Towards Efficient Privacy-Preserving Personal Information in User Daily Life

Hai Wang[1(✉)], Tong Feng[1], Zhe Ren[1], Ling Gao[2], and Jie Zheng[1]

[1] Northwest University, Xi'an, China
hwang@nwu.edu.cn
[2] Xi'an Polytechnic University, Xi'an, China

Abstract. The popularity of smart home has added a lot of convenience to people's lives. However, while users use these smart products, users' privacy data has also been leaked and it may cause some risks. Besides, because of untrusted third-party servers, we simply use traditional privacy-preserving methods could no longer protect users' private information effectively. In order to solve these problems, this paper proposes a privacy-preserving method for multi-private data: We first determine the privacy data format that needs to be protected, such as audio or text. Secondly, if the data format is text, we will use the local differential privacy method. We first obtain the key attributes of the user from the key information chain, and then select the appropriate localized differential privacy method according to the text characteristics of the key attributes. The user realizes the local disturbance of the data and then uploads it to the data collection center–the cloud platform. Finally, when an attacker attempts to obtain user information from the cloud platform, it uses the central differential privacy method to add noise and the noise-added data is transmitted to the attacker. If the data format is voice frequency, we first convert the voice information into binary code, then chaotically encrypt the binary code, and upload the encrypted binary code to the cloud platform. We verify the effectiveness of our methods by experiments, and it can protect users' privacy information better.

Keywords: Multi-privacy data · Key information chain · Localized differential privacy · Noise adding · Privacy-preserving

1 Introduction

As Internet of Things is deeply used in various industries, IoT devices have increased more connections with people. Smart home is a very common scene. When users want to play music, they only need to input voice commands to the smart speakers, when the users come back, the smart lock will open automatically; when the sun rises or falls, the smart curtain will follow, the smart camera allows users to keep an eye on the room and ensure the safety of the room. However, since the smart device can access very private data, for example, a smart camera can also be used by an attacker to photo the user's privacy, a smart speaker can obtain the user's voice information, and the smart

B. Li et al. (Eds.): IoTaaS 2019, LNICST 316, pp. 503–513, 2020.
https://doi.org/10.1007/978-3-030-44751-9_42

lock's password is related to the security of the entire room. Therefore, the Internet of Things has also raised concerns about the privacy of these digitally augmented spaces [1–3]. Most existing researches on IoT security issues focus on privacy data leakage detection. FlowFence [4] is a system that enforces streaming policies for IoT applications to protect sensitive data; ContexIoT [5] is a context-based licensing system for the IoT platform that collects contextual information to identify sensitive operations. Zawoad et al. [6] formally defined the Internet of Things forensics and proposed a forensic sensing Internet of Things (FAIoT) model to support forensic research in the Internet of Things infrastructure. ProvThings [7] seized a series of events through the security-sensitive system-level SmartThings API and used it for forensic reconstruction attacks. SAINT [8] is the first accurate system to carefully detect sensitive data streams in IoT applications to fully identify a complete set of pollution sources and sinks through analoging IoT-specific challenges, solutions for platform and specific language issues. Although the means of detection can make security personnel better understand the problems in the environment, the ultimate goal of detection is to better protect privacy. Therefore, it is also a feasible solution to solve this problem directly from the perspective of privacy protection.

