Estimation the Renewal Function based on Rayleigh distribution Using Firefly Algorithm

Muthanna Subhi Sulaiman
University of Mosul, College of Computer Science and Mathematics, Department of Statistics and Informatics.
E-mail: muthanna.sulaiman@uomosul.edu.iq
Telephone number: +964 7740889735
Shaimaa Waleed Mahmood
University of Mosul, College of Computer Science and Mathematics, Department of Statistics and Informatics.
E-mail: shaimaa.waleed@uomosul.edu.iq
Zakariya Yahya Algamal
University of Mosul, College of Computer Science and Mathematics, Department of Statistics and Informatics.
E-mail: zakariya.algamal@uomosul.edu.iq

Article Info
Page Number: 325-335
Publication Issue:
Vol 71 No. 3s2 (2022)

Article History
Article Received: 28 April 2022
Revised: 15 May 2022
Accepted: 20 June 2022
Publication: 21 July 2022

Abstract

Renewal process modeling is an important topic in stochastic processes to study deteriorating systems in maintenance problems. Rayleigh distribution is one of the most used reliability distributions. In this paper, a firefly algorithm, which is a metaheuristic continuous algorithm, is proposed to estimate the rate of the occurrence of events by estimating the renewal function. The proposed method will efficiently help to estimate the renewal function of the Rayleigh distribution with high performance. The experimental results show the favorable performance of the proposed method comparing with the maximum likelihood and moment estimation methods.
Keywords: Renewal function; Rayleigh distribution; firefly algorithm.

1. Introduction

Renewal process modeling is an important topic in stochastic processes to study deteriorating systems in maintenance problems. This model has been commonly studied in a variety of situations, such as the determination of the optimal replacement policy for deteriorate systems and the analysis of data with trend, which the system before failed would be as good as new, or back as if they were new when the replaced policy use in maintenance [1][2][3].

Renewal Process is useful learning process in the study of Systems Deteriorating. This model is observed in systems that can be repaired or maintained, which were return as if they were new when they were maintenance policy is used. It is also a useful tool for repairable systems in solving maintenance problems, as well as having several uses, such as the identification of the optimal replacement policy and the optimal inspection repair policy. It is also a major tool in the analysis of monotone trend data [4][5].

If-arrival times (i.e. the times between -successive inter process of pointa is $\{N(t); t \geq 0\}$ successive jumps of N), then the process $\{N(t); t \geq 0\}$ is a Renewal process where the random variables $\{N(t)\}$ are positive independent and identically distribution (i.i.d). This process can be determined by three main methods: the first method is by the joint distributions for arrival times. The second method is by the joint distributions of inter-arrival times. And, the last method is by the use of the joint distributions of counting process. In this case $\{N(t)\}$represents the number of arrival times in the time period (0, t].

2. Renewal Process

Renewal process has a function called renewal function. This function is useful for extracting a number of failure associated with the renewal process as a function of time, and it can be express as an iterative model through an integral equation [6][3][7].

Let$\{N(t); t \geq 0\}$ be a process of a series of non-negative independent and identically distribution random variables (i.i.d.), and refer to the inter-arrival times between the two incidents n-1th and nth, and all variables have a distribution function F, the renewal function represents the expected number of events that occurred during the time period, while the process$\{N(t); t \geq 0\}$ represents the run time after repair (n-1)th, the renewal function represents the expected number of cases failure in time period.

Thus, $H(t)$ is a renewal function can be defined as the expected number of accidents occurring during the time period $(0,t]$, i.e:

$$H(t) = \sum_{n=0}^{\infty} nP\{N(t) = n\} \tag{1}$$

Then renewal function becomes

$$\begin{aligned} H(t) &= F(t) + \int_0^t H(t-x)dF(x) \\ &= F(t) + \int_0^t H(t-x)f(x)dx \end{aligned} \tag{2}$$

To drive the renewal function using the direct method, Eq. (2) can be rewrite as:

$$\begin{aligned} H(t) &= E\left[N(t)\right] \\ &= \int_0^t E\left[N(t) \mid X_1 = x\right]dF(x) \end{aligned} \tag{3}$$

The integration equation in Eq. (3) can be estimated on the basis that the first renewal occurs after time t or that the first renewal occurs before time t . The first case can be expressed as follows:

$$E\left[N(t)\mid X_1 = x > t\right] = 0, x > t \tag{4}$$

That is, there is no renewal of a note before time t .

