Al-Azhar Bulletin of Science

Volume 29 | Issue 2 Article 23

12-1-2018

Section: Mathematics, Statistics, Computer Science, Physics and Astronomy

GREY WOLF OPTIMIZER WITH V-SHAPED TRANSFER FUNCTION FOR ATTRIBUTE SELECTION AND CLASSIFICATION

Heba Eid
Faculty of Science, Al - Azhar University, Cairo, Egypt, eid.heba@gmail.com

Follow this and additional works at: https://absb.researchcommons.org/journal

Part of the Life Sciences Commons

How to Cite This Article

Eid, Heba (2018) "GREY WOLF OPTIMIZER WITH V-SHAPED TRANSFER FUNCTION FOR ATTRIBUTE SELECTION AND CLASSIFICATION," *Al-Azhar Bulletin of Science*: Vol. 29: Iss. 2, Article 23.
DOI: https://doi.org/10.21608/absb.2018.61204

This Original Article is brought to you for free and open access by Al-Azhar Bulletin of Science. It has been accepted for inclusion in Al-Azhar Bulletin of Science by an authorized editor of Al-Azhar Bulletin of Science. For more information, please contact kh_Mekheimer@azhar.edu.eg.

Al Azhar Bulletin of Science Vol. 29, No. 2 (December), 2018, pp. 11-17.

GREY WOLF OPTIMIZER WITH V-SHAPED TRANSFER FUNCTION FOR ATTRIBUTE SELECTION AND CLASSIFICATION

Heba F. Eid

Faculty of Science, Al-Azhar University, Cairo, Egypt;
eid.heba@gmail.com, heba. fathy@azhar.edu.eg

Abstract

One of the most difficult challenges in pattern recognition is data attribute selection process. Attribute selection played a critical role in dealing with high-dimensional problems and can be modeled as an optimization problem. Gery Wolf optimization algorithm is a recent swarm based algorithm, which is competitive to other bio-inspired algorithms. In this paper, a grey wolf optimizer is proposed as a wrapper attribute selection method. A V-shaped transfer function is employ to map a continuous search space to a discrete binary search space. Experimental results on eight UCI datasets shows the ability of the novel wrapper attribute selection algorithm in selecting the most informative attributes for classification tasks when compared to three well-known bio-inspired algorithms.

Keywords: Bio-Inspired Optimization; Grey Wolf Optimizer (GWO); Attribute selection; Classification.

1. INTRODUCTION

Attribute selection is a critical procedure in data mining and pattern recognition, which contributes towards boosting the performance of a classification model. For large-scale datasets, large number of attributes may contain a lot of redundacy [1]. Therefor, attribute selection plays a pivotal role to increase the accuracy of the classification models as well as the learning speed.

Optimization literally means finding the best possible solution. To solve real world problems, optimization require enormous computational efforts, which tend to fail as the problem size increases. Swarm-based algorithm refers to the implementation of collective groups of agents that simulate the behavior of real world swarm. For which, the collection of agents interact locally with each other as well as with their environment. This motivate for employing swarm-based algorithms which show higher computational efficiency in avoiding local minima for attribute selection problems [2–5]. Also, the adaptability and self-learning capability of the swarm based algorithms attracted several real world application areas [6 ,7]. For instance, there are attribute selection methods based on bees [8], ants [9], partical swarm optimizatin [10], fishes [11], bats [12] and whale optimization [13].

Grey Wolf Optimizer (GWO) is a recent swarm optimization algorithm proposed by Mirjalili [14]. The GWO mimics the social leadership hierarchy and hunting mechanism of grey wolves swarms in nature. The GWO has been successfully employed for solving attribute selection problems [15].

However, a transfer function is an important part in the binary versions of the optimization. For which, it significantly impacts the balance between exploration and exploitation and the local optima avoidance. In this paper, a V-shaped transfer function is applied to the GWO algorithm to map a continuous search space to a discrete search space. The proposed V-GWO algorithm is employed for solving attribute selection problems. The proposed V-GWO attribute selection algorithm is tested on eight UCI datasets. Experimental results demonstrate the efficiency and superiors of the proposed algorithms in most cases.

