A New End-To-End Network Traffic Reconstruction Approach Based on Different Time Granularities

Wei Yang[1], Dingde Jiang[1](✉), Jianguang Chen[1], Zhihao Wang[1], Liuwei Huo[1], and Wenhui Zhao[2]

[1] University of Electronic Science and Technology of China, Chengdu 611731, China
[2] College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Abstract. End-to-end network traffic is an important input parameter for network planning and network monitoring, which plays an important role in network management and design. This paper proposes a new end-to-end network traffic reconstruction algorithm based on different time granularity. This algorithm reconstructs the end-to-end network traffic with fine time granularity by taking advantage of the characteristics of the link traffic which is easy to be measured directly in the network with coarse time granularity. According to the fractal and self-similar characteristics of network traffic found in existing studies, we first carry out fractal interpolation for link traffic measurement under coarse time granularity to obtain link traffic under fine time granularity. Then, by using the compressive sensing theory, an appropriate sparse transformation matrix and measurement matrix are constructed to reconstruct the end-to-end network traffic with fine time granularity. Simulation results show that the proposed algorithm is effective and feasible.

Keywords: End-to-end network traffic reconstruction · Fractal interpolation · Compression sense · Dictionary Learning Algorithm

1 Introduction

The end-to-end network Traffic Matrix (TM) is an overview of the entire network Traffic. The elements in the matrix represent the traffic in the network that starts at one node (source node) and ends at another node (destination node) [1]. The traffic can be in packets or bytes. Source and Destination Pairs are called OD Pairs. Traffic matrix is not only the key input of network management and traffic engineering tasks, but also can be applied to network management, network security and other fields.

In recent years, with the continuous expansion of network functions and the diversification of users' requirements for network performance [2–4], the end-to-end network traffic is very important for the network management [5], and it has become a hot topic for the network traffic measurement. There are many network traffic measurement scheme, such as SNMP (Simple Network Management Protocol), sending probing packet [6],

D. Jiang and H. Song (Eds.): SIMUtools 2021, LNICST 424, pp. 111–122, 2022.
https://doi.org/10.1007/978-3-030-97124-3_10

and so on. The end-to-end network traffic reconstruction in the traffic matrix has become an important research field that recover the end-to-end network traffic with measured coarse-granularity network traffic of link load.

For the end-to-end network traffic, it has become a hot research field that how to reconstruct the traffic matrix to meet the requirement of certain accuracy through limited measurement information [7]. At present, there are many reconstruction methods of the traffic matrix. Fan et al. sparsely represent the normal and anomalous traffic in wavelet and time domain, and use the convex program to estimate both the normal and anomalous components of the traffic matrix [8]. Pachuau et al. use the traffic to represent all traffic flows and use the Genetic algorithm (GA) based optimization method to further the solutions of the Gravity model [9], and authors use the Gravity model to obtain an initial solution and use GA model to solve link load-TM. Amoroso et al. [10] propose a super resolution technique for traffic matrix inference that does not require any knowledge on the structural properties of the matrix elements to infer, nor a large data collection. Jiang et al. [11] study the measurement and analysis technology for software-defined networking of IoV and propose a performance measurement and analysis method to measure and characterize its performance.

This paper studies the reconfiguration of end-to-end network traffic under the granularity of fine time measurement by using the information that is easy to be directly measured in the existing network: link load and routing matrix under the granularity of coarse time measurement. Whether the end-to-end network traffic can be reconstructed at different time granularity depends to a great extent on how to obtain the link load required by the traffic matrix at fine time granularity from the link load at coarse time granularity. The traditional mathematical interpolation of link load, such as Lagrange interpolation, linear interpolation, etc. cannot well reflect the characteristics of network traffic in time scale because these interpolation methods are only numerical interpolation, so the traditional interpolation cannot well meet the characteristics of traffic data. Given that network traffic has self-similarity characteristics in a larger time scale and fractal characteristics in a smaller time scale link load, as a linear weight of end-to-end network traffic, should also have fractal characteristics. In this paper, the application of fractal interpolation to link load is first proposed, that is, the link load under coarse time measurement granularity is fractal interpolation and the link load under fine time measurement granularity is obtained.

