TS-MC: TWO STAGE MATRIX COMPLETION ALGORITHM FOR WIRELESS SENSOR NETWORKS

Neha Jain *†, *Anubha Gupta*†, *Vivek Ashok Bohara**

*Wirocomm Research Group, †Signal processing and Bio-medical Imaging Lab,
IIIT-Delhi, New Delhi-110020, India.

ABSTRACT

Wireless sensor network (WSN) data is prone to huge losses and corruption. Hence, the existing matrix completion algorithms experience high estimation errors in such scenarios. Therefore, a robust matrix completion algorithm is required for WSN data to meet the above challenges. This paper proposes a robust "*two stage matrix completion (TS-MC)*" algorithm to recover data from missing and corrupted values. The proposed TS-MC algorithm consists of two stages. For the first stage, two different methods have been proposed for recovering the incomplete data that exploit the double DCT sparsity as WSN data varies smoothly in both time and spatial domain. In the second stage, the recovered data of the first stage is de-noised in the matrix factorization framework, wherein the rank of the data is estimated from the data recovered from the first stage. Simulations are performed on two real datasets of Intel Lab and Data Sensing Lab. Results demonstrate that the proposed TS-MC algorithm achieves high accuracy even when 90% of the data is missing.

***Index Terms*—** Matrix completion, wireless sensor network, data loss, DCT based recovery

1. INTRODUCTION

In recent years, wireless sensor networks (WSNs) are being widely used for both critical as well as non-critical applications. The data collected by sensor nodes face huge losses and corruption due to hardware impairments and severe environmental conditions such as deep fading. For critical applications such as detection of the forest fire, ocean currents, chemical pills and earthquake, it is essential to obtain complete data accurately at the fusion centre (FC) for making an appropriate decision. Previously, to recover the missing and corrupted data in WSN, various interpolation techniques such as *K*-nearest neighbours (KNN) [1], and Delaunay Triangulation (DT) [2] have been used. However, in huge losses, the accuracy of the above techniques are deemed insufficient.

Missing values are effectively recovered by using the low-rank constraint in various applications such as recommender systems, image inpainting and WSN, since most of the real world signals are low rank. However, minimizing the rank of the data is an NP-hard problem. Therefore, the sum of the singular values (nuclear norm) of a data is generally minimized by using the algorithm such as singular value thresholding (SVT) [3]. The rank of the data can also be minimized by using matrix factorization algorithm, in which the data matrix $\mathbf{X}$ of size, say, $n \times t$ is factorized into two matrices of size $n \times r$ and $r \times t$, where r is the rank of the data. However, the rank of an incomplete matrix can not be determined. Hence, a low-rank matrix fitting (LMaFit) algorithm [4] has been presented for matrix completion (MC), in which r is dynamically adjusted. Further, RPCA (robust principal component analysis) [5] is also one of the known method used for matrix completion in the noisy environment, in which the data matrix $\mathbf{X}$ is assumed to be the sum of a low rank matrix ($\mathbf{L}$) and a sparse matrix ($\mathbf{S}$). Hence, along with minimizing the nuclear norm of $\mathbf{L}$, the algorithm also minimizes the l_1 norm of $\mathbf{S}$. Several fast and efficient algorithms have been proposed to solve RPCA problem [6, 7]. In [6], Robust PCA has been solved via gradient descent method. This method is called as RPCA-GD. In [8], spatial and temporal correlation along with the rank constraint have been exploited to recover the missing data in WSN. However, this algorithm requires the rank of the data and topology of the sensor nodes, which is generally not available. Furthermore, during huge losses, the correlation factor also gets affected. Hence, the algorithm experiences high estimation errors.

In order to meet the above challenges, two algorithms employing "*two stage matrix completion (TS-MC)*" for WSN have been proposed in this paper. Both the proposed algorithms of TS-MC outperform various matrix completion algorithms and do not require any prior knowledge of the data such as the rank or the network topology. In the first stage of TS-MC, both algorithms named as *TS-MC-1* and *TS-MC-2* exploited the DCT sparsifying domain, as the real signals are sparse in the DCT domain [9]. However, the implementation of both algorithms is different. In the first stage of TS-MC-1, matrix completion is performed by minimizing the DCT coefficients jointly in both spatial and temporal domain. In the first stage of TS-MC-2, missing data recovery problem is formulated as a compressive sensing (CS) problem, where the sensing matrix will be the partial canonical identity (PCI) matrix which represents the missing data position.

