The Journal of Supercomputing (2020) 76:8391–8412
https://doi.org/10.1007/s11227-019-03104-0

A survey of local differential privacy for securing internet of vehicles

Ping Zhao[1] · Guanglin Zhang[1] · Shaohua Wan[2] · Gaoyang Liu[3] · Tariq Umer[4]

Published online: 9 December 2019

Abstract
Internet of connected vehicles (IoV) are expected to enable intelligent traffic management, intelligent dynamic information services, intelligent vehicle control, etc. However, vehicles' data privacy is argued to be a major barrier toward the application and development of IoV, thus causing a wide range of attentions. Local differential privacy (LDP) is the relaxed version of the privacy standard, differential privacy, and it can protect users' data privacy against the untrusted third party in the worst adversarial setting. Therefore, LDP is potential to protect vehicles' data privacy in the practical scenario, IoV, although vehicles exhibit unique features, e.g., high mobility, short connection times, etc. To this end, in this paper, we first give an overview of the existing LDP techniques and present the thorough comparisons of these work in terms of advantages, disadvantages, and computation cost, in order to get the readers well acquainted with LDP. Thereafter, we investigate the potential applications of LDP in securing IoV in detail. Last, we direct several future research directions of LDP in IoV, to bridge the gaps between LDP researches and the privacy preservation in IoV. The originality of this survey is that it is the first work to summarize and compare the existing LDP research work and that it also does an pioneering work toward the in-depth analysis of the potential applications of LDP in privacy preservation in IoV.

Keywords Internet of connected vehicles (IoV) · Data privacy · Local differential privacy (LDP) · Differential privacy (DP)

1 Introduction

Internet of connected vehicles (IoV) will significantly facilitate many applications, e.g., intelligent traffic management, intelligent dynamic information services, intelligent vehicle control, etc., since 5.2 million vehicles are on world's roads by 2017 [1,

✉ Shaohua Wan
shwanhust@gmail.com; shaohua.wan@ieee.org

Extended author information available on the last page of the article

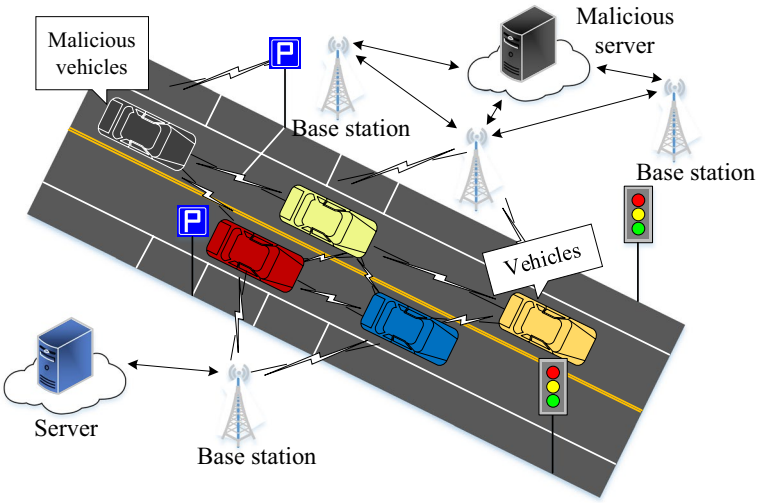


Fig. 1 Illustration of internet of connected vehicles (IoV) and the risk of disclosing data privacy

2] and the number of vehicles is expected to be doubled by 2040 [3]. The potential data disclosure of such a large number of unsecured vehicles is the bottleneck of the application and development of IoV. Therefore, the data privacy issues in IoV have caused a wide range of attentions.

IoV definitely facilitates human's life, but such benefits come with a huge price of data privacy [4, 5]. As shown in Fig. 1, vehicles in IoV are connected with each other and request various services via sharing data with other vehicles and servers. But it incurs potential risk of disclosing data privacy of vehicles to attackers, since other vehicles and the servers may be malicious. For example, security experts Kevin Mahaffey and Marc Rogers proved that attackers can get the data of a specific Tesla vehicle and further suddenly shut down the system engine to stop the vehicle [6]. For another example, two US cybersecurity experts have invaded the control system of Cherokee Jeep, and they could further control the jeep's wiper, speed, built-in air conditioner, and radio and even control the car's driving route at home [7]. To make matters worse, network security company UpGuard released a report that the confidential data of more than 100 car manufacturers have been leaked, including a number of traditional car companies such as GM, Ford, Toyota, and Volkswagen, as well as new forces Tesla [7]. Obviously, if vehicles' data in IoV are breached by attackers, it will cause serious personal and property safety issues [8, 9], e.g., interest, health status, religious beliefs, living habits, social relationships, etc.

To address the data privacy in IoV, many governments in the world have taken measures. For example, Australian Government established the Traffic and Facilities Standing Committee in 2011, aiming to process the security and privacy issues in smart transportation [10, 11]. Moreover, National Institute of Standards and Technology of United States has developed the "Safety Framework for Vehicle Network Attack Protection" in 2014 [10, 12]. In addition, General

Data Protection Regulation (GDPR) [10] enforced on May 25, 2018, requires that vehicle network applications should be implemented according to the privacy-preserving measures.

In academia, differential privacy (DP) is proposed and regarded as an extremely strong privacy standard. It formalizes both the degree of privacy preservation and data utility and can provably guarantee data privacy in the worst adversarial setting where adversaries know all users' data except one. But in the paradigm of DP, it relies on a third party that is assumed do not disclose users' data privacy during the process of perturbing users' data and releasing the noisy version of the data (cf. Fig. 2a). Therefore, DP suffers from a drawback that in many practical scenarios, the assumption does not hold, as the third party may disclose users' data for interests or be attacked by hackers.

An relaxed version of DP, local differential privacy (LDP), has the potential to guarantee the data privacy in IoV [13, 14]. Its main idea is that the third party does not collect the exact data of each individual, and yet still be able to compute the correct statistical results. In particular, in LDP, each individual locally perturbs his data with a differentially private mechanism and then sends the noisy version of the data (i.e., perturbed data) to the untrusted third party (cf. Fig. 2b). Upon receiving the noisy data of users, the untrusted third party computes the statistics and releases the statistical results. Thus far, LDP is being used in four different parts of iOS 10 [15]; moreover, Google open-source project called RAPPOR has cooperated LDP to protect Chrome users' privacy [16].

However, while there are a great number of studies focusing on LDP, there is a lack of comparisons of these work and analysis of their potentials for securing IoV. Work [13, 14, 17, 18] investigated the data publication based on local differential privacy. Moreover, the literature [19–23] extended local differential privacy to machine learning. Another set of studies [24–28] were devoted to processing users' queries using local differential privacy. Furthermore, studies [29–33] focused on the application of local differential privacy systems, employing local differential privacy to recommendation systems, social network, smart transportation system, search engine, etc. Nevertheless, there does not exist a survey that compares these work and analyzes their potentials to protect data privacy in IoV.