In the following, the paper is organized that Sect. 2 makes a problem elaboration in. Then, the end-to-end traffic is reconstructed and proposed the RLS-DLA algorithm in Sect. 3. Finally, we perform some simulations to verify our proposed methods in Sect. 4 and make a conclusion for our works in Sect. 5.

2 Problem Elaboration

The end-to-end network traffic matrix represents the traffic transmission between all OD pairs in the network. If there are n nodes in a network, then there are n^2 OD flows. The end-to-end network traffic matrix at time t can be expressed as $x(t) = [x_1(t), x_2(t), \ldots, x_k(t), \ldots, x_{n^2}(t)]^T$, where $x_k(t)$ represents the traffic of OD flow in article k at time t. Then, there are $k = (i-1) * n + j$ OD flows that $x_k(t)$ starts from the source node i and ends at the node j.

Link load matrix represents the change in traffic rate over each link at time t. The link traffic matrix can be expressed as $y(t) = [y_1(t), y_2(t), \ldots, y_r(t), \ldots, y_l(t)]^T$, where l represents the total number of links in the entire network. y_r represents the traffic rate change of time t on link r. The size of the routing matrix A is $l \times c(c = n^2)$. When OD flow j passes link i, route matrix element $a_{ij} = 1$, otherwise $a_{ij} = 0$. Each list in matrix A shows all the links that OD flows through as it traffic through the entire network. The routing matrix contains the actual routing information of the network. In this paper, it is considered that the routing matrix is stable and does not change in the period of traffic matrix reconstruction. The relationship between link load, routing matrix, and end-to-end network traffic matrix can be expressed as

$$y(t) = Ax(t) \tag{1}$$

The link load matrix can be directly measured through SNMP. There are many ways to obtain route matrix $y(t)$, usually by collecting the configuration information of IGP (Interior Gateway Protocol) or collecting the link state information of the interaction between routers and calculating the number of shortest paths. Therefore, the end-to-end network traffic reconfiguration problem is transformed into the problem of knowing $y(t)$ and A and solving $x(t)$. If A is a full rank square matrix, then the solution of Eq. (1) exists and is unique. However, in the actual network, the number of OD flow is often much greater than the number of links, that is $c \gg l$. The matrix A is not a full rank matrix. According to $x(t) = A^{-1}y(t)$, it is difficult to obtain the exact solution of end-to-end network traffic, so it is a great challenge to reconstruct the end-to-end network traffic.

More importantly, existing IP network devices do not always support direct measurements of end-to-end network traffic for a variety of reasons. And some devices support this kind of measurement. However, due to the rapid increase of network services, the equipment will need to collect a lot of measurement data. This will add additional burden to network equipment, which will affect network performance and affect the normal operation of the network.

In view of the above problems, this paper proposes to reconstruct the end-to-end traffic of the entire network under the fine time measurement granularity by directly measuring the link traffic under the coarse time granularity. The goal is to obtain end-to-end network traffic with as few measurements as possible to minimize the impact on network equipment and network performance. Fractal interpolation of link load $y_c(t)$ under coarse time measurement granularity is performed to obtain link load $y_f(t)$ under fine time granularity. Then, given $y_f(t)$ and A, the end-to-end solve the network traffic matrix $x(t)$.

Through the first time for coarse granularity fractal interpolation to get the link traffic $y_c(t)$ under the link traffic $y_f(t)$ under fine granularity of time, and then using the theory of compressed perception by reconstructing the corresponding time $y_f(t)$ under the granularity of end-to-end network traffic $x_f(t)$, and the $x_f(t)$ satisfy the following conditions:

$$y_f(t) = Ax_f(t) \tag{2}$$

Here, according to the above elaboration, we propose an end-to-end network traffic reconstruction model based on different time granularity.

The end-to-end network traffic $x_1(t), x_2(t), ..., x_k(t), ..., x_c(t)$ has significant self-similarity and fractal characteristics $y_1(t), y_2(t), ..., y_r(t), ..., y_l(t)$ also have obvious self-similarity and fractal characteristics. We propose to use fractal interpolation to solve the link traffic reconstruction problem under fine time granularity.

