CUR Decomposition for Compression and Compressed Sensing of Large-Scale Traffic Data

Nikola Mitrovic, Muhammad Tayyab Asif, Umer Rasheed, Justin Dauwels, and Patrick Jaillet

Abstract— **Intelligent Transportation Systems (ITS) often operate on large road networks, and typically collect traffic data with high temporal resolution. Consequently, ITS need to handle massive volumes of data, and methods to represent that data in more compact representations are sorely needed. Subspace methods such as Principal Component Analysis (PCA) can create accurate low-dimensional models. However, such models are not readily interpretable, as the principal components usually involve a large number of links in the traffic network. In contrast, the CUR matrix decomposition leads to low-dimensional models where the components correspond to individual links in the network; the resulting models can be easily interpreted, and can also be used for compressed sensing of the traffic network. In this paper, the CUR matrix decomposition is applied for two purposes: (1) compression of traffic data; (2) compressed sensing of traffic data. In the former, only data from a "random" subset of links and time instances is stored. In the latter, data for the entire traffic network is inferred from measurements at a "random" subset of links. Numerical results for a large traffic network in Singapore demonstrate the feasibility of the proposed approach.**

I. INTRODUCTION

Enhancements in sensor technologies such as GPS probes have led to rapid development in the field of Intelligent Transportation Systems (ITS). These systems are usually deployed in large metropolitan areas. Consequently, they have to deal with datasets from thousands of road segments with high temporal resolution [1]. The scale of the networks poses many challenges for ITS systems. One challenge is to infer meaningful spatial and temporal trends in large and diverse networks [2]. Such relations can be useful for many ITS applications, as we exploit them to develop highly accurate low-dimensional models for large and diverse road networks. Low-dimensional models can help to lower the strain on computational resources of ITS. Techniques such as Principal Component Analysis (PCA) have been proposed to create low-dimensional representations of road networks, for applications such as data compression [3]. Although methods such as PCA yield accurate low-rank approximations, they are hard to interpret in terms of individual links in the network. The principal components may involve hundreds or even thousands of links, and therefore, it is not straightforward to comprehend the role of individual links in the low-dimensional models.

In this paper, we explore the CUR matrix decomposition [4-6] to model large-scale traffic networks. Unlike other subspace methods such as PCA, CUR creates low-rank approximations where individual links of the traffic network correspond to basis vectors. In other words, the low-dimensional CUR model only depends on a subset of links (e.g., only 10% of the links). As a result, we obtain highly interpretable representation of large and diverse networks. Moreover, to infer the state of the entire network through CUR decomposition, only measurements of a limited number of links is required, paving the way to powerful compressed sensing algorithms for large-scale traffic networks. So far, CUR decomposition has not yet been utilized in the context of urban traffic modeling. CUR methods have found applications in many fields such as social networks [4], image processing [7, 8], and biology [8].

Previous studies related to low-dimensional representations for traffic networks mainly deal with PCA [3, 9-12]. Djukic et al. applied PCA on small network with OD pair data [9, 10]. In another study, Asif et al. applied different subspace methods for compression of traffic speed [11]. These studies showed that subspace methods such as PCA and discrete cosine transform (DCT) can efficiently compress traffic data [11]. However, they fail to provide insight about traffic behavior at specific roads and time periods. In contrast, the CUR decomposition involves data from individual links and time instance, and hence, it allows us to directly infer the underlying spatial and temporal patterns in large road networks.

To assess our approach, we consider a large urban network in Singapore comprising of more than 6000 road segments, including two expressways as well as parts of the downtown area. A road segment (link) is defined as the portion between two consecutive intersections (nodes). The Singapore Land Transportation Authority (LTA) provided speed data for several months in the year 2011, including the month of August 2011. For each segment, the average speed was reported at a 5-minute sampling interval. There is 3% of missing data in the data set at hand.

First we will apply CUR decomposition for the purpose of compression: we extract a low-dimensional representation of the road network by applying CUR decomposition to data of the entire traffic network [4-6]. To this end, we sample data from a subset of road segments and time instances, resulting in the $\mathbf{C}$ and $\mathbf{R}$ matrix respectively. By learning a suitable mixing matrix $\mathbf{U}$, we then extrapolate the state of the whole network by applying the CUR decomposition. The matrices $\mathbf{C}$, $\mathbf{U}$, and, $\mathbf{R}$ form a compact representation of the traffic

data, and can be used as a compression scheme. Our numerical results show that the resulting compression rates are slightly worse compared to PCA, while leading to more interpretable results.

Next, we apply the CUR decomposition for compressed sensing of the network-wide traffic state. In this setting, the matrices **U** and **R** are learned from historical data. The matrix **C** is filled with new data from a "small" subset of links. We infer the state of the entire traffic network through the CUR decomposition. In other words, the CUR method allows us to extrapolate the state of the whole network from measurements at a small number of links (arranged in matrix **C**). As demonstrated by our numerical results, the proposed compressed sensing scheme, derived from the CUR matrix decomposition, yields accurate estimates of the global traffic state, even from measurements at a limited number of links. Taken together, our results seem to suggest that the CUR matrix decomposition may find exciting applications in the realm of Intelligent Transportation Systems.

The paper is structured as follows. In Section II, we briefly describe the traffic network considered here. In Section III, we briefly review the CUR matrix decomposition, and Section IV and V, we explain how we apply CUR decomposition for compression and compressed sensing respectively of traffic network data. In Section VI, we provide numerical results. In Section VII, we summarize our contributions and suggest topics for future work.

II. EXPERIMENTAL SETUP

We consider here traffic speed data from the southeastern part of Singapore (Outram to Changi area). In Fig. 1, we show the macro and micro location of the network.

Figure 1: The traffic network studied in this paper, containing 6024 links. (Top left) Location of Singapore (red dot); (Top right) Traffic network in Singapore; the subnetwork studied here is marked by the red dotted line. (Bottom) Close-up of that same subnetwork.

The network includes n = 6024 links, including two expressways (East Coast Park and Pan Island Expressway) at its peripheries. The rest of the network consists of arterials and local roads in downtown area and around Changi Airport.

For this study, we selected traffic speed data for the month of August 2011, with a sampling interval of 5 minutes. The speed values represent the average speed of all vehicles which traverse a link during the given sampling interval. About 3% of the data is missing, due to issues with the sensors or other reasons. For the CUR analysis, we represent the data set as a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, where each column represents the average speed data from a particular road segment (link) during the month of August 2011. Each row of matrix **A** corresponds to a 5-minute time interval.

III. CUR MATRIX DECOMPOSITION

In this section, we briefly describe different CUR matrix decomposition strategies for low-rank approximation of a given matrix **A**. To carry out the CUR decomposition, we extract a subset of columns and rows of matrix **A** through random sampling. We store the subset of columns (road segments) and rows (time instances) in the matrices **C** and **R** respectively. Next we learn a suitable mixing matrix **U**, and create a low-dimensional representation of network matrix **A** as the product:

$$\mathbf{A} \approx \hat{\mathbf{A}} = \mathbf{CUR} . \qquad (1)$$

More specifically, we can divide CUR decomposition in following three steps:

(1) Calculate the importance of each column/row in the network matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$. Each row/column in the matrix **A** is assigned a probability of being selected.

(2) Construct the matrices $\mathbf{C} \in \mathbb{R}^{m \times c}$ and $\mathbf{R} \in \mathbb{R}^{r \times n}$ by randomized sampling, using the probabilities computed in the first step. We sample c columns from matrix **A**, without replacement, and arrange them in matrix **C**. We create the matrix **R** in similar manner from r randomly sampled distinct rows of **A**. We refer to corresponding index set of columns and rows as $\mathscr{C}$ and $\mathscr{R}$ respectively.

(3) Compute the regularization matrix **U**:

$$\mathbf{U} \cong \mathbf{C}^{+} \mathbf{A} \mathbf{R}^{+}, \qquad (2)$$

where $\mathbf{C}^{+}$ and $\mathbf{R}^{+}$ are the pseudo-inverses of **C** and **R** respectively [13]. Usually, the rows and columns of $\hat{\mathbf{A}}$ associated with the index sets $\mathscr{C}$ and $\mathscr{R}$ are not identical to the corresponding columns and rows in **A** (stored in the matrices **C** and **R** respectively). Therefore, after computing $\hat{\mathbf{A}}$, we replace those columns and rows of $\hat{\mathbf{A}}$ by the corresponding ones in **A** (contained in **C** and **R** respectively).

We propose two different strategies for the first step. The other two steps are the same for both strategies. In the first strategy, we assign sampling probabilities based on the energy (L_2 norm) of the columns and rows. This strategy is referred to as Energy-CUR (E-CUR). In the second strategy, we apply SVD to assign probabilities to each column and row; we refer to this approach as SVD-CUR.

A. E-CUR

We assign a selection probability to each column (link) and row (time instance) of **A** derived from their energy (L_2

norm). The underlying idea is that if the traffic speed is high at a link, that link will have high probability to be selected for matrix **C**. We define sampling probabilities p_i and p_j for each row i and column j respectively of the matrix **A** as:

$$p_i = \frac{\sum_{j=1}^{n} \mathbf{A}(i,j)^2}{\sum_{i=1}^{m}\sum_{j=1}^{n} \mathbf{A}(i,j)^2} \quad , \tag{3}$$

where $i = 1, 2, …, n$, and

$$p_j = \frac{\sum_{i=1}^{m} \mathbf{A}(i,j)^2}{\sum_{i=1}^{m}\sum_{j=1}^{n} \mathbf{A}(i,j)^2} \quad , \tag{4}$$

where $j = 1, 2, …, m$.

B. SVD-CUR

In this approach, we first perform Singular Value Decomposition (SVD) of the network matrix **A** [14]:

$$\mathbf{A} = \mathbf{X}\mathbf{S}\mathbf{Y}^{\mathbf{T}}, \tag{5}$$

where $\mathbf{X} \in \mathbb{R}^{m \times m}$ and $\mathbf{Y} \in \mathbb{R}^{n \times n}$ are unitary matrices whose columns are left and right singular eigenvectors respectively of **A**. The matrix $\mathbf{S} \in \mathbb{R}^{m \times n}$ contains the singular values of **A**. We obtain rank k SVD approximation of **A** by truncating original matrices (**X**, **S** and **Y**) by:

$$\mathbf{A}_{mxn} = \mathbf{X}_{mxk}\mathbf{S}_{kxk}\mathbf{Y}_{kxn}^{\mathbf{T}}. \tag{6}$$

We now calculate selection probability of column j of matrix **A** as:

$$p_j = \frac{1}{k}\sum_{i=1}^{k} \mathbf{y}_{ij}^2, \tag{7}$$

where $\mathbf{y}_{ij}$ is j-th coordinate of i-th right singular vector.

Since the matrix **Y** is unitary, the quantities p_j sum to one, and hence the p_j can indeed be considered as probabilities. Similarly, we define the selection probability p_i of row i of matrix **A** as:

$$p_i = \frac{1}{k}\sum_{j=1}^{k} \mathbf{x}_{ij}^2. \tag{8}$$

IV. COMPRESSION BY CUR DECOMPOSITION

As a first application, we use the CUR decomposition to compress large-scale traffic data. The traffic data is arranged in a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$. By sampling rows and columns from **A**, we obtain matrices **C** and **R** matrix respectively. By learning a suitable mixing matrix **U**, we then approximate the matrix **A** by its CUR decomposition. The matrices **C**, **U**, and, **R** form a compact representation of the traffic data in the matrix **A**, and can be used for compression. Instead of storing the large matrix **A**, the smaller matrices **C**, **U**, and, **R** are stored, leading to compression. To assess the resulting

compression scheme, we compute the Percent-Root-mean square Distortion (PRD):

$$PRD(\%) = \frac{\|\mathbf{A} - \mathbf{CUR}\|_F}{\|\mathbf{A}\|_F} \tag{9}$$

where

$$\|\mathbf{A}\|_F = \left(\sum_i \sum_j a_{i,j}^2\right)^{\frac{1}{2}} \tag{10}$$

