I.J. Modern Education and Computer Science, **2015, 1, 55-63**
Published Online January 2015 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2015.01.08

Multi Objective Optimization Problem resolution based on Hybrid Ant-Bee Colony for Text Independent Speaker Verification

J. Sirisha Devi
Research Scholar, Department of CSE, GITAM University, India
Email: siri.cse21@gmail.com

Dr. Srinivas Yarramalle
Professor, Department of IT, GITAM University, India
Email: sriteja.y@gmail.com

Abstract—Today major section of automatic speaker verification (ASV) research is focused on multiple objectives like optimization of feature subset and minimization of Equal Error Rate (EER). As such, numerous systems for feature dimension reduction are proposed. This includes framework coaching and testing analysis for every feature set that could be a time esurient trip. Because of its significance, the issue of feature selection has been researched by numerous scientists. In this paper, a new feature subset selection procedure is presented. Hybrid of Ant Colony and Artificial Bee Colony optimized the feature subset over 85% thereby decreased the computational complexity of ASV. Additionally an external record is maintained to store non-dominated solution vectors for which concept of Pareto dominance is used. An overall optimization of 87% is achieved thereby improved the recognition rate of ASV.

Index Terms—Ant Colony Optimization, Artificial Bee Colony, multi-objective Optimization, Gaussian Mixture Model.

I. INTRODUCTION

Speech processing is the investigation of speech signals and the transforming routines for these signals. Text independent speaker verification obliges no limitation on the sort of data discourse [1]. There are numerous applications for programmed speaker verification frameworks.

Late years have seen an expanding research in exploration on such frameworks [16][17][20]. These frameworks typically utilize high dimension feature vectors and consequently include high multifaceted nature with multiple objectives. Be that as it may, there is a general conviction that large portions of the features utilized within such frameworks are superfluous and repetitive. In this way, numerous strategies for features measurement diminishment have been proposed.

The majority of which are wrapper-based, which is costly since framework execution is utilized for peculiarity subset assessment which includes system preparing and execution assessment for each one peculiarity subset, which is a time taking process.

In the field of computational intelligence and especially the calculations focused around Swarm Intelligence (SI) are seriously studied and effectively applied for optimization issues [3][5]. Among these problems square measure people who incorporate multiple objectives, that usually square measure exceptionally traditional in varied application ranges [8]. SI-based algorithms involve many characteristics that create them notably appropriate for finding multi objective optimization issues (MOOPs), e.g., inherently localized, the members of the swarm will be answerable of various objectives, totally different levels and kinds of interactions will be outlined so as to share individual search expertise with the remainder of the swarm, etc [11]. The foremost representative and developed SI algorithms embrace Particle Swarm Optimization (PSO) and the Ant Colony Optimization (ACO) met heuristic [12]. Many of those algorithmic rules embrace the non-dominated sorting genetic algorithm II (NSGA-II), the strength economist organic process algorithmic rule a pair of (SPEA2), and therefore the multi objective particle swarm optimization (MOPSO) that is projected by Coello and Lechuga. MOEA's success is because of their ability to search out a collection of agent Pareto optimal solutions in a single run [14]. Artificial bee colony (ABC) algorithmic rule could be a new swarm intelligent algorithmic rule that was initially introduced by Karaboga in Erciyes University of Turkey in 2005, and therefore the performance of ABC is analyzed in 2007 [9]. The ABC algorithm imitates the behaviors of real bees to find food sources and sharing the knowledge with different bees [13]. Since ABC algorithm is straightforward in thought, to implement, and has fewer management parameters, it's been wide utilized in several fields [4][7][10]. For these benefits of the ABC and ACO algorithms, we tend to propose a unique algorithmic rule "Multi objective Hybrid Ant-Bee Colony" (MOHABC), that permits the Hybrid Ant-Bee Colony algorithmic rule to be ready to take care of multi objective optimization issues.

The remaining paper is organized as follows. Section 2 gives literature review of automatic speaker verification systems. Section 3, we will give a brief review of basic concepts involved in this work. Section 4, indispensable need to optimize the features in the feature selection phase is being studied. Section 5 presents the details of MOHABC algorithm. Section 6 presents the experimental results of the proposed algorithm. Section 7 summarizes our discussion with a brief conclusion.

II. Literature Review on ASV systems

Alejandro Bidondo in his paper discussed about a speaker recognition system which used r-ACF (running Autocorrelation Function) microscopic parameters and Euclidean distances vector's distance [18]. There was no considerable improvement in accuracy rate. Md. Jahangir Alam in his paper used MFCC and low-variance multitaper spectrum estimation methods for speaker recognition. Compared with the Hamming window technique, the sinusoidal weighted cepstrum estimator, multi-peak, and Thomson multitaper techniques provide a relative improvement of 20.25%, 18.73%, and 12.83 %, respectively, in equal error rate.

Taufiq Hasan in his paper discussed the usage of PPCA for acoustic factor analysis and i-vector system for speaker verification [20]. A relative improvement of16.52%, 14.47% and 14.09% in %EER, DCF (old) and DCF (new) respectively was bring into being. Pedro Univaso in his paper used 13 MFCC coefficients with delta and acceleration and achieved 25.1% equal error rate reduction relative to a GMM baseline system. Taufiq Hasan in his paper used Mean Hilbert Envelope Coefficients (MHEC), PMVDR Front-End, Rectangular Filter-Bank Cepstral Coefficients (RFCC), MFCC-QCN-RASTALP and attained a relative improvements in the order of 50 - 60% [21]. Gang Liu in his paper used Mel frequency Cepstral coefficients (MFCC) and several back-ends on i-vector system framework [22]. He could achieve a relative improvement in EER and minimum DCF by 56.5% and 49.4%, respectively.

