The Scientific Temper (2024) Vol. 15 (1): 1745-1754
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.1.23
https://scientifictemper.com/

RESEARCH ARTICLE

Modified firefly algorithm and different approaches for sentiment analysis

Radha K. Jana[1*], Dharmpal Singh[1], Saikat Maity[2]

Abstract

The firefly algorithm is the tool of swarm intelligence in almost all areas of optimization and engineering practice. Firefly algorithm and its variants have been used to solve diverse problems of real-world and motivate new researchers and algorithm developers to use this simple algorithm for problem-solving. The population of the initial solution, firefly light intensity rank selection and position update are important parameters of the firefly algorithm. The optimization result will vary in the variation of type of function selection by the researchers and developers. Therefore, here an effort has been made to propose a hybrid firefly algorithm where a population of initial solution generation done by the differential evolution algorithm (DEA) and selection of light intensity and change in position of firefly position will be done by position updating techniques of PSO. The hybrid firefly algorithm has been tested on five benchmark data set with nine benchmark algorithms and experiment result showed that firefly outperformed the other used algorithm. The aim of this study is also to motivate researchers to apply this algorithm for sentiment analysis also.

Keywords: Nature inspired algorithms, Swarm intelligence, Applications of bio-inspired algorithms, Image processing, Bio information and modified firefly, Sentiment analysis.

Introduction

Firefly is an insect that frequently produces short and rhythmic flashes that are formed by a process of bioluminescence. The function of the flashing light is to attract partners (communication) or probable prey. Thus, light intensity is the main factor in moving the fireflies toward the other fireflies.

The light intensity varies at the distance from the eyes of the beholder and decreases as the distance increases. The light intensity is also affected by the air absorbed by the surroundings, thus, the intensity becomes less appealing as the distance increases. Yang, X. S. (2010) opined that the firefly algorithm followed three idealize rules: 1) Fireflies are attracted toward each other regardless of gender. 2) The attractiveness of the fireflies is correlative with the brightness of the fireflies, thus the less attractive firefly will move forward to the more attractive firefly. 3) The brightness of fireflies depends on the objective function in Yang, X. S. (2010).

It is use to solve many problems, such as solving the vector quantization for image compression in Horng, M. H. (2012), economical emissions load dispatch problem [Apostolopoulos, T., & Vlachos, A. (2011)], solving traveling salesman problem [Kumbharana, S. N., & Pandey, G. M. (2013)], multilevel image thresholding selection [Horng, M. H., & Jiang, T. W. (2010, October)], object tracking [Gao, M. L., He, X. H., Luo, D. S., Jiang, J., & Teng, Q. Z. (2013)], finding optimal test sequence generation [Srivatsava, P. R., Mallikarjun, B., & Yang, X. S. (2013)] and sentiment analysis [Jana, R. K., & Maity, S. (2022)].

Related Work

In normal firefly algorithms, global best represent the current best solution of fireflies with the highest light intensity or attractiveness. The firefly's attractiveness will be lost in certain distance from the other fireflies due to its random movement to find the next best firefly for the next iteration. Therefore, this will cause ineffective performance in that particular iteration [Yang, X. S. (2010)], [Surafel, L. T., & Hong, C. (2012)] [Hassanzadeh, T., & Meybodi, M. R. (2012, May)]

[1]Department of Computer Science & Engineering, JIS University, Kolkata, West Bengal, India.

[2]Department of Computer Science & Engineering, Sister Nivedita University, New Town, West Bengal, India.

***Corresponding Author:** Radha K. Jana, Department of Computer Science & Engineering, JIS University, Kolkata, West Bengal, India., E-Mail: radhakrishnajana@gmail.com

How to cite this article: Jana, R. K., Singh, D., Maity, S. (2024). Modified firefly algorithm and different approaches for sentiment analysis. The Scientific Temper, **15**(1):1745-1754.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.1.23

Source of support: Nil

Conflict of interest: None.

Received: 12/01/2024 **Accepted:** 15/02/2024 **Published** : 15/03/2024

Surafel, L. T., & Hong, C. (2012) proposed a modification in the movement of fireflies by unit vector to determine the movement of firefly by the direction that leads toward the increases of brightness firefly. The current brightest firefly will stay in the current position if there is no firefly brighter than the current firefly. Wang, G., Guo, L., Duan, H., Liu, L., & Wang, H. (2012) also modified by Lévy flight for improvement in term of localized searching for closer solutions for firefly algorithm for improvement. The diversified authors Yu, S., Yang, S., & Su, S. (2013)., Olamaei, J., Moradi, M., & Kaboodi, T. (2013, April), Farahani, S. M., Abshouri, A. A., Nasiri, B., & Meybodi, M. (2011) have proposed to use adaptive formulation for randomization value α because when the value of α is large, it is better for the firefly to travel around the unknown place while small value of α will make firefly for local search.

Farahani, S. M., Abshouri, A. A., Nasiri, B., & Meybodi, M. (2011) proposed techniques to change the firefly's random movement from random to direct movement. They have opined that firefly will move randomly when there is no brighter firefly from the particular firefly and that random movement have to change to direct movement to find the best solution in that iteration. They have further opined that it will make firefly to reach a better position in next iteration and achieve global best.

Banati, H., & Bajaj, M. (2011) used the concept of firefly in compressing digital images, Yang, X. S. (2009, October) used to solving highly nonlinear, multimodal design, NP-hard problems, multi-objective load dispatch problems, scheduling problems [Das, G., Quraishi, M. I., & Barman, M. (2013)., Jati, G. K., & Suyanto. (2011, September)., Yousif, A., Abdullah, A. H., Nor, S. M., & Abdelaziz, A. A. (2011)], to solve flow shops and traveling salesman problem [Marichelvam, M. K., Prabaharan, T., & Yang, X. S. (2013)., Honarpisheh, Z., & Faez, K. (2013)] and local linear wavelet neural network [Dheeba, J., & Selvi, T. (2010)] for classifying the breast cancer. They have opined that firefly performed good in term of time and optimality of the solution.

