ANOMALY DETECTION IN IP NETWORKS BASED ON RANDOMIZED SUBSPACE METHODS

Maboud F. Kaloorazi[†] and Rodrigo C. de Lamare[†,‡]

[†] Centre for Telecommunications Studies (CETUC)
Pontifical Catholic University of Rio de Janeiro, Brazil
[‡] Department of Electronics, University of York, United Kingdom
Emails: {kaloorazi, delamare}@cetuc.puc-rio.br

ABSTRACT

In this paper we propose novel randomized subspace methods to detect anomalies in Internet Protocol networks. Given a data matrix containing information about network traffic, the proposed approaches perform a normal-plus-anomalous matrix decomposition aided by the randomized sampling scheme and subsequently detect traffic anomalies in the anomalous subspace using a statistical test. Simulation results demonstrate improvement over the traditional principal component analysis-based subspace methods in terms of robustness to noise and detection rate.

Keywords— network anomaly detection, PCA subspace method, randomized subspace methods, orthonormal basis, the Q-statistic.

1. INTRODUCTION

Network anomalies typically refer to abnormal behavior in the network traffic such as traffic volume, bandwidth and protocol use, which indicate a potential threat. Traffic anomalies may arise due to various causes ranging from network attacks such as denials-of-service (DoS) and network scans, to atypical circumstances such as flash-crowds and failures, which can have serious destructive effects on the performance and security of Internet Protocol (IP) networks [1, 2].

The seminal paper by Lakhina et al. [3] first employed Principal Component Analysis (PCA) [4] to detect network-wide traffic anomalies. Given a matrix of link traffic data $\mathbf{Y}$, the approach performs a normal-plus-anomalous matrix decomposition, i.e., $\mathbf{Y} = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}}$, using (a specific number of) its principal components and seeks anomalies in the anomalous subspace $\tilde{\mathbf{Y}}$. The emergence of this approach inspired researchers to improve its performance and to evaluate its sensitivity for detecting anomalies [5, 6]. Ringberg et al. [5] point out that since PCA does not consider the temporal correlation of the data, the normal subspace is contaminated with anomalies. To address this issue, Brauckhoff et al. [6] propose to apply the Karhunen-Loeve (KL) expansion [7], which considers both the temporal and spatial correlations. Recently, inspired by the well-established compressed sensing (CS) theory [8, 9] and also by robust principal component analysis (RPCA) [10–12], several works have approached network-wide traffic anomaly detection using these methods i.e., by solving a constrained optimization problem [13, 14].

The PCA-based methods [3, 6, 15] focus on link traffic covariance matrix and accordingly compute its singular value decomposition (SVD), a computationally expensive factorization, to separate the subspaces. In this paper, we present two novel randomized subspace approaches to detect anomalies in network traffic. In contrast to the works in [3, 6, 15], the proposed approaches do not form the covariance matrix and consequently obviate the computation of the SVD for the subspace separation. We validate the proposed approaches using synthetically generated data. Simulation results demonstrate that the proposed techniques can successfully diagnose network-wide anomalies with more effectiveness than PCA and robust PCA (RPCA).

The remainder of this paper is organized as follows. In Section 2 we introduce the data model that represents IP traffic and formulate the problem we are interested in solving. We review the method of PCA for network anomaly detection in Section 3. In Section 4, we describe our proposed methods in detail. In Section 5, we present and discuss our simulation results and our conclusion remarks are given in Section 6.

2. DATA MODEL AND PROBLEM FORMULATION

In this section, we describe a data model that represents the traffic in an IP network using linear algebra and state the problem of interest. Based on the structure of a network and the flow of data obtained by network tomography [16], we can model the link traffic as a function of the origin-destination (OD) flow traffic and the network-specific routing. Specifically, the relationship between the link traffic

$\mathbf{Y} \in \mathbb{R}^{m\times t}$ and OD flow traffic $\mathbf{X} \in \mathbb{R}^{n\times t}$, for a network with m links and n OD flows may be written as

$$\mathbf{Y} = \mathbf{R}\mathbf{X}, \quad (1)$$

where t is the number of snapshots and $\mathbf{R} \in \mathbb{R}^{m\times n}$ is a routing matrix. The entries of $\mathbf{R}$, i.e., $\mathbf{R}_{i,j}$, are assigned a value equal to one ($\mathbf{R}_{i,j} = 1$) if the OD flow j traverses link i, and are assigned a value equal to zero otherwise.

