# Low-Rank Representation for Internet Traffic Reconstruction using Compressive Sampling

Indrarini Dyah Irawati<sup>1, 2</sup>, Ian Joseph Matheus Edward<sup>2</sup> and Andriyan Bayu Suksmono<sup>2</sup>

<sup>1</sup>Telkom Applied Science School, Telkom University, Bandung, Indonesia. <sup>2</sup>School of Electrical and Informatics, Institut Teknologi Bandung, Indonesia.

indrarini@telkomuniversity.ac.id

Abstract—We study compressive sampling for internet traffic reconstruction. Compressive Sampling (CS) requires that the traffic satisfies the low-rank feature. Low-rank states that traffic matrix can be represented in the right domain which the entire necessary information is concentrated in a low number of coefficients. In this paper, we compared three low-rank representation, which are Principal Component Analysis (PCA), Singular Value Decomposition (SVD), and Singular Value Decomposition Mean (SVDM). This low-rank representation is applied to four CS reconstruction algorithms, namely: Sparsity Regularized Singular Value Decomposition (SRSVD), Singular Value Decomposition L1 (SVDL1), Iteratively Reweighted Least Square (IRLS), Orthogonal Matching Pursuit (OMP), and Interpolation. The SVD outperforms the others low-rank representation techniques when used together with SRSVD, SVDL1, IRLS, and Interpolation. The SVDM gives the best NMAE when applied to the OMP. The computational times is linear with the number of the rank matrix. For all reconstruction algorithms, SVDM takes the least computational times.

*Index Terms*—Compressive Sampling; Low-Rank; Internet Traffic Matrix; SVD.

#### I. INTRODUCTION

Compressive sampling is a novel sampling paradigm, that can recovery certain signal from a few number of samples [1] [2]. CS must fulfill three requirements, which are the sparse representation of the signal, Restricted Isometric Property (RIP) on measurement matrix, and signal reconstruction algorithm [3]. Sparsity expresses that the signal contains many elements of zero. A non-sparse signal can be represented as a sparse signal when expressed on the right basis. In matrix form, we call sparsity as low-rank.

Generally, internet traffic matrix is not sparse nor in a spatial domain or temporal domain, but it has sparse representation if present in a proper transform domain. There is an opportunity to achieve low-rank matrix on the right domain so that CS performs compression with minimal sample quantities and high accuracy.

Some research explore the low-rank traffic matrix representation methods [4], [5], [6], [7], [8]. Lakhina, et al. [4] proposed Principal Component Analysis (PCA) to reduce high dimensional traffic matrix by finding the most variance of dataset that represents data structure. However, this method is less accurate because it does not consider the temporal correlation between time instants [5]. Another research [6] explored Singular Value Decomposition (SVD) to represent low dimension matrix using the largest singular values. In [7], the authors modified the eigen values of SVD decomposition to improve the accuracy of noise signal. In

literature [8], the authors observed space-time wavelet transform to minimize wavelet coefficients. This technique only requires 3% of wavelet coefficients for traffic matrix reconstruction.

Related to the second requirement, the compressed signal can be restored to the original signal perfectly if the measurement matrix satisfied RIP. This property states that compressed signal has the same euclidean norm with the original signal [9], [10].

Signal reconstruction algorithms consist of greedy algorithm and convex optimization. A greedy algorithm can approach the original signal by selecting the solution of a local optimal in each iteration. The example of a greedy algorithm is Orthogonal Matching Pursuit (OMP) that developed by Tropp [11]. Convex problems can be solved by linear programming using  $l_1$  magic software to obtain the optimum solution [12].

In this paper, we apply Singular Value Decomposition Mean (SVDM) to achieve low-rank representation of internet matrix traffic. SVDM is a modification of SVD by averaging the singular values on diagonal matrix [7]. We compare the SVDM with others low-rank methods such as PCA [4] and SVD [13]. The binary matrix is used as measurement matrix that expresses routing matrix. The routing matrix elements are '0' and '1' [14], [15]. The reconstruction algorithms that we use are Sparsity Regularized Singular Value Decomposition (SRSVD) [6], l<sub>1</sub>-norm optimization [12], Iteratively Reweighted Least Square (IRLS) [16], Orthogonal Matching Pursuit (OMP) [11]. These algorithms are compared with Interpolation [17].

This paper is arranged as follows: Section II describes compressive sampling theory for internet traffix atrix. Lowrank representation techniques is given in Section III. Section IV contains analysis of experimentals. Conclusions and future works is given in Section V.

