Missing Internet Traffic Reconstruction using Compressive Sampling

Indrarini Dyah Irawati[1], Andriyan Bayu Suksmono[2], Ian Joseph Matheus Edward[3]

[1,2,3]School of Electrical and Informatics, Institut Teknologi Bandung, Indonesia
[1]Telkom Applied Science School, Telkom University, Bandung, Indonesia

Abstract: Missing traffic is a commonly problem in large-scale network. Because the traffic information is needed by network engineering task for network monitoring, there are several methods that recover the missing problem. In this paper, we proposed missing internet traffic reconstruction based on compressive sampling. The main contributions of this study are as follows: (i) explore the influence of the six missing patterns on the performance of the traffic matrix reconstruction algorithm; (ii) trace the link sensitivity; and (iii) detect the time sensitivity of the network. Using Abilene data, the simulation results show that compressive sampling can perform internet traffic monitoring such as reconstruction from missing traffic, finding link sensitivity, and detecting time sensitivity.

Keywords: missing traffic, compressive sampling, reconstruction, link sensitivity, time sensitivity

1. Introduction

Nowadays, the growth of internet network increases rapidly. As a result of this growth, network monitoring becomes extremely difficult. Network monitoring is a process of collecting and analyzing information from a network and utilize this information for managing the network resources effectively. The information that can be used for monitoring are, such as traffic, packet loss, delay, etc. This information can be obtained from direct measurement of network devices such as routers or switches. The measurement results indicate information flow in a set of links between source nodes to destination nodes. The links connect the nodes in the network topology. In reality, such information may be incomplete due to nodes or links damages. Therefore, it is important to recover accurately information of all links from incomplete measurement because many network engineering tasks involved these information that very sensitive to missing information.

A simple method to estimate the missing information is local interpolation. It can recover the missing information at any position based on the known information. Some research have participated in this area, such as K-Nearest Neighbors (KNN) [1], interpolation [2], [3]. These methods have good performance only for estimating low missing probability, but perform poorly when high missing probability occurs.

Nowadays, there is an emerging method for recovering data from incomplete information named Compressive Sampling (CS). CS can recover missing values of a signal as far as that signal is sparse [4], [5], [6]. Since sparse or compressible signals are involved in many applications, CS has been employed in various fields such as image reconstruction [7], direction of arrival estimation [8], radar detection [9], [10], waveform recovery [11], network traffic monitoring [1], etc.

In this paper, we concentrate on the observation of faults in the network and solve these problems with CS technique. There are three focusses of this research, which are, the reconstruction of the missing traffic matrix (TM), the detection of link sensitivity, and the detection of time sensitivity.

We compare 4 CS reconstruction methods, which are Sparsity Regularized Singular Value Decomposition (SRSVD) [1], L1 norm optimization [12], Iteratively Reweighted Least Square (IRLS) [13], Orthogonal Matching Pursuit (OMP) [14]. To compare with standard time series analysis method, we use Interpolation technique [1], [15].

We apply various missing patterns as cases on the actual network. This missing scenario is randomly chosen with missing probability from 0.01-0.98. This observation aims to obtain whole internet traffic information from the limited source.

In this paper, the row of TM represents a connection between nodes, whereas column represents the time series of observations. Link sensitivity indicates how a certain link will affect the traffic on the whole network. We remove a row of TM one by one in order to know the sensitive of a row that is currently influence the results of reconstruction. Detection link sensitivity has a purpose of determining the best path that can be passed over the network and isolate the bad links that cause the anomaly.

Time sensitivity indicates how a certain part of the time will influence the whole network. In this paper, we drop one by one the column of TM in order to know which is very dominant on the results of reconstruction. This information is very useful for decision-making in the network settings.

This paper is arranged in a systematic Section as follows. The related work on CS approaches for network monitoring is presented in Section 2. Section 3 points out the research methods. The experiment results obtained from simulation using Matlab are given in Section 4. Finally, conclusions are provided in the last Section, also the future work.

2. Related Work

Studies on CS applications for network monitoring had been done, for example, by Roughan et al. [1], and Chen, et al. [16]. Roughan et al. proposed Sparsity Regularized Matrix Factorization (SRMF) for solving various of TM issues covering network tomography, detection an anomaly, and prediction of traffic [1]. In [16], Chen, et al. proposed LEN decomposition technique that enables to present of missing data, calculation errors and traffic anomalies for network analytics.

Huibin et al. used spatial traffic similarity and temporal smoothness feature to reconstruct the missing traffic and reported that CS has the capability of reconstructing missing traffic up to 98% from the total traffic with error reconstruction below 32% [17]. Nie et al., developed a method to fix traffic from end to end increasingly accurate using Flow Sensing Reconstruction (FSR) [18].

In traffic anomaly (for example, unusual traffic spikes on several links), CS can detect, identify, and quantify the anomaly [19], [20], [21], [22]. Lakhina, et al., detected the anomaly of time series traffic matrix using Principal Component Analysis (PCA) [19]. This method does not associate temporal correlation between the time series. In another work [20], the authors improved by detecting and localizing congested links using greedy iteration algorithm. In [21], Bandara, et al., proposed adaptive CS that achieves 99% fault detection rate for network fault localization. Davina, et al., study a complex correlation to detect attacks on the large-scale network [23].

In network tomography, CS provides internal performance and QoS characteristics of a network using information from endpoint data [21], [24], [25]. Vardi created a term network tomography that indicates to the issue of approximating traffic matrix from link measurement on the network [24]. He used Poisson distribution to represent traffic matrix, but this method can not always model the real condition of network traffic. In [25], the authors solved the problems of link delay estimation in the congested network using Fast Reference-based Algorithm for Network Tomography via Compressive Sensing (FRANTIC) algorithm.

In traffic forecasting, Principal Component Analysis (PCA) predicted the traffic that flows between source-destination in the network for future user demands [26]. This paper showed low effective rank in measurements using temporal matrix traffic.

3. Research Method

There are several parameters that can be observed in a network, such as the traffic flow, packet delay, and packets loss. These parameters are used to monitor and analyze the quality of the network.

Traffic is the amount of data that traverses in the network. It often happens that due to nodes (routers or servers) shut down or links break, there is some missing traffic in certain nodes or links at a certain time. Missing traffic on node i at time j will correspond to a loss value on traffic matrix at (i,j) position. Detail description on traffic matrix will be discussed in the next section.

