Online recovery of missing values in vital signs data streams using low-rank matrix completion

Shiming Yang, Konstantinos Kalpakis
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, USA
{*shiming1, kalpakis*}*@umbc.edu*

Colin F. Mackenzie, Lynn G. Stansbury, Deborah M. Stein, Thomas M. Scalea and Peter F. Hu
R Adams Cowley Shock Trauma Center;
Shock Trauma and Anesthesiology Research Center
University of Maryland School of Medicine
Baltimore, USA
{*lstansbury,dstein,tscalea,phu*}*@umm.edu,*
cmack003@umaryland.edu

***Abstract*—Continuous, automated, electronic patient vital signs data are important to physicians in evaluating traumatic brain injury (TBI) patients' physiological status and reaching timely decisions for therapeutic interventions. However, missing values in the medical data streams hinder applying many standard statistical or machine learning algorithms and result in losing some episodes of clinical importance. In this paper, we present a novel approach to filling missing values in streams of vital signs data. We construct sequences of Hankel matrices from vital signs data streams, find that these matrices exhibit low-rank, and utilize low-rank matrix completion methods from compressible sensing to fill in the missing data. We demonstrate that our approach always substantially outperforms other popular fill-in methods, like k-nearest-neighbors and expectation maximization. Further, we show that our approach recovers thousands of simulated missing data for intracranial pressure, a critical stream of measurements for guiding clinical interventions and monitoring traumatic brain injuries.**

***Keywords*-low rank; matrix completion; Hankel matrix; vital signs; missing values; data imputation.**

I. Introduction

Traumatic brain injury (TBI) is the most common cause of admission to emergency care and of trauma-related death in the U.S. civilian population and is a major cause of death and disability in combat casualties [1], [2], [3]. Because of its fatality rate and profound effects on survivors' living quality, much research has focused on support of early-warning decision-assist systems that can maximize the potential for timely therapeutic interventions to improve long-term clinical outcomes [4], [5], [6], [7].

In most modern intensive care units (ICUs), vital signs (VS), such as heart rate (HR), blood pressure (BP), and oxygen saturation (SpO2), among others, are collected in high-quality, automated, continuous electronic data streams as basic evidence of patients physiological status. For TBI patients, intracranial pressure (ICP), the pressure measured inside of the closed box of the skull is of special importance. Even relatively brief periods of elevated ICP are associated with adverse outcome [7], and marked elevation of ICP or elevation unresponsive to medications may require surgery which may be lifesaving but which is also not without risk. However, directly measuring ICP is highly invasive and requires a surgical opening in the patient's skull through which a sensor is inserted. For various technical reasons, data streams from these systems often include gaps of missing data. Therefore, ICP data are often frustratingly intermittent compared to other less invasive VS. Missing values in data streams cause difficulties when applying many statistical or machine learning methods to VS analysis. Therefore, in dynamic data stream processing, we cannot wait to build regression models till all data are available. Hence, it is necessary to accurately recover or estimate missing values using the locally available information.

We describe a novel approach for imputing missing values for continuously collected clinical data streams. In our approach, we construct a sequence of Hankel matrices, find that they are of rather low-rank, and use low-rank matrix completion methods, recently proposed in compressible sensing, to fill in the missing values. We demonstrate that our approach provides improved accuracy with respect to other popular missing value fill-in methods.

II. Related Work

Incomplete data are a pervasive problem in many research fields, particularly those that depend heavily on observational or field data, such as clinical, social, and environmental studies. Missing data hinder the application of many statistical analysis and machine learning methods available in off–the–shelf software. To analyze datasets with missing values, researchers either remove incomplete data points or fill the vacuum with a 'reasonable guess'[8], [9]. [1] The first approach is simple but only acceptable when the dataset is sufficiently large such that the missing cases do not have noticeable impact on the analysis. Unfortunately, often we have limited measurements, and thus all available

[1]Filling–in missing values is also known as data imputation.

data must be used. Moreover, in data stream processing, data are handled in near–real time and in local temporal windows, which suggests that researchers use all available data points at each time step in order to capture the systems' characteristics to the maximum extent.