## II. COMPRESSIVE SAMPLING FOR INTERNET TRAFFIC MATRIX

The application of compressive sampling technique requires two processes, namely acquisition and reconstruction. The aim of the acquisition process is to descent into a small number of samples so that the data size is more proper for storage or delivery. The reconstruction process, on the other hand, recovers the original signal.

CS application on the Internet TM can solve the ill-posed problem of:

$$Y = AX \tag{1}$$

where  $Y \in \mathbb{R}^{m \times t}$  denotes the traffic measurements on the links,  $A \in \mathbb{R}^{m \times n}$  expresses routing matrix, and  $X \in \mathbb{R}^{n \times t}$  describes a matrix of the traffics.

A traffic matrix (X) expresses the total of traffic that flows from a source to a destination (SD) node of the network. We consider that there are c nodes in the network, then the number of SD pairs is  $n = c^2$ . If there is time-series TM measurement (t = 1, 2, ... T), hence  $X \in \mathbb{R}^{n \times t}$ . In CS, TM must satisfy low-rank property so that TM must be expressed in the proper domain.

The routing matrix **A** acts as measurement matrix in CS that defined as follows:

$$A_{ij} = \begin{cases} 1, \text{ if the link } i \text{ is the part of the path for SD pair } j \\ 0, \text{ otherwise} \end{cases}$$
(2)

In CS, measurement matrix must fulfill RIP condition. The minimum number of its rows is given as [14]:

$$m \ge k \left( r \log \frac{n}{r} \right), \tag{3}$$

where:

k : Determined by the range from 1 to 2

*r* : Number of ranks on the matrix *X*.

In the reconstruction process, the TM reconstruction  $(\hat{X})$  should be as close as possible to the original TM (X). Reconstruction algorithms work to find an optimal solution using norm  $l_n$  norm, as follows:

$$\min \left\| \widehat{X} \right\|_{n}, \quad \text{subject to } AX = Y \tag{4}$$

where:

 $\|.\|_p: l_p \text{ norm with } 0 , that used to determine the difference between <math>X$  and  $\hat{X}$ 

#### III. LOW-RANK REPRESENTATION

This section presents our proposed system that shown in Figure 1 [17]. Traffic internet is obtained from the direct measurement. This traffic is represented on a temporal domain so as forming a matrix with size  $n \times t$ , n denotes the number of links on the networks and t denotes the number of measurements at one time. We explore three low-rank representation methods, such as SVD, SVDM, and PCA. The system performs compressive sampling using binary measurement matrix. The reconstruction algorithms such as SRSVD, SVDL1, IRLS, OMP are compared with Interpolation technique. The scaling process aims to find proportional amplitude level.

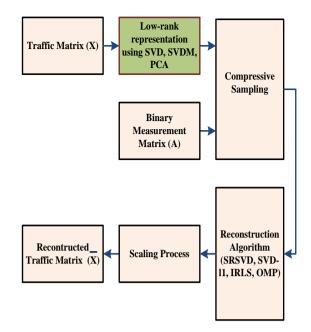


Figure 1: Block diagram of proposed system for low-rank representation on internet traffic reconstruction using CS [17]

#### A. Principal Component Analysis (PCA)

The principal component of  $X \in \mathbb{R}^{n \times t}$  is based on an eigenvector decomposition to obtain a low-rank traffic matrix approximation. PCA overcomes high dimensionality problem that described in Y = PX. This problem can be solved by finding some orthonormal matrix P, which  $S_Y = \frac{1}{t-1}YY^T = \frac{1}{t-1}PZP^T$ . The matrix  $Z = X^TX$  is a symmetric  $(n \times n)$  matrix. For the symmetric matrix Z, can be decomposed into [18]:

$$\boldsymbol{Z} = \boldsymbol{\mathsf{E}}\boldsymbol{\mathsf{D}}\boldsymbol{\mathsf{E}}^T \tag{5}$$

where:

**D** : Diagonal matrix of **Z** 

 $\mathbf{E}$ : Matrix composed by the eigenvector of  $\mathbf{Z}$  as columns.