3.1 Traffic Matrix Representation

Traffic flow on a network that has n nodes at a certain time can be represented as a matrix $\boldsymbol{\Psi}$ with dimension of $n \times n$. An element (i,j) on $\boldsymbol{\Psi}$ represents traffic flow from node i to node j. This matrix $\boldsymbol{\Psi}$ is called instantenous traffic matrix. Since the observation of traffic flow is done over a certain time interval (ΔT) and updated regularly, then it is convenience to rearrange the matrix $\boldsymbol{\Psi}$ of dimension $n \times n$ for observation time ΔT into a vector $\bar{\boldsymbol{\Psi}}$ of dimension $n^2 \times 1$. Mathematically, any element $\psi(i,j) \in \boldsymbol{\Psi}$ is translated into $\bar{\psi}(p,1) \in \bar{\boldsymbol{\Psi}}$,where $p = i + n \times (j-1)$; $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$. As the observation is done on regular time interval T_k $(k = 1, 2, \ldots, T)$, the traffic matrix $\bar{\boldsymbol{\Psi}}$ at each time interval T_k can be collected column-wise to produce $n^2 \times T$ link-time traffic matrix $\boldsymbol{X}$. Any column k of matrix $\boldsymbol{X}$ represents the traffic of network at time k.

Consider a simple example of a network consisting of three nodes as shown in Figure 1. Using this particular network, we have the instantenous traffic matrix $\boldsymbol{\Psi}$, reshaped instantenous traffic matrix $\bar{\boldsymbol{\Psi}}$, and traffic matrix $\boldsymbol{X}$ which are

$$\boldsymbol{\Psi} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix}, \tag{1}$$

$$\bar{\boldsymbol{\Psi}} = \begin{bmatrix} \psi_{11} \\ \psi_{21} \\ \psi_{31} \\ \psi_{12} \\ \psi_{22} \\ \psi_{32} \\ \psi_{13} \\ \psi_{23} \\ \psi_{33} \end{bmatrix}, \tag{2}$$

$$\boldsymbol{X} = \begin{bmatrix} \psi_{11,1} & \psi_{11,2} & \cdots & \psi_{11,T} \\ \psi_{21,1} & \psi_{21,2} & \cdots & \psi_{21,T} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \psi_{33,1} & \psi_{33,2} & \cdots & \psi_{33,T} \end{bmatrix}. \tag{3}$$

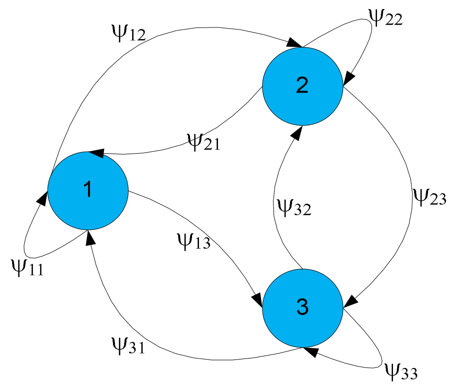


Figure 1. A network with 3 nodes

3.2 Link Numbering Assignment

As discussed previously, any $x(i,j)$ represents a traffic link from node i to node j. It is more convenience to renumbering all the link $(i,j), i = 1, 2, \ldots, n; j = 1, 2, \ldots, n$, into ke link number (LN) with $LN = 1, 2, \ldots, n^2$. Using this convention, the traffic matrix $\boldsymbol{X}_{\boldsymbol{ij,k}}$ can be denoted as $\boldsymbol{X}_{\boldsymbol{LN,k}}$ which indicates the traffic at link$-LN$ at time k. Table 1 shows an example of link number assignment of a network in Figure 1.

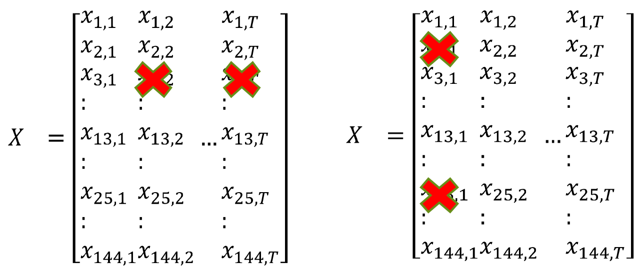
Table 1. Link number assignment of network with 3 nodes

Link (i,j)	Link Number Assignment
(1,1)	1
(2,1)	2
(3,1)	3
(1,2)	4
(2,2)	5
(3,2)	6
(1,3)	7
(2,3)	8
(3,3)	9

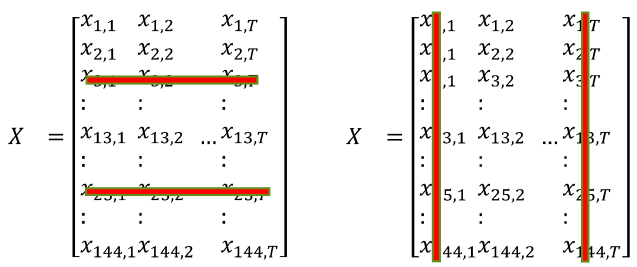
3.3 Missing Pattern

Generally, traffic measurement on a network may encounter incomplete data since links or nodes can fail, which yields missing values. In this paper, we evaluate six kinds of missing patterns on traffic matrix as shown in Figure 2.

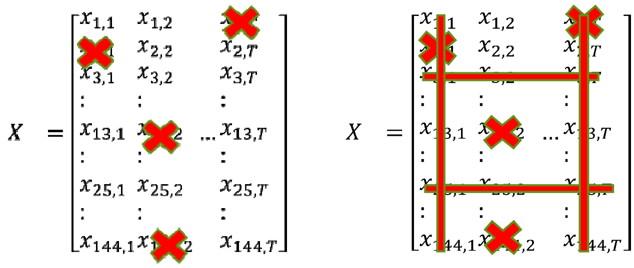
a) **Missing Row Elements (MRE)**: this event simulates the loss of randomly selected TM elements from a random row. The row is randomly chosen, and the TM elements in the row are randomly chosen with missing probability p. This simulation illustrates the quality of the certain link state at a specific time.
b) **Missing Column Elements (MCE)**: this simulates the loss by selecting a column randomly and missing TM elements from it at random with missing probability p. This simulation illustrates some missing data at a certain time. This event is caused, for example, by overloading at router data monitoring.
c) **Missing Rows at Random (MRR)**: this simulates the loss by selecting entire rows randomly with missing probability p. This event ilustrates some links failure or routers down for a long time.
d) **Missing Columns at Random (MCR)**: this simulates the entire column loss that selected randomly with missing probability p. There is no data at a certain time. In real situation, this case corresponds to the network breaks down or the software for monitoring in router fails.
e) **Missing Elements at Random (MER)**: this simulates missing elements of TM at random with missing probability p. This represents data loss on a particular link in a certain time.
f) **Combine Missing Patterns (CMP)**: this simulation is a combination of missing rows, missing columns, and missing elements. Each missing component is chosen randomly with missing probability p.



