ISSN: 0975-766X
CODEN: IJPTFI
Review Article
Available through Online
www.ijptonline.com

A SURVEY ON SWARM OPTIMIZATION TECHNIQUES FOR HEART DISEASE PREDICTION

Anbarasi M*, Saleem Durai M. A
School of Computer Science and Engineering, VIT University, Vellore, Tamilnadu, India
Email: manbarasi@vit.ac.in

Received on 20-05-2016 *Accepted on 15-06-2016*

Abstract

Heart syndrome is a most important reason of mortality and death for people in most of the countries all over the world. There are many factors of heart disease that affecting the structure or function of the heart. These might be challenging for doctors to predict heart disease accurately. Therefore, it is important to utilize automated innovations in coronary illness conclusion to help specialists to determine quicker to have higher precisely. Researchers have developed many techniques such as data mining, soft computing and optimization techniques for diagnosing heart disease. In this paper, we present a survey paper on the use of swarm intelligence for the detection of heart disease. This survey helps to understand the different swarm intelligence techniques that are involved in diagnosis for heart disease. This paper is relevant to swarm intelligent techniques which are involved in prediction of heart disease.

Keywords: Swarm intelligence, PSO,ACO, ABC, Firefly

I. Introduction

Heart disease has considerably increased for the last two decades and become the leading cause of death for people in most of the countries in the world. World Health Organization (WHO) reported that 30% of death is due to heart disease [1,2]. In 2008 17.3 million people died due to heart disease. More than 80% of passing away in world is because of coronary illness. WHO estimated by 2030 almost 23.6 million persons will pass away due to heart disease in [3]There are many features of heart disease affecting the function or structure of the heart. These might be difficult for the doctors to diagnosis effectively and accurately. So, it is necessary to computerized technologies in heart disease diagnosis to assist doctors to predict faster with higher accurately. In order to achieve goal to get quality social insurance the procedure of choice making is upgraded and supported by different frameworks such as expert system[4], decision support systems (DSS)[5], automatic diagnosis systems[6], Rule based support systems[7], diagnosis support system, medical record system, clinical information system, and hybrid intelligent system[8].

Several data mining, machine learning, artificial intelligence, pattern recognition, soft computing and optimization techniques have been proposed for heart disease diagnosis. Now a days there are many heart disease diagnosis systems depends on data mining, soft computing, optimization and hybrid approaches that have been proposed. In this paper we have discussed heart diagnosis prediction using swarm intelligent techniques.

Swarm optimization is based on the behavior of animals like birds, fish, honey bees etc. their characteristic wonder clarifies us the, communication and interaction with each entities, this minimizing so as to learn helps us to define an issue their aggregate expense. SI becomes vital research area for engineers, bioinformatics, computer scientists, operational researchers, economists, and many other disciplines. SI can solve the problems that the natural intelligent swarms can solve in several engineering areas of real world. Artificial bee colony is one of the most recently defined SI algorithm by Dervis araboga in 2005[9] motivated by the intelligent behavior of honey bees. Their natural sensation explains us the recognition, message and contact with each entities this information helps us to detail an issue by minimizing their overall expense. Firefly algorithm proposed by yang for global optimization problems[10] Fireflies emit light to attack the prey. The blinking light is expressed in such a way that it is related with the objective function to be optimized. In optimization, each fly is similar to a solution and the fly with the extreme power of brightness is considered as an optimal result. Pso[11] a population based algorithm has been useful typical data mining tasks such as clustering, classification, association analysis and regression. Cuckoo search recently developed by yang and deb in 2009, it is efficient in giving global optimization problems. This algorithm is enhanced by Lvy flights[12] rather than by simple isotropic random walks. [13] improved the algorithm by formulating a modified cuckoo search algorithm, [14] extended it to multiobjective optimization problem

II. Swarm Optimization Techniques

II.I Artificial bee colony(ABC)

Artificial Bee Colony (ABC) Optimization is used for solving various multidimensional optimization problems [15]. ABC algorithm plays intelligent performance of real honey bee colonies. The main advantage in ABC allows the results to converge to the optimal solution quickly and also it is simple and easy to implement because it has fewer control parameter to configure, it moderates the total cost and very effective in solving fuzzy problems.

This algorithm is iterative algorithm and consists of four levels, named as initialization level, employed bee level, onlooker bee level and scout bee level.

Initialization Level

REPEAT

Employed Bees Level

Onlooker Bees Level

Scout Bees Level

Memorize the best solution achieved so far

UNTIL (Cycle = Maximum Cycle Number or a Maximum CPU time)

In Initialization level it will find the quantity of employed bees (M), find the boundary value for the population, create food source positions, calculate the objective and fitness values of the solutions. In Employed bees level try to increase the food source which is feasible solution for the optimization problem and created in initialization of the algorithm. The employed bees silage food sources and move position information about them to the hive. Almost 50% of the bee population in a stockpile goes to this group. In onlooker bees level search all over the place the solutions of employed bees by seeing data shared by employed bees, estimate the food sources to choice a precise source with rich juice. In scout bee level, if an employed bee could not progress self-solution in a certain time or a limit which is determined for the population in initialization of the algorithm, this employed bee becomes a scout bee. After a novel solution is created for this scout bee, the scout bee turns into employed bee, again. The bee calculates the food source over the fitness rate of a certain food source is higher than the previous food source. The bee deletes the memory of the previous one and memorizes the existing one. These three steps are repeated until a termination criteria is satisfied, [16,] used hybrid algorithm ABC with KNN algorithm this could be an alternative way for heart disease prediction. [17]

Initially removed the redundant features using metaheuristic algorithm to get the optimal feature subset and then they have used ABC algorithm to identify the disease data and then evaluate the fitness of ABC using SVM approach. The results shows that the good classification accuracy with only seven features.[18] classify the attributes in to a tree structure, removing the fuzziness to get absolute rules for the prediction process and then applied optimization technique for coronary heart disease and achieved better predict results.

II.II Ant colony optimization (ACO)

ACO studies artificial structures that take motivation from the activities of real ant colonies and which are used to resolve discrete optimization issues. ACO was firsly proposed by Dorigo et al [19,20]. The inspiration of the ACO is the reflection of real ants capable to find the direct path from a source of food to their nest. [21] first they have used data mining concept to find the support and then considered support as a weight of symptom which will be the initial pheromone value of ant using these methods they increase detection rate in the early stage.

The key parameters involved in ACO are: number of ants n; number of iterations p; exponents α and β; pheromone evaporation rate ρ; and pheromone reward factor.

The process starts by generating m random ants(solution).

An ant k(k=1,2,..n) denotes a result string, with a chose value for each variable.

An ant is calculated according to an objective function.

Consequently, pheromone focus related with each possible route(variable value) is changed in a way to reinforce good solutions as follows:

$\tau ij(a) = \rho\tau_{ij}(t-1) + \Delta\tau_{ij}; t = 1,2,\ldots..N$

Where N is the number of iterations (generation cycles); $\tau_{ij}(a)$is the revised focus of pheromone associated with option lij at iteration a, $\tau_{ij}(a-1)$ is the focus of pheromone at the previous iteration(a-1); $\Delta\tau_{ij}$ = change in pheromone focus; and ρ = pheromone evaporation rate (0-1). The reason for permitting pheromone vanishing is to avoid too strong influence of the old pheromone to avoid premature solution stagnation.

Pseudo code for Ant Colony Optimization

Begin:

Initialize the pheromone trials and factors;

Generate population of m1 solutions (ants);

For each individual ant k ∈ m1: calculate fitness (k);

For each ant decide its best position;

Determine the greatest global ant;

Update the pheromone trail;

Check if termination = true;

End;

Pheromone focus Δπij is calculated as

$$\Delta\pi ij = \sum_{k=1}^{m} \begin{cases} R / fitness_k & if \ \ option \ \ l_{ij} \ is \ chosen \ by \ ant \ k \\ 0 & otherwise \end{cases}$$

Where R is a constant called the pheromone reward feature; and fitnessk is the value of the objective function that is calculated for ant k. It is well-known that the quantity of pheromone gets higher as the solution improves

Once the pheromone is updated then an iteration, the next iteration starts by changing the ants' paths (i.e. associated variable values) in a way that respects pheromone focus and also some heuristic first choice. As such, an ant k at iteration a will change the value for each variable according to the following probability

$$Pij(k,a) = \frac{[\pi_{ij}(a)]^{\alpha} \times [\eta_{ij}]^{\beta}}{\sum_{t_{ij}} [\pi_{ij}(a)]^{\alpha} \times [\eta_{ij}]^{\beta}}$$

Where Pij(k,a) = probability that option lij is chosen by ant k for variable I at iteration a;

πij(a) = pheromone focus associated with option lij at iteration a;

ηij =heuristic factor for selecting among existing choices and is an indicator of how good it is for ant k to select option lik.

α and β are exponent restrictions that control the relative importance of pheromone focus versus the heuristic factor.

Ants go through the food while laying down pheromone trails. Shortest path is learned via pheromone trails in the following ways: each ant moves at random, pheromone is deposited on path, shorter path, more pheromone rails (positive feedback sys) and ants follow the intense pheromone trails)

II.III Firefly algorithm (FA)

FA is one of the recent swarm intelligent methods proposed by Yang [22] and is a meta-heuristic algorithm that depends on blinking behavior of fireflies in nature to find global optimal solution in search space for special problems. FA is actually greater in solving noisy non-linear optimization problems. The FA looks like to be a good optimization tool in part due to the outcome of the attractiveness function which is an exclusive to the firefly behavior.

Flashing features of fireflies is used to develop firefly-inspired algorithm. Firefly Algorithm [23,24] use the following three idealized rules: •due All the fireflies are unisex so it means that one firefly is paying attention to other fireflies regardless of their sex.

• Attractiveness and brightness are comparative to each other, so for any two blinking fireflies, the less bright one will move near the one which is brighter. Attractiveness and illumination both decrease as their distance increases. If there is no one brighter than other firefly, it will move randomly.

• The illumination of a firefly is determined by the assessment of the objective function. For a maximization problem, the illumination is merely proportional to the value of the objective function. Other forms of the illumination could be defined in a same way to the fitness function in genetic algorithms.

Pseudo code of the firefly algorithm

Objective function f(x), x = (y1, ..., yd)T

Generate initial population of fireflies ym (m = 1, 2, ..., n)

Light intensity Im at ym is determined by f(ym)

Define light absorption coefficient γ

while (t <MaxGeneration)

for m = 1 : n all n fireflies

for p = 1 : i all n fireflies

if (Ip > Im), Move firefly m near p in d-dimension; end if

Attractiveness differs with distance r via exp[−γr]

Estimate new solutions and update light intensity

end for p

end for m

Rank the fireflies and discover the current best

end while

Postprocess results and visualization

FA has been effectively applied to a huge amount of challenging combinational optimization problems as well as NP-hard problems In addition, the firefly algorithm is greatly further effective in finding the global optima with higher success rates than genetic algorithms (GAs) and PSO [25]. Firefly algorithm is an effective and successful in optimizing continuous optimization problems [26] as well as discrete optimization problems.[27] introduce chaos into FA so as to increase its global search mobility for robust global optimization results shows that its better than ordinary FA. The Firefly algorithm seems to perform better for higher levels of noise. Firefly algorithm has some disadvantage such as getting trapped into several local optima. Firefly algorithm does local search as well and sometimes is unable to completely get rid of them. Firefly algorithm factors are set fixed and they do not change with the time. In addition Firefly algorithm does not remember any history of better situation for each firefly and this causes them to move irrespective of its previous better situation, and they may end up missing their situations.

II. IV Particle swarm optimization (PSO)

PSO has been used to solve numerous optimization problems developed by Kennedy and Eberhart in 1995 [28]. A population based optimization method stimulated by social behavior of bird flocking or fish schooling, PSO consists of a swarm of particles. Each particle exists in at a position in the search space. The fitness of each particle represents the quality of its position. The particles fly over the search space with a certain velocity. The velocity of each particle is based on its own best position found so far and the best possible solution that was found so far by its neighbours.

The pseudo code of the procedure is as follows

```
For each particle
   Initialize particle
END
Do
   For each particle
      Calculate fitness value
      If the fitness value is better than the best fitness value (pBest) in history
         set current value as the new pBest
   End
   Choose the particle with the best fitness value of all the particles as the gBest
   For each particle
      Estimate particle velocity according equation (a)
      Update particle position according equation (b)
   End
While when maximum number of iterations reached or lowest error conditions y is not achieved
```

When there was no noise on the procedure produces, the performance of both algorithms PSO and FA seems to be no different to method to the optimum.

Also Firefly is better than PSO in terms of the time taken for the optimum or near optimum value to be generated provided certain high level of noise where the difference in time taken becomes more evident with the increase in the level of noise.[25] specified the comparison about the performance of GA, PSO and FA in terms of efficiency and success rate.

II.V Cuckoo Search (CS)

CS algorithm is a novel meta-heuristic method proposed by XinShe Yang [29-34]. This algorithm was inspired by the make brood parasitism of some cuckoo species by placing their eggs in the nests of other host birds. Some host nest can keep direct difference. If an egg is discovered by the host bird and not its own, it will either throw the unknown egg away or simply abandon its nest and build a new nest elsewhere. CS can explore the search space more powerfully than algorithms by standard Gaussian procedure. This benefit combined with both local and search capabilities and guaranteed global convergence, makes cuckoo search very efficiently. The CS algorithm follows three idealized rules: a. Each cuckoo lays one egg at a time, and put its egg in arbitrarily selected nest; b. The finest nests with high quality of eggs will carry over to the next generations; c. The amount of existing host nests is fixed, and the egg placed by a cuckoo is discovered by the host bird with a likelihood. In this instance, the host bird can either throw the egg away or abandon the nest, and build an entirely new nest. The rule-c defined above can be approximated by the fraction probability of the n nests that are replaced by new nests (with new random solutions).

Psedo code for cuckoo algorithm

Objective Function:

f(y), Y $=(y_1, y_2, y_3......y_d)$

Create a first population of m host nests;

While(t<MaxCreation) or (stop criterion)

Get a cuckoo's nest i randomly and replace its solution by performing Levy flights;

Calculate its quality/fitness Fi

Select a nest j among n randomly;

if (Fi >Fj)

Replace j by the new solution;

End if

Fraction (Pa) of the not good nests is abandoned and new ones are built;

Keep the best solutions;

Rank the solutions and discover the current best;

Pass the current best solutions to the next generation;

End while

The main advantage of this algorithm is its simplicity. In fact, compared with other population- or agent-based metaheuristic algorithms such as PSO and harmony search, there is basically only a single parameter Pa in CS. Hence, it is very easy to be implemented. Recent studies show that CS is potentially far more efficient than PSO and genetic algorithm [36,37] suggested that cuckoo search and differential evolution algorithms deliver more robust outcomes than PSO and ABC.

III. Conclusion

In modern years the swarm intelligence model has received well-known consideration in research, mainly as Artificial Bee Colony Optimization (ABC), Ant Colony Optimization (ACO), Firefly Algorithm, Particle Swarm Optimization (PSO) and Cuckoo Search (CS).These are the best common swarm intelligence metaheuristics for Single and Multi-Objective Problems. This paper presents a comprehensive review of the various swarm intelligence that is used for predicting the heart disease. As part of this review, we have tried to identify the main features of each optimization technique. Among these SI techniques Firefly and Cuckoo both are better when compared with other techniques like ABC, ACO, PSO.

References

1. Huan Liu and hiroshi motoda, Rudy Setiono and Zheng Zhao (2010): Feature Selection: An Everlasting Fronteir in Data Mining:, JMLR: The 4th Workshop on Feature Selection and Data Mining.
2. Svagowry S.Durairaj and A. Persia(2013): An empirical study on applying data mining techniques for the analysis and prediction of heart disease. Information Communication and Embedded Systems(ICICES).
3. Miss. Chaitrali S. Dangare, Dr. Mrs. Sulabha S. Apte (2012): A data mining approach for prediction of heart disease using neural networks, international journal of computer engineering and technology.
4. H. Yan, Y. Jiang, J. Zheng, B. Fu, S. Xiao, C. Peng(2004): The internet-based knowledge acquisition and management method to construct large-scale distributed medical expertsystems, Comput. Methods Programs Biomed.1–10.
5. E. Salzsieder, L. Vogt, K.-D. Kohnert, P. Heinke, P. Augstein,Model-based decision support in diabetes care, Comput.Methods Programs Biomed, 206–218.

6. S.˙Ic¸er(2013): Automatic segmentation of corpus collasum using gaussian mixture modeling and fuzzy C means methods, Comput. Methods Programs Biomed, 38–46.

7. C. Sáez, A. Bresó, J. Vicente, M. Robles, J.M. García-Gómez(2013): AnHL7-CDA wrapper for facilitating semantic interoperabilityto rule-based clinical decision support systems, Comput.Methods Programs Biomed, 239–249.

8. M. Hariharan, K. Polat, R. Sindhu(2014): A new hybrid intelligentsystem for accurate detection of Parkinson's disease,Comput. Methods Programs Biomed. 904–913.

9. Karaboga(2005): An idea based on honey bee swarm for numerical optimization, technical report, Computer Engineering Department, Engineering Faculty, Erciyes University.

10. X.S. Yang (2010): Firefly algorithm, Levy flights and globaloptimization, Res. Dev. Intell. Syst. XXVI. 209–218.

11. J. Kennedy, R. Eberhart(1995): Particle Swarm Optimization, IEEE International Conference on Neural Networks –Conference Proceedings, 1942–1948.

12. Pavlyukevich I (2007): L´evy flights, non-local search and simulated annealing, *J. Computational Physics*, 1830–1844.

13. Walton, S., Hassan, O., Morgan, K., and Brown, M.R(2011): Modified cuckoo search: a new gradient free optimization algorithm. *Chaos, Solitons & Fractals*, 44(9), 710-718.

14. Yang, X.S. and Deb, S.(2012): Multiobjective cuckoo search for design optimization, *Computers and Operations Research*, doi:10.1016/j.cor.2011.09.02644.

15. Karaboga, C. Ozturk, N. Karaboga, and B. Gorkemli(2012): Artificial bee colony programming for symbolic regression, *Information Sciences*, 1-15.

16. Ismail Babaoglu, Mustafa Servet Kıran, Erkan Ulker, and Mesut Gunduz (2013): Diagnosis of Coronary Artery Disease Using Artificial Bee Colony and K-NN Algorithms, International journal of computer and communication Engineering,2(1), 56-59.

17. B. Subanya, Rajalaxmi, R.R (2014): Feature Selection using Artificial Bee Colony for Cardiovascular Disease Classification, International Conference on Electronics and Communication System(ICECS- 2014).

18. Sowmya N. R. Vijayabhanu (2015): Artificial Bee Colony optimization for the prediction of Coronary Heart Disease, International Journal of Trend in Research and Development, ISSN 2394-9333.

19. Y. Yang and M. Kamel(2003): Clustering ensemble using swarm intelligence, in Swarm Intelligence Symposium, SIS'03. Proceedings of the 2003 IEEE.
20. M. Dorigo, V. Maniezzo, and A. Colorni(1996): he ant system: Optimization by a colony of cooperating agents," IEEE Transactions on Systems, Man and Cybernetics -Part B, 29–41.
21. Animesh Dubey, Rajendra Patel and Khyati Choure(2014): An Efficient Data Mining and Ant Colony Optimization technique (DMACO) for Heart Disease Prediction, *International Journal of Advanced Technology and Engineering Exploration.*
22. Yang X. S(2008, Nature-Inspired Metaheuristic Algorithms, Luniver Press, 2008.
23. Sh. M. Farahani, A. A. Abshouri, B. Nasiri, and M. R. Meybodi(2011): A Gaussian Firefly Algorithm, International Journal of Machine Learning and Computing. 2011.
24. *Xin-She Yan(2011):* Chaos-Enhanced Firefly Algorithm with Automatic Parameter Tuning, International Journal of Swarm Intelligence Research, 2011.
25. X.-S. Yang(2009): Firefly algorithms for multimodal optimization, in: Stochastic Algorithms: Foundations and Applications, SAGA 2009, Lecture Notes in Computer Sciences,169-178.
26. I. Fister, I. Fister Jr., X.-S. Yang, and J. Brest(2013): A comprehensive review of firefly algorithms, Swarm and Evolutionary Computation, 34-46.
27. A. H. Gandomi, X.S. Yang, S. Talatahari, A. H. Alavi(2013): Firefly algorithm with chaos, Commun Non linear Sci numer simulat, 89-98.
28. Kennedy, J. and Eberhart, R.C(1995): Particle swarm optimization. Proc. IEEE Int'l. Conf. on Neural Networks, IV, 1942-1948. Piscataway, NJ: IEEE Service Center.1995.
29. Yang, X. S., and Deb, S(2010): Engineering Optimisation by Cuckoo Search, Int. J. of Mathematical Modelling and Numerical Optimisation, 330-343.
30. Tuba, M., Subotic, M., Stanarevic, N(2011): Modified cuckoo search algorithm for unconstrained optimization problems, In: European Computing Conference, 263-268.
31. Tuba, M., Subotic, M., Stanarevic, N(2012): Performance of a Modified Cuckoo Search Algorithm for Unconstrained Optimization Problems, J. Faculty of Computer Science, 62-74.
32. Pavlyukevich, I(2007): Levy flights, non-local search and simulated annealing, J. Journal of Computational Physics, 1830-1844.

33. Chaowanawate, K., and Heednacram, A(2012): Implementation of Cuckoo Search in RBF Neural Network for Flood Forecasting, In: Fourth International Conference on Computational Intelligence, Communication Systems and Networks, 22-26.
34. Walton, S., Hassan, O., Morgan, K., and Brown, M(2011): Modified cuckoo search: A new gradient free optimisation algorithm. J. Chaos, Solitons& Fractals,44 (9):710-718.
35. Yang. XS., and Deb. S(2009): Cuckoo search via Lévy flights, In: World Congress on Nature & Biologically Inspired Computing, India, 210-214.
36. Civicioglu, P., and Besdok, E(2011): A conception comparison of the cuckoo search, particle swarm optimization, differential evolution and artificial bee colony algorithms, *Artificial Intelligence Review*, doi:10.1007/s10462-011-92760.

Corresponding Author:
Anbarasi M*,
Email: *manbarasi@vit.ac.in*