Pastor Figueroa, Giancarlo

A Low-rank Tensor Model for Imputation of Missing Vehicular Traffic Volume

Published in:
IEEE Transactions on Vehicular Technology

DOI:
10.1109/TVT.2018.2833505

Published: 01/09/2018

Document Version
Publisher's PDF, also known as Version of record

Correspondence

A Low-Rank Tensor Model for Imputation of Missing Vehicular Traffic Volume

Giancarlo Pastor

***Abstract*—This paper presents a low-rank tensor model for vehicular traffic volume data. Contrarily to previous works, we capitalize on a definition of rank, called the tensor train, that is as effective as possible; so that it exploits all the correlation between local structures that are present in the multiple modes, but practical enough that efficient optimization algorithms still hold. From our model, a formulation to find balanced (higher order) tensors is derived. The resulting optimally-balanced tensor improves the imputation accuracy of the tensor train rank. Then, we design specific experiments, which are numerically evaluated using real-world traffic data from Tampere city, Finland. The experimental results are promising, our proposed approach outperforms existing algorithms in both imputation accuracy and, in some instances, computation time.**

***Index Terms*—Crowdsensing, crowdsourcing, data imputation, missing data, tensor completion, transportation systems.**

I. Introduction

Understanding vehicular traffic dynamics is essential to build next intelligent transportation systems. Therefore, different sensor technologies have been deployed to support traffic monitoring, from simple and low-cost inductive loops to complex and costly video cameras. In this paper, independently of the monitoring technology, we study the imputation of missing (or potentially corrupted) traffic volume data, which is inspired by two motivating scenarios [1]:

1) Data is missed (or corrupted) possibly due to sensor failure or by nearby events affecting the entire monitoring infrastructure, such as construction sites.
2) Data is intentionally missed with the aim to save in resources or to prevent poor user experience of mobile users contributing to crowdsourcing.

The first motivating scenario is not rare nowadays. However, the second scenario belongs to a new emerging design trend on large sensing systems. This new trend exploits efficient representations of real-world physical phenomena in appropriate information domains. In practice, both motivating scenarios are linked as they both require data recovery or imputation support. For vehicular monitoring systems, different imputation methodologies have been proposed, from classical compressive sensing [1] to completion of matrices [2] and tensors [3]; all of these sharing common information theoretical grounds. Thus, it is of significant interest to evaluate the feasibility of such innovative formulations that promise highly-efficient sensor designs.

In this paper, we focus on tensor completion [7] which may have practical advantages *with respect to* (w.r.t.) other imputation methodologies [8] (as we will discuss later in the paper). Its applications to vehicular systems have been introduced, and models progressed, notably by the same leading group. Table I summarizes our literature review. In the seminal work [3], and subsequent studies [4]–[6], authors explored the capabilities of tensor representations and solved their derived formulations using potent optimization algorithms (see Table I for details). All these works capitalized on the efficient Tucker factorization [9], or targeted minimization of the Tucker or trace rank (after matricization of tensors).

Manuscript received October 11, 2017; revised February 21, 2018; accepted April 23, 2018. Date of publication July 2, 2018; date of current version September 17, 2018. This work was supported in part by the Finnish Cultural Foundation under Grant 00160747, and in part by the Academy of Finland under Grant 294763. The review of this paper was coordinated by Dr. M. M. Atia.

The author is with the Department of Communications and Networking, Aalto University School of Electrical Engineering, Espoo 02150, Finland (e-mail: giancarlo.pastor@aalto.fi).

Digital Object Identifier 10.1109/TVT.2018.2833505

TABLE I
Summary of Literature Review on Tensor Completion for Vehicular Sensor Systems

Reference	Minimization objective	Optimization algorithm
[3], 2013	Tucker factorization error	gradient descent
[4], 2014	Tucker factorization error with model regularization	gradient descent on Grassmann manifold
[5], 2016	Tucker (matrix) rank of modes matricizations	fixed-point continuation
[6], 2016	Tucker (tensor) rank	alternating direction method of multipliers
This paper	Tensor train factorization error with balanced modes	block coordinate descent

Differently from previous approaches in tensor completion [10], the tensor train factorization and rank have been recently proposed in [11], as they can exploit the local structure of data by well-balanced matricization schemes. In this paper, we build on top of these recent advances to address the vehicular crowdsensing design challenge. We represent (incomplete) vehicular traffic volume data using tensors and conduct extensive imputation experiments on real-world traffic data from Tampere city, Finland. The following are our main contributions:

1) We introduce the **tensor train factorization and rank for imputation of missing vehicular traffic volume**, which reduces the error in ~40% w.r.t. state-of-the-art Tucker rank based algorithms for vehicular traffic.
2) We introduce a **simple methodology to find balanced high-order tensors**, which further reduces the error in ~15% (~50%) w.r.t. tensor train rank based algorithms (Tucker, resp.).