The rest of this paper is organized as follows: Section 2 briefly overviews the mathematical concepts of Grey Wolf optimizer.

While Section 3 presents the details of the proposed V-shaped GWO algorithm (V-GWO). Section 4, discusses the proposed V-GWO based attribute selection method. Experimentation design, results and comparative analysis occupy the remainder of the paper in Section. Finally, Section

6 summarizes the main findings of this study.

2 GREY WOLF OPTIMIZER (GWO)

Grey Wolf Optimizer (GWO) is a bio-inspired optimization algorithm; inspired by the grey wolves swarm [14]. It mimics the social leadership hierarchy and hunting mechanism of grey wolves swarms in nature. The population in GWO algorithm is divided into four types: alpha α, beta β, delta δ and omega ω. For which, the formal three types guide the final type toward the optimum areas in the search space.

During optimization process, the wolves update their positions according to:

$$\vec{X}(t+1) = \vec{X}_P(t) - \vec{A}.\vec{D} \qquad (1)$$

$$\vec{D} = \left|\vec{C}.\vec{X}_P(t) - \vec{X}(t)\right| \qquad (2)$$

where, t indicates the current iteration, $\vec{A}$ and $\vec{C}$ are coefficient vectors , $\vec{X}_P$ is the position vector of the prey, and $\vec{X}$ is the position vector of a grey wolf.

The coefficient vectors $\vec{A}$ and $\vec{C}$ are calculated by:

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a} \qquad (3)$$

$$\vec{C} = 2.\vec{r_2} \qquad (4)$$

where, a decreased linearly from 2 to 0, and r_1, r_2 are random vectors in the range [0,1].

During GW optimization, α, β and δ are assumed to be the first three best solutions respectively. While, ω update their positions with respect to α, β and δ, as follows:

$$\vec{X}(t+1) = \frac{\vec{X_1} + \vec{X_2} + \vec{X_3}}{3} \qquad (5)$$

where,

$$\vec{X_1} = \vec{X_\alpha} - \vec{A_1}.\left(\vec{D_\alpha}\right) \qquad (6)$$

$$\vec{X_2} = \vec{X_\beta} - \vec{A_2}.\left(\vec{D_\beta}\right) \qquad (7)$$

$$\vec{X_3} = \vec{X_\delta} - \vec{A_3}.\left(\vec{D_\delta}\right) \qquad (8)$$

and D is the behavior of grey wolves encircling the prey:

$$\vec{D_\alpha} = \left|\vec{C_1}.\vec{X_\alpha} - \vec{X}\right| \qquad (9)$$

$$\vec{D_\beta} = \left|\vec{C_2}.\vec{X_\beta} - \vec{X}\right| \qquad (10)$$

$$\vec{D_\delta} = \left|\vec{C_3}.\vec{X_\delta} - \vec{X}\right| \qquad (11)$$

3 PROPOSED V-SHAPE GREY WOLF OPTIMIZATION (V-GWO)

In the gery wolf optimization, the search agents continuously update their positions to whatever point in the search space; according to the best search agent obtained so far. In this paper, a binary version of the grey wolf optimization is proposed to suit the attribute selection and dimensionality reduction problems. For which, the continues to binary conversion is performed by applying a V-shaped transfer function. The V-shaped transfer function adopted is given by equation 12. The V-shaped transfer function define the probability of changing position vector's elements from "0" to "1" and vice versa, figure 1. Consequently, the V-shaped transfer function restricting the search agents to move in a binary space at any given iteration. For the proposed V-GWO, the search agents update their positions according to equation **Error! Reference source not found.**.

$$\mathbf{S}(\mathbf{x}) = \left|2\pi \arctan\left(\frac{\pi}{2}x\right)\right| \qquad (12)$$

$$X_{t+1}^{d} = \begin{cases} -X_{t'} & rand < S(X_{t+1}) \\ X_{t'} & atherwise. \end{cases} \qquad (13)$$