Balaji Vasan Srinivasan in his paper used 57 mel-frequency cepstral coefficients (MFCC) features and Kernel partial least squares (KPLS) for discriminative training in i-vector space [25]. He attained 8.4% performance improvement (relative) in terms of EER. Tomi Kinnunen in his paper used MFCCs and three Gaussian mixture model based classifiers with universal background model (GMM-UBM), support vector machine (GMM-SVM) and joint factor analysis (GMM-JFA). He achieved 20.4% (GMM-SVM), 13.7% (GMM-JFA) [26]. Tobias May in his paper used spectral features and universal background model [29]. He could achieve a substantial improvement in recognition performance, particularly within the presence of extremely non-stationary ground noise at low SNRs [28][30].

Shahla Nemati in her paper used 50 spectral features for speaker verification [2]. She was successful in reducing the feature vector size over 80% which led to less complexity of the system. No centralized processor to guide the ACO towards good solutions is the main drawback of this system. This system suffers premature convergence which leads to local optimum and shows less accurate results when real time speech signals are considered which consists of huge noise in background. Abdolreza Rashno in his paper used wrapper-based technique however makes use of Relieff weights so as to possess a lower victimization of system performance [36]. Therefore this technique has lower complexness compared to different wrapper-based strategies, will cause sixty nine feature dimension reduction and incorporates a one.25% of Equal Error Rate (EER) for the simplest case that appeared in RBF kernel of SVM. This technique showed lower EER and lower process overhead compared with 2 widespread population-based wrapper feature choice strategies, particularly ACO and GA but this system is a time consuming handbook tagging process and difficult to maintain. Monica Sood in his paper optimized speech features but increased the computational time [6].

III. Related Concepts

In this section a brief discussion on the basics of automatic speaker verification, multi objective optimization problem, ant colony and artificial bee colony is presented.

A. *Basics of ASV Model*

ASV depends mainly on the pitch frequency of the recorded voice in order to have an effective reorganization system, the speech samples are to be preprocessed before extracting features [31][32].

- Pre–emphasis is the process that helps to boosting the energy of speech signals to high frequency levels. In order to get high frequency ranges, every speech sample is processed using finite impulse response filter (FIR), outcome of which will be high order frequency curves. The equation of frustrated total reflectance FIR filter, which is of 1st order, is given by (1).

$$Y\,[n] = X\,[n] - 0.95\,X\,[n] \qquad (1)$$

This Pre–emphasis technique helps to remove silence parts and white noises [27].

- A frame of 36 ms (milliseconds) of every input speech signal is broken into, which ensures that the spectral characteristics remain the same within this time duration is known as framing. In our work, 50 % of overlapping window sizes is considered [19].
- Windowing is the next step where very frame, which is considered, will be given a shape by which the edge effects are removed. Hamming window is considered for this process, since they work better than other windows Fig 1.

The equation used for the calculation of Hamming Window is (2):

$$w(n) = 0.54 - 0.46\cos\frac{2\pi n}{n-1} \text{ where } 0 \leq N \leq n \quad (2)$$

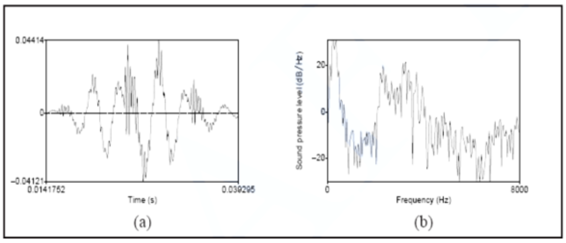

Fig. 1. A Hamming-windowed portion of a signal from a vowel and (b) its spectrum computed by a DFT

- Using fast fourier transform (FFT), log magnitude spectrum is obtained to determine MFCC.
- Then mel filter bank processing is done with 50% of overlapping Mel triangular filters are considered. First 13 coefficients are considered to obtain the first 13 features of MFCC. The following formula given by eq. (3), and is used to convert the obtained frequencies to Mel values.

$$f(mel) = 2595 * \log_{10}(1 + \frac{f}{700}) \quad (3)$$

- Discrete Cosine Transformation (DCT) decorelates and energy compaction of Mel frequency cepstral coefficients. A sequence of MFCC acoustic vector is obtained from every input speech signal which is used to generate the reference template [23].
- **T**hen delta energy and delta spectrum is calculated. The 13 delta coefficients represent the change in cepstral features over time along with an additional energy coefficient and 13 double delta or acceleration features. The 13 delta features represent the change between frames, while each of the 13 double delta features represent the change between frames in the corresponding delta features. In similar fashion all the total 39 MFCC feature are calculated for every frame which constitute feature vector [17][24].

B. Multi objective optimization

Formulation of a basic single objective optimization problem is given as follows in (4)

$$\min f(x), x \in S \quad (4)$$

where f is a scalar function and S is the set of constraints that can be defined as (5)

$$S = \{x \in R^m : h(x) = 0, g(x) \geq 0\} \quad (5)$$

In mathematical terms multi-objective optimization can be given as follows (6):

$$\min f_1(x), f_2(x), \ldots\ldots, f_n(x), x \in S \quad (6)$$

where $n > 1$ and S is the set of constraints. Objective space is the location where the objective vector belongs to. Attained set is the image of the feasible set under F and such a set will be denoted in the following with

$$C = \{y \in R^n : y = f(x), x \in S\} \quad (7)$$

The theory of Pareto optimality has been used as the notion of "optimality" does not apply directly in the multi-objective. Basically, a vector $x^* \in S$ is said to be Pareto optimal for a multi-objective problem if all other vectors $x \in S$ have a higher value for at least one of the objective functions f_i, with $i = 1, \ldots, n$, or have the same value for all the objective functions [37].