The paper is organized as follows: Section II includes fundamental concepts and algorithms, Section III describes the data and proposed imputation methodology, Section IV reports experiments, Section V discusses our planned future work, and Section VI concludes.

II. On Tensors

This Section covers general concepts on top of which we describe our data representation and imputation methodology.

A. *Tensors for High-Dimensional Data Representation*

Let $\mathcal{X} \in \mathbb{R}^{m_1 \times \cdots \times m_n}$ be a n-order or n-mode real-valued tensor, with mode configuration $\mathbf{m} = (m_1, \ldots, m_n)$, and elements $x_{i_1 \ldots i_n}$, where $i_k \in \{1, \ldots, m_k\}$ for $k = 1, \ldots, n$:

- A real number $x \in \mathbb{R}$ is a 0-order tensor (convention).
- A real-valued vector $\mathbf{x} \in \mathbb{R}^{m_1}$ is a 1-order tensor.
- A real-valued matrix $\mathbf{X} \in \mathbb{R}^{m_1 \times m_2}$ is a 2-order tensor.

It is sometimes convenient to *unfold* a tensor into a matrix. The k-th mode unfolding, matricization or flattening operation of a tensor $\mathcal{X}$ is defined as

$$\mathsf{unfold}_k(\mathcal{X}) = \mathcal{X}_{(k)} \in \mathbb{R}^{m_k \times (m_1 \times \cdots \times m_{k-1} \times m_{k+1} \times \cdots \times m_n)}. \quad (1)$$

The opposite operation, *fold*, is defined as

$$\mathsf{fold}_k(\mathcal{X}_{(k)}) = \mathcal{X}. \quad (2)$$

B. Low-Rank Tensor Completion (LRTC)

Let $\mathcal{Y} \in \mathbb{R}^{m_1 \times \cdots \times m_n}$ be a tensor containing observations of $\mathcal{X}$ in the index set $\Omega \subset \{1, \ldots, m_1\} \times \cdots \times \{1, \ldots, m_n\}$. The tensor completion problem consists of recovering $\mathcal{X}$ of minimum complexity, measured in terms of the (tensor) rank, given partial observations $\mathcal{Y}_\Omega$ of it,

$$\begin{cases} \min_{\mathcal{X}} & \operatorname{rank}(\mathcal{X}) \\ \text{s.t.} & \mathcal{X}_\Omega = \mathcal{Y}_\Omega. \end{cases} \quad (3)$$

The rationale behind (3) is the inherent structure of high-dimensional data in appropriate representations [12], which is revealed via tensor factorizations or decompositions. However, problem (3) is combinatorial hence NP-hard [13], thus it was relaxed in [10] to,

$$\begin{cases} \min_{\mathcal{X}} & \|\mathcal{X}\|_{\mathsf{T}} = \sum_{k=1}^{n} \|\mathcal{X}_{(k)}\|_{\mathsf{tr}} \\ \text{s.t.} & \mathcal{X}_\Omega = \mathcal{Y}_\Omega, \end{cases} \quad (4)$$

where $\|.\|_{\mathsf{T}}$ is the *Tucker rank* (also known as trace or nuclear tensor norm) [10], and $\|.\|_{\mathsf{tr}}$ is the trace or nuclear matrix norm. Essentially, the **Tucker rank is computed from the ranks of matrices constructed based on unbalanced matricizations**, that is one mode versus the rest [11].

