Convergence Analysis of Whale Optimization Algorithm

Wentao Feng*
National Key Laboratory of Science and Technology on Blind Signal Processing, Chengdu, Sichuan, 610041, China
*wentaofeng@163.com

Abstract. The whale optimization algorithm (WOA) has been widely used in different applications. It has simple control parameters and novel population updating mechanism. However, there is few theoretical analysis of WOA and the convergence property of WOA is ambiguity. This paper analyzes the convergence property of WOA by using the Markov chain of the stochastic process theory. The Markov chain model of the WOA algorithm is established. The one step transition probabilities and convergence properties of different population updating mechanisms in WOA are given. It’s proved that the convergence property of WOA is determined by its shrinking encircling mechanism. Finally, three algorithms with different population updating mechanisms are tested with thirteen benchmark functions on accuracy and convergence speed. The simulation results on benchmark functions verify the validity of the theoretical analysis of WOA.

Keywords: Whale Optimization Algorithm, Markov Chain, One Step Transition Probability, Convergence in Probability, Global Convergence

1. Introduction
Swarm intelligence algorithms [1] such as particle swarm algorithm [2], artificial bee colony [3] and ant colony optimization [4] are an important category of stochastic optimization algorithms. Swarm intelligence algorithms imitate the social behavior of animal organisms in the world. They are effective to deal with difficult optimization problems which are dynamic, nonlinear and high dimensional.

The whale optimization algorithm (WOA) is a novel algorithm introduced in 2016 [5]. WOA has been widely applied in different engineering applications such as breast X-ray image adaptive enhancement [6], 6 degree of freedom (DOF) robot manipulator [7], antenna design problem with a fractal structure [8], Virtual machine scheduling in the cloud [9], and so on. However, WOA has poor convergence speed and it is easily trapped into local optima [10]. Many researchers have developed different types of WOA to improve its accuracy and convergence speed. Luo et al. proposed a modified WOA called RWOA to speed up the convergence rate of WOA [11]. In [11], the time complexity and space complexity of RWOA are analyzed compared with original WOA algorithm. Sun et al. investigated the influences of control parameter \( a \) of WOA algorithm over the course of iterations [12]. They found that the exploitation ability and exploration ability are unbalanced which will leads to premature convergence. Guo et al. designed an improved WOA algorithm named SMWOA to overcome the shortcomings of WOA [13]. In [13], the control parameter \( p \) is linearly modified to control the proportion of exploration and exploitation process.
Though the WOA algorithm is modified with numerous improvement strategies, the convergence property of WOA is ambiguity and the mathematical analysis of WOA is still quite weak. In [14], a quantum whale optimization algorithm (QWOA) is proposed and applied in solving Job-Shop scheduling problem. The computational complexity analysis and global convergence proof of QWOA are provided. However, the convergence property of standard WOA algorithm is still ambiguity.

Markov chain in stochastic process theory has been used to analyze the global convergence property of bean optimization algorithm [15], fireworks algorithm [16] and brain storm optimization [17] successfully. In this paper, the convergence property of WOA is analyzed by using a discrete-time Markov chain of the stochastic process theory. The Markov chain modal of WOA is established at first. The one step transition probabilities and convergence properties of different population updating mechanisms in WOA are discussed. Finally, the convergence property of WOA is proved.

The rest of paper is organized as follows: the principle of standard WOA is described in Section II. In section III, we analyze the convergence property of WOA based on the Markov chain. The performances of three WOA algorithms with different population updating mechanisms are tested with 13 benchmark functions in Section IV. Finally, the main findings are provided in Section V.

2. Whale optimization algorithm

WOA uses three different mechanisms to update the individuals of the population. These mechanisms which are affected by a random number $p$ and coefficient vector $\text{A}$ are spiral updating position, encircling prey and search for prey.

The coefficient vector $\text{A}$ is calculated as follows:

$$\text{A} = 2a \cdot \bar{r} - a$$  \hfill (1)

Where $\bar{r} \in \text{random}[0,1]$, $a$ is decreased linearly from 2 to 0 with the iteration number increased to the end.

When $p \geq 0.5$ and $|\text{A}| \geq 1$, WOA will execute search for prey behavior:

$$\bar{X}(t + 1) = \bar{X}_{\text{rand}}(t) - \text{A} \cdot \bar{C} \cdot \bar{X}_{\text{rand}}(t) - \bar{X}(t)$$  \hfill (2)

Where $\bar{X}_{\text{rand}}(t)$ is a position vector of another individual which is selected from population randomly, $\bar{C}$ is a random vector from 0 to 2.