Since PCI sensing matrix is highly incoherent with the DCT matrix [10], this combination is able to recover the missing data with good accuracy. In the second stage of both the algorithms, the recovered data from the first stage is further de-noised using matrix factorization framework. The rank required for matrix factorization can be estimated from the data recovered from the first stage. To validate the performance of the proposed algorithm, results have been compared with various state-of-the-art matrix completion algorithms such as SVT [3], LMaFit [4], OptSpace [11] and RPCA-GD [6] on real dataset of Intel Lab and Data Sensing Lab.

Notations: Matrices are represented in capital and bold letters, vectors in small-case and bold letters, and variables are written in italics. Transpose of a matrix or vector is denoted as $(\cdot)^T$, vectorization of matrix is denoted as '(:)'. Further, $\mathbb{R}$ represents real number and '$\bullet$' denotes element wise multiplication. The element corresponds to i^{th} row and j^{th} column of a matrix $\mathbf{A}$ is represented as $A(i,j)$. Furthermore, $||.||_F$ represents Frobenius norm and $||.||_p$ represents the l_p norm, where $0 \leq p \leq 2$. $\mathbf{I}_a$ represents the identity matrix of size a and $\mathbf{0}_{a\times b}$ represents zero matrix of size $a \times b$. The noise considered in this paper is additive white Gaussian noise with zero mean and unit variance.

2. PROPOSED ALGORITHMS

Consider a wireless sensing network consisting of 'n' sensor nodes. Let us assume that 't' measurements are transmitted from each sensor node. Hence, the transmitted matrix $\mathbf{X}$ will have a dimension of $n \times t$. The received incomplete noisy matrix is given as $\mathbf{Y} = \mathbf{B} \bullet (\mathbf{X} + \mathbf{N})$, where $\mathbf{B}$ is the binary matrix consists of '1' and '0' at the position where data is present and absent, respectively, $\bullet$ is the element-wise product and $\mathbf{N}$ is the noise matrix.

2.1. First stage

In the first stage, we have exploited the fact that most of the real signal varies slowly. In [9], it has been shown that DCT acts like a Karhunen-Loève (KL) type basis for a large class of smooth signals and hence, may act as one of the best sparsifying basis for these signals. Further, the WSN data is slow-varying in both temporal and spatial domain and hence, DCT can be exploited for the both domains. Therefore, in the first stage, the missing data is recovered by exploiting the above fact by using the following methods.

2.1.1. *Method of TS-MC-1*

In this method, the matrix completion problem is formulated as

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{B} \bullet \mathbf{X}||_F^2 + \lambda_1 ||\mathbf{D}_1 \mathbf{X} \mathbf{D}_2||_1, \quad (1)$$

where $\mathbf{D}_1$ and $\mathbf{D}_2$ are the DCT matrices of size $n \times n$ and $t \times t$, respectively. The constraint $||\mathbf{D}_1 \mathbf{X} \mathbf{D}_2||_1$ used in (1) enforces the double DCT sparsity as data is slowly-varying in both spatial and temporal domains. Further, $\lambda_1 \in \mathbb{R}$ is the regularization parameter that controls the trade-off between data accuracy and the sparsity level. A proxy variable $\mathbf{W}$ is introduced to rewrite (1) as

$$\min_{\mathbf{X},\mathbf{W}} ||\mathbf{Y} - \mathbf{B} \bullet \mathbf{X}||_F^2 + \lambda_1 ||\mathbf{W}||_1 + \lambda_2 ||\mathbf{W} - \mathbf{D}_1 \mathbf{X} \mathbf{D}_2||_F^2, \quad (2)$$

