Cocktail parity problem solution based on modified blind extraction technique

Ahmed Kareem Abdullah[1], Hadi A. Hamed[1], Ali A. Abdullah Albakry[1], Ahmed Ghanim Wadday[2]
[1]Department of Electrical Power Engineering Techniques, Al-Mussaib Technical College, Al-Furat Al-Awsat Technical University, Babil, Iraq
[2]Department of Communication Engineering Techniques, Al-Najaf Technical College, Al-Furat Al-Awsat Technical University, Najaf, Iraq

Article Info

Article history:

Received Jul 9, 2022
Revised Aug 17, 2022
Accepted Aug 29, 2022

Keywords:

Blind source separation
Cocktail party problem
ICA
Speech separation
Wiener filter

ABSTRACT

The cocktail party problem solution is described as being the responsibility of isolating the voice signal in a noisy environment. Two popular methods for resolving this issue are blind source extraction and the Wiener filtering procedure. The blind source extraction approaches like fastICA, JADE and efficient fastICA are useful for extraction data from the mixed signals. The classical optimization techniques such as genetic algorithms or particle swarms for blind source extraction are mostly founded on the gradient and need the objective function, so the using of these techniques is very restricted. In the recent studies, the classical blind source separation techniques will not give the perfect solution for the cocktail parity problem. These methods struggle with convergence speed and accuracy issues as well. In order to enhance the separation process and get over these issues, this work adopted the glowworm swarm technique based on kurtosis as the objective function. The results show that the proposed technique produces good separation. The original and estimated signals are compared for similarity using the cross-correlation function.

Corresponding Author:

Ahmed Kareem Abdullah
Department of Electrical Engineering techniques, Al-Mussaib Technical College
Al-Furat Al-Awsat Technical University
Babil, Iraq
Email: ahmedalbakri2012@atu.edu.iq

1. INTRODUCTION

The human brain's acoustic system has the ability to pick and focus its attention on a single sound source in a noisy and loud setting, which is how the cocktail party problem (CPP) is solved [1]. According to prior study, because to the nature and settings of the CPP, it is difficult to come up with common solutions. The system cannot distinguish the audio signal or intended sources from interface noise and competing signals in an indoor noisy setting. The CCP occurs indoors, where the position and nature of signals are unknown, as well as the signal weights and ages in the mixing material [2]. Many scholars and academics have tried various ways to solve the CPP, with the Independent component analysis (ICA) technique being the most well-known.

The blind source extraction (BSE) or separation (BSS) approach is used to separate the various sources. The term "blind" refers to the fact that initial signals or the environment system's mixing are unknown or poorly understood [3]. Blind source extraction has a variety of applications in a variety of industries, including auditory signal processing systems, image separation techniques, the medical industry, and a variety of other fields. The CPP is an example of a problem that can be solved using BSE approaches.

Parande and Thomas [1] is provides a valuable and well-written review paper that examines various strategies developed by the authors to tackle the cocktail party problem and then suggests which technique is more beneficial based on applications. The review discovered that the problem can be solved using the Independent component analysis (ICA) technique, as well as principal component analysis (PCI) for information compression and the Wiener filter (WF) for de-noising. Stöter *et al.* [4] proposes an effective way for extracting the highest number of concurrent signals from speakers in a noisy circumstance utilizing a data-driven deep learning algorithm. When sufficient input and features are provided, the approach demonstrates that traditional neural networks (NN) can be utilized to solve and analyze the CPP.

Traditional blind source extraction approaches are used gradient method and obtain not sufficient results unless get primary solutions for the initial stage. However, it is not clear to choose the suitable parameters values duo to blind hypothesis approach. The classical blind techniques cannot operate well when the fitness function is incoherent, therefore the swarm methods is used for blind source techniques to solve this problem or limitations [5], [6]. Swarm techniques moves towards of natureinspired with flow different behavior of the member such as glow worm, bees, shark, bacteria and so on [7]. Glowworm swarm technique is optimization way created by Krishnanand based on manner that a glowworm load a luminescence magnitude named luciferin [8]-[10]. The research try to outcomes are matched and compared to human performance. Mavaddaty and Ebrahimzadeh [5] presents a novel model for speech separation based on the matching of the audio-visual network with the actions of the human lip with a wide range of audio separation, the suggested model is used.