C. Artificial Bee Colony Optimization (ABC)

Artificial Bee Colony algorithm was proposed by Karaboga for upgrading numerical issues. The calculation recreates the clever scavenging conduct of bumble bee swarms. It is an exceptionally straightforward, vigorous and populace based stochastic advancement calculation. In ABC calculation, the state of artificial bees contains three gatherings of 3 groups of bee: employed bees, onlookers and scouts. A honey bee tend to the move zone for settling on a choice to pick a sustenance source is called onlooker and one heading off to the nourishment source went to by it before is named employee bee. The other sort of honey bee is scout bee that completes arbitrary quest for finding new sources. The position of a nourishment source speaks to a conceivable answer for the advancement issue and the nectar measure of a sustenance source relates to the quality (fitness) of the related arrangement.

From the current source x_i, each employed bee finds a new food source v_i in its neighborhood, in employed bees' phase. The expression in (8) is used to evaluate the new solution.

$$v_{ij} = x_{ij} + \emptyset_{ij}(x_{ij} - x_{kj}) \quad (8)$$

Where $k \in (1,2,3, \ldots N)$ and $j \in (1,2,3, \ldots n)$ are randomly chosen indexes and $k \neq i$. $\emptyset_{ij}$ is a random number between [-1,+1]. It controls the production of a neighbor food source position around x_{ij}. Then greedy method is applied to compare new solution against the current solution.

Based on the probability which is related to fitness, onlooker bee selects a food source. If a food source cannot be improved through a predetermined cycles, called "limit", it is removed from the population, and the employed bee of that food source becomes scout. The scout bee finds a new random food source position using Eq. (9)

$$x_i^j = x_{min}^j + rand[0,1](x_{max}^j - x_{min}^j) \quad (9)$$

Where x_{min}^j and x_{max}^j are lower and upper bounds of parameter j, respectively.

D. Ant colony optimization (ACO)

Ant colony optimization was introduced by Dorigo in the early 1990s. It is stimulated by the nature of actual ants and provides a solution for hard combinatorial optimization problems. The ACO has been with success applied to improvement issues like data processing, telecommunications networks, vehicle routing [39, 40, 41]. Ants use an aromatic material (pheromone) for indirect communication. Ants lay some secretion to mark the trail, as soon as supply of food is found. The number of the ordered secretion depends upon the gap, amount and quality of the food supply.

While an ant moves at arbitrary discovers a laid pheromone, it is likely that it will choose to tail its way. This ant itself lays a certain measure of pheromone, and subsequently authorizes the pheromone trail of that particular way [15]. Appropriately, the way that has been utilized by more ants will be more appealing to take after. As such, the likelihood with which a ground dwelling insect picks way increments with the quantity of ants that at one time picked the same way. This methodology is thus described by a positive feedback loop [38]. But the disadvantage of ACO is it falls easily into local optima and thereby premature convergence.

IV. Need for Optimization of Feature Subset

Feature selection is a vital essential for classification. It is a methodology of extracting the numerous and useful options from the dataset by uprooting the repetitive, insignificant and boisterous options. thus feature choice has became a significant venture in various pattern classification issues. it's connected to settle on a set of options, from a far larger set, specified they selected set is adequate to perform the arrangement enterprise.

Generally, feature optimization is a methodology of searching for the optimal answers for a specific issue of investment, and this pursuit procedure can be completed utilizing different executors which basically structure an arrangement of developing operators. This framework can advance by emphases as per a set of standards or scientific mathematical statements. Thus, such a framework will demonstrate some eminent attributes, prompting sorting toward oneself out states which compare to some optima of the destination scene. Once the self composed states are arrived at, we say the framework meets. In this way, to plan a proficient enhancement calculation is proportional to impersonating the development of an orchestrating toward oneself framework.

Genetic Algorithms (GAs) are stochastic systems, in view of the arbitrary determination of an introductory populace, which may be utilized to take care of pursuit and advancement issues. They are focused around the hereditary methodologies of natural organic entities. Over numerous eras, regular populaces develop as indicated by the standards of characteristic choice and "survival of the fittest". By impersonating this procedure, hereditary calculations have the capacity "advance" answers for true issues, on the off chance that they have been suitably encoded. The essential standards of GAs were first set down thoroughly via Holland.

V. Proposed Algorithm

In this section multi objective hybrid ant- bee colony optimization algorithm is described along with its implementation for automatic speaker verification. A two folded design is chosen for these multiple objectives; to optimize the feature set and to minimize the equal error rate.

A. Initialization

Multi-objective Optimization using Hybrid Atnt-Bee Colony Optimization for Feature Selection is the proposed algorithm is described below. A fully connected graph with each node representing a feature is constructed. Graph is fully connected to prevent deadlocks. Population of the ants will be same size of features. Initial pheromone value $T_{ij}(0)$ is set to 1. Randomly assign an onlooker ant to each feature. Fitness of each feature is determined and memorized. Maximum number of iterations is set to 500.