(a). MRE (b). MCE



(c). MRR (d). MCR



(e). MER (f). CMP

Figure 2. Missing pattern, (a). MRE, (b). MCE, (c). MRR, (d). MCR, (e). MER, (f). CMP

3.4 Low-rank Representation using Singular Value Decomposition (SVD)

It is important to translate the traffic matrix into a basis such that the TM is sparse or compressible in that basis. One of such basis is Singular Value Decomposition (SVD), which can be used to decompose the TM ($\boldsymbol{X}$). Formally, the SVD of a $N \times T$ matrix $\boldsymbol{X}$ is denoted by [27]:

$$\boldsymbol{X} = \boldsymbol{U\Sigma V^T}, \tag{4}$$

where $\boldsymbol{U}$ is an $N \times N$ orthonormal matrix, i.e. $\boldsymbol{UU^T} = \boldsymbol{U^T U} = \boldsymbol{I}$. $\boldsymbol{V}$ is an $T \times T$ orthogonal matrix, i.e. $\boldsymbol{VV^T} = \boldsymbol{V^T V} = \boldsymbol{I}$, with $\boldsymbol{V^T}$ is the transpose of $\boldsymbol{V}$, and $\boldsymbol{\Sigma}$ is a $N \times T$ diagonal matrix that has content non-zero entries referred to as the singular values σ_i of $\boldsymbol{X}$. These values are structured with the result that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_i$. The total number of non-zero singular values indicate the number of rank matrix. If $i \ll min\,(N, T)$, then TM is a low-rank matrix. For low-rank approximation, Equation (4) can be rewritten as:

$$\boldsymbol{X} = \sum_{i=1}^{\min\,(N,T)} \sigma_i\, \boldsymbol{u_i v_i^T} = \sum_{i=1}^{\min\,(N,T)} \sigma_i\, \boldsymbol{B_i}\,, \tag{5}$$

where u_i and v_i are the i^{th} columns of U and V. The $\boldsymbol{B_i}$ is a matrix constructed by rank-1. Hence, a rank-r approaches to $\widehat{\boldsymbol{X}}$ using the SVD by holding the greatest singular values in the addition and discarding the others is equivalent to

$$\widehat{\boldsymbol{X}} = \sum_{i=1}^{\mathrm{r}} \sigma_i\, \boldsymbol{B_i} \tag{6}$$

Since $\widehat{\boldsymbol{X}}$ can be represented using i elements of singular value of SVD and $i \ll min\,(N, T)$ then $\widehat{\boldsymbol{X}}$ is sparse in SVD base.

3.5 Routing Matrix Representation

The routing matrix $\boldsymbol{A}$ is a $M \times N$ dimension, where M is the number of links and N is the number of source-destination pairs. The matrix $\boldsymbol{A} = a_{ij}$ is defined as $a_{ij} = 1$ if link i is part of the path for source-destination pair j, otherwise $a_{ij} = 0$. The routing matrix $\boldsymbol{A}$ is used as the sensing matrix for the compression of the approximation matrix $\widehat{\boldsymbol{X}}$ [17].

3.6 Compressive Sampling (CS)

CS is a new paradigm in signal processing [4], [28], with a capability to reconstruct data from an incomplete measurement. In order to perform successful reconstruction, CS exploit the special structure in the data structure which is the data sparsity. Data is called sparse if it can be expressed by a small number of significant components. In the TM case using SVD as the base, sparsity means low-rank as the energy spectrum composed of low-rank singular values of the matrix. It is well known that TM can be estimated from a low-rank matrix [19], [26].
Elements missing in traffic matrix can be considered as elements loss in compression. Therefore, it can be considered as a compression step in CS. On the TM reconstruction using CS, there is relationship between the measured traffic $\boldsymbol{Y}$ and the TM ($\boldsymbol{X}$) that expressed in a linear matrix equation as follows

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}\ , \qquad (7)$$

with $\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{X}$ are $M \times T$, $M \times N$, and $N \times T$ matrix respectively. $\boldsymbol{A}$ is a routing matrix which performs as measurement matrix in CS. It satisfies the Generalized Uncertainty Property (GUP) [29]. Corresponding with the theory of CS, the minimum number of rows in $\boldsymbol{A}$ is given as follows [30]

$$M \geq C\ (R\ log\frac{N}{R})\ , \qquad (8)$$

where C is determined empirically with a range in value from 1 to 2, while R is the total of rank matrix used.

3.7 Reconstruction Algorithm

3.7.1 SRSVD

SRSVD expressed SVD as a matrix factorization $\boldsymbol{X}$ and applied a regularization parameter to optimize the reconstruction of missing value. The factorization of $\boldsymbol{X}$ is equivalent to the form:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{L}\boldsymbol{R}^{\boldsymbol{T}}, \qquad (9)$$

where $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{R} = \boldsymbol{V}\boldsymbol{\Sigma}^{1/2}$. The low-rank feature has proved that it is possible to restore the missing values in a matrix. The minimization of low-rank composition can be expressed as follows:

$$\min rank(\boldsymbol{X}),\ \text{subject to } \boldsymbol{\mathcal{A}}(\boldsymbol{X}) = \boldsymbol{B}\ , \qquad (10)$$

where $rank(\cdot)$ as the rank of a matrix, $\boldsymbol{\mathcal{A}}(\cdot)$ is a linear operator that executes on matrix $\boldsymbol{X}$, and $\boldsymbol{B}$ denotes the collection of actual measurements. Low norm factorization from (9) satisfies Frobenius so that (10) is equivalent to:

$$\min rank\|\boldsymbol{L}\|_F^2 + \|\boldsymbol{R}\|_F^2\ , \text{subject to } \boldsymbol{\mathcal{A}}(\boldsymbol{L}\boldsymbol{R}^T) = \boldsymbol{B} \qquad (11)$$

In addition, because the TM ($\boldsymbol{X}$) is uncertain low-rank, then the regularized parameter (λ) is added to solve the optimization problem:

$$\min\|\boldsymbol{\mathcal{A}}(\boldsymbol{L}\boldsymbol{R}^T) - \boldsymbol{B}\| + \lambda(\|\boldsymbol{L}\|_F^2 + \|\boldsymbol{R}\|_F^2)\ , \qquad (12)$$

where λ is a parameter of regularization that keeps the tendency of precise measurement with the low-rank target. This technique is known as SRSVD that proposed in [1].

3.7.2 SVDL1

The reconstruction of TM is discovering optimal solution for $\widehat{\boldsymbol{X}}$ into equation (7) of the measured traffic $\boldsymbol{Y}$ that approaches the original $\boldsymbol{X}$ as near as possible, i.e.,

$$\min \left\|\boldsymbol{X} - \widehat{\boldsymbol{X}}\right\|_1^1, \text{subject to } \boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}\ , \qquad (13)$$

where $\|.\|_1^1$ states the l_1 norm that used to calculate the error between the original TM ($\boldsymbol{X}$) and the reconstructed TM ($\widehat{\boldsymbol{X}}$) [12].

