ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume X, Number X, XXXX, XXX–XXX

Supplementary material of the paper Umit Kilic, Esra Sarac Essiz, Mumine Kaya Keles

Umit KILIC[1, *], Esra SARAC ESSIZ[1], and Mumine KAYA KELES[1]

[1]Department of Computer Engineering, Adana Alparslan Turkes Science and Technology University, Saricam, 01230, Adana, Turkey

Email: ukilic@atu.edu.tr*, esarac@atu.edu.tr, mkaya@atu.edu.tr

* Corresponding author

This PDF file contains the tables and the figures created for the Binary Anarchic Society Optimization Algorithm for Feature Selection titled article published in the Romanian Journal of Information Science and Technology. You can access the article via this link: (the link will be added after publication)

1. Supplementary materials

Pseudo-code of the proposed Binary Anarchic Society Algorithm (BASO) can be seen below (Algorithm 1).

Algorithm 1: Pseudo-code of the proposed BASO

```
Create the initial population randomly and calculate the fitness value of each solution;
i* = the best solution;
GBest = the best solution;
while t < max iteration do
    for each member do
        Calculate FI, EI, and II using (8), (9), and (10);
    end
    Assigning β1, β2, β3 using mean of FI, EI, and II;
    for each member do
        Calculate fitnesses for the positions produced based on MP_i^Current,
         MP_i^Society, and MP_i^Past using (11), (12), (13) and save the positions.;
        Select the best fitness among them and assign its position to the member.;
    end
    Determine the iteration best and update i*;
    if i* > GBest then
        Update GBest as i*;
    end
    t=t+1;
end
return GBest;
```

The three indexes, fickleness index (FI), external irregularity index (EI), and internal irregularity index (II), used in BASO can be found using the following equations.

$$FI_i(k) = \frac{f(X_i(k)) - f(X_{ib}(k))}{f(X_{ib}(k))}, \tag{8}$$

where the fitness function is symbolized by f, and $X_i(k)$ represents the position of the i-th member in the k-th iteration. Likewise, $X_{ib}(k)$ denotes the position of the best member in the k-th iteration,

$$EI_i(k) = \frac{f(X_w(k)) - f(X_{ib}(k))}{f(X_{ib}(k))}, \tag{9}$$

where $X_w(k)$ represents the position of the member that has the worst fitness value in kth iteration,

$$II_i(k) = \frac{f(Pbest_i(k)) - f(GBest(k))}{f(GBest(k))}. \tag{10}$$

where $Pbest_i(k)$ denotes position of personal best of ith member and $GBest(k)$ symbolizes position of the global best member in iteration k.

There are three movement policies (MPs) in BASO that govern the update of members' situations. These MPs take into account the current situation, the society's situation, and the past situation. The modified MPs for feature selection problems are presented in (11), (12), and (13), respectively:

$$MP_i^{Current}(k) = \left\{ \begin{array}{ll} \text{Create a random member and apply} & i \neq i^*(k), \\ \text{ a single point crossover with it} & FI_i(k) \leq \beta_1 \\ & \\ \text{Apply single point crossover with} & i \neq i^*(k), \\ \text{a random number } j \text{ where } j \neq i, & FI_i(k) > \beta_1 \\ & \\ \text{Apply single point crossover with the } GBest, & i = i^* \end{array} \right\}, \quad (11)$$

$$MP_i^{Society}(k) = \left\{ \begin{array}{ll} \text{Apply single point crossover with the} & \\ \text{best member of the iteration } (i^*(k)), & EI_i(k) \leq \beta_2 \\ & \\ \text{Apply single point crossover with a} & \\ \text{random member j } (j \neq i), & EI_i(k) > \beta_2 \end{array} \right\}, \quad (12)$$

$$MP_i^{Past}(k) = \left\{ \begin{array}{ll} \text{Apply single point crossover} & \\ \text{with the personal best position,} & II_i(k) \leq \beta_3 \\ & \\ \text{Apply single point crossover} & \\ \text{with a random member j } (j \neq i), & II_i(k) > \beta_3 \end{array} \right\}. \quad (13)$$