Differential privacy technology [9, 10] is a hot research in the current academic world. Traditional differential privacy technology concentrates raw data into a data center and then publishes relevant statistical information that satisfies differential privacy, we call it centralized differential privacy technology. Centralized differential privacy protects sensitive information based on a premise: trusted third-party collectors, ensuring that third-party data collectors do not steal and reveal sensitive users' information. However, in practical applications, even the privacy of these trusted third-party collector users is still not guaranteed. In 2016, nearly 170 million accounts of American social networking site LinkedIn were publicly sold by hackers in the black market; personal information of nearly 5,000 W citizens in Turkey was leaked, and the President's personal information was hanged on the dark network platform; Yahoo happened the largest data breach, more than 5 billion users' account information was stolen by hackers. There are many similar examples, so it is very difficult to find a truly trusted third-party data collection platform in practical applications. Localized differential privacy [11, 12] came into being under such a background. Localized differential privacy has two major characteristics: (1) fully consider the background knowledge of any attacker and quantify the degree of privacy protection; (2) localize the perturbed data to protect against privacy attacks from untrusted third-party data collectors. In sensitive image feature extraction scenarios, localized differential privacy shows how important it is. Finn [13] took the medical images an example to illustrate the privacy problem contained in the image. Qin [14] proposes the encryption-based privacy protection method in the image feature extraction process in the cloud computing environment. Ren [15] pointed out that in the cloud computing environment local differential privacy had great potentiality in image processing. After reading much paper, we choose to apply the local differential privacy method to the smart home scene, and realize the adaptive privacy protection method for the privacy data in the smart home scene. Firstly, the user needs to judge the privacy data type, and then according to different data types user differently realizes the disturbance of ε-localized differential privacy, next transmit it to the third-party data collector.

The data collector receives the disturbed data and performs a series of queries and refinement to obtain effective statistical results.

Our contributions can be summarized as follows:

(1) Combining the differential privacy protection technology with the smart home scene, using the privacy protection method based on privacy data leakage detection method, so that the privacy of the user is better protected better.
(2) Using the localized differential privacy method to process the text type privacy data, using different methods for different text characteristics, and finally realizing the adaptive localized differential privacy protection for the text type privacy data.
(3) Using chaotic encryption algorithm to process speech type data to ensure the security of speech in the process of transmission, and using experiment to verify the effectiveness of methods we propose.

2 Model

Before implementing privacy protection, we need to understand what data needs to be protected. In the smart home scenario, this paper mainly studies text privacy data and voice privacy data. The text types include personal information (such as ID number and mobile phone number), room password, communication record and content, and whereabouts; the voice type is mainly the user's voice information, such as the voice command issued by the user to control the smart device, the voice when you call in the room or talk to other members of the family. Since these two types of private data are handled differently, we will use different privacy protection methods protect them.

2.1 Text Type Private Data

The privacy data of the text type includes two types: numerical data (such as age, telephone number) and non-numeric data (such as name and gender). For the user, the sensitive attributes are various, if users need to use the localized differential privacy method to protect every sensitive attribute, which will cause great trouble to users, and the user's sense of smart devices' convenience will be get bad. In this paper, each sensitive attribute of the user is connected to form a complete information chain of the user. In this information chain, some sensitive attributes are key attributes that can constitute the overall information of the user, while other sensitive attributes are subordinate to these key attributes. As shown in the Fig. 1 below, the blue circle represents the key sensitive attribute, and the white box represents other sensitive attributes. If all the key sensitive attributes represented by the circle are leaked, then a specific person can be located, so it is easy to infer other sensitive attributes. For example, If the text information of a user is 'age 27, Hangzhou, Zhejiang Province, China, square face, 1.9 m height', then the key information may be clearly distinguishable '1.9 m height, Hangzhou, Zhejiang Province, China'.

We can see that if we can hide the user's key sensitive attributes, the user's private information can be effectively protected. As in the above example, we blur the key

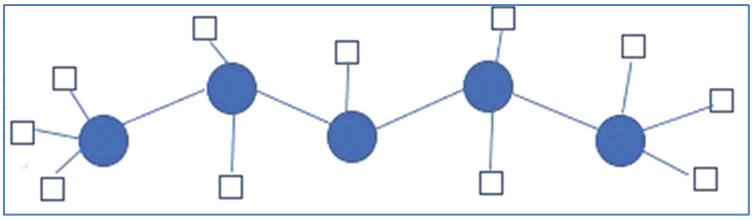

Fig. 1. Key information chain

information, then the user information becomes "age 27, a southern province in China, square face, 1.9 m height" or "age 27, Hangzhou, Zhejiang Province, China, square face", through this information it is not possible to locate a specific person, so it can be seen that the user's key attributes are very important to user privacy protection.