MOHABC Algorithm:

1. Create construction graph and determine the population of ants
2. Initialization
   2.1. Pheromone value
   2.2. Food source positions (solutions) (x_i, i=1, . . . , SN)
   2.3. Termination condition
   2.4. Size of External Record (ER)
3. Calculate the fitness of features (fit_i)
4. Based on non-domination the initialized solutions are sorted. The External Record(ER) is initialized with the sorted non-dominated solutions.
5. Repeat //Onlooker Ants' Phase
   5.1. For each onlooker ant
      5.1.1. Construct the solution for each onlooker ant using random proportional rule
      5.1.2. Feature subset's (fit_i) fitness is calculated.
      5.1.3. Probability of feature subset (P_f) is evaluated.
      5.1.4. Local pheromone updating
      5.1.5. To select the solution set which qualify to enter ER, Greedy Selection method is applied
   5.2. End For
   5.3. Evaluate the selected subset in ER using the chosen classification algorithm
      5.3.1. Based on classification result EER of every feature subset is calculated.
      5.3.2. Based on their EER sort the subsets
      5.3.3. Remember the feature subset with minimum EER as best-so-far trip

5.3.4. Global pheromone updation at nodes
5.4. Based on non domination the solutions are sorted in the ER
5.5. All the non domination solutions are perceived in the ER
5.6. If the number of non dominated solutions go beyond the allocated size of ER
5.6.1. Crowded members are removed by Crowding distance algorithm.
5.7. Increment number of iterations.
6. Until termination condition

B. Onlooker Ant Traversal

Each onlooker ant traverse based on the heuristic appeal and node pheromone levels which is expressed in the form of probabilistic transition rule (Eq.10) [33] [35].

$$p_{ij}^{k}(t) = \frac{[T_{ij}]^{\alpha}[\eta_{ij}]^{\beta}}{\sum_{l \epsilon N_i^k}[T_{il}]^{\alpha}[\eta_{il}]^{\beta}} \ ; if\ j\ \epsilon N_i^k \qquad (10)$$

Where k is the k^th onlooker ant at node i, T_{ij} is feasible neighbor of ant k at node i, η_{ij} is the amount of pheromone on an edge, $\eta_{ij} = \frac{1}{d_{ij}}$; d is the Euclidean distance between node i and node j as length of the edge and α and β two parameters determining the relative influence of T and η.

C. Fitness Evaluation

The basic parameters of an onlooker ant are fitness and probability, which are evaluated as follows. Fitness of feature subsets (11) [34]:

$$fit_i = \frac{1}{1+f_i}; \text{ if } f_i \geq 0 \text{ and}$$
$$= 1 + abs(f_i); \text{ if } f_i < 0 \qquad (11)$$

Probability of feature subset (P_f) (12)

$$P_f = \frac{fit_i}{\sum_{i=1}^{m} fit_i} \qquad (12)$$

Where fit_i fitness of the ith feature subset and m is is total number of onlooker ants.

D. Local Pheromone Updating

Local updating of pheromone is performed as per the local updating rule in ant colony system which is given as Eq.13.

$$T(i,j) = (1-\rho)T(i,j) + \rho T(0) \qquad (13)$$

Where ρ is pheromone trail decay coefficient or evaporation rate [ρϵ(0,1)] and it is taken as 0.2 and T(0) = $\frac{1}{nL_{nn}}$; where L_{nn} is trip length by nearest neighbor heuristic and ‘n’ is the number of features.

E. Subsequent Node Selection

Greedy selection method applied to decide which solution enters external record. Procedure of greedy method is to select between the paths constructed by onlooker ant. If new solution $x_{new} > x_i$, onlooker ant will inform to feature subset consisting of the feature it has been antecedently pointing and newly selected feature. If $x_{new} < x_i$, onlooker ant feature will be preserved and the newly selected feature is abandoned. A uniform distribution on the unit interval is used to generate a random number if $x_{new} = x_i$. x_i is replaced by x_{new} if the randomly generated value is less than 0.5.

F. Evaluation of Optimized Feature Subset

Evaluate equal error rate (EER) based on classification result of each subset of features. Sort the subsets based on their EER. Memorize the minimum EER and store the corresponding feature subset as best-so-far trip.

The above steps are repeated through a fixed number of iterations (T_{max}), or until a execution measure is satisfied.

G. Global Pheromone Updating

Then global pheromone updating is given by (14)

$$T(i,j) = (1-\delta)T(i,j) + \delta\Delta T(i,j) \qquad (14)$$

Where Δ T (i, j) is equal to 1/L; if (i, j) ϵ global best route and 0 otherwise. The constant δ is initialized to 0.8.

Pareto approach for non-domination is used to sort the solutions in the ER [37]. The non domination solutions are stored in ER.

H. Crowding Distance

In no dominated sorting area, after cycle, the arrangement inside the PL is sorted focused around no domination, and we keep the no domination arrangements of them staying inside the PL. In the event that the measure of no dominated arrangements surpasses the allocated size of PL, we have a tendency to utilize crowding distance to dispose of the excess features. Typically, the edge of the cuboid molded by exploitation the closest neighbors are called crowding distance. Before next iteration, the distance (uncertainty amplitude) and heading (converging rate) of every bee is bended towards its most empowering position. The significant parameters that are considered are convergence rate, β and learning rate γ. Convergence rate is gradually decreased to 0, before next iteration. Convergence rate is evaluated by the expression (15)

$$\beta = \beta_{max} - \left[\left\{\frac{\beta_{max} - \beta_{\min}}{T_{max}}\right\} \times T\right] \qquad (15)$$

where β_{max} is the initial value of the convergence rate, $\beta_{\min}$ is the ultimate convergence rate, T is the present iteration and T_{max} is the maximum number of iterations. Similarly learning rate γ is also attuned dynamically.

VI. Experimental Results

The following section gives the details about the data sets used, parameters and their initial values, classifier used and the results are shown in a tabular form.

A. Dataset

Two dissimilar datasets are used for this experimentation. One of them is BERLIN dataset which contains 535 sentences which includes male and female voices. Second one is telephone conversation dataset. It includes 48 male and 35 female voices. This telephone conversation data set consists of numerical data. Same dataset has been used for training as well as testing. In our experiment we have embedded 5db and 10db of noise to calculate the Equal Error Rate (EER) of the system.

B. Initialization of Parameters

The parameters used for the proposed work are

- Number of features, N=39
- Number of iterations, T_{max} =1000
- Convergence rate, β=1
- Learning rate, γ=1

C. Classification

The optimized feature vector thus obtained is send to Gaussian Mixture Model classifier for speaker verification. Gaussian Mixture Model (GMM) assumes the data to be ranging from the higher dimension having the ranges from -∞ to +∞ and the shape of the curve generated from the emotional speech samples are considered to be in bell shape distribution. This helps to model the high dimensional data. The PDF is given by (16) for the GMM

$$\mathrm{f(x)} = \frac{1}{\sqrt{2\pi\sigma}} \mathrm{e}^{(\frac{x-\mu}{\sigma})^2} \;;\; -\alpha < x < +\alpha \qquad (16)$$

Where "x" is considered to be MFCC values obtained for each speech signal, μ is the mean of each speech sample, σ is the variance.

With the component mean vector μ_k, and the diagonal covariance matrix Σ_k, Gaussian distribution modeled contingent likelihood p(f |k). For a given speech signal, expectation-maximization is used to obtain GMM based on an iterative process using a set of feature vectors. The successive likelihood of the features is maximized over all GMM densities for a speech signal. Log likelihood of an utterance, F = {f_1, f_2, . . . , f_T }, for speaker 'e' with a sequence of feature vector and GMM density model γ_e is given as in Eq. (17),

$$\rho\,\gamma_n = \log p(F \mid \gamma_e) = \sum_{t=1}^{T} \log p(\mathrm{f_t} \mid \gamma_\mathrm{e}) \qquad (17)$$

Where, p (F|sγ_e) is the GMM probability density for the speaker.

Then, the GMM density that maximizes posterior probability of the utterance is set as the verified speaker, which is given by Eq. (18).

$$\varepsilon = \arg\max\ \rho\,\gamma_n \qquad (18)$$

Where, ε is the result of verification.

VII. Results And Discussion

As the number of iterations increases, convergence rate reduced to 0.3 and there by learning rate reduced to 0.3. Detection cost function (DCF) is used for assessment, defined as (Reynolds & Rose, 1995) Eq. (19):

$$DCF = C_{miss}.FRR.P_{target} + C_{FA}.FAR\,.\left(1 - P_{target}\right) \qquad (19)$$

where P_{target} is the priori probability of objective test, FRR is false rejection rate and FAR is false acceptance rate, at a working point, the definite cost factors are C_{miss} and C_{FA}. Detection error tradeoff (DET) curve is also used for the assessment of the proposed system which shows the tradeoff between false rejection (FR) and false alarm (FA). Normally equal error rate (EER), which is the point on the curve where FA = FR, is chosen as assessment measure.

The routine measure is DCF and EER with a fixed threshold approach. The basic parameters used for MOHABC algorithm are shown in Table I. Experimental results of EER and DCF for different number of Gaussian (32 and 64) is shown in Table II.

DET curves for MOHABC based results for Berlin dataset, telephone conversation with 5db and 10db for 32 and 64 Gaussian are generated, shown in Fig. 2-3. The best results given by the proposed algorithm are with 32 Gaussian. The proposed algorithm has prevailing ability of steady search routine that approach most favorable solution by optimizing the multi objectives of the systems in parallel. The running time of any algorithm will be affected by the number of features in the feature subset, and the size of dataset. Working of the proposed system is shown in Fig. 4-7.

VIII. Conclusion

The research work presented in this paper focus on minimization of multiple objectives of speaker verification by optimizing the feature set which reduces the computational time and EER. When compared with the results of the whole feature set, the proposed optimized feature set gave superior accuracy rates. The work was initially focused on speech acquisition, Spectrogram analysis, Normalization, Features Extraction and Mapping using GMM. Moreover additionally multi objective optimization was focused to further reduce the complexity of the system.

As the real time data will contain noise embedded in it. So the proposed algorithm is tested with 5db, 10db noise embedded speech signals. It showed a better performance when compared to the existing systems. Hybrid of Ant Colony and Artificial Bee Colony optimized the feature subset over 85% thereby decreased the computational complexity of ASV. Additionally concept of Pareto dominance is used to preserve non-dominated results in an external record. An overall optimization of 87% is achieved thereby improved the recognition rate of ASV.

Table 1. Attribute Settings for MOHABC

Initial pheromone	α	β	γ	P_{target}	C_{FA}	C_{miss}
1	1	0.1	0.2	0.01	1	10

Table 2. For different number of Gaussians results of Speaker Verification

Number of Gaussians	EER			DCF		
	Berlin Dataset	Telephone Conversation		Berlin Dataset	Telephone Conversation	
		5db noise	10db noise		5db noise	10db noise
32	2.543	3.424	4.329	0.0235	0.0331	0.405
64	3.662	3.77	4.983	0.0296	0.0342	0.528

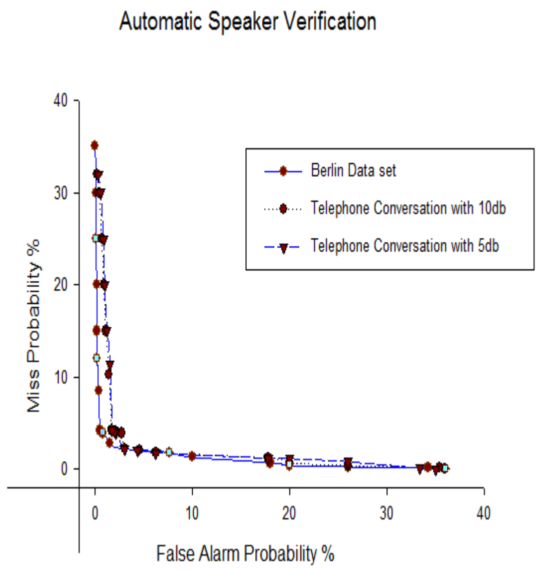

Fig. 2. DET curve with 32 Gaussians

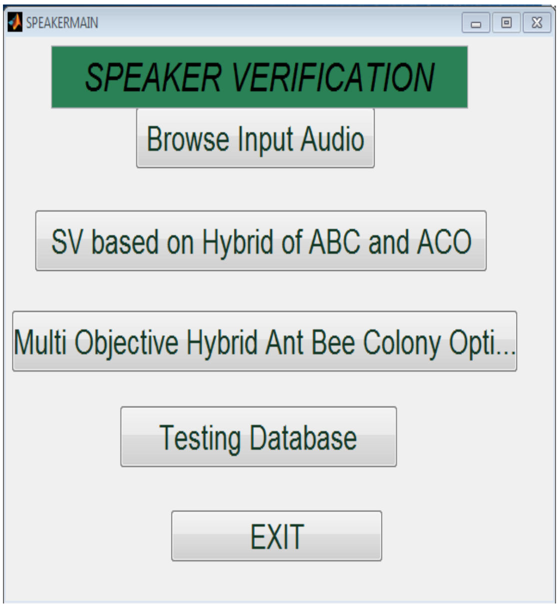

Fig. 4. Basic menu for ASV

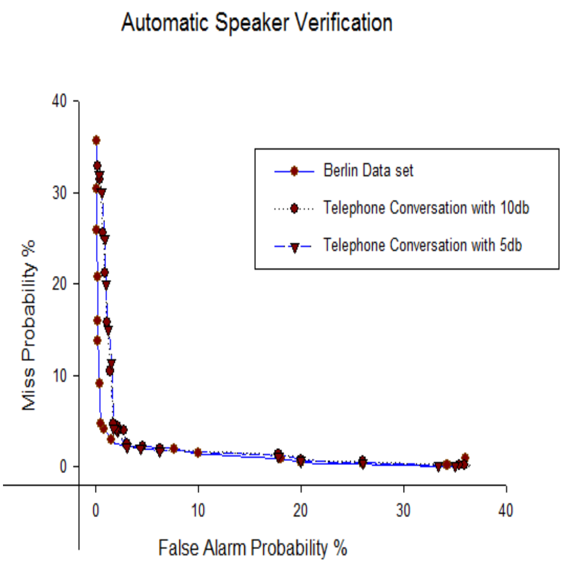

Fig. 3. DET curve with 64 Gaussians

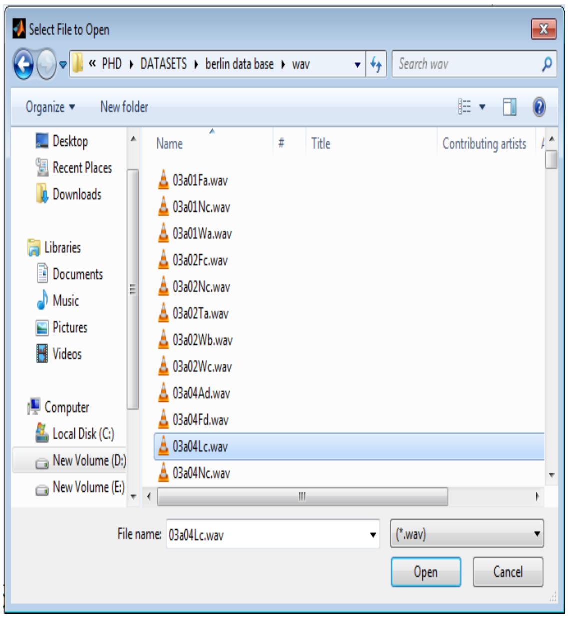

Fig. 5. Selection of Speech signal for verification

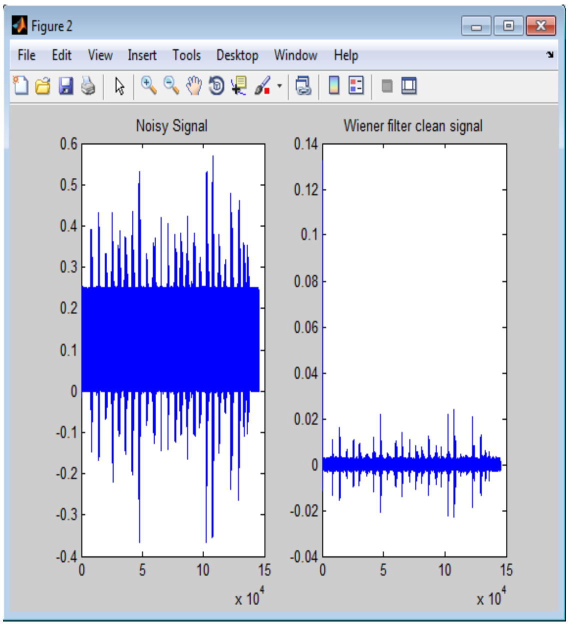

Fig. 6. Wiener Filtered Speech Signal

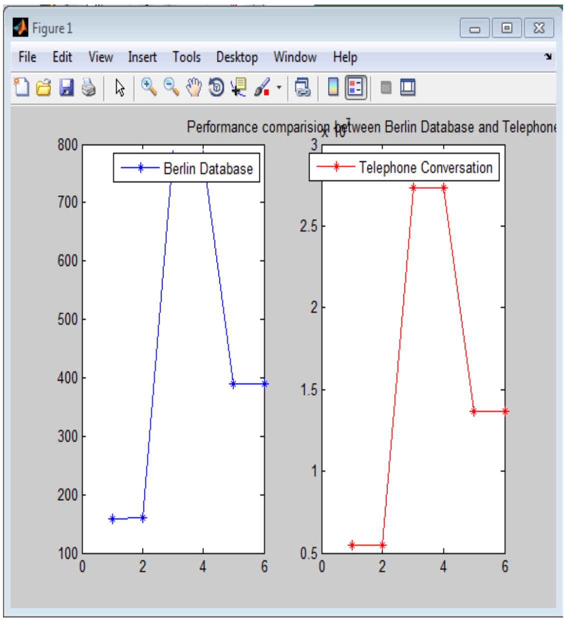

Fig. 7. Comparison of Hybrid Ant-Bee based ASV with MOHABC based ASV

REFERENCES

[1] Gish, H., Schmidt, M. *"Text-independent speaker recognition"*. IEEE Signal Process. Magazine (October), 18–32, 1994.

[2] Shahla Nemati, *"Text-dependent speaker verification using ant colony optimization-based selected features"*, Expert Systems with Applications 38 (2011) 620–630.

[3] Hazem Ahmed, "Swarm Intelligence: Concept, Models and Applications", *Technical Report 2012-585.*

[4] Sandeep Kumar *"Improved Onlooker Bee Phase in Artificial Bee Colony Algorithm"*, International Journal of Computer Applications (0975 – 8887) Volume 90 – No 6, March 2014.

[5] Swarm Intelligence and Bio-Inspired Computation Theory and Applications, First Edition 2013, Elsevier Inc.

[6] Monica Sood, "*Speaker Recognition Based On Cuckoo Search Algorithm*", International Journal of Innovative Technology and Exploring Engineering (IJITEE), April 2013.

[7] Nishant Pathak, "Travelling *Salesman Problem Using Bee Colony With SPV*", International Journal of Soft Computing and Engineering (IJSCE), July 2012.

[8] Falko Dressler, "*A survey on bio-inspired networking*", Computer Networks, Volume 54, Issue 6, 29 April 2010, Pages 881-900.

[9] Dervis Karaboga, "*A comparative study of Artificial Bee Colony algorithm*", Applied Mathematics and Computation 214 (2009) 108–132.

[10] Dervis Karaboga, "*A novel clustering approach: Artificial Bee Colony (ABC) algorithm*", Applied Soft Computing 11 (2011) 652–657.

[11] Frank Neumann," *Ant Colony Optimization and the minimum spanning tree problem*", Theoretical Computer Science 411 (2010) 2406_2413.

[12] Yoon-Teck Bau, "*Ant Colony Optimization Approaches to the Degree-constrained Minimum Spanning Tree Problem*", Journal of Information Science and Engineering 24, 1081-1094 (2008).

[13] Shunmugapriya Palanisamy, "*Artificial Bee Colony Approach for Optimizing Feature Selection*", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 3, May 2012.

[14] Song Zheng, "*Ant Colony Optimization based on Pheromone Trail Centralization*", Proceedings of the 6th World Congress on Intelligent Control and Automation, June 21 - 23, 2006, Dalian, China.

[15] Ahmed Al-Ani, "*Ant Colony Optimization for Feature Subset Selection*", proceedings of world academy of science, engineering and technology volume 4 february 2005 ISSN 1307-6884.

[16] Taufiq Hasan, John H. L. Hansen, "*Acoustic Factor Analysis for Robust Speaker Verification*", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 21, No. 4, April 2013.

[17] Khalid Saeed and Mohammad Kheir Nammous, "*A Speech-and-Speaker Identification System: Feature Extraction, Description, and Classification of Speech-Signal Image*", IEEE Transactions on Industrial Electronics Vol. 54, No.2, April 2007, pp. 887-897.

[18] Alejandro Bidondo, Shin-ichi Sato, Ezequiel Kinigsberg, Adri án Saavedra, Andr és Sabater, Agust ín Arias, Mariano Arouxet, and Ariel Groisman, "*Speaker recognition analysis using running autocorrelation function parameters*", POMA - ICA 2013 Montreal Volume 19, pp. 060036 (June 2013).

[19] J. Sirisha Devi, "Speaker Emotion Recognition Based on Speech Features and Classification Techniques", I.J. Computer Network and Information Security, 2014, 7, 61-77.

[20] Taufiq Hasan, Seyed Omid Sadjadi, Gang Liu, Navid Shokouhi, Hynek Boˇril, John H.L. Hansen," Acoustic Factor Analysis for Robust Speaker Verification" Audio, Speech, and Language Processing, IEEE Transactions on (Volume:21, Issue:4), 23 October 2012.

[21] Taufiq Hasan, Seyed Omid Sadjadi, Gang Liu, Navid Shokouhi, Hynek Boˇril, John H.L. Hansen," *CRSS SYSTEMS FOR 2012 NIST Speaker Recognition Evaluation*", ICASSP 2013.

[22] Gang Liu, Taufiq Hasan, Hynek Bořil, John H.L. Hansen," *An Investigation On Back-End For Speaker Recognition In Multi-Session Enrollment"*, ICASSP 2013.