3.7.3 IRLS

IRLS finds the optimal solution for rank minimization using l_p norm, where $0 \leq p \leq 1$. IRLS sets weight (W_p) to estimate TM ($\boldsymbol{X}$) and a parameter of regularization (γ) to guarantee that W_p is proper selected. The procedure of the algorithm are as follows [13]:

(i) Initialize the weight $\boldsymbol{W_p^0} = \boldsymbol{I}$, regularization parameter $\gamma_1 > 0$, and the iteration counter $k = 1$. While not convergen do:

(ii) Solve the minimum solution:

$$\boldsymbol{X}^k = \arg\min_{\boldsymbol{X}} \boldsymbol{T_r}(\boldsymbol{W_p^{k-1}}\,\boldsymbol{X}^T\boldsymbol{X}), \text{st } \boldsymbol{\mathcal{A}}(\boldsymbol{X}) = \boldsymbol{B} \qquad (14)$$

(iii) Find the new weight:

$$\boldsymbol{W_p^k} = [(\boldsymbol{X^{k^T}}\boldsymbol{X^k})^2 + \gamma^k\boldsymbol{I}]^{\frac{p}{2}-1}\ , \qquad (15)$$

(iv) Choose regularization parameter:

$$0 < \gamma^{k+1} < \gamma^k \qquad (16)$$

Increase k and return to (ii) until convergence. This algorithm reduces weighted using Frobenius norm of the traffic matrix ($\boldsymbol{X}$) at each iteration, which is

$$\boldsymbol{T}_r(\boldsymbol{W_p^k}\boldsymbol{X}^T\boldsymbol{X}) = \left\|(\boldsymbol{W_p^k})^{\boldsymbol{1/2}}\boldsymbol{X}\right\|_{\boldsymbol{F}}^{\boldsymbol{2}} \qquad (17)$$

3.7.4 OMP

OMP estimates TM ($\boldsymbol{X}$) which determines columns of A take part in the calculation of $\boldsymbol{Y}$. The concept of the algorithm is to choose columns in a greedy model. A column of $\boldsymbol{A}$ which has very high correlation with the remaining part of $\boldsymbol{Y}$ is chosen at each iteration. Then the $\boldsymbol{Y}$ residue is substracted with the result generate the new residue. The procedure of algorithm uses the columns to recognize the proper set after K iteration. The algorithm processes are given as follows: [14].

(i) Set the residual $\boldsymbol{R_0} = \boldsymbol{Y}$, index set $\boldsymbol{\Lambda_0} = \emptyset$, and the iteration counter $k = 1$.

(ii) Select the parameter λ_k that resolves the case of optimization:

$$\lambda_k = \arg max_{i=1..M} |\langle \boldsymbol{R_{k-1}}, A_i \rangle| \quad (18)$$

If the dot product produces maximum value then solve the tie deterministically.

(iii) Add λ_k to the the index set:

$$\Lambda_k = \Lambda_{k-1} \cup \{\lambda_k\} \quad (19)$$

The matrix of chosen atoms:

$$\boldsymbol{A_k} = [\boldsymbol{A_{k-1}} \;\; a_{\lambda_k}] \quad (20)$$

We take $\boldsymbol{A_0}$ initialization parameter as an empty matrix.

(iv) Calculate the estimation using a least squares procedure:

$$\boldsymbol{X_k} = \arg min_{\boldsymbol{X}} \|\boldsymbol{Y} - \boldsymbol{A_k X}\|_2 \quad (21)$$

(v) Count the new approach of the data and the residual:

$$a_k = \boldsymbol{A_k X_k}\ , \quad \boldsymbol{R_k} = \boldsymbol{Y} - a_k \quad (22)$$

Increment the iteration k and returns to equation step (ii) until $k < K, K$ is the sparsitas level. The estimate of $\widehat{\boldsymbol{X}}$ for the best possible matrix has non-zero denotes at the index listed Λ_k. The estimation result $\widehat{\boldsymbol{X}}$ in component λ_i is equivalent to the i_{th} component of $\boldsymbol{X_k}$.

3.7.5 Interpolation

Interpolation is an easy method that takes the value from closest neighbor to construct incomplete values. The TM ($\boldsymbol{X}$) interpolation is performed for each column with the length of $\boldsymbol{X}$ rows and for each row with the length of $\boldsymbol{X}$ columns. Interpolation formula is shown as follows [1] [15].

$$\widehat{\boldsymbol{X}}(\boldsymbol{i},\boldsymbol{j}) = mean\left(\overline{\boldsymbol{X}}_{\boldsymbol{row}}(\boldsymbol{i}) + \overline{\boldsymbol{X}}_{\boldsymbol{col}}(\boldsymbol{j})\right), \quad (23)$$

$$\overline{\boldsymbol{X}}_{\boldsymbol{row}}(\boldsymbol{i}) = mean\left(\boldsymbol{X}_{\boldsymbol{row}}(\boldsymbol{i}+\mathbf{1}) + \boldsymbol{X}_{\boldsymbol{row}}(\boldsymbol{i}-\mathbf{1})\right), \quad (24)$$

$$\overline{\boldsymbol{X}}_{\boldsymbol{col}}(\boldsymbol{j}) = mean\left(\boldsymbol{X}_{\boldsymbol{col}}(\boldsymbol{j}+\mathbf{1}) + \boldsymbol{X}_{\boldsymbol{col}}(\boldsymbol{j}-\mathbf{1})\right), \quad (25)$$

where $\widehat{\boldsymbol{X}}(\boldsymbol{i},\boldsymbol{j})$ as an estimation of $\boldsymbol{X}(\boldsymbol{i},\boldsymbol{j})$. $\overline{\boldsymbol{X}}_{\boldsymbol{row}}(\boldsymbol{i})$ is an estimation in row, where $\boldsymbol{X}_{\boldsymbol{row}}(\boldsymbol{i}+\mathbf{1})$ represents the nearest neighbor row above $\boldsymbol{X}(\boldsymbol{i})$, while $\boldsymbol{X}_{\boldsymbol{row}}(\boldsymbol{i}-\mathbf{1})$ represents the nearest neighbor row below $\boldsymbol{X}(\boldsymbol{i})$. $\overline{\boldsymbol{X}}_{\boldsymbol{col}}(\boldsymbol{j})$ is an estimation in column, where $\boldsymbol{X}_{\boldsymbol{col}}(\boldsymbol{j}+\mathbf{1})$ expresses the nearest neighbor column on the right of $\boldsymbol{X}(\boldsymbol{j})$, and $\boldsymbol{X}_{\boldsymbol{col}}(j-1)$ expresses the nearest neighbor on the left of $\boldsymbol{X}(\boldsymbol{j})$.

3.7.6 The Proposed Method of Internet Traffic Matrix Reconstruction

This section describes the proposed method of missing internet traffic reconstruction using compressive sampling. The process of traffic matrix reconstruction in this research is shown in Figure 3.

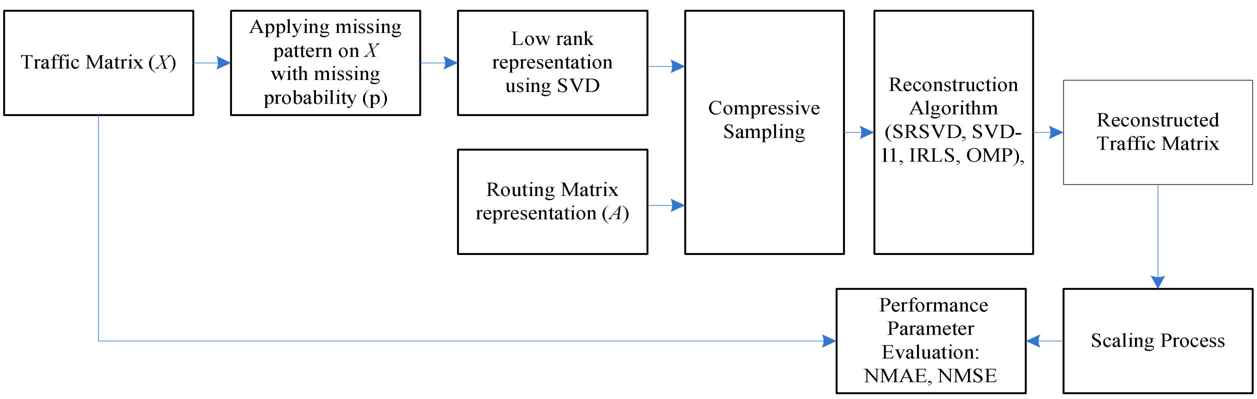


Figure 3. Missing internet traffic reconstruction processing using compressive sampling

Traffic that we used is actual traffic on the network. The TM is deleted with missing probability p. After SVD analysis, the significant singular values are used to construct the estimated TM as given in (6). The CS is applied on this estimated TM by using sensing matrix $\boldsymbol{A}$. The missing elements are recovered by CS reconstuction algorithms which are SRSVD, SVDL1, IRLS, OMP, and compared with Interpolation technique. The scaling function is applied on the reconstructed CS in order to get the comparable amplitude value for each algorithm. The performance of reconstruction algorithms are explored by comparing reconstructed value to the original TM in term of the NMAE and NMSE.

4. Experiment Results and Analysis

4.1 Data

We use the backbones of Abilene topology that is shown in Figure 4. The Abilene is a high speed network used for research and education which operates in the US and the information is available online [31]. This network consists of 12 nodes ($n = 12$) so that there are 12×12 links between source-destination ($N = 144$). Traffic measurement performed every 5 minutes, hence for a day, there are 288-time series ($T = 288$).

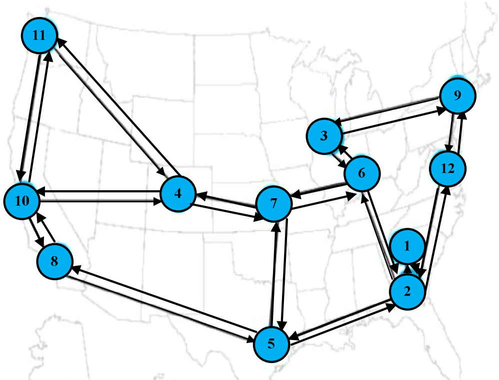


Figure 4. Abilene Topology [31]

4.2 Performance Parameter

The performance of CS algorithms is analyzed through parameters as given below which are commonly used in traffic matrix application [17], [1]. The metrics used to evaluate the accuracy of the error are Normalized Mean Absolute Error (NMAE) and Normalized Mean Square Error (NMSE).

$$NMAE(\boldsymbol{X}, \widehat{\boldsymbol{X}}) = \frac{\sum_{i,j:A(i,j)=0}^{N} \left| X(i,j) - \widehat{X}(i,j) \right|}{\sum_{i,j:A(i,j)=0}^{N} |X(i,j)|} \quad (26)$$

$$NMSE(\boldsymbol{X}, \widehat{\boldsymbol{X}}) = \frac{\left\| X(i,j) - \widehat{X}(i,j) \right\|_2^2}{\| X(i,j) \|_2^2} \quad (27)$$

where $\widehat{\boldsymbol{X}}$ is the reconstruction traffic matrix. $\boldsymbol{A}(\boldsymbol{i},\boldsymbol{j}) = 0$ represents matrix with the values eliminated on $\boldsymbol{X}(\boldsymbol{i},\boldsymbol{j})$. NMAE just counts the errors on the lost values.

4.3 Comparison of Reconstruction Algorithms in Different Missing Patterns

We analyze the impact of the missing patterns on the reconstruction algorithms performance. TM data in a day are discarded randomly with probability (p) from 0.02 to 0.98. We assume the missing pattern follows the uniform distribution so that all value of the TM has an identical probability to be discarded. For this particular experiment, we use Abilene traffic data on April 1st, 2004.

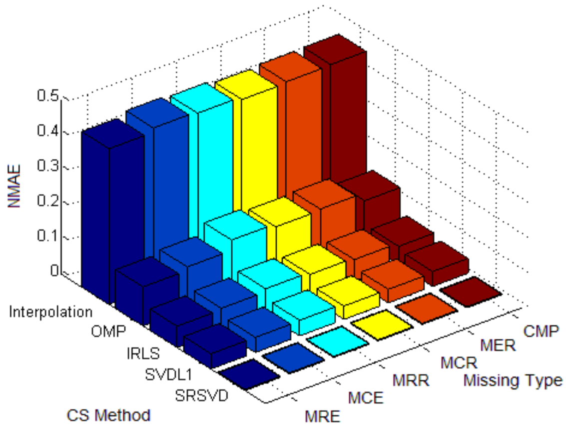
The simulation results of six data missing scenario using 5 CS reconstruction algorithms are shown in Figure 5. The X-axis expresses the CS method, the Y-axis states the missing pattern, and the Z-axis represents the value of NMAE.

In general, the reconstruction results depend on the position of the missing data and the amount of missing data. NMAE increases along with missing probability. The performance shows that SRSVD is the best method than others both in low and high missing probability.

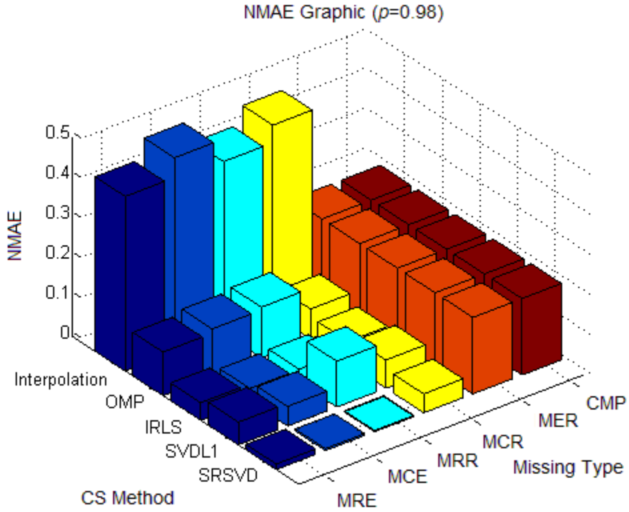
In MRE pattern simulation, the row is randomly chosen, then the elements in the row are randomly chosen with probability p. This missing pattern is easily reconstructed because the amount of data sample provided as more than other patterns.