Through the key information chain, we can understand what are the user's key attributes, and then users can localize differential privacy protection for these key attributes, which can save users' time and make users get better service. On the other hand, the user's critical privacy is protected from the risk of data leakage. Since the sensitive attributes of user in the smart home scene are mostly discrete, this paper mainly uses the frequency statistics method based on localized differential privacy to realize the privacy protection for text type data. The frequency statistics method is divided into single value and multi value. The RAPPOR [16] method is representative of single value frequency statistics. The value of a variable is represented as a string. Suppose there are a total of n users, the i_{th} user u_i corresponds to a certain sensitive value, $x_i \in X$ and $|X| = k$, and now it is desirable to count the frequency of the value $x_i (1 \leq i \leq k)$, the RAPPOR method is shown as Fig. 2 shows.

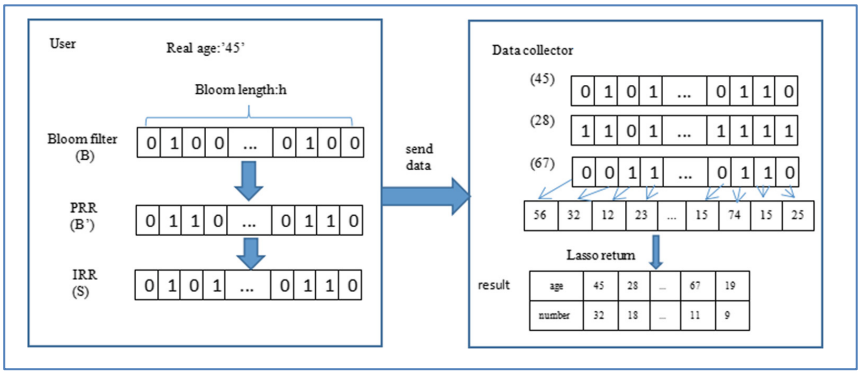

Fig. 2. Process of RAPPOR

In Fig. 2, X represents the age attribute, there is a certain value $x_i =' 45'$, we first used the Bloom Filter [17] technique to express it as a vector $B = (0, 1)^h$ that its length is h, and the mapping relationship matrix of the Bloom string is recorded at the same time. Then we perturbed to each bit of the vector B by using a random response technique

to obtain a permanent random response result B', the perturbation mode is performed according to the following formula, $f \in [0, 1]$ indicates the probability value:

$$p\left(B_i' = x\right) = \begin{cases} 0.5f & x = 1 \\ 0.5f & x = 0 \\ 1 - f & x = B_i \end{cases} \tag{1}$$

Then, the second perturbation of each bit of the vector B' is performed to obtain a transient random response result S, the second perturbation mode is performed according to the following formula, where $p \in [0, 1]$ and $q \in [0, 1]$ indicates the probability that S_i is 1 when B_i' takes a value of 1 or 0:

$$P(S_i = 1) = \begin{cases} p \; if \; B_i' = 1 \\ q \; if \; B_i' = 0 \end{cases} \tag{2}$$

After each user gets the disturbance result S, it is sent to the third-party data collector. The data collector counts the number of occurrences of each bit and corrects it, and then combines the mapping matrix to complete each frequency statistics corresponding to the age value by Lasso regression method [18].

In view of the high communication cost of the RAPPOR method, also in the single-value case, each user in the S-Hist [19] method encodes a character string, randomly selects one of the bits, and uses the random response technique to perform the disturbance, then send it to the data collector, thus greatly reducing the transmission cost. K-RR [20] is a gradient response technique proposed by Kairouz et al. It mainly overcomes the problem that the random response technique is for binary variables. For the case where the variable contains k(k > 2) candidate values, it can be directly make a random response. For any input $R \in X$, the response to the output $R' \in X$ is as follows:

$$P\left(R'|R\right) = \frac{1}{k - 1 + e^{\epsilon}} \begin{cases} e^{\epsilon} \; if \; R' = R \\ 1 \quad if \; R' \neq R \end{cases} \tag{3}$$

That is, the probability of $\frac{e^{\epsilon}}{k-1+e^{\epsilon}}$ is used to respond to the real result, and the probability of $\frac{1}{k-1+e^{\epsilon}}$ is used to respond to any of the remaining k – 1 results. Make it satisfy ε-localized differential privacy.

