Multiple Strategies Differential Privacy on Sparse Tensor Factorization for Network Traffic Analysis in 5G

Wang, Jin; Han, Hui ; Li, Hao; He, Shiming ; Sharma, Pradip Kumar ; Chen, Lydia

DOI
10.1109/TII.2021.3082576

Publication date
2021

Document Version
Final published version

Published in
IEEE Transactions on Industrial Informatics

Multiple Strategies Differential Privacy on Sparse Tensor Factorization for Network Traffic Analysis in 5G

Jin Wang, *Senior Member, IEEE*, Hui Han, Hao Li, Shiming He, Pradip Kumar Sharma, *Senior Member, IEEE*, and Lydia Chen, *Senior Member, IEEE*

***Abstract*—Due to high capacity and fast transmission speed, 5G plays a key role in modern electronic infrastructure. Meanwhile, sparse tensor factorization (STF) is a useful tool for dimension reduction to analyze high-order, high-dimension, and sparse tensor (HOHDST) data, which is transmitted on 5G Internet-of-things (IoT). Hence, HOHDST data relies on STF to obtain complete data and discover rules for real time and accurate analysis. From another view of computation and data security, the current STF solution seeks to improve the computational efficiency but neglects privacy security of the IoT data, e.g., data analysis for network traffic monitor system. To overcome these problems, this article proposes a multiple-strategies differential privacy framework on STF (MDPSTF) for HOHDST network traffic data analysis. MDPSTF comprises three differential privacy (DP) mechanisms, i.e., $\varepsilon-$ DP, concentrated DP, and local DP. Furthermore, the theoretical proof of privacy bound is presented. Hence, MDPSTF can provide general data protection for HOHDST network traffic data with high-security promise. We conduct experiments on two real network traffic datasets ($Abilene$ and $GÈANT$). The experimental results show that MDPSTF has high universality on the various degrees of privacy protection demands and high recovery accuracy for the HOHDST network traffic data.**

***Index Terms*—Differential privacy framework, multiple-strategies privacy protection, network traffic analysis, sparse tensor factorization.**

Manuscript received March 13, 2021; revised April 29, 2021; accepted May 15, 2021. Date of publication May 21, 2021; date of current version December 6, 2021. Paper no. TII-21-1173. *(Corresponding author: Hao Li.)*

Jin Wang, Hui Han, and Shiming He are with the School of Computer and Communication Engineering, ChangSha University of Science and Technology, Hunan 410004, China (e-mail: jinwang@csust.edu.cn; h_han@stu.csust.edu.cn; smhe_cs@csust.edu.cn).

Hao Li and Lydia Chen are with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 Mekelweg, The Netherlands (e-mail: H.Li-9@tudelft.nl; lydiaychen@ieee.org).

Pradip Kumar Sharma is with the Department of Computing Science, University of Aberdeen, AB24 3FX Aberdeen, U.K. (e-mail: pradip.sharma@abdn.ac.uk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2021.3082576.

Digital Object Identifier 10.1109/TII.2021.3082576

I. Introduction

With the advent of 5G technology, high speed and low latency 5G networks bring in huger transmission capacity than 4G, which can support the emerging application technologies relying on Internet-of-Things (IoT), e.g., virtual reality, augmented reality, and wearable devices linked with mobile phones. Meanwhile, thousands of connected devices on the 5G IoT network will generate a tremendous amount of data, and the network traffic data, which can capture the moving data across a network, usually be used to prevent the network jam and paralysis [1]. Hence, 5G IoT relies heavily on accurate and real-time network traffic analysis to maintain a steady, fluent, and high-speed network environment [2]. Due to distributed deployment of sensor devices, network traffic data has the form of temporal-spatial characteristic, and, own to downtime and some other crash problems for some devices linked by 5G IoT, the network traffic data presents the form of high-order, high-dimension, and sparse tensor (HOHDST) [3]. For HOHDST data, sparse tensor factorization (STF) can draw the low-rank feature of each order from HOHDST data simultaneously. Thus, STF plays a key role in the analysis of network traffic data [4], which can obtain complete data to measure more valuable information. Practitioners focus on mining the STF algorithm to conduct an accurate recovery for HOHDST network traffic data, while neglecting the data privacy in the transmission and analysis process.

Network traffic data involves the information about the location coordinate and network flow, and attackers can infer the individual private information. Meanwhile, there are three layers that consider how to address the network security problems, i.e., physical network security, technical network security, and administrative network security. It is designed to prevent unauthorized personnel from obtaining physical access to the network, viruses and other malicious software to manipulate the network, and protect the data.

But none of them can deal with the hidden danger that keeps high-security promise for the traffic data on data-level. Differential privacy (DP) is a strong privacy protection mechanism in data level, which can guarantee that anyone cannot make inference about the individual's private information by adding a kind of noise; Meanwhile, DP can ensure the availability of data [5].

DP can also provide mathematical provable privacy protection against common privacy attacks, i.e., linkage and reconstruction attacks [6].

More recently, due to the appetency to provide a privacy guarantee [7], DP naturally becomes a preferred tool for data privacy protection in the training process of machine learning (ML) models [8]–[10]. DP can provide privacy protection for the classic dimension reduction model, i.e., matrix factorization (MF) [11], [12] and the application fields have been extended to recommender systems and social networks [13], [14]. However, the DP for MF model (DPMF) can only handle the privacy protection for two-order matrix data and the DPMF cannot provide the same privacy security promise for third or higher order tensor data; Meanwhile, the current DPMF models can apply $\varepsilon-$ DP and local DP (LDP) individually, but do not have a general privacy protection framework.

There are several works that explore the DP mechanism for tensor factorization (TF). Wang and Anandkumar [15] proposed a DP framework for the tensor level method and the tensor data are symmetric and dense one. However, this method only considers the $\varepsilon-$ DP mechanism, and the memory overhead is not scalable. For distributed, large-scale, symmetric and dense tensor data, Imtiaz and Sarwate presented a distributed DP framework for orthogonal TF and the factorization process for the low-rank matrix of each order involves the singular value decomposition (SVD) for a huge symmetric matrix. Because the privacy proof on dense data does not apply to the asymmetric and HOHDST data generated in real-world applications, the abovementioned DP framework cannot provide a security promise, e.g., network traffic data and medical health data [16].

For HOHDST data, researchers focus on how to improve the computational efficiency and prediction accuracy for missing values, especially in traffic network data. Li *et al.* [17] proposed a high performance computation framework on GPU for sparse MF (SMF). Li *et al.* [17] presented stochastic gradient descent (SGD) based algorithm for STF to reduce the computational complexity [18] and in the face of accurate HOHDST network traffic data recovery, there have been many accurate processing technologies. Xie *et al.* [19] widely explored the low-rank structure of two-order matrix and high-order tensor generated from network traffic data and proposed an accurate sparse matrix and tensor recovery framework based on SMF and STF, respectively. The abovementioned methods do not solve the problem of privacy protection for HOHDST data, just improve the computational efficiency and recovery accuracy.

Recently, in order to overcome the abovementioned limitations and meet the requirements of real-time and accurate recovery of HOHDST network traffic data, Ma *et al.* [20] proposed a framework of canonical polyadic (CP) factorization under concentrated DP (CDP) protection for electronic health records. This method can compute the distributed STF, which can impute local missing diagnosis information and avoid direct data sharing, meanwhile, this model does not leak the local patient diagnosis information. Nie *et al.* [21] presented an STF analysis framework for IoT data generated from cloud and edge under $\varepsilon-$ DP protection. However, those methods do not solve the general privacy protection problem for HOHDST network traffic data, which means that those methods cannot ensemble various DP, i.e., $\varepsilon-$ DP, CDP, and LDP, and provide a general privacy protection framework for HOHDST network traffic data. To overcome the abovementioned limitations and meet the requirements of accurate recovery of HOHDST network traffic data under restrict and mathematically provable privacy protection, the following challenges should be solved.

1) Different DP has different system framework requirements on the STF.
2) LDP applications that do not apply to third-order tensors.
3) Achieve an optimal tradeoff between the degree of privacy protection and the precision of data recovery.

To handle these problems, A HOHDST network traffic data recovery framework under multiple-strategies differential privacy protection on STF (MDPSTF) is proposed. The HOHDST network traffic data recovery relies on the CP factorization, the factorization process of CP meets the requirements of real-time analysis in the recovery process of HOHDST network traffic data and only involves the low-rank matrix of each order. The main contributions of this article are summarized as follows.