Many methods have been developed to address the missing data problem. Traditional and intuitive methods include *filling with the mean or median* [9], [10] and *last observation carried forward* [9]. These methods are simple but tend to give more biased estimates. Moreover, they ignore information from other variables and knowledge of the system. Model–based methods have been developed to get around this problem. Regression models and neural networks are two examples. These methods treat the variables with missing values in deterministic ways. In many applications, it makes more sense to consider variables in a probabilistic approach, which is in better agreement with natural randomness. The maximum likelihood method, such as the expectation maximization is a very successful method applied in missing data analysis [8], [11].

Recently, the matrix completion problem [12], [13], [14] has attracted much attention in signal reconstruction with a sequence of partial observations. The matrix completion problem assumes that the matrix at hand has low rank, which allows a set of sparse and noisy observations to recover the matrix [13]. These methods have wide applications in many engineering and statistical modeling problems, where order, dimensionality, or complexity of a model can be evaluated by the rank of an appropriate matrix. Becker et al. [15] demonstrated an image processing application that recovers images from noisy observations using matrix completion. Candès et al. describe the use of matrix completion in recovering a signal from few elements of its Fourier spectrum, as well as in collaborative filtering in online recommendation systems [12], [13]. Another related problem is that of robust Principal Component Analysis (PCA) in which we recover partially observed low-rank matrices corrupted by a sparse error/artifact matrix. [16]. In many applications, such sparse matrix is actually a collection of artifacts that we hope to remove from the original observations, or some interesting but rare events that we want to separate from the background, like removing face illumination, or identifying moving objects from video frames[16].

The massive amounts of continuous electronic physiologic data collected in modern medical centers are beginning to come under scrutiny with advanced analytical methods. Efforts to collect, process, store, and use these data effectively have led to a number of innovative data mining and machine learning activities. The R Adams Cowley Shock Trauma Center (STC) associated with the University of Maryland School of Medicine (UMSOM) has been a leader in this process. The data used in our work was collected as part of a multi-pronged research effort aimed at improving outcomes for patients with severe TBI. With approval of the UMSOM Institutional Review Board (IRB)more than 50 types of vital signs are collected routinely on all patients admitted to the STC Neuro-Critical Care units. Demographic, injury, and outcome data are collated with the physiologic data and this dataset is used extensively in studies of estimating physiologic data from invasive versus non-invasive sources [4], and in studies predicting mortality and other outcomes using vital signs data collected early in the course of care [5], [6], [7].

III. Matrix Completion

A. Theoretical background

Consider a matrix $M \in \mathbb{R}^{n_1 \times n_2}$. Suppose that some of its elements are hidden (missing, unknown), and that we desire to recover them using its available elements. Let Ω be the set of indices (i, j) of the elements $M_{i,j}$ of M that are available/observed. Clearly, infinitely many solutions are possible unless we impose some additional constraints on M. One possibility is the matrix M has low rank or can be approximated by low rank, hence it is possible to seek an X of minimum rank that satisfies (1). The matrix completion problem is

$$\begin{aligned} \min \quad & \operatorname{rank}(X) \\ \text{s.t.} \quad & X_{ij} = M_{ij}, \quad (i,j) \in \Omega. \end{aligned} \tag{1}$$

Because this problem is NP–hard, its relaxation to the following convex optimization problem is solved instead

$$\begin{aligned} \min \quad & ||X||_* \\ \text{s.t.} \quad & X_{ij} = M_{ij}, \quad (i,j) \in \Omega, \end{aligned} \tag{2}$$

where $||X||_*$ is the nuclear norm of X and is defined as the sum of singular values of the matrix X. Hereafter, matrix completion will refer to the above convex optimization problem.

Candès et al. [12] show that not all low rank matrices can be recovered from a subset of their entries. The concept of matrix incoherence [2] is used to evaluate whether a low rank matrix can be exactly recovered. Intuitively, low coherence means that the singular vectors of the matrix are sufficiently spread, and it is less likely that the matrix resides in the null space of the sampling operators. This promises a high probability of recovering exacting a low rank matrix from a few observations. Suppose that there are m observed entries of M which are located uniformly at random in matrix M, and that $M \in \mathbb{R}^{n \times n}$ has rank r. Candès et al [12] show that there exist constants C and c, such that if $m \geq Cn^{5/4} r \log n$ then (2) has a unique optimal solution with probability at least $1 - cn^{-3}$, that is, we can recover the matrix M.