where λ_2 is the controlling parameter to control the degree of equality between the original term and the proxy variable. For small value of λ_1, the equality constraint is relaxed and for high value of λ_1, the constraint is enforced. The above problem (2) can be divided into the following sub-problems by using the alternating direction method of multipliers (ADMM):

$$\begin{aligned} &\mathrm{P}_1: \min_{\mathbf{W}} ||\mathbf{W} - \mathbf{D}_1 \mathbf{X} \mathbf{D}_2||_F^2 + \lambda_a ||\mathbf{W}||_1 \; \left(\because \lambda_a = \frac{\lambda_1}{\lambda_2}\right), \\ &\mathrm{P}_2: \min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{B} \bullet \mathbf{X}||_F^2 + \lambda_2 ||\mathbf{W} - \mathbf{D}_1 \mathbf{X} \mathbf{D}_2||_F^2 \end{aligned} \quad (3)$$

P_1 can be solved using soft-thresholding [12] and P_2 is a simple least squares problem.

Algorithm 1 TS-MC-1

Input: $\mathbf{B}$, $\mathbf{Y}$, $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{X}_{int}$, λ_a, μ, max_{iter1}, max_{iter2}
Initializing: $\mathbf{X} = \mathbf{X}_{int}$
Obtain $\mathbf{B}_c$ such that $B_c(i,j) = \begin{cases} 1 & \text{if } B(i,j) = 0 \\ 0 & \text{if } B(i,j) = 1 \end{cases}$
for $k_1 = 1 : max_{iter1}$
  $\mathbf{W} = sgn(\mathbf{D}_1 \mathbf{X} \mathbf{D}_2) \max(0, |\mathbf{D}_1 \mathbf{X} \mathbf{D}_2| - 0.5\lambda_a)$ [1]
  $\mathbf{X} = \mathbf{D}_1^T \mathbf{W} \mathbf{D}_2^T$
  $\mathbf{X} = \mathbf{Y} + \mathbf{B}_c \bullet \mathbf{X}$
end
$\begin{bmatrix}\mathbf{U} & \mathbf{D} & \mathbf{V}\end{bmatrix} = svd(\mathbf{X})$; svd is the singular value decomposition.
$\mathbf{d} = diag(\mathbf{D})$; $diag$ picks the diagonal elements.
r = number of highest values in $\mathbf{d}$
for $k_2 = 1 : max_{iter2}$
  $\mathbf{U} \leftarrow \min_{\mathbf{U}} \left\| \begin{bmatrix}\hat{\mathbf{X}} & \mathbf{0}_{n\times r}\end{bmatrix} - \mathbf{U} \begin{bmatrix}\mathbf{V} & \sqrt{\mu}\mathbf{I}_r\end{bmatrix} \right\|_F^2$
  $\mathbf{V} \leftarrow \min_{\mathbf{V}} \left\| \begin{bmatrix}\hat{\mathbf{X}} \\ \mathbf{0}_{r\times t}\end{bmatrix} - \begin{bmatrix}\mathbf{U} \\ \sqrt{\mu}\mathbf{I}_r\end{bmatrix} \mathbf{V} \right\|_F^2$
end
X=UV
Output: $\mathbf{X}$

2.1.2. *Method of TS-MC-2*

In this method, the problem of missing data recovery is formulated as a CS problem. According to the theory of CS, the original data vector $\mathbf{x}$ of length p can be sensed to a lower dimensional vector $\mathbf{y}$ of length q, $(q < p)$ using a sensing matrix $\mathbf{\Phi}$ of size $q \times p$ such that $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$. Assuming the vector $\mathbf{x}$ to