1) This is the first work to realize an ensemble MDPSTF framework. The framework MDPSTF combines three DP mechanisms, i.e., $\varepsilon-$ DP, CDP, and LDP, which can recover the HOHDST network traffic data and provide privacy protection.
2) Laplace mechanism and Gaussian mechanism add corresponding noise to the third-party trusted server, respectively, and realize the privacy protection following DP and zero-concentrated differential privacy (zCDP) definition.
3) The LDP protects the data collection source, i.e., adding noise to the whole tensor data on the user's device, which can protect the data from the attack of the untrusted third-party server.
4) The theoretical proof of privacy under various mechanisms is given. Combined with the experimental results, the data recovery accuracy of each mechanism under different privacy degrees is analyzed.

The rest of this article is organized as follows. The problem formulation and preliminaries are presented in Section II. Section III introduces the framework of MDPSTF for the recovery problem of HOHDST network traffic data under general privacy protection. The experimental results are given in Section IV. Finally, Section V concludes this article.

II. Problem Formulation and Preliminaries

A. Sparse Tensor Factorization

The main notation, including scalars, vectors, matrices, and tensors, as well as the slice format of tensors are listed in Table I. In the following sections, for simplicity, the STF refers to the CP factorization.

Definition 1 (Tensor approximation): Given a N-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_n \times \cdots \times I_N}$, the low-rank tensor approximation problem can be formalized as $\mathcal{X} = \widetilde{\mathcal{X}} + \zeta$, where $\widetilde{\mathcal{X}}$ is the low-rank

TABLE I
DEFINITION OF SYMBOLS

Symbol	Definition
I_n	The size of row in the nth factor matrix;
R	The rank of CP factorization;
$\mathcal{X}$	N-order tensor $\in \mathbb{R}^{I_1 \times \cdots \times I_n \times \cdots \times I_N}$;
$\mathbf{A}^{(n)}$	The n-th factor matrix $\in \mathbb{R}^{I_n \times R}$;
$x_{(i_1,\cdots,i_n,\cdots,i_N)}$	The $(i_1,\cdots,i_n,\cdots,i_N)$-th entry of $\mathcal{X}$;
$a^{(n)}_{i_n,:}$	i_n-th row vector $\in \mathbb{R}^R$ of $\mathbf{A}^{(n)}$;
$\bar{a}^{(n)}_{:,j_n}$	j_n-th column vector $\in \mathbb{R}^{I_n}$ of $\mathbf{A}^{(n)}$;
$a^{(n)}_{i_n,j_n}$	The (i_n,j_n)-th element of $\mathbf{A}^{(n)}$;
$\| \bullet \|_2$	L_2 norm;
Ω	Index $(i_1,\cdots,i_n,\cdots,i_N)$ of a tensor;
$\mathbf{X}_{:,i_n,:}$	The i_n-th lateral slice matrix $\in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N}$ of tensor $\mathcal{X}$;
$\circ$	Outer production of vectors;

tensor and ζ is noise data. The optimization problem can be formalized as $\arg\min_{\widetilde{\mathcal{X}}} \|\mathcal{X} - \widetilde{\mathcal{X}}\|_2^2$.

In this article, we only consider the accurate recovery and privacy protection problem of HOHDST network traffic data. Hence, besides tensor approximation, the definition of sparse tensor factorization should be presented.

Definition 2 (Sparse tensor factorization): The approximation tensor can be obtained by sum of outer product rank-1 tensors $\widetilde{\mathcal{X}} = \sum_{r=1}^{R} \lambda_r \bar{a}^{(1)}_{:,r} \circ \cdots \circ \bar{a}^{(n)}_{:,r} \circ \cdots \circ \bar{a}^{(N)}_{:,r}$. The constant λ_r, $r \in \{1,\ldots,R\}$ can be omitted. Factor matrices $\mathbf{A}^{(n)}$, $n \in \{1,\ldots,N\}$ are obtained by following the sparsity pattern of the sparse tensor $\mathcal{X}$.

Definition 3 (μ strongly-convex): For any $x_1, x_2 \in \mathbb{R}^r$, if there is a constant $\mu > 0$, $f(x_1) \geq f(x_2) + \bigtriangledown f(x_2)(x_1 - x_2)^T + \frac{1}{2}\mu\|x_1 - x_2\|_2^2$, then the continuously differentiable function $f(x)$ satisfies the μ strongly convexity.

Definition 4 (L-Lipschitz continuity): Suppose there is a continuously differentiable and L-smooth function $f(x)$, $x \in \mathbb{R}^r$. Take any two gradient values $\nabla f(x_1), \nabla f(x_2) \in \mathbb{R}^r$ of the two variables $x_1, x_2 \in \mathbb{R}^r$, and if these two gradient values satisfy $\| \nabla f(x_1) - \nabla f(x_2) \|_2 \leq L \| x_1 - x_2 \|_2$, then the gradient $\nabla f(x)$ is L-Lipschitz continuous for any $x \in \mathbb{R}^r$.

Definition 5 (Stochastic gradient descent): The optimization loss function $f(w)$ is μ strongly-convex and L-Lipschitz continuity as:

$$\arg\min_{w \in \mathbb{R}^R} f(w) = \underbrace{L\left(w \middle| y_i, x_i, w\right)}_{Loss\ Function} + \underbrace{\lambda_w R(w)}_{Regularization} \tag{1}$$

where $y_i \in \mathbb{R}^1$, $x_i \in \mathbb{R}^R$, $i \in \{1,\ldots,N\}$, $w \in \mathbb{R}^R$, and $L(w|y_i, x_i, w) + \lambda_w R(w) = \sum_{i=1}^{N} L_i(w|y_i, x_i, w) + \lambda_w R_i(w)$. The original optimization model needs gradient, which should select all the samples $\{x_i | i \in \{1,\ldots,N\}\}$ from the dataset Ω and the gradient descent is presented as $w \leftarrow w - \gamma \frac{\partial f_\Omega(w)}{\partial w}$, where $\frac{\partial f_\Omega(w)}{\partial w} = \frac{1}{N}\sum_{i=1}^{N} \frac{\partial (L_i(w) + \lambda_w R_i(w))}{\partial w}$. An M entries set Ψ is randomly selected from the set Ω, and the SGD [22] is presented as $w \leftarrow w - \gamma \frac{\partial f_\Psi(w)}{\partial w}$, where $\frac{\partial f_\Psi(w)}{\partial w} = \frac{1}{M}\sum_{i \in \Psi} \frac{\partial (L_i(w) + \lambda_w R_i(w))}{\partial w}$ and γ is the learning rate.

In our work, the HOHDST network traffic data can be formalized as a third-order tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ [23], [24]. Thus, the value N in this article is set as three and we substitute $\mathbf{A}^n$, $n \in \{1,2,3\}$ as $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$, respectively. The objective function $f(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is represented as

$$f(\mathbf{A},\mathbf{B},\mathbf{C}) = \underbrace{\frac{1}{2}\sum_{(i,j,k)\in\Omega}\left(x_{ijk} - \sum_{r=1}^{R} a_{ir}b_{jr}c_{kr}\right)^2}_{\text{Loss Function}} + \underbrace{\frac{1}{2}\lambda_1\|\mathbf{A}\|_2^2 + \frac{1}{2}\lambda_2\|\mathbf{B}\|_2^2 + \frac{1}{2}\lambda_3\|\mathbf{C}\|_2^2}_{\text{Regularization}}. \tag{2}$$

B. Differential Privacy

Some preliminary knowledge about DP mechanism, i.e., $\varepsilon-$ DP, CDP, and LDP [7]–[14], [20], [21], are presented.

Definition 6 (Neighbor datasets): Supposed that for any two datasets on the database $\mathcal{D}$, deemed as $\widetilde{\mathcal{D}}$ and $\widehat{\mathcal{D}}$, and the two datasets have the same structure. Entry record difference is denoted as $\widetilde{\mathcal{D}} \triangle \widehat{\mathcal{D}}$, and $|\widetilde{\mathcal{D}} \triangle \widehat{\mathcal{D}}|$ means the number of entries record. If $|\widetilde{\mathcal{D}} \triangle \widehat{\mathcal{D}}| = 1$, the two datasets $\widetilde{\mathcal{D}}$ and $\widehat{\mathcal{D}}$ are called neighbor datasets.