When $p \geq 0.5$ and $|\text{A}| < 1$, the individual will be updated according to encircling prey behavior. The mathematical model is described by

$$\bar{X}(t + 1) = \bar{X}_{\text{best}}(t) - \text{A} \cdot \bar{C} \cdot \bar{X}_{\text{best}}(t) - \bar{X}(t)$$  \hfill (3)

Where $\bar{X}_{\text{best}}(t)$ is the best individual in current population.

Finally, the mathematical model of spiral updating position is defined as follows:

$$\bar{X}(t + 1) = \bar{X}_{\text{best}}(t) - \bar{X}(t) \cdot e^{bl} \cdot \cos(2\pi l) + \bar{X}_{\text{best}}(t)$$  \hfill (4)

Where $b$ is a constant, $l$ is a random number between $[-1,1]$.

The flowchart of population updating mechanism of standard WOA algorithm is shown in Fig.1.
3. Convergence analysis of WOA

WOA will not execute the search for prey behavior in the later of iteration because the value of $|A|$ is always smaller than 1[12]. Therefore, the convergence property of WOA is influenced by encircling prey and spiral updating position behavior together.

Same as other population-based swarm intelligence algorithms, WOA can be discretized and analyzed in discrete states similar to DE [18].

Let’s suppose that $X^{(t)} = x^{t}, x^{t} \in D$ represents a population state while $D$ is the state space at iteration $t$.

Theorem 1: The population serial $\{X^{(t)}, t = 1, 2, \ldots\}$ of WOA is a finite homogeneous Markov chain.

Proof: The population serial $\{X^{(t)}, t = 1, 2, \ldots\}$ of WOA is a stochastic process in discrete state. The state of $X^{(t+1)} = x^{(t+1)}$ is dependent only on the state of $X^{(t)} = x^{(t)}$ according to the population updating mechanism of WOA. Thus $\{X^{(t)}, t = 1, 2, \ldots\}$ is a Markov chain.

It’s obviously that the one step transition probability of $\{X^{(t)}, t = 1, 2, \ldots\}$ has nothing to do with the iteration number thus this Markov chain is homogeneous. In addition, the solution space of WOA algorithm in discrete state is constrained in a finite set. The population size of WOA is finite as well. Therefore, this Markov chain is finite.

Theorem 2: The population serial $\{X^{(t)}, t = 1, 2, \ldots\}$ of WOA is an absorbing Markov process.

Proof: The individual of current population updates its position around the global optimal solution $x_{best}^{(t)}$. If the optimal solution $x_{best}^{(t+1)}$ in next iteration is better than $x_{best}^{(t)}$, the optimal solution $x_{best}^{(t)}$ will be
replaced by $x_{best}^{(t+1)}$. Suppose that The WOA algorithm has found the global optimal solution $x_{best}^{(t)} \in G^*$ at iteration $t$ where $G^*$ is the global optimal solution set, the one step transition probability is:

$$P[X^{(t+1)} \in G^* \mid X^{(t)} \in G^*] = 1$$  (5)

Thus this population serial is an absorbing Markov process.

The WOA algorithm which contains the spiral updating mechanism alone is named WOASU. Hence we have

**Theorem 3:** If the $ith$ individual of WOASU is trapped in the local optima $lp^{(i)}$ in the $tth$ iteration, the one step transition probability of population serial $\{X^{(t)}, t=1,2,\ldots\}$ is

$$P(X^{(t+1)} = lp^{(i+1)} \mid X^{(t)} = lp^{(i)}) = \begin{cases} 1, & lp^{(i)} = lp^{(i+1)} \\ 0, & lp^{(i)} \neq lp^{(i+1)} \end{cases}$$  (6)

**Proof:** If the WOA algorithm only contains the spiral updating mechanism, the $ith$ individual in population serial $\{X^{(t)}, t=1,2,\ldots\}$ can be updated by

$$X_i^{(t+1)} = lp^{(i)} + |lp^{(i)} - X_i^{(t)}|e^{bi} \cos(2\pi l)$$  (7)

Thus the one step transition probability is given by

$$P(X_i^{(t+1)} = lp^{(i+1)} \mid X_i^{(t)} = lp^{(i)}) = \frac{P(X_i^{(t)} = lp^{(i)})}{P(X_i^{(t)} = lp^{(i)})}$$

$$= \frac{P(X_i^{(t)} = lp^{(i)} + |lp^{(i)} - X_i^{(t)}|e^{bi} \cos(2\pi l), X_i^{(t)} = lp^{(i)})}{P(X_i^{(t)} = lp^{(i)} + |lp^{(i)} - X_i^{(t)}|e^{bi} \cos(2\pi l), X_i^{(t)} = lp^{(i)})}$$

$$= \begin{cases} 1, & lp^{(i)} = lp^{(i+1)} \\ 0, & lp^{(i)} \neq lp^{(i+1)} \end{cases}$$  (8)

It can be seen that the one step transition probability from the individual $X_i^{(t)}$ in state $lp^{(i)}$ to the individual $X_i^{(t+1)}$ in state $lp^{(i+1)}$ of WOASU is 1.

Thus WOASU is easily trapped in the local optima.

**Theorem 4:** The WOASU algorithm can’t converge in probability to the global optimal solution.

**Proof:** Let $G^*$ be a closed global optimal solution set, $G$ is a closed local optimal solution set and $G^* \cap G = \Phi$. Suppose that all of the individuals in WOASU are trapped in the local optima $lp^{(i)}$ where $lp^{(i)} \in G$ at iteration $t$. According to the Theorem 3, the one step transition probability of the Markov chain that the individual $X_i^{(t)}$ to the individual $X_i^{(t+1)}$ is $P[X_i^{(t+1)} \in G \mid X_i^{(t)} \in G] = 1$. It means that WOASU can’t jump out of $G$ because it is an absorbing Markov process. Thus the convergence
probability that the population serial \( \{X^{(t)}, t=1,2,\ldots\} \) reaches the global optimal solution set after infinite iterations is

\[
\lim_{t \to \infty} P\{X^{(t)} \cap G^* \neq \emptyset\} < 1
\]  

(9)

Therefore, WOASU can’t converge in probability to the global optimal solution.

The WOA algorithm which only contains the encircling prey mechanism is named of WOAEP. Hence we have:

Theorem 5: If the \( i \)th individual of WOAEP is trapped in the local optima \( lp^{(t)} \) in the \( k \)th iteration, the one step transition probability of population serial \( \{X^{(t)}, t=1,2,\ldots\} \) is

\[
P(X^{(t)} = lp^{(t+1)} | X^{(t)} = lp^{(t)}) = \begin{cases} 
1, & \text{if } A = 0 \text{ or } C = 1 \\
0, & \text{if } A \neq 0 \text{ and } C \neq 1
\end{cases}
\]  

(10)

Proof: If the WOA algorithm only contains the encircling prey mechanism, the \( i \)th individual can be updated by

\[
X^{(t+1)} = lp^{(t)} - A \cdot |C| \cdot lp^{(t)} - X^{(t)}
\]  

(11)

Thus the one step transition probability of \( X^{(t)} \) is given by

\[
P(X^{(t+1)} = lp^{(t+1)} | X^{(t)} = lp^{(t)}) = \frac{P(X^{(t+1)} = lp^{(t+1)} \, | \, X^{(t)} = lp^{(t)}, X^{(t)} = lp^{(t)})}{P(X^{(t)} = lp^{(t)})}
\]  

(12)

According to the equation (12), we have \( lp^{(t+1)} = lp^{(t)} \) when \( A = 0 \) or \( C = 1 \). If \( A \neq 0 \) and \( C \neq 1 \), we have \( lp^{(t+1)} \neq lp^{(t)} \) which means that the WOAEP will jump out of local optima in next iteration. Thus the one step transition probability of WOAEP is

\[
\frac{P(X^{(t+1)} = lp^{(t+1)}, X^{(t)} = lp^{(t)})}{P(X^{(t)} = lp^{(t)})} = \begin{cases} 
1, & \text{if } A = 0 \text{ or } C = 1 \\
0, & \text{if } A \neq 0 \text{ and } C \neq 1
\end{cases}
\]  

(13)

The proof is complete.

Theorem 6: The WOAEP can converge in probability to the global optimal solution.

Proof: Let \( G^* \) is the closed global optimal solution set, suppose that the state of the \( i \)th individual \( X^{(t)} \) of WOAEP is in the optimal set \( G^* \). Hence we have

\[
P\{X^{(t)} \in G^*\} = P\{X^{(t)} \in G^* \mid X^{(t-1)} \in G^*\} P\{X^{(t-1)} \in G^*\} + P\{X^{(t)} \in G^* \mid X^{(t-1)} \notin G^*\} P\{X^{(t-1)} \notin G^*\}
\]  

(14)

According to the Theorem 2, the population serial \( \{X^{(t)}, t=1,2,\ldots\} \) of WOA is an absorbing Markov process. So we have \( P(X^{(t)} \in G^* \mid X^{(t-1)} \in G^*) = 1 \). The one step transition probability of the \( i \)th
individual $X^{(t)}_i$ from a normal state out of optimal set $G^*$ to the state of optimal set $G^*$ at iteration $t$ is defined as follows:

$$e_t(t - 1) = P[X^{(t)}_i \in G^* | X^{(t-1)}_i \notin G^*]$$

(15)

Hence, we have $0 < e_t(t - 1) < 1$ according to the Theorem 5. The equation (14) can be modified as

$$1 - P[X^{(t)}_i \in G^*] = 1 - P[X^{(t-1)}_i \in G^*] - e_t(t - 1)*P[X^{(t-2)}_i \notin G^*] = (1 - e_t(t - 1))(1 - P[X^{(t-1)}_i \in G^*])$$

(16)

Then

$$1 - P[X^{(t-1)}_i \in G^*] = (1 - e_t(t - 2))(1 - P[X,(t - 2) \in G^*])$$

$$1 - P[X^{(t-2)}_i \in G^*] = (1 - e_t(t - 3))(1 - P[X,(t - 3) \in G^*])$$

... 

(17)

Applying the equation (17) in (16), we have

$$1 - P[X^{(t)}_i \in G^*] = (1 - e_t(t - 1))(1 - e_t(t - 2))...(1 - P[X^{(0)}_i \in G^*])$$

(18)

Then we can conclude that

$$\lim_{t \to \infty} (1 - P[X^{(t)}_i \in G^*]) = \lim_{t \to \infty} (1 - e_t(t - 1))(1 - e_t(t - 2))...(1 - P[X,(0) \in G^*]) = 0$$

(19)

Where $0 < e_t(k) < 1, \ k = 1, 2, ..., t - 1$.

Hence we have $\lim_{t \to \infty} P[X^{(t)}_i \in G^*] = 1$. So it can be concluded that WOAEP can converge in probability to the global optimal solution.

Theorem 7: WOA can converge to the global optimal solution in probability.

Proof: WOA algorithm executes spiral updating position or encircling prey behavior with a probability of 50% in the second half of iteration according to the Fig.1. Even though WOA is trapped into the local optimal due to the operation of spiral updating mechanism, it can jump out of local optimum soon by executing the encircling prey mechanism. Therefore, WOA can converge to the global optimal solution in probability in approximate infinite search iteration.

4. Simulation and results discussion

The validity of the theoretical analysis of WOA is verified on a set of 13 benchmark functions from $f_1$ to $f_{13}$ in [5]. These functions include unimodal and multimodal functions simultaneously. The performance of original WOA is compared with WOAEP and WOASU algorithms.

The common parameter settings of three algorithms are: the population size is 30, maximum iteration number is 500, number of dimensions is 30, and number of independently runs is 30. The simulation results based on the best, worst, mean, and standard deviation values are displayed in Table 1.

It can be seen from Table 1 that the WOAEP algorithm outperforms WOA and WOASU on 12 benchmark functions except $f_5$. WOASU can’t converge to the global optima at all times. WOAEP obtains lower results of best value on function $f_5$ than WOA algorithm while the mean values of them are similar.
Fig.2 to Fig.4 show the convergence curves of three algorithms on some benchmark functions. WOASU algorithm is easily trapped in the local optima. WOAE algorithm has the advantage of high convergence accuracy and fast convergence speed. The performance of WOA is medium between that of WOAE and WOASU algorithms.

| Iteration Number | WOAEP | WOASU | WOA |
|------------------|-------|-------|-----|
|                  | Best  | Worst | Mean | Best  | Worst | Mean | Best  | Worst | Mean |
| 500              | 9.5e+03 | 9.76e+04 | 4.95e+04 | 4.08e+04 | 1.08e+05 | 7.06e+04 | 3.01e+04 | 6.84e+04 | 4.14e+04 |
| 500000           | 1.16e+06 | 5.40e+15 | 2.75e+16 | 2.39e+04 | 7.08e+04 | 4.30e+04 | 2.37e+14 | 8.64e-05 | 7.27e-06 |

Tab.1 The simulation results of WOA, WOAE and WOASU on 13 benchmark functions.

Fig.2 Convergence plots of WOA, WOAE and WOASU on function $f_1$ (left) and $f_2$ (right).
Fig. 3 Convergence plots of WOA, WOAEP and WOASU on function $f_7$ (left) and $f_8$ (right).

Fig. 4 Convergence plots of WOA, WOAEP and WOASU on function $f_{10}$ (left) and $f_{13}$ (right).

Three algorithms can’t find the global optimum on function $f_3$ within 500 iterations as can be seen in Tab.1. The performances of these algorithms are simulated when the iteration number is increased to 50000. The simulation results are shown in Tab.2 and Fig.5. The convergence accuracy of WOAEP is improved obviously with the iteration number increased from 500 to 50000 while the performance of WOASU is invariant.

Fig. 5 Convergence plots of WOA, WOAEP and WOASU on function $f_3$ of iteration 500 (left) and 50000 (right).

5. Conclusion
The convergence property of WOA is analyzed based on the Markov chain in this paper. We have proved that WOA can converge in probability to the global optimal solution. The encircling prey mechanism is critical for WOA to get rid of the local optima. The simulation results on thirteen benchmark functions show that there are significant differences in performance between those WOA algorithms which have different population updating mechanism.

Research on the convergence property of WOA in this paper can provide some guidelines for the improvement of WOA. Further study on how to balance the exploration ability and exploitation ability by adjusting the scope of control parameters is studied in the future.

Acknowledgements

This study was supported by the Foundation of National Key Laboratory of Science and Technology on Blind Signal Processing (No. 614241302070417).

References

[1] Kennedy J, Eberhart R C and Yuhui S 2001 Swarm Intelligence (San Francisco: Morgan Kaufmann)
[2] Kennedy J and Eberhart R 1995 Particle Swarm Optimization Proc. Int. Conf. on Neural Networks (Piscataway NJ: IEEE Service Center) pp 1942-1948
[3] Karaboga D and Basturk B 2007 A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm J Glob Optim 39 pp 459–471
[4] Dorigo M, Birattari M and Stutzle T 2006 Ant colony optimization IEEE Comput Intell M 1 pp 28–39
[5] Mirjalili S and Lewis A 2016 The whale optimization algorithm Adv Eng Softw 95 pp 51-67
[6] Changjiang Z and Huanhuan N 2019 An adaptive enhancement method for breast X-ray images based on the nonsubsampled contourlet transform domain and whale optimization algorithm Med boil Eng Comput 57 pp 2245-2263
[7] Libin Z, Jiacai W, Jiaoliao C, Kang C, Bangyang L and Fang X 2019 Dynamic modeling for a 6-DOF robot manipulator based on a centrosymmetric static friction model and whale genetic optimization algorithm Adv Eng Softw 135 pp 1-9
[8] Singh G and Singh A P 2019 On the design of planar antenna using Fibonacci word fractal geometry in support of public safety Int. J. RF Microw C E 29 pp 1-7
[9] Rana N, Latiff M S A and Abdulhamid S M 2018 A cloud-based conceptual framework for multi-objective virtual machine scheduling using whale optimization algorithm Int J. Innov comput i 8 pp 53-58
[10] Mohammed H M, Umar S U and Rashid T A 2019 A systematic and meta-analysis survey of whale optimization algorithm Comput Intel Neurosc pp 1-25
[11] Qifang L, Ying L and Yongquan Z 2019 Modified whale optimization algorithm for infinitive impulse response system identification Arab J. Sci Eng 45 pp 1-14
[12] Yongjun S, Tong Y and Zujun L 2019 A whale optimization algorithm based on quadratic interpolation for high-dimensional global optimization problems Appl Soft Comput 85 pp 1-20
[13] Wenyan G, Ting L, Fang D and Peng X 2020 An improved whale optimization algorithm for forecasting water resources demand Appl Soft Comput 86 pp 1-18
[14] Xu Y, Chunming Y and Yuanyuan Y 2019 Solving job-shop scheduling problem by quantum whale optimization algorithm Application Research of Computers 36 pp 975-979
[15] Xiaoming Z, Halei W, Bingyu S, Wenbo L and Rujing W 2013 The Markov model of bean optimization algorithm and its convergence analysis Int. J. Comput Int Sys 6 pp 609-615
[16] Jianhua L, Shaqiu Z and Ying T 2014 Analysis on global convergence and time complexity of fireworks algorithm IEEE Congress on Evolutionary Computation (Beijing, China: IEEE) pp 1-7
[17] Ziwei Z and Haibin D 2016 Convergence analysis of brain storm optimization algorithm IEEE Congress on Evolutionary Computation (Vancouver, BC, Canada: IEEE) pp 1-6
[18] Zhongbo H 2014 *The Theoretical Researches and Algorithmic Design of Convergent Differential Evolution Algorithm in Probability* (Wuhan: Wuhan University of Technology)