The matrix Z has  $r \le n$  orthonormal eigenvectors, where r is the rank of matrix Z. If each row  $(p_i)$  of the matrix P is an eigenvector  $X^T X$  so that  $P = E^T$  and  $Z = P^T DP$ . The diagonal term of  $S_Y$  as the covariance matrix can be rewritten as:

$$S_{Y} = \frac{1}{n-1} P Z P^{T}$$
(6)  
$$= \frac{1}{n-1} P (P^{T} \mathbf{D} P) P^{T}$$
$$= \frac{1}{n-1} (P P^{-1}) \mathbf{D} (P P^{-1})$$
$$= \frac{1}{n-1}$$

Performing PCA is quite simple: (1) the dataset X is subtracted with the mean of row  $x_r$ , (2) The principal component of X are the eigenvectors of  $XX^T$  or the row of P. The PCA solution is given in Equation (7).

$$\widehat{X} = E^T (X_{m \times n} - \overline{X}_{m \times n})$$
(7)

#### B. Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) is a potential technique for representation the low-rank internet traffic matrix [6]. Consider an internet traffic matrix  $X \in \mathbb{R}^{n \times t}$ . The eigenvalue  $(\lambda_r)$  and eigenvector  $(\boldsymbol{v_r})$  of  $X^T X$  fulfills [19]:

$$(\boldsymbol{X}^T \boldsymbol{X}) \boldsymbol{v}_r = \lambda_r \boldsymbol{v}_r \tag{8}$$

A set of  $(v_1, v_2, ..., v_R)$  represents orthonormal  $(t \times 1)$  eigenvectors with associated eigenvalues  $(\lambda_1, \lambda_2, ..., \lambda_R)$  of the  $X^T X$ . The eigenvalues are ordered from largest to smallest  $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_R$ . Each principal component can be calculated by  $v_r$  dan  $u_r$ , as follows:

$$\boldsymbol{X}\boldsymbol{v}_{\boldsymbol{r}} = \sigma_{\boldsymbol{r}}\boldsymbol{u}_{\boldsymbol{r}} \tag{9}$$

The set of  $(u_1, u_2, ..., u_R)$  is an orthonormal  $(n \times 1)$  of X. The singular value  $\sigma_r = \sqrt{\lambda_r}$  is a positive real value that represents the quantity of energy captured by principal component r. The new diagonal matrix ( $\Sigma$ ) is arranged by the singular values with the result that  $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_R$ ,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\sigma}_{1} & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{\sigma}_{2} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{\sigma}_{R} & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{0} \end{bmatrix}$$
(10)

For the orthogonal matrices U and V, where  $U = [u_1, u_2, ..., u_n, V = [v_1, v_2, ..., v_t]$ ,  $U = U^{-1}$ , and  $V = V^{-1}$ . SVD decomposes X according to Equation (11).

$$XV = U\Sigma$$
(11)  

$$XVV^{T} = U\Sigma V^{T}$$
  

$$X = U\Sigma V^{T}$$

The maximum number of linearly independent rows or column in a matrix is equivalent to the number of non-zero singular values that termed as the rank of the matrix. The matrix X is low-rank if it satisfies  $R \ll \min(n, t)$ . For the low-rank matrix, it can be represented as:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{T} = \sum_{r=1}^{\min(n,t)} \sigma_{r} \, \boldsymbol{u}_{r} \boldsymbol{v}_{r}^{T}$$
(12)

where:

 $\begin{array}{ll} \boldsymbol{u_r} & r^{th} \text{ columns of } \boldsymbol{U} \\ \boldsymbol{v_r} & : r^{th} \text{ columns of } \boldsymbol{V} \end{array}$ 

The low-rank approximation of X is equivalent to [6]:

$$\widehat{\boldsymbol{X}} = \sum_{r=1}^{R} \sigma_r \, \boldsymbol{u}_r \boldsymbol{v}_r^T = \sum_{r=1}^{R} \sigma_r \, \boldsymbol{B}_r \tag{13}$$

where:

 $\widehat{X}$  : Top rank estimate of X

 $B_r$  : Matrix that composed by rank-1

### C. Singular Value Decomposition Mean (SVDM)

SVDM is a modified SVD technique. In SVD, the TM is

decomposed to produce  $U, \Sigma$ , and  $V^T$ . On the other hand, SVDM is obtained by modifying the singular value matrix ( $\Sigma$ ) from the Equation (12) [7]. The modified singular value is an average of  $\Sigma = diagonal(\sigma_1, \sigma_2, ..., \sigma_r)$ , which in  $\sigma_1 \ge \sigma_2 \ge$  $\dots \ge \sigma_R$ ,  $Mean = \overline{\sigma} = \frac{\sigma_1 + \sigma_2 + \dots + \sigma_R}{R}$  so that  $\overline{\Sigma} =$  $diagonal(\overline{\sigma_1}, \overline{\sigma_2}, ..., \overline{\sigma_R}) = diagonal(\overline{\sigma}, \overline{\sigma}, ..., \overline{\sigma})$ , which is  $\overline{\sigma_1} = \overline{\sigma_2} = \dots = \overline{\sigma_R} = \overline{\sigma}$ . The modified singular value is shown in Equation (14).

