ISSN 2229-5518

Parameter Ranking and Reduction in Communication Systems

M.H. Azmol, M.H. Biswas, and A. Munnujahan

Abstract—Parameter reduction from experimental data is an important issue arising in many frequently encountered problems with different types of applications in communications engineering. However, the computational effort grows drastically with the number of parameters in such types of applications. This paper proposes a technique that reduces the performance parameters of a communication system based on eigenvalues of covariance matrix as well as providing a weighted rank of parameters by an approach called non-negative matrix factorization (NMF). The factorization of original matrix provides a weight metric that offers a means of ranking and selecting meaningful important parameters. The relative importance of each parameter is measured from the sequentially ordered eigenvalues. The main aims of this paper are to determine, identify and reduce the reasonable number of performance parameters which will reflect the best measurement system of a describing network scenario.

Index Terms—Eigenvalues, Parameter reduction, Non-negative matrix factorization

1 Introduction

IN the modern era, recent technological developments of telecommunication devices, computer based instrumentation and electronic data storage have resulted in increasing quantities of data and in a consequence, overwhelm many of the classical data analysis tools available. Processing these huge amounts of data has created many challenges and new concerns of data representation with a few numbers of dimensions. Many researchers have been exploring in these fields and a large portion of these efforts are concerned with the challenging task of evaluating the performances of a network by measuring some quality of service (QoS) support parameters such as bandwidth, load, delay, jitter, error rate etc. However, measuring all of these variables together is very challenging due to the growing complexity of telecommunication systems as well as some constraints on information gathering devices. The collected data from these complex phenomena represent the integrated result of several interrelated variables since they act together. The original data becomes noisy, overlap and ambiguous. In most phenomena, some parameters may have little effect on model predictions, and the effects of some parameters can be highly correlated with the effects of others, making parameter estimation difficult. As a result, modelers often use simplified models that contain a reduced number of parameters or they estimate only a subset of the parameters in their complete model. Moreover, being ignorant experimenters we don't know which, let alone how many, axes and dimensions are important to measure. Even we often don't know which measurements best reflect the dynamics of our system. Generally, in many applications, the data set is gathered from a huge number of dimensions (variables). This *curse of the dimensionality* (variables) degrades the query efficiency and accuracy of a system. So, the question is raised, how can we reduce these dependent parameters as well as provide a compact system model which will provide fidelity near the significant level of QoS of the original system without much loss of information's. A suitable parameter reduction technique may address all of these issues. The problem of parameter reduction has received a broad attention in different areas of research such as computer vision [2], information retrieval [3], machine learning, pattern recognition etc. Moreover, it has also widespread application in other areas, ranging from ecological systems, power systems, production systems, chemical reactions and biochemical networks to wastewater treatment processes. There exist many parameter reduction techniques such as Principal Component Analysis [6], [7], Singular Value Decomposition [8], Feature Selection [4], Non-negative Matrix Factorization (NMF) [5], and Factor Analysis [9]. The main disadvantage of feature selection and factor analysis approaches are that some combination of parameters may give better results, but could be excluded because of parameters selected at earlier steps. In, principal components, there are some limitations on non-linear cases. We use NMF algorithms as it is simple-nonparametric method which reveals easily underlying structures in complex data set. By using, NMF algorithms of the original data matrix, we derive a weight metric to obtain their ranking. The relative importance of each parameter is obtained from the sequentially ordered eigenvalues of covariance matrix, which in turn expresses the attributive energy of the data. Eigenvalues quantifies the importance of each dimension by describing the variability of a data set. In particular, the measurement of variance along each dimension provides a means for comparing the relative importance of each variable. An implicit hope behind employing this method is that a small number of dimensions (i.e. less than the number of measurement types) will provide a reasonable characterization of the complete data set. In fact, categorizing eigenvalues in this manner,

ISSN 2229-5518

a significant insight of the whole network can be obtained. For an illustration, in our network data matrix, we have explored six parameters of a WLAN network scenario and compare them. In general, we know that “throughput” and “network load” are the most important parameters for assessing the performance of a network. In our method, we have proved that these two parameters among six exhibit 94% characteristic of the whole network. In summary, we have proposed two measurement metrics in this paper: (i) a weighted metric for selecting the parameters that reveals important characteristics of a network (ii) and stipulated eigenvalues, measure the relative importance of each parameter. The primary objectives of our work are to determine, identify and reduce significant number of performance parameters in a systematic and representative way, so that a describing network scenario will characterize and reflect the best measurement system. The rest of this paper is organized as follows. The notations used throughout the rest of this paper are given in section 2. Related works are described in section 3. We outline the steps taken to collect the data in section 4. A brief overview of mathematical measurement technique of non-negative matrix factorization is discussed in section 5. In section 6, we discussed our experimental results. Concluding remarks are presented in sections 7.

2 BACKGORUND NOTATIONS

In order to facilitate discussion in subsequent sections we introduce relevant notation first. We use X is a m-by-n measurement matrix in which each column denotes the predicted variables and each row represents the sample observations. $\Re$ is a set of non-negative numbers, P and Q are the factorization of X of size m-by-k and k-by-n respectively. Y is the new transformation of X , $\|.\|$,is the Euclidean norm, λ_i and v_i are the corresponding eigenvalues and eigenvectors of covariance matrix X respectively. In our experiment data, we have taken six variables and 300 observations. So the size of X is *300 by 6*. Unless otherwise stated, all the vectors in this paper are column vectors.

3 RELATED WORKS

Our ideas behind parameter reduction in communication systems are motivated by a variety of past studies. Although to our knowledge, parameter reduction in a network using NMF and eigenvalues has not been previously applied directly. Author’s in paper [11] and [12] present a framework for nonlinear systems analysis based upon controllability and observability covariance matrices. This covariance matrix is used for reduction of the nonlinear model. Sanjay L. et al. [13] introduced a new method of model reduction for nonlinear system with inputs and outputs. A new technique based on Hankel singular values for reducing the parameter set of a fundamental model is introduced in paper [14]. The main drawback of this method is that the physical meaning of the parameters is lost during this procedure. There is a little prior work that is closely related to ours. Authors in paper [16] measured the performance of wireless networks by state space reduction. They propose the stochastic method based on the comparison of Markov chains by following three steps: first, choice of the state space of comparison, second, choice of the relation order on the chosen space and finally, construction of bounding models by applying aggregation. Paper [4] describes feature selection algorithm based on measuring similarity between features. A. Lakhina et al. in [17] analyzed the network flow and found that, in a network with over a hundred of origin to destination flows, these flows can be accurately modeled in time using a small number (10 or less) of independent components or dimensions. In [18] and [19] authors proposed different functional test areas to verify next generation broadband network. They classified the equipments into different groups by some scores such as MUST, SHOULD, MAY and CAN. This method can be used in order to compare any device of different vendors with each other but not their own devices, since all tests don’t have the same value. Hence this method will not be as reliable as needed. Most of the existing reduction techniques based on model reduction and focus on to reduce the number of states, whereas relatively few techniques are available for reduction of the number of parameters in these models.

4 NETWORK ENVIRONMENT AND EXPERIMENTAL DATA

For network simulations and raw data of different parameters, we use OPNET modeler which is one of the most leading and powerful simulation tools for the analysis, planning, optimization and evaluating of communication networks, devices and protocols. The measurement platforms and environment in which we conduct our experiments (Fig-1) are described in this section. We use two WLAN fixed nodes separated by 80 meter distance in our test bed. We perform experiments with high resolution video conference as an application in IEEE 802.11b WLAN. The attributes of different performance parame-

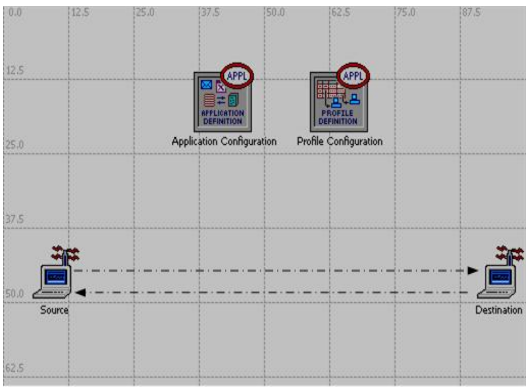

Fig. 1. WLAN network scenario.

ISSN 2229-5518

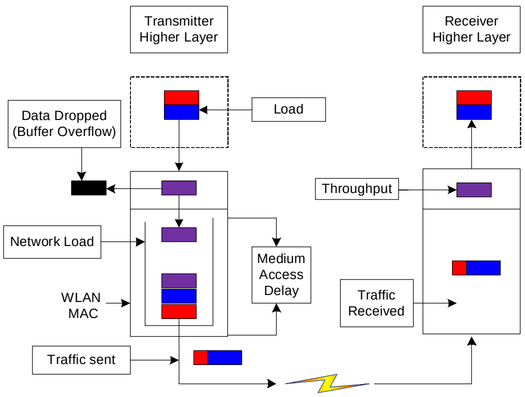

Fig. 2. Attributes of different parameters.

TABLE 1
SIMULATION PARAMETERS OF WLAN

Wireless LAN parameters	**Values**
Physical Characteristics	DSSS
Data Rate	11 Mbps
Channel Settings	Auto Assigned
Transmit Power	0.001w
Packet Reception-Power Threshold	-95
RTS Threshold	None
Fragmentation Threshold	None
CTS-to-Self Option	Enabled
Short Retry Limit	7
Long Retry Limit	4
Buffer Size	256000
Large Packet Processing	Drop
PCF Parameters	Disabled
HCF Parameters	Not Supported

TABLE 2
MEASUREMENT UNITS OF PARAMETERS

Collected parameters	**Units**
Data Dropped (Buffer Overflow)	bits/sec
Delay	sec
Load	bits/sec
Media Access Delay	sec
Network Load	bits/sec
Throughput	Bits/sec

ters are presented in Fig-2. We collect the global statistics of six parameters: data dropped (data overflow), delay, load, media access delay, network load and throughput. The configuration of WLAN parameters and measurement unit of each parameter are shown in Table-I and Table-II.

5 MATHEMATICAL APPROACH

In this section, we discuss our mathematical approach. Suppose X is our original data set, where each row corresponds to a set of measurements from one particular trial (a single sample) and each column of X corresponds to all measurements of a particular type. In our data set X is an *m-by-n* matrix where *m = 300* and *n = 6*. Our first approach is to find some orthonormal matrix v such that $Xv = Y$, and $Y^T Y$ is a diagonal matrix. Here generally, v is the eigenvector of the covariance matrix $X^T X$ and Y is a new transformed presentation of X, derived from $Xv = Y$. Generally, the primary motivation behind our first approach is to de-correlate the data set, i.e. to remove second-order dependencies. Y is arranged with the stipulated eigenvalues $\lambda_1 > ... > \lambda_r$. Our second approach is to factorize the original data X so that we can find a weight matrix to make some rank of the variables in X. Here, we discuss the factorization technique.

5.1 Non-negative Matrix Factorization

In an article in *Nature*, Lee and Seung, 1999, [20] started a flurry of research and proposed the notion of non-negative matrix factorization algorithm. Prior to its publication several lesser known paper published and actually they deserve more credit for the factorization of non-negative matrix. Given a data matrix $X = [x_1, x_2, ..., x_n] \in \Re^{m \times n}$, each column of X is a sample vector and a positive integer $k < \{m, n\}$, non-negative matrix factorization aims to find two non-negative matrices $P = [p_{ij}] \in \Re^{m \times k}$ and $Q = [q_{ij}] \in \Re^{k \times n}$ which minimize the following function

$$f(P,Q) = \frac{1}{2} \|X - PQ\|_F^2 \quad (1)$$

where $\|.\|_F$ denotes the Frobenius norm. The Frobenius norm of a matrix $X = [x_{ij}]$ is defined by $\|X\|_F = \sqrt{\sum_{ij} X_{ij}^2}$.