Based on the K-RR method, Kairouz et al. proposed the O-RR method for the case where the value of the variable is unknown [21]. The O-RR method is an improvement of the K-RR method. Hash mapping and grouping operations are also introduced on the basis of K-RR. The hash mapping makes the method no longer pay attention to the string itself, so that no candidate characters string list need to be collected in advance. And the probability of hash map value collisions can be further reduced by grouping operations. The above describes several frequency statistics methods for localized differential privacy. In the actual situation, we will select the most suitable localized differential privacy protection method for different text types, which can reduce the communication cost and save the overhead.

After selecting the appropriate localized differential privacy method for the text, we can upload the disturbed data to the cloud platform, and the cloud platform will perform the second noise addition and then release the data by the exponential mechanism. The general idea of the exponential mechanism is to select an output value r from the output field based on the score of the usability q, and the probability of selecting the value is exponentially proportional to the score. The function q needs to be insensitive to the variation of a single record, that is, the function sensitivity is low, and its sensitivity can be expressed as $\Delta q = \max_{\forall r, D1, D2} |q(D1, r) - q(D2, r)|$. The privacy protection algorithm can be designed by an exponential mechanism, as shown by the theorem.

Given a data set D and an availability function $q = (D \times R) \rightarrow R$, the privacy protection mechanism A satisfies ε-differential privacy if and only if the following expression holds:

$$A(D, q) = \left\{ reutrn\ with\ probability \infty exp\left(\frac{\epsilon q(D, r)}{\Delta q} \right) \right\} \tag{4}$$

The exponential mechanism is the operation of the cloud platform. It is not our main research method, so it is not detailed. Through the above method, we can effectively protect the text data, and finally realize the privacy protection of the text type data.

2.2 Voice Type Private Data

The speech signal is a simple and succinct signal with a large amount of information, such as accent, language, emotion, gender identity and speech content. All kinds of information are expressed in one-dimensional signals. From the perspective of structure, human language information can be divided into three layers: the first layer is language information, including speech content and sentences; the second layer is sub-language information, including the speaker's attitude, emotion, intention, etc. Pitch, rhythm, volume, and tone, etc. can show intentions and attitudes; the third layer is non-verbal information, such as physical condition, age, and gender. The linguistic information is conducive to prevent forgery, and is also conducive to the preservation of living evidence; the paralinguistic information is useful for detecting true intentions, and the non-verbal information can be partially traced back to evidence. Therefore, we need to start from the above single layer to achieve privacy protection of voice data. Speech recognition is also a hot topic of current research, so if we want to protect the privacy of voice type data, then the starting point is speech recognition. Speech recognition includes two processes of recognition and text translation. If the user speaks a local dialect, then the approximate area of the user can be determined according to the dialect type, and the user's personal preference can be understood according to the user's speech during the translation stage. After translating the voice information into text, it can be processed according to the above-mentioned text type privacy data, so we mainly protect the privacy of the speech recognition process in this section.

The goal of speech recognition is to convert the vocabulary content in human speech into computer-readable input, such as buttons, binary codes or character sequences. This paper encrypts the digitized code of the speech signal to protect the user's voice information. Both compressed sensing and chaos can be applied to the field of data

encryption. This paper combines the two methods to implement the chaotic encryption method based on compressed sensing. The mathematical model is as follows:

$$C = M\left(\dot{X}\right) \tag{5}$$

M() is Arnold map, and C is a ciphertext information that needs to be transmitted. In Eq. (5), $\dot{X}$ is the superimposed n × n dimensional matrix of X, X is:

$$X = AS + L\left(A\bar{S}\right) \tag{6}$$

A is the key matrix, we choose the random Gauss matrix as the key matrix, A is a random Gauss matrix, and $A \in R^{M\times N}(M \ll N)$, the construction process of the key matrix A: construction a matrix A of $M \times N$ size, each element in A independently obeys a Gaussian distribution with a mean of 0 and a variance of 1/M, that is:

$$A_{i,j} = N\left(0, \frac{1}{M}\right) \tag{7}$$

S is a speech signal that needs to be encrypted. According to the short-term stationary analysis of the speech signal, the speech signal needs to be pre-processed in the encryption process, that is, the encryption process is actually after the speech signal pre-processing framing, for each frame. The signal is encrypted. The voice signal frame length is selected to be 400, that is, the plaintext S length is 400, and the key matrix A has a dimension of 256 × 400. L is the Lorenz transform. This encryption model uses the third-order Lorenz equation as the transformed chaotic system, and $\bar{S}$ is the fixed random signal in the same dimension as the plaintext signal. The information is known to both parties, that is $\bar{S} \in R^N$. The key matrix A and the random signal $\bar{S}$ are transformed as parameters in the Lorenz system, and the key space is added directly to the voice signal encrypted by the compressed sensing, which has better security. The signal X processed by the Eq. (6) is a 256-dimensional column signal, which is sequentially superimposed into $\dot{X}$, $\dot{X}$ is a 16-order square matrix, and then the position of each element in the $\dot{X}$ is performed using Arnold map. Scrambling. The Arnold mapping can be expressed as:

$$\begin{bmatrix} X_{n+1} \\ Y_{n+1} \end{bmatrix} = U \begin{bmatrix} X_n \\ Y_n \end{bmatrix} (mod \quad 1) \tag{8}$$

U is:

$$U = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \tag{9}$$

As shown in Eq. (8), mod1 means that only the fractional part is taken. In order to scramble $\dot{X}$ and spread the formula (7), there is:

$$\begin{bmatrix} X_{n+1} \\ Y_{n+1} \end{bmatrix} = U \begin{bmatrix} X_n \\ Y_n \end{bmatrix} (mod \quad N) \tag{10}$$

After the diffusion, in order to ensure its feasibility, the matrix U in Eq. (9) is processed, and the parameters a, b(a, b < N) are introduced, then the matrix U can be expressed as:

$$U_d = \begin{bmatrix} ab+1 & a \\ b & 1 \end{bmatrix} \tag{11}$$

a, b need to satisfy the condition 0 < a < 256, 0 < b< 256. After scrambling it with Arnold map, the encryption operation is completed, and the encrypted information becomes ciphertext, which is transmitted to the receiver through the channel. Decryption is the inverse of encryption. It is not described here. The signal is reconstructed from the decrypted code to obtain the original signal.

3 Experiment and Analysis

For the above methods, we have carried out experimental verification, and the experimental results prove the effectiveness of our proposed method.

3.1 Text Type Private Data Experiment

In this paper, experiments are carried out by adopting different privacy budgets ε. The communication cost, progressive error boundary and computational cost are compared and analyzed among the above methods. The communication cost refers to the data transmission overhead from each user to the data collector. Here, we approximate that the communication cost is proportional to the amount of data. In the progressive error boundary, n refers to the total number of users, k refers to the number of attribute candidates, and h represents the length of the Bloom Filter string. The calculation cost refers to the calculation cost when the data collector counts the user data, and it is divided into three levels: high, medium, and low.

It can be seen from Table 1 that the availability based on the RAPPOR method is higher, and it also brings higher computational overhead; The S-Hist-based method greatly reduces the communication cost, but its computational cost is positively correlated with the number of users, the computational overhead is huge, and the sampling process also brings a certain precision loss; the K-RR-based method simplifies the data perturbation process and sacrifices certain release precision.