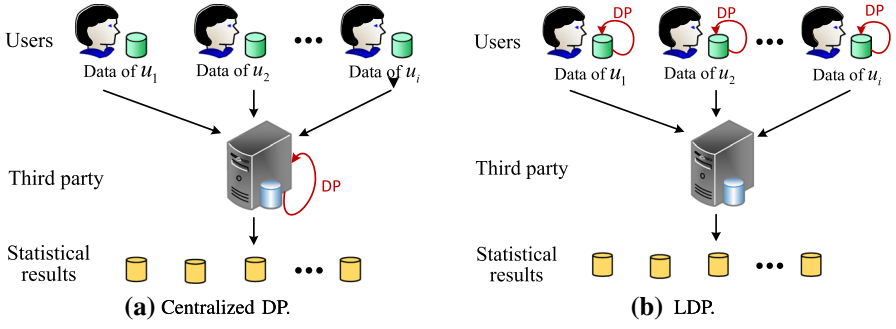


Fig. 2 Illustration of frameworks of centralized DP and LDP

To this end, in this paper, we first give an overview of the existing LDP studies and compare the advantages, disadvantages, computation complexity, and the privacy budget of these work, to get the readers well acquainted with LDP. Thereafter, we bring forth the applications of LDP in IoV in detail. Last, we discuss several open research directions about LDP in IoV to bridge the gaps between LDP researches and the privacy preservation in IoV. In summary, the main contributions of this paper are as follows:

- To the best of our knowledge, it the first work to survey the existing work about LDP in terms of advantages, disadvantages, computation complexity, and the privacy budget, and so on.
- It is also the first efforts toward investigating the potential applications of LDP in securing IoV in several typical scenarios and highlighting the new challenges when LDP is applied to IoV applications.

The remainder of this paper is organized as follows: Sect. 2 summarizes and compares the existing work about LDP. Section 3 introduces the attack model and the potential applications of LDP in IoV, followed by the future research opportunities in Sect. 4. Finally, Sect. 5 concludes the paper.

2 LDP researches

Existing work mainly focused on applying LDP to the data publication, machine learning, and query processing. In the following, we overview the existing work in detail.

2.1 Data publication

Data publication based on local differential privacy mainly uses non-interactive framework (cf. Fig. 3) to release statistical information of sensitive data and enables the released data to meet the needs of data analysis at the same time. The commonly

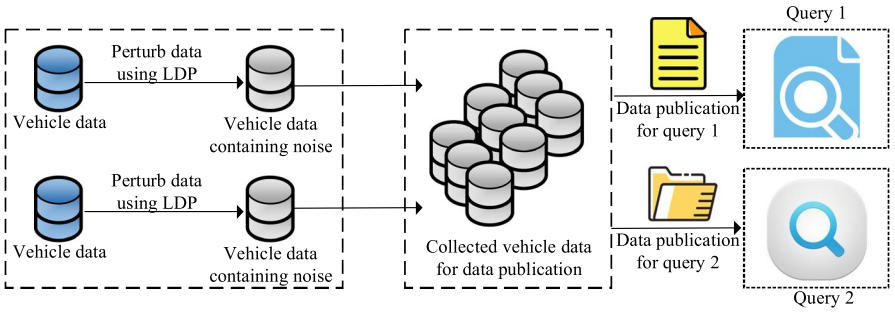


Fig. 3 Illustration of the applications of LDP in data publication

used data publication technologies mainly include histogram, partitioning, and sampling filter [34–38].

The earliest research RAPPOR [17] proposed to collect crowd-sourced statistics from users' data without violating users' data privacy. But it has two disadvantages: (1) the communication cost between users and the data collector is relatively high, because each user needs to transmit a vector of length h to the data collector and (2) data collector is required to collect candidate string lists in advance for frequency statistics.

To address the first drawback of RAPPOR, i.e., the high communication cost, S-Hist [18] proposed to randomly select bits using randomized response techniques to perturb users' data and send the perturbed data to the collector. Although S-Hist greatly reduces the communication overhead, the accuracy of the S-Hist is not stable in practical applications, mainly because the positive and negative values of each element in the random projection matrix are random. To make the inner product of any two column vectors to be zero, the candidate value k of the string is supposed to be large enough. Moreover, Nguyn et al. [39] proposed Harmony for dependent data, which can be applied to the mean statistics of continuous data and the frequency statistics of discrete data. In addition, LDPMiner [19] is a combined method based on RAPPOR [17] and S-Hist [18], which can effectively deal with the stream frequent item mining on set-valued data. To contrate, LDPMiner mechanism divides tasks into two sub-processes, thus improving the utility of data. LoPub [40–46] is a multi-value frequency statistics combining the RAPPOR [17] and the probability graph model, where users locally perturb its data with RAPPOR and send perturbed data to the collector for data dimensionality reduction, and finally, the collector synthesizes dataset.

To deal with the second drawback of RAPPOR [17], i.e., candidate string list collected in advance, Kairouz et al. [47] proposed O-RAPPOR based on the unknown values of variables and designed hash mapping and grouping operations. Specifically, the hash function first performs one-time value mapping for strings in different groups to generate Bloom Filter strings. Then, the hash value data are perturbed by RAPPOR. As such, there is no need to collect candidate string list in advance, thus guaranteeing the data utility. Likewise, to address the second disadvantage of RAPPOR, an improved method for multi-valued frequency statistics, RAPPOR-unknown, was proposed [48], which extracts r substrings of the same length from the perturbed data according to the n-gram. Thereafter, the perturbed data and the substring information are sent to the data collector. Although this method improves the RAPPOR, the communication cost is higher, and thus, it is not applicable to the case where there are a large number of substrings. However, the studies [17, 18, 47, 49] are only applicable to the case where the value of the variables is binary and not to the case where the value of the variables is k (k greater than 2).

To this end, the gradient response technique k-RR was proposed [53], which randomly responds to queries between multiple values of the variable. In contrast, work [50, 51] focused on the mean calculation under LDP. The main idea is to map the ith tuple in d-dimensional dataset containing n tuples to the tuple with a binary variable, according to a certain probability distribution. Nevertheless, the literature [50, 51, 54] is not applicable to the high-dimensional data, since it

takes extremely high cost to deal with the high-dimensional data. To this end, harmony mean was proposed in [39], which reduces the communication cost through sampling, simplifies the previous work, improves data utility, and reduces errors. Another method k-Subset [52] was proposed, which considered the multiple output values after data perturbation, that is, the output is a set. As a result, k-Subset reduces the error and improves the accuracy of data publication.

The latest research [55] proposed RescueDP and designed real-time crowdsourced data publication framework for social networks, which dynamically groups regions with small statistics and adds Laplace noise to each group, eliminating the effect of perturbation error. Moreover, it applied Kalman filter to guarantee the data utility. Based on RescueDP, it further proposed an enhanced RescueDP scheme with neural networks, to get the optimal grouping strategy and finally improve RescueDP.

The data publication under local differential privacy is introduced above, and the aforementioned techniques are compared in detail in Table 1.

2.2 Machine learning

The main idea of LDP-based machine learning is that users locally perturb parameter updates of machine learning on their vehicle data using LDP and the server gets the global parameter updates via collecting local parameter updates from users, which is shown in Fig. 4.

Regression analysis based on LDP is another research topic. Regression analysis is a commonly used data classification method in machine learning, which determines the quantitative relationship of two or more attributes in the input datasets [56–60]. Regression analysis consists of two kinds of functions: One is the prediction function; the other is the objective function, or the risk function. It will leak the prediction function and the data information in the datasets when publishing the weight vector. To protect such data privacy in machine learning, a variety of work has applied LDP to regression analysis.

Smith et al. [20] proposed to solve the weight vector by minimizing the risk function, thereafter add Laplacian noise to the weight vector, and finally calculate the prediction function with the help of the weight vector containing noise. However, due to the correlations between the inputs of regression analysis (i.e., datasets) and the outputs (i.e., prediction function), it is extremely expensive to compute the sensitivity of the weight vector, resulting in low prediction accuracy. In contrast, work [21, 61–63] proposed a perturbation objective function method, which adds noise to the mean of the n objective functions in the datasets. Based on the perturbed objective function, it calculates the weight vector via minimizing its error. However, the scale of the noise is still determined by the sensitivity of the weight vector. So, the cost of calculating the sensitivity of the weight vector is still high. Moreover, the method is only applicable to objective functions with strong constraints, convex function, and dual differentiable.