[2]The incoherence of a matrix M is defined as $\mu(M) = \max_{i \neq j} \frac{M_i^T M_j}{||M_i|| \cdot ||M_j||}$, where M_i denotes the ith column of M.

B. Block Hankel matrices for data streams

We construct block Hankel matrices from data streams with missing values, continuously estimate the missing values within a sliding window using matrix completion, and perform moving-window smoothing.

Consider a sequence of observations S_{t-w+1}, S_{t-w}, ..., S_t for m data stream variables within a sliding window of length w that ends at time t, so that $S_i \in \mathbb{R}^m$ is the vector of values of these variables at time i. The $km \times \ell$ block Hankel matrix H_t at time t is constructed through a partition-and-stacking process as follows

$$H_t = \begin{bmatrix} S_{t-w+1} & S_{t-w+2} & \cdots & S_{t-w+\ell} \\ S_{t-w+2} & S_{t-w+3} & \cdots & S_{t-w+\ell+1} \\ \vdots & \vdots & \vdots & \vdots \\ S_{t-w+k} & S_{t-w+k+1} & \cdots & S_t \end{bmatrix},$$

where $\ell = w - k + 1$. In constructing matrix H_t, a moving window of size w over the data stream variables is used. Within this window, the sliding sequences of length k (e.g. slices) form the columns of H_t. Matrix $H_t \in \mathbb{R}^{km \times \ell}$ can be made square if $\ell = k \cdot m$. Note that missing values in the data stream variables result into Hankel matrices with missing entries.

Hankel matrices have important applications in system identification [17] and signal detection and estimation [18]. A low rank Hankel matrix implies a simple system model. However, real observations are contaminated by noise and result in Hankel matrices that tend to have full rank. Moreover, due to missing values, we are faced with Hankel matrices with missing entries. Hence, the matrix completion methods reviewed in the previous section may fail for such high rank matrices. Fortunately, we find that for the Hankel matrices for the vital signs dataset at hand, a few of their singular values capture most of their nuclear norm. Thus, the matrix completion methods could be an effective method to recover the missing entries of these Hankel matrices.

Suppose that an observation S_i is missing value(s) for some variables. Since S_i appears in multiple Hankel matrices, namely $H_i, \ldots, H_{i+w-1}$, we will have multiple surrogate estimates for the missing values. Hence, there is opportunity for smoothing for the estimated fill-in values for S_i as subsequent future observations become available. We use the simple arithmetic average of such multiple estimates as the smoothed estimate for the missing values of S_i.

IV. Experimental Evaluation

We identified 12 hours of vital sign data streams from a TBI patient without any missing ICP values and less than 5% missing values for the heart rate (HR), systolic blood pressure (SBP), mean BP (MBP), and CO_2. The sampling period for these vital sign variables is 6 seconds. Note that there are five variables in our data stream. We constructed a sequence of 30×30 Hankel matrices, one matrix for each sampling step using the sliding window method discussed in the last section. [3] Note that each Hankel matrix contains data spanning about 3 minutes. In determining the size of Hankel matrices for experiments, we tested different sizes, 15×15, 30×30, and 60×60. We chose 30×30 Hankel matrix since there is a noticeable improvement when going from 15 to 30, but not from 30 to 60.

In using a sliding window, previously missing points may become available after they have been estimated by the matrix completion. When such entries enter into the next newly constructed Hankel matrix, we can either treat them as still being missing; or as being observed. The latter could be useful if there is a long sequence of missing values for one variable. Using the previously estimated values can avoid the situation of constructing a matrix with a complete row missing, which prevents the use of matrix completion methods.

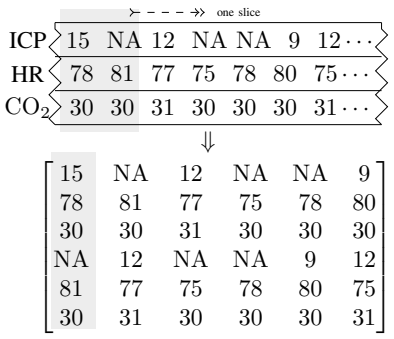

Figure 1. Example Hankel matrix for a sequence of $m = 3$ vital signs using a sliding window of size $w = 7$ ($k = 2$).