Definition 7 ($\varepsilon-$ differential privacy ($\varepsilon-$ DP)): Suppose any two neighbor datasets $\widetilde{\mathcal{D}}, \widehat{\mathcal{D}} \in \mathcal{D}$, and the output set is $S \subset \mathbb{R}$. If a random mechanism $f : \mathcal{D} \to \mathbb{R}$ satisfies $\Pr[f(\widetilde{\mathcal{D}}) \in S] \leq e^{\varepsilon}\Pr[f(\widehat{\mathcal{D}}) \in S] + \delta$. Then, f satisfies $(\varepsilon - \delta)-$ DP. The ε here is called the privacy. δ can relax and guarantee on a very small probability. The larger the privacy, the higher privacy protection, but the worse the data availability [8]–[13].

Definition 8 (Global sensitivity): Suppose there is a query function $f : \mathcal{D} \to \mathbb{R}^d$. If any pair of neighbor datasets $\widetilde{\mathcal{D}}$ and $\widehat{\mathcal{D}}$, their global sensitivity is given by $GS_{f(\mathcal{D})} = max\|f(\widetilde{\mathcal{D}}) - f(\widehat{\mathcal{D}})\|_1$, where $\|f(\widetilde{\mathcal{D}}) - f(\widehat{\mathcal{D}})\|_1$ is the Manhattan distance between $f(\widetilde{\mathcal{D}})$ and $f(\widehat{\mathcal{D}})$. L_1 norm is available. Global sensitivity has nothing to do with datasets, only with query results [8], [13].

Definition 9 (Laplace mechanism): The Laplace mechanism derives $\varepsilon - DP$ is often used in numerical output functions. It is basically to add a noise tensor of the same size as the original tensor data, where the noise element conforms to the Laplace distribution. Defined a function $f : D \to R^d$. The probability density function of tensor Laplace distribution : $f(x_{i_1,i_2,\ldots,i_N}) = \frac{exp(-\frac{|x_{i_1,i_2,\ldots,i_N}|}{\lambda})}{2\lambda}$, when λ is the noise parameter [21].

Definition 10 (zero-concentrated differential privacy): Supposed that there is a random mechanism M, If any two neighbor databases $\widetilde{\mathcal{D}}$ and $\widehat{\mathcal{D}}$, and $\widetilde{\mathcal{D}}$ differs from $\widehat{\mathcal{D}}$ by at most one entry. If the $\alpha-$R*ényi* divergence between the distributions of $M(\widetilde{\mathcal{D}})$ and $M(\widehat{\mathcal{D}})$ of these two databases with respect to the distribution satisfies $D_\alpha(M(\widetilde{\mathcal{D}})||M(\widehat{\mathcal{D}})) \triangleq \frac{1}{\alpha-1}log(\mathbb{E}[e^{(\alpha-1)Z}]) \leq \xi + \rho\alpha$, where Z is the privacy loss random variable denoted as $Privloss(M(\widetilde{\mathcal{D}})||M(\widehat{\mathcal{D}}))$. Then, the random mechanism M is ρ-zCDP in $\alpha \in (1, \infty)$. Define the privacy loss random variable

Z between $\widetilde{D}$ and $\widehat{D}$. Z is distributed according to $f(\widetilde{D})$, where the function $f : \widetilde{\mathcal{D}} \to \mathbb{R}$ by $f(d) = log\frac{\mathbb{P}(\widetilde{D}=d))}{\mathbb{P}(\widehat{D}=d)}$ [20].

Definition 11 (Local differential privacy): Given a privacy mechanism M and its domain and range are defined as $Dom(M)$ and $Ran(M)$, respectively. If the mechanism M satisfied the following inequality on any two records x and x' $(x, x' \subset Dom(M))$ that obtained the same output $x^*(x^* \subseteq Ran(M)$, then M satisfies the $\varepsilon - LDP$ as $\Pr[M(x) = x^*] \leq e^{\varepsilon}\Pr[M(x') = x^*]$. The randomized response is the primary disturbance mechanism of LDP. This mechanism mainly consists of two steps: perturbation statistics and correction [7], [14].

Current DP mechanisms have been applied successfully in dense TF communities. However, due to the lack of data recovery and general privacy protection strategies, those solutions cannot make privacy protection for HOHDST network traffic data. In Section III, the general privacy protection framework MDPSTF for HOHDST network traffic data is presented.

Both DP and CDP work in a supposedly trusted third-party server. The difference lies in that DP accepts the entire tensor dataset transmitted by the user, while CDP accepts matrix data from each user and aggregates it into the tensor dataset. The next step: decomposing the tensor dataset through CP factorization, and three factor matrices can be obtained. As for the factor matrix $\mathbf{A}$, it contains a large amount of user information. So Laplace mechanism and Gaussian mechanism are, respectively, used to add noise in this article. LDP is based on the assumption that the third-party server is not trusted. In order to protect the data at the source, the randomized response mechanism is used to add noise on each user before the server aggregates the data.

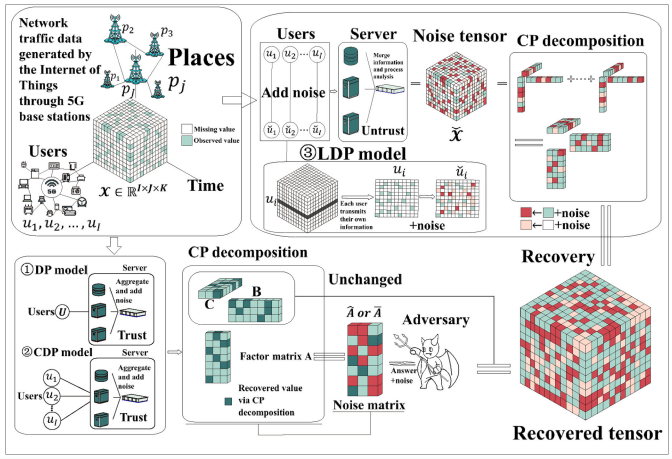

Fig. 1. Framework of MDPSTF.

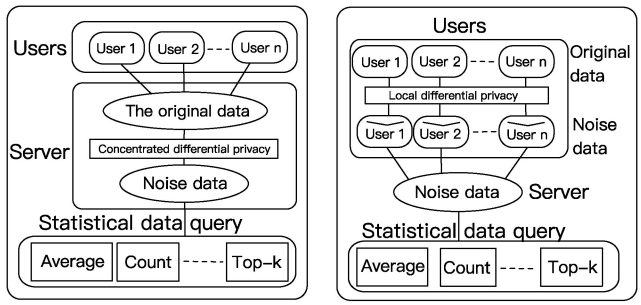

Fig. 2. Data processing framework for CDP and LDP.

III. Multiple Strategies Differential Privacy for Sparse Tensor Factorization

The research background of this article is the data recovery of HOHDST network traffic data generated from 5G network. Due to equipment failure, signal missing, and other crash problems, the network traffic data are commonly sparse and the transmission process of the HOHDST network traffic data cannot leak privacy information. Therefore, in this article, we conduct the MDPSTF framework to make data recovery under privacy protection for HOHDST network traffic data.

As Fig. 1 shows that MDPSTF can protect the HOHDST network traffic data from two views: 1) the obtained factor matrices and 2) the original HOHDST data. DP is used to ensure the privacy security of users and to achieve a balance between recovery accuracy and privacy protection. To consider the accuracy of the data recovery, the HOHDST network traffic data are first modeled as a tensor, i.e., a third-order tensor $\mathcal{X} \in \mathbb{R}^{I\times J\times K}$. The tensor $\mathcal{X}$ is generated from a dynamic 5G network, which comprises user and location coordinate. By combining the low-rank representation of CP factorization and its factor matrix with various DP strategies, a general framework MDPSTF is proposed, which can recover the HOHDST network traffic data under privacy protection.