[1] $sgn = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$

be sparse in $\mathbf{\Psi}$ domain, the compressive measurement vector $\mathbf{y}$ in the presence of noise can be written as

$$\begin{aligned}\mathbf{y} &= \mathbf{\Phi}(\mathbf{x}+\mathbf{n}_p)\\ &= \mathbf{\Phi}\mathbf{\Psi}^{-1}\mathbf{s}+\mathbf{n}_q \quad (\because \mathbf{s}=\mathbf{\Psi}\mathbf{x}, \mathbf{n}_q=\mathbf{\Phi}\mathbf{n}_p)\,, \end{aligned} \tag{4}$$

where $\mathbf{s}$ is the sparse vector representation of original data $\mathbf{x}$ and, $\mathbf{n}_p$ and $\mathbf{n}_q$ are the noise vectors of length p and q, respectively. The original data vector $\mathbf{x}$ of length p can be recovered from $\mathbf{y}$ if $\mathbf{\Phi}$ and $\mathbf{\Psi^{-1}}$ are mutually incoherent. The coherence can be calculated as

$$\mu(\mathbf{\Phi},\mathbf{\Psi}^{-1}) = \sqrt{p}\max_{\forall i,j}\frac{\left|\left\langle \mathbf{\Phi}_i, \mathbf{\Psi}_j^{-1}\right\rangle\right|}{\left\|\mathbf{\Phi}_i\right\|_2\left\|\mathbf{\Psi}_j^{-1}\right\|_2}, \tag{5}$$

with $\mu(\mathbf{\Phi},\mathbf{\Psi}^{-1}) \in [1,\sqrt{p}]$. Here, $\mu = 1$ is the best case which represents maximum incoherence between the matrices. It is observed that the random matrices like Gaussian and Bernoulli follow the above constraint with coherency of $\sqrt{\log p}$. Signal $\mathbf{x}$ can be recovered from $\mathbf{y}$ by solving the following optimization problem

$$\hat{\mathbf{s}} = \min_{\mathbf{s}}||\mathbf{y}-\mathbf{A}\mathbf{s}||_2^2+\lambda||\mathbf{s}||_1, \tag{6a}$$

$$\hat{\mathbf{x}} = \mathbf{\Psi}^{-1}\hat{\mathbf{s}}, \tag{6b}$$

where λ_3 is the regularization parameter to control the level of sparsity and the data accuracy. Iterative soft thresholding algorithm (ISTA) [12] can be used to solve (6a).

In this paper, above theory has been utilized in the context of matrix completion as follows. The data matrix $\mathbf{X}$ is vectorized to a vector $\mathbf{x}$ of length nt. A sub-sampled vector $\mathbf{y}$ of length m is obtained from the known entries of $\mathbf{x}$ such that the signal $\mathbf{y}$ can be written as $\mathbf{y}=\mathbf{\Phi}\mathbf{x}$, where $\mathbf{\Phi}$ is the partial canonical identity (PCI) sensing matrix of size $m\times nt$. This matrix will contain single '1' in each row corresponding to the available data and all remaining entries are zero. For example, assume $\mathbf{x} = \begin{bmatrix}x_1 & x_2 & x_3 & x_4 & x_5 & x_6\end{bmatrix}^T$ where only x_1 and x_5 entries are received, then vector $\mathbf{y}$ will be equal to $\begin{bmatrix}x_1 & x_5\end{bmatrix}^T$ and hence, the sensing matrix will be written as $\mathbf{\Phi} = \begin{bmatrix}1&0&0&0&0&0\\0&0&0&0&1&0\end{bmatrix}$. The sparse representation of $\mathbf{X}$ is given by $\mathbf{S}=\mathbf{D}_1\mathbf{X}\mathbf{D}_2$. Since $\mathbf{D}_1$ and $\mathbf{D}_2$ are DCT matrices, $\mathbf{X}=\mathbf{D}_1^T\mathbf{S}\mathbf{D}_2^T$. Therefore, the observed entries of transmitted data matrix $\mathbf{X}$ can be re-written as

$$\mathbf{y}=\mathbf{\Phi}\mathbf{X}(:)=\mathbf{\Phi}\mathbf{x}=\mathbf{\Phi}(\mathbf{D}_2\otimes\mathbf{D}_1^T)\mathbf{s}, \tag{7}$$

where $\mathbf{s}=\mathbf{S}(:)$. From (7), we observe that the sparsifying matrix is $\mathbf{\Psi}^{-1}=\mathbf{D}_2\otimes\mathbf{D}_1^T$. To recover the transmitted data $\mathbf{x}$ using (6), PCI sensing matrix and $(\mathbf{D}_2\otimes\mathbf{D}_1^T)$ must be incoherent. The coherency between above matrices is calculated using (5) and is observed to be $\sqrt{2}$, that is indeed small. Hence, this method ensure good data reconstruction. Compared to TS-MC-1, this method also carries out de-noising of data.