A. Fractal Interpolation of Link Flow Signal

Network traffic reflects the activities of network devices and users accessing the network. It has obvious long correlation, short correlation, self-similarity and heavy tail distribution. Reflect the local and the whole have obvious similarity. For objects with this fractal feature, we can use a fractal interpolation algorithm. The link traffic with fine time granularity can be obtained by interpolating the measured link traffic with coarse time granularity.

Suppose the known coarse time granularity is N times the fine time granularity (where $N \geq 2$ is an integer). For the link traffic $y_c(t)$ under coarse time granularity, we can get the following equation:

$$\begin{cases} \{(t, y_{1c}(t))|y_{1c}(t) \in R; t = 1, 1+N, \ldots, 1+(T-1)N\} \\ \{(t, y_{2c}(t))|y_{2c}(t) \in R; t = 1, 1+N, \ldots, 1+(T-1)N\} \\ \ldots \\ \{(t, y_{lc}(t))|y_{lc}(t) \in R; t = 1, 1+N, \ldots 1+(T-1)N\} \end{cases} \tag{3}$$

where T represents the number of moments under coarse granularity of link traffic.

For each link traffic, we hope to construct the following function to obtain the link traffic data at $[t, t+N]$ time points within the measurement interval T:

$$f_i : \begin{cases} y_{ic}(t) \rightarrow y_{ic}(\hat{t}) \\ t \rightarrow \hat{t} \end{cases} \tag{4}$$

where $i = 1, 2, \ldots, l, t = 1, 1+N, \ldots, 1+(T-1)N$.

According to Eq. (4), we use the fractal theory to construct the function f_i through iteration, that is:

$$w_{ij}\begin{pmatrix} \hat{t} \\ y_{ic}(\hat{t}) \end{pmatrix} = \begin{pmatrix} a_{ij} & 0 \\ c_{ij} & s_{ij} \end{pmatrix}\begin{pmatrix} t \\ y_{ic}(t) \end{pmatrix} + \begin{pmatrix} d_{ij} \\ e_{ij} \end{pmatrix} \tag{5}$$

where $\hat{t}$ represents the corresponding moment calculated through the transformation of equation t (5). Similarly, $y_{ic}(\hat{t})$ represents the link traffic value at the corresponding $\hat{t}$ moment calculated by $y_{ic}(t)$ transformation. $i = 1, 2, \ldots, l$ and $j = 2, 3, \ldots, T$, w_{ij} represents the functional relationship within the measurement interval $[t, t+N]$, and $f_i = (w_{i2}, w_{i3}, \ldots, w_{iT})$. Therefore, f_i represents the iterative function system composed of $(T-1)$ functions.

For $t = 1, 1+N, \ldots, 1+(T-1)N$, at $\hat{t} = 1, 1+(T-1)N$, Eq. (5) must satisfy:

$$\begin{cases} w_{ij}\begin{pmatrix} 1 \\ y_{ic}(1) \end{pmatrix} = \begin{pmatrix} t_{j-1} \\ y_{ic}(t_{j-1}) \end{pmatrix} \\ w_{ij}\begin{pmatrix} 1+(T-1)N \\ y_{ic}(1+(T-1)N) \end{pmatrix} = \begin{pmatrix} t_j \\ y_{ic}(t_j) \end{pmatrix} \end{cases} \tag{6}$$

where t_{j-1} and t_j represent the $j-1$ and j values in t respectively.

According to Eqs. (5) and (6):

$$\begin{cases} a_{ij} + d_{ij} = t_{j-1} \\ a_{ij}(1 + (T-1)N) + d_{ij} = t_j \\ c_{ij} + s_{ij} y_{ic}(1) + e_{ij} = y_{ic}(t_{j-1}) \\ c_{ij}(1 + (T-1)N) + s_{ij} y_{ic}(1 + (T-1)N) + e_{ij} = y_{ic}(t_j) \end{cases} \tag{7}$$

where $i = 1, 2, \ldots, l$ and $j = 2, 3, \ldots, T$. In general, s_{ij} is chosen as a constant, and $0 \le s_{ij} < 1$. Then, we can calculate a_{ij}, d_{ij}, c_{ij} and e_{ij} according to Eq. (7), and uniquely determine the iterative function system f_i represented by Eqs. (5)–(6). In general, in order to ensure that the function w_{ij} (where $j = 2, 3, \ldots, T$) can well describe the fractal characteristics of link traffic, a_{ij}, c_{ij} and s_{ij} meet the following constraints:

$$\max(\sqrt{a_{ij}^2 + (1+\varepsilon)c_{ij}^2}, \sqrt{\frac{1}{\varepsilon} + 1}|s_{ij}|) < 1 \tag{8}$$

where ε is an constant that $\varepsilon = 0$.