We conducted three experiments of missing values for a single patient dataset, and one experiment for a 10 patients dataset.

We compare our approach with three other commonly techniques: the k-nearest neighbor for $k = 1$ and for $k = 5$, and the regularized expectation maximization [19]. The k-nearest neighbor is a simple yet effective data imputation method, and is widely available in many statistical analysis software packages. The expectation maximization (EM) algorithm maximizes the likelihood and the information of the missing data in the partially observed matrix. By penalizing the likelihood with the mutual information between the missing ones and the incomplete data, the Regularized EM reduces the uncertainty of missing data [19]. For brevity, hereafter we call our approach the **MC** approach.

We use the relative error and the normalized mean square error (NMSE) to assess the performance of methods in

[3] We choose sliding windows instead of tumbling (non–overlapping) windows when constructing our Hankel matrices, in order to have the temporal neighbors of observations to be in the close–by columns and in the same row of the Hankel matrix.

filling-in missing values. The NMSE is the mean square error divided by the product of the means of the estimated value and the true value for the missing values.

A. Rank of our Hankel matrices

We examine the rank of the Hankel matrices constructed by our approach, since their rank is critical in the number of missing values the matrix completion methods can tolerate and still recover a unique Hankel matrix.

Almost all of our Hankel matrices are of full rank. We also calculate the least k so that the sum of the largest k singular values of a Hankel matrix is at least 95% of its nuclear norm. Figure 2 shows that for most Hankel matrices, k is 10 or less, while Figure 3 shows that the top 5 singular values contain over 90% of the nuclear norm for almost all Hankel matrices. This suggests that our Hankel matrices originate from low rank matrices perturbed with some random noise.

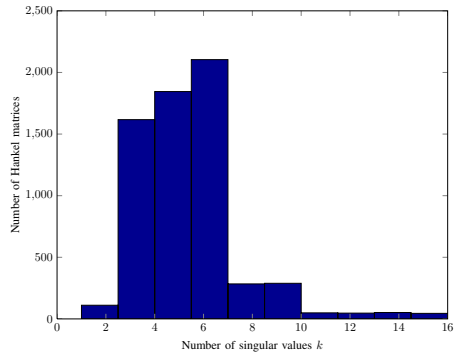

Figure 2. Histogram of the number of Hankel matrices whose top k singular values sum to at least 95% of their nuclear norm, vs. k.

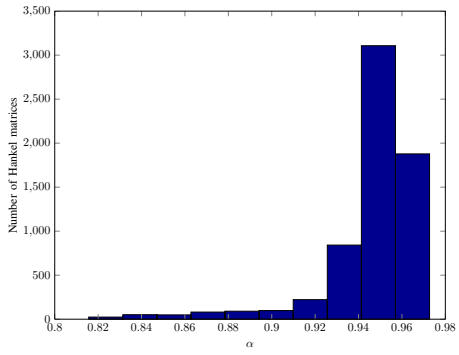

Figure 3. Histogram of the number of Hankel matrices whose top 5 singular values sum to at least α of their nuclear norm, vs. α.

B. Detailed experimental evaluation

Experiment 1: missing ICP values for random sampling steps. In this setting, 25% ICP data points are randomly selected and flagged as missing, while other vital signs remain unaffected. This situation happens when the devices occasionally stop working and recover back to normal immediately. A sequence of 30×30 Hankel matrices are constructed and all missing values are estimated. If one missing ICP value appears multiple times in the same Hankel matrix, its final estimated value is calculated as the arithmetic average of all its occurrences. [4]

Table I
RELATIVE ERROR (RE) AND NMSE FOR MISSING VALUE FILL-IN METHODS: OUR APPROACH(MC), k-NN ($k = 1, 5$), AND REGULAR EXPECTATION MAXIMIZATION (REGEM).

Method	Exp 1		Exp 2		Exp 3	
	RE	NMSE	RE	NMSE	RE	NMSE
MC	16.44%	0.035	32.05%	0.149	18.70%	0.046
KNN1	23.98%	0.074	52.10%	0.368	NA	NA
KNN5	20.93%	0.056	47.45%	0.295	NA	NA
RegEM	19.90%	0.051	40.06%	0.214	20.75%	0.056

Figure 4 shows a scatter plot of the true vs. the estimated value obtained from the four different fill-in methods. Points closer to the $y = x$ line indicate smaller difference between estimated and true values.