2.2. Second stage

The data recovered from the first stage $(\hat{\mathbf{X}})$ is de-noised in the second stage using matrix factorization such as $\hat{\mathbf{X}}=\mathbf{U}\mathbf{V}$. The dimension of $\mathbf{U}$ and $\mathbf{V}$ are chosen to be $n\times r$ and $r\times t$, respectively, where r is the rank of the data that can be estimated from $\hat{\mathbf{X}}$. The problem can be formulated as

$$\min_{\mathbf{U},\mathbf{V}}||\hat{\mathbf{X}}-\mathbf{U}\mathbf{V}||_F^2+\mu||\mathbf{U}||_F^2+\mu||\mathbf{V}||_F^2. \tag{8}$$

The problem in (8) can be divided into two sub-problems using ADMM as

$$\begin{aligned}\mathrm{P}_3:\mathbf{U}&\leftarrow\min_{\mathbf{U}}||\hat{\mathbf{X}}-\mathbf{U}\mathbf{V}||_F^2+\mu||\mathbf{U}||_F^2,\\ \mathrm{P}_4:\mathbf{V}&\leftarrow\min_{\mathbf{V}}||\hat{\mathbf{X}}-\mathbf{U}\mathbf{V}||_F^2+\mu||\mathbf{V}||_F^2.\end{aligned} \tag{9}$$

The above sub-problems can be re-written as

$$\begin{aligned}\mathrm{P}_3:\mathbf{U}&\leftarrow\min_{\mathbf{U}}\left\|\begin{bmatrix}\hat{\mathbf{X}} & \mathbf{0}_{n\times r}\end{bmatrix}-\mathbf{U}\begin{bmatrix}\mathbf{V} & \sqrt{\mu}\mathbf{I}_r\end{bmatrix}\right\|_F^2\\ \mathrm{P}_4:\mathbf{V}&\leftarrow\min_{\mathbf{V}}\left\|\begin{bmatrix}\hat{\mathbf{X}}\\ \mathbf{0}_{r\times t}\end{bmatrix}-\begin{bmatrix}\mathbf{U}\\ \sqrt{\mu}\mathbf{I}_r\end{bmatrix}\mathbf{V}\right\|_F^2.\end{aligned} \tag{10}$$

The above sub-problems are the simple least squares problems. Further, the transmitted matrix $\mathbf{X}$ can be recovered as $\tilde{\mathbf{X}}=\mathbf{U}\mathbf{V}$. This is note that the second stage for both algorithms *TS-MC-1* and *TS-MC-2* is same.

Algorithm 2 TS-MC-2

Input: $\mathbf{\Phi}$, $\mathbf{y}$, $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{x}_{int}$, n, t, μ, λ_3, max_{iter1}, max_{iter2}
Initializing: $\mathbf{x}=\mathbf{x}_{int}$
$\mathbf{H}=\mathbf{\Phi}(\mathbf{D}_2\otimes\mathbf{D}_1^T)$
$\alpha=\max(eig(\mathbf{H}^T\mathbf{H}))$
for $k_1=1:max_{iter1}$
  $\mathbf{x}=sgn(\mathbf{x}+\frac{1}{\alpha}\mathbf{H}^T\mathbf{y}-\mathbf{H}\mathbf{x})\max(0,|\mathbf{H}^T\mathbf{y}-\mathbf{H}\mathbf{x}|-\frac{\lambda_3}{2\alpha})^1$
end
$\mathbf{X}$ = reshape($\mathbf{x}$, (n,t)); reshape converts the vector $\mathbf{x}$ into a matrix of size $n\times t$.
$\begin{bmatrix}\mathbf{U} & \mathbf{D} & \mathbf{V}\end{bmatrix}=svd(\mathbf{X})$; svd is the singular value decomposition.
$\mathbf{d}=diag(\mathbf{D})$; $diag$ picks the diagonal elements.
r = number of highest values in $\mathbf{d}$
for $k_2=1:max_{iter2}$
  $\mathbf{U}\leftarrow\min_{\mathbf{U}}\left\|\begin{bmatrix}\hat{\mathbf{X}} & \mathbf{0}_{n\times r}\end{bmatrix}-\mathbf{U}\begin{bmatrix}\mathbf{V} & \sqrt{\mu}\mathbf{I}_r\end{bmatrix}\right\|_F^2$
  $\mathbf{V}\leftarrow\min_{\mathbf{V}}\left\|\begin{bmatrix}\hat{\mathbf{X}}\\ \mathbf{0}_{r\times t}\end{bmatrix}-\begin{bmatrix}\mathbf{U}\\ \sqrt{\mu}\mathbf{I}_r\end{bmatrix}\mathbf{V}\right\|_F^2$
end
X=UV
Output: **X**