According to coarse time, the link load $y_{ic}(t)$ under granularity was measured, and the iteration coefficient $a_{ij}, c_{ij}, d_{ij}, e_{ij}$ was first obtained through formula (7). When $j = 2, \ldots, T$, in formula (5), $t = 1, 1+N, 1+2N, \ldots, 1+(T-1)N$, the corresponding w_{ij} is obtained in turn, and the interpolation process of fractal interpolation is completed. The attractor $G^{'} = \cup_{j=2}^{T} w_j(G)$ of IFS is a graph of continuous function $f[1, 1+(T-1)N] \to R$ through all interpolation points $(t, y(t))$, $t = 1, 1+2N, \ldots, 1+(T-1)N$. The continuous Function is called Fractal Interpolation Function (FIF).

B. Link Flow Interpolation Reconstruction

In this paper, fractal interpolation is applied successively to link load under coarse time measurement granularity of l bar to obtain link load under fine time measurement granularity of l bar. According to the principle of fractal interpolation, it can be seen that, if the coarse time measurement granularity is known, the link load t has T moments, then $\hat{t}$ in the obtained fractal interpolation function $G^{'}$ has $T(T-1)$ moments. Obviously, the number of moments obtained is far more than the required moment $1 + (T-1)N$.

There are only two moments that are closest to the required $\tilde{t}$ moment are selected from the obtained $\hat{t}$ moment, and the corresponding two link traffic values are averaged as the link traffic values at the required moments.

In summary, it is known that the link load of t at T moments under coarse time measurement granularity is known, and the interval is N. Through fractal interpolation and selection of required moments, the link load of $1 + (T-1)N$ at $\tilde{t}$ moments under fine time measurement granularity is finally obtained. As Eq. (9) shows:

$$\begin{cases} \{(\tilde{t}, y_{1f}(\tilde{t})) | y_{1f}(\tilde{t}) \in R; \tilde{t} = 1, 2, 3, \ldots, 1 + (T-1)N\} \\ \{(\tilde{t}, y_{2f}(\tilde{t})) | y_{2f}(\tilde{t}) \in R; \tilde{t} = 1, 2, 3, \ldots, 1 + (T-1)N\} \\ \ldots \\ \{(\tilde{t}, y_{lf}(\tilde{t})) | y_{lf}(\tilde{t}) \in R; \tilde{t} = 1, 2, 3, \ldots, 1 + (T-1)N\} \end{cases} \tag{9}$$

3 End-To-End Network Traffic Reconfiguration

In the third part of this paper, fractal interpolation is used to obtain the link load with fine time measurement granularity corresponding to the traffic matrix. Then, we use the compressive sensing framework to reconstruct end-to-end network traffic $x_f(\tilde{t})$.

A. Sparse Transformation Matrix Construction

According to the compressive sensing theory, when the original traffic signal is not k-sparse signal, the original signal $x_f(\tilde{t})$ in Eq. (2) needs to be sparse represented by the transformation matrix Ψ, that is:

$$x_f(\tilde{t}) = \Psi * s(\tilde{t}) \tag{10}$$

where $s(\tilde{t})$ after transformation is sparse signal or compressible signal, $\tilde{t} = 1, 2, 3, \ldots, 1 + (T-1)N$.

Signal sparse representation is widely used in various fields of signal processing and image processing. It is also the basis of compressed sensing theory. Only by selecting an appropriate transformation matrix, can the sparse representation coefficient $s(\tilde{t})$ be ensured to have sufficient sparsity or attenuation and the reconstruction accuracy of compressed sensing. The most prominent advantage of this algorithm is that in the algorithm, the forgetting factor λ and $0 \leq \lambda < 1$ are introduced to gradually get rid of the dependence on the initial dictionary in the training process and enhance the sparse representation ability of the RLS learning dictionary.