First, we note that most often the estimates of our approach (red stars) are more accurate than those of the other methods.

Second, we partition the range of ICP values into three intervals with clinical importance: ICP values less than 20 mmHg (normal), ICP between 20 and 30 mmHg (high pressure), and ICP greater than 30 mmHg (potentially lethal; requires prompt corrective action). Clinical protocols for TBI aim to keep ICP $<$ 20 mmHg. Therefore should remain low ($<$ 20), otherwise necessary medical intervention will be provided. Therefore, observations of ICP $\geq$ 20 mmHg will be infrequent and even fewer will be $\geq$ 30 mmHg. Figure 4 shows that our method still performs better than other methods for the rare cases of high ICP values. Table I shows that both the relative error and the NMSE of our method is smaller than that of the other three methods.

Experiment 2: missing ICP values for random long time periods. Instead of missing ICP values at random sample steps, ICP values are flagged as missing in random blocks of sampling steps. These non–overlapping blocks consists have their total duration 25% of the duration of the whole ICP stream, while each of them constitutes of 50 sampling steps and lasts span 5–minutes. Because the length of this missing gap is longer than the dimension of the Hankel matrix we construct, we may end with a whole row of the Hankel matrix missing. If this happens, the matrix completion method may work in an unexpected way. To mitigate this problem, we use estimated values for

[4] After each completion on a Hankel matrix, all the estimates are stored at different location for future analysis, instead of directly filling in the vacuums. Hence, previously estimated values will not affect following calculation.

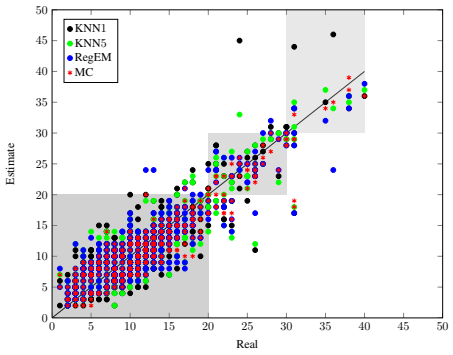

Figure 4. Scatter plot of performance of data imputation methods when missing ICP values at random sampling steps.

missing values in previous sampling steps when constructing the Hankel matrices for subsequent sampling steps.

Long duration of missing values lead to a deterioration of performance for the fill-in methods. But the Bland-Altman plots (Figure 6) show that our method provides better estimates, more (Estimate − Real) clustering at 0, and smaller standard deviation than other methods used.

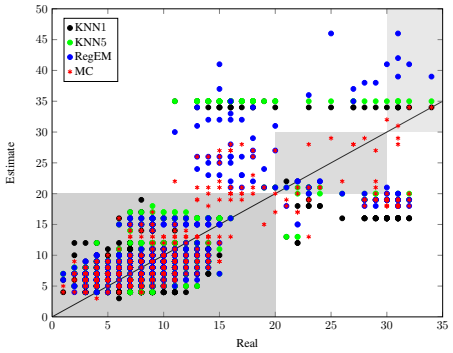

Figure 5. Scatter plot of performance of data imputation methods when missing ICP values for random long time periods.

Experiment 3: missing values for any vital sign at random sampling steps. In this experiment, a total 25% of all vital sign values are selected uniformly at random and are flagged as missing. This corresponds to common situations in the clinical environment. Because the k–nearest neighbor methods can not compute the distance between two observations when they are missing values on more than one variable, we only compare our approach with the regularized EM method. Figure 7 shows that both methods have reasonable performance, having their estimates crowded around the line $y = x$. Table I also shows that our approach gives smaller relative error and NMSE than the regularized EM method.