TABLE II
Definition of Symbols

Symbol	Definition
$\widetilde{\mathcal{X}}$	The recovered tensor by CP factorization;
$\hat{\mathcal{X}}$	The recovered tensor after $\varepsilon-$DP;
$\bar{\mathcal{X}}$	The recovered tensor after CDP;
$\check{\mathcal{X}}^*$	The recovered tensor after LDP;
α	The learning rate of SGD;
ρ	Privacy budget;
η	The noise matrix $\in \mathbb{R}^{I\times R}$;
μ	The mean of a Gaussian distribution;
σ	The variance of a Gaussian distribution;
L	The Lipschitz Constant;
u	A Bernoulli random variable;
R	Rank of CP factorization (CP-ranks);
$\mathcal{X}^n$	Each of these elements represents information about a user.

The privacy strategies can be divided into two categories: 1) after CP factorization, the servers will aggregate information and add noise processing for $\varepsilon-$DP and CDP; 2) In LDP, the server will aggregate the information and CP factorization after users add noise information. In Fig. 1, u_i denotes the information matrix of the ith user, and $\check{u}_i$ represents the noise information matrix of the ith user. Table II records the variables which are used in MDPSTF.

A. $\varepsilon-$ DP Mechanism

$\hat{\mathbf{A}}$ is the noise factor matrix which means adding the noise $\hat{\mathbf{A}} = \mathbf{A} + \eta$, and then the disturbed objective function is $f(\hat{\mathbf{A}}, \mathbf{B}, \mathbf{C}) = f(\mathbf{A} + \eta, \mathbf{B}, \mathbf{C})$.

Theorem 1: Let $\mathcal{X}$ be the range of user network traffic data values. If each noise value η is independent and randomly selected from the density function of Laplace distribution, where $\triangle = \mathcal{X}_{max} - \mathcal{X}_{min}$, the derived factor matrix $\hat{\mathbf{A}}$ is deduced to satisfy the $\varepsilon-$DP.

Proof: We assume that there is only one record difference between the two HOHDST tensor $\left\{ \widetilde{\mathcal{X}} = \{ \widetilde{x}_{(1,1,1)}, \ldots, \widetilde{x}_{(i,j,k)}, \ldots, \widetilde{x}_{(I,J,K)} \}, \hat{\mathcal{X}} = \{ \hat{x}_{(1,1,1)}, \ldots, \hat{x}_{(\hat{i},\hat{j},\hat{k})}, \ldots, \hat{x}_{(I,J,K)} \} \right\}$, (i,j,k) and $(\hat{i},\hat{j},\hat{k}) \subset \Omega$. N and $\hat{N}$ are, respectively, expressed as the noise matrix of $\mathcal{X}$ and $\hat{\mathcal{X}}$. We observe that $f(\hat{\mathbf{A}}, \mathbf{B}, \mathbf{C})$ are differentiable anywhere. After obtaining the gradient, we set the equality $\frac{\partial f(\mathbf{A},\mathbf{B},\mathbf{C})}{\partial a_{i,r}} = \frac{\partial f(\hat{\mathbf{A}},\mathbf{B},\mathbf{C})}{\partial \hat{a}_{i,r}}$, $\{a_{i,r}, \hat{a}_{i,r}\}$ are (i,r)-element of the factor matrix $\{\mathbf{A}, \hat{\mathbf{A}}\}$, respectively. The abovementioned equation is expanded as: $\eta_{i,r} - \sum_{(i,j,k)\in\Omega} \big(\widetilde{x}_{(i,j,k)} - \sum_{r=1}^{R} a_{i,r} b_{j,r} c_{k,r}\big)(-b_{j,r} c_{k,r}) = \hat{\eta}_{i,r} - \sum_{(i,j,k)\in\Omega} \big(\hat{x}_{(i,j,k)} - \sum_{r=1}^{R} a_{i,r} b_{j,r} c_{k,r}\big)(-b_{j,r} c_{k,r})$. If $(i,j,k) \neq (\hat{i},\hat{j},\hat{k})$, $\eta_{i,r} - \hat{\eta}_{i,r} = 0$. Else $(i,j,k) = (\hat{i},\hat{j},\hat{k})$, $\eta_{i,r} - \hat{\eta}_{i,r} = (\sum_{r=1}^{R} a_{i,r}\ b_{j,r}^2\ c_{k,r}^2)\ (\widetilde{x}_{(i,j,k)} - \hat{x}_{(i,j,k)})$. Then, the global sensitivity is defined as $GS(a_{p,q}) = \big(\mathcal{X}_{\max} - \mathcal{X}_{\min}\big) \times \big(\max_{(i,j,k)} \| \sum_{l=p+1}^{I} \sum_{l=1}^{p-1} \sum_{r=q+1}^{R} \sum_{r=1}^{q-1} a_{l,r} b_{j,r}^2 c_{k,r}^2 \|_F\big)$. Because the factor matrices $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ are randomly sampled from the uniform distribution [0,1], so $\max_{(i,j,k)} \| \sum_{l=p+1}^{I} \sum_{l=1}^{p-1} \sum_{r=q+1}^{R} \sum_{r=1}^{q-1} a_{l,r} b_{j,r}^2 c_{k,r}^2 \|_F \leq 1$. Hence, we can infer $GS(a_{\hat{i},r}) \leq \triangle$, then, $\|\eta_{\hat{i},r} - \hat{\eta}_{\hat{i},r}\|_F \leq GS(a_{\hat{i},r}) \leq \triangle$. The density function defined as: $p(\eta_{i,r}) \propto e^{-\frac{\varepsilon \|\eta_{i,r}\|}{2\triangle}}$. Finally, $\frac{\Pr[\widetilde{\mathcal{X}}]}{\Pr[\hat{\mathcal{X}}]} = \frac{\prod_{i\in\{1,\ldots,I\}} p(\eta_{i,r})}{\prod_{i\in\{1,\ldots,I\}} p(\hat{\eta}_{i,r})} = e^{-\frac{\varepsilon(\sum_{i\in\{1,\ldots,I\}} \|\eta_{i,r}\| - \sum_{i\in\{1,\ldots,I\}} \|\hat{\eta}_{i,r}\|)}{2\triangle}} = e^{-\frac{\varepsilon(\|\eta_{\hat{i},r}\| - \|\hat{\eta}_{\hat{i},r}\|)}{2\triangle}} \leq e^{-\varepsilon}$. □

B. CDP Mechanism

The privacy degree of CDP is tighter than that of ε-DP, which provides a more explicit analysis for many calculations that retain privacy. Dwork and Rothblum came up with this CDP [25], Bun and Steinke propose an alternative formulation of CDP called "zero-CDP (zCDP)" [26]. The realization mechanism of zCDP is the same as the Gaussian mechanism of $(\varepsilon - \delta)$-DP. The general Gaussian distribution as follows $p(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$, where μ represents the mean and σ represents the variance.

Theorem 2 (Gaussian mechanism): Supposed that there is a random algorithm M. Let $\varepsilon \in (0,1)$ be a arbitrary variable. For $c^2 > 2In(\frac{1.25}{\delta})$, the Gaussian mechanism with the parameter $\sigma \geqslant c \triangle_2 (M)/\varepsilon$, and adding noise scaled to $\mathcal{N}(0, \sigma^2)$ to each component of algorithm M output, is $(\varepsilon - \delta)$-DP.

The following propositions will be used together in the proof of zCDP [25].

Proposition 1: The relation between Gaussian mechanism and zCDP. If for all $x, x' \in \mathcal{X}^n$ differing in a single entry, and define a function $q : \mathcal{X}^n \to \mathbb{R}$ be a sensitivity$-\triangle$. We have $|q(x) - q(x')| \leq \triangle$. Suppose there have a Gaussian mechanism $M : \mathcal{X}^n \to \mathbb{R}$ releases a sample from $\mathcal{N}(q(x), \sigma^2)$ on the input x, then M satisfies $(\frac{\triangle^2}{2\sigma^2}) - zCDP$. $(\frac{\triangle^2}{2\sigma^2})$ can also be written as ρ, namely $\sigma = \frac{\triangle}{\sqrt{2\rho}}$.

Proposition 2 (): The transformation between DP and zCDP. A randomized mechanism M is a $\rho - zCDP$. If any δ with $\varepsilon = \rho + 2\sqrt{\rho ln(\frac{1}{\delta})}$, then M is a $(\varepsilon' - \delta)$-DP. On the contrary, the mechanism M satisfies $(\varepsilon - \delta)$-DP. If $\rho \approx \frac{\varepsilon^2}{4ln(\frac{1}{\delta})}$, it suffices to satisfy $\rho - zCDP$.