In this paper, under the fine-time measurement granularity, the real end-to-end traffic data at the first $T^{'}$ moments is taken as the training data $x\_train(t), t = 1, 2, ..., T^{'}$, where $T^{'} < 1 + (T-1)N$, and RLS-DLA is adopted for sparse representation or sparse approximation, then according to Eq. (10):

$$x\_train(t) = D * s\_train(t) \tag{11}$$

where, $c \times 1$ is the size of $x\_train(t)$, $c = n^2$ is the total number of OD flow in the network with n nodes, D is the over-complete dictionary, and its size is $c \times c$. $s\_train(t)$ is the sparse representation coefficient of D under the dictionary of the training data $x\_train(t)$, and its size is $c \times 1$. The pseudo code for RLS-DLA as Table 1 shows.

B. Obtain End-To-End Traffic

In this paper, the dictionary D obtained by training with some real traffic data is taken as the sparse transformation matrix Ψ in the reconstruction of all end-to-end network traffic. Then, combining Eqs. (2) and (10) can be written as follows:

$$y_f(\tilde{t}) = A * x_f(\tilde{t}) = A * D * s(\tilde{t}) \tag{12}$$

where $\tilde{t} = 1, 2, 3, \ldots, 1 + (T-1)N$.

The problem is transformed into: given the link load $y_f(\tilde{t})$, routing matrix A and sparse transformation matrix D under the granularity of fine time measurement, how to solve $s(\tilde{t})$ through the compressed sensing reconstruction algorithm. In the compressive sensing theory, a sufficient condition for the reconstruction algorithm to accurately reconstruct the original signal is that $\Theta = \Phi * \Psi$ meets the RIP criterion, which is equivalent to that the measurement matrix Φ is irrelevant to the transformation matrix Ψ.

In Eq. (12), A is the fixed routing matrix, and the matrix element is 0 or 1; D is a fixed sparse transformation matrix, and obviously $A * D$ cannot meet the RIP criterion. We generate a random Gaussian matrix $R \sim N(0, \frac{1}{n^2})$ of size $m \times l$, l is the total number of links under link load, and $m < l \ll n^2$. For formula (12), multiply left and right sides by the matrix R at the same time:

Table 1. The pseudocode for RLS-DLA

Algorithm: Recursive Least Squares Dictionary Learning Algorithm
Input: training data: $x\_train(t)$
Output: overcomplete dictionary: D
Initialization:
$D = x\_train(:,1:n^2)$ {use first n^2 training vectors as initial dictionary and then normalize D}
$S = \phi$ {initial coefficient matrix}
$m = 1$ {the number of iterations}
$K = 10$ {sparsity of OMP}
repeat
get the forgetting factor: $0 \le \lambda_m < 1$
$x = X\_train(:,m)$ {every training vector}
$s = OMP(x, D, K)$ {coefficient vector}
$r = x - D * s$ {signal residual}
$S = S \bigcup s$
$C = S * S^T$
$C_\lambda = \lambda^{-1} * C$ {apply forgetting factor}
$u = C_\lambda * s$
$\alpha = 1/(1 + w^T * u)$
$D = D + \alpha * r * u^T$ {update the dictionary}
$C = C_\lambda - \alpha * u * u^T$ {update C from m = 2}
$m = m + 1$
until $m > T'$ {T' training vectors}
Output: D

$$R * y_f(\tilde{t}) = R * A * D * s(\tilde{t}) = \Phi' * D * s(\tilde{t}) \quad (13)$$

Therefore, the new measurement matrix $\Phi^{'} = R * A$ constructed in this way is still an independent co-distributed Gaussian random matrix, which is highly irrelevant to the fixed sparse transformation matrix D, so $\Theta^{'} = \Phi^{'} * D$ can meet the RIP criterion. The new observation vector $z(\tilde{t}) = R * y_f(\tilde{t})$ has a size of $m \times 1$. Finally, for the observation vector $z(\tilde{t})$ of each time, the OMP (Orthogonal Matching Pursuit) algorithm is called to reconstruct the sparse representation coefficient $s(\tilde{t})$ of the corresponding time. Note: the desired end-to-end network traffic $x_f(\tilde{t}) = D * s(\tilde{t})$. The measurement matrix $\Phi^{'}$ constructed in this way not only meets the RIP criterion but also reduces the dimension of the observation vector $z(\tilde{t})$ while ensuring the reconstruction accuracy. IPFP constraint optimization is carried out on the obtained initial traffic matrix $x_f(\tilde{t})$ to obtain the network end-to-end traffic $x(\tilde{t})$ under the fine time measurement granularity.