To address the problems above, FM (functional mechanism) [64] was proposed, which achieved the regression analysis while meeting local differential privacy. FM first perturbed the sum of objective functions corresponding to each data tuple in the

Table 1 Comparisons of work about data publication under LDP

Studies	Advantages	Disadvantages	Computation cost	Communication cost	Budget allocation
[17]	Small publishing error; high data utility	The parameter setting of Bloom Filter should be considered	High; additional regression calculations	$O(h)$	Null
[18]	Sampling technology reduces communication costs	Query accuracy is not stable	Medium; extra encoding matches the string	$O(1)$	Null
[39]	Query accuracy is stable; small publishing error	Sampling technology reduces accuracy	Medium; extra coding and string matching overhead	$O(m)$	ϵ assigned to the sampling variable
[19]	Small publishing error; high data availability	Only for heavy hitter query	High; two rounds of string decoding cost	$O(h) + O(m)$	ϵ divided into two stages
[40, 41]	Reduce the influence of high dimension on precision	High communication and computing cost	High; additional graph model construction and reasoning cost	$O(d)$	Partition ϵ by attributes
[47]	Suitable for situations where variable values are unknown	The parameter setting of Bloom Filter should be considered	Low; only frequency statistics are involved	$O(1)$	Null
[48]	No need to predict the list of attribute candidate values	No need to predict the list of attribute candidate values	High; extra and string matching overhead	$O(h) + O(r)$	Partition ϵ by the number of substrings
[50, 51]	Small publishing error; high data availability	The spatial and temporal complexity is high; only applicable to low dimensional data; and individual data deviate from the original data to large extent	High; need to iterate through all the combinations of variables	$O(d)$	Divide ϵ equally based on the number of variables
[52]	Small matching error of input and output; high data publishing accuracy	Counting single-valued frequency results in additional errors	Medium; calculate the element frequency according to the set frequency	$O(1)$	Null

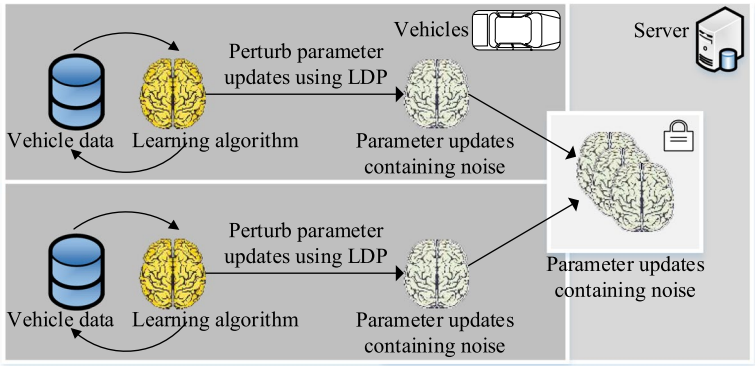


Fig. 4 Illustration of the application of LDP in machine learning

datasets and then obtained the weight vector by minimizing the target function. In this process, the noise scale is not determined by the sensitivity of the weight vector, but by expressing the objective function as a polynomial, avoiding the computation of the sensitivity of the weight vector.

However, the objective function of logistic regression cannot be expressed in polynomial in some cases. To this end, Zhang et al. [64] combined the truncated Taylor expansions to express the objective function as an approximate polynomial. In addition, Lei et al. [65] proposed another method that does not calculate the weight vector sensitivity. It uses the Laplace mechanism to generate a multidimensional noise histogram on the training dataset and synthesizes the training dataset using the histogram. Then, it calculates the weight vector. This method is only applicable to training datasets with small dimensions. Unfortunately, it is only applicable to training dataset with smaller dimensions, and the regression prediction accuracy significatively decreases with the dimensions.

In summary, the above three regression analysis methods suffer from different shortcomings. The regression analysis method based on Laplace mechanism has lower regression classification accuracy, and the noise error is relatively high [20]; the regression analysis method based on the perturbation mechanism is only applicable to specific objective functions [21, 61]; although the performance and efficiency of work [64–67] are better, it is only applicable to linear representation of the objective function. With regard to the complex objective functions in practical applications, such as Cox Regression, it achieves poor performance. So, how to design regression analysis meeting LDP for the complex objective functions is the future research direction.

Support vector machines (SVM), another data classification technique, typically deals with classification problems where the input space is nonlinear. It maps the input space to a feature space by using the nonlinear transformation in Kernel trick and then computes the support vector in the feature space using the solutions of the linear problem. Since the training dataset in the input space contains sensitive information, it will also disclose sensitive information when publishing weight vector directly. Smith et al. [20] proposed a method of SVM classification, PrivateSVM,

via combining with LDP, which uses Laplace noise to perturb the weight vector so that PrivateSVM satisfies ϵ-differential privacy. But, due to the high sensitivity of the weight vector, the scale of the noise is very large, which greatly reduces the classification accuracy. Different from PrivateSVM, Lei et al. [65] proposed ObjectiveSVM, a classification via adding noise to the objective function, which uses Laplace to generate random noise, adds noise to the risk function to obtain the perturbed objective function, and finally calculates the weight vector. Although the classification accuracy of ObjectiveSVM is higher than that of PrivateSVM, ObjectiveSVM is only applicable to objective functions with convex function characteristics. Furthermore, Chaudhuri et al. [68] adapted local differential privacy to the convex optimization problem. In contrast, Zhang et al. [69] applied local differential privacy to general optimization problem, such as logistic regression and support vector machine. However, study [68] is only applicable to the case where the absolute values of the primary and secondary derivatives of the objective function are constrained by constants, and the literature [69] achieves better performance. The latest work [70] proposed a certified defense that was applicable to both large networks and datasets, using differential privacy and cryptographically. Another work [71] used public key encryption to encrypt users' data and used differential privacy to protect data privacy when users' data are analyzed by cloud server. Study [72] allowed a trainer to train a Naive Bayes classifier over the dataset provided by multiple providers, providing differential privacy for each provider. The literature [73] proposed a numerically stable procedure for precise computation of SGM's Ryi Differential Privacy and proved a nearly tight closed-form bound.

In summary, the SVM classification based on Laplace has low classification accuracy and larger noise, while SVM classification based on perturbing objective functions is only applicable to specific objective functions. Therefore, how to design a perturbation mechanism with high precision and applications to multiple objective functions is the future research topic.

We have made a detailed comparison of the above work in Table 2.