Another type of tensor rank is the *tensor train* (TT) [14], denoted $\|.\|_{\mathsf{TT}}$. The **tensor train rank consists of ranks of matrices formed by a well-balanced matricization scheme**, i.e. matricize the tensor along permutations of modes. Hence, instead of problem (4), authors of [11] formulated,

$$\begin{cases} \min_{\mathcal{X}} & \|\mathcal{X}\|_{\mathsf{TT}} = \sum_{k=1}^{n-1} \|\mathcal{X}_{[k]}\|_{\mathsf{tr}} \\ \text{s.t.} & \mathcal{X}_\Omega = \mathcal{Y}_\Omega, \end{cases} \quad (5)$$

where $\mathcal{X}_{[k]} = U\Lambda^{[k]}V^T \in \mathbb{R}^{\tilde{m}_1 \times \tilde{m}_2}$ is the $\{1, \ldots, k\}$-mode matricization of tensor $\mathcal{X}$, with $\tilde{m}_1 = \prod_{j=1}^{k} m_j$ and $\tilde{m}_2 = \prod_{j=k+1}^{n} m_j$, matrices $\mathbf{U} \in \mathbb{R}^{\tilde{m}_1 \times r_k}$ and $\mathbf{V} \in \mathbb{R}^{\tilde{m}_2 \times r_k}$ are orthogonal, and rank $r_k \leq \min(\tilde{m}_1, \tilde{m}_2)$ is the number of non-vanishing singular values of $\Lambda^{[k]}$.

C. Two Efficient LRTC Algorithms via the Tucker Rank

This last part of the Section describes two efficient LRTC algorithms, both based on the Tucker rank.

1) Tucker Rank + (Smoothing via) Nesterov (FaLRTC [10]): Convert (4) into a smooth problem following,

$$\begin{cases} \min_{\mathcal{X}} & \sum_{k=1}^{n} \max_{\|\mathcal{Z}_{k(k)}\| \leq 1} \langle \mathcal{X}, \mathcal{Z}_i \rangle - \frac{\mu}{2} \|\mathcal{Z}_i\|_2 \\ \text{s.t.} & \mathcal{X}_\Omega = \mathcal{Y}_\Omega, \end{cases} \quad (6)$$

whose gradient can be efficiently computed [10].

2) Tucker Rank + ADMM (HaLRTC [10]): First, write the equivalent problem to (4),

$$\begin{cases} \min_{\mathcal{X}, \{\mathcal{M}_k\}} & \sum_{k=1}^{n} \|\mathcal{M}_{k(k)}\|_{\mathsf{tr}} \\ \text{s.t.} & \mathcal{X}_\Omega = \mathcal{Y}_\Omega, \\ & \mathcal{X} = \mathcal{M}_k, \quad k = 1, \ldots, n, \end{cases} \quad (7)$$

and apply the Alternating Direction Method of Multipliers (ADMM) to the augmented Lagrangian function, whose terms can be efficiently updated [10].

The above algorithms, both based on the Tucker rank, produce excellent results both in terms of accuracy and computing time. Next Section however will describe a recently proposed LRTC algorithm based on the tensor train rank and which has reported improved performance. Other advanced LRTC algorithms, applied to different fields, will be also evaluated, namely STDC: Tucker factorization error via maximum a posteriori with Inexact Augmented Lagrange Multiplier (IALM) [15], and SPC: PARAFAC factorization error via model regularization with Hierarchical Alternating Least Squares (HALS) [16].

III. Imputation of Missing Vehicular Traffic Volume

This Section describes the dataset, a recently-proposed efficient LRTC algorithm based on the tensor train rank, and a key component of it.

A. The Tampere (City) Traffic Dataset

The data consist of vehicular traffic volume measurements from Tampere city, Finland. The measurements correspond to 21 roads around the city. In each road, inductive loop sensors are deployed to count the volume of vehicles circulating the roads in both directions, in sampling intervals of five minutes. The measurements are collected along three years, January 2014 to December 2016.

The data exploration reveals moderate to significantly long intervals of missing measurements and anomalies, the latter possibly due to construction sites nearby the roads. Consequently, the dataset is pruned to keep as many consecutive weeks of uninterrupted, and possibly regular, measurements along year 2016. The resulting dataset can be summarized as

- $m_1 = 32$ weeks, week 1 to week 32
- $m_2 = 24$ hours per day
- $m_3 = 16$ roads
- $m_4 = 7$ days per week
- $m_5 = 2$ directions per road

which is represented by a 5-order tensor $\mathcal{X} \in \mathbb{R}^{32 \times 24 \times 16 \times 7 \times 2}$, with (order $n_0 = 5$ and) mode configuration $\mathbf{m}_0 = (32, 24, 16, 7, 2)$. Fig. 1 illustrates the measurements corresponding to one road and one direction, during the first seven days of year 2016.