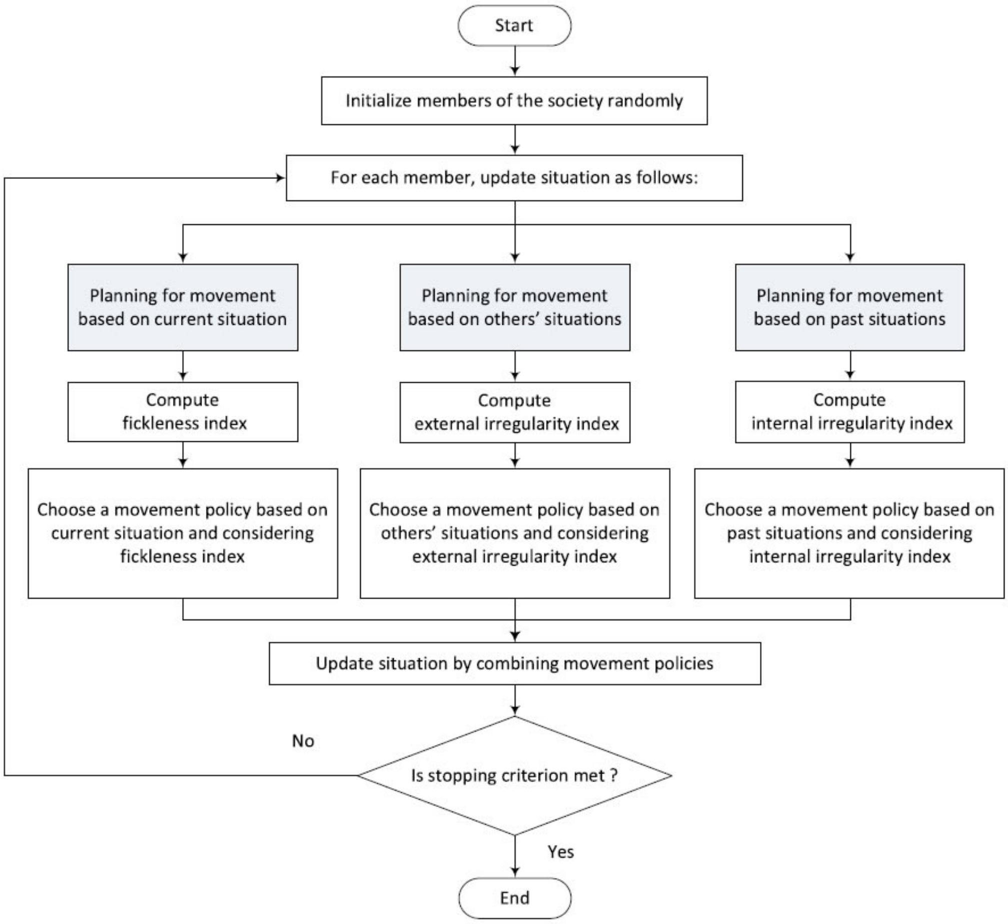

Fig. 1. The general framework of the ASO algorithm [1]

Figure 1 depicts the general framework of the Anarchic Society Optimization (ASO) algorithm. The algorithm begins with the random initialization of members. Their situations are subsequently updated considering the current situations, the situations of others, and past situations. The relevant index is calculated based on the selected situation, and movement policies are employed to update the members' situations.