Experiment 4: missing ICP values for random sampling steps for a group of 10 patients. This experiment extends

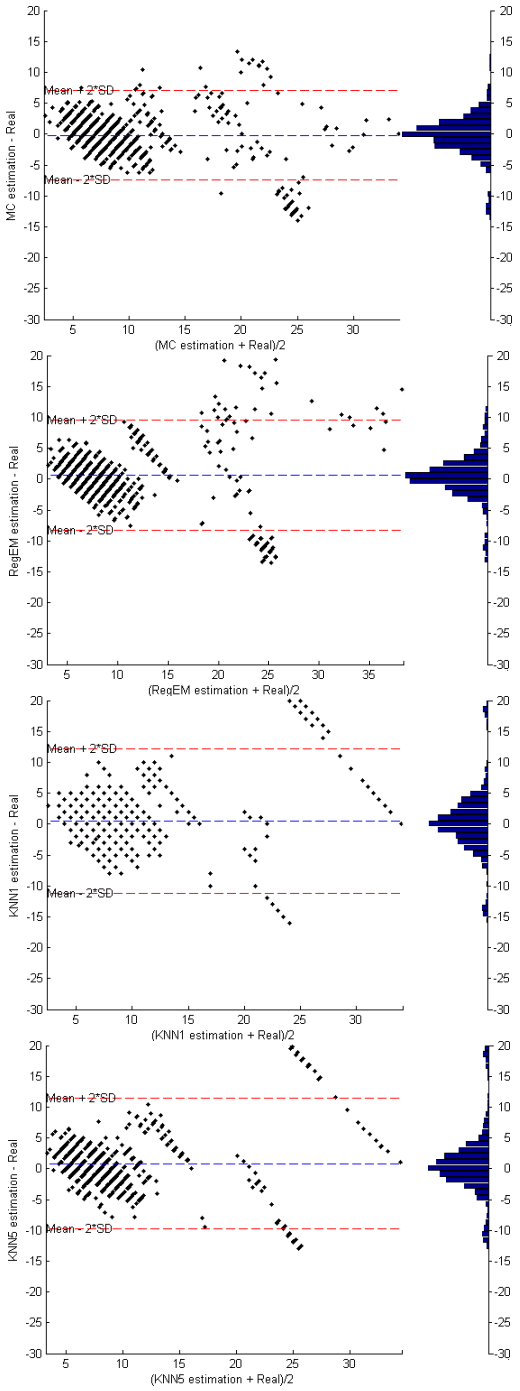

Figure 6. Bland-Altman plots of performance of data imputation methods when missing ICP values for random long time periods.

the population to 10 TBI patients (patients P1–P5 survived and P6-P10 died in the hospital). Patient P1 is the same patient used in experiments 1–3. A total of 160 hours of vital signs data (96,000 sampling steps) with complete ICP are identified and processed with the same parameter settings as in experiment 1. In Figure 8, each cluster of bars

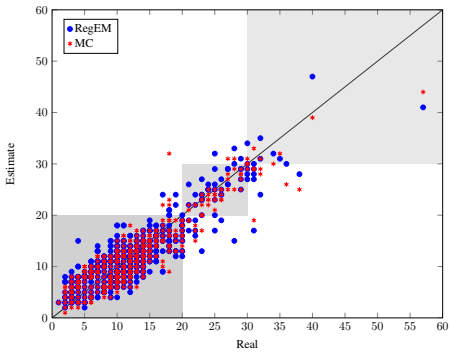

Figure 7. Scatter plot of performance of data imputation methods when missing values for any VS at random sampling steps.

represents the relative errors of the four methods for the corresponding single patient. Table II summarizes the overall relative error and NMSE for missing data estimated for all 10 patients. Despite different mortality outcomes and various vital sign patterns, we observe that our approach always outperforms the other fill-in methods. We also note that the mortality outcome does not seem to affect the accuracy of our approach.

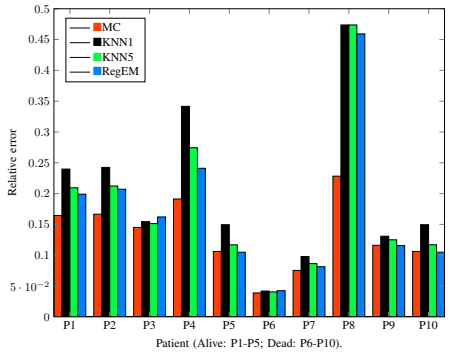

Figure 8. Performance of data imputation methods when missing ICP values for random sampling steps for 10 patients.