In MCE pattern simulation, the column is randomly chosen, then elements in a column are randomly chosen with probability p. Basically, MCE pattern is similar to the MRE pattern where the number of samples provided is more than other so that it has a good value of NMAE.

In MRR and MCR pattern, the reconstruction results are worse than MRE and MCE pattern. This is due to less amount of sample data so that the CS reconstruction has difficulty to convert to the correct value.



(a). Missing probability (p=0.02)



(b). Missing probability (p=0.98)

Figure 5. Comparison between reconstruction algorithms for different missing patterns, (a). Missing probability (p=0.02), (b). Missing probability (p=0.98)

In MER pattern simulation, the missing elements are randomly chosen. The MER is a missing pattern that has difficult structure to be reconstructed for all reconstruction algorithms, especially at a high probability of missing. It is because the random missing pattern in MER has no correlation in the available sample elements.

In CMP pattern simulation, the missing pattern is set by the combination of MRE, MCE, MRR, MCR, and MER. Because the entire missing patterns are selected at random, there is a chance that a few patterns yield the same missing data. The pattern can still be reconstructed, although the results are worse as compared to those of MRE, MCE, MRR, and MCR.

4.4 Link Sensitivity Detection

The link sensitivity detection is done by eliminating one by one the row of TM sequentially. The row represents the link connection between a source node to destination node. The link connection number is shown in Table 2. The connections consist of loop connection, direct connection, and indirect connection as shown in Figure 6.

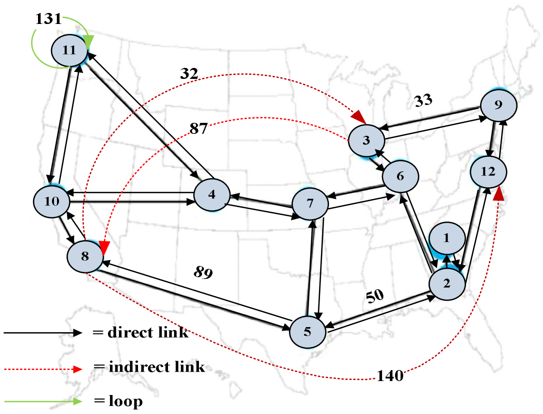


Figure 6. Representation of link connection

We observed the effect of the missing link to NMSE parameter for a month in April 2004. Figure 7 shows link sensitivity detection for a particular day in different reconstruction algorithms. The X-axis represents the missing of link connection, and the Y-axis represents the NMSE value. All reconstruction algorithms have the same results for link sensitivity detection.

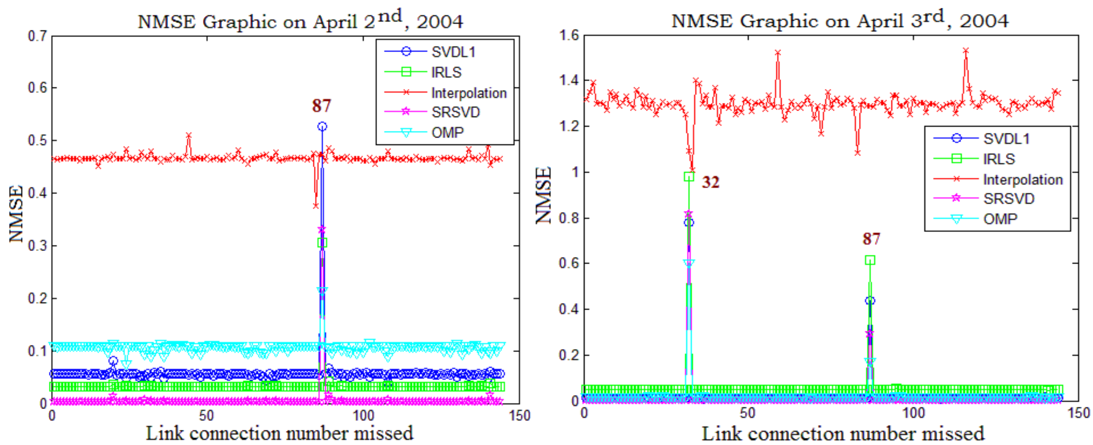
Table 3 shows the five highest sensitive links observed during that particular month. Link number 87 is an indirect link that connecting between node-3 to node-8, while link number 32 is also an indirect link that connecting between node-8 to node-3.

Table 2. Representation of link connection number between source node to destination node

		Destination Node											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Source Node	**1**	1	13	25	37	49	61	73	85	97	109	121	133
	2	2	14	26	38	50	62	74	86	98	110	122	134
	3	3	15	27	39	51	63	75	87	99	111	123	135
	4	4	16	28	40	52	64	76	88	100	112	124	136
	5	5	17	29	41	53	65	77	89	101	113	125	137
	6	6	18	30	42	54	66	78	90	102	114	126	138
	7	7	19	31	43	55	67	79	91	103	115	127	139
	8	8	20	32	44	56	68	80	92	104	116	128	140
	9	9	21	33	45	57	69	81	93	105	117	129	141
	10	10	22	34	46	58	70	82	94	106	118	130	142
	11	11	23	35	47	59	71	83	95	107	119	131	143
	12	12	24	36	48	60	72	84	96	108	120	132	144

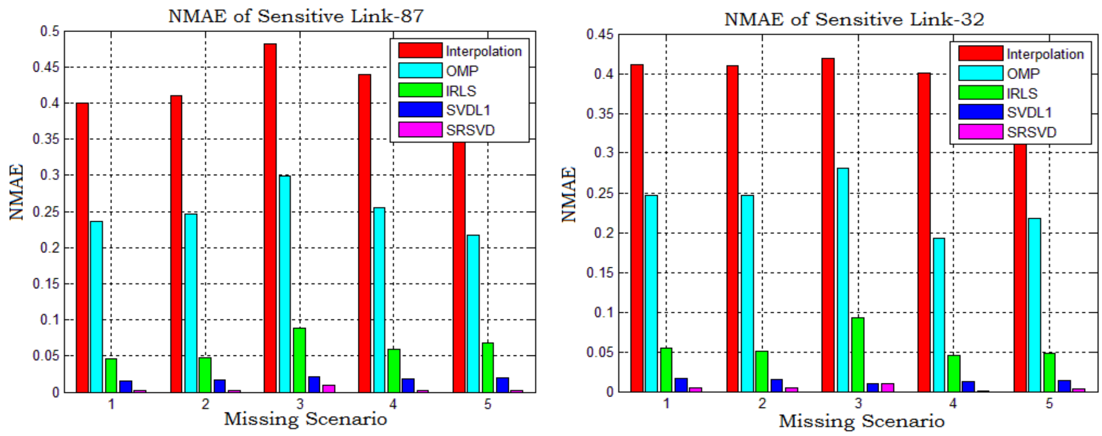
Table 3. The 5 highest link sensitivity detection for a month (April 1^{s}t, 2004 – April 30th, 2004)

Number	**Link Connection Number**	**Source-Destination Node Representation**	**Link Connectivity**
1	87	3-8	Indirect
2	32	8-3	Indirect
3	89	5-8	Direct
4	134	2-12	Direct
5	141	9-12	Direct



(a). NMSE on April 2nd, 2004 (b). NMSE on April 3rd, 2004

Figure 7. Detection of link sensitivity for a particular day, (a). NMSE on April 2nd, 2004, (b). NMSE on April 3rd, 2004



(a). NMAE of sensitive link – 87 (b). NMAE of sensitive link – 32

Figure 8. Missing scenario for tracing sensitive link composed, (a) NMAE of sensitive link number 87, (b) NMAE of sensitive link number 32

Because the link number 87 and 32 are indirect links, then we can trace for any links that composed these links. Some possible links compositions are as shown in Table 4. To find out the problems as well as solutions for each sensitive link, we compare the scenarios by looking at the effect of the missing scenario to NMAE value.