$$\overline{\Sigma} = \begin{bmatrix} \overline{\sigma_1} & 0 & 0 & 0 & 0 \\ 0 & \overline{\sigma_2} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \overline{\sigma_R} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$$
(14)

The modified form of the SVDM decomposition from Equation (12) and (13) can be expressed as follow:

$$\boldsymbol{X} = \boldsymbol{U}\overline{\boldsymbol{\Sigma}}\boldsymbol{V}^{T} = \sum_{r=1}^{\min(n,t)} \overline{\sigma_{r}} \, \boldsymbol{u}_{r} \boldsymbol{v}_{r}^{T}$$
(15)

$$\widehat{X} = \sum_{r=1}^{R} \overline{\sigma_r} \, \boldsymbol{u}_r \boldsymbol{v}_r^T = \sum_{r=1}^{R} \overline{\sigma_r} \, \boldsymbol{B}_r \tag{10}$$

#### IV. EXPERIMENTAL RESULTS

A. Abilene Data

We use the actual traffic data that is taken from the Abilene network. The Abilene consists of 12 nodes (c = 12) so that there are 144 pairs of source destination (n = 144) [20]. The Abilene data set contains link count measured at every 5 minutes interval. For a day, there are 288 values of five minutes aggregates. The traffic matrix is shown in Figure 2 for a period of 1<sup>st</sup> April 2004.

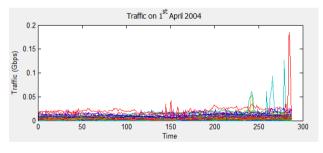


Figure 2: Traffic measured for a day on 1st April 2004

#### B. Performance Metric

The performance metrics that used to measure the accuracy of the TM reconstruction are Normalized Mean Absolute Error (NMAE) and computational times. These metrics are commonly used for CS in network application [6] [21]. NMAE is defined as follows:

$$NMAE(\boldsymbol{X}, \widehat{\boldsymbol{X}}) = \frac{\sum_{i,j}^{N} |\boldsymbol{X}(i, j) - \widehat{\boldsymbol{X}}(i, j)|}{\sum_{i,j}^{N} |\boldsymbol{X}(i, j)|},$$
(17)

where: X(i, j)

X(i, j): Original TM $\hat{X}(i, j)$ : Reconstruction TM

Computational time is the amount of time required by an algorithm to perform the computing processes. The time measurement is started since the input inserted to the output produced [22].

#### C. Comparing of Low-rank Type in Different Reconstruction Algorithms

We evaluate the effect of low-rank types on various reconstruction algorithms. These algorithms use the same regularization, which is k = 2 for RIP condition, input rank parameters are r = 30 and r = 100, and binary value as measurement matrix.

In Figure 3, The X-axis represents the CS method, the Yaxis represents the low-rank type, and the Z-axis represents the value of NMAE. SVD gives the best NMAE value when used in conjunction with reconstruction algorithms, namely SRSVD, SVDL1, and IRLS. Whereas SVDM shows a good performance when applied to OMP algorithm. PCA, on the other hand, gives the best results when combined with IRLS and SRSVD.

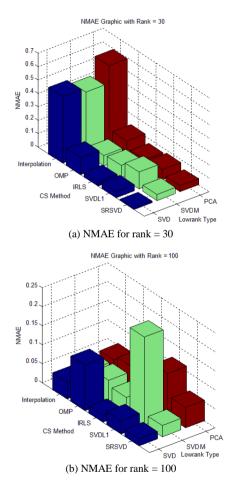
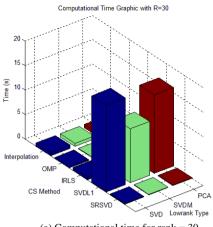


Figure 3: NMAE of Low-rank type in different reconstruction algorithms (a) for rank=30, (b) for rank=100.

SVD performs low-rank approximation with the number of rank-R. The SVD concept finds optimum R vectors that together reach a subspace in which the most of the data sample data located on until a small reconstruction error. The rank-R represents the number of singular values that illustrate the significant information of matrix.

PCA, on the other hand, reduces the dimension of a matrix by projecting the matrix to a different matrix such as diagonalized covariance matrix. The PCA procedure finds the deviation between the data sample and means data sample that depends on the proper distribution around the mean.

SVDM principle finds the average of singular value on the matrix and replaces the old singular value with the new ones. This case illustrates that significant information of matrix spread so that the use of minimum rank-R can not represent the actual matrix.



(a) Computational time for rank = 30

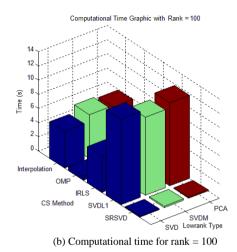
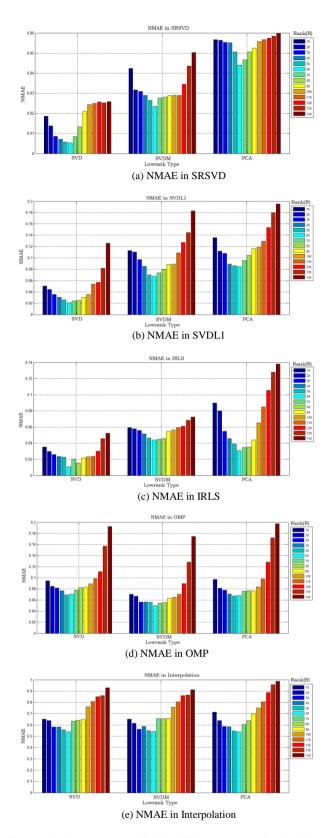


Figure 4: Computational time of Low-rank type in different reconstruction algorithms (a) for rank=30, (b) for rank=100.

In Figure 4, The X-axis represents the CS method, the Yaxis represents the low-rank type, and the Z-axis represents the value of computational time. Figure 3 showed that SVDM yields the lowest computational time when used together with all of the reconstruction algorithms. The SVDL1 algorithm takes the longest computational time than others.

#### D. Rank Parameter Sensitivity

This simulation aims to determine the effect of the number of rank on the performance of all algorithms. The observation is done with rank value (r = 10, 20, ..., 140). Figure 5 shows the performance results associated with a rank parameter. The X-axis represents the low-rank type and the Y-axis represents the NMAE value. For all low-rank types, the NMAE value increases with the increasing number of rank since r = 60. We find that the rank input (r = 60) gives the best performance. In CS, rank is an input parameter that is used to calculate the number of samples as shown in Equation (3). This is a requirement to satisfy RIP condition. It can be concluded that the more of samples of CS processing, the performance of algorithms will be an increase.



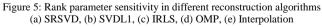
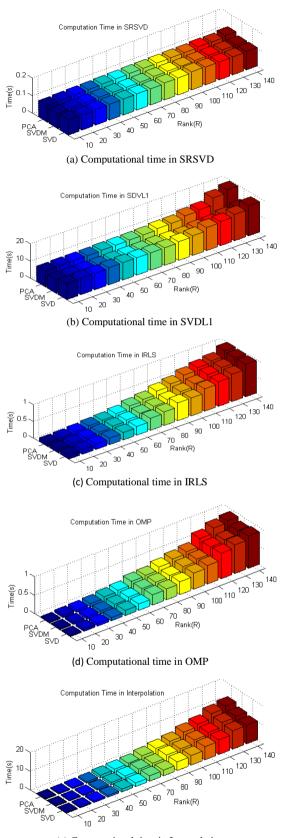


Figure 6 shows the influence of input parameter rank to the computational time. The X-axis represents the low-rank type, Y-axis represents rank value and the Z-axis represents computational time. We conclude that the greater number of rank, the computational time is getting longer. For all algorithms, SVDM has lowest computational time. In SRSVD algorithms, the low-rank representation types show that the computational process needs in a short time (0.08-

0.12 s). Whereas in SVDL1 algorithms, all types of low-rank representation have in a long computational time, which is 6-19s.



(e) Computational time in Interpolation

Figure 6: The effect of rank parameter to computational time in different reconstruction algorithms (a) SRSVD, (b) SVDL1, (c) IRLS, (d) OMP, (e) Interpolation

#### V. CONCLUSIONS

CS in network application provides the good TM reconstruction results and fast computational time especially if the TM represents the proper base so that the TM has lowrank property. The SVD low-rank type has good performance when it is applied together with appropriate reconstruction algorithms, ie: SRSVD, SVDL1, IRLS, and Interpolation. SVDM, on the other hand, can increase the performance when works with OMP algorithm. The number of ranks is very influential to the number of sample for the reconstruction process. The best performance is obtained when the input rank of r = 60. The computational time is increasing as increasing the number of ranks. The SVDM produces the fastest computational time compared to the others on the use of all reconstruction algorithms. The computational time in the SVDL1 algorithm is the worst for all low-rank types, which is 8-20 seconds.

Further advances optimizes the measurement matrix and modifies the reconstruction algorithms simultaneously based on right low-rank representation. We can expand future work by using static and dynamic routing protocol.

#### REFERENCES

- D. Donoho, "Compressed sensing," IEEE Transaction Information Theory, vol. 52, pp. 1289-1306, 2006.
- [2] E. Candes, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information.," IEEE Transaction Information Theory, vol. 52, pp. 489-506, 2006.
- [3] E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," IEEE Transaction Information Theory, vol. 52, no. 12, pp. 5406-5425, 2006.
- [4] A. Lakhina; K. Papagiannaki; M. Crovella; C. Diot; E. Kolaczyk; N. Taft, "Structural analysis of network traffic flows," in The Joint International Conference on Measurement and Modeling of Computer Systems, New York, 2004.
- [5] A. Soule; A. Lakhina; N. Taft; K. Papagiannaki; K. Salamatian; A. Nucci; M. Crovella; C. Diot, "Traffic matrices: balancing measurements, inference and modeling," in ACM SIGMETRICS International Conference on Measurement, Banff, AB, 2005.

- [6] M. Roughan, Y. Zhang, W. Willinger and L. Qiu, "Spatio-temporal compressive sensing and internet traffic matrices (extended version)," IEEE/ACM TRANSACTIONS ON NETWORKING, vol. 20, no. 3, pp. 662-676, 2012.
- [7] Yulou PENG; Yigang HE, "A reconstruction algorithm for compressed sensing noise signal," Journal of Computational Information Systems, vol. 8, no. 14, pp. 6025-6031, 2012.
- [8] J. Wang and M. Rabbat, "Wavelet-based traffic matrix modeling," in INFORMS Telecommunications Conference, Montreal, QC, 2010.
- [9] E. Candes, "The restricted isometry property and its implications for compressed sensing," Compte Rendus de l'Academie des Sciences, vol. 1, pp. 589-592, 2008.
- [10] R. Baraniuk, "Compressed sensing, lecture note," IEEE Signal Processing Magazine, pp. 118-121, 2007.
- [11] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurement via orthogonal matching pursuit," IEEE Transactions on Information Theory, vol. 53, no. 12, pp. 4655-4666, 2007.
  [12] E. Candes, J. Romberg and Caltech, "L1magic: recovery of sparse
- [12] E. Candes, J. Romberg and Caltech, "L1magic: recovery of sparse signals via convex programming," 2005, available: https://www.acm.edu/11magic/download/11magic.pdf4,14, accessed at October 21st 2015.
- [13] R. E. Bryant, "Graph-based algorithms for Boolean function manipulation," IEEE Trans. on Computers, vol. 35, pp. 677-691, 1986.
- [14] W. Xu, E. Mallada and A. Tang, "Compressive Sensing over Graph," in IEEE Proceedings of Infocom, 2011.
- [15] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sampling," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p. 3869–3872, 2008.
- [16] 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.
- [17] I. D. Irawati and I. J. M. E. Andriyan Bayu Suksmono, "Missing internet traffic reconstruction using compressive sampling," International Journal of Communication Networks and Information Security (IJCNIS), vol. 9, no. 1, pp. 57-66, 2017.
- [18] D. Lay, Linear algebra and it's applications, New York: Addison-Wesley, 2000.
- [19] D. Yang, "Singular value decomposition for high dimensional data," University of Pennsylvania, 2012.
- [20] "The Abilene Research Network," http://abilene.internet2.edu/.s, accessed at 21st October 2015
- [21] Zhou Huibin; Zhang Dafang; Xie Kun; Wang Xiaoyan, "Data reconstruction in internet traffic matrix," IEEE Journal and Magazine, China Communication, vol. 11, pp. 1-12, 2012.
- [22] M. Sipser, Introduction to the theory of computation, Cambridge, Massachusetts: Course Technology Inc., 2006.