Speaker Recognition Based on Cuckoo Search Algorithm

Monica Sood, Gurline Kaur

Abstract- Today's world sees a lot of changes being done. These are a result of some modification or some innovation. This research is being done in the field of Swarm Intelligence or SI. It deals with studying the behaviour of organisms or swarms. Swarms are individual entities which are working on their own, yet their combined or aggregate behaviour yields some great results. This begins by studying the behaviour of any organism as fish, ants, bees, cuckoo bird or something like water drops. When the behaviour is understood it is then converted in the form of an algorithm. It is this algorithm that is of utmost importance; it not only studies the behaviour of these organisms but also provides some principles which can help in providing solutions to real world applications. This research is based on an algorithm of Swarm Intelligence called Cuckoo Search. This is an algorithm which is aimed at understanding the breeding behaviour of the cuckoo bird. In this research, it is applied in the field of Biometrics. Biometrics is used to identify an individual as per their some special characteristics as finger print, voice, iris, handwriting, typing speed. In this Cuckoo Search has been applied on Speaker Recognition systems and voice. Thus by applying this algorithm, the process of Speaker Recognition is optimized by a fitness function by matching of voices being done on only the extracted optimized features produced by the Cuckoo Search algorithm.

Keywords- Correlation, mean, Fitness Function, Swarm Intelligence.

I. INTRODUCTION

This dissertation is based on CS or cuckoo search algorithm. Some applications of Swarm Intelligence include job scheduling [1], biometrics, text document classification and so on. Cuckoo Search algorithm studies the brood parasitism of the cuckoo bird. This behavior is then converted to an algorithm which can be used to solve actual real life problems. Cuckoo Search has its applications in optimization problems. This algorithm also works in a similar way. It studies the behavior of cuckoo bird. This bird lays egg in some host bird's [2] nest. The host bird will take care of these eggs; some eggs can even mimic the color [3] of the host bird's eggs and can also produce sounds similar to its chicks in order to dupe the host bird. These cuckoo eggs are taken care and are passed on to the next generations. In real life this algorithm can be implemented in optimization problems as feature reduction. Speaker recognition falls under the category of Voice Recognition. The input voice sample will be matched with the extracted features of pre stored voices from the database. For this the important features are extracted or are optimized, a feature subset is created and the speaker is matched as per these features.

Manuscript received on April, 2013.

Monica Sood, Assistant Professor, Department of CSE/IT, Lovely Professional University, Phagwara, India.

Gurline Kaur, M. Tech. CSE, Student, Department of CSE/IT, Lovely Professional University, Phagwara, India.

II. METHODOLOGY

A. Rules of Cuckoo Search Algorithm

Each cuckoo will lay 1 egg at a time and dump it in any next, chosen at random.

1) Best nest with high probability will be passed on to the next generations.
2) Available hosts are fixed, egg is discovered with a probability of (0, 1), it will either be accepted by the host birds or thrown out of its nest. In other words, the worst discovered nests are dumped.
3) Best nest with high probability will be passed on to the next generations.
4) Available hosts are fixed, egg is discovered with a probability of (0, 1), it will either be accepted by the host birds or thrown out of its nest. In other words, the worst discovered nests are dumped.
5) New cuckoo or a new solution is generated.
6) In each iteration, solution is compared and only the best solution is passed further. This is done with a "fitness function".

These steps can be matched any optimization problem. Each cuckoo egg and cuckoo nest plays an important role in such problems. Therefore as per this algorithm:

Egg in a Nest → Solution or the Available Voice Samples

Cuckoo Egg → New Solution or the Voice Sample to be matched

Each Nest → One Egg, One Solution or Features of Voice Samples

B. Steps of Speaker Recognition using Cuckoo Search Algorithm

The following steps will elaborate how a speaker is verified using Cuckoo Search algorithm.

1) Voice Recording

In this the speaker speaks and his voice is recorded for either feature matching or feature extraction. This voice sample if matches with the already existing voice samples, then the speaker will be authenticated.