The reason to prune the dataset is to evaluate the algorithms' performance against ground-truth and regular measurements (read stable traffic conditions). For the anomalies however, although these measurements could be detected and pruned for subsequent imputation (with solutions following regular traffic patterns), their analysis is postponed, as well as the effect of yearly data, to our next work.

B. An Efficient LRTC Algorithm via the Tensor Train Rank

Very recently, TMac-TT [11] reformulated problem (5) as,

$$\begin{cases} \min_{\{U_k, V_k\}, \mathcal{X}} & \sum_{k=1}^{n} \|U_k V_k - \mathcal{X}_{[k]}\|_2^2 \\ \text{s.t.} & \mathcal{X}_\Omega = \mathcal{Y}_\Omega, \end{cases} \quad (8)$$

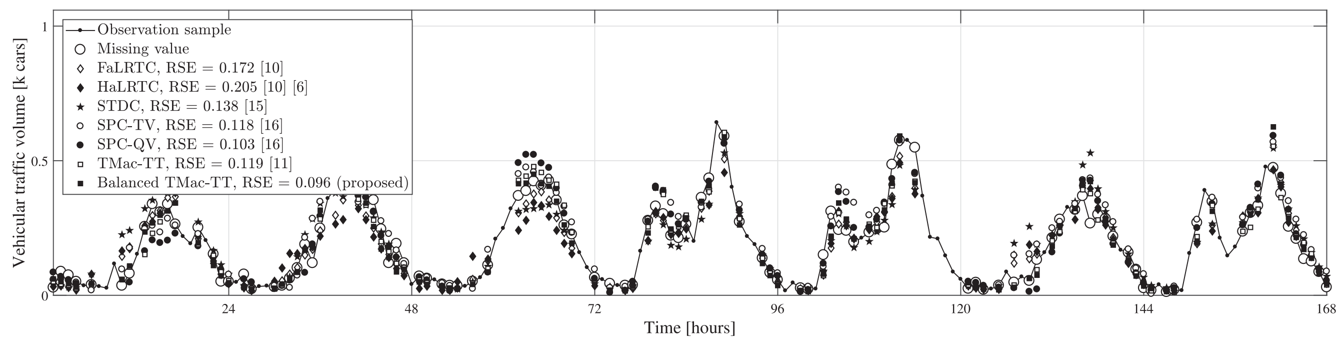

Fig. 1. Example of data imputation. Tensor completion algorithms based on the Tucker, PARAFAC and tensor train are compared. Data correspond to one road and one direction, during the first seven days of year 2016. The total amount of missing data is 70%.

which resembles a model previously proposed by [4] for the analysis of vehicular traffic. The new problem (8) is convex when its variables are separately optimized, thus Block-coordinate Descent (BCD) can be applied [17].

At iteration t and mode $k = 1, \ldots, n-1$, solve on separate variables of each summation term in (8), following [18],

$$U_k^{(t)} = \mathcal{X}_{[k]}^{(t-1)} (V_k^{(t-1)})^T \tag{9}$$

$$V_k^{(t)} = ((U_k^{(t)})^T U_k^{(t)})^{\dagger} (U_k^{(t)})^T \mathcal{X}_{[k]}^{(t-1)} \tag{10}$$

$$\mathcal{X}_{[k]}^{(t)} = U_k^{(t)} V_k^{(t)} \tag{11}$$

where operators $(.)^T$ and $(.)^{\dagger}$ are the transpose and Moore-Penrose pseudoinverse, respectively.

At the end of iteration t, compute the optimal $\mathcal{X}^*$ as,

$$x_{\mathbf{i}}^* = \begin{cases} (\sum_{k=1}^{n-1} \text{fold}(\mathcal{X}_{[k]}^{(t)}))_{\mathbf{i}}, & \mathbf{i} \notin \Omega \\ y_{\mathbf{i}}, & \mathbf{i} \in \Omega \end{cases} \tag{12}$$

where vector index $\mathbf{i} = (i_1 \ldots i_n)$.

C. Tensor Order Augmentation

The tensor train rank, expression (5), suggests two controlling parameters:

1) The order n related to the depth of exploration, or overall granularity.
2) The mode configuration $\mathbf{m}$ connected to the balance of such granularity.