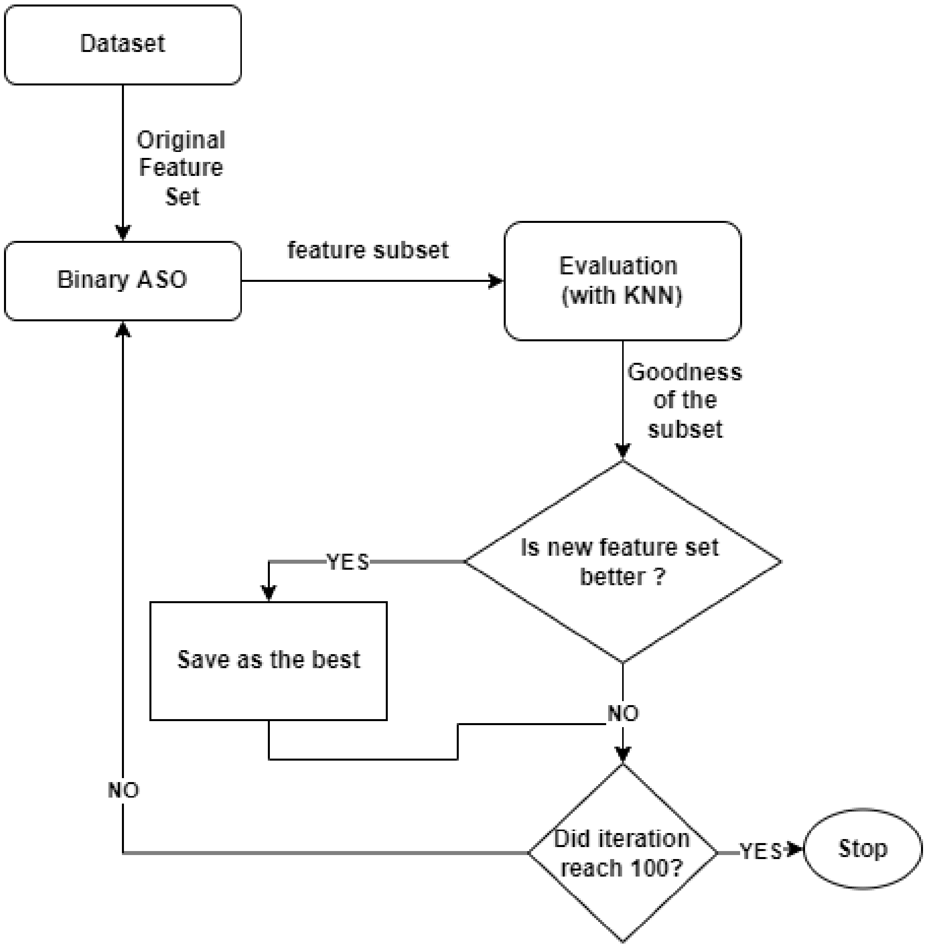

Fig. 2. Flowchart of FS process

Figure 2 presents an overview of the feature selection process utilized in this study. Initially, the original data sets are provided as input to the proposed method. The method evaluates the quality of subsets to determine if they are an improvement. The termination criteria are checked at the end of each iteration.