Table 4. The missing scenarios for detecting link sensitivity compositions

Sensitive Link	Missing Scenario	Composed by link number
87	1	63-78-55-89
	2	63-18-50-89
	3	99-141-24-50-89
	4	63-78-43-112-94
	5	63-78-43-124-119-94
32	1	56-77-67-30
	2	56-17-62-30
	3	56-17-134-108-33
	4	116-46-76-67-30
	5	116-130-47-76-67-30

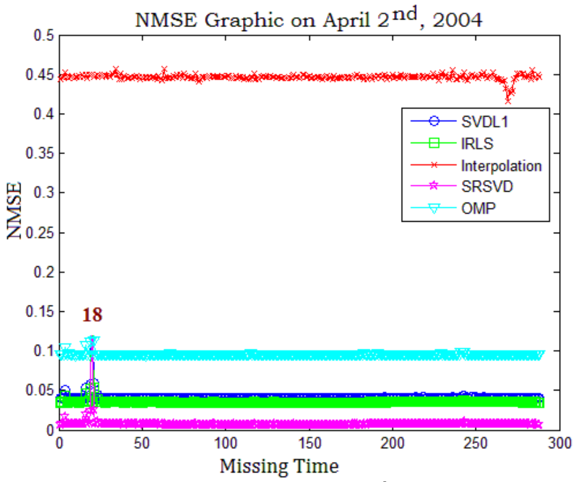
Figure 8 shows the comparison between all missing scenarios using the NMAE value. It can be concluded that the worst link composition on the sensitive link number 87 is scenario 3 (arranged by link number 99-141-24-50-89). The best link composition is scenario 4 which is arranged by link number 63-78-43-112-94.

Whereas on the sensitive link number 32, the worst composition is scenario 3 that composed by link number 56-17-134-108-33. The best forming is scenario 4 that arranged by link number 116-46-76-67-30. The best composition can be selected as a selected route for improving the performance of the network

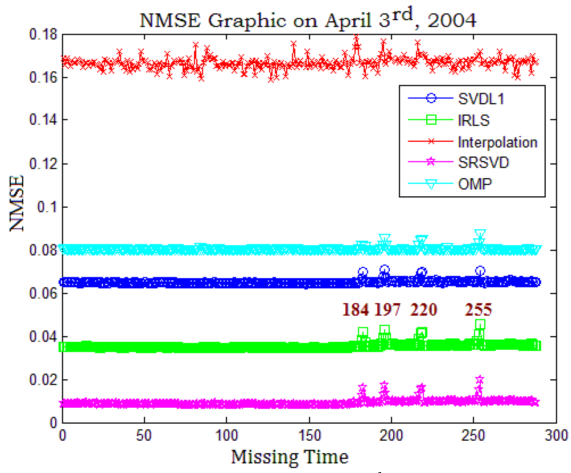
4.5 Detection of Time Sensitivity

Time sensitivity detection is done by eliminating one by one the column of TM sequentially. As discussed before, a column in TM represents a 5-minute traffic snapshot, hence one-day measurement will contain 288 columns of TM.

We investigate the effect of the missing time to the NMSE parameter for TM matrix that collected from April 1st, 2004 to April 30th, 2004. Figure 9 shows an example of time sensitivity detection for a day for each different reconstruction algorithms. The X-axis represents the missing of time, the Y-axis represents the NMSE value. Test results show that the missing of one time does not affect the value of NMSE, this is because the time series values of traffic matrix have a high correlation with each other. This applies to all the reconstruction algorithms.



(a). NMSE on April 2nd, 2004



(b). NMSE on April 3rd, 2004

Figure 9. Detection of time sensitivity for a day, (a). NMSE on April 2nd, 2004, (b). NMSE on April 3rd, 2004

The next simulation is to test the effect of missing block pattern of time. Missing block pattern indicates missing traffic during a certain period of time. We use the range of missing block probability from $0 < p_t \leq 0.9$. The missing block of time is chosen randomly.

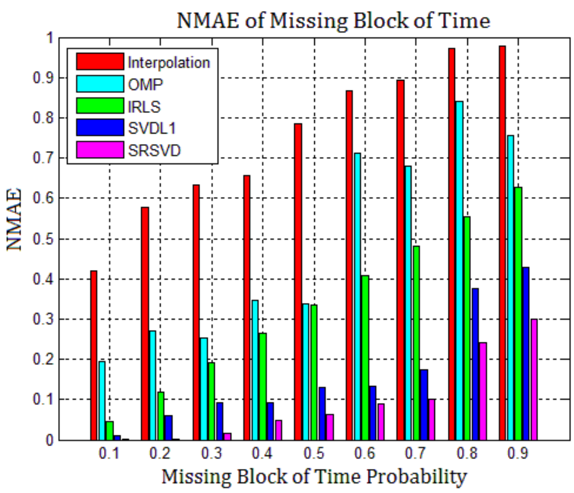


Figure 10. The missing block of time simulation

Figure 10 shows the missing block of time for a day using each different reconstruction algorithms. The X-axis represents the probability of missing block of time and the Y-axis represents the NMAE value. While the probability of missing block of time is more than 0.8, the NMAE increases significant, especially on SRSVD and SVDL1 algorithm. The probability of missing 0.8 means there are missing data during 1152 s or 19.2 hours. Due to a large block time of missing data, the available data is insufficient for an accurate estimation. The sample of data does not correlate each other.

5. Conclusions

CS can be used in internet traffic monitoring, such as, for reconstruction from missing traffic, detection of link sensitivity, and time sensitivity. The results show that SRSVD outperforms other CS reconstruction algorithms and Interpolation technique for estimating TM in various missing patterns. In addition, CS is also able to detect the sensitive link, thus it can increase the performance of the network, by choosing the best link. In the missing block of time, the CS reconstruction algorithms fail to estimate the correct TM if the missing probability is greater than 0.8 or the missing data occur more than 19.2 hours.