2) Feature Extraction

As Cuckoo Search will work on only few best features, so there is a need to initially extract the features from the voices. In this step the features will be extracted from the inputted voice. This voice pattern will be in the form of spectrograms consisting of various frequencies as per time. For feature extraction,

the unwanted signal as voice or other disturbances have to be removed as these will create hindrance in the next process. Hence, for this purpose a threshold value is used which will extract only the voice signal and therefore neglecting the other induce unwanted signal. The following formula is used:

$$i>0.3 \,||\, i<-0.3 \qquad (1)$$

where

i = voice sample or values in it

The signal values in this range are selected. The optimized or the best features will be fetched from these values.

Cuckoo Search will reduce the number of these extracted features to get the best available features from this voice. The best features are selected as per a Fitness Function.

This Fitness Function is based on mean. From every voice the mean is calculated and it is subtracted from each value of that sample. This will fetch the required best values from each voice sample. The required features thus extracted are stored in the database and will be used for speaker recognition. Hence, this Fitness Function is providing the required optimization. The following formula is used:

$$FF = a_i-(\Sigma^{N}{}_{i=1}\, a_i/N) \qquad (2)$$

where

FF = Fitness Function

a = Voice Signal

N= Total values in the voice sample

a_i/N = Mean of Voice

3) Pattern Matching

Till this step the system will not only be having the stored voice patterns, but also its extracted features. These extracted features will be matched with the inputted voice's features. For matching purpose, Correlation is used. The extracted features closest to the stored features will be the one that will be matched. To avoid the voice matching in all scenarios, even in case of un- authenticated speaker, a threshold value is used to increase security and to correctly authenticate or neglect a speaker. This threshold value specifies a likelihood ratio, which will signify the extent of match of speaker recognition

4) Decision

Then the voice will either be accepted or rejected. Acceptance means that the speaker is authenticated as the voice is matched otherwise it will be rejected. This decision is taken as per correlation as described above. Match above the threshold will be accepted. The matched voice will have a high correlation otherwise a low value below the threshold is neglected; hence the speaker is not allowed the access.

In this research text dependent speaker recognition is used, in which the enrolment and test security codes are same [4].

The following diagram explains the process of Text Independent Speaker Recognition using Cuckoo Search Algorithm.

C. Flowchart of Speaker Recognition using Cuckoo Search Algorithm

The following diagram clearly explains the process of speaker recognition. This begins by inputting a voice sample and following the steps as shown to recognize or authenticate a speaker as:

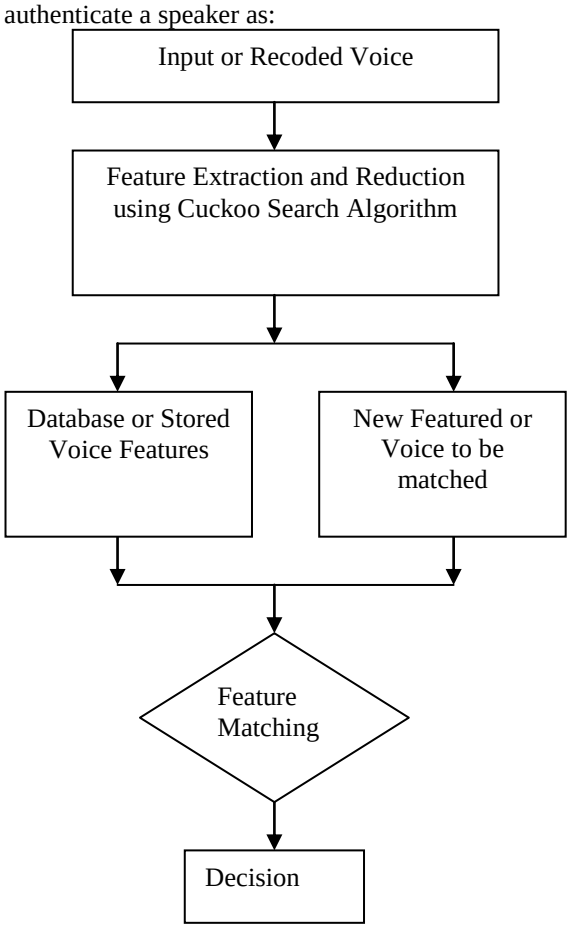

Fig. 1 Process of Speaker Recognition

D. Algorithm for Speaker Recognition using Cuckoo Search Algorithm

1) Input the voice for speaker recognition.

2) Extract features using threshold as described in the above section 2.2.

3) Generate host nests n.

n= number of solutions or features extracted from the voice pattern.

4) Repeat while (j<MG)

i = loop variable

MG = maximum generation

ck = number of cuckoos or feature subset

(i) for(i=0;i<ck;i++)

a) Get a cuckoo and move it to nest.

b) Choose a Fitness Function F

if (Fi>Fm)

Fm=Fi;

End for

(ii) Abandon the nest and build new ones or build new feature set.

(iii) Keep the remaining solutions.

(iv) Rank these as per F.

(v) Pass these to next generations or iterations.

5) Get the nest with best solution as per F.

6) This is the result or the best optimized features.

7) This optimized feature subset will be matched with the already stored best features and as per the correlation and threshold a speaker will be authenticated or recognized.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Hence with these steps cuckoo search can recognize a speaker.

III. RESULTS

The system was tested for stored voices and then for voice input in real time. After the implementation it is very clear that a speaker is recognized only if his voice sample is already present in the database. As per the following snapshot, the maximum value is 9.54, so the voice will be matched with the third voice sample.

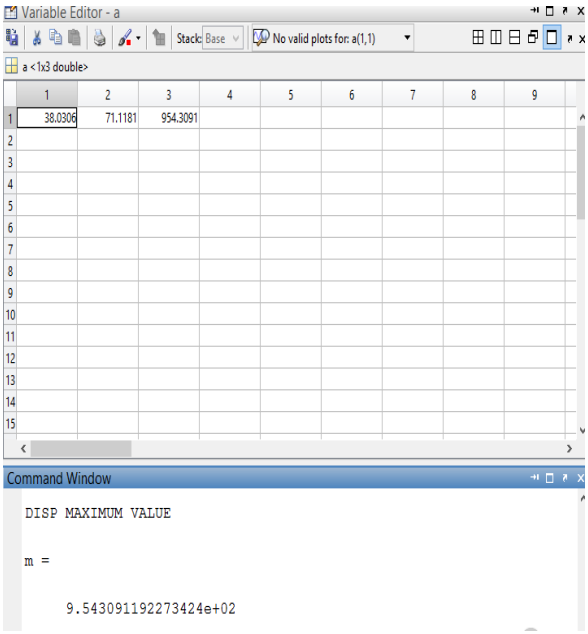

Figure 2 Result of Voice Matching

IV. CONCLUSION

In this paper speaker recognition is optimized using a Swarm Intelligence algorithm, Cuckoo Search. This algorithm will aim at finding and short listing the features from voice which can uniquely identify it. After feature extraction using a threshold to remove the unwanted signal or disturbance and considering only the voice sample, a Fitness Function is applied based on mean of the individual sample, which will fetch few unique and best features and will discard the remaining. As a result, there is no need to match a speaker's voice through all the features. To increase the security a threshold value is added to correlation in the matching phase. As only optimized features are extracted, so this will not only optimize this technique but will also save resources.

REFERENCES

1. R. G. Babukartik and P. Dhavachelvan (2012) "Hybrid Algorithm using the advantage of ACO and Cuckoo Search for Job Scheduling", International Journal of Information Technology Convergence and Services (IJITCS) Vol.2, No. 4, August 2012.
2. Vipinkumar Tiwari, "Face Recognition Based on Cuckoo Search Algorithm", Indian Journal of Computer Science and Engineering (IJCSE), June-July 2012.
3. Xin-She Yang and Suash Deb, "Cuckoo Search via Levy Flights", In: Proc. Of World Congress son Nature & Biologically Inspired Computing (NaBIC 2009), December 2009, India. IEEE Publications, USA, pp. 210-214(2009).
4. H.B. Kekre, Vaishali Kulkarni, Prashant Gaikar and Nishant Gupta, "Speaker Identification using Spectrograms of Varying Frame Sizes", International Journal of Computer Applications (0975-8887) Volume 50- No. 20, July 2012.