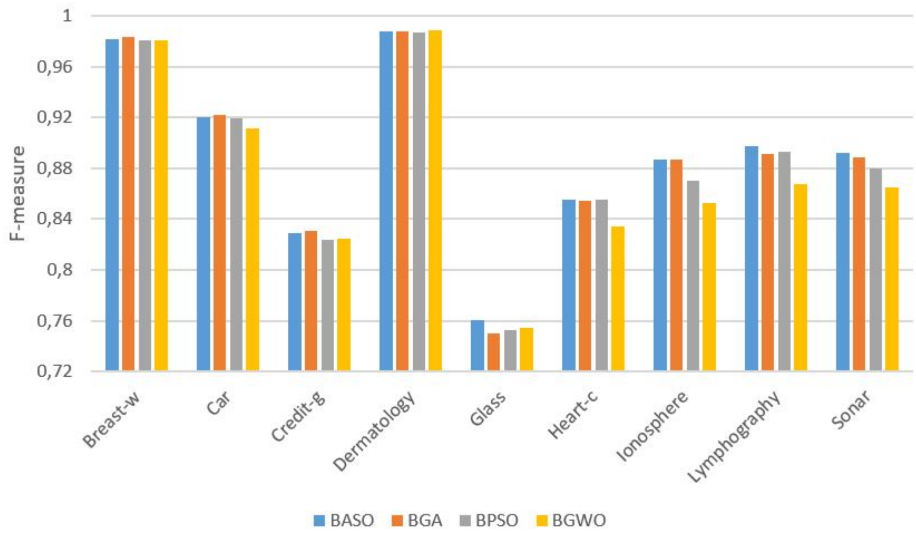

Fig. 3. Average f-measure value for different FS algorithms

Figure 3 and Figure 4 present alternative representations of the average f-measure and average number of selected features. Figure 3 illustrates the competitive nature of the proposed method through the close proximity of f-measure values among the algorithms. BASO performs similarly to other algorithms when dealing with datasets with fewer features. However, when confronted with datasets containing a larger number of features and samples, BASO demonstrates superior performance. This advantage can be attributed to BASO's utilization of past member behavior information, which becomes more effective with an increased number of features and samples. In Figure 4, it is evident that the number of selected features is comparable across the methods, with the exception of BGWO, which tends to select a higher number of features. Although this occasionally places BGWO at the top of the performance list, it may be beneficial to sacrifice a small degree of performance in order to deal with a reduced number of features.

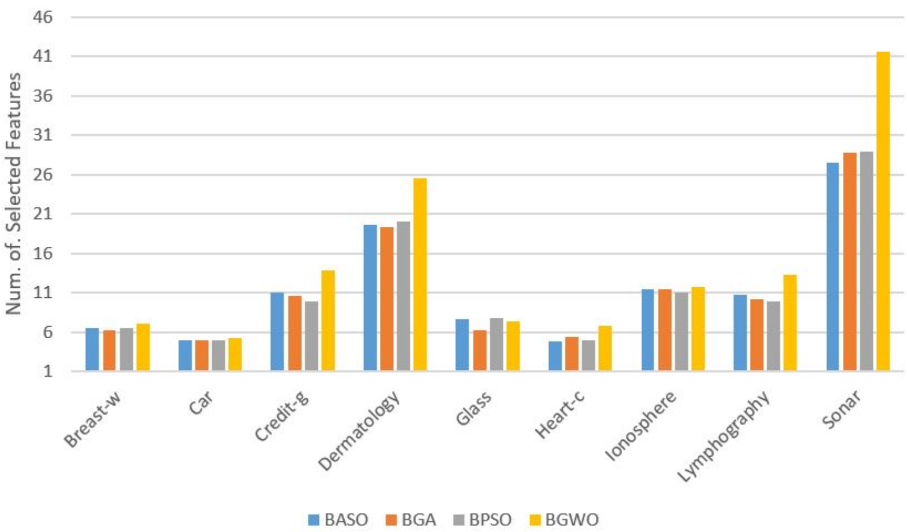

Fig. 4. Average number of selected features for different FS algorithms

Table 4. The best and worst F-measure values acquired from the different algorithms