3. SIMULATION RESULTS

In this section, the proposed algorithms (TS-MC-1, TS-MC-2) have been compared with various other matrix completion algorithms such as SVT [3], LMaFit [4], OptSpace [11] and RPCA-GD [6]. To validate the performance of the proposed algorithms, results have been shown on real datasets of temperature and humidity, taken from the Intel lab. We have also compared the proposed algorithms on temperature dataset of data sensing lab.

3.1. Intel Lab dataset

The data from 53 sensor nodes at every minute has been considered for 200 minutes in the simulation Hence, $n = 53$ and $t = 200$. This is to note that out of $nt = 10600$ entries only $k = 7901$ entries are available, hence for creating the ground truth, few entries say g have been manually removed. Hence, nt entries have been recovered from $m = k - g$ entries for all the algorithms. However, as ground truth is available only for the g entries. Therefore, data loss percentage is computed as $\frac{g}{k} \times 100\%$ and the normalized mean square error (NMSE) [2] is also computed only for g values. In Fig. 1 and 2, NMSE has been plotted against data loss percentage at signal-to-noise power ratio (SNR) of 10 dB for temperature and humidity dataset, respectively. From results, we observe that both proposed algorithms of TS-MC outperforming the various MC algorithms. Furthermore, the performance of TS-MC-2 is much better than TS-MC-1 at high data loss. From Fig. 1, we observe that at 90% data loss, TS-MC-1 is providing around 6 dB improvement, while TS-MC-2 is providing around 12.5 dB improvement as compared to LMaFit, OpTspace and RPCA-GD. Therefore, at higher data losses, the proposed algorithm is outperforming very well, and hence, in Fig. 3 we have plotted the NMSE with respect to SNR for 90% data loss.

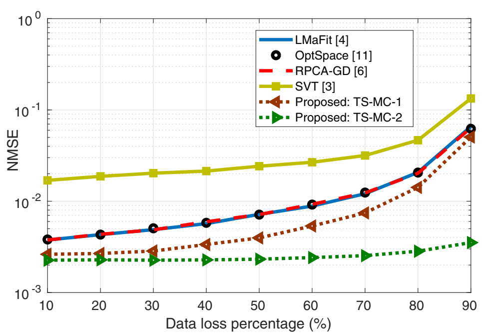

Fig. 1: *NMSE against data loss percentage at SNR = 10 dB for humidity dataset taken from Intel lab*

3.2. Data Sensing Lab

For further verification, we have also compared the algorithms using the dataset of another lab (data sensing lab). In this dataset, $n = 38$ sensor nodes are present and similar to above $t = 200$ timestamps have been taken. Further, in this dataset, 9.16% of entries are not available, as $nt = 7600$ and

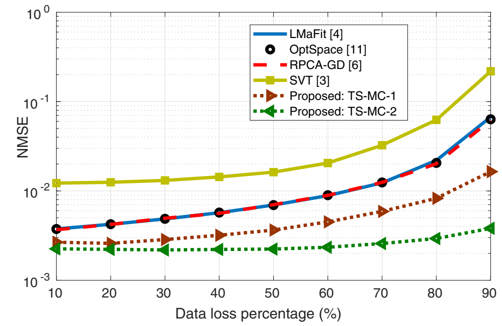

Fig. 2: *NMSE against data loss percentage at SNR = 10 dB for temperature dataset taken from Intel lab*

$k = 6904$. Similar to above, in Fig. 4, NMSE for g entries has been plotted against data loss percentage. The proposed algorithms are consistently outperforming the conventional MC algorithms on this dataset as well.