Table II
OVERALL RELATIVE ERROR (RE) AND NMSE OF DATA IMPUTATION METHODS WHEN MISSING ICP VALUES FOR RANDOM SAMPLING STEPS FOR THE 10 PATIENT DATASET.

	MC	KNN1	KNN5	RegEM
RE	12.69%	18.42%	16.20%	15.48%
NMSE	0.019	0.041	0.032	0.029

C. Discussion

In simulating random missing data, we identified a single subsection for each patient where all ICP measurements are available and most of the rest vital signs are also available. We selected 25% of data inside these selected segments, either for only ICP or for all vital signs. In reality, non-invasive vital signs (heart rate, blood pressure, etc.) are available $\geq 98\%$ of the time. Also due to the great effort of the medical staff and the benefit of automated electronic collection, ICP values are also available for rather long periods of time for many patients. Among the 10 patient we worked on, during the time period when ICP values were recorded, actual mean availability of complete data was ICP, 97.9%; HR, 98.2%; SBP, 96.6%; MBP, 96.6%; and CO_2, 30%. Thus, our hypothetical situation of 25% missing data grossly exaggerates actual data loss in our clinical data collection systems. However, this approach increases confidence in the capabilities of the methods.

Our approach of online fill-in of missing values consumes little CPU time. An Intel Core i5 2.67GHz CPU with 8.00GB memory running the 64-bit Windows 7 operating system, the Matlab 2012b implementation of our approach takes an average of 30 seconds to process a 30×30 Hankel matrix and do the fill-in, which is acceptable for real-time processing for clinical use. The matrix completion routine is adapted from the TFOCS package (Templates for First-Order Conic Solvers) [15].

V. CONCLUDING REMARKS

Continuous, automated, electronic patient vital sign data streams are important for clinical evaluation of the physiologic status of critical ill patients and the support of timely therapeutic interventions. Appropriately collected, complete vital signs sequences allow statistical and machine learning software to uninterruptedly supply VS needed by physicians for decision-making about therapeutic interventions. Such support can maximize the potential for saving lives and improving long-term clinical outcomes.

We proposed a novel method for the commonly encountered problem of missing values in vital signs data streams. Our method builds a sequence of Hankel matrices for the data stream, and utilizes recently developed matrix completion methods.

We demonstrated that a few singular values are sufficient to capture most of the nuclear norm of our Hankel matrices, which suggests that they are might be low rank matrices perturbed with high rank noise.

We compared our approach with commonly used data imputation methods, k–nearest neighbors ($k = 1, 5$) and the regularized EM method. We demonstrated that our approach provides better estimates of missing values for important vital signs variables like ICP than these other methods.

In future work, we plan to integrate our approach (as a prepossessing step) with the method of [6] to forecast near–term clinical outcomes, such as mortality and massive blood transfusion needs. We also plan to study the impact on the imputation performance of our approach of constructing Hankel matrices by including additional vital sign data stream variables. Motivated by the fact that VS data

streams are contaminated by artifacts caused by unstably attached sensors, or patient movement, we will use the robust PCA method to simultaneously perform data imputation and artifact removing.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their careful review and comments. This research was funded in part by the following grants: USAF: (FA8650-11-2-6D01) Continuing non-invasive monitoring and the development of predictive triage indices for outcomes following trauma; W81XWH-07-2-0118 Early Support of Intracranial Perfusion USAF: (FA8650-12-2-6D06) Noninvasive Intracranial Pressure Monitoring Using Advanced Machine Learning Techniques and USAF: (FA8650-11-2-6D03)-TimePO.

The opinions expressed in this paper are those of Authors and do not reflect the opinions of USAF.

In addition, the authors wish to thank the Physician and Nurses in Neuro ICU and the ONPOINT investigators group* for their support. *The ONPOINT investigators group includes: Amechi Anazodo, Steven Barker, John Blenko, Patrick Boyle, Chein-I Chang, Theresa Dinardo, Joseph duBose, Raymond Fang, Yvette Fouche, Lisa Gettings, Linda Goetz, Tom Grissom, Victor Guistina, George Hagegeorge, Anthony Herrera, John Hess, Peter Hu, Cris Imle, Matthew Lissauer, Colin Mackenzie, Jay Menaker, Karen Murdock, Tim Oates, Sarah Saccicchio, Thomas Scalea, Stacy Shackelford, Robert Sikorski, Lynn Smith, Lynn Stansbury, Deborah Stein, and Chris Stephens.