The theory of zCDP can be proved by the combination of the abovementioned two key propositions [20], [25], [26].

Proposition 3 (Serial composition): Let $M : \mathcal{D}^n \to \mathcal{Y}$ and $M' : \mathcal{D}^n \to \mathcal{Z}$ are any two random algorithms. If M is $\rho - zCDP$ and M' is $\rho' - zCDP$. In addition to define a new random algorithm $M'' : \mathcal{D}^n \to \mathcal{Y} \times \mathcal{Z}$ by $M'' = (M, M')$. Then, M'' is $(\rho + \rho') - zCDP$.

Proposition 4 (Parallel composition): Supposed that a mechanism A consist of a sequence of T adaptive mechanisms, $\{A_1, \ldots, A_T\}$, where each $A_t : \prod_{j=1}^{iter-1} \mathcal{O}_j \times \mathcal{D}_t \to \mathcal{O}_{iter}$ and A_t satisfies $\rho_t - zCDP$. Let $\{\mathcal{D}_1, \ldots, \mathcal{D}_T\}$ be a randomized partition of the input $\mathcal{D}$. The mechanism $A(\mathcal{D}) = (A_1(\mathcal{D})_1, \ldots, A_T(\mathcal{D})_T)$ satisfies $\frac{1}{T} \sum_{t=1}^{T} \rho_t - zCDP$.

Theorem 3: Set a random mechanism as M, and we assume that the added Gaussian noise parameter setting conforms to Theorem 2, then, the random mechanism satisfies $(\varepsilon - \delta)$-DP. Let's assume that the total privacy budget each iteration input of a factor matrix is $\rho_a = \frac{\varepsilon^2}{4Tln(\frac{1}{\delta})}$, where T is the total number of achieve converges.

Proof: First, Theorem 2 shows that the random mechanism M satisfies $(\varepsilon - \delta)$-DP

$$f(\mathbf{A}^{[t]}) = \frac{1}{2} \sum_{(i,j,k)\in\Omega} \left(x_{(i,j,k)} - \bar{x}_{(i,j,k)} \right)^2 + \frac{\lambda_1}{2} \|\mathbf{A}^{[t]}\|_2^2 \quad (3)$$

where $\bar{x}_{(i,j,k)} = \sum_{r=1}^{R} a_{i,r} b_{j,r} c_{k,r}$, and the L_2-sensitivity of this problem is $\triangle_2 = \|2TL\lambda_1\|$ (L is the Lipschitz constant), $\triangle_2 = 2TL\lambda_1 = 2T\| \sum_{j,k:(i,j,k)} \sum_{r=1}^{R} (b_{jr}^2 c_{kr}^2) \|_2 \lambda_1$. The base zCDP parameter is ρ_a. According to the Proposition 3, $\bar{\mathbf{A}}$ costs $T\rho_a$ in total. According to the Proposition 4, then the total privacy budget cost of all N users is $\frac{\sum_{n=1}^{N} T\rho_n}{N} = \frac{T(\rho_1 + \rho_2 + \cdots + \rho_N)}{N} = T\rho_a$. Finally, according to Proposition 1 and 2, the random mechanism M satisfies $(\varepsilon - \delta)$-DP that can be converted to zCDP by setting as $\rho = T\rho_a$, where $\rho_a = \frac{\varepsilon^2}{4Tln(\frac{1}{\delta})}$. □

Algorithm 1 consists of three modules. The module ① is the original network traffic data tensor $\mathcal{X}$ decomposed by CP factorization to obtain three factor matrices, namely **A**, **B**, and **C**.

Algorithm 1: DPSTF($\mathcal{X}$,R,ε,λ,ρ).

Input: A third-order tensor $\mathcal{X}$, CP-ranks R, privacy ε and parameters λ, ρ.

Output: ① Factor matrices $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, and $\mathbf{C} \in \mathbb{R}^{K\times R}$, ② Noise factor matrix $\hat{\mathbf{A}} \in \mathbb{R}^{I\times R}$ and recovered tensor $\hat{\mathcal{X}}$, ③ Noise factor matrix $\bar{\mathbf{A}} \in \mathbb{R}^{I\times R}$ and recovered tensor $\bar{\mathcal{X}}$,

①Initialize $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, $\mathbf{C} \in \mathbb{R}^{K\times R}$;
While convergence threshold is not reached or $t < iter_{max}$ do;
t=t+1;
Randomly sample element x_{ijk} from $\mathcal{X}$;
$x^*_{ijk} = a_r \circ b_r \circ c_r$;
for $(i,j,k) \in \Omega$ **do**
  Compute the gradient $\frac{\partial f}{\partial a_{i,r}}, \frac{\partial f}{\partial b_{j,r}}, \frac{\partial f}{\partial c_{k,r}}$;
  Update $\widetilde{x}_{i,j,k} \leftarrow \sum_{r=1}^{R} a_{i,r} b_{j,r} c_{k,r}$;
end;
If the convergence condition meets or reaches the maximum number of iterative steps, then outputs the factor matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$.;

② **for** $i = 1, ..., I; r = 1, ...R$ **do**
  Generate noise that corresponds to the Laplace distribution $f(a_{ir}) = \frac{exp(-\frac{|a_{ir}|}{\lambda})}{2\lambda}$;
  Update $\hat{a}_{ir} \leftarrow a_{ir} + f(a_{ir})$;
end;
$\hat{\mathcal{X}} = \sum_{r=1}^{R} \hat{a}_{:,r} \circ b_{:,r} \circ c_{:,r}$;
Return $\hat{\mathcal{X}}$ and $\hat{A}$;

③ **for** $i = \{1, \cdots, I\}$, $r = \{1, \cdots, R\}$ **do**
  Update $\bar{A} \leftarrow A$+Gaussian noise matrix $\mathcal{N}(0, \frac{\triangle_2^2}{2\rho})$;
end;
$\bar{\mathcal{X}} = \sum_{r=1}^{R} \bar{a}_r \circ b_r \circ c_r$;
Return $\bar{\mathcal{X}}$ and $\bar{A}$;

Then, the corresponding noise is added by module ② or module ③. Module ② is to add noise matrix with the same size as factor matrix **A** under DP strategy based on Laplace mechanism. After updating the elements in factor matrix **A**, Algorithm 1 returns a recovery tensor $\hat{\mathcal{X}}$. Module ③ add a noise matrix of the same size as factor matrix **A** based on Gaussian mechanism under zCDP strategy, and Algorithm 1 returns a recovery tensor $\bar{\mathcal{X}}$ after updating the elements in factor matrix **A**.

C. LDP Mechanism

Because the HOHDST network traffic data presents the tensor form, the DP communities try to extend the DP mechanism from MF to STF. So far, only DP and zCDP have been applied to TF just for the privacy protection on dense and symmetric tensor. Thus, the application domain is limited. The theoretic base of both approaches is the assumption that the third-party servers are trustworthy and cannot provide LDP protection mechanism, LDP can address this assumption well. We want to combine the STF for HOHDST network traffic data recovery and privacy protection with LDP. Fig. 2 illustrates that LDP is used to collect the individual user information and used for data publishing or querying.

LDP inherits the combined characteristics of the zCDP and can extend it furthermore. LDP mainly uses the randomized response mechanism to the input noise disturbance. Thus, LDP can resist the privacy attack from the source from the third-party data collector, which is assumed to be untrusted. CDP has serial and parallel propositions (Propositions 3 and 4). The serial combination can allocate privacy budget under different iteration times of the algorithm. The parallel combination ensures that private datasets on disjoint confidential subsets satisfy differential privacy respectively. From the definition, the DP is defined on the neighboring datasets, while the LDP is specified on the two records. However, the form of a privacy guarantee does not change. Therefore, the LDP also continues the serial composition of DP.

Theorem 4 (Serial composition on the LDP): Supposed that a method consists of m independent random functions $\mathcal{M} = \{\mathcal{M}_1, \ldots, \mathcal{M}_m, \ldots, \mathcal{M}_M\}$, and each function $\mathcal{M}_m$ satisfies ε_m-LDP. Then, the method $\mathcal{M}$ satisfies $\sum_{m=1}^{M} \varepsilon_m$-LDP [27].