[23] Balaji Vasan Srinivasan, Yuancheng Luo, Daniel Garcia-Romero, Dmitry N. Zotkin, and Ramani Duraiswami," *A Symmetric Kernel Partial Least Squares Framework for Speaker Recognition*", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 21, No. 7, July 2013.

[24] Tomi Kinnunen, Rahim Saeidi, Filip Sedlák, Kong Aik Lee, Johan Sandberg, Maria Hansson-Sandsten, Haizhou Li," *Low-Variance Multitaper MFCC Features: A Case Study in Robust Speaker Verification*", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 20, No. 7, September 2012.

[25] Khalid Saeed and Mohammad Kheir Nammous, "*A Speech-and-Speaker Identification System: Feature Extraction, Description, and Classification of Speech-Signal Image*", IEEE Transactions on Industrial Electronics Vol. 54, No.2, April 2007, pp. 887-897.

[26] J. Sirisha Devi," Automatic Speech Emotion and Speaker Recognition based on Hybrid GMM and FFBNN", International Journal on Computational Sciences & Applications (IJCSA) Vol.4, No.1, February 2014.

[27] Nitisha and Ashu Bansal, "*Speaker Recognition Using MFCC Front End Analysis and VQ Modelling Technique for Hindi Words using MATLAB*", Hindu College of Engineering, Haryana, India.

[28] Alejandro Bidondo, Shin-ichi Sato, Ezequiel Kinigsberg, Adrián Saavedra, Andrés Sabater, Agustín Arias, Mariano Arouxet, and Ariel Groisman, "*Speaker recognition analysis using running autocorrelation function parameters*", POMA - ICA 2013 Montreal Volume 19, pp. 060036 (June 2013).

[29] Tobias May, Steven van de Par, and Armin Kohlrausch," *Noise-Robust Speaker Recognition Combining Missing Data Techniques and Universal Background Modeling*", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 20, No. 1, January 2012.

[30] Wen Wang, Andreas Kathol, Harry Bratt, "*Automatic Detection of Speaker Attributes Based in Utterance Text*", *INTERSPEECH, page 2361-2364. ISCA, (2011).*

[31] Mohamed Abdel Fattah, "*Speaker recognition for Wire/Wireless Communication Systems*", The International Arab Journal of Information Technology, Vol.3, No.1, January 2006.

[32] N. Murali Krishna, P.V. Lakshmi, Y. Srinivas, J.Sirisha Devi, "Emotion Recognition using Dynamic Time Warping Technique for Isolated Words", IJCSI International Journal Of Computer Science Issues, Vol. 8, Issue 5, No 1, September 2011.

[33] C. Garcı´a-Martı´nez, "*A taxonomy and an empirical analysis of multiple objective ant colony optimization algorithms for the bi-criteria TSP*", European Journal of Operational Research 180 (2007) 116–148.

[34] Wenping Zou, "*Solving Multiobjective Optimization Problems Using Artificial Bee Colony Algorithm*", Hindawi Publishing Corporation Discrete Dynamics in Nature and Society Volume 2011.

[35] Guillermo Leguizamón, "*Multi-Objective Ant Colony Optimization: A Taxonomy and Review of Approaches"*, World Scientific Review Volume - 9in x 6in, May 25, 2010.

[36] Abdolreza Rashno, "*Highly Efficient Dimension Reduction for Text-Independent Speaker Verification Based on Relieff Algorithm and Support Vector Machines"*, International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 6, No. 1, February, 2013.

[37] Ahmed Elhossini, "*Strength Pareto Particle Swarm Optimization and Hybrid EA-PSO forMulti-Objective Optimization"*, 2010 by the Massachusetts Institute of Technology Evolutionary Computation 18(1): 127–156.

[38] M. Dorigo, "Ant System: Optimization by a colony of cooperating agents". *IEEE Transactions on Systems, Man, and Cybernetics* – Part B, 26:29–41, 1996.

[39] G. Di Caro and M. Dorigo. "AntNet: Distributed stigmergetic control for communications networks". *Journal of Artificial Intelligence Research*, 9:317–365, 1998.

[40] R.S. Parpinelli; H.S. Lopes; A.A. Freitas, "Data mining with an ant colony optimization algorithm", *IEEE Transactions on Evolutionary Computation*, 6: 321 – 332, 2002.

[41] R. Montemanni, L.M. Gambardella, A.E. Rizzoli and A.V. Donati. "A new algorithm for a Dynamic Vehicle Routing Problem based on Ant Colony System". *Proceedings of ODYSSEUS 2003*, 27-30, 2003.

Authors' Profiles

Dr. Srinivas Yarramalle was awarded M. Tech (Computer Science & Technology) from Andhra University -1999. He was awarded Doctorate in Computer Science with Specialization in Image Processing from Acharya Nagarjuna University, Guntur.-21-1-2008. He received Best Teacher Award from JNTU University, Sept-5-2010 and from B.C. Corporation in 2005. He prepared 4 Monographs for School of Correspondence Courses, Andhra University. He received SASTRA Award from Vignan's Institute of Information Technology, 10-Jan-2008, for Research publications. Received SASTRA Award from Vignan's Institute of Information Technology, 10-Jan- 2009, for Research publications.

How to cite this paper: J. Sirisha Devi, Srinivas Yarramalle,"Multi Objective Optimization Problem resolution based on Hybrid Ant-Bee Colony for Text Independent Speaker Verification", IJMECS, vol.7, no.1, pp.55-63, 2015.DOI: 10.5815/ijmecs.2015.01.08