3.2 Voice Type Private Data Experiment

This paper verifies the feasibility of the above method for speech encryption through experiments. The experiment selected a male voice from our laboratory with a total of 23,455 points and a sampling rate of 16 K. Different frame lengths and compression ratios (M/N) were chosen to test the results. The results are shown in Table 2 and Fig. 3:

The average frame signal-to-noise ratio AFSNR is used to evaluate the recovery quality of the signal. The larger the AFSNR, the smaller the difference between the signal after decryption and the original signal. It can be seen that different frame lengths

Table 1. Existing four methods of single-valued frequency estimation

Methods	Communication cost	Progressive error boundary	Computational cost
RAPPOR	$o(h)$	$o\left(\frac{k}{\epsilon\sqrt{n}}\right)$	High
S-Hist	$o(1)$	$o\left(\frac{\sqrt{\log k}}{\epsilon\sqrt{n}}\right)$	High
K-RR	$o(1)$	$o\left(\frac{\sqrt{k^2}}{\epsilon\sqrt{n}}\right)$	Low
O-RR	$o(1)$	$o\left(\frac{\sqrt{k^2}}{\epsilon\sqrt{n}}\right)$	Low

Table 2. Voice signal test

Frame length	Compression ratios M/N				
	0.2	0.3	0.4	0.5	0.6
160	5.461	8.950	13.637	17.877	20.872
240	5.822	8.154	12.554	16.928	19.296
320	6.697	9.775	13.656	17.513	18.219
400	6.767	9.954	12.037	15.952	15.972
480	6.708	9.235	11.780	13.102	24.773

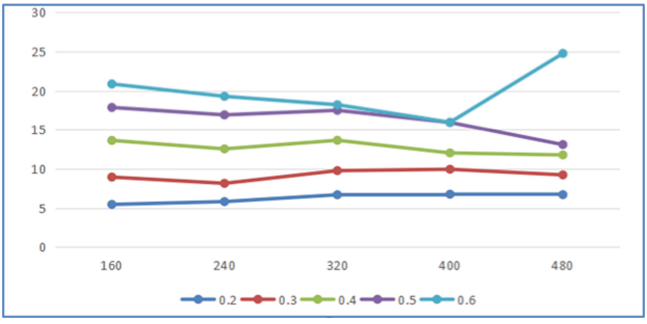

Fig. 3. Voice signal test

and compression ratios have an impact on the accuracy of the decryption. And the accuracy is best when the frame length is 400 and the compression ratio is 0.6.

Through the above experiments, we use the existing localized differential privacy method for text type private data, and select the appropriate privacy protection method according to the data characteristics (such as the number of attribute values), and can estimate the cost according to the importance of privacy, then we select the better method;

For the speech type, we use the chaotic encryption algorithm based on compressed sensing. It is verified by experiments that this method can effectively protect the privacy of users information.

4 Conclusion

This paper focuses on the privacy security issues in the smart home scene. Different from the existing privacy disclosure detection, this paper takes a different approach. We divided the privacy data types in the smart home scene into two types: text type and voice type. For the text type, we first integrate and analyze the user's text privacy data, and put forward the idea of a key information chain, so that users do not need to protect every sensitive attribute, only need to protect key attributes. In addition, considering the problem of untrusted third-party servers, it is proposed to introduce the localized differential privacy method into the scenario, and describe and analyze several existing localized differential privacy methods. The specific experiments verify our analysis. For the voice type, this paper combines the compressed sensing and chaotic encryption methods, and uses these two methods to encrypt the speech-coded binary code. The experimental results show that the proposed method can effectively protect voice information and prevent threats caused by interception of voice information.

This paper protects the user privacy from the two data types of voice and text, but the method proposed in this paper is not used in the real smart home scene, the effectiveness of the method is just verified by simulation experiments. In the case of the effect, further exploration is needed. On the other hand, the privacy in this scene also includes information such as pictures and videos, and it will be more complicated. We need to study more deeply, and we can complete the comprehensive privacy protection work under the smart home scene in the future.