Lemma 1: Let u_i and u'_i be any two nonprivate information matrix as inputs. u and u' is the Bernoulli variable generated by the given input u_i and u'_i, respectively. $\breve{u}_i$ is the output of the algorithm. Then, each submission of the user's information matrix satisfies $\varepsilon/I-$LDP. So the noisy information matrix submitted by all the users satisfies the ε-LDP.

Proof: Sample a Bernoulli variable u such that $\Pr[u = 1] = \frac{u_{j,k}(e^{\varepsilon}-1)+e^{\varepsilon}+1}{2e^{\varepsilon}+2}$. Assume that $\frac{\Pr[\breve{u}_i|u_i]}{\Pr[\breve{u}_i|u'_i]} = \frac{Pr[u=1|u_i]}{Pr[u'=1|u'_i]} \leq \frac{\max_{u_i} Pr[u=1|u_i]}{\min_{u'_i} Pr[u'=1|u'_i]} = \frac{\max_{u_i}((u_i)_{j,k}(e^{\varepsilon/I}-1)+e^{\varepsilon/I}+1)}{\min_{u'_i}((u'_i)_{j,k}(e^{\varepsilon/I}-1)+e^{\varepsilon/I}+1)} = \frac{2e^{\varepsilon/I}}{e^{\varepsilon/I}+1} \leq e^{\varepsilon/I}$. It can prove that the noise matrix of each user to submit satisfies $\varepsilon/I-$LDP. By Theorem 4, the total noise matrix meets $\varepsilon-$LDP. □

Algorithm 2 adds noise to the user information matrix $u_i, i = 1, \ldots, I$ through the Laplace mechanism or Gaussian mechanism. This disturbs the user's network traffic data at the source and avoids the attack of untrusted third-party servers. The server receives the disturbing user information matrix $\breve{u}_i, i = 1, \ldots, I$, aggregates them into a third-order noise tensor $\breve{\mathcal{X}}$, and decomposes it through CP factorization and recovers it into $\breve{\mathcal{X}}^*$.

IV. Experimental Results and Analysis

The experiments are carried on the public traffic trace data *Abilene* [23] and the pan-European research backbone network *GÈANT* [24] to evaluate the performance of MDPSTF framework. *Abilene* network consists of 12 nodes; Thus, it will generate 144 original-source pairs. Thus, *Abilene* network contained 144 users, 288 locations, and 168 time points. So *Abilene* contains a network traffic tensor data with a size of $\mathbb{R}^{144\times288\times168}$, and the *GÈANT* is a large-scale, symmetric, and dense tensor. It records monitoring data vary in 112 days and containing a network traffic tensor data with a size of $\mathbb{R}^{272\times96\times112}$. Mean

Algorithm 2: LDPSTF($\mathcal{X}$,R,ε,λ,ρ).

Input: A third-order tensor $\mathcal{X}$, CP-ranks R, privacy ε and parameters λ, ρ.

Output: Factor matrices $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, and $\mathbf{C} \in \mathbb{R}^{K\times R}$, noise user information matrix $\check{u}_i \in \mathbb{R}^{J\times K}, i = 1, ..., I$ and recovered tensor $\check{\mathcal{X}}^*$.

for $i = 1, ..., I$ **do**
  Initialize $(\check{u}_i)_{j,k} \in \{0\}^{J\times K}$;
  Randomly generate a probability rnd_i;
  $p = \frac{e^{\varepsilon}}{e^{\varepsilon}+n+1}$;
if $rnd_i < p$: *The user transmits* u_i *directly;*
then
  Generate noise by the Laplace or Gaussian distribution:
  for $j = \{1, \cdots, J\} \in \Omega, k = \{1, \cdots, K\} \in \Omega$ **do**
    $f(u_{j,k}) = \frac{exp(-\frac{|u_{j,k}|}{\lambda})}{2\lambda}$ or $p(u_{j,k}|\varepsilon) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{u_{j,k}^2}{2\varepsilon}}$;
  Update $\check{u}_{j,k} \leftarrow u_{j,k} + f(u_{j,k})$ *or* $p(u_{j,k}|\varepsilon)$;
end;
Aggregate all user information matrix into a third-order noise tensor $\check{\mathcal{X}} \in \mathbb{R}^{I\times J\times K}$;
Initialize $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, $\mathbf{C} \in \mathbb{R}^{K\times R}$;
While convergence threshold is not reached or $t < iter_{max}$ do;
t=t+1;
Randomly sample element $\check{x}_{ijk}$ from $\check{\mathcal{X}}$;
for $(i, j, k) \in \Omega$ **do**
  Compute the gradient $\frac{\partial f}{\partial \check{a}_{i,r}}, \frac{\partial f}{\partial \check{b}_{j,r}}, \frac{\partial f}{\partial \check{c}_{k,r}}$;
  Update $\check{x}_{i,j,k} \leftarrow \sum_{r=1}^{R} \check{a}_{i,r}\check{b}_{j,r}\check{c}_{k,r}$;
end;
If the convergence condition meets or reaches the maximum number of iterative steps, then outputs the factor matrices $\check{\mathbf{A}}, \check{\mathbf{B}}, \check{\mathbf{C}}$;
$\check{\mathcal{X}}^* = \sum_{r=1}^{R} \check{a}_r \circ \check{b}_r \circ \check{c}_r$ for $\check{\mathbf{A}}, \check{\mathbf{B}}, \check{\mathbf{C}}$;
Return $\check{\mathcal{X}}^*$.

squared error (MSE), root-mean-squared error (RMSE), and fitness error are defined as a follows.

Definition 12 (Mean squared error): The formula for MSE is $MSE(\mathcal{X}) = \frac{\sum_{(i,j,k)\in\Gamma}(x_{ijk}-\tilde{x}_{(i,j,k)})^2}{|\Gamma|}$; where $\tilde{x}_{(i,j,k)} = \sum_{r=1}^{R} a_{i,r}b_{j,r}c_{k,r}$ and Γ is the test set.

Definition 13 (Root-mean-squared error): The formula for RMSE is $RMSE(\mathcal{X}) = \sqrt{\frac{\sum_{(i,j,k)\in\Gamma}(x_{ijk}-\tilde{x}_{(i,j,k)})^2}{|\Gamma|}}$ where $\tilde{x}_{(i,j,k)} = \sum_{r=1}^{R} a_{i,r}b_{j,r}c_{k,r}$ and Γ is the test set.

Definition 14 (Fitness error): Fitness error can measuring the recovery error of entries in the tensor [28]. The formula is defined as $\frac{\sqrt{\sum_{(i,j,k)\in\Omega}(x_{(i,j,k)}-\tilde{x}_{(i,j,k)})^2}}{\sqrt{\sum_{(i,j,k)\in\Omega} x_{(i,j,k)}^2}}$.

TABLE III
DESCRIBE OF MODELS

Model	Describe
CP-SGD	CP factorization recovery based on SGD algorithm;
CP-DP	Laplace mechanism of DP based on CP factorization with SGD algorithm;
CP-zCDP	Gaussian mechanism of zCDP based on CP factorization with SGD algorithm;
DPFacT[20]	DPFacT is a distributed tensor factorization method, which enhances differential privacy;
CP-ALS[29]	CP factorization recovery based on Alternating Least Squares optimization algorithm;

First, the performance of CP factorization is tested on the two datasets, i.e., *Abilene*, and *GÈANT*. To investigate the impact of the privacy, this value is set to between 0 and 2, while the other parameters are set to default values(we choose $\delta = 10^{-4}$, sampling ratio=50%, CP-ranks=10, regularization coefficient=0.1 and number of iterations=200).

We compare the difference between our methods and other methods in Table III.

We study the effects of sampling ratio, CP-ranks, regularization coefficient, and the number of iterations under five different privacy degree schemes. First, with the increase of sampling ratio, regularization coefficient, and the number of iterations in Fig. 3(a)–(c), the RMSE of the five privacy schemes also decreases. And basically keep the order RMSE($\varepsilon = 0.5$) < RMSE($\varepsilon = 1.0$) < RMSE($\varepsilon = 1.5$) < RMSE($\varepsilon = 2.0$). In Fig. 3(d), the RMSE of the five privacy schemes increases with the increase of the CP-ranks.