4 The Simulation Analysis

In this paper, the simulation experiment used the real traffic data from the Abilene backbone network for one week to verify. Under different time measurement granularity, the link load with coarse time measurement granularity was used to reconstruct the network end-to-end traffic with fine time granularity, that is, the traffic matrix. Abilene network has a total of $n = 12$ nodes, that is, there is $c = n^2 = 144$ OD flow. The Abiline network has 30 links, plus the two outgoing and incoming links of each node, a total of 54 links, so $l = 54$. In the actual construction of the measurement matrix, a random Gaussian matrix R of size 20×54 is generated, that is $m = 20 < l = 54 < c = 144$. TomoG, SRSVD, and PCA methods are considered to be more accurate methods for end-to-end network traffic reconstruction. This paper will compare the new method of reconstructing end-to-end traffic of finer-grained network from coarse-grained time measurement link load with the above three methods, and analyze and compare the effect of OD flow reconfiguration, Spatial relative errors (SREs) and Temporal relative errors (TREs) of the four methods.

A. Reconstruction Effect of Traffic

In this paper, the OD20 stream was randomly selected from 144 OD streams. The reconstruction effect of OD20 under different coarse time measurement granularity is compared with TomoG, SRSVD and PCA respectively. The reconstruction effect with coarse time measurement granularity of 10 min is demonstrated. When the time measurement granularity of link load is 10 min, the OD20 reconstructed and the OD20 reconstructed through SRSVD, TomoG and PCA are compared with the true value of OD20, as shown in Fig. 1. Real20 represents the true values of OD20 under 5 min measurement granularity; In Fig. 1, NEW 10 min represents the OD20 reconstructed when the measurement granularity is 10 min. It can be seen from the reconstruction results of OD20 that OD flow has the best reconstruction effect when the coarse-time measurement granularity is 10 min, which is better than the other three reconstruction methods and other coarse-time measurement granularity. As the granularity of time measurement increases, it can be seen that the accuracy of reconstruction decreases gradually, which is related to the decrease of known link load time.

B. Error Assessment

The SREs of traffic matrix estimation represent the spatial relationship of the estimation error with OD flow, which reflects the estimation accuracy of the estimation method in space. The TREs represent the variation of the estimation error with time, reflecting the estimation accuracy of the estimation method in time. SREs and TREs of traffic matrix are important indicators to measure the quality of estimation methods. They are respectively defined as follows:

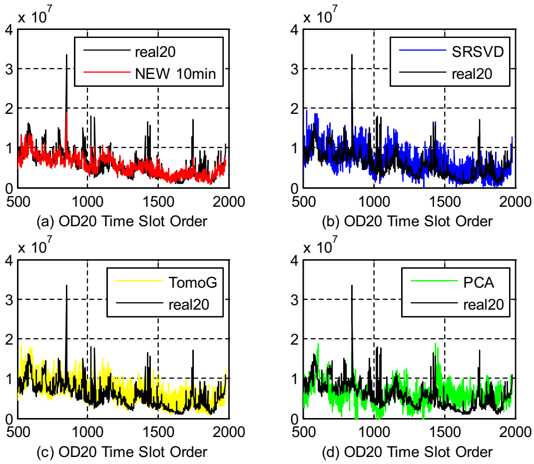

Fig. 1. The reconstruction effect of OD20 at particle size was measured in 10 min

$$\left\{ \begin{array}{ll} err_s(i) = \frac{||\hat{x}_M(i) - x_M(i)||_2}{||x_M(i)||_2}, & t = 1, 2, \ldots, M \\ err_t(t) = \frac{||\hat{x}_c(t) - x_c(t)||_2}{||x_c(t)||_2}, & i = 1, 2, \ldots, c \end{array} \right\} \tag{14}$$

where $c = n^2$ represent the total number of OD streams in IP backbone network, and M represents the number of moments of network traffic reconstructed under fine time measurement granularity; $err_{sp}(i)$ represents the SREs of OD flow $err_{sp}(i)$ in article i relative to all measuring moments; $x_M(i)$ and $\hat{x}_M(i)$ respectively represent the true value and estimated value of OD flow in article i relative to all measuring moments. $err_{tm}(t)$ represents the TREs of all OD flow at time t. $x_c(t)$ and $\hat{x}_c(t)$ represent the true and estimated values of all OD flow at time t, respectively.