2. COCKTAIL PARITY PROBLEM

In ordinary lifetime, persons are regularly confronted to attend to definite sound signals or track specific chats in the middle of challenging related conversations occurrence mentioned to cocktail party problem (CPP) [11]. The sounds that reach a listener's ears from a specific sound source virtually never occur in isolation, whether the listener is at a genuine cocktail party, strolling down a bustling street, or conversing in a packed coffee shop. They repeatedly happen when there are other competing sources and distractions in the auditory surroundings of the listener. The act of organizing this soundscape into meaningful percepts is known as "auditory scene analysis" (ASA) [12], [13]. The ASA problem is not just for people. Similar challenges must be addressed by animals as well to traverse their intricate auditory environments, escape predators, mate, and find their infants. This includes mammals, penguins, songbirds, and fish [14], [15]. Engineering systems, including anything from smart phones to military communication and surveillance equipment, face a similar difficulty. Kumar and Jayakrishnan [6] showed and discussed this problem in 1953 using the auditory human system to distinguish the speaker's gender, intensity, and location. The cocktail parity problem solution is used to build a hearing aid system for the hearing impaired. Because the human has two ears to determine the position and surroundings of desired sources, the CPP algorithms are utilized to assist the impaired human in recognizing the useful sound from the mixture. Figure 1 depicts how the brain detects people's voices among the mixed inputs.

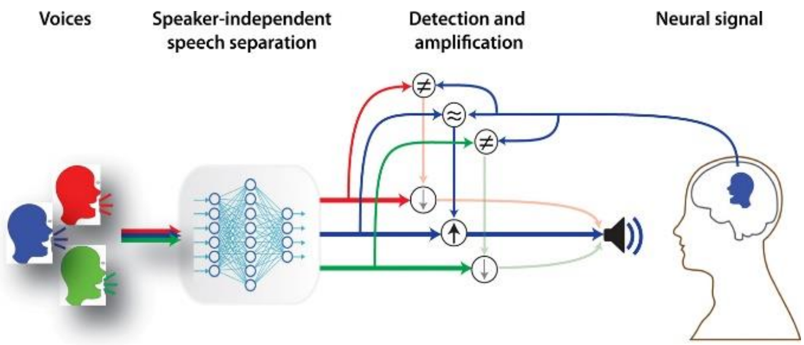

Figure 1. Sound recognition in the brain for various sounds

3. BLIND SOURCE SEPARATION

The primary goal for any blind source extraction technique is to extract or cover the original sources from the combination without relying. On sources or mixing media for guidance [3]. The primary block diagram of the blind source extraction approaches [6] is shown in Figure 2.

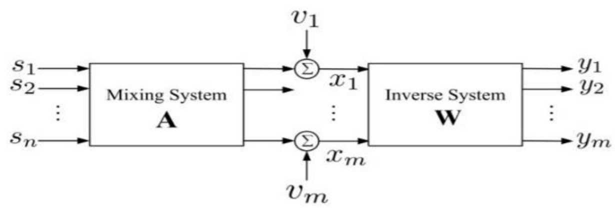

Figure 2. Blind source extraction block diagram [6]

The mathematical model for the blind source separation is (1).

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{Nr}(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1Ns} \\ a_{21} & a_{22} & \cdots & a_{2Ns} \\ \vdots & \vdots & \cdots & \vdots \\ a_{No1} & a_{No2} & \dots & a_{NrNs} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_{Ns}(t) \end{bmatrix} \tag{1}$$

The following rules regulate the system:

$$x(t) = A\, s(t) \tag{2}$$

where A is the mixing matrix, x(t) denotes the received vector, and s(t) denotes the original vector. Because there is no information about the A matrix or the original signal in this system, the goal is to obtain the original signal s(t) from only the detected signal x. (t). The final step is to get the unmixing matrix W so that (2) can be applied:

$$\hat{}\mathrm{s(t)} = \mathrm{W\, x(t)} \tag{3}$$

where W is the A^{-1} used to estimate the independent signals' sources s(t).