Fig. 4 shows the performance in terms of fitness error and comparison with others algorithms. With the increase of the sampling ratio, thus, sample data, the fitness error decrease and, thus, better recovery performance is obtained; Meanwhile, with the gradual increase of privacy and the reduction of noise value, the fitness error decreases gradually.

In Fig. 5, RMSE grows as the number of algorithm iterations increases. By Theorem 3, the total privacy budget is $(3.035, 10^{-4}), (2.575, 10^{-4})$ under the $(\varepsilon - \delta)$ − DP for *Abilene* and *GÈANT* dataset, respectively, when DP-zCDP converges. As can be seen from Fig. 6(a), we used mean estimates to observe the impact of privacy. When the privacy is small, the statistical result of the estimated mean value has greatly deviated from the true value. According to Figs. 6(b) and (c), when the privacy is large, the statistical result of the estimated mean value is very close to the true value. Because the privacy ε directly affects the probability of getting a true answer under the randomized response mechanism. As shown in Fig. 7, the larger ε is, the higher the probability p of a true answer. As the privacy increases, the noise generated by the Laplace mechanism increases first and then decreases gradually in Fig. 7(a). In Fig. 7(b), the noise generated by the Gaussian mechanism gradually increases. In addition, it can be seen that when the given privacy is larger, the probability of a true answer is greater in Fig. 7(c). And also shows that how various privacy can affect the tradeoff between data availability and privacy protection. We need to set the privacy parameters appropriately according to the actual different applications.

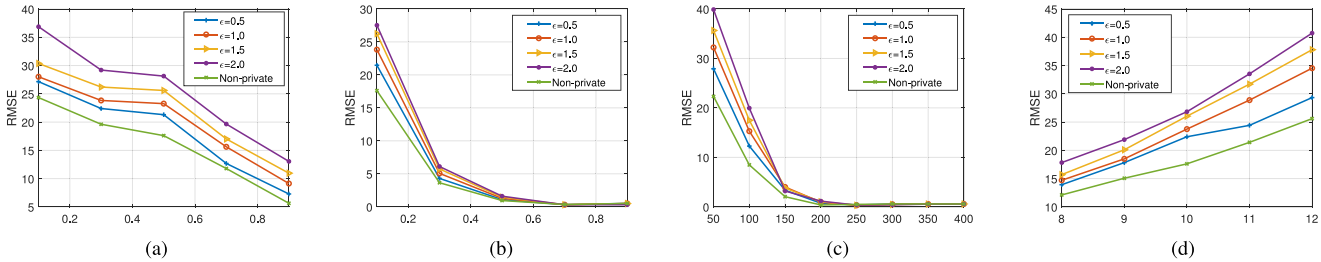

Fig. 3. Various parameter settings for DP in $Abilene$ dataset.

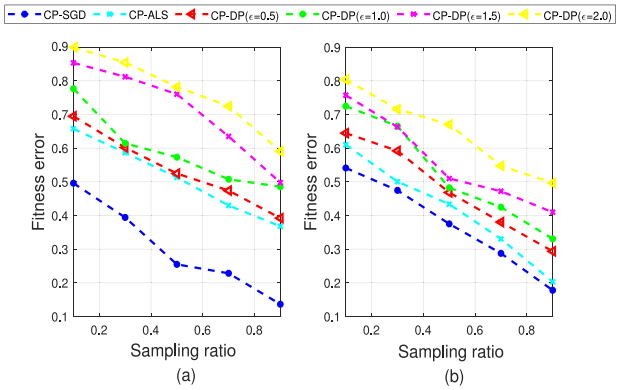

Fig. 4. Fitness error on (a) $Abilene$ and (b) $GÈANT$.

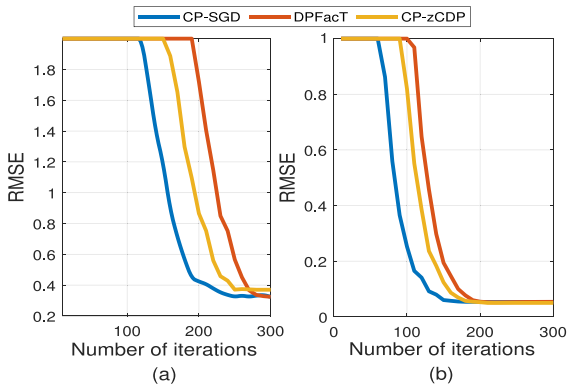

Fig. 5. RMSE on (a) $Abilene$ and (b) $GÈANT$.

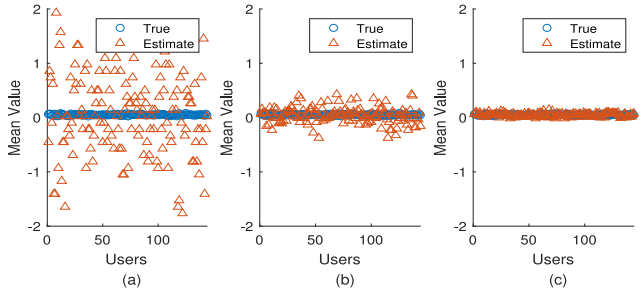

Fig. 6. Mean value estimation under different privacy budget in $Abilene$ dataset.

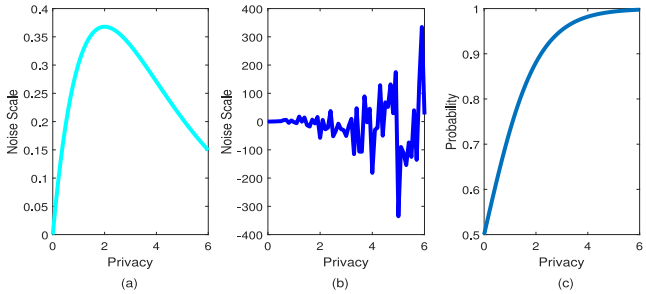

Fig. 7. Functional relation between ε and Noise scale or p.

V. Conclusion

Secure use of network traffic data in the future of 5G networks to provide user privacy is a new concern. Many effective tensor factorization method was proposed, but previous efforts to recover the network traffic data tensor have focused on increasing the computational rate issue. However, there is no privacy protection method for the disclosure of user information in the data center. So the framework for network traffic tensor data privacy protection (MDPSTF) was proposed. The multiple-strategies differential privacy was used for network traffic tensor data, and the experiment was carried out on two real datasets ($Abilene$ and $GÈANT$). The experimental results showed that various privacy budgets have different effects on RMSE and other evaluation indexes. In general, the framework can achieve privacy protection according to different privacy budgets and maintain data availability to a certain extent.

References

[1] J. Liu, F. Liu, and N. Ansari, "Monitoring and analyzing big traffic data of a large-scale cellular network with hadoop," *IEEE Netw.*, vol. 28, no. 4, pp. 32–39, Jul./Aug. 2014.

[2] J. M. Batalla *et al.*, "Security risk assessment for 5G networks: National perspective," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 16–22, Aug. 2020.

[3] H. Zhou, D. Zhang, K. Xie, and Y. Chen, "Spatio-temporal tensor completion for imputing missing internet traffic data," in *Proc. IEEE 34th Int. Perform. Comput. Commun. Conf.*, 2015, pp. 1–7.

[4] H. Xiao, J. Gao, D. S. Turaga, L. H. Vu, and A. Biem, "Temporal multi-view inconsistency detection for network traffic analysis," in *Proc. 24th Int. Conf. World Wide Web*, 2015, pp. 455–465.

[5] J. Xiong *et al.*, "Enhancing privacy and availability for data clustering in intelligent electrical service of IoT," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1530–1540, Apr. 2019.