2.3 Query processing

The query processing technologies based on local differential privacy mainly focus on how to respond to queries with less privacy budget and lower error [57, 74]. For example, linear queries in the interactive framework, batch processing of linear queries includes matrix mechanisms and low-rank mechanisms. The earliest matrix mechanism used the Query Plan and optimization strategies to refine batch queries [24–26, 75, 76]. The mechanism first generated the corresponding query plan A based on the weight vector (A is the full rank matrix), and then, the result of the query of the weight vector is estimated based on the noise result of the A request training datasets. In addition, the mechanism considers the association between linear queries. We can optimize the noise error generated by the mechanism through matrix decomposition [24, 25], while another method gives the lowest lower bound of the noise error through singular value analysis [26]. These matrix mechanisms require less noise, but are only suitable for small-scale datasets and

Table 2 Comparisons of work about machine learning under LDP

Studies	Advantages	Disadvantages
[20]	Satisfying differential privacy; preventing sensitive information from leaking	Low accuracy of regression classification; high noise error
[21]	Satisfying differential privacy; preventing sensitive information from leaking	The cost of calculate the sensitivity of weight vector is high; only applicable to objective function with strong constraint conditions, convex function, and dual differentiable
[64]	High classification accuracy; small noise error; generically	Only applicable to the objective function of linear representation
[61]	Satisfying differential privacy; preventing sensitive information from leaking	The cost of calculating the sensitivity of weight vector is high; only being applied to objective function with strong constraint conditions
[65]	High precision; low noise	Only applicable to specific objective function
[68]	Small noise error	Only applicable to the case where the absolute values of the primary and secondary derivatives of the objective function are constrained by constants
[69]	Output global optimal solution; an optimization method for selecting local optimal solutions for the exponential mechanism under local differential privacy	High computation overhead

query processing, and this mechanism typically produces sub-optimized query plans with less accuracy in the return results than those generated by adding noise directly to the training dataset. For this reason, it combined with the weight vector is a low-rank matrix and proposed a low-rank mechanism to improve the matrix mechanism. The mechanism first decomposes the weight vector into a full rank matrix and a full row rank matrix and finally determines the query sensitivity of the weight vector by the full row rank matrix. The low-rank mechanism avoids the disadvantages of the matrix mechanism producing suboptimal results and has a linear convergence speed.

Another kind of work [77–79, 85, 86] focused on process users' queries based on division methods. The study *kd*-standard [77] first combined the *kd*-tree and local differential privacy to divide the basic spatial data, according to the median number of the data space. The follow-up study [78] randomly segment the taxonomy tree based on generalization technology in a top-down manner, to publish set-valued numbers. In addition, Li et al. [79] designed an adaptive partitioning strategy to avoid the problem that the unit is too dense and too sparse. Peng et al. [80] designed a tree segmentation method that uses the nested tree to index multidimensional data to support range counting query, but it is susceptible to fan out of the tree. In work [81], the original isometric histogram was transferred using Haar wavelet, and it can accurately respond to long range count queries. Hay et al. [82] proposed to reorganize the isometric histogram using the m-ary tree to reduce query sensitivity; Xu et al. [83] used V-optimized histograms to merge similar neighboring buckets in the original isometric histogram. Acs et al. [84] adaptively segmented the original buckets from top to bottom combining the greedy halving strategy,

In summary, the matrix mechanism is prone to suboptimal results in practice; the low-rank mechanism only considers the correlation of the load matrix and does not take into account the relevance of the data itself. Therefore, how to design a general batch query processing mechanism from the actual relevance of the data itself is a future research direction. We have made a detailed comparison of the above work in Table 3.

2.4 Application of LDP system

Recommendation systems enable users to find required information from a large amount of data [87]. Since recommendation systems need to use a large amount of users' data for collaborative filtering, the privacy protection of data has attracted a wide range of attentions [88, 89]. Frank et al. [29] first introduced the local differential privacy method to the recommendation system. They assumed that the recommendation system is not trusted. The attacker can estimate the user's private information by analyzing the historical data of the recommendation system. So it is necessary to interfere with the input of the recommended system. When analyzing the relationship between projects, they first establish a project similarity covariance matrix and add Laplace noise to the matrix to implement interference, and then submit it to the recommendation system to implement the conventional recommendation algorithm. Another work [30] used a local differential privacy method in recommendation system based on social network data. They used the number

Table 3 Comparisons of work about query processing under LDP

Studies	Advantages	Disadvantages	Budget allocation	Computation complexity
[24]	Small size of noise; considering the correlation among linear queries	Easy to produce suboptimal results	Reasonable	$O(h)$
[77]	Applicable to data independent range count query	Unbalanced noise error and uniform hypothesis	Reasonable	$O(n \log n)$
[78]	Applicable to multidimensional data dependent range count query; high query accuracy	Without considering the semantic association of items in set-valued data.	Reasonable, adaptive allocate ϵ	$O(nI)$
[79]	Applicable to range count query; balanced noise error and uniform hypothesis	Without considering the sparseness of data distribution	Reasonable	$O(n)$
[25]	Avoid suboptimal results; linear convergence rate	Ignoring the correlation of the data	Reasonable	$O(I)$
[26]	Proving the low bound of the noise error	Producing suboptimal results; low query accuracy	Reasonable	$O(n)$
[80]	Applicable to multidimensional data dependent range count query; high query accuracy	Susceptible to fadeout of the tree	Reasonable	$O(n \log n^3)$
[81]	Applicable to longer range count queries; higher precision	Difficult to implement in practise	Reasonable	$O(n + k)$
[82]	Applicable to range queries of unit length and longer range queries	Only one-dimensional isometric histograms	Reasonable	$O(n + \log n)$
[83]	Applicable to longer range count queries; higher precision	Unable to balance reconstruction and noise errors; unable to handle outliers	Reasonable; using planning strategy to finding the best ϵ configuration	$O(n^2 k)$
[84]	Applicable to longer range count queries; higher efficiency	Only applicable to one-dimensional histogram conversion; poor scalability	Reasonable; ϵ being divided into two parts	$O(n^2)$

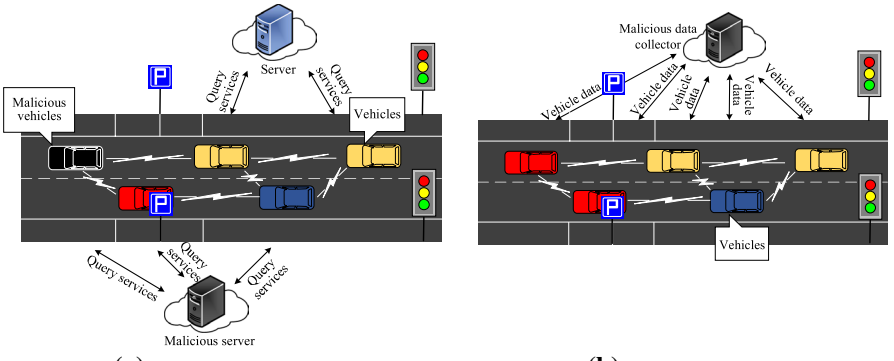
of neighbors of a specific node as a function of utility and adopted an exponential mechanism to randomly construct the edges in the graph. Finally, it achieved protection for all edges of the diagram. In addition, Zhu et al. [90] proposed a neighbor collaborative filtering algorithm based on local differential privacy protection so that the user's history cannot be inferred from the user's recommendation selection. Since the anonymization method cannot protect the privacy of network data, the concept of local differential privacy is introduced into the network trace analysis in work [31]. The local differential privacy method for network data statistical analysis is implemented on the PINQ platform. Furthermore, Chen et al. [91] used local differential privacy in public transportation systems to protect the transfer information of passengers. According to the characteristics of the data, the PrefixTree is used to represent the transportation information data, and Laplace noise is added to provide differential privacy protection for the users. Moreover, local differential privacy is also used in the situations where search engine companies publish high-frequency keywords, queries, and clicks [92].

3 Applications of LDP in IoV

We first introduce the attack models in IoV and then investigate the potential applications of LDP in several typical scenarios in IoV.