Our future study will be develop our approach to solve large amount missing value, especially in the case of Missing Element at Random (MER), Combine Missing Patterns (CMP), and Missing Block of Time (MBT). Because of random missing patterns in large quantities makes the TM element samples may have not any correlation with each other. The next projects are to conduct correlation approach by considering local data structures using interpolation technique and combining with global data structures through low-rank approach on TM

References

[1] M. Roughan, Y. Zhang, W. Willinger and L. Qiu, "Spatio-Temporal Compressive Sensing and Internet Traffic Matrices (Extended Version)," *IEEE/ACM Transaction on Networking*, vol. 20, no. 3, pp. 662-676, 2012.

[2] D. Guo, X. Qu and L. Huang, "Sparsity-Based Spatial Interpolation in Wireless Sensor Networks," *Sensors*, vol. 11, pp. 2386-2407, 2011.

[3] A. B. Suksmono, "Interpolation of PSF based on compressive sampling and its applicationin weak lensing survey," *Monthly Notices of the Royal Astronomical Society (MNRAS)* , vol. 443, pp. 919-926, 2014.

[4] D. L. Donoho, "Compressed Sensing," *IEEE Transaction Information Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

[5] E. Candes, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, pp. 1-20, 2006.

[6] E. Candes, J. Romberg and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," *Communications on pure and*

applied mathematics, vol. 59, no. 8, pp. 1207-1223, 2006.

[7] I. D. Irawati, A. B. Suksmono, "Image reconstruction based on compressive sampling using IRLS and OMP algorithm," *Jurnal Teknologi,* vol. 78, no. 5, pp. 309-314, 2016.

[8] K. Usman, A. B. Suksmono, H. Gunawan, "Uniform non-exhaustive search on sparse reconstruction for direction of arrival estimation," in *IEEE Proceedings of Asia Pacific Wireless and Mobile (APWiMob),* Bandung, Indonesia, pp.20-24, 2015.

[9] A. B. Suksmono, "Improved target detection in phase-coded radar by Constrained Matching Pursuit Algorithm," in *Proceedings of Instrumentation Communications Information Technology and Biomedical Engineering (ICICI-BME),* Bandung, Indonesia, pp. 337-380, 2013.

[10] A. B. Suksmono, "Improved Compressive Sampling SFCW Radar by Equipartition of Energy Sampling," *International Journal on Electrical Engineering and Informatics (IJEEI),* vol. 6, no. 3, pp. 553-561, 2014.

[11] M. Zeeshan and S. A. Khan, "Robust Sampling Clock Recovery Algorithm for Wideband Networking Waveform of SDR," *International Journal of Communication Networks and Information Security (IJCNIS)* , vol. 5, no. 1, pp. 10-18, 2013.

[12] E. Candes and J. Romberg, "l1-magic: Recovery of Sparse Signals via Convex Programming," *available: http://www. acm. caltech. edu/l1magic/downloads/l1magic. pdf 4, 14., 2005,* accessed at: october 21st, 2015.

[13] K. Mohan and M. Fazel, "Iterative Reweighted Algorithms for Matrix Rank Minimization," *Journal of Machine Learning Research* , vol. 13, pp. 3441-3473, 2012.

[14] J. A. Tropp and A. C. Gilbert, "Signal Recovery from Random Measurement via Orthogonal Matching Pursuit," *IEEE Transaction on Information Theory,* vol. 53, no. 12, pp. 4655-4666, 2007.

[15] E. Kreysig, Advanced Engineering Mathematics, New York: John Wiley and Sons Inc., 1972.

[16] Y. C. Chen, L. Qiu, Y. Zhang, G. Xue and Z. Hu, "Robust Network Compressive Sensing," in *Proceedings of the 20th annual international conference on Mobile computing and networking,* Hawaii, USA, pp. 545-556, 2014.

[17] Z. Huibin, Z. Dafang, X. Kun and W. Xiaoyan, "Data Reconstruction in Internet Traffic Matrix," *IEEE Journal and Magazine, Chinese Communication,*, vol. 11, pp. 1-12, 2014.

[18] L. Nie, D. Jiang and L. Guo, "A Compressive Sensing-Based Reconstruction Approach to End-to-End Network Traffic," in *IEEE Proceedings of Wireless Communications, Networking and Mobile Computing (WiCOM),* Shanghai, China, pp. 1-4, 2012.

[19] A. Lakhina, M. Crovella and C. Diot, "Diagnosing Network-wide Traffic Anomalies," *ACM SIGCOMM*, pp. 219-230, 2004.

[20] J. Chen, X. Qi, Y. Wang, "An efficient solution to locate sparsely congested links by network tomography," in *IEEE Proceedings of International Conference on Communications (ICC),* Sydney, Australia, pp. 1278-1283, 2014.

[21] V. W. Bandara, A. P. Jayasumana and R. Whitner, "An adaptive compressive sensing scheme for network tomography based fault localization," in *IEEE Proceedings of International Conference Communications (ICC),* Sydney, Australia, pp. 1290-1295, 2014.

[22] M. Mardani and G. B. Giannakis, "Robust Network Traffic Estimation via Sparsity and Low Rank," in *IEEE Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* Canada, pp. 4529-4533, 2013.

[23] D. Lavrova and A. Pechenkin, "Applying Correlation and Regression Analysis to Detect Security Incidents in the Internet of Things," *International Journal of Communication Networks and Information Security (IJCNIS)* , vol. 7, no. 3, pp. 131-137, 2015.

[24] Y. Vardi, "Network tomography: Estimating Source Destination Traffic Intensities From Link Data," *Journal of the American Statistical Association,* vol. 91, no. 433, pp. 365-377, 1996.

[25] S. Cai, M. Bakshi, S. Jaggi and M. Chen, "FRANTIC: A Fast Reference-based Algorithm for Network Tomography via Compressive Sensing," in *IEEE Proceedings of International Conference on Communication Systems and Networks (COMSNETS),* Bangalore, India, pp. 1-7, 2014.

[26] A. Lakhina, K. Papagiannaki, M. Crovella, C. Diot and E. D. Kolaczyk, "Structural Analysis of Network Traffic Flows," *ACM SIGMETRICS/Perform.,* pp. 61-72, 2004.

[27] D. Yang, "Singular Value Decomposition for High Dimensional Data," in *University of Pennsylvania,* 2012.

[28] E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Transaction on Information Theory,* vol. 52, no. 12, pp. 5406-5425, 2006.

[29] M. Elad and A. Bruckstein., "A Generalized Uncertainty Principle and Sparse Representation in Pairs of Bases," *IEEE Transaction on Information Theory,* vol. 48, pp. 2558-2567, 2002.

[30] W. Xu, E. Mallada and A. Tang, "Compressive Sensing over Graph," in *IEEE Proceedings of International Conference on Computer Communications (INFOCOM),* Shanghai, China, pp. 2087 - 2095, 2011.

[31] -, "The Abilene Research Netwok," *http://abilene.internet2.edu/.s.,* accessed at: January 13rd, 2015.