These two parameters should be carefully designed to enhance the local structure exploration.

Thus, before applying BCD to solve problem (8), we apply *Ket Augmentation* [19], where a tensor is folded into a high-order one, with the aim to exploit the local structure of data [14]. In short, ket augmentation casts an n_0-order tensor $\mathcal{X} \in \mathbb{R}^{m_1 \times \cdots \times m_{n_0}}$ into a n-order tensor $\tilde{\mathcal{X}} \in \mathbb{R}^{\tilde{m}_1 \times \cdots \times \tilde{m}_n}$, where $n_0 \le n$ and $\prod_{k=1}^{n} \tilde{m}_k = \prod_{k=1}^{n_0} m_k$. After ket augmentation, the tensor train rank will be more efficient because the local structure of the data can be exploited effectively in terms of computational resources [14] [19].

For the $n_0 = 5$ order traffic tensor $\mathcal{X} \in \mathbb{R}^{32 \times 24 \times 16 \times 7 \times 2}$ several orders and mode configurations are feasible. For completeness, we will search all orders n, not necessarily greater than n_0, so that tensor $\tilde{\mathcal{X}}$ is as balanced as possible. Therefore, we formulate the **Balanced TMac-TT as searching the mode configuration with maximum fairness**, or minimum sparsity, for a given order $n = 2, \ldots$,

$$\begin{cases} \min_{\mathbf{m}} & \mathsf{s}_{12}(\mathbf{m}) \\ \text{s.t.} & \prod_{k=1}^{n} \mathbf{m}(k) = \prod_{k=1}^{n_0} \mathbf{m}_0(k) \end{cases} \tag{13}$$

where function s_{12} measures sparsity [12]. The results are listed in Table II. The computed balance levels of resulting mode configurations suggest that tensor train rank should deliver improved results if $n = \{2, 3, 6\}$. We evaluate this intuition in the next section.

TABLE II
OPTIMAL MODE CONFIGURATIONS

Order n	Mode Configuration $\mathbf{m}$	Balance
2	$(448, 384)$	**0.997**
3	$(64, 56, 48)$	**0.993**
4	$(28, 24, 16, 16)$	0.970
$n_0 = 5$	$\mathbf{m}_0 = (32, 24, 16, 7, 2)$	0.829
5	$(16, 14, 12, 8, 8)$	0.964
6	$(8, 8, 8, 8, 7, 6)$	**0.994**
7	$(8, 8, 7, 6, 4, 4, 4)$	0.959
8	$(7, 6, 4, 4, 4, 4, 4, 4)$	0.972
9	$(7, 4, 4, 4, 4, 4, 4, 3, 2)$	0.955
10	$(7, 4, 4, 4, 4, 4, 3, 2, 2, 2)$	0.930
11	$(7, 4, 4, 4, 4, 3, 2, 2, 2, 2, 2)$	0.911
12	$(7, 4, 4, 4, 3, 2, 2, 2, 2, 2, 2, 2)$	0.898
13	$(7, 4, 4, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2)$	0.890
14	$(7, 4, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$	0.885
$n^* = 15$	$\mathbf{m}^* = (7, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$	0.886

In the Table, n_0 and $\mathbf{m}_0$ corresponds to original tensor order and mode configuration.

IV. Numerical Experiments

This Section compares the Tucker-, PARAFAC- and tensor train-based LRTC algorithms. Figures 2-3 illustrate the results.

A. TMac-TT With Balanced Mode Configuration

Formulation (13) is an integer factorization, which is at least NP. However, the largest mode corresponds to the number of weeks per year. Thus $\|\mathbf{m}\|_\infty \le 52$ and a simple search of prime factors (of components of $\mathbf{m}$) will do it. Then, for fixed order n, we select the mode configuration of highest balance, in terms of sparsity s_{12} [12]. Table II reports the results.