The network traffic model that takes into account the traffic anomalies and the measurement noise over the links can be expressed by

$$\mathbf{Y} = \mathbf{R}(\mathbf{X} + \mathbf{A}) + \mathbf{V}, \quad (2)$$

where $\mathbf{R} \in \mathbb{R}^{m\times n}$ is a fixed routing matrix, $\mathbf{X} \in \mathbb{R}^{n\times t}$ is the clean traffic matrix, $\mathbf{A} \in \mathbb{R}^{n\times t}$ is the matrix with traffic anomalies and $\mathbf{V} \in \mathbb{R}^{m\times t}$ denotes the link measurement noise samples. The problem we are interested in this work is how detect anomalies by observing $\mathbf{Y}$.

3. PRINCIPAL COMPONENT ANALYSIS FOR NETWORK ANOMALY DETECTION

Given the link traffic $\mathbf{Y}$, in order to detect anomalies the work in [3] performs a normal-plus-anomalous matrix decomposition such that $\mathbf{Y} = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}}$, where $\hat{\mathbf{Y}}$ is the modeled traffic and $\tilde{\mathbf{Y}}$ is the residual traffic.

The modeled traffic represented by $\hat{\mathbf{Y}}$ is the projection of $\mathbf{Y}$ onto the normal subspace $\mathcal{S}$ and the residual traffic modeled by $\tilde{\mathbf{Y}}$ is the projection of $\mathbf{Y}$ onto the anomalous subspace $\tilde{\mathcal{S}}$, both, using a selected number of principal components of $\mathbf{Y}$. Specifically, the modeled traffic can be obtained by

$$\hat{\mathbf{Y}} = \mathbf{P}\mathbf{P}^T\mathbf{Y} = \hat{\mathbf{C}}\mathbf{Y}, \quad (3)$$

and the residual traffic can be obtained by

$$\tilde{\mathbf{Y}} = (\mathbf{I} - \mathbf{P}\mathbf{P}^T)\mathbf{Y} = \tilde{\mathbf{C}}\mathbf{Y}, \quad (4)$$

where $\mathbf{P} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_r]$ is formed by the first r singular vectors $\mathbf{W}$ of the covariance of the centered traffic data $\hat{\mathbf{\Sigma}} = \frac{1}{t-1}(\mathbf{Y} - \mu)(\mathbf{Y} - \mu)^T$, where μ contains the mean of $\mathbf{Y}$, and $\hat{\mathbf{\Sigma}} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^T$ is a singular value decomposition.

Typically, a traffic anomaly results in a large change to the residual traffic $\tilde{\mathbf{Y}}$ [3]. To detect abnormal changes in $\tilde{\mathbf{Y}}$, a statistic referred to as the Q-statistic [17] is applied by computing the squared prediction error (SPE) of the residual traffic:

$$\text{SPE} = \|\tilde{\mathbf{Y}}\|_2^2 = \|\tilde{\mathbf{C}}\mathbf{Y}\|_2^2. \quad (5)$$

The network traffic is considered to be normal if

$$\text{SPE} \le Q_\beta, \quad (6)$$

where Q_β is a threshold for the SPE defined as

$$Q_\beta = \theta_1\Big[\frac{c_\beta\sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\Big]^{\frac{1}{h_0}}, \quad (7)$$

where

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}, \quad (8)$$

and

$$\theta_i = \sum_{j=k+1}^{m} \lambda_j^i, \text{for } i = 1, 2, 3, \quad (9)$$

with λ_j denoting the j-th singular value of $\hat{\mathbf{\Sigma}}$ and c_β is the $1-\beta$ percentile in a standard normal distribution.

The singular vectors of $\hat{\mathbf{\Sigma}}$ (or principal components of $\mathbf{Y}$) maximize the variance of the projected data. Thus, for instance, the j-the singular value of $\hat{\mathbf{\Sigma}}$ (or the variance captured by the j-the PC) can be expressed as $\lambda_j = \mathbb{V}\text{ar}\{(\mathbf{w}_j{}^T\mathbf{Y})^T\}$ [4]. Note that, each column in $\mathbf{Y}$, $Y_i \in \mathbb{R}^m$.

4. RANDOMIZED SUBSPACE METHODS FOR ANOMALY DETECTION

This section describes our proposed approaches termed Randomized Basis for Anomaly Detection (RBAD) and Switched Subspace-Projected Basis for Anomaly Detection (SSPBAD). Similar to the works in [18] and [3], given the data traffic matrix $\mathbf{Y}$, RBAD and SSPBAD perform a normal-plus-anomalous matrix decomposition. However, instead of using the principal components of $\mathbf{Y}$, they employ a matrix with a set of orthonormal basis $\mathbf{Q} \in \mathbb{R}^{m\times m}$ whose range approximates the range of $\mathbf{Y}$. The orthonormal basis $\mathbf{Q}$ is constructed employing the randomized sampling scheme [19–21], and once constructed, as will be explained in the next subsections, $\mathbf{Y}$ is represented as a linear superposition of normal and anomalous components, i.e., $\mathbf{Y} = \hat{\mathbf{Y}} + \hat{\mathbf{Y}}$, as given by

$$\hat{\mathbf{Y}} = \mathbf{P}\mathbf{P}^T\mathbf{Y} = \hat{\mathbf{C}}\mathbf{Y}, \quad (10)$$

and

$$\tilde{\mathbf{Y}} = (\mathbf{I} - \mathbf{P}\mathbf{P}^T)\mathbf{Y} = \tilde{\mathbf{C}}\mathbf{Y}, \quad (11)$$

where the matrix $\mathbf{P} = [\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_r]$ contains the first r columns of $\mathbf{Q}$. In words, the modeled traffic $\hat{\mathbf{Y}}$ is the projection of $\mathbf{Y}$ onto the normal subspace $\mathcal{S}_Q$ spanned by $\mathbf{q}_1, ..., \mathbf{q}_r$, and the residual traffic $\tilde{\mathbf{Y}}$ is the projection of $\mathbf{Y}$ onto the anomalous subspace $\tilde{\mathcal{S}}_Q$, the subspace orthogonal to $\mathcal{S}_Q$, i.e., $\tilde{\mathcal{S}}_Q = \mathcal{S}_Q^\perp$.

In order to detect abnormal behavior in the anomalous component using the Q-statistic [17], the variances captured by the orthonormal basis must be known, as stated in Section 3. The variances are computed as follows

$$\mathbf{\Lambda_Q} = \mathbb{V}\text{ar}\{(\mathbf{Q}^T\mathbf{Y})^T\}. \quad (12)$$

The Q-statistic, then, is applied to diagnose traffic anomalies. As stated, in the proposed approaches the orthonormal bases obtained and the variances captured serve as surrogates to the basis of principal components and the singular values, respectively, used in the PCA-based approach [3, 18]. As a result, the estimation of the covariance matrix from the traffic data is not required to be formed and, accordingly, its SVD is not required to be computed.

4.1. Randomized Basis for Anomaly Detection (RBAD)

To separate normal and anomalous subspaces as described in (3) and (4), RBAD uses orthonormal basis whose range approximates the range of the traffic matrix $\mathbf{Y}$, instead of the singular vectors of $\hat{\mathbf{\Sigma}}$ used in [3, 18]. To construct the basis, the product $\mathbf{B} = \mathbf{Y\Phi}$ is first formed using a random matrix $\mathbf{\Phi} \in \mathbb{R}^{t\times m}$, e.g., drawn from the standard Gaussian distribution, and a QR factorization is then performed on $\mathbf{B}$, i.e., $\mathbf{QR} = \mathbf{B}$, [21]. To improve the approximation accuracy of the basis, the work in [20] multiplies $\mathbf{B}$ with $\mathbf{Y}$ and $\mathbf{Y}^T$ alternately. Once the basis is obtained, the subspaces are separated as described in (10) and (11). Next, the variances captured by $\mathbf{Q}$ are calculated as described in (12). To detect abnormal behavior in the anomalous component (11), the Q-statistic is applied afterwards. A pseudocode for RBAD is given in Table 1.

Table 1: Pseudocode for the proposed RBAD technique.

Input: traffic matrix $\mathbf{Y} \in \mathbb{R}^{m\times t}$, rank r, an exponent q (e.g., $q = 1$ or $q = 2$);

1: Generate a random matrix $\mathbf{\Phi}$;

2: Form $\mathbf{B} = (\mathbf{YY}^T)^q\mathbf{Y\Phi}$;

3: Perform a QR factorization on $\mathbf{B}$ to build an orthonormal basis: $\mathbf{B} = \mathbf{QR}$;

4: Separate the subspaces with rank r: $\mathbf{Y} = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}}$;

5: Compute the variances: $\mathbf{\Lambda_Q} = \mathbb{V}\text{ar}\{(\mathbf{Q}^T\mathbf{Y})^T\}$;

6: Apply the Q-statistic to $\tilde{\mathbf{Y}}$: if SPE $> Q_\beta \rightarrow$ anomalies;

7: **return** anomalies in $\mathbf{A}$

4.2. Switched Subspace-Projected Basis for Anomaly Detection (SSPBAD)

The proposed SSPBAD technique, similar to RBAD, also constructs basis with orthonormal columns whose range approximates the range of $\mathbf{Y}$ which based on projects the traffic data $\mathbf{Y}$ onto two subspaces orthogonal to each other ($\mathcal{S}_Q$ and $\tilde{\mathcal{S}}_Q$). To construct the basis, first, the product $\mathbf{T}_1 = \mathbf{Y}^T\mathbf{T}_2$ is formed using a random matrix $\mathbf{T}_2 \in \mathbb{R}^{m\times m}$. Next, $\mathbf{T}_2$ is updated by $\mathbf{T}_1$ such that $\mathbf{T}_2 = \mathbf{YT}_1$. A QR factorization is performed to construct the orthonormal basis for the range of $\mathbf{T}_2$ afterwards. This orthonormal basis, a surrogate to the basis of principal components used in [3, 18], is employed to separate normal and anomalous subspaces (10), (11). Subsequently, the variances captured by $\mathbf{Q}$ are computed as described in (12). To detect traffic anomalies in the anomalous component (11), the Q-statistic follows.

A similar approach to constructing the orthonormal basis as in SSPBAD was proposed in [23] to approximate a rank-r matrix, but they construct the basis for the range of $\mathbf{T}_1$. To increase robustness of the algorithm for detecting anomalies, we employ different random matrices $\mathbf{T}_2$ as in [24], [25]. The random matrices generated include

- a matrix with i.i.d Gaussian entries i.e., $\mathcal{N}(0, 1)$,
- a matrix whose entries are i.i.d. random variables drawn from a Bernoulli distribution with probability 0.5,
- a Markov matrix whose entries are all nonnegative and the entries of each column add up to 1,
- a matrix whose entries are independently drawn from $\{$-1, 1$\}$.

Therefore, SSPBAD switches among different random matrices and chooses the best one in order to obtain the maximum number of traffic anomalies. A pseudocode for SSPBAD is given in Table 2.

Table 2: Pseudocode for the proposed SSPBAD technique.

Input: traffic matrix $\mathbf{Y} \in \mathbb{R}^{m\times t}$, rank r;

1: Generate N random matrices $\mathbf{T}_2$;

2: **for** $i = 1$: N **do**

3: Form $\mathbf{T}_1$: $\mathbf{T}_1 = \mathbf{Y}^T\mathbf{T}_2$;

4: Update $\mathbf{T}_2$: $\mathbf{T}_2 = \mathbf{YT}_1$;

5: Perform a QR factorization on $\mathbf{T}_2$ to build an orthonormal basis: $\mathbf{T}_2 = \mathbf{QR}$;

6: Separate the subspaces with rank r: $\mathbf{Y} = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}}$;

7: Compute the variances: $\mathbf{\Lambda_Q} = \mathbb{V}\text{ar}\{(\mathbf{Q}^T\mathbf{Y})^T\}$;

8: Apply the Q-statistic to $\tilde{\mathbf{Y}}$: if SPE $> Q_\beta \rightarrow$ anomalies;

9: **end for**

10: Choose the best random matrix with maximum number of anomalies;

11: **return** anomalies in $\mathbf{A}$

5. SIMULATION RESULTS

To validate the proposed approaches, we conduct experiments on synthetically generated data and compare the results with those of PCA and RPCA. The data matrix $\mathbf{Y}$ is generated according to the model in (2) with dimensions $m = 120, n = 240, t = 640$. The low-rank matrix $\mathbf{X}$ is formed by a matrix multiplication $\mathbf{U}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{n \times r}$ and $\mathbf{V} \in \mathbb{R}^{t \times r}$ have Gaussian distributed entries $\mathcal{N}(0, 1/n)$ and $\mathcal{N}(0, 1/t)$, respectively and $r = 0.2 \times m$. The routing matrix $\mathbf{R}$ is generated by entries drawn from a Bernoulli distribution with probability 0.05. The sparse matrix of anomalies has $s = 0.001 \times mt$ non-zero entries drawn randomly from the set $\{-1, 1\}$, and the noise matrix $\mathbf{V}$ has independent and identically distributed (i.i.d) Gaussian entries with zero mean and variance σ^2. We set the confidence limit $1 - \beta = 99.5\%$ for the value of the Q-statistic for all three approaches.

In Fig. 1, we compare the variances captured by the orthonormal bases of the proposed approaches with those of the principal components since they play a crucial role in the statistical test (the Q-statistic) used to detect anomalies. As can be seen, returned variances by RBAD and SSPBAD are very close to those returned by the SVD.

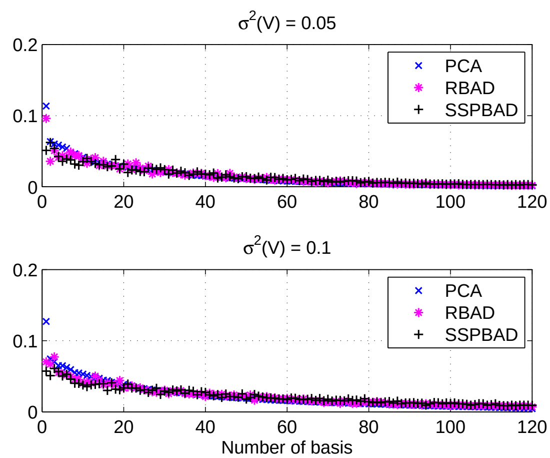

Figure 1: A comparison of variances for PCA, RBAD, SSPBAD.

Fig. 2 compares the detection rate against the number of basis for different approaches. The detection rate combines false-alarm rate and detection probability into one measure and obviates the need for showing these two probabilities in one versus the other manner [2]. As can be seen, the proposed RBAD and SSPBAD approaches outperform PCA when the measurement noise has a higher variance. Furthermore, RPCA [10–12] performs poorly. The reason is that, since we consider measurement noises $\mathbf{V}$ in our data model (2), by increasing the rank, these noise samples contaminate the matrix of outliers returned by RPCA and as a result, the abnormal patterns of the network (anomalies) cannot be recovered.

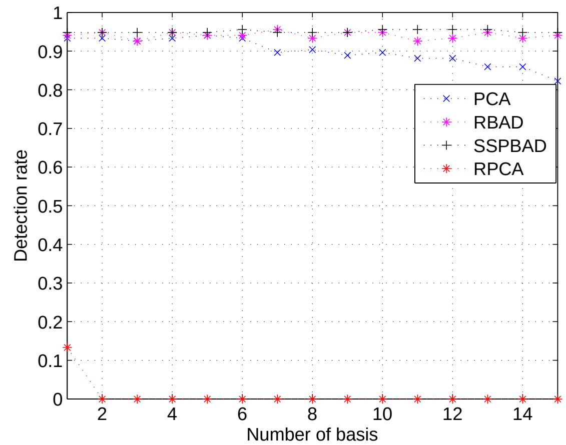

Figure 2: A comparison of detection rate for PCA, RBAD, SSPBAD and RPCA. Variance of the measurement noises $\sigma^2 = 0.1$

5.1. Computational Complexity

The PCA-based methods operate on the link traffic covariance matrix $\hat{\mathbf{\Sigma}}$ to separate the subspaces [3, 6, 15]. In particular, PCA employs the SVD which requires $O(tm^2)$ floating-point operations (flops) [26]. RBAD and SSPBAD operate on the link traffic data directly, but employ the QR factorization, which requires $O(tm^2)$ flops as well. Although, the computational complexity of RBAD and SSPBAD is roughly the same as PCA in the context of anomaly detection, in certain applications where the SVD cannot be efficiently used, an extension of the proposed approaches can be employed. For instance, they can be used to build a direct solver for contour integral equations with nonoscillatory kernels where the computational cost for a QR factorization is considerably less prohibitive than that of the SVD [27].