Despite that the firefly algorithm is widely used to solve the problem, the standard firefly algorithm have some short-coming in term of trapping into several local optima for complex problems [Farook, S., & Raju, P. S. (2013)][Yu, S., Yang, S., & Su, S. (2013)]. Pseudo code for firefly algorithm is shown in Figure 1. Moreover, the function and techniques used to generate the initial population of firefly, evaluate the rank of light intensity and update the firefly's position vary with function variation. It has been observed that authors of diversified fields got different optimization results for the same data set. Furthermore, this is a disadvantage of using a single method only, which will be overly restrictive for high-dimensional and nonlinear problems. Thus, some modification and hybridization is suggested to overcome the shortcomings of single method [Abdullah, A., Deris, S., Mohamad, M. S., & Hashim, S. Z. M. (2012)] Figure 2.

Researchers have used the firefly algorithm in diversified filed of image processing and routing problems such as multilevel image thresholding selection [Horng, M. H., & Jiang, T. W. (2010, October)], active contour model for medical image segmentation [Sahoo, A., & Chandra, S. (2013, August)] and vector quantization using the firefly algorithm for image compression [Horng, M. H. (2012)] and traveling salesman problem [Jati, G. K., & Manurung, R. (2013)], Kumbharana and Pandey, [Kumbharana, S. N., & Pandey, G. M. (2013)], Wang *et al*.,([Wang, G., Guo, L., Duan, H., Liu, L., & Wang, H. (2012)] [Taş, D., Dellaert, N., Van Woensel, T., & De Kok, T. (2013)] [Pan, F., Ye, C., Wang, K., & Cao, J. (2013)]).

The world of people are busy in social media. They post their comment on different contexts in social media [Jana, R. K., & Maity, S. (2022)]. They also make groups in social media. They share their views in their group [Jana, R. K., Maity, S., & Maiti, S. (2022, May)]. There are various techniques for analyzing their comments. A genetic algorithm was applied for sentiment analysis [Radha, K., Dharmpal, S., Saikat, M., & Hrithik, P. (2023)].

There are various approaches for text categorization. Hadni, M., & Hassane, H. (2023) proposed a model. The proposed model showed better performance than other related models. The proposed algorithm is used for nature-inspired cybersecurity [Shandilya, S. K., Choi, B. J., Kumar, A., & Upadhyay, S. (2023)].

In this paper, an effort has been made to provide a general rule to generate the initial population of fireflies, evaluate the rank of light intensity, and update the position of the firefly using the concept of particle swarm optimization (PSO) and differential evolution algorithm.

In this hybrid firefly algorithm, the population of initial solution is generated by the initial population of the differential evolution algorithm (DEA) and selection of light intensity and change in firefly position will be done by position updating techniques of PSO.

Methodology

The firefly algorithm mainly works on two important variables: the light intensity and attractiveness and the second one attraction of firefly toward the other firefly which have a brighter flash than itself. The attractiveness of firefly depends on the light intensity.

```
Firefly Algorithm
Objective function f(x),      x = (x_1, ..., x_d)^T
Generate initial population of fireflies x_i (i = 1, 2, ..., n)
Light intensity I_i at x_i is determined by f(x_i)
Define light absorption coefficient γ
while (t < MaxGeneration)
for i = 1 : n all n fireflies
    for j = 1 : n all n fireflies (inner loop)
    if (I_i < I_j), Move firefly i towards j; end if
        Vary attractiveness with distance r via exp[−γr]
        Evaluate new solutions and update light intensity
    end for j
end for i
Rank the fireflies and find the current global best g_*
end while
Postprocess results and visualization
```

Figure 1: Pseudo code for firefly algorithm (Yang, 2010)

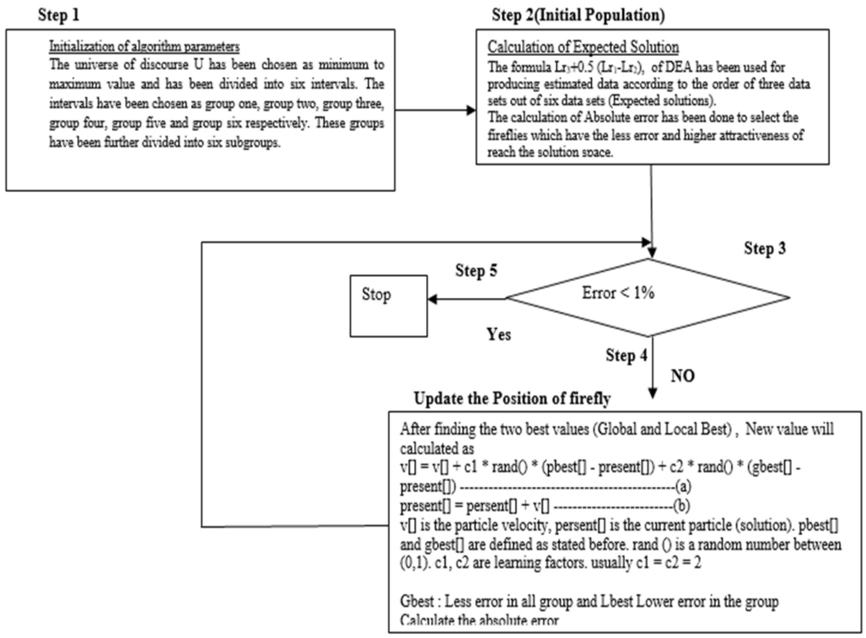

Figure 2: Proposed modified firefly algorithm

The light intensity thus attractiveness, is inversely proportional with the particular distance *r* from the light source. Thus, the light and attractiveness is decrease as the distance increases.

$$I(r) = I_0 e^{-\gamma r^2} \tag{1}$$

I = light intensity,

I_0 = light intensity at initial or original light intensity,

γ = the light absorption coefficient

r = distance between firefly i and j

Attractiveness is proportionally to the light intensity seen by another firefly, thus attractiveness is β

$$\beta = \beta_0 e^{-\gamma r^2} \tag{2}$$

β_0 = Attractiveness at r is 0

The distance between two fireflies can define using the Cartesian distance

$$r_{ij} = |x_i - x_j| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \tag{3}$$