Data sets	The Best (number of selected features)				The Worst (number of selected features)			
	BASO	BGA	BPSO	BGWO	BASO	BGA	BPSO	BGWO
Breast-w	**0.984(5)**	0.984(6)	0.982(7)	0.982(8)	0.980(8)	**0.982(7)**	0.980(7)	0.980(8)
Car	**0.923(5)**	**0.923(5)**	0.921(5)	0.920(5)	0.919(5)	**0.921(5)**	0.918(5)	0.872(6)
Credit-g	0.835(8)	**0.837(9)**	0.831(9)	0.836(14)	**0.825(12)**	0.819(10)	0.819(14)	0.814(17)
Dermato.	0.992(23)	**0.992 (18)**	0.989 (18)	0.992 (27)	**0.984 (21)**	0.983 (16)	0.984(23)	0.984(26)
Glass	**0.855(8)**	0.771(7)	0.762(8)	0.771(7)	**0.757(8)**	0.724(6)	0.738(7)	0.720(8)
Heart-c	0.866(5)	**0.870(4)**	0.860(4)	0.861(6)	0.849(5)	0.840(6)	**0.852(6)**	0.736(8)
Ionosph.	**0.906(9)**	0.897(10)	0.880(8)	0.882(15)	**0.877(15)**	0.862(18)	0.856(12)	0.838(19)
Lymph.	0.905(10)	**0.905(9)**	**0.905(9)**	0.899(14)	0.878(10)	0.865(13)	**0.878(7)**	0.851(13)
Sonar	0.916(27)	**0.918(23)**	0.889(25)	0.888(36)	**0.877(32)**	0.859(33)	0.871(29)	0.851(47)

The presence of both close best and worst cases, as observed in Table 4, further underscores the competitiveness of the BASO method. The table displays the number of features utilized to achieve the respective performance, indicated in parentheses. While the BGA method appears to have achieved most of the best cases according to the table, this does not necessarily reflect its overall performance. Randomness inherent in the algorithms can occasionally lead some methods to achieve the best score by chance. To ensure a comprehensive assessment and avoid reliance on chance, the methods are executed multiple times, and the average score of the runs is considered. Across nearly all datasets, BGWO selects the largest number of features. The best of the best cases and the best of the worst cases are highlighted in bold.

Table 5. Average f-measure and the number of selected features using selected feature by standard FS methods (CHI, IG, GR) and BASO

Data sets	Average F-measure					Average No. of selected features				
	BASO	CHI	IG	GR	ReliefF	BASO	CHI	IG	GR	ReliefF
Breast-w	**0.982**	0.961	0.971	0.960	0.967	6.5	7	7	7	7
Car	0.920	**0.937**	**0.937**	**0.937**	**0.937**	5	5	5	5	5
Credit-g	**0.829**	0.719	0.719	0.703	0.707	11.1	12	12	12	12
Dermato.	**0.988**	0.857	0.943	0.914	0.918	19.7	20	20	20	20
Glass	**0.778**	0.748	0.748	0.748	0.735	7.6	8	8	8	8
Heart-c	**0.857**	0.775	0.775	0.791	0.821	4.8	5	5	5	5
Ionosph.	0.887	0.879	0.879	**0.897**	0.865	11.4	12	12	12	12
Lymph.	**0.896**	0.763	0.774	0.777	0.843	10.8	11	11	11	11
Sonar	**0.892**	0.870	0.874	0.846	0.859	27.5	28	28	28	28

To facilitate a performance comparison between traditional FS methods and BASO, we have created Table 5. The table reveals that, with the exception of the car and ionosphere datasets, BASO demonstrates superior performance over CHI, IG, GR, and ReliefF. It is worth noting that when the number of features selected by BASO is not an integer, we have rounded it up for the traditional FS methods. This rounding may confer certain advantages to the traditional methods.