The Percent-root-mean square distortion (PRD) is a common performance measure to compare the reconstruction performance of low-dimensional models [15]. Interestingly, after decompression, we obtain the exact measurements without errors for some links (corresponding to **C**); therefore, it is crucial to select the links associated with **C** in a careful manner. As a rule of thumb, the main highways should be included in **C.** In other subspace methods such as PCA on the other hand, the reconstruction error generally affects all links.

We calculate the Compression Ratio (CR) as the ratio of the total number of elements in the original (uncompressed) matrix **A** and the total number of elements in the low-rank approximation. In the CUR decomposition, we approximate the network matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ by matrices $\mathbf{C} \in \mathbb{R}^{m \times c}$, $\mathbf{U} \in \mathbb{R}^{c \times r}$ and $\mathbf{R} \in \mathbb{R}^{r \times n}$**.** The compression ratio in this case equals:

$$CR = \frac{mn}{mc + cr + rn} \tag{11}$$

where c and r represent the number of road segments (columns) and time instances (rows) stored in the low-rank approximation. Interestingly, the CR (11) for CUR decomposition can be improved as follows. In CUR decomposition, we store a subset of the original columns and rows of the data matrix **A**. It is noteworthy that the matrices **C** and **R** have cr elements in common. We can improve the compression efficiency by storing the $r+c$ positions of these columns and rows, instead of the rc redundant elements. As a result, the compression ratio CR for the CUR decomposition becomes:

$$CR = \frac{mn}{mc + cr + (rn - cr) + (c + r)} \tag{12}$$

We will use the definition (12) of CR for CUR in the rest of the paper.

V. COMPRESSED SENSING BY CUR DECOMPOSITION

As a second application, we use the CUR decomposition for compressed sensing of a traffic network. The objective is to infer the traffic speed in the whole traffic network from measurements of the traffic speed at a few links. Also for this problem, the CUR matrix decomposition is directly applicable. Assume that measurements are available from a small set of links. The matrix **C** is filled with those measurements, where each column of **C** corresponds to a

link. The matrices **U** and **R** are determined from historical data; those matrices are *not* learned from the new measurements. We infer the state of the entire traffic network through the CUR decomposition (1). In other words, the CUR method allows us to extrapolate the state of the whole network from measurements at a small number of links (arranged in matrix **C**).

There are two main differences with the problem of compression (cf. Section IV):
1. We now do *not* have access to traffic speed data at all links. Both matrices are learned from data set at hand.
2. The matrices **U** and **R** are *not* learned each time from the new measurements. They are fixed for the purpose of compressed sensing. The product **UR** can be viewed as an extrapolation matrix, that maps matrix **C** to **A**.

Obviously, this same method can also be used for compression, where the matric **C**, obtained by compressed sensing, is stored instead of **A**. It is noteworthy that other subspace methods such as PCA cannot be applied for compressive sensing in this manner, since the basis vectors usually involve a large number of links.

The proposed CUR-based procedure for compressed sensing may substantially reduce the resources required for data acquisition, as it only requires data to be collected at a small subset of links. The proposed compressed sensing procedure is especially relevant for real-time acquisition and processing of large-scale traffic data. Specifically, compressed sensing may allow us to infer, monitor, and predict the state of a large traffic network in real-time, while avoiding potential bottlenecks due to limited bandwidth and computational resources.

VI. RESULTS

In this section, we analyze the performance of the proposed CUR method for compression and compressive sensing of large-scale traffic data. We consider here the traffic network depicted in Fig. 1; the traffic data consists of the average traffic speed at each link in that network during the month of August 2011, acquired at 5-minute intervals. First we investigate the underlying spatial and temporal patterns associated with the low-dimensional CUR models. Next we assess the performance of the proposed CUR-based compression scheme, and compare it to PCA. At last, we evaluate the proposed CUR-based compressed sensing scheme, and again use PCA as benchmark.

A. Temporal patterns in CUR decomposition

In Fig. 2, we show the importance (selection probability) of each time instance as determined by the E-CUR and SVD-CUR strategies. The colors represent the selection probabilities, such that black corresponds to a value of one. On the *x*-axis, we display different days for the month of August 2011. We limit ourselves here to the first four weeks (28 days). The *y*-axis shows different time instances for each day, starting from midnight (00:00 AM). Each time instance represents an interval of 5 minutes.

Fig. 2a shows the selection probability computed by E-CUR. The E-CUR strategy gives significant importance to time instances at which the average traffic speed is high.

Fig.2a shows that large values of traffic speed occur during the early morning and late night hours. These are usually the time periods with lowest traffic. We also observe that the traffic tends to be smooth during the weekends (6th, 7th, 13th day, etc.). Interestingly, we observe a slightly unique pattern on the 9th of August, which is a public holiday is Singapore. For the sake of simplicity, we represented the missing entries as zero in the data set. (In a more extended study in the future, we will impute the missing data.) These instances are represented by blue color (see Fig. 2a). E-CUR correctly assigned least importance to such time instances.
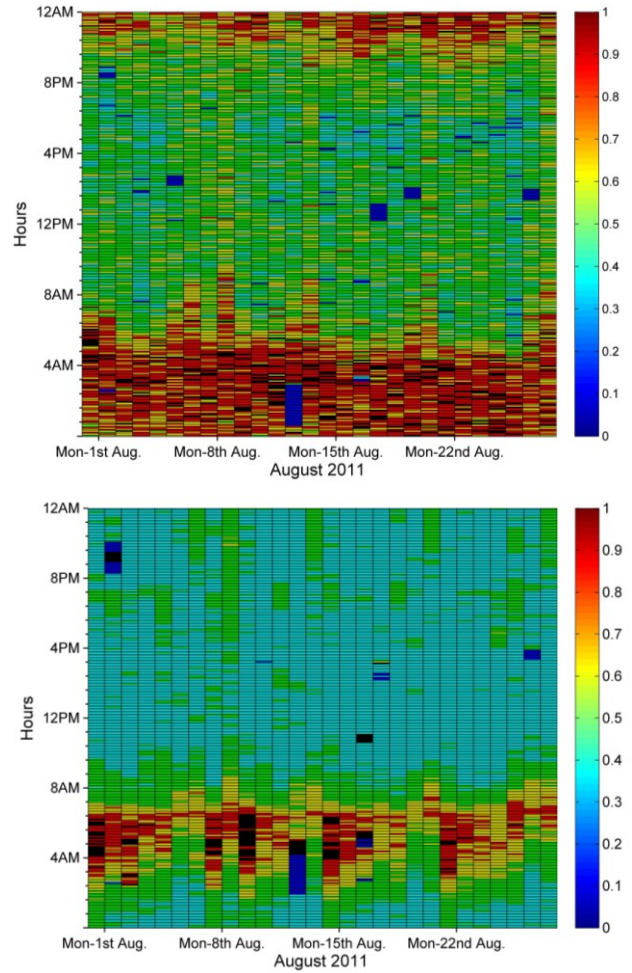

Figure 2: Selection probability of time instances (normalized to one). The colors represent the selection probability for each time instance assigned by E-CUR (top) and SVD-CUR (bottom).

We show the temporal patterns for SVD-CUR in Fig. 2b. The SVD-CUR strategy tries to find correlations between different time instances. For instance, if large a number of time instances follow similar traffic patterns then we can easily represent them by a few components. However, if a certain time instance does not conform to normal traffic behavior, then we assign a higher weight to that time instance. From Fig. 2b, we can see that the traffic is mostly irregular during the morning transition period, when traffic conditions change from free flow traffic to high-density traffic during the morning peak period. During the rest of the

day, the traffic patterns are more stable, and therefore, the corresponding time instances receive smaller weights. We also observe that time instances during evening peak hours are given higher importance. Fig. 2b shows that the SVD-CUR method can also reveal uncommon traffic patterns such as the one during the public holiday on the 9th of August. At the top of Fig. 2b, green regions appear only on Sundays and the public holiday, suggesting that traffic conditions during these days follow unusual patterns. Similarly, anomalies such as missing data can also be easily spotted: SVD-CUR also assigns least importance to time periods with missing data.

B. Spatial patterns in CUR decomposition

In Fig. 3, we display the importance (selection probability) of each link in the network, as determined by the E-CUR and SVD-CUR strategies. The colors represent the normalized weights assigned by each algorithm.
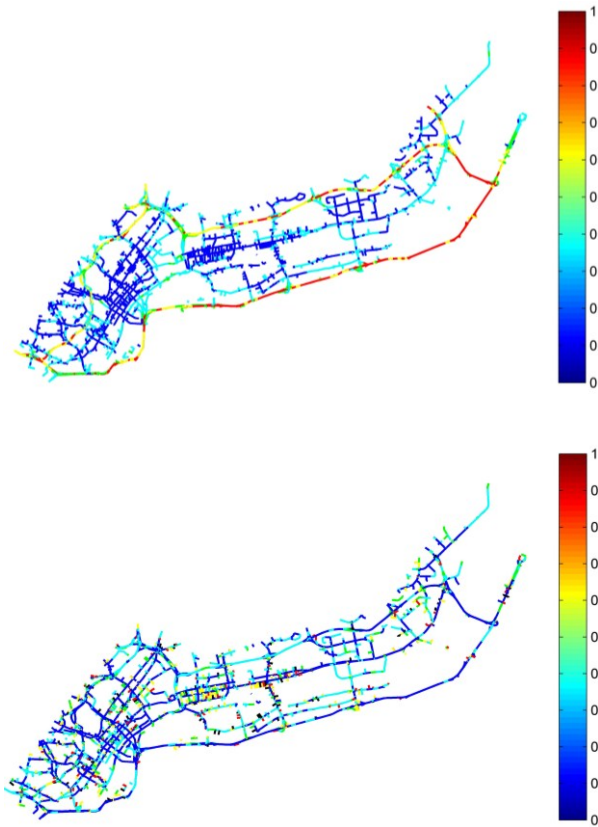

Figure 3: Selection probability of links (normalized to one). The colors represent the selection probability for each link assigned by E-CUR (top) and SVD-CUR (bottom).

In Fig. 3a we show the importance of the different links in the network as determined by E-CUR. The links with high values of traffic speed are given higher importance. Most of these links belong to expressways. Arterials and roads in the vicinity of the central business district are considered less important by E-CUR. We can use such reduced network representation to make applications such as route guidance more scalable. On the other hand, if arterials and roads in the vicinity of the central business district are strongly relevant for the application at hand, the SVD-CUR method is recommended. In Fig. 3b we show the results for the SVD-CUR method. Links with unusual temporal patterns are given higher scores, which tend to be the roads of lower hierarchical level. On the other hand, most expressways and arterial streets often have regular temporal patterns, and are therefore given less weight. However, there are some expressway segments with larger temporal variations (and hence weights) than others. These segments are typically located in the vicinity of the central business district (CBD).

C. Compression by CUR decomposition

We now compare the compression efficiency of the proposed CUR-based low-dimensional models. To this end, we compute the Percent-Root-mean square Distortion (PRD) for different compression ratios (CR). We use PCA as a benchmark, as it is considered as the optimal subspace transformation.
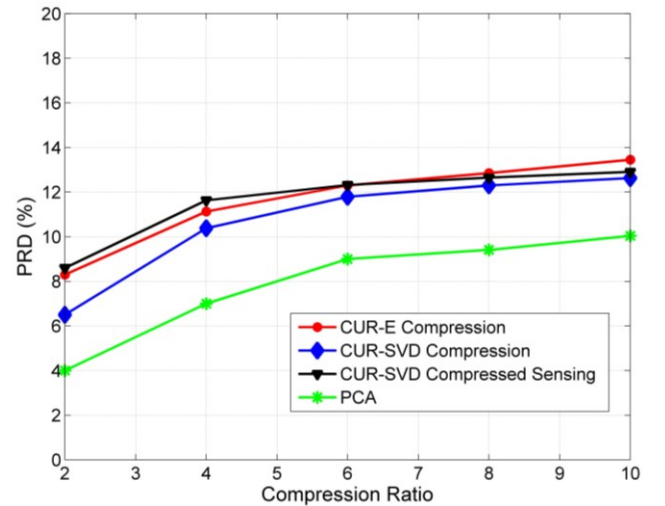

Figure 4: Performance of CUR for compression and compressed sensing. Results are shown for traffic speed data of August 2011, for the following methods: E-CUR and SVD-CUR compression scheme, PCA, and SVD-CUR compressed sensing.

Fig. 4 shows that the two CUR-based compression schemes perform slightly worse than PCA. SVD-CUR provides lower construction error than E-CUR, which may be explained by the fact that the former takes various kinds of links into account, whereas the latter mostly retains expressways. We observe from Fig. 4 that the CUR methods provide accurate low-dimensional representation of large networks. The main advantage of such methods compared to PCA lies in the fact that they offer us insight into the underlying spatial and temporal patterns in large networks. Another advantage is their potential for compressed sensing, which is discuss in the next section.

D. Compressive Sensing by CUR decomposition

Following the procedure outlined in Section V, we use the SVD-CUR method to reconstruct the traffic data of the month of August 2011 from measurements at a small set of links. Measurements at a small number of links are arranged in a matric **C**. The matrices **U** and **R** are computed from an

earlier month, with the same choice of column and row indices $\mathscr{C}$ and $\mathscr{R}$ respectively.

Fig. 5 shows the reconstruction error of the SVD-CUR method in the setting of compressed sensing. The results are only slightly worse compared to the setting of compression, where data for the entire network is available and the matrices **U** and **R** are learned from the same month (August 2011). As pointed out earlier, compressed sensing can greatly reduce the amount of sensors required to infer the state of large traffic networks, and consequently, it may prove to be worthwhile to tolerate the incurred minor loss in reconstruction accuracy.

In the formulation of compressive sensing, the compression ratio (CR) for CUR-SVD is inversely proportional to the number of sampled road segments. Table 1 shows the relationship between the fraction of measured links and the achieved compression efficiency for CUR-SVD. Similarly, Table II provides the corresponding results for PCA. For the latter approach, all links need to be probed, and hence it cannot be used as such for compressed sensing.

TABLE I: Results for compressed sensing through the CUR-SVD method.

Number of the road segments	300	600	1000	1500	3000
Portion of the network	5%	10%	16%	25%	50%
CR	20	10	6	4	2
PRD (%)	13.71	12.91	12.32	11.63	8.60

TABLE II: Compression efficiency of PCA

Number of the road segments	6024	6024	6024	6024	6024
Portion of the network	100%	100%	100%	100%	100%
CR	20	10	6	4	2
PRD (%)	11.51	10.05	8.50	7.00	4.00

VII. Conclusions and Future Work

In this paper, we applied the CUR matrix decomposition for generating low-dimensional representations of large-scale traffic networks. In those subspace representations, the basis vectors correspond to individual links and time instances. Consequently, those low-dimensional models can easily be interpreted, in contrast to PCA and related extensions. We experimented with two strategies to select individual links and time instances, leading to intuitive outcomes: These strategies provide details about traffic patterns on weekdays, weekends and on public holidays. Moreover, they provided useful spatial information.

We derived a compression scheme from the CUR matrix decomposition, and applied it to speed data from a large traffic network in Singapore. Our numerical results suggest that the scheme has favorable compression. Furthermore, we

designed an algorithm for compressed sensing from the CUR matrix decomposition. Also for that approach, we have obtained encouraging numerical results.

Taken together, the results presented in this paper pave the way for further applications of the CUR matrix decomposition for Intelligent Transportation Systems (ITS). In future work, we will explore applications to traffic prediction and routing. Moreover, interesting alternatives and extensions of CUR matrix decomposition have been proposed, which could be investigated in the context of ITS.

References

[1] J. Aslam, S. Lim, X. Pan, and D. Rus. "City-scale traffic estimation from a roving sensor network." in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, pp. 141-154. ACM, 2012.

[2] Y. F. Han and F. Moutarde. "Analysis of Network-Level Traffic States Using Locality Preservative Non-Negative Matrix Factorization." in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC),* pp. 501-506. IEEE, 2011.

[3] Q. Li, H.Jianming, and Z. Yi. "A flow volumes data compression approach for traffic network based on principal component analysis." in *ITSC 2007 IEEE Intelligent Transportation Systems Conference,* pp. 125-130. IEEE, 2007.

[4] M. W. Mahoney and P. Drineas. "Cur Matrix Decompositions for Improved Data Analysis." *Proceedings of the National Academy of Sciences of the United States of America 106*, no. 3 (Jan 2009): 697-702.

[5] P. Drineas, M. W. Mahoney, and S. Muthukrishnan. "Relative-error CUR matrix decompositions." *SIAM Journal on Matrix Analysis and Applications 30*, no. 2 (2008): 844-881

[6] J. Sun, Y. Xie, H. Zhang, and C.Faloutsos. "Less is more: Compact matrix decomposition for large sparse graphs." (2007). CMU Computer Science Department Paper, Paper 532.

[7] Y. Liu and J. Shao. "High dimensionality reduction using CUR matrix decomposition and auto-encoder for web image classification." in *Advances in Multimedia Information Processing-PCM 2010*, pp. 1-12. Springer Berlin Heidelberg, 2011.

[8] S. Wang, Z. Zhang, and J. Li. "A Scalable CUR Matrix Decomposition Algorithm: Lower Time Complexity and Tighter Bound." *arXiv preprint arXiv:1210.1461* (2012).

[9] T. Djukic, , J. WC van Lint, and S. P. Hoogendoorn. "Application of Principal Component Analysis to Predict Dynamic Origin-Destination Matrices." *Transportation Research Record: Journal of the Transportation Research Board 2283*, no. 1 (2012): 81-89.

[10] T. Djukic, G.Flotterod, H. van Lint, and S.Hoogendoorn. "Efficient real time OD matrix estimation based on Principal Component Analysis." in *2012 15th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 115-121. IEEE, 2012.

[11] M. T. Asif, K.Srinivasan, J.Dauwels, and P.Jaillet. "Data Compression Techniques for Urban Traffic Data." SCCI 2013, 2013 *IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems,* accepted.

[12] X. Jin, Y. Zhang, and D. Yao. "Simultaneously prediction of network traffic flow based on PCA-SVR." in *Advances in Neural Networks–ISNN 2007*, pp. 1022-1031. Springer Berlin Heidelberg, 2007.

[13] G. Golub and W. Kahan. "Calculating the singular values and pseudo-inverse of a matrix." *Journal of the Society for Industrial & Applied Mathematics, Series B: Numerical Analysis 2*, no. 2 (1965): 205-224.

[14] G. H. Golub and C.Reinsch. "Singular value decomposition and least squares solutions." *NumerischeMathematik 14*, no. 5 (1970): 403-420.

[15] J. Dauwels, K. Srinivasan, M. R. Reddy, and A. Cichocki. "Multi-channel EEG compression based on matrix and tensor decompositions." in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* pp. 629-632. IEEE, 2011.