Firefly *i* is attracted toward the more attractive firefly *j*, the movement is defined as

$$\Delta x_i = \beta_0 e^{-\gamma r_{ij}^2} (x_j^t - x_i^t) + \alpha \varepsilon_i, \quad x_i^{t+1} + \Delta x_i \tag{4}$$

In equation (4), the first term is for attraction, γ is the limitation when the value is tend to zero or too large. If γ approaching zero ($\gamma \to 0$), the attractiveness and brightness become constant, $\beta = \beta_0$. In other words, a firefly can be seen in any position, easy to complete global search. If the γ is nearing infinity or too large ($\gamma \to \infty$), the attractiveness and brightness become decreased. The firefly movements become random. The implementation of firefly algorithm can be done in these two asymptotic behaviors. The second term is for randomization, as α is the randomization parameter. The ε_i can be replaced by ran -1/2, which is a random number generated from 0 to 1.

Hybrid Firefly Algorithm

Single metaheuristic algorithm is restrictive to solve the problem and reach the optimal solution within a reasonable time. Therefore, a hybrid metaheuristic called hybridization will be used to solve the high dimensional and nonlinear problems. The concept of hydride metaheuristic can be show more efficient behavior with higher flexibility to deal the problem of real-world and large scale [Abdullah, A., Deris, S., Mohamad, M. S., & Hashim, S. Z. M. (2012)], [El-Sawy, A., Zaki, E. and Rizk-Allah, R.(2012)], [Rizk-Allah, R. M., Zaki, E. M., & El-Sawy, A. A. (2013)].

Farahani, S. M., Abshouri, A. A., Nasiri, B., & Meybodi, M. (2011) proposed a hybridized firefly algorithm using genetic algorithm to overcome the problem of global optima of firefly. El-Sawy, A., Zaki, E. and Rizk-Allah, R. (2012) proposed hybridized firefly algorithm with the ant colony optimization on a parallel approached to sort the result in descending order of constrain violation of the feasibility rule. (Abdullah *et al.*, [Abdullah, A., Deris, S., Mohamad, M. S., & Hashim, S. Z. M. (2012)] proposed a combined firefly algorithm with differential evolution where the population of firefly will be creating into two group; one is in the with the potential fitness values, and will proceed to calculate the distance between solution using Euclidean distance and attractiveness. The others sub-population with less significant fitness value will use the evolutionary operation of differential evolution.

Modified Firefly Algorithm

In case of the numeric data *set all* the elements have to sort in ascending order and the universe of discourse U has been chosen as minimum to maximum value and divided into six

Table 1: Group and subgroup range

Group and subgroup	*Subgroup*					
Group one (A1)	*A11*	*A12*	*A13*	*A14*	*A15*	*A16*
(3300-4386)	(3300-3481)	(3481-3662)	(3662-3843)	(3843-4024)	(4024-4205)	(4205-4386)
Group Two(A2)	A21	A22	A23	A24	A25	A26
(4386-5472)	(4386-4567)	(4567-4748)	(4748-4929)	(4929-5110)	(5110-5291)	(5291-5472)
Group Three(A3)	A31	A32	A33	A34	A35	A36
(5472-6588)	(5472-5653)	(5653-5834)	(5834-6015)	(6015-6196)	(6196-6377)	(6377-6558)
Group Four(A4)	A41	A42	A43	A44	A45	A46
(6558-7644)	(6588-6739)	(6739-6920)	(6920-7101)	(7101-7282)	(7282-7463)	(7463-7644)
Group Five(A5)	A51	A52	A53	A54	A55	A56
(7644-8730)	(7644-7825)	(7825-8006)	(8006-8187)	(8187-8386)	(8386-8549)	(8549-8730)
Group Six(A6)	A61	A62	A63	A64	A65	A66
(8730-9816)	(8730-8911)	(8911-9092)	(9092-9273)	(9273-9454)	(9454-9635)	(9635-9816)

intervals. The intervals have been chosen as groups one, two, three, four, five, and six, respectively. These groups have been further divided into six subgroups (Figure 2).

Generation of Initial Population

The expected solution has been formed by continuously adding the step-up value to the starting value of every subgroup of each group until it has reached the end value of that subgroup.

From the available data, initially, it is necessary to see the group number where the available data belongs. After selecting the group number, the expected solution has to be formed.

The expected solution has been formed by continuously adding the step up value 30.16 to the starting value of every subgroup of each group until it has reached the end value of that subgroup. As for example, suppose Pa represents the data element 3331.597 belonging to serial number one of group one at the beginning of the iteration. The expected solution has been formed as Pa (t) 3330.16, 3360.32, 3390.48, 3420.64, 3450.8 and 3480.96, respectively, by adding the step-up value as 30.16 to 3300 (Initial value from Table 1 group A11) to get expected solution 1 as 3330.16 (Pa (t+1)). Now the step-up value as 30.16 has been added to 3330.16 (Pa (t+1)) to get expected solution 2 as 3360.32 (Pa (t+2)). This addition continued until it reached the finishing value of the range at 3481.

The said ranges of solutions have been formed as expected solutions of data item belonging to serial number one of group one. Accordingly, all expected solutions corresponding to all data elements have been formed. The initial population of data has been formed using the data values of the expected solution. At a time, three of six expected data sets have been used in a particular order using the available data to produce estimated data. The formula $Lr_3+0.5\ (Lr_1-Lr_2)$ of DEA has been used to produce estimated data according to the order of three data sets out of six data sets (Expected solutions).

It is noteworthy that three expected solutions out of six have been used in four ways (orders) for producing estimated data. These orders of usage of three expected solutions are (ES_1, ES_2, and ES3), (ES_2, ES_3, ES_4), (ES_3, ES_4, ES_5) and (ES_4, ES_5, ES_6).