[6] C. Dwork *et al.*, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3/4, pp. 211–407, 2014.
[7] M. Gong, Y. Xie, K. Pan, K. Feng, and A. K. Qin, "A survey on differentially private machine learning," *IEEE Comput. Intell. Mag.*, vol. 15, no. 2, pp. 49–64, May 2020.
[8] Y. Wang, Y.-X. Wang, and A. Singh, "Differentially private subspace clustering," *in Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1000–1008.
[9] D. Su, J. Cao, N. Li, E. Bertino, and H. Jin, "Differentially private k-means clustering," in *Proc. 6th ACM Conf. Data Appl. Secur. Privacy*, 2016, pp. 26–37.
[10] M. Abadi *et al.*, "Deep learning with differential privacy," in *Proc. 2016 ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 308–318.
[11] J. Hua, C. Xia, and S. Zhong, "Differentially private matrix factorization," in *Proc. 24th Int. Conf. Artif. Intell.*, 2015, pp. 1763–1770.
[12] J. Upadhyay, "The price of privacy for low-rank factorization," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 4180–4191.
[13] H. Shin, S. Kim, J. Shin, and X. Xiao, "Privacy enhanced matrix factorization for recommendation with local differential privacy," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 9, pp. 1770–1782, Sep. 2018.
[14] B. Ermiş and A. T. Cemgil, "Data sharing via differentially private coupled matrix factorization," *ACM Trans. Knowl. Discov. Data*, vol. 14, no. 3, pp. 1–27, 2020.
[15] Y. Wang and A. Anandkumar, "Online and differentially-private tensor decomposition," in *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 3539–3547.
[16] H. Imtiaz and A. D. Sarwate, "Distributed differentially private algorithms for matrix and tensor factorization," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 6, pp. 1449–1464, Dec. 2018.
[17] H. Li, K. G. Li, J. An, and K. G. Li, "An online and scalable model for generalized sparse non-negative matrix factorization in industrial applications on multi-GPU," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2019.2896634.
[18] H. Li, Z. Li, K. Li, J. S. Rellermeyer, L. Y. Chen, and K. Li, "SGD_tucker: A bovel stochastic optimization strategy for scalable parallel sparse tucker decomposition," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 7, pp. 1828–1841, Jul. 2021.
[19] K. Xie, L. Wang, X. Wang, G. Xie, and J. Wen, "Low cost and high accuracy data gathering in wsns with matrix completion," *IEEE Trans. Mobile Comput.*, vol. 17, no. 7, pp. 1595–1608, Jul. 2018.
[20] J. Ma, Q. Zhang, J. Lou, J. C. Ho, L. Xiong, and X. Jiang, "Privacy-preserving tensor factorization for collaborative health data analysis," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, 2019, pp. 1291–1300.
[21] X. Nie, L. T. Yang, J. Feng, and S. Zhang, "Differentially private tensor train decomposition in edge-cloud computing for SDN-based Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5695–5705, Jul. 2020.
[22] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. New York, NY, USA: Springer Sci. Bus Media., 2013.
[23] "The abilene observatory data collections," Accessed: Jul. 20, 2004. [Online]. Available: http://abilene.internet2.edu/observatory/data- collections.html
[24] S. Uhlig, B. Quoitin, J. Lepropre, and S. Balon, "Providing public intradomain traffic matrices to the research community," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 83–86, 2006.
[25] C. Dwork and G. N. Rothblum, "Concentrated differential privacy," 2016, *arXiv:1605.02065*.
[26] M. Bun and T. Steinke, "Concentrated differential privacy: Simplifications, extensions, and lower bounds," in *Proc. Theory Cryptography Conf.*, 2016, pp. 635–658.
[27] D. Kortenkamp, T. Milam, R. Simmons, and J. L. Fernandez, "Collecting and analyzing data from distributed control programs," *Electron. Notes Theor. Comput. Sci.*, vol. 55, no. 2, pp. 236–254, 2001.
[28] K. Xie, L. Wang, X. Wang, G. Xie, J. Wen, and G. Zhang, "Accurate recovery of Internet traffic data: A tensor completion approach," in *Proc. IEEE 35th Annu. Int. Conf. Comput. Commun.*, 2016, pp. 1–9.
[29] C. Battaglino, G. Ballard, and T. G. Kolda, "A practical randomized CP tensor decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 39, no. 2, pp. 876–901, 2018.

Jin Wang (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from the Nanjing University of Posts Telecommunications, Nanjing, China, in 2002 and 2005, respectively, and the Ph.D. degree in computer engineering from Kyung Hee University, Seoul, South Korea, in 2010.

He is currently a Professor with the Changsha University of Science and Technology, Changsha, China. He has authored or coauthored more than 400 international journal and conference papers. His research interests mainly include wireless ad hoc and sensor network, and network performance analysis and optimization.

Prof. Wang is a Fellow of IET.

Hui Han received the B.S. degree in mathematics and applied mathematics from Changsha University, Changsha, China, in 2019. She is currently working toward the M.S. degree in soft engineering with the Changsha University of Science and Technology, Changsha.

Her research interests include tensor decomposition and differential privacy.

Hao Li is currently working toward the Ph.D. degree in computer science and technology with Hunan University, Changsha, China.

From 2019 to 2021, he is a Visiting Ph.D. Student with TU Delft, Delft, The Netherlands. He has authored or coauthored several journal and conference papers in IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, InforSci, IEEE-TII, ACM-TDS, ACM CIKM, and IEEE ISPA. His research interests mainly include large-scale sparse matrix and tensor factorization, recommender systems, machine learning, and parallel and distributed computing.

Dr. Li is a Reviewer of the top-tier conferences and journals, such as HPCC, IJCAI, WWW, *Neurocomputing*, IEEE ACCESS, JPDC, InforSCI, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, IoT, ACM TKDD, and IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING.

Shiming He received the B.S. degree in information security and the Ph.D. degree in computer science and technology from Hunan University, Changsha, China, in 2006 and 2013, respectively.

She is currently an Associated Professor with the School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha. Her research interests include machine learning, data analysis, and anomaly detection.

Pradip Kumar Sharma (Senior Member, IEEE) received the Ph.D. degree in CSE from the Seoul National University of Science and Technology, Daejeon, South Korea, in 2019.

He is currently an Assistant Professor of cybersecurity with the Department of Computing Science, University of Aberdeen, Aberdeen, U.K. He was a Postdoctoral Research Fellow with the Department of Multimedia Engineering, Dongguk University, Seoul, South Korea. He was a Software Engineer with MAQ Software, Mumbai, India, and involved on variety of projects, proficient in building largescale complex data warehouses, OLAP models, and reporting solutions that meet business objectives and align IT with business. He has authored or coauthored many technical research papers in leading journals from IEEE, Elsevier, Springer, and MDPI. Some of his research findings are published in the most cited journals. His current research interests are focused on the areas of cybersecurity, blockchain, edge computing, SDN, and IoT security.

Dr. Sharma has been an Expert Reviewer for IEEE transactions, Elsevier, Springer, and MDPI journals and magazines. He is listed in the world's Top 2 Scientists for citation impact during the calendar year 2019 by Stanford University. He was the recipient of the Top 1 Reviewer in computer science by Publons Peer Review Awards 2018 and 2019, Clarivate Analytics. He has also been invited as the Technical Programme Committee Member and Chair in several reputed international conferences, such as the IEEE DASC 2021, IEEE CNCC 2021, CSA 20202, IEEE ICC2019, IEEE MENACOMM'19, and 3ICT 2019. He is currently an Associate Editor for the *Peer-to-Peer Networking and Applications*, *Human-centric Computing and Information Sciences*, *Electronics* , and *Journal of Information Processing Systems*. He has been the Guest Editor of international journals of certain publishers, such as IEEE, Elsevier, Springer, MDPI, and JIPS.

Lydia Chen (Senior Member, IEEE) received the B.A. degree from National Taiwan University, Taipei, Taiwan, in 2002, and the Ph.D. degree from Pennsylvania State University, State College, PA, USA, in 2006.

She is currently an Associate Professor with the Department of Computer Science, Delft University of Technology, Delft, The Netherlands. She has authored or coauthored more than 80 papers in journals, such as the IEEE TRANSACTIONS ON DISTRIBUTED SYSTEMS and IEEE TRANSACTIONS ON SERVICE COMPUTING, and conference proceedings, such as INFOCOM, Sigmetrics, DSN, and Eurosys. Her research interests include dependability management, resource allocation, privacy enhancement for large scale data processing systems and services, developing stochastic and machine learning models, and applying these techniques to application domains, such as datacenters and AI systems.

Dr. Chen was the corecipient of the Best Paper Awards at CCgrid'15 and eEnergy'15. She was the recipient of the TU Delft Professor Fellowship in 2018. She was the Program Co-Chair of Middleware Industry Track 2017 and IEEE ICAC 2019, and the Track Vice-Chair of ICDCS 2018. She was on the Editorial Boards of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, IEEE TRANSACTIONS ON SERVICE COMPUTING, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTIN, and IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS.