Journal of Theoretical and Applied Information Technology
30th April 2013. Vol. 50 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ON THE APPLICATION OF BLIND SOURCE SEPARATION ALGORITHM IN SPEECH SEPARATION

[1]NIQIN JING
[1]Beijing Polytechnic, Beijing 100015, CHINA
E-mail: [1]jingniqin@163.com

ABSTRACT

Blind source separation technology refers to the process for observing the recovery of source signals by mixed signals through statistical analysis on the characteristics of source signals under the situation that the source signals and signal transmission signals are unknown, which is applied in many fields, particularly used extensively in processing speech signals, array signals, images and medical signals, etc. It is pretty important for separating speech signals, regarded as a key technology in current research field of speech signals. This paper plainly studied the blind source separation, including technical status quo and development as well as principles of blind source separation technologies and solutions, etc. Meanwhile, it roughly introduced the application of separation algorithm in processing digital signals. In particular, the feasibility of the algorithm was validated by analyzing basic theories and simulating speech signals. Moreover, this paper examined the effect of separation algorithm when it was practically used for blind source separation of speech signals.

Keywords: *Blind Source Separation, Speech Signals, Separation Algorithm*

1. INTRODUCTION

In daily life, various signals emerge around us, in which a great deal of pretty useful information is included. After obtaining signals by devices, useful information can be acquired by processing the data. Nevertheless, some signals are noise and useless, among which some may be predictable and known while others might be completely unknown. Hence, it is difficult to gain real and useful information from signals. Particularly, it will be of greater difficulties to acquire accurate and true information which can generally be obtained by detecting and processing signals with sensors. Thus, new methods are constantly emerging to improve the detection of sensors, e.g. detecting the authenticity of source signals by forming array. After analyzing received signals, it is convenient to extract necessary and useful information from large amounts of data. Besides, the information is used by receivers as reference, which is the major task of processing signals.

Blind source separation, rising in 1980s, is one of the hot research topics in the field of information processing at present, and actually refers to blind identification of speech signals. It is used for processing images and signals in multiple fields including telecommunication, vibration engineering, biomedical engineering and array. It is also utilized in the field of remote sensing, particularly for SONAR, telecommunication, radar, speech and image processing, playing a crucial role in military development and national defense. In fact, blind source separation is most typically used for processing speech signals. The most representative examples are "issues about birthday parties". In other words, people talk with each other at a birthday party in which various sounds are mixed. Under this situation, the recordings recorded in this occasion are analyzed to separate certain person's voices. Hence, blind source separation refers to the separation and extraction of source signals from such mixed signals. "Blind" means the separation is implemented under the situation that source signals and the features of mixed system are unknown. Of course, the solutions somewhat differ for different types of signals. Due to the features of blind source separation, there are diversified separation methods.

2. STATUS QUO OF THE RESEARCH ON BLIND SOURCE SEPARATION TECHNOLOGIES

2.1. Origin of the Research on Blind Source Separation Technologies

It is generally believed that it was French people that first studied blind source separation of signals from 1985 or so. In 1991, several articles of great significance for improving blind source separation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

were published by a journal named Signal Processing, in which well-known H-J Algorithm for Blind Separation was proposed and special algorithm for CMOS chips was developed. These researchers imitated the features of auricular nerve in the nerve network by bionic thoughts, which impacted subsequent researchers' studies. In 1995, Bell and Sejnowski made landmark contributions to independent component analysis, which laid a foundation for algorithms. Hyvarinen H1 put forward the fixed-point training algorithm suitable for a single source signal with positive or negative kurtosis. Later, Cichocki brought forth the natural gradient algorithm effective for separation and convenient for calculation. Amari proved the effectiveness of this algorithm after further analysis and research. Candoso and Laholdwl defined relative gradient, from which the algorithms obtained were regarded as randomized. In fact, this algorithm and natural gradient algorithm are equivalent. [1-2] Domestic researchers conducted relevant studies at a later period, but there have been an increase in domestic studies over the past few years. "Time Series Analysis-High Order Statistics Method", firstly published by Professor Zhang Xianda from Tsinghua University, introduced the basic theories of blind source separation. Subsequently, more studies about blind source separation begun to emerge gradually.

On the development prospects of blind source separation. After nearly two decades of development, basic theoretical framework and related algorithms of blind source separation have been generally perfected, while bright prospects have been gradually presented in the application field. However, some difficulties still remain to be handled, among which the major difficulties covers blind source separation in noisy mixtures, nonlinear mixing, underdetermined and non-stationary signals, etc. A general survey was made on the development of processing blind signals particularly in mid 1990s when researchers actively conducted relevant studies, many algorithms in combination with mathematical tools, and some processing thoughts.

3. BASIC DESCRIPTION OF BLIND SOURCE SEPARATION

As regards blind source separation, emphasis was mainly laid on introducing mathematical model and the model classified according to hybrid system. The first introduced was mathematic model for blind source separation. The blind source separation was described through the schematic diagram as Figure 1

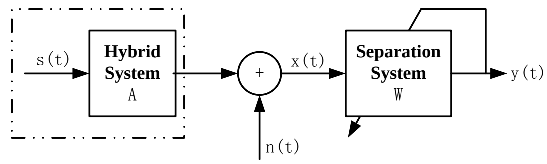

Figure 1: Principles of Blind Source Separation

Where, S(t)=[S(1), s(2)---s(n)] were unknown n-dimensional source signals, A represented unknown hybrid system, and x(t)=[X(1), X(2)---X(M)] were m-dimensional signals impacted by n(t)=[n(1), ---n(m)] (noises). With w as the separation system, then y was the estimation of source signals. Blind sources were classified according to concrete conditions because the hybrid system was unknown. One was linear mixing while the other was convolutive mixing. In terms of linear instantaneous mixing, the mixing method (the impacts of noise were not considered) is presented in the form of following formulas:

$$x_1(t) = a_{11}s(t) + a_{12}s(t) + \cdots + a_{1n}s(t)$$
$$x_2(t) = a_{21}s(t) + a_{22}s(t) + \cdots + a_{2n}s(t)$$
$$\cdots$$
$$x_m(t) = a_{m1}s(t) + a_{m2}s(t) + \cdots + a_{mn}s(t)$$

(1)

Written in the form of matrix:

$$x(t) = As(t)$$

(2)

Where, A represents hybrid matrix. Formula (2) can also be expressed in the form as follows:

$$x_i(t) = \sum_{j=1}^{N} a_{ij} s_j(t)$$

(3)

Thus, in terms of linear instantaneous mixing, blind source separation was described as follows: when "S(t)" (wave of source signals) and "a" (hybrid system" were unknown, "S(t)" (the vector of source signals) or "A" (hybrid matrix) is estimated according to "X(t)" (the mixed signals received by a sensor). Meanwhile, the blind source separation of source signals was presented as: when "s(t)" (vector of source signals) and "A" (hybrid matrix) were unknown, the mixing matrix of n*m solutions ("W") was solved. Finally, "X(t)" (vector of mixed signals) was linearly transformed by "W" into:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Y(t) = W_x(t) \tag{4}$$

To estimate "s(t)" (vector of source signals) more accurately, matrix "W" was generally marked as separation matrix.

When Formula (2) and (4) are integrated, it is discovered that

$$Y(t) = W_x(t) = W_{As}(t) = C_s(t) \tag{5}$$

Where, C=WA was the matrix of n*n, called as compound matrix.

The second type is convolutive mixed model. Although instantaneous linear mixed models have been covered in the research on blind source separation technologies for 20 years, generated many algorithms and widely applied, there are still some limitations for this algorithm. Then, what are convolutive mixed models? Under experimental environment, voices, in fact, are partially transmitted directly, while some are finally received by recording devices after being reflected by wall and refracted by air (this is actually multipath effect). To better depict this situation, experts have eventually put forward convolutive mixed modes which can be expressed as follows:

$$x_i(t) = \sum_{j=1}^{n} \sum_{k=0}^{K-1} a_{ij}(k) s_j(t-k) + v_i(t)$$

$i=1,2\cdots n$

(6)

From the formula, it can be seen that the signal "x" was the impulse response, which went through the process of convolutive mixing. Aij (k) was the coefficient of the filter between the jth sensor and ith sensor and k was the length of response pulse of the filter. Thus, Formula (6) could also be expressed as follows:

$$x(t) = \sum_{k=0}^{K-1} A(k) s(t-k) + v(t) \tag{7}$$

Where, "A(k)" was the hybrid matrix of m*n, including the coefficient of the kth filter, v(t) was the noise vector and the dimension was m*l.

4. THE APPLICATION OF BLIND SOURCE SEPARATION ALGORITHM IN SPEECH SEPARATION

In this paper, the blind source separation algorithm for speech signals was implemented on the basis of DSP (Digital Signal Processor). Before the algorithm implementation, signals were properly preprocessed, including zero mean normalization and whitening. Then, the hardware architecture, experimental environment, building and experimental processes of DSP were discussed in detail. Next, specific cases on speech separation were exemplified.

4.1. Preprocessing Speech Signals

Signals should be preprocessed before algorithm implementation no matter the simulation was fulfilled in Matlab or the separation was realized in DSP, while it appeared more necessary to preprocess signals in DSP the computational capacity of which was not so strong as PC. [3-4]

At first, zero mean normalization was implemented for signals, which was the most necessary for mixed signals and was also the most fundamental preliminary step. Assuming a=E(x), the normalization results of mixed signals were obtained by X-a, for the purpose that the mean of mixed signals equaled to zero. Generally, there is always a hypothesis in the blind source separation algorithm at present, namely the mean of all components of signal sources is 0. Hence, to satisfy the hypothesis, signals were treated by zero mean normalization. The mathematical expectation of mean can usually be replaced by arithmetic mean.

Assuming X(t)=[x_1(t), x_2(t), ……, x_n(t)], t=l, 2, ……n indicated several samples of random variables, then the process could be reflected by the Formula as follows:

$$x_{oi}(t) = x_i(t) - \frac{1}{n}\sum x_i(t) \tag{8}$$

$i = 1, 2, \cdots n$

Then, signals were whitened. Observed signals were whitened, which was a pretty useful preprocessing step. The whitening of a vector refers to linear transformation.

$$\tilde{x} = Tx \tag{9}$$

Supposed relevant matrix of transformed vectors satisfied $R\tilde{x} = E\left[\tilde{x}\tilde{x}^T\right] = I$. In other words, each component met the requirement that $E\left[\tilde{x}_i \tilde{x}_i\right] = \partial_{ii}$, then whitening of mixed signals could be realized by decorrelation of each signal component. Namely, the compoents of whitened signals should be statistically independent at least at the second order. In other words, before blind source separation algorithm was used (after zero mean normalization, observed vector (x) was linearly transformed to obtain a new vector to be whitened.

Journal of Theoretical and Applied Information Technology
30th April 2013. Vol. 50 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

That is to say, each component was uncorrelated and their standard difference was 1, or the covariance matrix was the unit matrix (I). Provided U=TA, following formula was obtained:

$$E\left[\tilde{x}\tilde{x}^T\right]=E\left[(Us)(Us)^T\right] \tag{10}$$
$$=UE\left[ss^T\right]U^T=UU^T=I$$

As shown in the above formula, matrix U was presented, namely an orthogonal matrix. After signals were whitened, the independence of each component of mixed source signals generally still couldn't be recovered, which indicated that the signals might not be separated by blind source even if the second-order irrelevance between signals was recovered. As discussed above, although it was still impossible to completely realize the blind source separation of signals by whitening signals, the separation algorithm became simper and the algorithm convergence was facilitated. Generally, whitening could be realized in two ways. On one hand, mixed signals could be linearly transformed by iterative algorithm. On the other hand, eigenvalues of a matrix related to mixed signals were decomposed. The whitening process could be completed by decomposing eigenvalues, minimizing cost functions and online whitening of signals.

4.2. Examples of Preprocessed Simulation

Comparing aforementioned three whitening algorithms, the first could be realized most easily with comparatively ideal whitening effect. Thus, hereunder the third algorithm would be applied. Before the blind source separation algorithm was implemented for the data of mixed signals, the processes of zero mean normalization and whitening were completed. [5-8] Through the simulation experiment, processing effects were apparently seen.

With different sinusoidal signals of four periods as source signals of experiments, the wave diagrams were obtained as Figure 2:

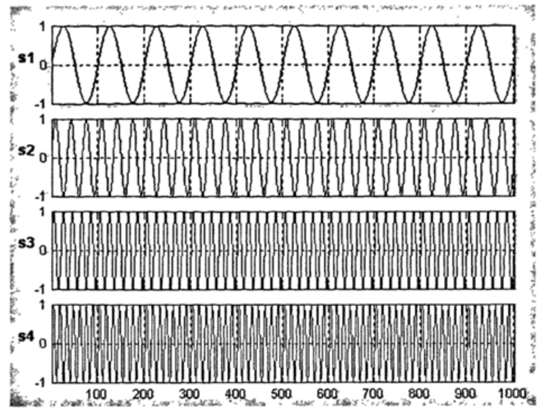

Figure 2: Wave Diagram of 4 Sinusoidal Source Signals

By random mixing, the hybrid matrix was obtained as follows:

$$A=\begin{bmatrix} 0.900259 & 0.782598 & 0.642814 & 0.843626 \\ -0.537723 & 0.524194 & -0.110593 & 0.476414 \\ 0.213685 & -0.087065 & 0.230865 & -0.647468 \\ -0.028035 & -0.962993 & 0.583874 & -0.188588 \end{bmatrix}$$

After mixing, the wave was Figure 3:

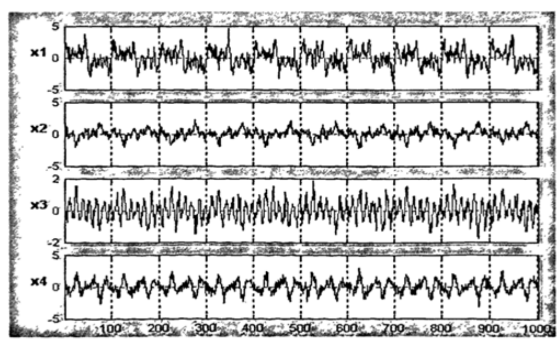

Figure 3: Wave Diagram after Mixing

After whitening, the wave was changed as Figure 4:

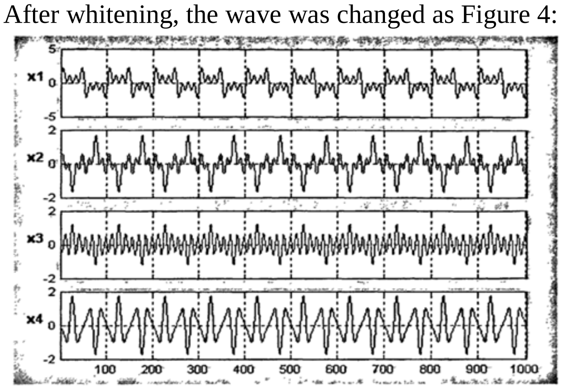

Figure 4: Wave Diagram after Whitening

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3 Analysis on Experimental Process and Results of Blind Source Separation Algorithm in DSP

From above description of blind source separation algorithm, it can be known that the authors only transplanted the algorithm operated when signals were linearly mixed to DSP. [9] Some materials have indicated that, variable step-size blind source separation algorithm based on the maximum entropy is operated under convolutive mixing environment, but the operation isn't ideal enough. The authors realized operational separation based on a hypothesis that signals were linearly and instantaneously mixed. As inspired by various research materials, to make sure of satisfying the hypothesis before experiment, it is necessary to deal with mixed signals.

(1) Linear and instantaneous mixing. Linear mixing corresponds to convolutive mixing, which means that signals from multiple paths interact with each other several times rather than once. In this experiment, experimenters' voice was only allowed to be transmitted to the sensor (namely mike) once. In other words, it should be ensured that the voice was directly transmitted to the mike without being interrupted by any obstacles and there were no echoes. Apparently, the requirements of linear and instantaneous mixing couldn't be satisfied if above conditions were not met. Inspired by other researchers, the authors did something to the experimental environment, in order that it could satisfy their hypothesis. Firstly, a relatively large experimental room was selected, on the wall of which sound absorbing materials were stuck. Secondly, it was ensured that there were no obstacles between experimental personnel and mike. Besides, the sundries should be moved away from the experimental room as much as possible, to make the room as spacious as possible. After taking these measures, it could be regarded that the environment approximately met the requirements of linear mixing studied in the algorithm.

(2) Invariability of hybrid models. Regarding the simulation of algorithm in Matlab in this paper, a hybrid system was applied to simulate signals by linear hybrid matrix. The matrix was a constant. To simulate such situation, the distance between speaker and mike should keep unchanged. In other words, experimenters were not allowed to move and they should maintain a posture as far as possible.

(3) Disposal of noise. Experiments were conducted under two situations when there was noise and there was no noise. When noise existed, the experiment was carried out in a room where many people spoke during rest, while noiseless experiment could only be done in a room where there were only experimenters and efforts should be made to prevent people from walking around the corridor as far as possible. [10] Furthermore, amplifier and mike of comparatively higher signal to noise ratio were used.

After taking aforementioned measures, the experimental environment that has satisfied the requirement for algorithm implementation was created. After the system was debugged, two experimenters sent speech signals through two microphones. A was 10cm away from a mike and 30cm from the other, while the distance from B to those mikes was just contrary to A's. For these two experimenters, some of them were homosexual, speaking in similar voices, while some were heterosexual, speaking in highly different voices. Besides, experimenters were selected from different age groups. Two experimenters started speaking simultaneously, the voice of which was recorded. For example, in one group of experiments, A said "open the conditioner" and B said "turn off the television". Then, two discreet speech signals were obtained by 8khz sampling through the voice input interface of demoboard. Then, the signals were separated by the variable step-size blind source separation algorithm based on the maximum entropy. At last, the separation results were displayed through a loudspeaker by transforming mathematical models. The transplantation of algorithm was a heavy workload, during which following discoveries were made. When two people said the same contents, the separation effect was not so ideal as the situation when the contents were different. It was guessed that there would be higher correlation when the contents were the same, which conflicted with the basic hypothesis of blind source integration algorithm. Nonetheless, the separation effect was somewhat indicated even if those two people said the same words, which proved that each person's tone was comparatively independent although it was not completely the same. The effect would be poorer if people stayed farther away from the microphone (e.g. 1m) when they spoke, because the farther the distance between them, the more convulsive the speech signals. Thus, the algorithm applied in this paper based on linear mixing isn't adaptable any more. The separation effect would be quite poor when the entire experiment was conducted in a noisier place and other conditions kept unchanged.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. RESEARCH AND PROSPECTS OF BLIND SOURCE SEPARATION

Blind source separation, as an important technology processing signals, is extensively used at present. A good many scholars conduct studies in this field and they devote themselves to studying the theories and application technologies of blind source separation. Currently, some exciting outcomes have been achieved, whereas by now many problems still remain to be further explored, deserve being further examined by experts and scholars.

At present, in the theoretical and algorithm fields, many scholars extensively and deeply study blind source separation and solve relevant problems by the knowledge of different fields such as information theories and artificial neural network, etc, who have achieved some results. [11-15] Blind source separation has been widely applied in many fields such as multi-user communication, acoustic processing, array processing, biomedical engineering and earthquake, etc. However, lots of problems still remain to be further studied and solved, particularly reflected from following aspects.

Firstly, it is hard to realize blind source separation when noise is included in mixed signals. There are some difficulties to deal with such problem which has been studied in some literatures; whereas some of them only examined Gaussian noise but seldom studied other kinds of noise such as non-Gaussian noise, colored multiplicative noise and impulse noise, etc. Furthermore, there were no great differences between the separation results obtained by processing noise in this paper and source signals. In fact, in most blind source separation or algorithm for blind deconvolution, a hypothesis is usually made that no noise is in the system or noise is treated as an independent source signal. It has always been a difficulty to realize blind source separation if noise is mixed in signals, so the use of existing blind source separation algorithm in common model mixed with noise deserves being examined. [16-20]

Secondly, underdetermined problems exist in blind source separation, which remain to be solved as well. In other words, in independent component analysis, there would be more signal sources than sensors. Lewicki and Sejnowski took the first step to solve such difficulties, proposing excessively complete ICA which has indicated that although it is highly hopeful, it is more complicated to realize and need to be further studied. [21] People have studied blind source separation for almost three decades, due to which rapid development has been gained in terms of the theories and algorithms about blind source separation. Nevertheless, it shall also be noted that, no matter from the perspective of the depth of theoretical research and algorithm implementation, there is still a long way to go in terms of blind source integration. There is still much space for conducting studies in this field, so it is necessary to be persistent.

REFRENCES:

[1] Khurram Waheed, FathiM. Salem. "State-Space Blind Source Recovery of Non-Minimum Phase Environments". *45th IEEE International Midwest Symposium on Circuit and Systems*. Tulsa, Oklahoma August 2-7, 2012, VOL. 2, pp.422-425

[2] D . T . Pham. "Blind Separation of Instantaneous mixture of sources based on order statistics". *IEEE Trans. On Signal Processing*, V01.48, No. 2, 2011, pp.363-375.

[3] Acharyya, R. , Scott, N. and Teal, P. "Non-invasive foetal heartbeat rate extraction from an underdetermined single signal". *Health*, 1, 111-116. doi: 10.4236/health.2011.12018.

[4] W. HUANG, S. WU, F. KONG and Q. WU, "Research on Blind Source Separation for Machine Vibrations," *Wireless Sensor Network*, Vol. 1 No. 5, 2012, pp. 453-457. doi: 10.4236/wsn.2012.15054.

[5] Zha, D. "New blind estimation method of evoked potentials based on minimum dispersion criterion and fractional lower order statistics". *Journal of Biomedical Science and Engineering*, 1, 91-97. doi: 10.4236/jbise.2008.12015.

[6] H. JEONG, Y. KIM and H. JANG, "Adaptive Parallel Computation for Blind Source Separation with Systolic Architecture," *Intelligent Information Management*, Vol. 2 No. 1, 2010, pp. 46-52. doi: 10.4236/iim.2010.21006.

[7] Fan, H. and Wang, F. "Using schema transforation pathways for biological data integration". *Journal of Biomedical Science and Engineering*, 1, 204-209. doi: 10.4236/jbise.2012.13035.

[8] B. Tazehkand and M. Tinati, "Underdetermined Blind Mixing Matrix Estimation Using STWP Analysis for Speech Source Signals," *Wireless Sensor Network*, Vol. 2 No. 11, 2010, pp. 854-860. doi: 10.4236/wsn.2010.211103.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[9] S. McNeill, "Considerations for Application of SOBI to Order Tracking," *Journal of Signal and Information Processing*, Vol. 2 No. 1, 2011, pp. 33-36. doi: 10.4236/jsip.2011.21005.

[10] A. Hidri, S. Meddeb and H. Amiri, "About Multichannel Speech Signal Extraction and Separation Techniques," *Journal of Signal and Information Processing*, Vol. 3 No. 2, 2012, pp. 238-247. doi: 10.4236/jsip.2012.32032.

[11] B. Kusumoputro, A. Buono and L. Na, "Identification of Noisy Utterance Speech Signal using GA-Based Optimized 2D-MFCC Method and a Bispectrum Analysis," *Journal of Software Engineering and Applications*, Vol. 5 No. 12B, 2012, pp. 193-199. doi: 10.4236/jsea.2012.512B037.

[12] S. RYOO, K. LEE and I. HWANG, "Performance Analysis of an AMC System with an Iterative V-BLAST Decoding Algorithm," *Int'l J. of Communications, Network and System Sciences*, Vol. 1 No. 2, 2008, pp. 119-123. doi: 10.4236/ijcns.2011.12015.

[13] Liu, H. , Liu, J. and Wang, L. "Searching maximum quasi-bicliques from protein-protein interaction network". *Journal of Biomedical Science and Engineering*, 1, 200-203. doi: 10.4236/jbise.2012.13034.

[14] J. Bruce, "A Bioinformatics-Inspired Adaptation to Ukkonen's Edit Distance Calculating Algorithm and Its Applicability Towards Distributed Data Mining," *Journal of Software Engineering and Applications*, Vol. 1 No. 1, 2008, pp. 8-12. doi: 10.4236/jsea.2011.11002.

[15] S. Kim, "Stochastic Design of Enhanced Network Management Architecture and Algorithmic Implementations," *American Journal of Operations Research*, Vol. 3 No. 1A, 2013, pp. 87-93. doi: 10.4236/ajor.2013.31A008.

[16] M. Al-Towaiq, "Parallel Implementation of the Gauss-Seidel Algorithm on k-Ary n-Cube Machine," *Applied Mathematics*, Vol. 4 No. 1A, 2013, pp. 177-182. doi: 10.4236/am.2013.41A028.

[17] M. Khan, A. Khan, S. Shah and A. Abdullah, "An Energy Efficient Color Based Topology Control Algorithm for Wireless Sensor Networks," *Wireless Sensor Network*, Vol. 5 No. 1, 2013, pp. 1-7. doi: 10.4236/wsn.2013.51001.

[18] Y. Yang, Y. Zhou, Z. Sun and H. Cruickshank, "Heuristic Scheduling Algorithms for Allocation of Virtualized Network and Computing Resources," *Journal of Software Engineering and Applications*, Vol. 6 No. 1, 2013, pp. 1-13. doi: 10.4236/jsea.2013.61001.

[19] S. Tabatabaee and V. Vakili, "A New Method for Sensing Cognitive Radio Network under Malicious Attacker," *Int'l J. of Communications, Network and System Sciences*, Vol. 6 No. 1, 2013, pp. 60-65. doi: 10.4236/ijcns.2013.61007.

[20] A. Ahmadi, F. El Bouanani, H. Ben-Azza and Y. Benghabrit, "A Novel Decoder Based on Parallel Genetic Algorithms for Linear Block Codes," *Int'l J. of Communications, Network and System Sciences*, Vol. 6 No. 1, 2013, pp. 66-76. doi: 10.4236/ijcns.2013.61008.

[21] Panuganti, S. , John, P. , Devraj, D. and Dash, S. "Voltage Stability Constrained Optimal Power Flow Using NSGA-II". *Computational Water, Energy, and Environmental Engineering*, 2, 1-8. doi: 10.4236/cweee.2013.21001.