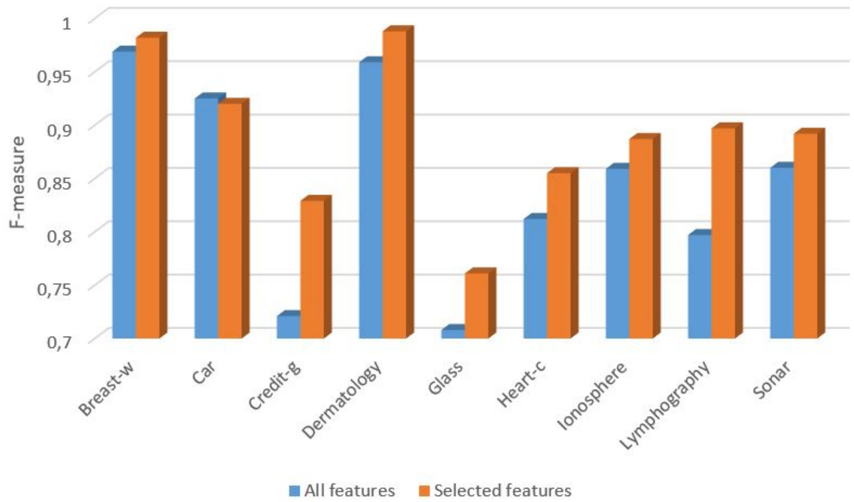

Fig. 5. F-measure values without FS and with FS using proposed BASO method

Figure 5 illustrates the f-measure obtained from both the dataset with all features and the dataset with selected features. The figure demonstrates that the proposed method enhances the f-measure value while reducing the number of features for all datasets, except for the car dataset.

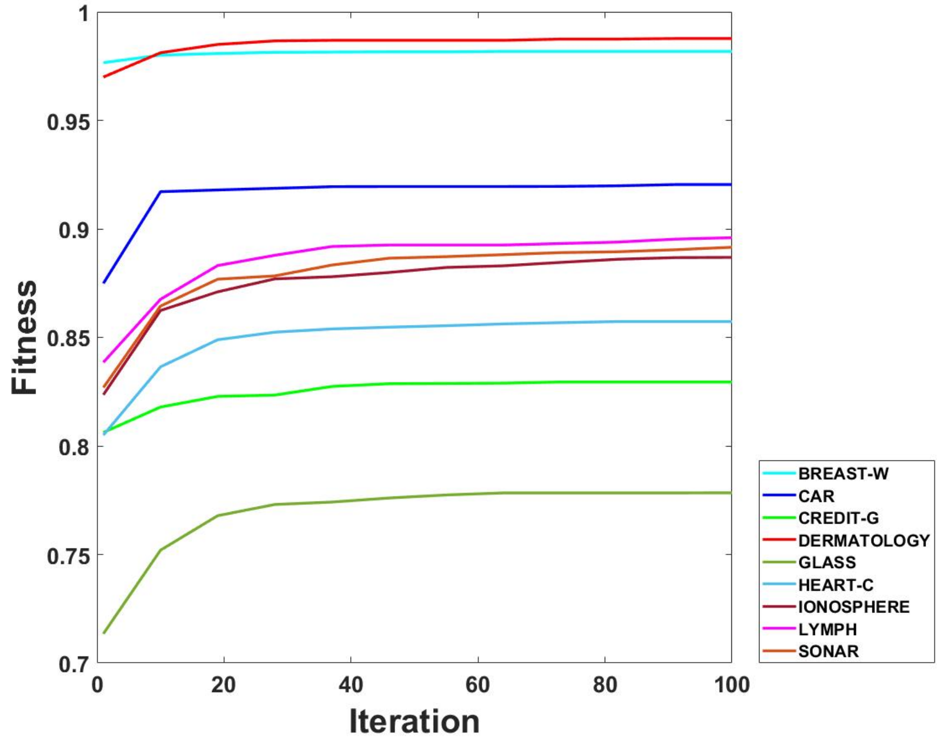

Fig. 6. The convergence curve of proposed method

The convergence of the proposed method was also evaluated and depicted in Figure 6. The figure demonstrates the rapid convergence of the method in terms of the number of iterations. Furthermore, it highlights the point that selecting an excessively high iteration number may not be beneficial, as the F-measure values plateau after approximately 20 iterations.

References

[1] A. AHMADI-JAVID and P. HOOSHANGI-TABRIZI, *Integrating employee timetabling with scheduling of machines and transporters in a job-shop environment: A mathematical formulation and an Anarchic Society Optimization algorithm*, Computers & Operations Research, **84**, pp. 73-91, 2017.