The formula $Lr_3+0.5\ (Lr_1-Lr_2)$ and the order of three expected solutions for producing estimated data have been used. As for an example, the available data for serial number one of group one is 3331.597. Six expected solutions are 3330.16, 3360.32, 3390.48, 3420.64, 3450.8 and 3480.96, corresponding to the expected solution (Table 2). They are 3330.16 as expected solution number one, 3360.32 as expected solution number two, 3390.48 as expected solution number three, 3420.64 as expected solution number four, 3450.8 as expected solution number five and 3480.96 as expected solution number six as furnished in Table 1.

If the expected solution order (ES1, ES2, and ES3) is used, the value of Lr3 becomes 3390.48, the expected solution three of data element as 3331.597. The value of Lr_2 becomes 3360.32 which is the expected solution two of data element as 3331.597. The value Lr_1 becomes 3330.16, which is the expected solution one of the data item as 3331.597. Now using the formula, the estimated data has become 3375.40. The estimated data has become 3405.56 for the expected solution order (ES_2, ES_3, ES_4), 3435.72 for (ES_3, ES_4, ES_5), and 3465.88 for (ES_4, ES_5, ES_6).

Thereafter, the estimated error will be calculated for each data element to select the closest of the element of their optimal solution.

This estimated error is calculated as follows:

Estimated Error= Absolute ((Estimated Error-Actual)/ctual))*100 -------------------- (1)

Work as the intensity of the light of the firefly. Furthermore, based on the estimated error values, firefly decides their next move and changes their position accordingly.

Table 2: Recent approaches for sentiment analysis

Ref.	*Year*	*Method*	*Dataset*	*Evaluation metrics (Accuracy)*
T. Sreenivasulu *et al*.	2020	FFLY+EDBN	Product review in Amazon (13871)	98.60
Kiran Sahu *et al*.	2021	FFOGA	Twitter dataset	91.52 (Joy), 96.61 (Love), 94.91 (Sad)
Samik Datta *et al*.	2021	FF-MVO-RNN	demonetization tweets	89.79
D. Elangovan *et al*.	2022	FFL-MLP	Canon (500 dataset), iPod (1000)	Canon (98% accuracy) and iPod (99% accuracy).
H. Swapnarekha *et al*.	2022	Firefly Algorithm + LSTM	Covid-19 tweets (5016)	99.59%
D. Elangovan *et al*.	2023	FF-MLP	DVD Database	97.97

Change of the position

Proposed FA changes their position according to the formula of PSO as furnished below.

v[t+1] = v[t] + c1 × rand()× (pbest[t] - present[t]) + c2 × rand()× (gbest[t] - present[t]) -------(2)

present[t+1] = present[t] + v[] -------------------------(3)

v[] is the initial velocity of firefly assume to 0, persent[] is the current particle (solution). pbest[] and gbest[] are defined as global best (Represent the minimum error of fireflies in all group) and local best (Minimum error of in their group) of the firefly, rand () is a random number between (0,1), c1, c2 are learning factors assume as c1 =0.3 c2 =0.4.

Limitation of Other Used Algorithm

Most swarm intelligence algorithms viz. artificial bee colony (ABC), ant colony optimization (ACO), BA and PSO used different types of populations, structured to perform the different jobs with the population cooperating to search in the solution space. However, proposed FA used the fixed rule for the initial population generation and updated firefly position to achieve the optimal result for the numeric data set.

The authors of diversified used the different functions to generate the initial population and update the firefly's position, which will cause a different result for the same data set. The modified firefly algorithms used the fixed rule based to generate initial population and move the position. The firefly easily uses these techniques for the numeric dataset to find the optimal solution compared to other algorithms.

Although authors have used DEA algorithm and PSO with firefly to generate the hybrid structure to optimize the problem domain but in this paper, initial population generation and movement of next position has been inspired by the initial population DEA and next position movement of PSO, respectively. This will help the firefly algorithm generate the same optimized result for the same data set and avoid it from being struck in the local optimal.

Experimental Result on Benchmark Data

Experimental results have been performed by famous algorithms, namely genetic algorithm [Singh, D. P., Choudhury, J. P., & De, M. (2014, May)], particle swarm optimization, artificial bee colony optimization [Singh, D., Choudhury, J. P., & De, M. (2012)], harmony search [P. Singh, J. P. Choudhury and M. De. (2012)], ant colony optimization [Singh, D. P., Choudhury, J. P., & De, M. (2012)], tabu search [P. Singh, J. P. Choudhury and M. De.(2012)], simulated annealing [Singh, D. P., Choudhury, J. P., & De, M. (2013)] and differential evolution algorithm [Singh, D. P., Choudhury, J. P., & De, M. (2013)] on four data set like Iris flower data set [http://en.wikipedia.org/wiki/Iris_flower_data_set], Wine data set [http:archive.ics.uci.edu/ml/datasets/wine + quality], Boston city data set [http://www.ics.uci.edu/~mlearn/MLRepository.html,http://lib.stat.cmu.edu/datasets] and Concrete slump test [https://archive.ics.uci.edu/ml/datasets/Concrete+Slump+Test].

In PSO, both positions of particles and the velocity of particle have been considered. Based on the particle positions, local best position as pbest and global best position as gbest have been decided. Optimization using the method of PSO depends on the change of velocity; change of position and on its local best position pbest and global best position gbest. It may be difficult to decide the best local and global positions. Therefore, particle swarm optimization may not be suitable soft computing model. In DEA, the expected solutions (initial populations) depend on the maximum and minimum value of data items. In mutation operation under DEA, new off springs are created using three off springs and using a certain constant random number. Based on the fitness function value of the existing offspring, the said offspring may be incorporated into the population or the said offspring may be dropped. Here the fitness function, the way of combination of off springs may not be accurate. It may take quite a number of iterations to achieve optimization. Therefore, the DEA may not be suitable. In artificial bee colony optimization, the fitness function value of the population and recruitment of bees for selected sites has to be decided based on the fitness function value of the key at the beginning. Therefore, ABC optimization may not be feasible for knowledge extraction. In tabu search (TS) and simulated annealing, the initial solution must be obtained using a certain standard curve. The finalization of standard curve may be difficult. That is the reason for not preferring tabu search and simulated annealing methods. In the genetic algorithm (GA), the success of the model depends on proper selection of the fitness function. The fitness function can be chosen based on the trend of data values or some adhoc theoretical concept. That is the reason GA may not be suitable for optimization tool. Ant colony optimization allows the ants to build solutions incrementally through the use of a probabilistic transition rule, based on the amount of pheromone in the trial and on a local problem-dependent heuristic. The items can be added to the current partial solution based on some heuristic function. That heuristic function is problem dependent. Since the set of solutions initially formed and the set of solutions to be added depend totally on some heuristic function, it may happen that ACO may not produce good results in some practical problem.