The product PQ is called the non-negative matrix factorization of X, though X is not necessarily equal to the product of PQ. The product PQ is an approximate factorization of X of rank at most k. Usually, an appropriate decision on the value of k is critical, but the choice of k is often problem dependent. The objective function (1) is convex in both P and Q only, it is not convex in both

ISSN 2229-5518

variables together. Therefore, to expect an efficient algorithm for finding global minima of $f(P,Q)$ is unrealistic. This important problem affects the numerical minimization of (1) since it includes the local minima due to the non convexity. Various minimization methods for the solution of (1) have been proposed in an effect to speed up the convergence. Lee and Seung's iterative multiplicative update algorithms have become a balance algorithm as follows:

$$P_{ij}^{t+1} \leftarrow P_{ij}^{t} \frac{\left(XQ^T\right)_{ij}}{\left(PQQ^T\right)_{ij}}$$

$$Q_{ij}^{t+1} \leftarrow Q_{ij}^{t} \frac{\left(X^T P\right)_{ij}}{\left(Q^T P^T P\right)_{ij}}$$

Multiplicative Update Algorithms:

Lee and Seung [20] proposed the prototype of multiplicative algorithm with the mean squared error objective function which is provided below:

P = rand (m, k) % initialize P as random dense matrix

Q = rand (k, n) % initialize Q as random dense matrix

for i = 1: maxiter

$$Q = Q.*\left(P^T X\right)./\left(P^T PQ + 10^{-9}\right);$$
$$P = P.*\left(XQ^T\right)./\left(PQQ^T + 10^{-9}\right);$$

end

To avoid division by zero, 10^{-9} is added in each update rule.

5.2 Basic Geometric Review of NMF

NMF essentially tries to find two non negative matrices P and Q such that $X \approx PQ$. This approximation can be viewed column by column as $x_i \approx \sum_{i=1}^{k} p_i q^i$ where p_i denotes *i-th* column vector of P and q^i denotes *i-th* row of matrix Q. Thus each data vector x_i is approximated by a linear combination of the columns of P weighted by the components of Q. Therefore, P can be considered as basis matrix which optimized the linear transformation of X. The outer product of PQ demonstrates how the rows of Q essentially specify the weights of each of the column-vector of P. If we look geometrically, NMF is a conical coordinate transformation. Graphical interpretation of NMF is shown in Figure-3. The two basis vectors P_1 and

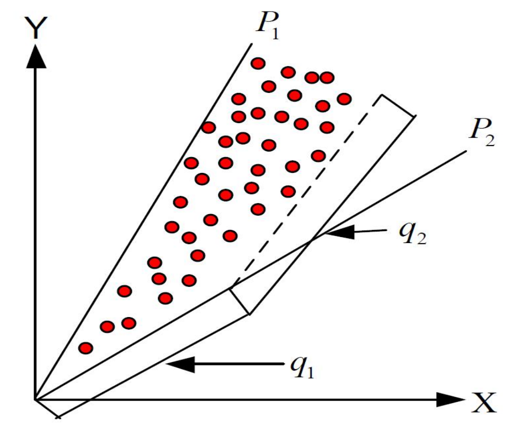

Fig. 3. Geometric behavior of NMF.