REFERENCES

[1] J. J. DuBose, G. Barmparas, K. Inaba, D. M. Stein, T. Scalea, L. C. Cancio, B. Eastridge, and L. Blackboume, "Isolated severe traumatic brain injuries sustained during combat operations: Demographics, mortality outcomes and lessons to be learned from contrasts to civilian counterparts," *J. Trauma*, vol. 70, pp. 11–18, 2011.

[2] R. P. Dutton, L. G. Stansbury, S. Leone, E. Kramer, J. R. Hess, and T. M. Scalea, "Trauma mortality in mature trauma systems: are we doing better? an analysis of trauma mortality patterns," *J. Trauma*, vol. 69, pp. 620–626, 2010.

[3] B. J. Eastridge, L. G. Stansbury, H. Stinger, L. Blackboume, and J. B. Holcomb, "Forward surgical teams provide comparable outcomes to combat support hospitals during support and stabilization operations on the battlefield," *J. Trauma*, vol. 66, no. 4 Suppl, pp. 48–50, 2009.

[4] S. Kahraman, R. P. Dutton, P. Hu, L. Stansbury, Y. Xiao, D. M. Stein, and T. M. Scalea, "Heart rate and pulse pressure variability are associated with intractable intracranial hypertension after severe traumatic brain injury," *J. Neurosurg Anesthesiol*, vol. 22, no. 4, pp. 296–302, 2010.

[5] S. Kahraman, R. P. Dutton, P. F. Hu, Y. Xiao, B. Aarabi, D. M. Stein, and T. M. Scalea, "Automated measurement of "pressure times time dose" of intracranial hypertension best predicts outcome after severe traumatic brain injury," *J. Trauma*, vol. 69, no. 1, pp. 110–118, 2010.

[6] K. Kalpakis, S. Yang, P. F. Hu, C. F. Mackenzie, L. G. Stansbury, D. M. Stein, and T. M. Scalea, "Outcome prediction for patients with severe traumatic brain injury using permutation entropy analysis of electronic vital signs data," in *8th International Conference on Machine Learning and Data Mining, MLDM 2012* (P. Perner, ed.), (Berlin), July 2012.

[7] D. M. Stein, P. F. Hu, M. Brenner, K. N. Sheth, K. Liu, W. Xiong, B. Aarabi, and T. M. Scalea, "Brief episodes of intracranial hypertension and cerebral hypoperfusion are associated with poor functional outcome after severe traumatic brain injury," *J. Trauma*, vol. 71, no. 2, pp. 364–373, 2011.

[8] C. K. Enders, *Applied Missing Data Analysis*. The Guilford Press, 2010.

[9] J. S. Haukoos and C. D. Newgard, "Advanced statistics: Missing data in clinical research - part 1: An introduction and conceptual framework," *Academic Emergency Medicine*, vol. 14, no. 7, pp. 662–668, 2007.

[10] M. J. Daniels and J. W. Hogan, *Missing Data in Longitudinal Studies: strategies for Bayesian modeling and sensitivity analysis*. Chapman & Hall/CRC, 2008.

[11] F. V. Nelwamondo, S. Mohamed, and T. Marwala, "Missing data: A comparison of neural network and expectation maximisation techniques," *eprint arXiv:0704.3474*, 2007.

[12] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[13] E. J. Candès and Y. Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.

[14] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[15] S. Becker, E. J. Candès, and M. Grant, "Templates for convex cone problems with applications to sparse signal recovery," *Mathematical Programming Computation*, vol. 3, no. 3, pp. 165–218, 2011.

[16] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?." arXiv:0912.3599v1.

[17] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, "Hankel matrix rank minimization with applications in system identification and realization." faculty.washington.edu/mfazel/Hankelrm9.pdf.

[18] H. H. Yang and Y. Hua, "On rank of block hankel matrix for 2-d frequency detection and estimation," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 1046–1048, 1996.

[19] H. Li, K. Zhang, and T. Jiang, "The regularized em algorithm," in *Proceedings of the 20th National Conference on Artificial Intelligence*, (Pittsburgh), pp. 807–812, 2005.