Clearly, we obtained more balanced mode configurations with low orders, as we have more prime factors to work with, i.e. the balance levels decrease as the order increases. Moreover, the number of days per week, 7, prohibits to obtain more squared configurations. We illustrate the performance of tensor train rank with these resulting balanced mode configurations in Fig. 2. **Tensor train rank delivers its higher**

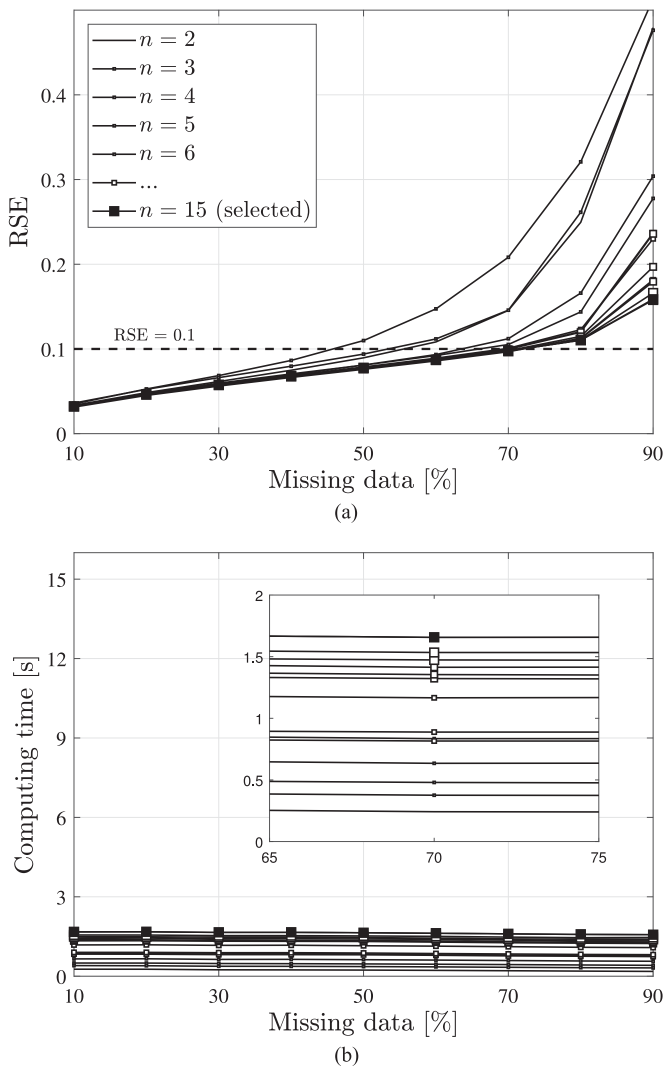

Fig. 2. Searching the optimal order and mode configuration in terms of the imputation accuracy error. (a) Imputation error benchmark. (b) Computing time benchmark.