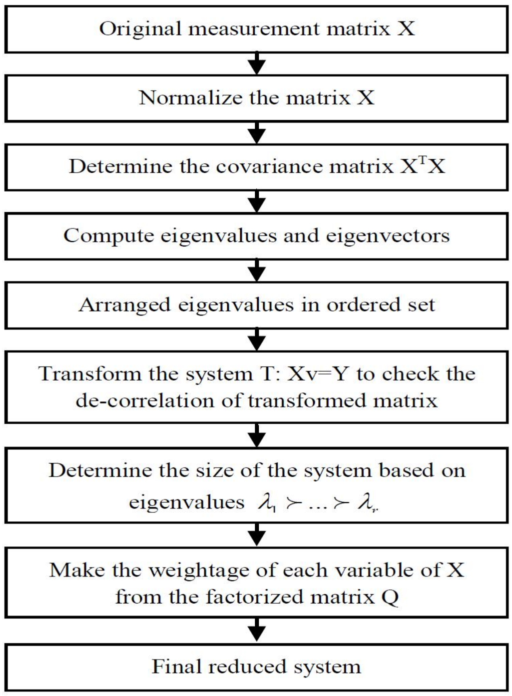

Fig. 4. Flowchart of reduction procedure.

TABLE 3
WEIGHT MATRIX Q

0.0552	0.1336	0.8836	0.1362	0.2894	0.3099
0.0734	0.1522	0.8868	0.1546	0.2737	0.2938

TABLE 4
EIGENVALUES AND THEIR PERCENTAGE

Eigenvalues	*Percentage*	*Cumulative sum*
4.6262	75.0049	0.7500
1.0827	18.6317	0.9364
0.2366	04.2039	0.9784
0.0528	01.4876	0.9933
0.0017	00.6498	0.9998
0.0002	00.0221	1.0000

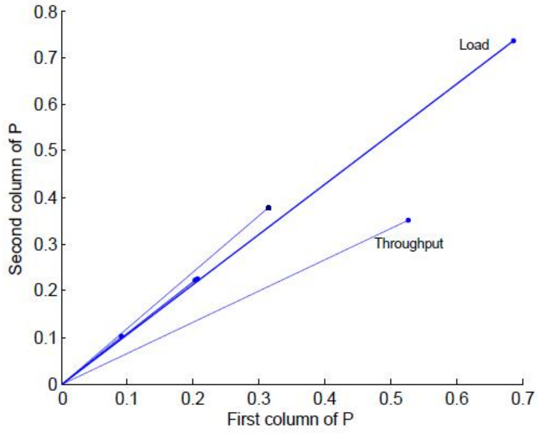

Fig. 5. Relative weight of variables in X w.r.t. to varianles in P.

P_2 describe a cone which encloses the approximation data set of X . Due to the non-negative constraints the points inside this cone can be reconstructed through linear combination of these basis vectors: $x = (p_1, p_2).(q_1.q_2)^T$

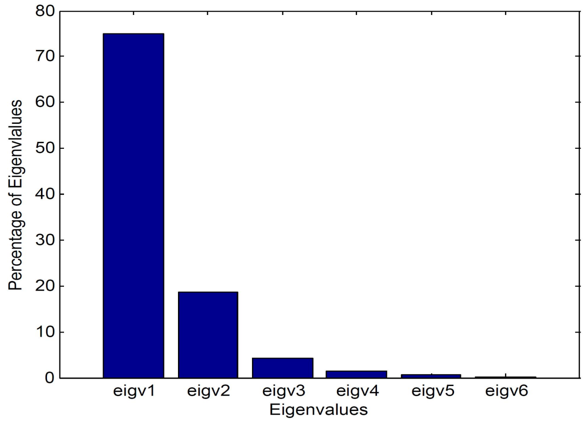

Fig. 6. Plot of eigenvalues versus their percentage.