Given the real traffic data of the Abilene network for one week, the time measurement granularity is 5 min, the data of 2016 real end-to-end network traffic moments. In this paper, link loads with coarse time measurement granularity of 10 min are used to reconstruct the network end-to-end traffic with fine time measurement granularity of 5 min. For coarse time measurement, the known number of time T corresponding

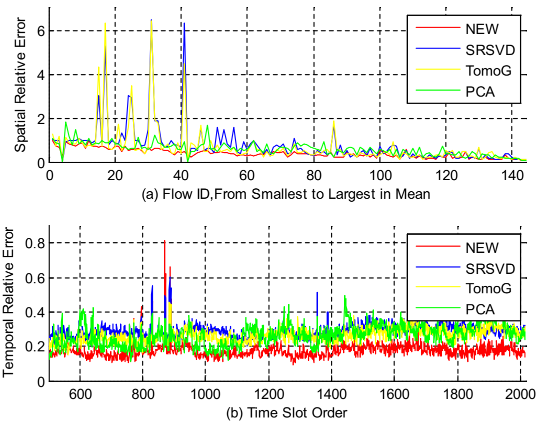

Fig. 2. SREs (a) and TREs (b) of network end-to-end traffic reconstructed when the coarse-time measurement granularity is 10 min.

to particle size 10 min is 1008, and the corresponding number of time interval N is 2. Therefore, the number of moments of network end-to-end traffic measured at 5 min by fractal interpolation and compressed sensing reconstruction algorithm: $1 + (T - 1) * N$ corresponds to 2015. Due to measurement of particle size for 5 min before the real traffic data $T^{'} = 500$ a moment as the RLS-DLA constructed over complete dictionary when training data *x_train*, therefore, in the final reconstruction of traffic data to remove its 500 times before, then the results with TomoG, SRSVD and PCA method carries on the analysis comparison, the reconstruction error is shown in Fig. 2.

Figure 2 shows SREs and TREs of end-to-end traffic of the Abilene network reconstructed from time measurement granularity of 10 min. It is not difficult to find that for different coarse time granularity, SREs of the new method proposed in this paper are all smaller than TomoG, SRSVD and PCA. However, with the increase of time granularity, the TREs of the new method increases obviously, and the performance of end-to-end traffic reconstruction becomes worse. The reason is that, with the increase of time measurement granularity, the number of moments of known link load becomes less and less, the network traffic error obtained by reconstruction will inevitably increase, and the reconstruction effect will also become worse and worse.

To evaluate the reconstruction accuracy of the four methods more accurately, the CDFs (Cumulative Distribution Function) of SREs and TREs are analyzed below. That is, CSRE (Cumulative distribution function of Spatial Relative Errors) and STRE (Cumulative distribution function of Temporal Relative Errors). Figure 3 is the CSRE and CTRE of end-to-end traffic of Abilene network reconstructed by time measurement granularity of 10 min.

In Fig. 3, coarse time measurement granularity is 10 min, that is, when the number of moments to be obtained is about twice the number of known moments. When SREs is 0.71, the new method proposed in this paper-reconstructing fine-grained traffic matrix from coarse-grained link load can accurately reconstruct 90.3% OD flow. When TREs is 0.21, the new method can accurately reconstruct 90.0% OD flow.

It is not difficult to find that the CDF curve of spatial relative error changes little with the increase of time measurement granularity. The performance of the new method proposed in this paper is superior to the other three existing methods. But time relative error of the CDF curves in rough time measuring particle size for 10 min while performance gradually decline, but are better than the other three kinds of existing methods. This also reflects that, as the number of known moments decreases, the known information decreases, so the effect of reconstructing the entire network end-to-end traffic gradually becomes worse.