3.1 Attack model in IoV

- Malicious vehicle. In IoV, vehicles are connected with each other and require Navigation, air quality navigation, traffic information, etc., via sharing data with other vehicles. However, as shown in Fig. 5a, other vehicles may be malicious. On the one hand, they may deliberately disclose vehicles' data to advertisers, illegal organizations, etc. [93–95]. On the other hand, they may collude with others, send poisoning data, deliberately drop out during the process of completing



(a) Query services in IoV **(b)** Crowdsourcing in IoV.

Fig. 5 Illustration of frameworks of query services and crowdsourcing in IoV

IoV services, etc., aiming to disclosing other users' data privacy or paralyzing the IOV systems.

- Malicious server. In IoV, vehicles may have to send their data to the server to receive the corresponding services (cf. Fig. 5b). But the server may also be malicious and discloses vehicles' data to illegal organizations for commercial interest [96–98]. Specifically, malicious server may be passive or active adversary. The passive adversary is curious about users' data, but it honestly performs the protocols. In contrast, the active adversary may tamper the protocols or actively launch attacks to disclose users' data privacy.

To make matters worse, vehicles are vulnerable to serious attacks in the event of data privacy disclosure [99, 100]. Specifically, when the location data of a specific vehicle are breached by the malicious vehicles or the server, sensitive information such as the driver's life style, social relationship, and political beliefs, etc., can be easily revealed, thereby exposing the driver to spams, or even blackmails and physical violence [101], etc.

3.2 Query services in IoV

Query services will be an important kind of services in IoV, as recent years have witnessed the growing popularity of query services, e.g., Google Maps, HERE-CITY, Shopkick, MyTown, Foursquare, Gowalla, etc., and the revenue of query services in the world is USD 10.86 billion in 2014 and is projected to increase to USD 34.8 billion by 2020 according to Berg Insight [102]. Nevertheless, vehicles have to send their data to the malicious server whenever they request for services (cf. Fig. 5a), and more seriously, many kinds of sensitive information can be inferred from the disclosed information, e.g., religious activities, social relationships, living habit, etc. Therefore, it raises severe privacy concerns about vehicles data privacy in query services in IoV.

A number of techniques have been proposed to protect users' data privacy in academia. Specifically, rule protocol-based privacy-preserving techniques [103, 104] are first proposed. However, it is high overhead for server to obtain vehicles' authorization before utilizing vehicles' data in IoV, because of the unique features vehicles exhibit, e.g., high mobility, short connection times, etc. Encryption techniques [105–108] allow the server to collect and process the encrypted data of vehicles. Nevertheless, it is not applicable to large amounts of vehicles' data due to the extremely high computation overhead. Heuristic algorithms [109–112], e.g., dummy, *k*-anonymity, pseudonym, cloaking, m-invariance, etc., are quite lightweight compared to encryption techniques. But all these heuristic algorithms are vulnerable to the side information-based attacks. Furthermore, the evolvement of wireless communication technologies and mobile devices enables attackers to get access to various side information. In summary, LDP that is relatively lightweight and thoroughly considers side information of attackers is promised to protect vehicles' data privacy in query services in IoV.

3.3 Crowdsourcing in IoV

The ever-growing trend of mobile computation and communication technologies has enabled a variety of crowdsourcing applications in IoV. Crowdsourcing is a data collection paradigm where a large group of vehicles measure various data, e.g., traffic information, motor vehicle information, environmental information, etc., and then send their data to the data collector, which is shown in Fig. 5b. For example, systems, e.g., Uber, Waze, and Amazon Mechanical Turk, etc., provide various crowdsourcing platforms [113]. However, sharing data with the data collector may result in the serious privacy disclosure, as the collector may be malicious. What's more, more sensitive personal information can be inferred from the shared data.

Many existing work has focused on protecting data privacy in crowdsourcing. Work [114] introduces a third party, the cloud, that is responsible for storage and computation burden. Study [115] proposed a general feedback-based k-anonymity scheme to cloak users' data. The literature [116] utilized a random perturbation to mask users' data and employed the error-correcting codes to guarantee data utility. However, all these existing work ignores the side information of attackers and therefore is susceptible to side information-based attacks. Furthermore, these work is not applicable to the data privacy preservation in IoV, as density of vehicles is varying, and vehicles move with high speed and can only be connected in short time. In such a case, LDP is applicable to protect vehicle users' data privacy in crowdsourcing applications in IoV, as it thoroughly considers the available side information of attackers and is a lightweight privacy-preserving method.

4 Future research opportunities

In IoV, the complexity, diversity, and large-scale nature of vehicle data will add new data privacy risks. Therefore, we believe that LDP will face many new challenges, when it is applied to IoV applications.

- LDP for complex data types in IoV.
  The size of vehicle data is significantly growing in IoV, and the correlations among vehicle data make the vehicle data types more complex. At present, the research of LDP mainly focused on simple data types, e.g., frequency statistics or mean value statistics on set-valued data that only contains one attribute. However, in IoV, the structural characteristics of vehicle data make the global query sensitivity extremely high and bring in excessive noise. Moreover, it is still a big challenge to ensure the correlation among the original vehicle data according to the independently perturbed vehicle data in data publication, although the sensitivity is acceptable.
- LDP for various query and analysis tasks in IoV.
  At present, the existing work about LDP only investigated the privacy preservation in the two types of simple aggregate queries, i.e., counting queries of the discrete data and mean queries of the continuous data. Furthermore, the way of

data perturbation is generally depended on the types of queries. In IoV, a variety of query services are provided, and thus, LDP faces many new challenges.

- High-dimensional vehicle data publication based on LDP.

  In IoV, the set-valued data contains many attributes, and the existing studies about LDP will not work, as they only focused on simple data types. Thus, it is a new challenge for LDP to deal with the set-valued data with large dimensionality. To contrate, the set-valued data with large dimensionality will increase the size of vehicle data and decrease the signal-to-noise ratio. Moreover, communication overhead increases linearly or exponentially with the increasing data dimensions, which largely limits the applications of LDP in IoV.

  In summary, the high-dimensional data publication under LDP should consider the following three aspects: 1) How to measure the correlations among attributes within a certain privacy budget to reduce the dimensionality of vehicle data; 2) How to design an inference model that minimizes the approximation error of edge distribution and the joint distribution, to improve the data utility; 3) How to reduce the communication cost resulted from the high-dimensional vehicle data.
- Improvements in the LDP model for IoV.

  As the research about local differential privacy becomes more and more important, we need to consider the improvement in the local differential privacy model. In practice, the value of the privacy-preserving parameter ϵ still does not have a standard. Although the physical meaning of parameters in k-anonymity and l-diversity is intuitive, the privacy preservation provided by ϵ-differential privacy is relatively vague, which indicates that the problem is still up in the air.
- Considering correlations among vehicle data with LDP.

  Local differential privacy assumes that vehicle data are independent of each other, i.e., ignoring the correlations among vehicle data. However, in practice, vehicle data may be dependent. Therefore, considering the correlation among vehicle data in IoV will become a hot topic of LDP in future research.
- Combinations the LDP model with other techniques, e.g., machine learning, AI, and so on.

  The vast amount of available data and the advancement of machine learning methods and AI techniques are constantly reshaping research and industry in many aspects. In such a case, algorithms combining LDP, machine learning methods, and AI techniques are expected to provide potential solutions to, e.g., smart city, intelligent transportation, travel route recommendation, environment monitoring, air quality navigation, map navigation, etc., in IoV.

5 Conclusion

In this work, we have presented the overview and the thorough comparisons of the existing LDP techniques to get the readers well acquainted with LDP. Thereafter, we investigated the potential applications of LDP in securing IoV in detail. Last, we highlighted several future research directions of LDP in IoV, to bridge the gaps between LDP researches and the privacy preservation in IoV. It is the first work to

summarize and compare the existing LDP research work, and it also does an pioneering work toward the in-depth analysis of the potential applications of LDP in privacy preservation in IoV.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under Grant 61902060, part by Shanghai Sailing Program 19YF1402100, part by the Fundamental Research Funds for the Central Universities 2232019D3-51, 2722019PY052, and part by "Chenguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission.

References

1. 5.2 million vehicles will be on world's roads. Available at 5.2 Million Electric Vehicles Will Be On World's Roads By 2017. https://www.sogou.com/link?url=DSOYnZeCC_obPuaw6mFPWe4YBIs76YGqSEjsBEpORElYSw3r_2XEntVyzmvXbimxEqJPGVypItYeQ5gq3b3_vyra29P1rkT7iRI32N2X1uAxCie5vXk7hv0KmOA4bUvlT_PAMheO8n5Sj3w-WsG0SA. Accessed 6 Dec 2019
2. Corrigan-Gibbs H, Boneh D (2017) Prio: private, robust, and scalable computation of aggregate statistics. In: Proceeding of NSDI
3. The number of cars worldwide is set to double by 2040. https://www.weforum.org/agenda/2016/04/the-number-of-cars-worldwide-is-set-to-double-by-2040. Accessed 6 Dec 2019
4. Zhao P, Li J, Zeng F, Xiao F, Wang C, Jiang H, Jiang H (2018) ILLIA: enabling k-anonymity-based privacy preserving against location injection attacks in continuous LBS queries. IEEE Internet Things J 99:1–10
5. Jiang H, Zhao P, Wang C (2018) RobLoP: towards robust privacy preserving against location dependent attacks in continuous LBS queries. IEEE/ACM Trans Netw 26(2):1018–1032
6. DEF CON 23: Digital certificates key to mobile security, says researcher. https://www.computerweekly.com/news/4500251370/Def-Con-23-Digital-certificates-key-tomobile-security-says-researcher . Accessed 6 Dec 2019
7. Privacy leaks in IOV. http://www.afzhan.com/news/detail/74983.html. Accessed 6 Dec 2019
8. Zhang Y, Chen Q, Zhong S (2017) Efficient and privacy-preserving min and k th min computations in mobile sensing systems. IEEE Trans Dependable Secure Comput 14(1):9–21
9. Ben-Sasson E, Chiesa A, Genkin D (2013) SNARKs for c: verifying program executions succinctly and in zero knowledge. In: Advances in Cryptology-CRYPTO, pp 90–108
10. Home page of EU GDPR. http://www.eugdpr.org/. Accessed 6 Dec 2019
11. Araki T, Furukawa J, Lindell Y, Nof A, Ohara K (2016) High-throughput semi-honest secure three-party computation with an honest majority. In: Proceedin of ACM SIGSAC Conference on Computer and Communications Security
12. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, Ramage D, Segal A, Seth K (2017) Practical secure aggregation for privacy-preserving machine learning. In: Proceeding of ACM SIGSAC Conference on Computer and Communications Security
13. Nguyen T, Xiao T, Yang X, Hui Y, Shin SC, Shin H (2016) Collecting and analyzing data from smart device users with local differential privacy. arXiv preprint arXiv:1606.05053
14. Qin Z, Yu T, Yang Y, Khalil I, Xiao X, Ren K (2017) Generating synthetic decentralized social graphs with local differential privacy. In: Proceeding of ACM SIGSAC Conference on Computer and Communications Security, pp 425–438
15. This is what Apple's differential privacy means for ios 10. https://www.theverge.com/2016/6/17/11957782/apple-differential-privacy-ios-10-wwdc-2016. Accessed 6 Dec 2019
16. How Google tricks itself to protect Chrome user privacy. https://www.cnet.com/au/news/how-google-tricks-itself-to-protect-chrome-user-privacy/. Accessed 6 Dec 2019
17. Erlingsson, ulfar, Pihur V, Korolova A (2014) Rappor: randomized aggregatable privacy-preserving ordinal response. In: Proceeding of ACM SIGSAC conference on computer and communications security, pp 10–23

18. Bassily R, Smith A (2015) Local, private, efficient protocols for succinct histograms. In: Proceeding of ACM symposium on Theory of computing, pp 10–19
19. Qin Z, Yang Y, Yu T, Khalil I, Xiao X, Ren K (2016) Heavy hitter estimation over set-valued data with local differential privacy. In: Proceeding of ACM SIGSAC Conference on Computer and Communications Security, pp 192–203
20. Smith A (2011) Privacy-preserving statistical estimation with optimal convergence rates. In: Proceeding of ACM Symposium on Theory of Computing
21. Samet S (2015) Privacy-preserving logistic regression. J Adv Inf Technol 6(3):1–8
22. Shokri R, Shmatikov V (2015) Privacy-preserving deep learning. In: Proceeding of ACM SIGSAC Conference on Computer and Communications Security
23. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
24. Li C, Hay M, Rastogi V, Miklau G, McGregor A (2010) Optimizing linear counting queries under differential privacy. In: Proceeding of ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, pp 123–134
25. Yuan G, Zhang Z, Winslett M, Xiao X, Yang Y, Hao Z (2012) Low-rank mechanism: optimizing batch queries under differential privacy. In: Proceeding of VLDB Endowment, pp 1352–1363
26. Li C, Miklau G (2011) Efficient batch query answering under differential privacy. arXiv preprint arXiv:1103.1367
27. Wan S, Li X, Xue Y, Lin W, Xu X (2019) Efficient computation offloading for internet of vehicles in edge computing-assisted 5G networks. J Supercomput. https://doi.org/10.1007/s11227-019-03011-4
28. Xu X, Gu R, Dai F, Qi L, Wan S (2019) Multi-objective computation offloading for internet of vehicles in cloud-edge computing. Wirel Netw. https://doi.org/10.1007/s11276-019-02127-y
29. McSherry Frank, Mironov I (2009) Differentially private recommender systems: Building privacy into the netflix prize contenders. In: Proceeding of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 1–9
30. Machanavajjhala Ashwin A, Korolova, Sarma AD (2011) Personalized social recommendations: accurate or private. In: Proceeding of the VLDB Endowment, pp 440–450
31. Frank M, Mahajan R (2011) Differentially-private network trace analysis. ACM SIGCOMM Comput Commun Rev 41(4):123–134
32. Wan S, Zhao Y, Wang T, Gu Z, Abbasi QH, Choo K-KR (2019) Multi-dimensional data indexing and range query processing via voronoi diagram for internet of things. Future Gener Comput Syst 91:382–391
33. Xu X, Xue Y, Qi L, Yuan Y, Zhang X, Umer T, Wan S (2019) An edge computing-enabled computation offloading method with privacy preservation for internet of connected vehicles. Future Gener Comput Syst 96:89–100
34. Purohit S, Smith W, Chappell A, West P, Lee B, Stephan E, Fox P (2016) Effective tooling for linked data publishing in scientific research. In: Proceeding of IEEE Tenth International Conference on Semantic Computing, pp 24–31
35. Ye Q, Meng X, Zhu M, Huo Z (2018) Survey on differential privacy. Ruan Jian Xue Bao/Journal of Software, no. 7
36. Xu X, He C, Xu Z et al (2019) Joint optimization of offloading utility and privacy for edge computing enabled IoT. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2019.2944007
37. Gao Z, Xuan H-Z, Zhang H, Wan S, Choo K-KR (2019) Adaptive fusion and category-level dictionary learning model for multi-view human action recognition. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2019.2911669
38. Niari AK, Berangi R, Fathy M (2018) ECCN: an extended CCN architecture to improve data access in vehicular content-centric network. J Supercomput 74(1):205–221
39. NT, Xiao T, Yang X, Hui Y, Shin SC, SJ (2016) Collecting and analyzing data from smart device users with local differential privacy. arXiv preprint arXiv:1606.05053
40. Ren X, Yu CM, Yu W, Yang S, Yang X, Mccann JA, Yu PS (2018) LoPub: High-dimensional crowdsourced data publication with local differential privacy. IEEE Trans Inf Forensics Secur 13(9):2151–2166
41. Ren X, Yu C, Yu W, Yang S, Yang X, McCann J (2016) High-dimensional crowdsourced data distribution estimation with local privacy. In: Proceeding of IEEE International Conference on Computer and Information Technology (CIT), pp 226–233