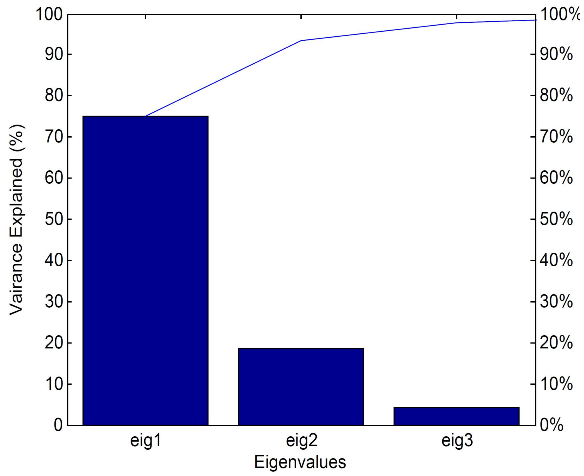

Fig. 7. Scree plot of the percent variability explained by eigenvalues.

6 Experimental Results

Experimental results are discussed in this section. The flowchart of our experiment is shown in Fig 4. For our experiment, we use a multidimensional data of six variables collected from a WLAN network scenario. The sample size of our data is *300 by 6*. When we analyze our data set, an important and unmet challenge arises against different units or scales. In real world data set, variables are measured in different units and these discrepancies can distort the calculations also the analysis may lead inaccurate. We mitigate this problem by converting and normalizing all the values of data set by subtracting each column from the minimum value for moving towards center of the data and dividing by corresponding standard deviation of each variable for scaling. We factorize our original matrix $[X]_{300\times6}$ to $[P]_{300\times2}$ and $[Q]_{2\times6}$ so that $X \approx PQ$. Curious reader may ask why we chose the value of k is 2. This issue is explained in later. For the space constraints we show only the weight matrix Q in Table-III. The matrix Q represents the coefficients of the linear combinations of the original six variables in X that generate the transformed new variables in P i.e. rows of Q provide the relative contributions of the six variables in X to produce the new variables in P. The third variable "load" in X (weight .8836) strongly influences the first predictor in P so we can consider it as the highest contributed variable in producing the matrix P. From the second row of Q the third variable "load" and the sixth variable "throughput" provide relatively strong weights. Since we have chosen "load" as the most contributed variable, we can select the sixth variable "throughput" as the second highest contributed variable. Finally, we can rank all the variables by providing weight in this manner. We visualize the relative contributions of the variables in the column space of P in Fig-5. The relative magnitude of the vectors provides the corresponding weights. The derived eigenvalues and their percentage of the total variability explained are given in Table IV. The bar plot Fig-6 shows the eigenvalues versus their explained variation. The first eigenvalue exhibits the maximum variation compare to others. We plot only three components in Fig-7 and it clearly shows that the first two components explain 94% energy of the data. Only two variables amongst six variables represent almost whole of the characteristic of original data matrix X. This is the reason why we

ISSN 2229-5518

choose the values of k is 2.

7 Conclusion

In this paper, we have proposed a weight metric, derived from a data matrix of a network scenario to select lesser parameters that consequently leads to a parameter reduction. This method is easy to implement and the results are easily interpreted. The method gives a simple matrix manipulation and the ordered eigenvalues is an inherent property of the method itself. We emphasize the eigenvalues of the covariance matrix which believed to be important. Additionally, we provide the importance as well as quantifying contribution of the parameters by discounting those eigenvalues that are believed to be unimportant. The weighted metric and eigenvalues of the covariance matrix lead to parameter choices that appear finally parameter reduction. The primary benefit of our method is that, it quantifies the importance of each dimension for describing the variability of a data and provides a means for comparing the relative importance of each dimension. This method is not used just for WLAN but it is more general. From our experiment we have found that only two parameters "network load" and "throughput" among six, exhibit 94% percent contributions of whole network. So to evaluate the performance of a network, we emphasize more on these two parameters. We expect that, our approach will help to demystify the parameter reduction technique of a network and of great use in future.

Acknowledgment

The authors greatly acknowledge the financial support sponsored, in part, by EU-FP7 project OneLab2. We would also like to thank Joao Paulo Barraca and Diogo Gomes for numerous helpful discussions.

References

[1] IEEE 802.11 WG, Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, Aug. 1999.