Blind source extraction techniques have numerous applications in a variety of fields, which are discussed below [3]-[6]:

- Medical application: EEG signal processing extracts information from the brain that is measured by EEG as a mixture of electrical brain cell signals from various places with artifacts or noise to reflect brain activity. To isolate the useful brain signal from the mixture, independent component analysis is usually utilized.
- Signal and image processing: the BSS is used to extract features from audio and video sources and find relevant renderings.
- Music: there are numerous applications in this area to extract singing and contribute to a combination, such as karaoke.
- Cocktail party: for the design of a hearing aid tool for disabled persons, blind source separation is employed to isolate each sound according to position and intensity.

4. GLOWWORM SWAM TECHNIQUE

Swarm techniques are established and used in different ways through current years. Like bee colony (BC) technique is interested from bee's performances [15]. Genetic algorithm (GA) is motivated by genetic appliance [16]. Glowworm swarm technique is suggested and based on glowworms' performance [17]. Matched with old-style optimization techniques, the swarm procedures are more operative and used to solve complex system [18], [19]. To begin, the glowworms approach is first randomly distributed over the solution space. Each glowworm carries a certain quantity of luciferin and identifies a basis functions solution that may be found in the region being searched. There is a correlation between the quantity of luciferin and the agent's current location. A more luminous one is equivalent to a more favorable opportunity glowworm swarm optimization [20]-[22] by Krishnan and Ghost mimics the glowworm's behavior of carrying a brightness quantity known as luciferin to communicate with its friends. Within the local-decision domain, an individual can only be lured by a neighbor whose luciferin brightness is greater than their own, and they can only migrate towards that neighbor by a probabilistic procedure [23]. When the density of a glowworm's neighbors is not high, the glowworm's local-decision field will extend to discover more neighbors; otherwise, it will constrict to allow the swarm to fracture into tiny teams. Density of the glowworm's neighbors calculate the size of its local-decision field. The stages that came before it are repeated over and over again until the approach determinate the ultimate final decision. Glowworm technique may be broken down into the following stages: the luciferin updating stage, the neighborhood selection stage, the moving probabilistic stage, the action stage, and the judgment radius maintain updating stage [24].

5. PROPOSED WORK AND RESULTS

One of the most important tasks for the blind source extraction is to separate the mixed signals without any information about the original sources and the cocktail parity problem is one of the famous problem solved by BSS [25]-[28]. In the model presented in Figure 3, two separate microphones are used, each of which is a weighted sum of all sources, with the weights based on the distances between the microphones. The microphones and receivers are situated in a distinct location from separate and original sources.

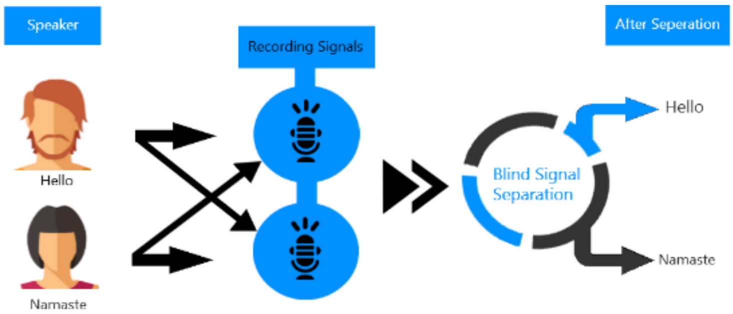

Figure 3. A solution of cocktail parity problem based on BSE

The suggested algorithm's primary steps are based on a hybridization of glowworm optimization and blind source extraction to provide the optimal separation for the task, as illustrated below:
Step 1. Collect and measure the received signals x
Step 2. Centering and whitening process to simplify the difficulties in kurtosis function. Over-all, this function is so hard to calculate from measured signal when have large information.

$$\text{Kurt}(x) = E[x^4] - 3(E[x^2])^2 \tag{4}$$

Where E is the mean vector for the received signal.
Step 3. Set the required glowworm factors (ρ, γ, β, n_t, l_0) as shown in Table 1.

Table 1. The parameters values for glowworm algorithm used in the proposed work based on [21]

Parameter	Value
luciferin decay constant (ρ)	0.4
luciferin enhancement constant (γ)	0.6
constant (β)	0.08
neighbor number (n_t)	5
luciferin (l_0)	5

Step 4. Produce or generate separation vectors as initial solution and then prepare the position $x(t)$ and decision radius $rd(t)$ of the individual.
Step 5. Compute the preliminary fitness function of each individual.
Step 6. Compute the best location x_{opt} and best fitness (x_{opt}) for the individual.
Step **7**. Calculate new luciferin $l(t)$ values based on (5), settle the neighbors $N(t)$ based on (6), and calculate compute probability $P(t)$ based on (7) [21].

$$l_i(t+1) = (1-\rho)\, li(t) + \gamma \text{ Fitness}(xi(t+1)) \tag{5}$$

$$Ni(t) = \{j : d_{ij}(t) < ri\, d(t);\ li(t) < lj(t)\} \tag{6}$$

The $dij(t)$ is Euclidean distance and $r_i\, d(t)$ is the decision radius.

$$Pij(t) = lj(t) - li(t) \sum k \in Ni(t)\, lk(t) - li(t) \tag{7}$$

Step 8. Calculate the new position $x\ (t + 1)$ and new decision radius $rd(t + 1)$ of individuals based on (8).

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t)-x_i(t)}{\|x_j(t)-x_i(t)\|}\right) \quad (8)$$

The $\|\cdot\|$ is the norm and s is the step.

Step 9. Compute the best fitness. If the new fitness is better than previous, then will updated. For every updating the radius is governed by (9):

$$r_d^i(t+1) = min\left\{rs, max\left\{0, r_d^i(t) + \beta(n_t - |N_i(t)|)\right\}\right\} \quad (9)$$

The rs is the sensory radius. Figure 4 shows the fleshly radius and judgement radius of glowworm i.

Step 10. Check the stopping criteria. Figure 5 shows the main flowchart for the proposed work. To catch the estimate original signals based on received mixtures, the blind technique approach is employed, in order to solve the Cocktail Party Problem.

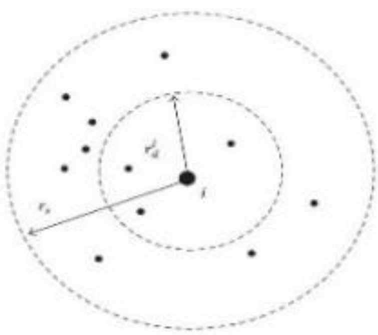

Figure 4. Sensory radius and decision

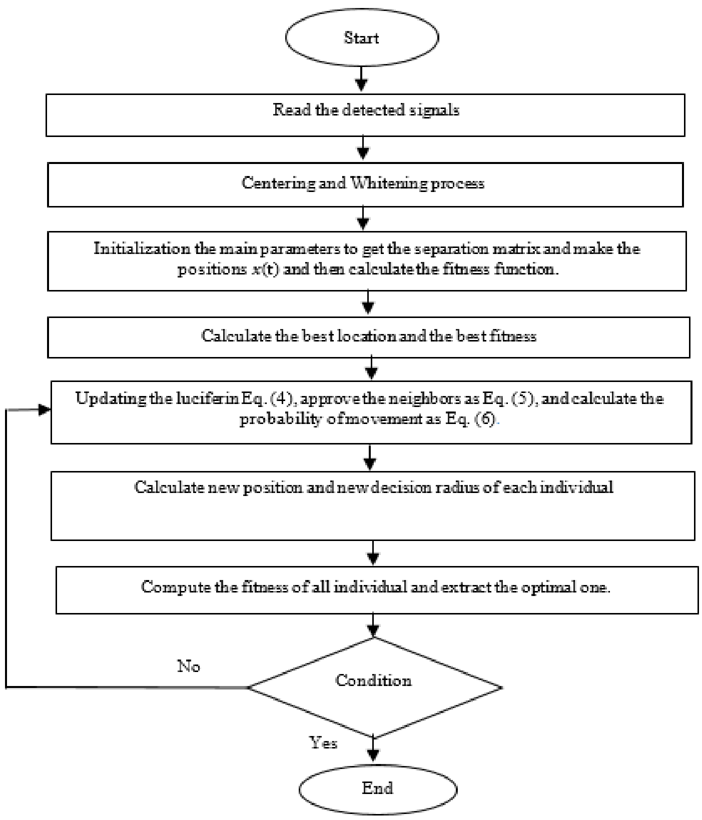

Figure 5. Flowchart for the proposed work

Also, some simple assumptions must be met, such as that the original sources are statistically independent and that the mixing medium matrix is a square. As indicated in Figure 6, two input auditory actual data files are acquired from [29]. The square matrix mixes these input audios at random to produce mixture signals, as shown in Figure 7. The proposed approach is used to extract and isolate the original signals from each other after the centering and whitening processes, as shown in Figure 8.

The preceding numbers demonstrate that the separation method, which is based on the stone BSE technique, is quite efficient, as evidenced by listening to the separated sound files. Two blind source separation approaches (rapid ICA and JADE) are also used to test the proposed approach's validity against other methods. Table 1 illustrates the performance criteria for several types of BSS; the suggested work is the best in this table since it has a high SNR and a short convergence time assessed in MATLAB using the (tic-toc) function. The similarity between original and extracted signals is calculated using the MATLAB 'corrcoef' function, as shown in the last column of Table 2. The values in the diagonal matrix (0,0) & (1,1) represent the autocorrelation function of the source with itself, so the value should be equal '1', but other values in the correlation matrix for proposed approach closed to '1' with high accuracy when compared to FICA and JADE different blind source extraction techniques.

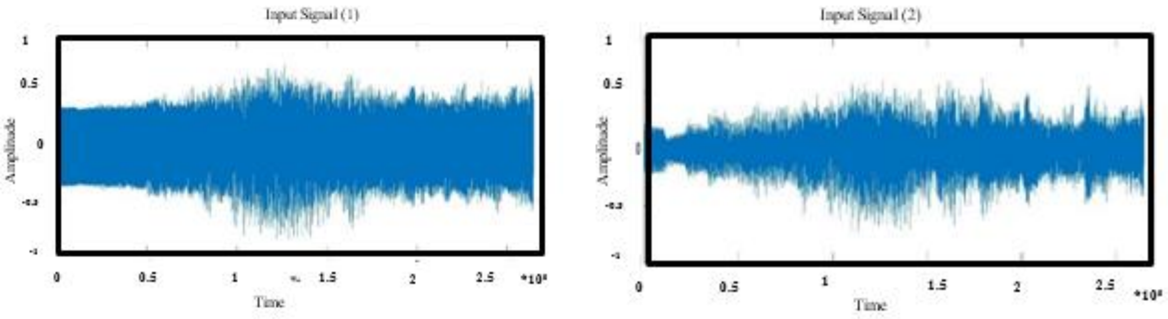

Figure 6. The utilization of two real input signals

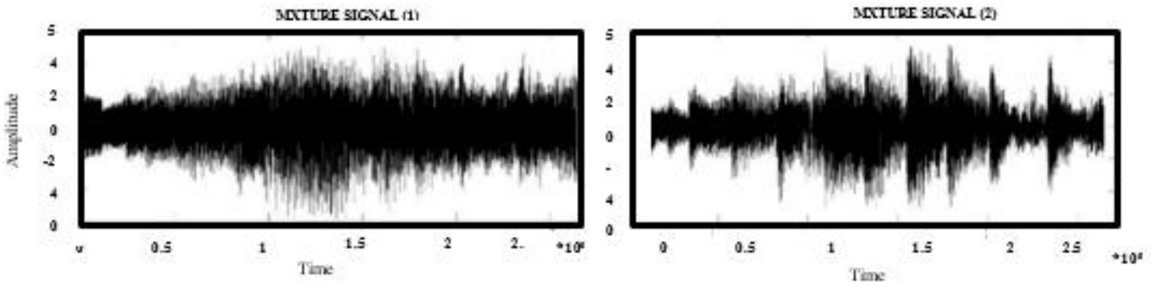

Figure 7. Two mixed signals

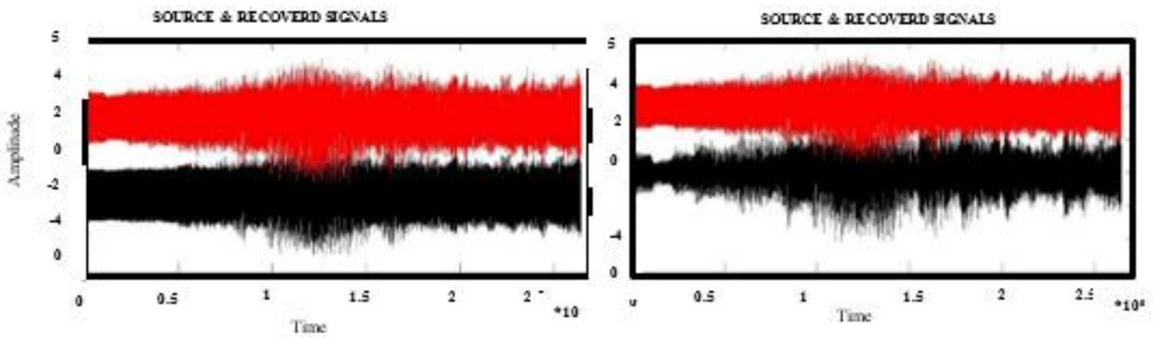

Figure 8. Original and extracted signals using the suggested method, vertically shifted for comparison

Table 2. The compression results

	Proposed Work			FICA			JADE		
Signal	SNR	Elapsed time	correlation matrix (r)	SNR	Elapsed time	correlation matrix (r)	SNR	Elapsed time	correlation matrix (r)
Signal (1)	7.887	0.47187	$\begin{bmatrix} \mathbf{1} & \mathbf{0.999} \\ \mathbf{0.998} & \mathbf{1} \end{bmatrix}$	7.108	0.994963	$\begin{bmatrix} \mathbf{1} & \mathbf{0.998} \\ \mathbf{0.899} & \mathbf{1} \end{bmatrix}$	5.988	0.890705	$\begin{bmatrix} \mathbf{1} & \mathbf{0.981} \\ \mathbf{0.971} & \mathbf{1} \end{bmatrix}$
Signal (2)	16.138			12.123			13.563		
Mean	**12.0125**			**9.6155**			**9.7755**		

6. CONCLUSION

To propose the good optimization approach for blind source extraction techniques, an innovative step change approach for the glowworm swarm technique is used to enhance the separation process. The experiment results show the GSO has robust universal search capability and quicker convergence. thus, the this algorithm is more appropriate for BSE. Glowworm optimization technique is presented as a new technique for solving the cocktail parity problem, and it is used as a first in this field, where practically all cocktail parity problems are solved using classic source separation techniques. This technique is a better choice and more effective than the other blind source extraction strategies now in use. In addition, this research proposes a comparison examination of three blind strategies for solving the problem and evaluating their performance. The goal of this research is to employ it in the 'computational auditory scene analysis' (CASA) system in the future. Soft computing techniques can also be utilized to improve the separation process and isolate high-performance signals. Finally, the findings show that the suggested approach is capable of efficiently separating the original source while requiring less convergence time.

REFERENCES

[1] P. G. Parande and T. G. Thomas, “A study of the cocktail party problem,” *International Conference on Electrical and Computing Technologies and Applications, IEEE, Ras-Al-Khaimah, United Arab Emirates*, 2017, doi:10.1109/ICECTA.2017.8251979.

[2] K. Wiklund, “The cocktail party problem: solutions and applications,” *MSC, McMaster University*, 2009. doi: 10.1162/0899766054322964.

[3] Y. Yang, Z. Li, X. Wang, and D. Zhang, “Noise source separation based on the blind source separation,” *Chinese Control and Decision Conference (CCDC), Mianyang, China*, 2011, doi: 10.1109/CCDC.2011.5968579.

[4] F. R. Stöter, S. Chakrabarty, B. Edler, and E. A. P. Habets. “CountNet: estimating the number of concurrent speakers using supervised learning,” *IEEE/ACM Transactions on Audio, Speech and Language Processing, Institute of Electrical and Electronics Engineers*, vol. 27, no. 2, pp. 268-282, 2019. doi: 10.1109/TASLP.2018.2877892.

[5] S. Mavaddaty and A. Ebrahimzadeh, “Blind signals separation with genetic algorithm and particle swarm optimization based on mutual information,” *Radio electronics and Communications Systems*, vol. 54, no. 6, pp. 315–324, 2011, doi: 10.3103/S0735272711060045.

[6] A. N. Kumar and G. Jayakrishnan, “Application of ICA for separation of artifacts in meg/eeg signals by blind source separation using improved particle swarm optimizer,” *Procedia Engineering*, vol. 30, pp. 1020–1028, 2012, doi: 10.1016/j.proeng.2012.01.959.

[7] K. N. Krishnanand and D. Ghose, “Glowworm swarm optimization for searching higher dimensional spaces,” in *Innovations in Swarm Intelligence, C. P. Lim, L. C. Jain, and S. Dehuri, Eds., vol. 248 of Studies in Computational Intelligence*, pp. 61–75, 2009, doi: 10.1155/2016/5481602.

[8] R. Dogra and N. Gupta “Glowworm swarm optimization technique for optimal power flow,” *Advance in Electronic and Electric Engineering*, pp. 155-160, vol. 4, no. 2, 2014.

[9] K. N. Krishnanand and D. Ghose, “Glowworm swarm optimization for multimodal search spaces,” in *Handbook of Swarm Intelligence*, pp. 451–467, 2011, doi: 10.1007/978-3-642-17390-5_19.

[10] N. Zainal, A. M. Zain, N. H. M. Radzi, and A. Udin, “Glowworm swarm optimization (GSO) algorithm for optimization problems: a state-of-the-art review,” *Applied Mechanics and Materials,* vol. 421, pp. 507-511, 2013.

[11] E. C. Cherry, “Some experiments on the recognition of speech, with one and with two ears,” *The Journal of the Acoustical Society of America*, vol. 25, no. 5, pp. 975–979, 1953, doi: 10.1121/1.1907229.

[12] E. C. Cherry, “On human communication,” Cambridge, MA: MIT Press. 1957, doi: 10.2307/1292231.

[13] A. S. Bregman, “Auditory scene analysis: the perceptual organization of sound,” CambridgeMA: MIT Press, 1990, doi: 10.1121/1.408434.

[14] A. Izumi, “Auditory stream segregation in Japanese monkeys,” *Cognition*, vol. 82, no. 3, pp. B113– B122, 2002, doi: 10.1016/S0010-0277(01)00161-5.

[15] X. Xu, Q. Hou, C. Wu, and Z. Fan, “Improved GSO algorithms and their applications in multi-target detection and tracking field," *IEEE Access*, vol. 4, 2016, doi: 10.1109/ACCESS.2020.2986492.

[16] M. G. Li, M. Li, G. P. Han, N. Liu, Q. M. Zhang, and Y. O. Wang, “Optimization analysis of the energy management strategy of the new energy hybrid 100% low-floor tramcar using a genetic algorithm," *Applied sciences,* vol. 8, pp. 1144, 2018, doi: 10.3390/app8071144.

[17] Y. Sun, W. Xiong, Z. Yao, K. Moniz, and A. Zahir, “Analysis of network attack and defense strategies based on pareto optimum," *Electronics,* vol. 7, 2018. doi: 10.3390/electronics7030036.

[18] Y. Zhou, X. Li, and L. Gao, “A differential evolution algorithm with intersect mutation operator," *Applied Soft Computing,* vol. 13, no. 1, pp. 390-401, 2008, doi: 10.1016/j.asoc.2012.08.014.

[19] W. B. Hu, L. P. Yan, K. Z. Liu and H. Wang. “A short-term traffic flow forecasting method based on the hybrid PSO-SVR,’’ *Neural Processing Letters*, vol. 43, pp. 155-172, 2016, doi: 10.1007/s11063-015-9409-6.

[20] K. N. Krishnanand and D. Ghose, "Glowworm swarm optimization for searching higher dimensional spaces," in *Innovations in Swarm Intelligence, C. P. Lim, L. C. Jain, and S. Dehuri, Eds., vol. 248 of Studies in Computational Intelligence*, pp. 61–75, Springer, Berlin, Germany, 2009.