6. CONCLUSION

In this paper, we have proposed the RBAD and SSPBAD random subspace methods to detect traffic anomalies in IP networks. Both approaches form normal and anomalous randomized subspaces by orthonormal bases constructed for the range of the traffic data. A statistical test is then applied and detects anomalies in the network traffic measurements. Simulations show that RBAD and SSPBAD outperform PCA and RPCA. Future work will concentrate on mathematical analysis of RBAD and SSPBAD.

7. REFERENCES

[1] M. Thottan and C. Ji, "Anomaly detection in IP networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2191 – 2204, aug 2003.

[2] Y. Zhang, Z. Ge, A. Greenberg, and M. Roughan, "Network Anomography," in *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement (IMC 05)*, oct 2005.

[3] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing Network-Wide Traffic Anomalies," in *proceedings of ACM SIGCOMM*, aug 2004.

[4] I. T. Jolliffe, "Principal Component Analysis," 2nd ed, Springer, 2002.

[5] H. Ringberg, A. Soule, J. Rexford, and C. Diot, "Sensitivity of PCA for traffic anomaly detection," in *Proceedings of the 2007 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, jun 2007, pp. 109–120.

[6] D. Brauckhoff, K. Salamatian, and M. Martin, "Applying PCA for Traffic Anomaly Detection: Problems and Solutions," in *Proceedings of INFOCOM 2009*, apr 2009, pp. 2866 – 2870.

[7] R. M. Gray and L. D. Davisson, "An Introduction to Statistical Signal Processing," Cambridge University Press, 2005.

[8] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289 – 1306, apr 2006.

[9] E. J. Candès, J. Romberg, and T. Tao, "Exact Signal Reconstruction From Highly Incomplete Frequency Information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489 – 509, feb 2006.

[10] V. Chandrasekaran, S. Sanghavi, P. a. Parrilo, and A. S. Willsky, "Rank-Sparsity Incoherence for Matrix Decomposition," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2009.

[11] J. Wright, Y. Peng, Y. Ma, A. Ganesh, and S. Rao, "Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices," *Advances in Neural Information Processing Systems (NIPS)*, pp. 2080–2088, 2009.

[12] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, pp. 1–37, may 2011.

[13] 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, jun 2012.

[14] M. Mardani and G. B. Giannakis, "Estimating Traffic and Anomaly Maps via Network Tomography," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1533 – 1547, jun 2016.

[15] L. Huang, X. Nguyen, M. Garofalakis, M. Jordan, A. D. Joseph, and N. Taft, "In-network pca and anomaly detection," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2007-10, Jan 2007. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-10.html

[16] Y. Vardi, "Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data," *Journal of the American Statistical Association*, no. 433, pp. 365–377, feb 1996.

[17] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, aug 1979.

[18] R. Dunia and S. J. Qin, "A Subspace Approach to Multidimensional Fault Identification and Reconstruction," *American Institute of Chemical Engineers (AIChE) Journal*, vol. 44, no. 8, pp. 1813–1831, aug 1998.

[19] J. Demmel, I. Dumitriu, and O. Holtz, "Fast linear algebra is stable," *Numerische Mathematik*, vol. 108, no. 1, pp. 59–91, Nov 2007.

[20] V. Rokhlin, A. Szlam, and M. Tygert, "A randomized algorithm for principal component analysis," *SIAM Journal on Matrix Analysis and Applications (SIMAX)*, vol. 31, no. 3, pp. 1100 – 1124, 2009.

[21] N. Halko, P.-G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Rev*, vol. 53, no. 2, pp. 217–288, Jun 2011.

[22] J. E. Jackson, "A User's Guide to Principal Components," New York: Wiley, 1991.

[23] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," in *ICML*, 2011.

[24] M. F. Kaloorazi and R. C. de Lamare, "Switched-Randomized Robust PCA for Background and Foreground Separation in Video Surveillance," in *IEEE SAM 2016*, jul 2016.

[25] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2503–2514, jul 2009.

[26] L. N. Trefethen and D. Bau III, "Numerical Linear Algebra," SIAM, 1997.

[27] H. Cheng, Z. Gimbutas, P. G. Martinsson, and V. Rokhlin, "On the compression of low rank matrices," *SIAM Journal on Scientific Computing (SISC)*, vol. 26, no. 4, p. 1389 1404, 2005.