Xiao, X. (2018) 'Privacy enhanced matrix factorization for recommendation with local differential privacy', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 30, No. 9, pp.1770–1782.

Soria-Comas, J., Domingo-Ferrer, J., Sanchez, D. and Megias, D. (2017) 'Individual differential privacy: a utility-preserving formulation of differential privacy guarantees', *IEEE Transactions on Information Forensics and Security*, Vol. 12, No. 6, pp.1418–1429.

Taamallah, A., Khemaja, M. and Sami, F. (2019) 'A web-based platform for strategy design in smart cities', *International Journal of Web-Based Communities*, Vol. 15, No. 1, pp.62–84.

Tian, X., Song, Q. and Tian, F. (2018) 'Multidimensional data aggregation scheme for smart grid with differential privacy', *International Journal of Network Security*, Vol. 20, No. 6, pp.1137–1148.

Ugolotti, R., Sassi, F. et al. (2013) 'Multi-sensor system for detection and classification of human activities', *Journal of Ambient Intelligence and Humanized Computing*, Vol. 4, No. 1, pp.27–41.

Ul Hassan, M., Rehmani, M.H. et al. (2019) 'Differential privacy for renewable energy resources based smart metering', *Journal of Parallel and Distributed Computing*, September, Vol. 131, pp.69–80.

Wang, Q., Chen, D., Zhang, N. et al. (2017) 'PCP: a privacy-preserving content-based publish-subscribe scheme with differential privacy in fog computing', *IEEE Access*, 2 September, Vol. 5, pp.17962–17974.

Wang, S. and Sinnott, R.O. (2017) 'Protecting personal trajectories of social media users through differential privacy', *Computers and Security*, 1 June, Vol. 67, pp.142–163.

Wang, Y., Huang, M., Jin, Q. and Ma, J. (2018) 'DP3: a differential privacy-based privacy-preserving indoor localization mechanism', *IEEE Communications Letters*, Vol. 22, No. 12, pp.2547–2550.

Xiong, P., Zhang, L. and Zhu, T. (2017) 'Reward-based spatial crowdsourcing with differential privacy preservation', *Enterprise Information Systems*, Vol. 11, No. 10, pp.1500–1517.

Yang, M., Zhu, T., Xiang, Y. and Zhou, W. (2018) 'Density-based location preservation for mobile crowdsensing with differential privacy', *IEEE Access*, 16 March, Vol. 6, pp.14779–14789.

Yu, K. and Cao, M. (2018) 'Differential privacy and qualitative privacy analysis for nonlinear dynamical systems', *IFAC-PapersOnLine*, Vol. 51, No. 23, pp.52–57.