[21] Z. Li and X. Huang "Glowworm swarm optimization and its application to blind signal separation," *Mathematical Problems in Engineering*, 2016, doi: 10.1155/2016/5481602.

[22] K. N. Krishnanand and D. Ghose, "Detection of multiple source locations using a glowworm metaphor with applications to collective robotics," *Proceedings of the IEEE Swarm Intelligence Symposium (SIS '05),* pp. 87–94, IEEE, June 2005, doi: 10.1109/SIS.2005.1501606.

[23] X. Cui, C. T. Hardin, R. K. Ragade, and A. S. Elmaghraby, "A swarm approach for emission sources localization," *Proceedings of the 16th International Conference on Tools with Artificial Intelligence (ICTAI 2004),* 2004, pp. 424-430.

[24] H. Ando, Y. Oasa, I. Suzuki, and M. Yamashita, "Distributed memoryless point convergence algorithm for mobile robots with limited visibility," *IEEE Trans. on Robotics and Automation,* pp. 818-828, vol. 15, no. 5, October 1999.

[25] P. C. Loizou, "Speech enhancement: Theory and practice," 2nd ed. Boca Raton, FL: CRC Press. 2013, doi: 10.1201/9781420015836.

[26] R. Lu, Z. Duan, and C. Zhang, "Listen and look: audio-visual matching assisted speech source separation," *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1315-1319, 2018, doi: 10.1109/LSP.2018.2853566.

[27] E. C. Cherry, "Some experiments on the recognition of speech with one and two ears," *Journal of the Acoustical Society of America*, vol. 25, no. 5, pp. 975-979, 1953, doi: 10.1121/1.1907229.

[28] B. Kour, R. Singh, P. Lehana, and P. Rajput, "Effect of singular value decomposition based processing on speech perception," *International Journal on Natural Language Computing (IJNLC)*, vol. 2, no. 1, 2013, doi: 10.5121/ijnlc.2013.2102.

[29] R. K. Olsson, "Algorithms for source separation- with cocktail party applications," *MSC, Technical University of Denmark/ Informatics and Mathematical Modelling*, 2007.

BIOGRAPHIES OF AUTHORS

Ahmed Kareem Abdullah is a lecturer in Electrical power engineering techniques at the Al-Furat Al-Awsat technical university, Al-Mussaib city, Iraq. He received BSc. Eng. degree Mosul- Iraq in 2000 and MSc. Eng. degrees in electronic and communication Engineering from Al-Mustansiriya University-Iraq in 2002, and then from 2010 to 2015 received PhD in information and signal processing from Harbin engineering university-China. He has been an Associate Professor in Al-Furat Al-Awsat technical university, Iraq, since 2017. He is currently head of departments in electrical engineering technique department. His research interests include the field of soft computing technique, EEG signal, brain computer interface (BCI) system, renewable energy, smart home application, artificial intelligence. He can be contacted at email: ahmedalbakri2012@atu.edu.iq.

Hadi A. Hamed is a lecturer in Electrical power engineering techniques at the Al-Furat Al-Awsat technical university, Al-Mussaib city, Iraq. He received BSc. Eng., MSc. Eng. And PhD degree university of technology-Iraq. He has been an Associate Professor in Al-Furat Al-Awsat technical university, Iraq. He is currently lecture in electrical engineering technique department. His research interests include the field of soft computing technique, communication, renewable energy, artificial intelligence. He can be contacted at email: hadihamed@atu.edu.iq.

Ali A. Abdullah Albakry is a lecturer in Electrical power engineering techniques at the Al-Furat Al-Awsat technical university, Al-Mussaib city, Iraq. He received BSc. And MSc. From Baghdad university and PhD. Eng. degrees in electrical power Engineering from University of Technology-Iraq. His research interests include the field power system, renewable energy and soft computing technique He can be contacted at email: draalia@atu.edu.iq.

Ahmed Ghanim Wadday is a lecturer in communication engineering techniques at the Al-Furat Al-Awsat technical university, Al-Najaf Technical College, Iraq. He received BSc, MSc. and PhD from University of Technology-Iraq. His research interests include the field communication system, internet of things (Iot). Renewable energy and soft computing technique. He can be contacted at email: bahmadghw@atu.edu.iq.