42. Chen C, Liu L, Qiu T, Yang K, Gong F, Song H (2018) ASGR: an artificial spider-web-based geographic routing in heterogeneous vehicular networks. IEEE Trans Intell Trans Syst 20(5):1604–1620
43. Chen C, Liu L, Qiu T, Ren Z, Hu J, Ti F (2018) Driver's intention identification and risk evaluation at intersections in the internet of vehicles. IEEE Internet Things J 5(3):1575–1587
44. Wan S, Gu Z, Ni Q (2019) Cognitive computing and wireless communications on the edge for healthcare service robots. Comput Commun 149:99–106
45. Zhang R, Xie P, Wang C, Liu G, Wan S (2019) Classifying transportation mode and speed from trajectory data via deep multi-scale learning. Comput Netw 162:106861
46. Park JJ (2018) Fusion algorithms and high-performance applications for vehicular cloud computing. J Supercomput 74(3):995–1000
47. Ye M, Barg A (2018) Optimal schemes for discrete distribution estimation under locally differential privacy. IEEE Trans Inf Theory 64(8):5662–5676
48. Fanti G, Pihur V, Erlingsson l (2016) Building a rappor with the unknown: privacy-preserving learning of associations and data dictionaries. Proc Priv Enhanc Technol 3:41–61
49. Liu J, Wang W, Li D, Wan S, Liu H (2019) Role of gifts in decision making: an endowment effect incentive mechanism for offloading in the IoV. IEEE Internet Things J 6(4):6933–6951
50. MIJ, Duchi, John C, Wainwright MJ (2013) Local privacy, data processing inequalities, and statistical minimax rates. arXiv preprint arXiv:1302.3203
51. Duchi JC, Jordan MI, Wainwright MJ (2014) Privacy aware learning. J ACM 61(6):1–57
52. Ye M, Barg A (2018) Optimal schemes for discrete distribution estimation under locally differential privacy. IEEE Trans Inf Theory 64:5662–5676
53. Kairouz P, Oh S, Viswanath P (2014) Extremal mechanisms for local differential privacy. Adv Neural Inf Process Syst 4:2879–2887
54. Balouchzahi N-M, Fathy M, Akbari A (2016) An efficient infrastructure based service discovery in vehicular networks using P2P structures. J Supercomput 72(3):1013–1034
55. Wang Q, Zhang Y, Lu X, Wang Z, Qin Z, Ren K (2018) Real-time and spatio-temporal crowd-sourced social network data publishing with differential privacy. IEEE Trans Dependable Secure Comput 15(4):591–606
56. Jagielski M, Oprea A, Biggio B et al (2018) Manipulating machine learning: Poisoning attacks and countermeasures for regression learning. IEEE Symp Secur Priv (SP): 19–35
57. Zhang X, Meng X (2014) Differential privacy protection for data publishing and analysis. J Comput 4:927–949
58. Chen C, Pei Q, Li X (2016) A GTS allocation scheme to improve multiple-access performance in vehicular sensor networks. IEEE Trans Veh Technol 65(3):1549–1563
59. Jinna H, Qiu T, Atiquzzaman M et al (2018) CVCG: Cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular ad-hoc networks. IEEE Trans Mob Comput 18(12):2811–2828
60. Khosravi MR, Samadi S (2019) Reliable data aggregation in internet of ViSAR vehicles using chained dual-phase adaptive interpolation and data embedding. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2019.2952284
61. Chaudhuri K, Monteleoni C, Sarwate AD (2011) Differentially private empirical risk minimization. J Mach Learn Res 12:1069–1109
62. Khosravi MR, Basri H, Rostami H, Samadi S (2018) Distributed random cooperation for VBF-based routing in high-speed dense underwater acoustic sensor networks. J Supercomput 74(11):6184–6200
63. Chen C, Hu J, Qiu T, Atiquzzaman M, Ren Z (2019) CVCG: cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular ad-hoc networks. IEEE Trans Mob Comput 18(12):2811–2828
64. Zhang J, Zhang Z, Xiao X, Yang Y, Winslett M (2012) Functional mechanism: regression analysis under differential privacy. In: Proceeding of VLDB Endowment, pp 1364–1375
65. Lei J (2011) Differentially private m-estimators. Adv Neural Inf Proc Syst 2011:361–369
66. Khosravi MR, Basri H, Rostami H (2018) Efficient routing for dense UWSNs with high-speed mobile nodes using spherical divisions. J Supercomput 74(2):696–716
67. Wu W-C (2017) A secret push messaging service in VANET clouds. J Supercomput 73(7):3085–3097
68. Chaudhuri K, Monteleoni C, Sarwate AD (2011) Differentially private empirical risk minimization. J Mach Learn Res JMLR 12(2):1069