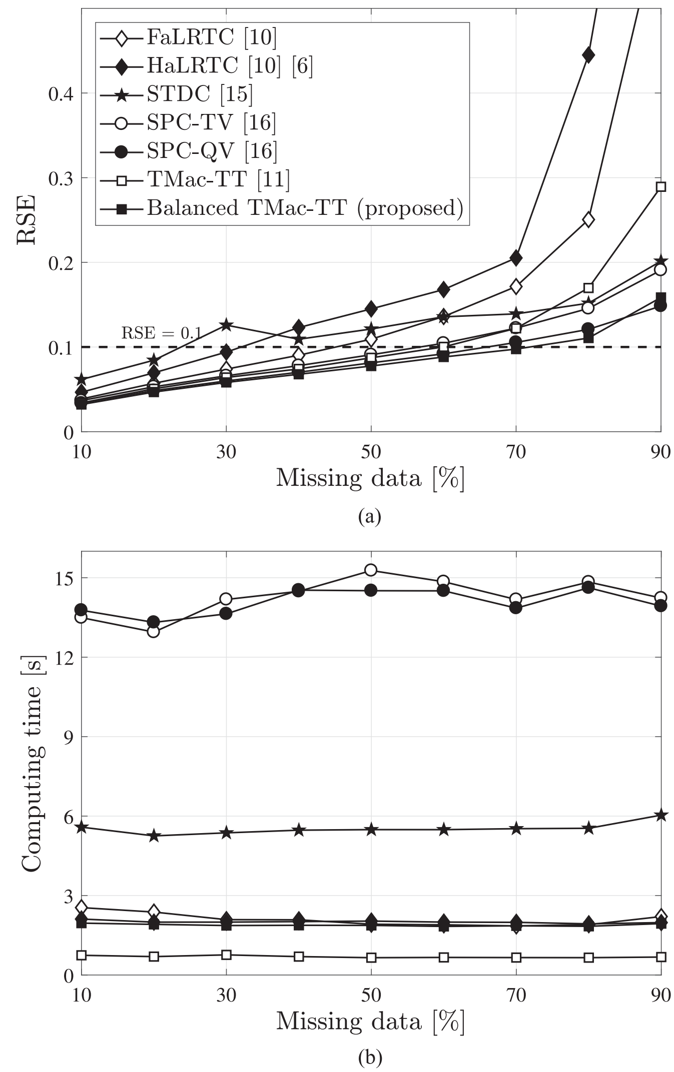

Fig. 3. Comparing the Tucker-, PARAFAC- and tensor train-based tensor completion algorithms. (a) Imputation error benchmark. (b) Computing time benchmark.

accuracy for the largest order tensor, $n^* = 15$, although this will not correspond to the most balanced mode configuration.

B. Tucker and PARAFAC vs. (Balanced) Tensor Train Rank

With the found optimal order $n^* = 15$ and mode configuration $\mathbf{m}^* = (7, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$ (see Fig. 2(a)), we now compare the performance of algorithms.

First, as reported in [11], the TMac-TT outperforms the results obtained by the FaLRTC and HaLRTC (by $\sim$40%). More importantly, when the order and mode configuration are optimized (see Section IV-A), the Balanced TMac-TT improves the completion accuracy (by another $\sim$15%). See Fig. 3(a). Its computing time is however affected ($\times$3 times slower now) which is in the same level of the FaLRTC and HaLRTC algorithms. See Fig. 3(b). The STDC offers an RSE $> 10\%$ in most cases. Finally, the Balanced TMac-TT slightly outperforms the SPC, except at 90% missing data rate. However, its computing time is one order less than the one for the SPC.

Concretely, the **Balanced TMac-TT is able to reach an RSE of 10% even when 70% of the data is missing**. The Relative Squared Error (RSE) is given by, $\mathsf{RSE} = \frac{\|\mathcal{X} - \mathcal{X}^*\|_2}{\|\mathcal{X}\|_2}$. Finally, it is important to note that the use of more balanced tensors strategy is only valid for the tensor train rank.

V. Discussion and Future Work

In this work, we study the benefits of data imputation of balanced higher-order tensors via the tensor train factorization:

1) w.r.t. **compressed sensing**, tensor completion could be advantageous, because of its capability to replace the computationally-heavy dictionary learning, by efficient tensor factorization [20], [21].
2) w.r.t. **matrix completion**, tensor completion could be advantageous, because of its capability to model different data sources in multiple modes, and exploit the correlation between them [22].
3) w.r.t. **Tucker rank and PARAFAC tensor completion**, the tensor train rank could be advantageous, because of its capability to exploit the local structure of data [11].
4) w.r.t. **tensor train rank completion**, our proposed balanced mode configuration could be advantageous, because of its capability to exploit refined local structures.

All in all, the present study allowed us to conceive possible next research questions:

- A main use of tensor completion is on replacing identifiable anomalies by normal patterns. A potential extension work could include the analysis of traffic anomalies and its effect while trying to characterize vehicular traffic.
- Previous works on traffic modelling have not exploited the information summarized in the actual tensor factorization. We plan to study the different tensor components (or factors) to derive interpretable vehicular traffic models.
- From a probabilistic view, when the estimate of the missing value is a linear combination of the existing values, then its addition keeps the rank untouched. Thus, the method does not assume that the rank of the data as such be small at all, rather it is a way to compute linear estimates of conditional expectations (they are projections). Thus, formulations of tensor train rank with sparsity constraints would be desirable.
- We have selected the missing points randomly, however note that in reality the omissions rather appear in blocks. It would be interesting to generate omissions at each station by independent on/off processes.

We will explore some of these aspects in our next works.

VI. Conclusion

In the past recent years, tensor rank, traditionally in the form of the trace (matrix) norm of the multiple modes, has been utilized to find efficient data representations. In this paper, a new data imputation methodology using a recent definition of rank, called the tensor train, has enabled an efficient recovery of missing vehicular traffic volume.

In practice, the possibilities for future research on tensor representation and imputation are extensive in view of the current trends whereby (1) multiple data sources require integration and mathematical modelling, (2) interpretable representations of underlying real-world phenomena are desired, and (3) undemanding crowdsensing systems are favored by users. An extension of our methodology to build models using the core tensor would be desirable. It would also be of interest to explore new scalable formulations.