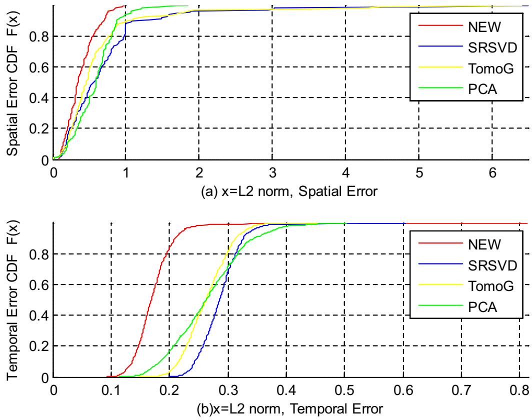

Fig. 3. CSRE (a) and CTRE (b) of network end-to-end traffic reconstructed when the coarse-time measurement granularity is 10 min.

5 Conclusions

Network end-to-end traffic reconstruction is a highly ill-posed linear inverse problem. How to obtain more accurate estimation of network end-to-end traffic through appropriate methods is a hot and difficult issue in current research. In this paper, the link load under coarse time measurement granularity, which is easy to be directly measured in the network, is used to reconstruct the network end-to-end traffic under fine time measurement granularity under the precondition that the routing matrix is known. Then, a new sparse transformation matrix and measurement matrix are constructed, and the compressive sensing framework is used. The network end-to-end traffic under fine time measurement granularity is obtained from the link load reconstruction under fine time measurement granularity. Simulation results show that with the increase of the time measurement granularity, the reconstruction performance of the new method becomes worse gradually.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (No. 61571104), the Sichuan Science and Technology Program (No. 2018JY0539), the Key projects of the Sichuan Provincial Education Department (No. 18ZA0219), the Fundamental

Research Funds for the Central Universities (No. ZYGX2017KYQD170), the CERNET Innovation Project (No. NGII20190111), the Fund Projects (Nos. 2020-JCJQ-ZD-016–11, 61403110405, 315075802, JZX6Y202001010161), and the Innovation Funding (No. 2018510007000134). The authors wish to thank the reviewers for their helpful comments.

References

1. Hashemi, H., Abdelghany, K.: End-to-end deep learning methodology for real-time traffic network management. Comput.-Aided Civ. Infrastruct. Eng. **33**(10), 849–863 (2018)
2. Jiang, D., Wang, Z., Wang, W., et al.: AI-assisted energy-efficient and intelligent routing for reconfigurable wireless networks. IEEE Trans. Netw. Sci. Eng. **9**, 78–88 (2021)
3. Clemm, A., Zhani, M.F., Boutaba, R.: Network management 2030: operations and control of network 2030 services. J. Netw. Syst. Manage. **28**(4), 721–750 (2020)
4. Jiang, D., Wang, F., Lv, Z., et al.: QoE-aware efficient content distribution scheme for satellite-terrestrial networks. IEEE Trans. Mob. Comput. (2021)
5. Petrov, V., Lema, M.A., Gapeyenko, M., et al.: Achieving end-to-end reliability of mission-critical traffic in softwarized 5G networks. IEEE J. Sel. Areas Commun. **36**(3), 485–501 (2018)
6. Zhang, H., Cai, Z., Liu, Q., et al.: A survey on security-aware measurement in SDN. Secur. Commun. Netw. **2018**, 1–15 (2018)
7. Kumar, A., Vidyapu, S., Saradhi, V.V., et al.: A multi-view subspace learning approach to internet traffic matrix estimation. IEEE Trans. Netw. Serv. Manage. **17**(2), 1282–1293 (2020)
8. Fan, X.B., Xu, X.: Sparse representation for network traffic recovery. Comput. Commun. **160**, 547–553 (2020)
9. Pachuau, J.L., Roy, A., Krishna, G., et al.: Estimation of traffic matrix from links load using genetic algorithm. Scalable Comput.: Pract. Exp. **22**(1), 29–38 (2021)
10. Amoroso, R., Esposito, F., Merani, M.: Estimation of traffic matrices via super-resolution and federated learning. In: Proceedings of the 16th International Conference on emerging Networking EXperiments and Technologies, pp. 560–561 (2020)
11. Jiang, D., Wang, Z., Huo, L., et al.: A performance measurement and analysis method for software-defined networking of IoV. IEEE Trans. Intell. Transp. Syst. **22**, 3707–3719 (2020)