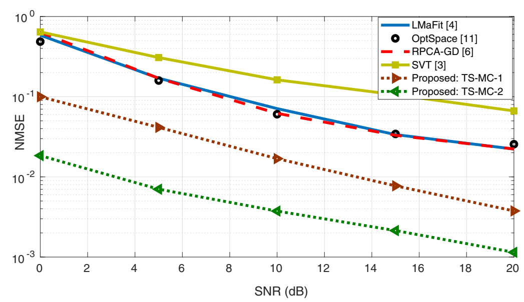

Fig. 3: *NMSE against SNR at 90% data loss for temperature dataset taken from Intel lab*

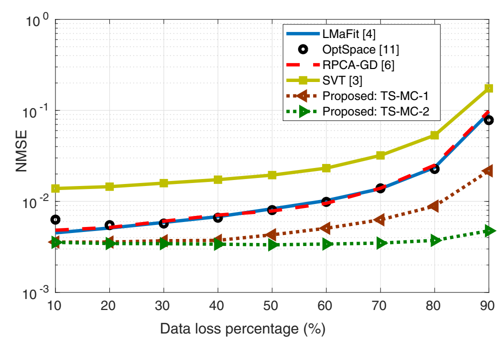

Fig. 4: *NMSE against data loss percentage at SNR = 10 dB for temperature dataset taken from Data sensing lab*

4. CONCLUSION AND FUTURE WORK

The proposed TS-MC algorithm is outperforming various MC algorithms as tested on two datasets. At higher data loss of 90%, TS-MC is performing better than the conventional MC algorithms by almost 12 dB. The proposed method can also be explored for missing data recovery in a recommender system and image inpainting applications, which will be a subject of future work.

[2]NMSE=$\frac{||x-\hat{x}||_2^2}{||x||_2^2}$, where x is the original data and $\hat{x}$ is the recovered data.

5. REFERENCES

[1] Thomas Cover and Peter Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[2] L. Kong, D. Jiang, and M. Wu, "Optimizing the spatio-temporal distribution of cyber-physical systems for environment abstraction," in *2010 IEEE 30th International Conference on Distributed Computing Systems*, June 2010, pp. 179–188.

[3] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[4] Zaiwen Wen, Wotao Yin, and Yin Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Mathematical Programming Computation*, vol. 4, no. 4, pp. 333–361, Dec 2012.

[5] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 11:1–11:37, June 2011.

[6] Xinyang Yi, Dohyung Park, Yudong Chen, and Constantine Caramanis, "Fast algorithms for robust pca via gradient descent," in *Advances in neural information processing systems*, 2016, pp. 4152–4160.

[7] Zhouchen Lin, Minming Chen, and Yi Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, 2010.

[8] L. Kong, M. Xia, X. Y. Liu, M. Y. Wu, and X. Liu, "Data loss and reconstruction in sensor networks," in *2013 Proceedings IEEE INFOCOM*, April 2013, pp. 1654–1662.

[9] Anubha Gupta, Shiv Dutt Joshi, and Pushpendra Singh, "On the approximate discrete KLT of fractional brownian motion and applications," *Journal of the Franklin Institute*, 2018.

[10] N. Jain, V. A. Bohara, and A. Gupta, "iDEG: Integrated data and energy gathering framework for practical wireless sensor networks using compressive sensing," *IEEE Sensors Journal*, vol. 19, no. 3, pp. 1040–1051, Feb 2019.

[11] Raghunandan H Keshavan, Andrea Montanari, and Sewoong Oh, "Matrix completion from noisy entries," *Journal of Machine Learning Research*, vol. 11, no. Jul, pp. 2057–2078, 2010.

[12] Ivan Selesnick, "Sparse signal restoration," 10 2018.