Acknowledgments

The author would like to thank Ilkka Norros (VTT Technical Research Centre of Finland) for his careful reading and comments on the manuscript, and the reviewers for their feedback which enabled to present a more updated and complete benchmark.

References

[1] X. Hao, L. Xu, N. D. Lane, X. Liu, and T. Moscibroda, “Density-aware compressive crowdsensing,” in *Proc. 16th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, Vienna, Austria, Apr. 2017, pp. 29–40.

[2] R. Du, C. Chen, B. Yang, and X. Guan, “VANET based traffic estimation: A matrix completion approach,” in *Proc. IEEE Global Commun. Conf.*, Dec. 2013, pp. 30–35.

[3] H. Tan, G. Feng, J. Feng, W. Wang, Y.-J. Zhang, and F. Li, “A tensor-based method for missing traffic data completion,” *Transp. Res. Part C: Emerg. Technol.*, vol. 28, pp. 15–27, Mar. 2013.

[4] H. Tan, J. Feng, Z. Chen, F. Yang, and W. Wang, “Low multilinear rank approximation of tensors and application in missing traffic data,” *Adv. Mech. Eng.*, vol. 6, pp. 1–12, 2014.

[5] B. Ran, H. Tan, J. Feng, W. Wang, Y. Cheng, and P. Jin, “Estimating missing traffic volume using low multilinear rank tensor completion,” *J. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 152–161, 2016.

[6] B. Ran, T. Tan, Y. Wu, and P. J. Jin, “Tensor based missing traffic data completion with spatial temporal correlation,” *Phys. A: Statist. Mech. Appl.*, vol. 446, pp. 54–63, Mar. 2016.

[7] N. D. Sidiropoulos, L. D. Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning,” *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.

[8] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[9] L. R. Tucker, “Some mathematical notes on three-mode factor analysis,” *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1966.

[10] J. Liu, P. Musialski, P. Wonka, and J. Ye, “Tensor completion for estimating missing values in visual data,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[11] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, “Efficient tensor completion for color image and video recovery: Low-rank tensor train,” *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.

[12] G. Pastor, I. Mora-Jimenez , R. Jantti, and A. J. Caamano, “Sparsity-based criteria for entropy measures,” in *Proc. 10th Int. Symp. Wireless Commun. Syst.*, Ilmenau, Germany, Aug. 2013, pp. 1–5.

[13] C. J. Hillar and L.-H. Lim, “Most tensor problems are NP-hard,” *J. ACM*, vol. 60, no. 6, pp. 45:1–45:39, Nov. 2013.

[14] I. V. Oseledets, “Tensor-train decomposition,” *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, Sep. 2011.

[15] Y. L. Chen, C. T. Hsu, and H. Y. M. Liao, “Simultaneous tensor decomposition and completion using factor priors,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 577–591, Mar. 2014.

[16] T. Yokota, Q. Zhao, and A. Cichocki, “Smooth PARAFAC decomposition for tensor completion,” *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5423–5436, Oct. 2016.

[17] P. Tseng, “Convergence of a block coordinate descent method for nondifferentiable minimization,” *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[18] Y. Xu, R. Hao, W. Yin, and Z. Su, “Parallel matrix factorization for low-rank tensor completion,” *Inverse Probl. Imag.*, vol. 9, no. 2, pp. 601–624, May 2015.

[19] J. I. Latorre, “Image compression and entanglement,” unpublished paper, 2005. [Online]. Available: https://arxiv.org/abs/quant-ph/0510031

[20] G. Pastor, I. Norros, R. Jantti, and A. J. Caamano, “Compressive data aggregation from Poisson point process observations,” in *Proc. 12th Int. Symp. Wireless Commun. Syst.*, Brussels, Belgium, Aug. 2015, pp. 106–110.

[21] G. Pastor, “Probabilistic models and algorithms for energy-efficient wireless sensor networks,” Ph.D. dissertation, Dept. Commun. Netw., Aalto Univ. School Elect. Eng., Helsinki, Finland, 2016.

[22] M. Signoretto, R. V. de Plas, B. D. Moor, and J. A. K. Suykens, “Tensor versus matrix completion: A comparison with application to spectral data,” *IEEE Signal Process. Lett.*, vol. 18, no. 7, pp. 403–406, Jul. 2011.