As their overall performance in terms of best fitness values achievable and convergence speed is comparable to modify FA and due to the size of the paper, the detailed simulation results will not be presented in this paper.

Results

Preprocessing is necessary to simplify subsequent operations without losing relevant information. Since all the data sets have been considered reliable source and have not contained any missing value, any inconsistent data, i.e., any abnormally low or any abnormally high value (all the data values are regularly distributed within the range of that data item), the data cleansing techniques are not available to the applicable data. Moreover, the available data have been taken from a single source. Therefore, the data integration technique is not required. In the preprocessing stage, only smoothing aggregation, normalization, and decimal scaling techniques are required.

After getting the proper data, antecedent and consequent items have been formed. The methods of multivariate analysis (factor analysis and principal component analysis) have been applied on the available input data to make a decision to select the optimal model for the formation of association rule. The cumulative antecedent item has been formed using these methods. Furthermore, it has been observed that factor analysis is more effective than principal component analysis after the formation of clusters using Iris data. Moreover, factor analysis is more effective as compared to principal component analysis after the formation of clusters using other data items i.e., wine data set. Breast cancer data set and concrete slump test. It is noted that principal component analysis is effective for Boston city-data.

Soft computing models viz. GA, PSO, ABC, ACO, HS, DEA, SA TS, and modified firefly algorithms have been applied to produce estimated data on the Iris flower, wine, Boston City and concrete slump data. The error analysis and residual analysis have been applied on the estimated data with respect to the available data. The computation of average error and residual analysis methods are like the sum of absolute residual, mean of absolute residual, mean of absolute residual, median of absolute residual, maximum of absolute residual, and standard deviation of absolute residual.

Furthermore, it has been observed that the modified firefly algorithm is more effective than other models in Iris data set, wine data, Boston city and concrete slump test data set, respectively.

Experimental Comparison with Other Algorithms

It has been observed that proposed firefly algorithm is the best prediction model among all the models for the extraction of knowledge-based information. The performance of any model can be evaluated on the basis of two major components viz. diversification and intensification. Proper diversification makes sure the search in the parameter space which can explore many locations and regions as possible in an efficient and effective manner so that the evolving system will not be tapped in biased local optima. The appropriate intensification intends to ensure to speed up the convergence by reducing randomness and limiting diversification.

In proposed firefly algorithm, diversification is controlled by the estimated error of the firefly and randomization. New solutions are generated *via* PSO technique. In addition, by considering the estimated error and randomization of the proposed algorithm, the algorithm effectively reaches the global search space.

The intensification is represented by estimated error and generation of the new solution as stated in section 3. Such interactions between various components may be the reason for the success of the proposed spider algorithm over other algorithms.

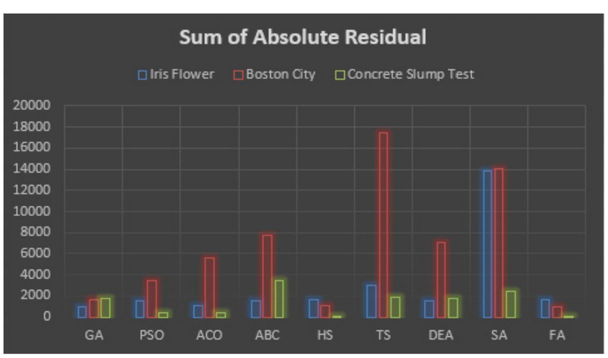

Figure 3: Sum of absolute residual

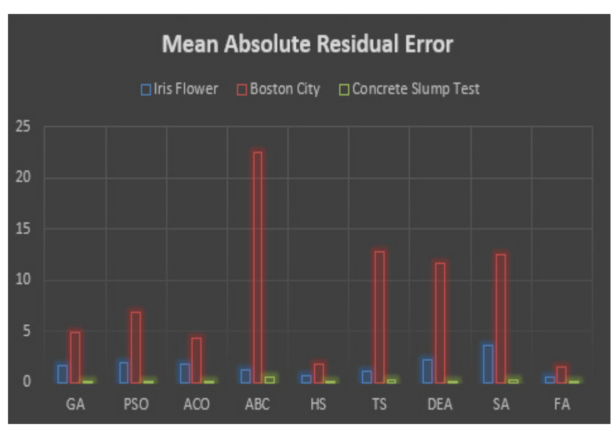

Figure 4: Maximum absolute residual

Scalability Test

In this section, the simulation results of FA along with other benchmark algorithms identified in section 4 have been furnished in Figures 3-8, respectively.

Computational Complexity

The computational test was applied to the proposed firefly and other aforesaid algorithms to check the computational complexity as shown in Figure 9.

Reliability Test

The reliability test was applied to the proposed firefly algorithm along with another aforesaid algorithm to check the visualizing performance of all algorithms on all benchmark data set success rate test of the used algorithm is shown in Figure 10.

Discussion

According to the no-free-lunch (NFL) theorem [Wolpert, D. H., & Macready, W. G. (1997)], meta-heuristics produce exactly the same extremum search for evaluating and average objective function. Moreover, it has further opined that it is theoretically impossible to have a better general-purpose universal optimization technique [Ho, Y. C., & Pepyne, D. L. (2002)] but classes of problems [Ho, Y. C., & Pepyne, D. L. (2002)] or general, but real-world ones [Garcia-Martinez, C., Rodriguez, F. J., & Lozano, M. (2012)] may produce the superior performing. Furthermore, the high number of possible problems provides much room to develop the new algorithm.

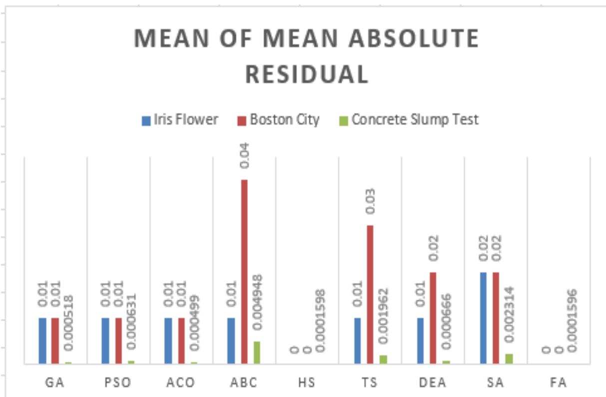

Figure 5: Mean of mean absolute residual

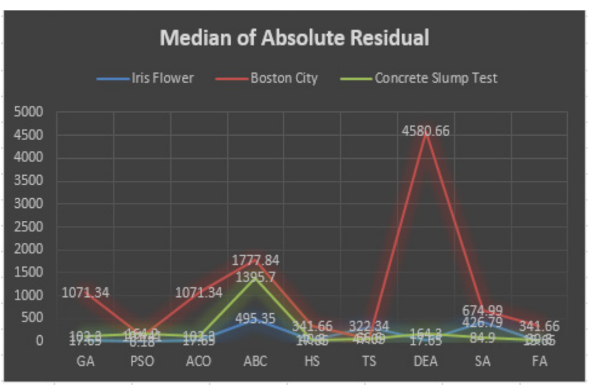

Figure 6: Median of absolute residual

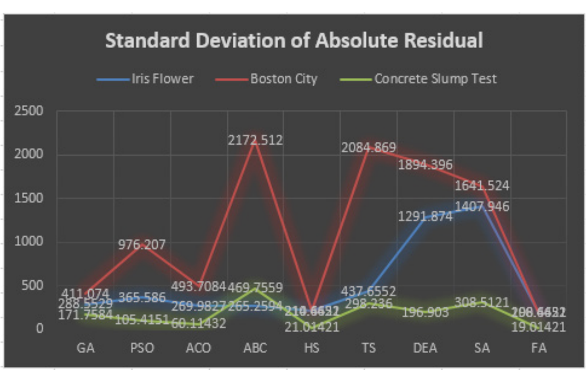

Figure 7: Standard deviation

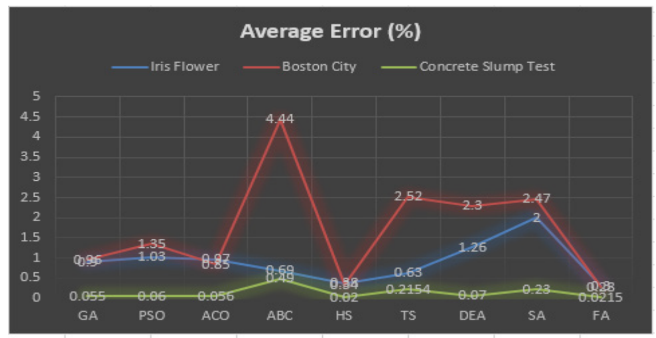

Figure 8: Average error (%)

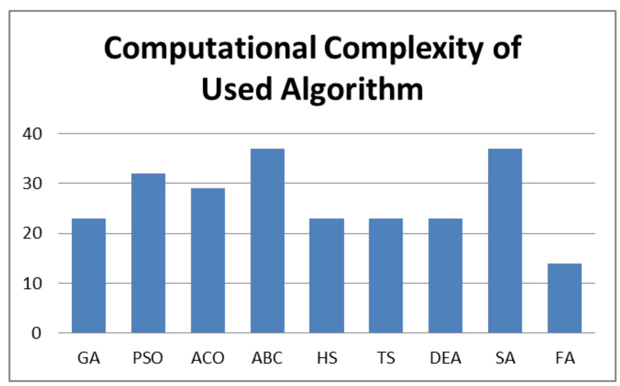

Figure 9: Computational complexity of used algorithm

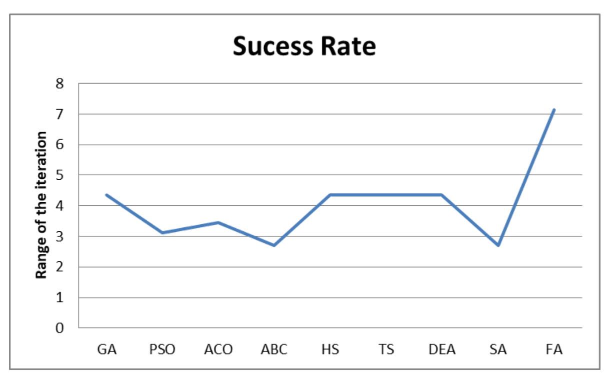

Figure 10: Success rate test of the used algorithm

Though existing meta-heuristics successfully solve many optimization problems, it is always to develop a new searching algorithm with superior performance in other class of problems [Lam, A. Y., & Li, V. O. (2009)].

This is the motivation for us to propose a modified firefly algorithm for solving numerical optimization problems. This modified algorithm is applied on social tweets for sentiment analysis.

Different Firefly Algorithm Approaches for Sentiment Analysis

Sentiment analysis

Sentiment analysis is a process of surveying people's sentiments, feelings, etc. Different types of content post on social media. These are text, audio, video and picture. This type of data is to be prepared into functional statistics with the help of the latest techniques. The sentiment are negative sentiment, positive sentiment or neutral sentiment [Jana, R. K., Maity, S., & Maiti, S. (2022, May)].

Different techniques for sentiment analysis

There are various algorithms for sentiment analysis. We showed different firefly approaches for sentiment analysis. We also showed the performance of different approaches in sentiment analysis in Table 2.

Conclusion

The firefly algorithm is considered new algorithm in the swarm intelligence family. Despite that, the usage of the firefly algorithm in the various types of problems shows the anticipation from the researcher to use this algorithm.

This algorithm already proves that it is superior to the previously introduced swarm intelligence from the previous research. Even though the firefly algorithm has proven superior to the previous swarm intelligence, some modification can be done to improve the local search and global search to ensure the solution obtained is optimum and not premature.

The firefly algorithm is also suitable for high-dimensional and nonlinear problems.

The downside is that the single metaheuristic is hard to reach an optimal solution within a reasonable time. Thus, combining the metaheuristic will help overcome the shortcomings of the single metaheuristic algorithm.

In the proposed firefly algorithm, diversification is controlled by the estimated error of the firefly and randomization. New solutions are generated *via* PSO technique. In addition, by considering the estimated error and randomization of the proposed algorithm, the algorithm effectively reaches the global search space.

In future work, the researcher should tackle firefly in various problems such as finding an optimum route for newly built trains rails route, which has multiple constraints such as preserving nature as much as possible and multi-tracking for object tracking.

This is the motivation for us to propose a modified firefly algorithm for solving numerical optimization problems.

In the future we also proposed a sentiment analysis and opinion mining model for better performance than another existing model.

Acknowledgment

I sincerely appreciate my two co-researchers, Dharmpal Singh and Saikat Maity, for their research contribution.

References

Abdullah, A., Deris, S., Mohamad, M. S., & Hashim, S. Z. M. (2012). A new hybrid firefly algorithm for complex and nonlinear problem. In *Distributed Computing and Artificial Intelligence: 9th International Conference* (pp. 673-680). Springer Berlin Heidelberg.

Apostolopoulos, T., & Vlachos, A. (2011). Application of the firefly algorithm for solving the economic emissions load dispatch problem. *International journal of combinatorics*, *2011*.

Banati, H., & Bajaj, M. (2011). Fire fly based feature selection approach. *International Journal of Computer Science Issues (IJCSI)*, *8*(4), 473.

Das, G., Quraishi, M. I., & Barman, M. (2013). Firefly algorithm based mammographic image analysis. *Asian Journal of Computer Science and Information Technology*, *3*(4), 56-59.

Dheeba, J., & Selvi, T. (2010). Bio inspired swarm algorithm for tumor detection in digital mammogram. In *Swarm, Evolutionary, and Memetic Computing: First International Conference on Swarm, Evolutionary, and Memetic Computing, SEMCCO 2010, Chennai, India, December 16-18, 2010. Proceedings 1* (pp. 404-415). Springer Berlin Heidelberg.

Datta, S., & Chakrabarti, S. (2021). Aspect based sentiment analysis for demonetization tweets by optimized recurrent neural network using fire fly-oriented multi-verse optimizer. *Sādhanā*, *46*(2), 79.

El-Sawy, A., Zaki, E. and Rizk-Allah, R.(2012). A Novel Hybrid Ant Colony Optimization and Firefly Algorithm for Solving Constrained Engineering Design Problems. *Journal of Natural Sciences and Mathematics*, 6(1),pp.1-22.

Elangovan, D., & Subedha, V. (2023). Firefly with Levy Based Feature Selection with Multilayer Perceptron for Sentiment Analysis. *Journal of Advances in Information Technology*, *14*(2).

Elangovan, D., & Subedha, V. (2023). Automatic Sentimental Analysis by Firefly with Levy and Multilayer Perceptron. *Computer Systems Science & Engineering*, *46*(3).

Farahani, S. M., Abshouri, A. A., Nasiri, B., & Meybodi, M. (2011). A Gaussian firefly algorithm. *International Journal of Machine Learning and Computing*, *1*(5), 448.

Farook, S., & Raju, P. S. (2013). Evolutionary hybrid genetic-firefly algorithm for global optimization. *IJCEM International Journal of Computational Engineering & Management*, *16*(3), 37-45.

Gao, M. L., He, X. H., Luo, D. S., Jiang, J., & Teng, Q. Z. (2013). Object tracking using firefly algorithm. *IET Computer Vision*, *7*(4), 227-237.

Garcia-Martinez, C., Rodriguez, F. J., & Lozano, M. (2012). Arbitrary function optimisation with metaheuristics: No free lunch and real-world problems. *Soft Computing*, *16*, 2115-2133.

Ho, Y. C., & Pepyne, D. L. (2002). Simple explanation of the no-free-lunch theorem and its implications. *Journal of optimization theory and applications*, *115*, 549-570.

Horng, M. H. (2012). Vector quantization using the firefly algorithm for image compression. *Expert Systems with Applications*, *39*(1), 1078-1091.

Horng, M. H., & Jiang, T. W. (2010, October). Multilevel image thresholding selection based on the firefly algorithm. In *2010 7th International Conference on Ubiquitous Intelligence & Computing and 7th International Conference on Autonomic & Trusted Computing* (pp. 58-63). IEEE.

Hassanzadeh, T., & Meybodi, M. R. (2012, May). A new hybrid algorithm based on firefly algorithm and cellular learning automata. In *20th Iranian conference on electrical engineering (ICEE2012)* (pp. 628-633). IEEE.

Honarpisheh, Z., & Faez, K. (2013). An Efficient Dorsal Hand Vein Recognition Based on Firefly Algorithm. *International Journal of Electrical & Computer Engineering (2088-8708)*, *3*(1).

Horng, M. H., & Jiang, T. W. (2010, October). Multilevel image thresholding selection based on the firefly algorithm. In *2010 7th International Conference on Ubiquitous Intelligence & Computing and 7th International Conference on Autonomic & Trusted Computing* (pp. 58-63). IEEE.

Hadni, M., & Hassane, H. (2023). New model of feature selection based chaotic firefly algorithm for arabic text categorization. *Int. Arab J. Inf. Technol.*, *20*(3A), 461-468.

http://en.wikipedia.org/wiki/Iris_flower_data_set.

http:archive.ics.uci.edu/ml/datasets/wine + quality.

http://www.ics.uci.edu/~mlearn/MLRepository.html,http://lib.stat.cmu.edu/datasets.

https://archive.ics.uci.edu/ml/datasets/Concrete+Slump+Test.

Jana, R. K., & Maity, S. (2022). An Accuracy based Comparative Study on Different Techniques and challenges for Sentiment Analysis. *Pervasive Computing and Social Networking: Proceedings of ICPCSN 2022*, 601-619.

Jana, R. K., Maity, S., & Maiti, S. (2022, May). An Empirical Study of Sentiment and Behavioural Analysis using Homophily Effect in Social Network. In *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1508-1515). IEEE.

Jati, G. K., & Manurung, R. (2013). Discrete firefly algorithm for traveling salesman problem: A new movement scheme. In *Swarm Intelligence and Bio-Inspired Computation* (pp. 295-312). Elsevier.

Jati, G. K., & Suyanto. (2011, September). Evolutionary discrete firefly algorithm for travelling salesman problem. In *International conference on adaptive and intelligent systems* (pp. 393-403). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kumbharana, S. N., & Pandey, G. M. (2013). Solving travelling salesman problem using firefly algorithm. *International Journal for Research in science & advanced Technologies*, *2*(2), 53-57.

Lam, A. Y., & Li, V. O. (2009). Chemical-reaction-inspired metaheuristic for optimization. *IEEE transactions on evolutionary computation*, *14*(3), 381-399.

Marichelvam, M. K., Prabaharan, T., & Yang, X. S. (2013). A discrete firefly algorithm for the multi-objective hybrid flowshop scheduling problems. *IEEE transactions on evolutionary computation*, *18*(2), 301-305.

Olamaei, J., Moradi, M., & Kaboodi, T. (2013, April). A new adaptive modified firefly algorithm to solve optimal capacitor placement problem. In *18th Electric power distribution conference* (pp. 1-6). IEEE.

Pan, F., Ye, C., Wang, K., & Cao, J. (2013). Research on the Vehicle Routing Problem with Time Windows Using Firefly Algorithm. *J. Comput.*, *8*(9), 2256-2261.

P. Singh, J. P. Choudhury and M. De.(2012) A Comparative Study to discover the knowledge in data mining by soft computing model, Proceeding of CCSN-20121st International Conference on Computing, Communication and Sensor Network, Raurkela, pp. 90-94.

Rizk-Allah, R. M., Zaki, E. M., & El-Sawy, A. A. (2013). Hybridizing ant colony optimization with firefly algorithm for unconstrained optimization problems. *Applied Mathematics and Computation*, *224*, 473-483.

Radha, K., Dharmpal, S., Saikat, M., & Hrithik, P. (2023). Perception Analysis of Social Network Tweets Using a Genetic Algorithm Based Approach s. *Journal of Software Engineering Tools &Technology Trends*, *10*(02), 27-38.

Srivatsava, P. R., Mallikarjun, B., & Yang, X. S. (2013). Optimal test sequence generation using firefly algorithm. *Swarm and Evolutionary Computation*, *8*, 44-53.

Surafel, L. T., & Hong, C. (2012). Modified firefly algorithm. *Journal of Applied Mathematics*, *39*(2), 01-12.

Sahoo, A., & Chandra, S. (2013, August). L'evy-flight firefly algorithm based active contour model for medical image segmentation. In *2013 Sixth International Conference on Contemporary Computing (IC3)* (pp. 159-162). IEEE.

Singh, D. P., Choudhury, J. P., & De, M. (2014, May). A comparative study on principal component analysis and factor analysis for the formation of association rule in data mining domain. In *Proceedings of the 2 nd International Conference on Mathematical, Computational and Statistical Sciences (MCSS'14), Gdansk* (pp. 442-452).

Singh, D., Choudhury, J. P., & De, M. (2012). A comparative study on the performance of fuzzy logic, Bayesian logic and neural network towards decision-making. *International Journal of Data Analysis Techniques and Strategies*, *4*(2), 205-216.

Singh, D. P., Choudhury, J. P., & De, M. (2012). A comparative study to select a soft computing model for knowledge discovery in data mining. *International Journal of Artificial Intelligence and Knowledge Discovery*, *2*(2), 6-19.

Singh, D. P., Choudhury, J. P., & De, M. (2013). An effort to developing the knowledge base in data mining by factor analysis and soft computing methodology. *International Journal of Scientific & Engineering Research (IJSER)*, *4*(9), 1912-1923.

Shandilya, S. K., Choi, B. J., Kumar, A., & Upadhyay, S. (2023). Modified Firefly Optimization Algorithm-Based IDS for Nature-Inspired Cybersecurity. *Processes*, *11*(3), 715.

Sreenivasulu, M. T., & Jayakarthik, R. (2020). A Reinforced Deep Belief Network with Firefly Optimization for Sentiment Analysis of Customer Product Review.

Sahu, M. T. S. K., & Sharma, A. (2021). Fire Fly Optimization Based Sentiment Mining Using Pattern Feature.

Swapnarekha, H., Dash, P. B., & Pelusi, D. (2023). An optimistic firefly algorithm-based deep learning approach for sentiment analysis of COVID-19 tweets. *Mathematical Biosciences and Engineering*, *20*(2), 2382-2407.

Taş, D., Dellaert, N., Van Woensel, T., & De Kok, T. (2013). Vehicle routing problem with stochastic travel times including soft time windows and service costs. *Computers & Operations Research*, *40*(1), 214-224.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), 67-82.

Wang, G., Guo, L., Duan, H., Liu, L., & Wang, H. (2012). A modified firefly algorithm for UCAV path planning. *International Journal of Hybrid Information Technology*, *5*(3), 123-144.

Yu, S., Yang, S., & Su, S. (2013). Self-adaptive step firefly algorithm. *Journal of Applied Mathematics*, *2013*, 1-8.

Yang, X. S. (2010). *Nature-inspired metaheuristic algorithms*. Luniver press.

Yang, X. S. (2009, October). Firefly algorithms for multimodal optimization. In *International symposium on stochastic algorithms* (pp. 169-178). Berlin, Heidelberg: Springer Berlin Heidelberg.

Yousif, A., Abdullah, A. H., Nor, S. M., & Abdelaziz, A. A. (2011). Scheduling jobs on grid computing using firefly algorithm. *Journal of Theoretical and Applied Information Technology*, *33*(2), 155-164.