References

1. Ronen, E., Shamir, A., Weingarten, A.-O., O'flynn, C.: IoT goes nuclear: creating a ZigBee chain reaction. In: IEEE Security and Privacy (SP) (2017)
2. Fernandes, E., Jung, J., Prakash, A.: Security analysis of emerging smart home applications. In: IEEE Security and Privacy (SP) (2016)
3. Celik, Z.B., Mcdaniel, P., Tan, G.: SOTERIA: automated IoT safety and security analysis. In: USENIX ATC (2018)
4. Fernandes, E., Paupore, J., Rahmati, A., Simionato, D., Conti, M., Prakash, A.: FlowFence: practical data protection for emerging IoT application frameworks. In: USENIX Security (2016)
5. Jia, Y.J., et al.: ContexIoT: towards providing contextual integrity to appified IoT platforms. In: NDSS (2017)
6. Zawoad, S., Hasan, R.: FAIoT: towards building a forensics aware eco system for the internet of things. In: SCC, pp. 279–284 (2015)
7. Wang, Q., Hassan, W.U., Bates, A., Gunter, C.: Fear and logging in the internet of things. In: NDSS (2018)
8. Celik, Z.B., et al.: Sensitive information tracking in commodity IoT. In: USENIX Security (2018)

9. Dwork, C., Lei, J.: Differential privacy and robust statistics. In: Proceedings of the 41st Annual ACM Symposium on Theory of Computing, pp. 371–380. ACM (2009). https://doi.org/10.1145/1536414.1536466
10. Smith, A.: Privacy-preserving statistical estimation with optimal convergence rates. In: Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, pp. 813–822. ACM (2011). https://doi.org/10.1145/1993636.1993743
11. Kasiviswanathan, S.P., Lee, H.K., Nissim, K., Raskhodnikova, S., Smith, A.: What can we learn privately. In: Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 531–540. IEEE (2008)
12. Duchi, J.C., Jordan, M.I., Wainwright, M.J.: Local privacy and statistical minimax rates. In: Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 429–438. IEEE (2013). https://doi.org/10.1109/focs.2013.53
13. Finn, R.L., Wright, D., Friedewald, M.: Seven types of privacy. In: Gutwirth, S., Leenes, R., de Hert, P., Poullet, Y. (eds.) European Data Protection: Coming of Age, pp. 3–32. Springer, Dordrecht (2013). https://doi.org/10.1007/978-94-007-5170-5_1
14. Qin, Z., Yan, J., Ren, K., Chen, C.W., Wang, C.: Towards efficient privacy-preserving image feature extraction in cloud computing. In: Proceedings of the 22nd ACM International Conference on Multimedia, pp. 497–506. ACM (2014). https://doi.org/10.1145/2647868.2654941
15. Ren, K.: Privacy-preserving image processing in cloud computing. Chin. J. Netw. Inf. Secur. **1**, 12–17 (2016)
16. Erlingsson, Ú., Pihur, V., Korolova, A.: RAPPOR: randomized aggregatable privacy-preserving ordinal response. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1054–1067. ACM (2014). https://doi.org/10.1145/2660267.2660348
17. Bloom, B.H.: Space/time trade-offs in hash coding with allowable errors. Commun. ACM **13**(7), 422–426 (1970). https://doi.org/10.1145/362686.362692
18. Tibshiran, R.: Regression shrinkage and selection via the Lasso. J. Roy. Stat. Soc. (Ser. B-Methodol.) **58**, 267–288 (1996)
19. Bassily, R., Smith, A.: Local, private, efficient protocols for succinct histograms. In: Proceedings of the 47th Annual ACM on Symposium on Theory of Computing, pp. 127–135. ACM (2015). https://doi.org/10.1145/2746539.2746632
20. Kairouz, P., Oh, S., Viswanath, P.: Extremal mechanisms for local differential privacy. In: Advances in Neural Information Processing Systems, pp. 2879–2887 (2014)
21. Kairouz, P., Bonawitz, K., Ramage, D.: Discrete distribution estimation under local privacy. In: Proceedings of the 33rd International Conference on Machine Learning, New York, pp. 2436–2444 (2016)