The second case can be express as:

$$E\Big[N(t)\mid X_1 = x \in \big[0,t\big]\Big] = 1 + H\left(t - x\right), x \le t \tag{5}$$

Meaning that the first renewal occurs at time t , and from the definition of the renewal function, the expected number of renewals between case x and time t is $H\left(t - x\right)$. Therefore, the equation of the renewal function using Laplace transformation is defined follows:

$$\begin{aligned} H(t) &= \int_0^t \left[1 + H\left(t - x\right)\right] dF(x) \\ &= \int_0^t dF(x) + \int_0^t H\left(t - x\right) dF(x) \\ &= F(t) + \int_0^t H\left(t - x\right) dF(x) \end{aligned} \tag{6}$$

3. Estimation of Renewal Function

Most of the studies were based on non-parametric estimator in the estimation of renewal function based on the real data of the process of renewal is destined to reach the approximate this function [4][2][8][9].

One of the suggested methods for estimating the $H(t)$ renewal function is by using the Laplace transform, as in the following formula:

$$\hat{H}(t) = \frac{t}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + O(1) \tag{7}$$

where the mean μ and variance σ^2 are two important parameters of the renewal process, and the estimator of these two parameters using the moment method are

$$\hat{\mu} = \overline{x} \tag{8}$$

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \tag{9}$$

4. Estimation the renewal function based on Rayleigh distribution

It is common practise to utilise the Rayleigh distribution to simulate the behaviour of systems with rising failure rates. The two-parameter Rayleigh distribution offers a straightforward but helpful model for lifetime analysis. A continuous probability density function distribution is the Rayleigh distribution. It bears Lord Rayleigh's surname from England. It is essentially a chi distribution with two degrees of freedom. The density function of Rayleigh distribution is:

$$f\,(x\,,\gamma,\delta)=\frac{x-\gamma}{\delta^2}\exp\left(-\frac{(x-\gamma)^2}{2\delta^2}\right),\, x\geq\gamma,\ \delta>0,\ \gamma\in R \tag{10}$$

where γ is the location parameter and δ is the scale parameter. The distribution function is defined as

$$F(x)=1-\exp\left(-\frac{(x-\gamma)^2}{2\delta^2}\right) \tag{11}$$

The mean and variance of Rayleigh distribution is:

$$\mu=\sqrt{\frac{\pi}{2}}\delta+\gamma \tag{12}$$

$$\sigma^2=\left(2-\frac{\pi}{2}\right)\delta^2 \tag{13}$$

From the approximation (non-parametric) solution for renewal function (Eq. (7)), and by compensating the μ and σ^2 for Rayleigh distribution in renewal function, we get

$$H\,(t)=\frac{t}{\sqrt{\frac{\pi}{2}}\delta+\gamma}+\frac{\left(2-\frac{\pi}{2}\right)\delta^2-\left(\sqrt{\frac{\pi}{2}}\delta+\gamma\right)^2}{2\left(\sqrt{\frac{\pi}{2}}\delta+\gamma\right)^2} \tag{14}$$

5. Parameter Estimation of Rayleigh Distribution:

Maximum Likelihood Method

The likelihood function of Rayleigh distribution is

$$L=\prod_{i=1}^{n}\frac{x_i-\gamma}{\delta^2}\exp\left(-\frac{(x-\gamma)^2}{2\delta^2}\right) \tag{15}$$

$$\log L=-2n\log\delta+\sum_{i=1}^{n}\log(x_i-\gamma)-\sum_{i=1}^{n}\frac{(x_i-\gamma)^2}{2\delta^2} \tag{16}$$

$$\frac{\partial\log L}{\partial\delta}=\sum_{i=1}^{n}\frac{(x_i-\gamma)^2}{\delta^3}-\frac{2n}{\delta} \tag{17}$$

$$\frac{\partial \log L}{\partial \gamma} = \frac{\sum_{i=1}^{n} x_i - n\gamma}{\delta^2} - \sum_{i=1}^{n} \frac{1}{(x_i - \gamma)} \tag{18}$$

Then the maximum likelihood estimator of δ is:

$$\hat{\delta}_{Ml} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \hat{\gamma})^2}{2n}} \tag{19}$$

Moment Method:

$$\mu_1 = E(X) = \sqrt{\frac{\pi}{2}}\delta + \gamma \tag{20}$$

$$\mu_2 = E(X^2) = 2\delta^2 + \sqrt{2\pi}\gamma\delta + \gamma^2$$

$$V(X) = \left(2 - \frac{\pi}{2}\right)\delta^2$$

Then the estimator of δ and γ using moment method are

$$\hat{\delta}_{Mo} = \frac{s}{\sqrt{\left(2 - \frac{\pi}{2}\right)}} \tag{21}$$

$$\hat{\gamma}_{Mo} = \bar{X} - \hat{\delta}_{MO}\sqrt{\frac{\pi}{2}} \tag{22}$$

Firefly algorithm

The firefly algorithm (FA) is a type of stochastic, naturally inspired meta-heuristic algorithm that can be used for solving the most challenging optimization problems. It was developed by Xin She Yang at Cambridge University in late 2007 and early 2008 and was based on how a firefly responds to the light of other fireflies [10][11]. The attraction and the fluctuation in light intensity are two crucial factors in this algorithm.
Three idealised rules were adhered to by FA[12].

The brightness (light intensity) of a firefly, which is connected to the objective function, determines how enticing it is. Both light intensity and attractiveness diminish with increasing distance from the light source, which is inversely proportional to distance. We know the light intensity I varies according to the inverse square law

$$I(r) = \frac{I_s}{r^2} \quad (23)$$

where I_s is the light intensity at the source. For a given the medium with a fixed light absorption coefficient γ , the light intensity varies can be determined

$$I = I_0 e^{-\gamma r} \quad (24)$$

where I_0 is the light intensity at initial or original light intensity. Combining the inverse square law and the absorption can be approximated as

$$I(r) = I_0 e^{-\gamma r^2} \quad (25)$$

The attractiveness β is proportional to the light intensity seen by adjacent fireflies as determined by

$$\beta = \beta_0 e^{-\gamma r^2} \quad (26)$$

where β_0 is attractiveness at distance $r = 0$. The distance between two fireflies i and j at positions x_i can be defined using Cartesian distance

$$ri_j = \left\| x_i - x_j \right\| = \sum_{k=1}^{d} \sqrt{(x_i - x_j)^2} \quad (27)$$

The movement of a firefly i is attracted toward the more attractive firefly j is determined by

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2} (x_j - x_i) + \alpha(rand - 0.5) \quad (28)$$

where the first term is the current position of firefly i , "while the second term is due to the firefly's attractiveness and the third term is used for the random movement if there are no any brighter firefly, where $rand$ is a random number generator from 0 to 1 and the coefficient α is randomization parameter where $\alpha \in (0.1)$. For most our implementation we can take $\beta_0 = 1$ and γ parameter which is the light absorption coefficient from 0.1 to 10 [12][13][14]. The firefly algorithm FA can be defined in the following pseudo code as [15][14]:

- Define the objective function $f(x)$, $x = (x_1, x_2, ..., x_d)^T$.
- Generate initial population of fireflies x_i , $(i = 1, 2, ..., n)$.
- Formulate light intensity of firefly I_i at x_i is determined by $f(x)$.
- Define light absorption coefficient γ .
- While $(t < MaxGeneration)$.
- foe $i = 1 : n$ all n fireflies.

- foe $j = 1:n$ all n fireflies (inner loop).
- If $(I_i < I_j)$, move firefly i toward firefly j .
- end if.
- Vary attractiveness with distance r via a $e^{-\gamma r}$.
- 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.
- Post process results and visualization and end procedure".

In this paper we used $\alpha = 0.5, \beta_0 = 0.2, \gamma = 1$ and

$$fitness = \min\left(\max_{1 \le i \le n} \left\{ \frac{\left| T_i - \hat{T}_i \right|}{T_i} \right\} \right) \tag{29}$$

$$T_i = \sum_{j=1}^{i} X_j \qquad and \quad \hat{T}_i = \sum_{j=1}^{i} \hat{X}_j$$

6. Real Application

In this paper, four datasets with different sample sizes for the Mosul gas power station for the period 2010-2013 were used. The operating times are in days. To test that if the four datasets follow Rayleigh distribution according to the hypothesis and the Kolmogorov goodness of fit

$H_0:$ Data follow Rayleigh distribution.

$H_1:$ Data not follow Rayleigh distribution.

Depending on the Kolmogorov goodness of fit test, Table 1 shown that the datasets follow Rayleigh distribution under significant level of 0.05.

Table 1: Test the fit of the data according the Rayleigh distribution

Data	n	Kolmogorov	Critical value	P-value
D1	39	0.20675	0.21273	0.06123
D2	35	0.18489	0.22425	0.16083
D3	42	0.20218	0.20517	0.05558
D4	53	0.17822	0.18311	0.06061

To test the datasets whether they are appropriate for the renewal process according to the following hypothesis

H_0 : Data follow renewal process.

H_1 : Data not follow renewal process.

To apply this test, we find the values of R and L statistical test as shown in Table 2

Table 2: Results test for the renewal process

Data	R	P_R	L	P_L
D1	1.5612	0.1188	1.7982	0.0734
D2	-0.5358	0.5962	0.3897	0.7040
D3	1.3257	0.1868	1.2132	0.2262
D4	1.7110	0.0872	0.5116	0.61

As we can see from Table 2, the P_R , P_L values are greater than 0.05, therefore the datasets are appropriate for the renewal process.

Using maximum likelihood, moment methods, and firefly algorithm for estimating Rayleigh distribution parameters, the results are summarized in Table 3.

Table 3: ML, MO and FA estimators of Raleigh distribution parameters

Data	$\hat{\delta}_{ML}$	$\hat{\gamma}_{ML}$	$\hat{\delta}_{MO}$	$\hat{\gamma}_{MO}$	$\hat{\delta}_{FA}$	$\hat{\gamma}_{FA}$
D1	26.2429	-11.4340	31.4375	-19.7857	18.0759	-11.1656
D2	35.7990	-13.0040	42.4258	-23.6015	27.7680	-25.0000
D3	39.3734	-16.8930	49.0359	-32.6002	36.6072	-34.0000
D4	26.6909	-10.4420	31.7873	-18.6697	27.1601	-20.0000

To find the estimation of the renewal function, the approximate solution of the three methods used to Rayleigh distribution was used and shown with the following plots in Figures 1 - 3.

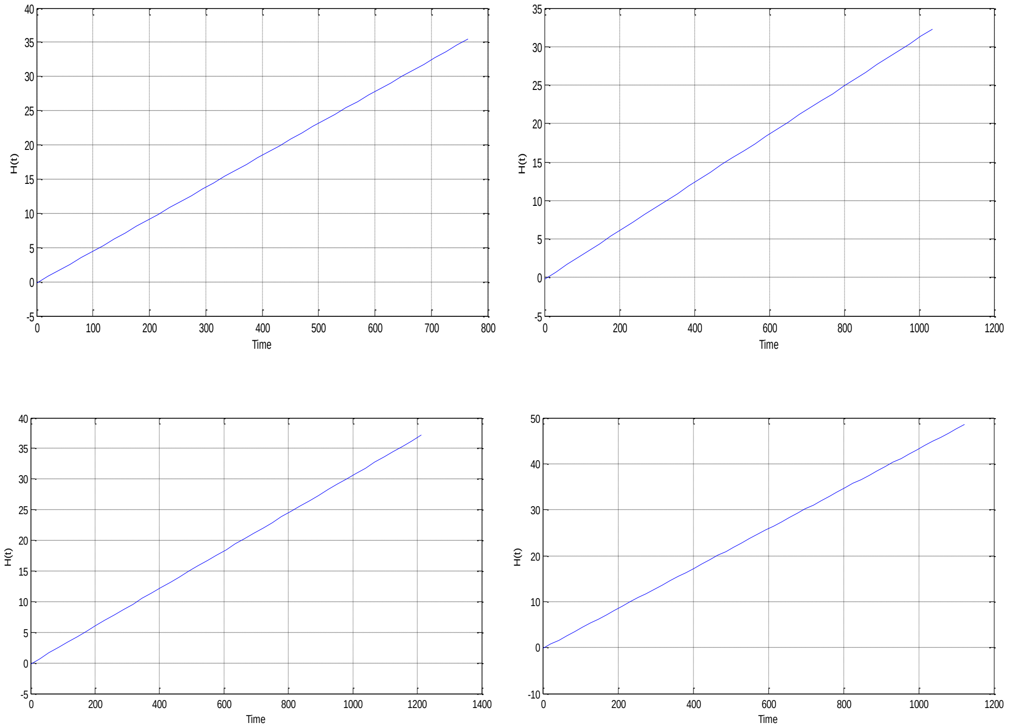

Figure 1: Renewal Function for Rayleigh distribution by using ML method

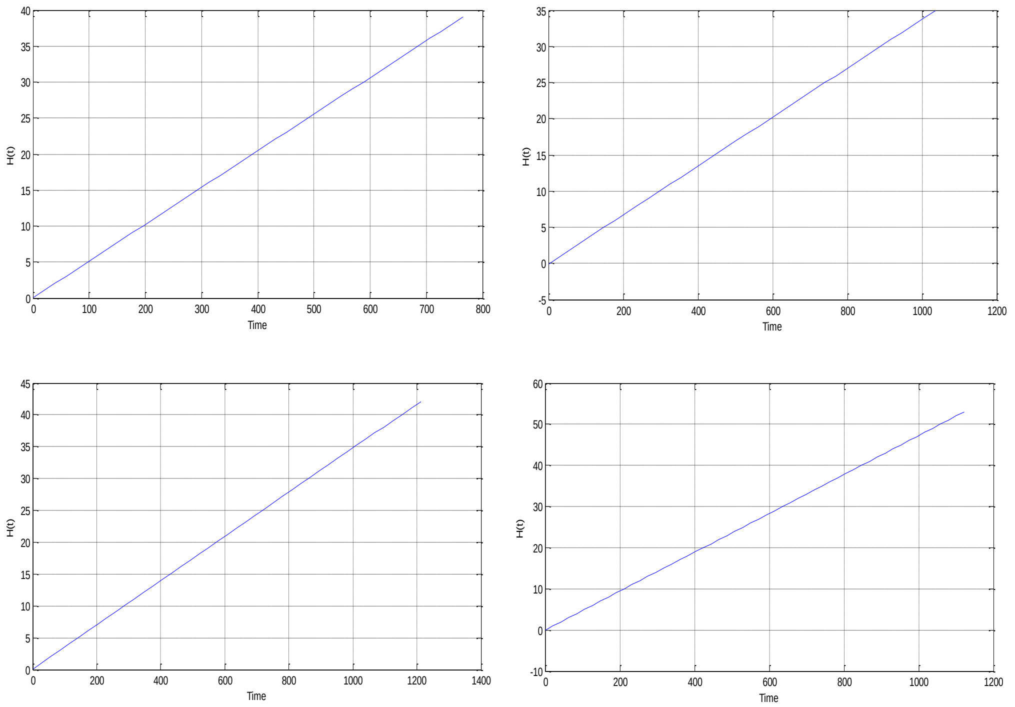

Figure 2: Renewal Function for Rayleigh distribution by using MO method

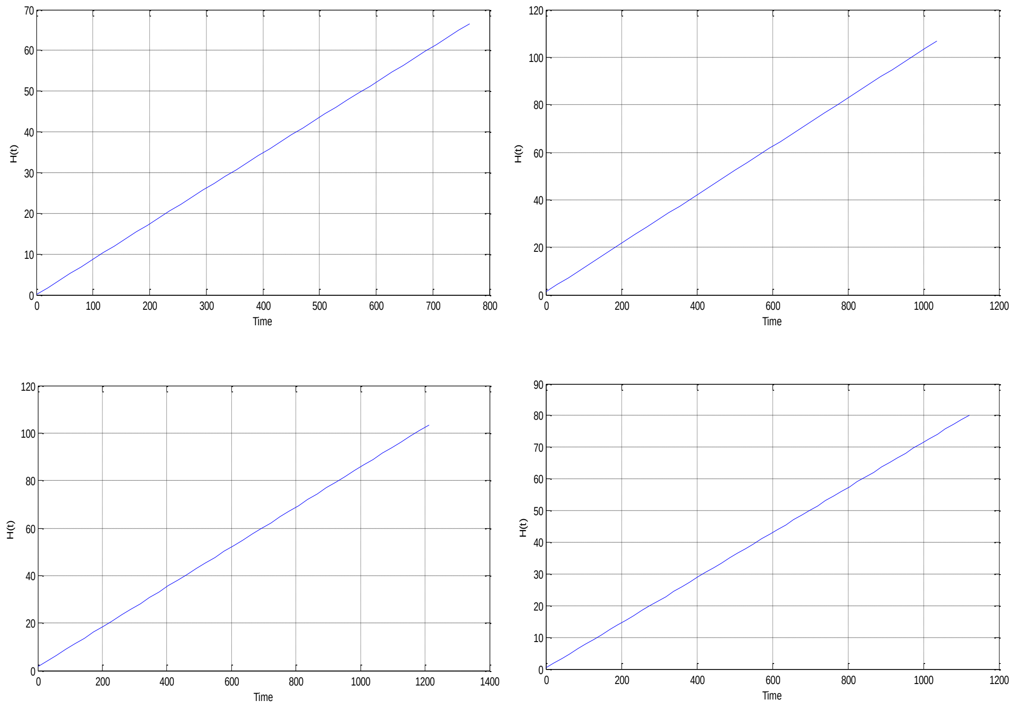

Figure 3: Renewal Function for Rayleigh distribution by using FA algorithm

For the purpose of comparing the three methods used in estimating the renewal function, the MPE criterion was used and the results are reported in Table 4.

Table 4: MPE values of the renewal function

Data	ML	MO	FA
D1	0.8705	0.8478	0.7426
D2	0.9753	0.9693	0.8777
D3	0.9757	0.9614	0.8632
D4	0.9715	0.9638	0.9374

From Table 4, we can be seen that MPE values for the FA is less than the ML and MO for all the used datasets. In data 1, for example, the FA reduced the MPA by 14.68% and by 12.40% using ML and MO methods, respectively.

7. Conclusions

In this present study, a firefly algorithm method is proposed to estimate the rate of the occurrence of events by estimating the renewal function. Depending on the Rayleigh distribution, the FA outperformed the two other estimation methods: MLE and MO in terms of MSE criterion. Moreover, the results obtained, in terms of MAP, demonstrated that the FA can produce a reliable renewal function.

References

[1] A. Jokiel-Rokita and R. Magiera, "Estimation of parameters for trend-renewal processes," Stat. Comput., vol. 22, no. 2, pp. 625–637, 2012.
[2] R. Lefevere, M. Mariani, and L. Zambotti, "Large deviations for renewal processes," Stoch. Process. their Appl., vol. 121, no. 10, pp. 2243–2271, 2011.
[3] M. N. M. van Lieshout, "Likelihood based inference for partially observed renewal processes," Stat. \& Probab. Lett., vol. 118, pp. 190–196, 2016.
[4] Y. Guédon and C. Cocozza-Thivent, "Nonparametric estimation of renewal processes from count data," Can. J. Stat., vol. 31, no. 2, pp. 191–223, 2003.
[5] S. G. From, "Approximating the distribution of a renewal process using generalized Poisson distributions," J. Stat. Comput. Simul., vol. 74, no. 9, pp. 667–681, 2004.
[6] S. Wang and Y.-K. Liu, "Modeling renewal processes in fuzzy decision system," Appl. Math. Model., vol. 39, no. 5–6, pp. 1536–1553, 2015.
[7] I. Hadi and others, "A method for computing the autocovariance of renewal processes," J. Korean Stat. Soc., vol. 47, no. 4, pp. 491–508, 2018.
[8] M. L. Gámiz and B. H. Lindqvist, "Nonparametric estimation in trend-renewal processes," Reliab. Eng. \& Syst. Saf., vol. 145, pp. 38–46, 2016.
[9] M. D. Pandey and J. A. M. Van Der Weide, "Stochastic renewal process models for estimation of damage cost over the life-cycle of a structure," Struct. Saf., vol. 67, pp. 27–38, 2017.
[10] X. S. Yang and N.-I. M. Algorithms, "pp. 83-96." Luniver Press, London, UK, 2008.
[11] I. Fister, I. Fister Jr, X.-S. Yang, and J. Brest, "A comprehensive review of firefly algorithms," Swarm Evol. Comput., vol. 13, pp. 34–46, 2013.
[12] X.-S. Yang, Nature-inspired metaheuristic algorithms. Luniver press, 2010.
[13] J. Kwiecień and B. Filipowicz, "Firefly algorithm in optimization of queueing systems," Bull. Polish Acad. Sci. Tech. Sci., vol. 60, no. 2, pp. 363–368, 2012.
[14] N. Ali, M. A. Othman, M. N. Husain, and M. H. Misran, "A review of firefly algorithm," ARPN J. Eng. Appl. Sci., vol. 9, no. 10, pp. 1732–1736, 2014.
[15] A. Hashmi, N. Goel, S. Goel, and D. Gupta, "Firefly algorithm for unconstrained optimization," IOSR J Comput Eng, vol. 11, no. 1, pp. 75–78, 2013.