69. Zhang J, Xiao X, Yang Y, Zhang Z, Winslett M (2013) Privgene:differentially private model fitting using genetic algorithms. In: Proceeding of ACM SIGMOD International Conference on Management of Data, pp 665–676
70. Lecuyer M, Atlidakis V, Geambasu R et al (2019) Certified robustness to adversarial examples with differential privacy. IEEE Symp Secur Priv (SP) 2019:656–672
71. Li P, Li T, Ye H, Li J, Chen X, Xiang Y (2018) Privacy-preserving machine learning with multiple data providers. Future Gener Comput Syst 87:341–350
72. Li T, Li J, Liu Z, Li P, Jia C (2019) Differentially private Naive Bayes learning over multiple data sources. Inf Sci 444:89–104
73. Mironov I, Talwar K, Zhang L (2019) Rényi differential privacy of the sampled Gaussian mechanism. arXiv:1908.10530 [cs.LG]
74. Wang J, Cai Z, Ai C, Yang D, Gao H, Cheng X (2016) Differentially private k-anonymity: achieving query privacy in location-based services. In: Proceeding of International Conference on Identification, Information and Knowledge in the Internet of Things, pp 475–480
75. Zhuang Y, Fong S, Yuan M, Sung Y, Cho K, Wong RK (2017) Predicting the next turn at road junction from big traffic data. J Supercomput 73(7):3128–3148
76. Karimi V, Mohseni R, Samadi S (2019) Ofdm waveform design based on mutual information for cognitive radar applications. J Supercomput 75(5):2518–2534
77. Cormode G, Procopiuc C, Srivastava D, Shen E, Yu T (2012) Differentially private spatial decompositions. IEEE 28th Int Conf Data Eng 41(4):20–31
78. Chen R, Mohammed N, Fung BCM, Desai BC, Xiong L (2012) Publishing setvalued data via differential privacy. VLDB 4(4):1087–1098
79. Li N, Yang W, Qardaji W (2013) Differentially private grids for geospatial data. In: Proceeding of IEEE International Conference on Data Engineering, pp 757–768
80. Peng S, Yang Y, Zhang Z et al (2012) DP-tree: indexing multi-dimensional data under differential privacy. In: Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data. ACM, pp 864–864
81. Xiao X, Wang G, Gehrke J (2011) Differential privacy via wavelet transforms. IEEE Trans Knowl Data Eng 23(8):1200–1214
82. Hay M, Rastogi V, Miklau G, Dan S (2010) Boosting the accuracy of differentially private histograms through consistency. Proc VLDB Endow 3(1–2):1021–1032
83. Xu J, Zhang Z, Xiao X, Yang Y, Yu G (2012) Differentially private histogram publication. In: IEEE International Conference on Data Engineering, pp 32–43
84. Acs G, Castelluccia C, Chen R (2013) Differentially private histogram publishing through lossy compression. In: Proceeding of IEEE International Conference on Data Mining, pp 1–10
85. Karimi V, Mohseni R (2019) Intelligent target spectrum estimation based on ofdm signals for cognitive radar applications. J Intell Fuzzy Syst 36(3):2557–2569
86. Nkenyereye L, Park Y, Rhee K-H (2018) Secure vehicle traffic data dissemination and analysis protocol in vehicular cloud computing. J Supercomput 74(3):1024–1044
87. Ping X, Tianqing Z, Xiaofeng W (2014) A survey on differential privacy and applications. Chin J Comput 37(1):101–122
88. Patel AA, Dharwa JN (2017) An integrated hybrid recommendation model using graph database. In: Proceeding of International Conference on ICT in Business Industry and Government, pp 1–5
89. Xiong P, Zhu T, Wang X (2014) Differential privacy protection and its application. J Comput 37(1):101–122
90. Zhu T, Li G, Ren Y, Zhou W, Xiong P (2013) Differential privacy for neighborhood-based collaborative filtering. In: Proceeding of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp 752–759
91. Chen R, Fung B, Desai BC, Sossou NM (2012) Differentially private transit data publication: a case study on the montreal transportation system. In: Proceeding of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 213–221
92. Gotz M, Machanavajjhala A, Wang G, Xiao X, Gehrke J (2012) Publishing search logsa comparative study of privacy guarantees. IEEE Trans Knowl Data Eng 24(3):520–532
93. Lindell Y, Pinkas B, Smart NP, Yanai A (2015) Efficient constant round multi-party computation combining BMR and SPDZ. In: Proceeding of Annual Cryptology Conference
94. Corrigan-Gibbs H, Wolinsky DI, Ford B (2013) Proactively accountable anonymous messaging in verdict. In: Proceeding of USENIX Security Symposium

95. Elahi T, Danezis G, Goldberg I (2014) Privex: private collection of traffic statistics for anonymous communication networks. In: Proceeding of ACM SIGSAC Conference on Computer and Communications Security
96. Goryczka S, Xiong L (2017) A comprehensive comparison of multiparty secure additions with differential privacy. IEEE Trans Dependable Secure Comput 14(5):463–477
97. Rastogi V, Nath S (2010) Differentially private aggregation of distributed time-series with transformation and encryption. In: Proceeding of ACM SIGMOD International Conference on Management of Data
98. Qi L, Zhang X, Li S, Wan S, Wen Y, Gong W (2019) Spatial-temporal data-driven service recommendation with privacy-preservation. Inf Sci. https://doi.org/10.1016/j.ins.2019.11.021
99. Wang X, Liu Y, Shi Z, Lu X (2015) A privacy-preserving fuzzy localization scheme with csi fingerprint. In: IEEE Global Communications Conference
100. Higuchi T, Martin P, Chakraborty S, Srivastava M (2015) Anonycast: privacy-preserving location distribution for anonymous crowd tracking systems. In: ACM International Joint Conference on Pervasive and Ubiquitous Computing
101. Primault V, Mokhtar S. Ben, Brunie L (2015) Privacy-preserving publication of mobility data with high utility. In: IEEE International Conference on Distributed Computing Systems
102. Berg Insight: LBS Revenue to Grow to 34.8 billion in 2020. http://www.gpsbusinessnews.com/Berg-Insight-LBS-Revenue-to-Grow-to-34-8-billion-in-2020-a5627.html. Accessed 6 Dec 2019
103. Ietf "geographic location/privacy (georriv) working group". https://www.ietf.org/. Accessed 6 Dec 2019
104. W3C, platform for privacy preferences (P3P) project. https://www.w3.org/P3P/. Accessed 6 Dec 2019
105. Manickam P, Shankar K, Perumal E, Ilayaraja M, Sathesh Kumar K (2019) Secure data transmission through reliable vehicles in VANET using optimal lightweight cryptography. Cybersecurity and secure information systems. Springer, Cham, pp 193–204
106. Lu Z, Wang Q, Chen X, Qu G, Lyu Y, Liu Z (2019) Leap: a lightweight encryption and authentication protocol for in-vehicle communications. IEEE Intell Transp Syst Conf 2019:1158–1164
107. Brousmiche K. Leo, Durand A, Heno T, Poulain C, Dalmieres A, Hamida EB (2018) Hybrid cryptographic protocol for secure vehicle data sharing over a consortium blockchain. In: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), pp 1281–1286
108. Kang J, Lin D, Jiang W, Bertino E (2018) Highly efficient randomized authentication in vanets. Pervasive Mob Comput 44:31–44
109. Liu L, Chen C, Qiu T, Zhang M, Li S, Zhou B (2018) A data dissemination scheme based on clustering and probabilistic broadcasting in vanets. Veh Commun 13:78–88
110. Xu D, He X, Xu B, Wang Y, Zhang C, Li F (2012) L2P2: location-aware location privacy protection for location-based services. IEEE Infocom. https://doi.org/10.1109/INFCOM.2012.6195577
111. Lv N, Chen C, Qiu T, Sangaiah AK (2018) Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in sar images. IEEE Trans Ind Inf 14(12):5530–5538
112. Karimi V, Mohseni R, Samadi S (2019) Adaptive OFDM waveform design for cognitive radar in signal-dependent clutter. IEEE Syst J 99:1–12
113. Chen Y, Li B, Zhang Q (2016) Incentivizing crowdsourcing systems with network effects. In: IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications. IEEE, pp 1–9
114. Zhuo G, Jia Q, Guo L, Li M, Li P (2016) Privacy-preserving verifiable data aggregation and analysis for cloud-assisted mobile crowdsourcing. In: IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications. IEEE, pp 1–9
115. Wu S, Wang X, Wang S, Zhang Z, Tung AK (2014) K-anonymity for crowdsourcing database. IEEE Trans Knowl Data Eng 26(9):2207–2221
116. Varshney LR, Vempaty A, Varshney PK (2014) Assuring privacy and reliability in crowdsourcing with coding. In: Information Theory and Applications Workshop (ITA), 2014. IEEE, pp 1–6

Affiliations

Ping Zhao[1] · Guanglin Zhang[1] · Shaohua Wan[2] · Gaoyang Liu[3] · Tariq Umer[4]

Ping Zhao
pingzhao2014ph@gmail.com

Guanglin Zhang
glzhang1981@gmail.com

Gaoyang Liu
husterlgy@gmail.com

Tariq Umer
t_umer@yahoo.com

1 College of Information Science and Technology, Donghua University, Shanghai 201620, China

2 School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China

3 School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China

4 Department of Computer Science, COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan