

Ant Colony Optimization Based Subset Feature Selection in Speech Processing: Constructing Graphs with Degree Sequences

R. Rajesvary Rajoo^{**}, Rosalina Abdul Salam[#]

[#]Faculty of Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai, Negeri Sembilan, Malaysia
E-mail: rajes_e@nilai.edu.my, rosalina@usim.edu.my

^{**}Faculty of Engineering and Technology, Nilai University, No 1, Persiaran Universiti, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

Abstract— Feature selection or the process of selecting the most discriminating feature subset is an essential practice in speech processing that significantly affects the performance of classification. However, the volume of features that presents in speech processing makes the feature selection perplexing. Moreover, determining the best feature subset is a NP-hard problem (2^n). Thus, a good searching strategy is required to avoid evaluating large number of combinations in the whole feature subsets. As a result, in recent years, many heuristic based search algorithms are developed to address this NP-hard problem. One of the several meta heuristic algorithms that is applied in many application domains to solve feature selection problem is Ant Colony Optimization (ACO) based algorithms. ACO based algorithms are nature-inspired from the foraging behavior of actual ants. The success of an ACO based feature selection algorithm depends on the choice of the construction graph with respect to runtime behavior. While most ACO based feature selection algorithms use fully connected graphs, this paper proposes ACO based algorithm that uses graphs with prescribed degree sequences. In this method, the degree of the graph representing the search space will be predicted and the construction graph that satisfies the predicted degree will be generated. This research direction on graph representation for ACO algorithms may offer possibilities to reduce computation complexity from $O(n^2)$ to $O(nm)$ in which m is the number of edges. This paper outlines some popular optimization based feature selection algorithms in the field of speech processing applications and overviewed ACO algorithm and its main variants. In addition to that, ACO based feature selection is explained and its application in various speech processing tasks is reviewed. Finally, a degree based graph construction for ACO algorithms is proposed.

Keywords— feature selection; speech processing; heuristic algorithms; ant colony optimization; degree sequences.

I. INTRODUCTION

Most speech processing tasks employ Machine Learning paradigm in which a classifier must undergo a proper learning process. The performance of classification in machine learning is strongly associated with salient features. Therefore, selecting salient features from the feature vectors, which consist of large set of feature values, is very crucial. Extracting salient features from the given set of features will reduce the dimensionality of the data set and consequently raise the accuracy and runtime performance of the classifiers [1]. Feature selection process is aimed to generate a reduced set of most discriminative features from the existing feature set by eliminating redundant and irrelevant features. The feature selection process is comprised of two main parts; the searching strategy that explores the search space that select a subset of features and a measurement procedure that evaluates the quality of these subsets of features and makes the best subset to be selected [2].

Determining the most appropriate feature subset is a NP-hard problem (2^n) where n denotes the number of features. Thus, a good searching strategy is required to avoid evaluating large number of combinations in the whole feature subsets. As a result, many searching strategies have been proposed in the literature such as Sequential Backward Selection, Sequential Forward Selection, Bidirectional Selection and Complete Search. These search processes are mainly categorized into two main approaches; filter and wrapper methods. Filter-based approaches categorize features or subset of features independently of the classifier. However, wrapper approaches use a classifier to evaluate the subset of features. Some researches use embedded method to take advantages of both approaches [3, 4, 5].

Most of the searching techniques mentioned above use local search instead of global search throughout the entire process, and therefore, it is difficult to achieve near optimal to optimal solutions. Hence, in recent years there is a lot of drive from computational intelligent community for

developing heuristic based search algorithms to address NP hard problems that focus on global search algorithms by utilizing local search appropriately. These metaheuristics algorithms are based on multi agent systems and can address the problem of finding quality solutions in polynomial time [4, 6]. A metaheuristic is a set of algorithmic theories that based on heuristic design methods. These methods are applicable to a wide set of diverse optimization problems with slight modifications [4]. Several metaheuristic optimization algorithms have been presented in the literature. Among them, Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization Algorithms (ACO), Artificial Bee Colony Optimization (ABC), Tabu Search (TS), Firefly Algorithm (FA) and Simulated Annealing (SA) have been used widely in different applications such as Telecommunications, Job scheduling, Vehicle Routing, Signal Processing, Data Mining and Protein Folding [7],

It is worth noting that several researchers in the field of speech processing have explored metaheuristic optimization algorithms. Speech processing denotes to the study of speech signals and their processing techniques in various speech applications. Feature selection is an important step in speech processing in which subsets of relevant acoustic features are selected for classification. In addition to that, feature selection also plays an essential role in acoustic data analysis to show how these features are related [8]. Feature selection has been established as an important step in many practical applications of speech processing. Among them are; Speech Recognition, Speech Enhancement, Speaker Recognition, Voice Analysis, Speech Recognition and Speech Coding [10]. This is presented in Table I.

TABLE I
FEATURE SELECTION USING OPTIMIZATION TECHNIQUES IN SPEECH PROCESSING TASKS

Author	Optimization Algorithm	Speech Processing Application
LB Asl, and VM Nezhad [11]	PSO	Speech Enhancement
Hassanzadeh, K Faez and G Seyfi [12]	FA	Speech Recognition
GK Venayagamoorthy and W Zha [13]	PSO	Speech Coding
A Shahzadi et al.[14]	GA	Speech Emotion Recognition
Lee JS, Park CH [15]	SA	Visual Speech Recognition
Ermilov A.V [16]	SA	Modelling Speech Feature
Ahmed Al-Hmouz et al. [17]	PSO	Speaker Identification
Fabiola Araújo, José Filho, Aldebaro Klautau [18]	GA	Speech Syntheses
Reza Arefi, E. Tahami [19]	GA	Voice Analysis

The studies presented in Table I underlined the importance of subset feature selection in speech processing tasks and evidenced that the optimization based feature selection techniques have improved the performance of the

classification.

In addition to the optimization techniques, which have been presented in Table I, the available literature also noted that ACO based feature selection algorithms have been effectively adapted in several speech processing tasks. Among the advantages of ACO based feature selection algorithms is that they are seldom affected by the problem of feature interaction. In this paper, the authors proposed a feature selection algorithm based on ACO that is expected to reduce the computation time.

This paper is organized as follows. In section II, ACO approach and its main invariants are described, while ACO for feature selection is elucidated in section III. Section IV presents an overview of ACO based feature subset selection in various speech processing applications and section V explains ACO algorithm using degree based graph. Lastly, conclusions are drawn in section VI that includes a proposed method.

II. MATERIAL AND METHOD

A. General Description of ACO

The ACO algorithm is a population-based metaheuristic algorithm, which was originally suggested by Marco Dorigo in 1991 for addressing combinatorial optimization problems [20]. The ACO algorithm is regarded as part of Swarm Intelligence (SI), a growing discipline in Artificial Intelligence that is inspired by social behaviours of swarms that consist of a group of simple agents with no central control structure. Even though each agent (ant) is measured as unintelligent, but they synchronize with each other to achieve a coordinated intelligent behaviour [20].

ACO algorithms are based on the foraging behavior of certain ant species. During foraging, ants deposit a substance called pheromone along the trail they travel. This pheromone will influence the other ants of the colony to a large extent. However, due to their stochastic behavior, some ants opt for paths that are not yet explored. As more ants choose the shorter paths, the shorter paths will obtain more pheromone over the time. Besides that, the amount of pheromone on every trail will decrease over time due to the evaporation. Thus, at the advanced stage, the shortest path will have the utmost concentration of pheromone [20,21].

B. Algorithmic Structure of ACO

ACO algorithms represent feature space as a construction graph, where the features are represented as nodes. Each ant creates a candidate solution (feature subset) by adding the node (feature) it visits during its traversal. The general algorithmic structure of ACO approach is given in Fig. 1.

The algorithm starts with the initialization process in which the number of ants and the pheromone trails are set. This is followed by the repetition of the optimization process. At each repetition, ants construct candidate solutions depending on the pheromone values and heuristic information. Quality of these solutions is refined through a local search which is problem specific and optional. Lastly, the pheromone value is revised and the best solution will be returned if the termination condition is met [20].

```

Input: problem's construction graph
Output: best solution
Initialise()
while termination condition not met do
    ConstructAntSolutions()
    ApplyLocalSearch ()
    UpdatePheromones ()
endwhile
return best solution

```

Fig. 1 General algorithmic structure of ACO. Adopted from [20]

C. Main ACO algorithms

The first Ant Colony Optimization (ACO) algorithm was introduced in 1996 called as Ant System (AS) by Dorigo et al. to address the Travelling Salesman Problem (TSP) [20]. Most ACO methods are based on this AS. In AS, each ant creates a complete tour by moving from a node to another in the construction graph according to the probabilistic transition rule. The pheromone will be updated using pheromone updating rule after all ants have completed their tours. The process is then iterated [22].

The success of the AS has motivated researchers to introduce some extensions and improvements in the original AS. The first improvement of ACO was Elitist Ant System (EAS) [23], which is based on elitist strategy. In EAS, when all ants complete their solutions in each iteration step, only the best solution will be used to update the pheromone trail. In this way, the search is even more focused around the best so far solutions. The subsequent amended ACO algorithm was proposed by Gambardella and Dorigo [24] is Ant-Q which is a link between reinforcement learning and ACO. However, Ant-Q was replaced by a new improved version, Ant Colony System (ACS) [24]. ACS introduces local pheromone update phase in addition to the pheromone update which is done at the end of the solution construction process. This local pheromone update is carried out by all ants after each candidate solution construction step and reduces the premature convergence problem of ACO algorithms.

The Rank-Based Ant System (RAS) technique [25] was introduced to exploit the success of the elitist strategy of EAS to improve computational performance of the ACO algorithm. In RAS, the ants are sorted and given a rank according to their constructed solution. Thus, only a several elitist ants are considered in pheromone updating phase. Another ACO algorithm, Max-Min Ant System (MMAS) was proposed by Stutzle and Hoos [26] as an improved version to original AS. In this algorithm, the pheromone values updated by the best ant are bounded (e.g. [0, 1]) The Table II shows the basic characteristics of the ACO variants.

TABLE II
ACO VARIANTS

ACO variant	Authors	Improvements	Year
Ant-Q	Gambardella & Dorigo [24]	A link between reinforcement learning and ACO	1996

Elitist AS (EAS)	Dorigo, Maniezzo, and Colomi [23]	More emphasis on global-best tour	1996
Ant Colony System (ACS)	Dorigo and Gambardella [25]	Local pheromone update was introduced	1997
Rank-based AS (RAS)	Bullnheimer, Hartl, and Strauss [26]	The ants are sorted and trail update is done based on the rank	1999
MAX-MIN Ant System (MMAS)	Stützle T, Hoos HH. [27]	The pheromone values are bounded	2000

In addition to the various ACO algorithms presented in the Table II, the ACO research team has established additional algorithmic structures for refining the search process performed by ACO algorithms such as Recursive Ant Colony Optimization (RACO), AntNet, Ant Colony Optimization for Continuous Domain (ACO_R), AS Local Best Tour (AS-LBT) and Population Based ACO (P-ACO) [7].

In general, it can be concluded that majority of ACO algorithms are varying in pheromone update rule. ACO algorithms solve optimization problems by creating candidate solutions using pheromone update rule and these candidate solutions are consequently used to revise the pheromone values in order to achieve high quality solutions.

ACO algorithms have the advantage of performing high quality solutions for optimization problems, which are seldom affected by the problem of feature interaction compared to other optimization algorithms [2]. In order to use an ACO algorithm for a feature selection problem, the following points need to be addressed [28, 29]:

I. Graph Formation

The problem domain must be represented as a graph, where features are encoded as nodes and the edges between them denote the possible choices for following feature. Each ant represents a subset of features (nodes) that the ant traverses during its solution construction.

II. Transition scheme

The transition scheme helps ants to select features during its solution construction by using the pheromone trail (τ) and the heuristic measure (η) which could be entropy based measure or the rough set dependency measure. Formula (1) indicating the possibility that ant k will select feature i in its solution construction:

$$P_i^k(t) = \begin{cases} \frac{[\tau_i(t)]^\alpha [\eta_i]^\beta}{\sum_{u \in J^k} [\tau_u(t)]^\alpha [\eta_u]^\beta} & \text{if } i \in J^k \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where J^k are the possible features that can be included in the feature set and α and β are two parameters may take real positive values that related with heuristic information and pheromone trails.

III. Subset Evaluation

A method to determine the quality of the solutions is

needed to select salient features in which the resulting subsets are gathered and assessed to find the optimal subset. Subset evaluation is done either using filter based approaches that employ statistical analysis or wrapper based approaches that use classifiers or predetermined learning model to assess feature subsets [2, 3].

IV. Pheromone Update

A pheromone update scheme is needed to update the pheromone levels on edges by both depositing and evaporating. The pheromone values are updated in proportion to the quality of the trails. Thus, this defines the learning directions that lead ants to explore optimal solution in the subsequent iterations. In every repetition, the pheromone trail is updated in accordance to the formula given in (2).

$$\tau_i(t+1) = (1-\rho)\tau_i(t) + \sum_{k=1}^m \Delta\tau_i^k(t) + \Delta\tau_i^g(t), \quad (2)$$

where m is the number of ants in each iteration, ρ is the pheromone evaporation rate and g denotes the best ant at each iteration.

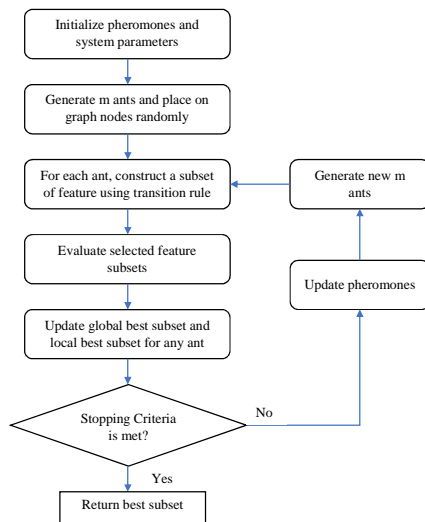


Fig. 2 ACO-based feature selection [1, 28,30]

Fig. 2 depicts the comprehensive process of ACO feature selection. In the ACO representation for feature selection process, features are coded as nodes to construct a graph model. The process starts by initializing the pheromone value and initiating a number of ants that will be positioned randomly on the graph. Every ant constructs a potential feature subset, where the features selected are the nodes it visited. The selected features are gathered from each ant and assessed. If the best subset of features has been encountered or the process has been accomplished for a definite number of times, the selection process stops and outputs the optimal feature subset found. If neither condition is met, then the pheromone value is updated. This is followed by the generation of new set of ants and the repetition of the whole process [27, 28].

III. RESULT AND DISCUSSION

Speech is the most accepted form of human communication and research in speech processing has been one of the most relevant and challenging area in signal processing [29]. Fig. 3 shows general functional block diagram of speech processing system.

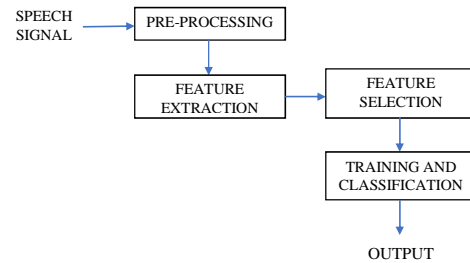


Fig. 3 Functional block diagram of speech processing

The speech samples will undergo a pre-processing task. This will be followed by feature extraction process in which the relevant features will be extracted from the speech signal. These extracted features will go through a selection process to determine the minimum number of salient features that contribute to the classification accuracy and efficiency the most.

As shown in the Fig. 3, selecting salient feature is one of the important phases in speech processing and ACO based feature subset selection has been successfully applied in several speech processing tasks.

Poonkuzhali et al. [30] applied ACO algorithm to address the optimization problem of acoustic feature set for Automatic Speech Recognition (ASR). The study proposed a subset size establishment strategy that leads the ants to assemble a reduced form of feature subsets. The experimental results show that the dimensionality of feature set gets reduced if the number of MFCC coefficients and the number of iterations are increased. ACO is able to select the most informative features to increase the performance of ASR.

Mehdi Hosseinzadeh Aghdam [28] presented ACO based method for selecting the most discriminative features to enhance the performance of Automatic Speaker Verification (ASV) system. In this study, the Equal Error Rate (EER) is used as the evaluation criteria. The results of experiments using TIMIT data set indicate that the performance of the ASV system has improved compared to Genetic Algorithm based feature selection method.

Xing Wei and Xiaojin Yang [31] have proposed ACO based algorithm to solve the dynamic time warping (DTW) problem in speech. The proposed algorithm uses adaptive evaporation coefficient in accordance to the roulette rule selection. The experimental results show that the suggested algorithm improved the accuracy compared to the traditional ant colony algorithm and the DTW and has improved global search capability. Consequently, it increases the accuracy of the speech recognition rate.

Lihui DU and Yueguang Li [32] applied an enhanced quantum ant colony algorithm in parameter optimization problem to increase the learning ability in practical English

Speech Emotion Recognition. The main idea of the algorithm is: i) a random moving ant with no load when encounter an object it compares it with surrounding objects to pick up the object with highest probability, ii) a random moving load ant will drop the load and pick up the object with highest probability from the surrounding objects. The findings from the experiments showed that the proposed algorithm improved practical English Speech Emotion Recognition.

In order to increase the performance of ACO algorithms, number researches combined it with some other optimization algorithms. J.Sirisha Devi and Srinivas Yarramalle [30] have proposed a hybrid approach of Ant Colony and Artificial Bee Colony optimized new feature subset selection procedure for ASV. In this study, a fully connected graph with each node representing a feature is constructed. An onlooker ant is randomly assigned to each feature. The proposed method showed increases in performances when tested with two different dataset; Berlin dataset and telephone conversation data set.

A fuzzy and ACO based method had been proposed for Speech Recognition by Foad Jalili and Milad Jafari Barani in [34]. In this research, speech samples are fed to fuzzy system for dimensionality reduction. The ACO algorithm is used to cluster these signals using city-block distance measure. This method showed better quality results with reduced time complexity with regard to other fuzzy systems.

The hybrid of ACO and GA is used for feature selection in [35] by Mansour Sheikhan in order to reduce the number of inputs of Dynamic Neural Network (DNN). The simulation results showed that the proposed model offers low root mean square error (RMSE).

Table III shows ACO based feature selection in various speech applications.

TABLE III
ACO BASED FEATURE SELECTION ALGORITHMS IN VARIOUS SPEECH APPLICATIONS

Author	Speech Processing Applications	Results and Discussions
Mehdi Hosseinzadeh Aghdam [28]	Speaker Verification	ASV system is improved compared to Genetic Algorithm based feature selection method.
Poonkuzhali et al. [30]	Speech Recognition	ACO is capable to gathering the most discriminative features without losing the performance.
Xing Wei and Xiaojin Yang [31]	Speech Recognition	Proposed algorithm has improved global search ability and increases the performance in the speech recognition rate.
Lihui DU and Yueguang Li [32]	Speech Emotion Recognition	Improve the learning ability in practical English speech emotion recognition.
J.Sirisha Devi and Srinivas Yarramalle [33]	Speaker Verification	Hybrid of Ant Colony and Artificial Bee Colony optimized the feature subset over 85% and decreased the computational complexity of ASV.

Milad Jafari Barani [34].	Speech Recognition	Fuzzy and ACO based method has lower time complexity regard to fuzzy system.
Mansour Sheikhan [35]	Speech Synthesis	Hybrid of genetic algorithm (GA) and ant colony optimization (ACO) has reduced root mean square error (RMSE)

In order to apply ACO based feature selection in a searching problem, the search space has to be represented as a graph. The success of an ACO based feature selection algorithm depends on the choice of the construction graph with respect to runtime behavior. In the search of optimal feature subset, a solution is constructed by letting an ant traverse through the construction graph by visiting a minimum number of nodes that fulfills the stopping criteria. While most ACO based feature selection algorithms use fully connected graphs, the authors proposed ACO based algorithm that uses graphs with prescribed degree sequences [36]. In this method, the degree of the graph to represent the search space will be predicted and the construction graph that satisfied the predicted degree will be generated. The artificial ants traverse this degree based graph to construct subset features. The main phases of the proposed algorithm are as follows:

A. Graph formation

In this phase, the search space is represented as s degree based graph using degree driven approach. In this approach, every node is linked to other nodes based on the degrees of the specified node. Therefore, the search space is represented by a degree based graph with only $O(NM)$ in which N is the number of features and M is the number of edges. The heuristic information and pheromone update scheme of ACO is applied to each node based on degree based graph representation.

B. Initialization of algorithmic parameters

In this phase, the algorithms starts and parameters such as number of ants, pheromone value and number of iterations are set.

C. Construction of Ant Solutions

In this phase, each ant traverses on the degree based graph. Ants require to construct solutions by selecting the next node based on the degree of each node using pheromone value and heuristic information calculated based on formula (1) and (2). The solution constructed by each ant denotes its own feature subset. At the end of this phase, a number of feature subsets are formed.

D. Selection of Optimum Solution

The feature subsets constructed from the previous phase are evaluated for quality solution using a subset evaluation scheme. Optimum solutions are selected and updated in accordance to the selection scheme.

E. Pheromone Update

After all ants constructed their solutions, the pheromone values of all features are updated according to the

pheromone update scheme and heuristic information. Step C to E are repeated until the stopping condition is reached. In principle, the graph constructed realizing a given sequence of degrees is expected to reduce the computational complexity and increase the performance of ACO algorithm.

IV. CONCLUSIONS

From the available literature it is known that ACO based feature selection approach has been employed in several speech processing tasks. Evidence from the previous studies indicates that ACO based feature selection algorithms differ in terms of number of iterations, number of ants, choice of guiding solution, subset construction and evaluation, pheromone update rule and also hybridization of ACO algorithms. A prevalent view on the previous studies also indicates that most of ACO based feature selection algorithms often use complete graphs with $O(n^2)$ edges; features/nodes are fully connected to each other in the graph.

Typically, ACO for feature selection is represented as a graph where the ants traverse the nodes in the graph to construct a graph model. In most of the ACO based feature selection studies, ACO algorithms use complete graphs in which each node is connected to every other node. An ant traverse from one node to the next node based on the pheromone value and heuristic information assigned to the edge connects the nodes. However there are some studies in the literature which have established connection from one node to another two nodes only, thus reducing the computation complexity from $O(n^2)$ to $O(n)$ [2, 37].

The research direction on graph representation for ACO algorithms may offer possibilities to reduce the computation complexity. Thus as future work of this review, the authors proposed a degree based graph representation for ACO algorithm in which each ant will construct a graph model with a given degree sequence. In this approach, first the degree of the graph to be generated will be predicted by extrapolation from the available data and then generate a graph that satisfies the target degree sequence. The ACO algorithm on a search space represented by a degree based graph will have computation complexity of $O(NM)$ in which M is the number of edges [36]. The proposed method will have an advantage over complete graph in term of reduced complexity and expected to give more flexibility on the search space compared to binary connected graph model.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Universiti Sains Islam Malaysia (USIM) for the supports and facilities provided. This research study is sponsored by University Sains Islam Malaysia (USIM) under USIM Competitive Grant [PPP/UTG-0114/FST/30/11414].

REFERENCES

- [1] E. Saraç and S. A. Özel, "An Ant Colony Optimization Based Feature Selection for Web Page Classification", *The Scientific World Journal*, vol. 2014 (2014), Article ID 649260, 16 pages.
- [2] Bing Xue, Mengjie Zhang, Will N. Browne, Xin Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection", *IEEE Transactions On Evolutionary Computation*, vol. 20, No. 4, August 2016.
- [3] Sina Tabakhi and Parham Moradi, "Relevance– redundancy feature selection based on ant colony optimization", *Pattern Recognition*, vol. 48 (2015) pp. 2798–2811.
- [4] Rizwan Mehmood, Waseem Shahzad and Ejaz Ahmed, "Maximum Relevancy Minimum Redundancy Based Feature Subset Selection using Ant Colony Optimization", *Journal of Applied Environmental and Biological Sciences*, 7(4)118-130, 2017.
- [5] Roiss Alhutaish and Nazlia Omar, "Feature Selection for Multi-Label Document Based on Wrapper Approach through Class Association Rules", *IJASET*, vol.7 (2017), no.2.
- [6] M. Dorigo, M. Birattari and T.Stutzle, "Ant Colony Optimization: Artificial Ants as Computational Intelligent Technique", *IEEE Computational Intelligent Magazine*, Nov 2006.
- [7] M. M Kabir, M. Shahjahan and K Murase, "A new hybrid ant colony optimization algorithm for feature selection", *Expert Systems with Applications*, vol. 39 (2012) pp. 3747–3763.
- [8] Adubi A. Stephen and Sanjay Misra, "A comparative Study on the Ant Colony Optimization Algorithms", *Electronics, Computer and Computation (ICECCO)*, 11th International Conference 2014.
- [9] Jianbo Jiang, Zhiyong Wu, Mingxing Xu, Jia Jia and Lianhong Cai, "Comparing feature dimension reduction algorithms for GMM-SVM based speech emotion recognition", *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2014 Asia-Pacific, Jan 2014.
- [10] T. Drugman, "Advances in Glottal Analysis and its Applications", *PHD THESIS*, University of Mons, Belgium, 2011.
- [11] LB Asl and VM Nezhad, "Speech Enhancement Using Particle Swarm Optimization Techniques", *Measuring Technology and Mechatronics Automation (ICMTMA)*, 2010 International Conference on, Changsha City, China, China, March 2010.
- [12] T. Hassanzadeh, K. Faez and G Seyfi, "A speech recognition system based on structure equivalent fuzzy neural network trained by firefly algorithm", *Biomedical Engineering (ICoBE)*, 2012 International Conference on, Penang, Malaysia, Feb 2012.
- [13] G.K Venayagamoorthy and W. Zha, "Comparison of Nonuniform Optimal Quantizer Designs for Speech Coding With Adaptive Critics and Particle Swarm", *IEEE Transactions on Industry Applications*, vol.43, Issue: 1, Jan-Feb. 2007.
- [14] A. Shahzadi, A. Ahmadyard and A. Harim, "Speech emotion recognition using nonlinear dynamics features", *Turkish Journal of Electrical Engineering & Computer Sciences*, *Turk J Elec Eng & Comp Sci* (2015) vol. 23, pp. 2056-2073.
- [15] J.S. Lee and C.H. Park, "Hybrid simulated annealing and its application to optimization of hidden Markov models for visual speech recognition", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40(4), pp. 1188-96. Aug 2010.
- [16] A. V. Ermilov, "Modeling Speech Features Via Simulated Annealing Algorithm", vol (2) (2014), pp. 354-358.
- [17] Ahmed Al-Hmouz , Khaled Daqrouq, Rami Al-Hmouz and Jaafar Alghazo, "Feature Reduction Method for Speaker Identification Systems Using Particle Swarm Optimization", *International Journal of Engineering and Technology (IJET)*, vol 9 (3), Jun-Jul 2017.
- [18] Fabiola Araújo, José Filho and Aldebaro Klautau, "Genetic algorithm to estimate the input parameters of Klatt and HLSyn formant-based speech synthesizers", *BioSystems*, vol. 150 (2016), pp. 190–193..
- [19] R. Arefi Shirvan and E. Tahami, "Voice analysis for detecting Parkinson's disease using genetic algorithm and KNN classification method", *Biomedical Engineering (ICBME)*, 2011 18th Iranian Conference of, 14-16 Dec. 2011.
- [20] M.Dorigo and K.Socha, "An introduction to ant colony optimization," *Handbook of Metaheuristic*, Brussels: IRIDIA, 2006, vol. 26, no 1. ISSN 1781-3794
- [21] Yusuf Hendrawan and Dimas Firmanda Al Riza, "Machine Vision Optimization using Nature-Inspired Algorithms to Model Sunagoke Moss Water Status", *IJASEIT*, Vol.6(2016), no.1.
- [22] Stephan A. Adubi and Sanjay Misra, "A comparative Study on the Ant Colony Optimization Algorithms", *Conference: 11th International Conference on Electronics, Computer and Computation (ICECCO)*, 2014.
- [23] Dorigo M, Maniezzo V, Colomi A. Ant System: Optimization by a colony of cooperating agents. *IEEE Trans Syst Man Cybernet Part B* 1996, vol. 26(1), pp. 29–41.
- [24] Gambardella L. and M. Dorigo, "Ant-Q: A Reinforcement Learning approach to the traveling salesman problem", *Proceedings of ML-95, Twelfth International Conference on Machine Learning*, Tahoe City, CA, A. Prieditis and S. Russell (Eds.), Morgan Kaufmann, 252–260, 1995