Where X_{t+1}^{d} is the agent position in dimension *d* at iteration *t*

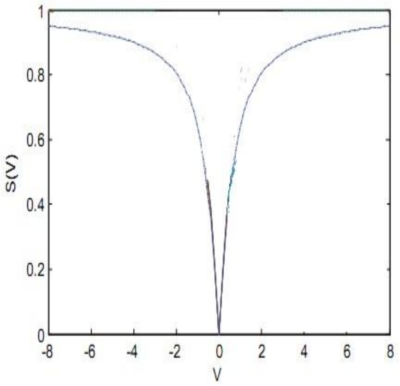

Figure 1: V-shaped transfer function

4 V-SHAPED GWO BASED ATTRIBUTE SELECTION

For attribute selection problems, a solution is limited to the binary space {0,1} and is represented as an N-sized binary vector, where N is the total number of attributes in a dataset. Subsequently, the proposed V-GWO is an appropriate algorithm for attribute selection problems. Since, the wolves position is represented by binary vectors; either "1" indicates the corresponding attribute is selected or "0" indicates the non selection of the attribute.

Attribute selection algorithm has two main criteria; maximizing the classification accuracy and minimizing the number of attributes. The V-GWO is used to adaptively search for the best attribute subset, which considers these two criteria. The fitness function adopted in the proposed V-GWO to evaluate each individual wolf positions is modeled in:

$$Fitness = \alpha ER + (1-\alpha)\frac{|S^*|}{|S|} \qquad (14)$$

where E_R is the classification error rate of the selected attribute, S^* is the number of selected attributes and *S* is the total number of attributes. α and (1−α) represent the relative importance of the classification accuracy and the number of selected attributes, $\alpha \in (0.5,1]$.

5 EXPERIMENTAL DESIGN AND RESULTS

Algorithm 1shows the flow of the proposed V-GWO attribute selection:

Algorithm 1 Proposed V-shaped GWO Attribute Selection Algorithm

Input:
Number of wolves n
Number of iterations for optimization Max_Iter
Output:
Optimal wolves binary position X_α

```
Initialize the n wolves population positions at random ∈ [0, 1].
Initialize a, A and C.
Calculate the fitness of each α, β and δ.
t=1
while t ≤ Max_Iter do
    for each search agent do
        Update the position of the current search agent to a binary position by equation 13
    end for
    Update a, A and C
    Evaluate the fitness of each search agent by equation 14.
    Update α, β and δ if there is better solution
    t=t+1
end while
return X_α
```

The initial controlling parameters of the optimization algorithm GWO, PSO and GA are listed in Table 1.

Table 1: Initial Parameters setting for Experiments.

Algorithm	Parameter	Value
GWO	No of iterations	70
	Population size	5
PSO	Cognitive constant C_1	0.8
	Social constant C_2	0.8
	Inertia constant w	1.2
	No of iterations	150
	Population size	100
GA	Crossover probability	0.9
	Mutation probability	0.1
	Selection mechanism	Roulette wheel
ACO	Information heuristic factor α	1.0
	Expectation heuristic factor β	2.0
	Number of ants	10

5.1 Datasets Description

Eight datasets from the UCI machine learning repository [16] were chosen to evaluate the performance of the proposed V-GWO algorithm; as given in Table2. The 8 datasets were chosen to have various numbers of attributes, classes and instances.

Table 2: Datasets description

Dataset	Attributes no.	Instances no.	Classes no.
Australian	14	690	2
German Credit	24	1000	2
Sonar	60	208	2
Zoo	17	101	7
Diabetic	19	1151	2
Heart Disease	13	270	2
Segment	19	2310	7
Liver Disorders	6	345	2

5.2 Evaluation Criteria

The K-nearest neighbour (KNN) [17] was used in the experiments where K is set to 5 (5NN). During the training phase, the KNN classification error rate used in the fitness function is calculated using 10-fold cross-validation to evaluate the performance of each single attributes subset. For which, the training set is used to evaluate the KNN on the validation set throughout the optimization. Then, the selected attributes are evaluated on the test set to obtain the final evaluation of the selected attributes.

5.3 Results and Analysis of The Proposed V-GWO

To evaluate the performance of the proposed V-GWO attribute selection algorithm, 8 datasets instances are randomly divided into a cross validation manner to three sets: training, validation and test sets. Table 3 represents the experimental results of the proposed V-GWO based attribute selection method on the unseen test sets. According to table 3, it can be seen that the F-measure of the attribute subsets selected by V-GWO is significantly better than using the full attributes. As well as, the number of attributes selected by V-GWO is about half of the total number of original attributes; or less than half of the total number of attributes, as for australian and heart disease dataset.

Table 3: Comparison Results of V-GWO attribute selection Algorithm on different Datasets

	All		BWO	
Dataset	Attributes no.	F-measure	Attributes no.	F-measure
Australian	14	84.34%	6	87.68%
German Credit	24	71.40%	15	80.2%
Sonar	60	83.65%	37	88.94%
Zoo	17	94.05%	9	95.88%
Diabetic	19	61.59%	11	79.41%
Heart Disease	13	78.88%	4	85.18%
Segment	19	95.36%	10	97.27%
Liver Disorders	6	56.23%	3	85.92%

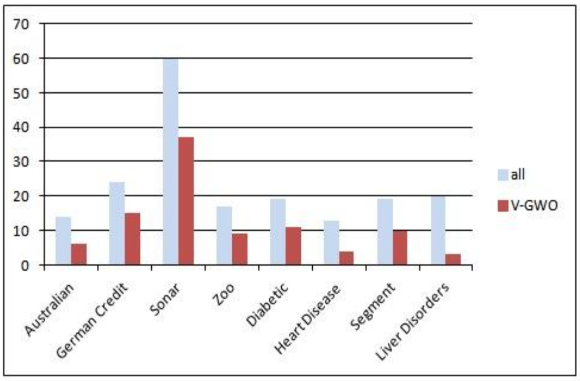

Figure 2: Attributes numbers Comparison Results of V-GWO Algorithm on different Datasets

5.4 Comparative Analysis

To examine the overall performance of the proposed V-GWO attribute selection method, it is compared with other standard optimization algorithms; such as genetic algorithm (GA), particle swarm optimization (PSO) and ant colony optimization (ACO). Table 4 reports the statistical evaluation for 30 independent runs, and for each run the classification accuracy; best accuracy (*Best_Acc*), average accuracy (*Av_Acc*); as well as the standard deviation (*Std*) on the eight datasets are recorded.

As shown in table 4, the V-GWO outperforms all GA, PSO and ACO attribute selection algorithm for the average accuracy on all datasets, except for the segment and liver disorders dataset. Meanwhile, in almost all cases, the standard deviation values of the proposed V-GWO is smaller than that of GA, PSO and ACO, which indicates the stability of the V-GWO attribute selection algorithm compared to the others optimizers.

Table 4: Performance Results of V-GWO, GA , PSO and ACO attribute Selection algorithm on different Datasets

Dataset		V-GWO	GA	PSO	ACO
Australian	Av_Acc	**0.85319**	0.8289	0.8246	0.8390
	Std	**0.0001**	0.0228	0.0731	0.0240
	Best_Acc	**0.94783**	0.8553	0.8744	0.8530
German Credit	Av_Acc	**0.7219**	0.7133	0.6889	0.7081
	Std	**0.0098**	0.0200	0.0207	0.0168
	Best_Acc	**0.8020**	0.7451	0.7333	0.7240
Sonar	Av_Acc	**0.9542**	0.7540	0.7857	0.8130
	Std	**0.0012**	0.0691	0.0346	0.0255
	Best_Acc	**0.9931**	0.8720	0.8571	0.8751
Zoo	Av_Acc	**0.94118**	0.8550	0.9512	0.9406
	Std	**0.0001**	0.0690	0.0646	0.0324
	Best_Acc	**0.9767**	0.9601	0.9714	0.9730
Diabetic	Av_Acc	**0.7556**	0.7504	0.6931	0.6451
	Std	**0.0041**	0.0169	0.0347	0.0394
	Best_Acc	**0.78125**	0.7748	0.6897	0.6681
Heart Disease	Av_Acc	**0.8667**	0.7801	0.7700	0.8260
	Std	**0.0007**	0.0210	0.0360	0.0240
	Best_Acc	**0.9185**	0.9102	0.9059	0.8871
Segment	Av_Acc	0.9407	0.9150	**0.9431**	0.9152
	Std	**0.0008**	0.0177	0.0147	0.0167
	Best_Acc	**0.9688**	0.9515	0.9521	0.9462
Liver Disorders	Av_Acc	0.6879	0.6780	**0.7030**	0.6120
	Std	**0.0012**	0.0524	0.1263	0.0460
	Best_Acc	0.7052	0.7373	**0.7573**	0.6551

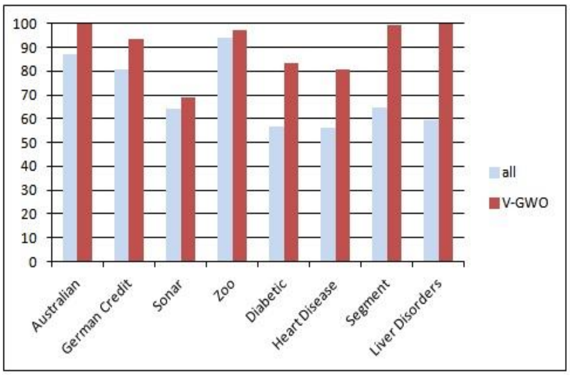

Figure 3: SVM Comparison Results of V-GWO Algorithm on different Datasets

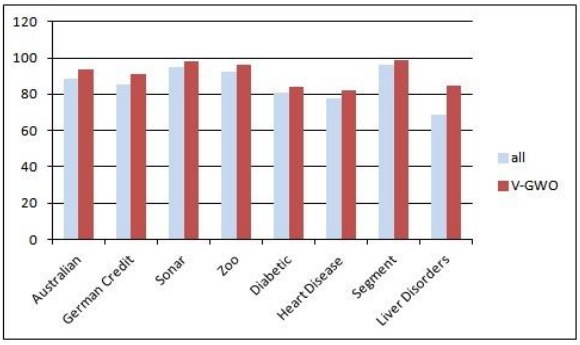

Figure 4: J48 Comparison Results of V-GWO Algorithm on different Datasets

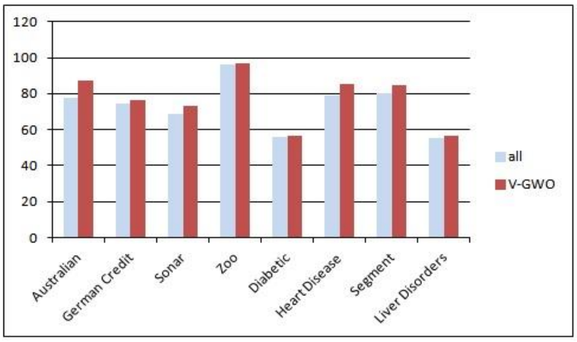

Figure 5: NB Comparison Results of V-GWO Algorithm on different Datasets

6 CONCLUSION

This paper proposed a bio-inspired algorithms named V-GWO based on grey wolf optimization for solving attribute selection problems. In the proposed V-GWO, a V-shaped transfer function is adopted to map a continuous search space to a binary search space. Eight UCI datasets were employed to verify the performance of the proposed V-GWO algorithm. Experimental results illustrates that the proposed V-GWO algorithms provide highly competitive results when tested using SVM, NB and J48 classifiers. Moreover, comparative results on the eight UCI datasets reveals that the proposed V-GWO is able to out perform three state-of-the-art attribute selection algorithms; PSO, GA and ACO.

REFERENCES

[1] B. Xue, M. Zhang, W. Browne, X. Yao, A survey on evolutionary computation approaches to feature selection, IEEE Transactions on Evolutionary Computation 20 (2016) 606–626.

[2] M. Crepinek, S.-H. Liu, M. Mernik, Exploration and exploitation in evolutionary algorithms: a survey., ACM Comput Surv (CSUR) 45 (2013) 35–42.

[3] X. Gang, An adaptive parameter tuning of particle swarm optimization algorithm., Appl Math Comput 219 (2013) 4560–4569.

[4] Y. Zhang, S.Wang, G. Ji, A comprehensive survey on particle swarmoptimization algorithm and its applications.,Math Probl Eng (2015) doi:10.1155/2015/931256.

[5] S. Nesmachnow, An overview of metaheuristics: accurate and efficient methods for optimisation, International Journal of Metaheuristics 3 (2014) 320–347.

[6] A. H. Gandomi, X. Yang, S. Talatahari, A. H. Alavi, Metaheuristic Algorithms in Modeling and Optimization, Metaheuristic Application in Structures and Infrastructures, 2013.

[7] X.-S. Yang, Random walks and optimization. In: Nature-inspired optimization algorithms, first edition Edition, Elsevier, Oxford, 2014.

[8] L. T. Wang, J., R. A. Ren, A real time idss based on artificial bee colony-support vector machine algorithm, in: Proc. of the the Advanced Computational Intelligence (IWACI), IEEE, 2010, pp. 856–863.

[9] M. Dorigo, M. Birattari, Ant colony optimization, Encyclopedia of Machine Learning, Springer, 2010.

[10] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: Proc. of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 1995, pp. 39–43.

[11] P. Miller, The Smart Swarm: How Understanding Flocks, Schools, and Colonies Can Make Us Better at Communicating, Decision Making, and Getting Things done, Avery Publishing Group, Inc., 2010.

[12] X.-S. Yang, Bat algorithm for multi-objective optimisation., Int. J. Bio-Inspired Comput. 3 (2011) 267–274.

[13] H. F. Eid, Binary whale optimization: an effective swarm algorithm for feature selection, International Journal of Metaheuristics (2018) 67–79.

[14] S.Mirjalili, S.Mirjalili, A. Lewis, Grey wolf optimizer., Advances in Engineering Software 69 (2014) 46–61.

[15] H. F. Eid, A. Abraham, Plant species identification using leaf biometrics and swarm optimization: A hybrid pso, gwo, svm model, International Journal of Hybrid Intelligent Systems 14 (2017) 155–165.

[16] A. Frank, A. Asuncion, Uci machine learning repository, Retrieved from https://archive.ics.uci.edu/ml/index.php; November 2018.

[17] L. Chuang, H. Chang, C. Tu, C. Yang, Improved binary pso for feature selection using gene expression data., Computational Biology and Chemistry 29 (2008) 29–38.

استمثال الذئب الرمادى باعتماد داله نقل V لتخفيض بعد البيانات و التصنيف

هبه فتحى عيد
مدرس علوم الحاسب – كليه العلوم (بنات) – جامعه الازهر

فى هذا البحث تم اقتراح خوارزميه معتمده على grey wolf optimization ، الخوارزميه المقترحة تدعى V-GWO وتستخدم فى تخفيض بعد البيانات عن طريق اختيار الموصفات ذات الصله للبيانات . حيث تم اعتماد V-transfer function لنقل مساحه البحث المستمره الى مساحه بحث ثنائيه ملائمه لمسائل تخفيض ابعاد البيانات. و قد تم استخدام ثماني مجموعات بيانات مختلفه من قاعده البيانات UCI للتحقق من أداء الخوارزمية المقترحه V-GWO واوضحت النتائج التجريبية حصولها على نتائج تنافسية للغاية عند اختبارها باستخدام ثلاث مصنفات SVM و NB و J48 ، كما تم تعضيد فعاليه الخوارزميه المقترحه بمقارنتها بخوارزميات تخفيض البيانات PSO, GA , ACO .