[2] K.V.R. Kanth, D. Agarwal, A.E. Abbadi, & A. Singh, "Dimensionality Reduction for Similarity Searching in Dynamic Databases," *ACM SIGMOD Conference Proceeding,* pp. 166-176, 1998.

[3] Ye J. Janardan, R. Li, "An Efficient Dimension Reduction Scheme for Image Compression and Retrieval," *ACM SIGKDD Conference Proceedings*. 2004.

[4] P. Mitra, C.A. Murthy, K. Sankar Pal, "Unsupervised Feature Selection Using Feature Similarity," *IEEE transaction on pattern analysis and machine intelligence*. vol. 24, no 3, march 2002.

[5] M. Berrya, M. Brownea, A.N. Langvilleb, V.P. Paucac, R.J. Plemmonsc, "Algorithms and Applications for Approximate Nonnegative Matrix Factorization," *Computational Statistics and Data Analysis* 52. pp. 155-173, 2007.

[6] J. Shlens, "A Tutorial on Principal Component Analysis", *Center for Neural Science*, New York City, NY 10003-6603, April 22, 2009.

[7] I.T. Jolliffe, "*Principal Component Analysis*," Springer-Verlag, New York, 2002.

[8] M.A. Hasan, "Low Rank Approximations with Applications to Principal Singular Component Learning Systems," *Proceedings of the 47th IEEE Conference on Decision and Control, Cancun Mexico*, Dec. 9-11, 2008.

[9] H.H. Harman, *Modern Factor Analysis*. 3rd Ed. Chicago: University of Chicago Press, 1976.

[10] Y. Jieping , "Generalized Low Rank Approximations of Matrices," *Proceedings of the 21st International Conference on Machine Learning*. 2004.

[11] B.C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," *IEEE Transactions on Automatic Control* 26 (1), 17-32, 1981.

[12] J. Hahn, T.F. Edgar, W. Marquardt, "Controllability and Observability Covariance Matrices for the Analysis and Order Reduction of Stable Nonlinear Systems," *Journal of Process Control* , 13, 115-127, 2003.

[13] S. Lall, J.E. Marsden, S. Glavaski, "Empirical Model Reduction of Controlled Nonlinear Systems," *14th IFAC World Congress*, Beijing, 1999.

[14] C. Sun, J. Hahn, "Reduction of Differential-Algebraic Equation Systems Via Projections and System Identification," *Journal of Process Control*, 15, 639-650, 2005.

[15] C.L. Sun, J. Hahn, "Parameter Reduction for Stable Dynamical Systems Based on Hankel Singular Values and Sensitivity Analysis," *Chem. Eng.Sci*., 61(16), pp. 5393–5403, 2006.

[16] H. Castel-Taleb, L. Mokdad, "Performance Measure Bounds in Mobile Networks by State Space Reduction," *Proceedings of the 13th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems* (MASCOTS'05), 2005.

[17] A. Lakhina, K. Papagiannaki, M Crovela, C. Diot, Eric D. Kolaczyk, N. Taft, "Structural Analysis of Network Traffic Flows," *SIGMETRICS/Performance.* New York, NY, USA, June 12–16, 2004.

[18] Y. Nasr Harandi, M. Yaghoubi Waskasi, M. Pirhadi, M. Mirzabaghi, A. Iravani Tabrizipoor, " A Test Methodology for Testing Next Generation Broadband IP Access Services," *International Conference on Advanced Communication Technology* (ICACT 2007), Seoul, Korea, Feb. 2007.

[19] S.M.R. Sadri, M.Pirhadi, Y. N. Harandi, M.Y. Waskasi, A.I. Tabrizipoor, M. Mirzabaghi, " Test Strategy For DSL Broadband IP Access Services," *Highcapacity Optical Networks and Enabling Technologies* (HONET 2007), Dubai UAE, Nov. 2007

[20] D. Lee, H. Seung, "Algorithms for Non-Negative Matrix Factorization," *Adv. Neural Inform. Process. Systems 13*, pp. 556-562, 2001.