- [25] Dorigo M. and Gambardella LM. "Ant colony system: A cooperative learning approach to the traveling salesman problem", *IEEE Trans Evolutionary Comput* 1997, vol. 1(1), pp.53–66.
- [26] Bullnheimer B, Hartl R, Strauss C, "A new rank-based version of the Ant System: A computational study", *Central European J Operations Res Econom* 1999, vol. 7(1), pp. 25–38.
- [27] Stützle T. and Hoos HH, "MAX–MIN Ant system", *Future Generation Computer System* 2000, vol. 16(8), pp. 889–914.
- [28] Mehdi Hosseinzadeh Aghdam. "An Improved Ant Colony Optimization Algorithm And Its Application To Text-Independent Speaker Verification System", *JAISCR*, 2012, vol. 2 (4), pp. 301-315.
- [29] H.R. Kanan, K. Faez, and S.M. Taheri, "Feature Selection Using Ant Colony Optimization (ACO): A New Method and Comparative Study in the Application of Face Recognition System", P. Perner (Ed.): *ICDM 2007, LNAI 4597*, pp. 63–76, 2007.
- [30] C.Poonkuzhali, R.Karthiprakash, Dr.S.Valarmathy, and M.Kalamani, "An Approach to Feature Selection Algorithm Based On Ant Colony Optimization For Automatic Speech Recognition", *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 2(11), November 2013.
- [31] Xing Wei and Xiaojin Yang, "Speech Dynamic Time Warping Based on Ant Colony Optimization Algorithm", *Consumer Electronics, Communications and Networks (CECNet)*, Jan 2014.
- [32] Lihui DU and Yueguang Li, "Recognition of practical English speech emotion using improved Quantum Ant Colony Algorithm", *International Symposium on Computers & Informatics (ISCI 2015)*.
- [33] J.Sirisha Devi and Srinivas Yarramalle, "Multi Objective Optimization Problem resolution based on Hybrid Ant-Bee Colony for Text Independent Speaker Verification", *IJ. Modern Education and Computer Science*, vol. 2015 (1), pp. 55-63.
- [34] Foad Jalili and Milad Jafari Barani, "Speech Recognition Using Combined Fuzzy and Ant Colony Algorithm", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6(5), October 2016, pp. 2205-2210.
- [35] Mansour Sheikhan, "Synthesizing Supra segmental Speech Information Using Hybrid of GA-ACO and Dynamic Neural Network", *2013 5th Conference on Information and Knowledge Technology (IKT)*.
- [36] H Kim, C I Del Genio, K E Bassler and Z Toroczkai, "Constructing and sampling directed graphs with given degree sequences", *New Journal of Physics*, Volume 14, February 2012.
- [37] Bolun Chen, Ling Chen and Yixin Chen, "Efficient ant colony optimization for image feature selection", *Signal Processing* 93 (2031), pp 1566-1576.