Global Journal of Mathematical Analysis, 2 (2) (2014) 58-64
©Science Publishing Corporation
www.sciencepubco.com/index.php/GJMA
doi: 10.14419/gjma.v2i2.2404
Research Paper

Parameter estimation of a tumor growth model using the real-valued genetic algorithm

Tanuja Agrawal [1]*, Vinti Agarwal [2]

[1] Department of Applied Mathematics, Z.H. College of Engineering and Technology, AMU, Aligarh, UP 202 002, India
[2] School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi 110067, India
**Corresponding author E-mail: tanu.18aug@gmail.com*

Abstract

This paper presents the use of real-valued Genetic Algorithm (GA) to evolve set of unknown parameters and initial conditions for tumor growth model using data extracted from El-Gohary [1]. The main focus of this work is to reach beyond the possibilities of traditional optimization methods in obtaining the far situated global optimum solutions with the help of arithmetic crossover and uniform mutation operators. Experimental results show the effectiveness of our approach by comparing the results obtained against the one mentioned in El-Gohary [1].

Keywords: *Real-Valued Genetic Algorithms; Tumor Growth Model; Parameter Estimation*.

1. Introduction

Mathematical models based on a system of ordinary differential equations are widely used in many applications. The solutions of such models cannot be expressed by elementary functions, in most cases where system contains unknown parameters which are usually estimated by experimental data obtained from well-defined standard conditions. This type of problem is usually called parameter estimation in the literature and is often solved by deterministic optimization methods [2], [3], [4]. Unfortunately, the solution to the problem using these methods is usually around the local minima if there are more than one minimum available.

However, better solutions may be found by one of stochastic optimization methods to obtain global minima if they exist in the given search space. The genetic algorithms (GA) method, the best known stochastic optimization technique, imitates the natural evolution process [5], [6], [7]. One of the main advantages of this method is that it doesn’t require any additional information of the problem.

Recently there has been an increasing interest towards real-valued GA method for parameter estimation in various applications to overcome the drawbacks of the earlier developed binary coded GA (BCGA) such as high precision and faster convergence rate. In his study, Tutkun [8] .has employed Real-valued GA approach for estimating parameters of enzyme diffusion model. The results shows considerably improved with respect to previous findings in many aspects such as accuracy, computation time etc. Parameter estimation is a useful method to obtain the parameters of a mathematical model through knowledge of data on a given system [9]. It focuses on adjusting the model parameters to get a final model that delivers outputs similar to the observed data; however, in many cases the process can gain complexity in the presence of disturbances or badly taken data. The application to a nonlinear system, as will be done in this paper, is a bit more complicated than doing so with linear systems, hence heuristic methods are useful in these situations [10]. Mathematical modeling tools can be applied to problems of any topic; for example, sciences like biology are full with hundreds of interesting problems to be modeled [11], like the predator-prey model, which has been a topic of much research because of its characteristics and flexibility [12]. Even though, there has been several attempts towards using evolutionary approaches for parameter estimation, but no one has considered the same idea to generate set of different parameters and initial conditions that satisfy the same feasibility criteria as the already used set does, which is the focus of our work. Here, our idea is to use real-valued GA for parameter estimation in predator-prey type tumor growth model by obtaining outputs similar to data obtained by El-Gohary [1].

The paper is organized as follows. Section II represents predator-prey type tumor growth model from a mathematical point of view; section III describes the genetic algorithm and its implementation; section IV shows the experimental results; and section V presents a summary with some conclusions

2. Predator-prey type tumor growth model

El-Gohary [1] considered the following model for cancer self-remission and tumor growth

$$\frac{dM}{dt} = q + rM\left(1 - \frac{M}{k_1}\right) - \alpha MN$$

$$\frac{dN}{dt} = \beta NZ - d_1 N$$

$$\frac{dZ}{dt} = sZ\left(1 - \frac{Z}{k_2}\right) - \beta NZ - d_2 Z \qquad (2.1)$$

In (2.1), different variables and parameters have the following interpretations:

$M(t), N(t), Z(t)$: Densities of tumor cells, hunting predator cells and resting cells at time t

q: Conversion of normal cells to malignant cells

r: Growth rate of tumor cells

k_1: Maximum carrying capacity of tumor cells

α: Rate of killing of tumor cells by hunting cells

β: Conversion rate of the resting cells to tumor cells

d_1: Natural death rate of hunting cells

s: Growth rate of resting cells

d_2: Natural death rate of resting cells

Using non-dimensional variables and parameters as

$\tau = qk_1^{-1}t,\ x_1 = k_1^{-1}M,\ x_2 = \alpha k_1 q^{-1}N,\ x_3 = k_2^{-1}Z$,

$a_1 = rk_1q^{-1},\ a_2 = \beta k_1 k_2 q^{-1},\ a_3 = d_1 k_1 q^{-1},\ a_4 = sk_1q^{-1},\ a_5 = \beta\alpha^{-1},\ a_6 = d_2 k_1 q^{-1}$,

El-Gohary [1] obtained the following non-dimensional form of model (2.1)

$$\frac{dx_1}{d\tau} = 1 + a_1x_1(1 - x_1) - x_1x_2 \qquad (2.2a)$$

$$\frac{dx_2}{d\tau} = a_2x_2x_3 - a_3x_2 \qquad (2.2b)$$

$$\frac{dx_3}{d\tau} = a_4x_3(1 - x_3) - a_5x_2x_3 - a_6x_3, \qquad (2.2c)$$

Equilibriums of the model (2.2) have been reported as

$$E_1 = \left\{\frac{1}{2}\left(1 + \sqrt{1 + \frac{4}{a_1}}\right), 0, 0\right\},$$

$$E_2 = \left\{\frac{1}{2}\left(1 + \sqrt{1 + \frac{4}{a_1}}\right), 0, \left(1 - \frac{a_6}{a_4}\right)\right\} \text{ And}$$

$$E_3 = (\overline{x_1}, \overline{x_2}, \overline{x_3}) = \left[\frac{1}{2a_1}\left\{(a_1 - \overline{x_2}) + \sqrt{(a_1 - \overline{x_2})^2 + 4a_1}\right\}, \left\{\frac{a_4}{a_5}\left(1 - \frac{a_3}{a_2}\right) - \frac{a_6}{a_5}\right\}, \frac{a_3}{a_2}\right]$$

In El-Gohary [1], under the biologically feasible conditions as

$$a_4 > a_6 \text{ And } \frac{a_3}{a_2} + \frac{a_6}{a_4} < 1 \qquad (2.3)$$

The first condition states that the ratio of the natural death rate of resting cell to its growth rate is less than one. The second condition implies that the ratio of the natural death rate of the hunting cell to its growth rate is also less than one.

3. Learning parameters and initial conditions using real-valued genetic algorithms

In order to explore the problem, we have considered 3-dimensional tumor growth model with the following set of parameters and initial conditions: a_1, a_2, a_3, a_4, a_5, a_6 are unknown parameters and x_1, x_2, x_3 are initial conditions at t =0, used in model defined in eq. (2.2a-2.2c). Before proceedings with the optimization operations, a performance criteria or an objective function should be first defined. In this study, the sum of squared errors between real and estimated responses for a given number of samples is considered as fitness of estimated model parameters. So the fitness function is defined as follows:

$$\text{Fitness} = \frac{1}{M}\left\{\sum_{k=1}^{M}\|x_1(k) - \hat{x}_1(k)\| + \sum_{k=1}^{M}\|x_2(k) - \hat{x}_2(k)\| + \sum_{k=1}^{M}\|x_3(k) - \hat{x}_3(k)\|\right\}$$

Where M is the length of data used for parameter identifications, $x_1(k)$, $x_2(k)$, $x_3(k)$and $\hat{x}_1(k)$ $\hat{x}_2(k)$, $\hat{x}_3(k)$ are real and estimated values of tumor, hunting and resting cells at time t, respectively. Fitness function validates the diversion of process towards its optimization goal by allowing best individuals to breed that lead to good recommendation [13]. Recursively the individuals that achieve a higher fitness are more likely to be selected as parents for next generation. For this, we rerun our experiment on entire set including parameters generated after crossover and mutation operators until it reaches to the minimum values of fitness or iterations are exhausted.

3.1. Genetic algorithms

The idea behind the origin of Genetic Algorithms is its robustness that creates balance between efficiency and efficacy, which is necessary for survival in competitive environment-"Darwin principle of evolution". The widely adoption of GA, in various applications across a large and growing number of disciplines indicates its simplicity and power of search in complex spaces. Unlike other search techniques, the scope of GA is not limited by some constraints over the search space. In nature, competition among individuals for scanty resources results in the fittest individuals dominating over the weaker ones. The numbers of phases involved in a simple GA are as

i. Chromosome Representation The choice of a way to represent chromosomes is one of the most important challenges in GA. Basically, this step involves the representation of population according to the given problem.

ii. Targeted Objective Function Each chromosome, mapped into a potential solution to an optimization problem, and is evaluated by using an objective function that decides virtue of the solution.

iii. Reproduction is a process which begins with a series of cycle of replacing its current population of chromosomes by new population. Each cycle includes three vital components:

Parent selection, Mutation, and Crossover [14]. The goal of parent selection method is to give more opportunity to those population candidates which are best among others. One the most commonly used technique for this purpose is "Roulette wheel Selection" also known as "fitness proportionate selection", where chance of an individual being selected is proportional to its fitness value. Crossover and mutation are the two most common genetic operators used to generate new individuals from the ones in the current generation. Crossover works by letting a pair of chromosomes exchange meaningful information to create two offspring, while mutation involves a random manipulation of a single chromosome to create a new individual.

3.2. Chromosome representation in real-valued GA

Traditional GA where chromosome representation technique is bit strings suffers from slow convergence due to the large chromosome structure, especially in optimization problems involving several parameters. We have used real-valued GA, to learn the parameters as real numbers and therefore a floating point representation of chromosomes is most suited for our purpose. In our GA each chromosome is represented as a vector [θ_1, θ_2 …θ_9], where each θ has two real variables which indicates maximum and minimum values in the valid range [15]. The chromosome search space for real-valued GA is defined as

$$\Omega = \{\theta \in R^m \mid \theta_1^{\min} \leq \theta_1 \leq \theta_1^{\max}, \ldots\ldots\ldots \ldots\ldots\ldots\ \theta_9^{\min} \leq \theta_9 \leq \theta_9^{\max}\} \tag{3.1}$$

Figure 1 shows an example of two parent chromosomes of a particular user. Initially, we assume that each weight lies between 0 and 30.

3.3. Selection

The selection process we used in our work is based on spinning the roulette wheel pop- size times, and each time the probability $p_{i\ \text{that}}$ a chromosome is selected as a new offspring is given by:

$$\text{Pi} = \frac{\mathbf{F(Xi)}}{\sum_{\mathbf{j=1}}^{\mathbf{pop-size}} \mathbf{F(Xj)}} \tag{3.2}$$

Where F (X_i) is the fitness of X_i individual.

3.4. Crossover and mutation operator

Genetic operators are being used to drive search process in GA but may differ from traditional operators described in [16]. Crossover and mutation are the two basic genetic operators. While crossover allows creation of two new individuals by allowing two parent chromosomes to exchange meaningful information, mutation is used to maintain the genetic diversity of the population by introducing a completely new member into the population. Crossover and mutation operators for real-valued GA were introduced by [16]. From generation to generation, the good newly generated individuals (if exist) replace the current bad individuals to form the new population for the next generation. Given real-valued parents,

$$X_1 = \{x_{11}, x_{12} \ldots x_{1m}\} \tag{3.3}$$

$$X_2 = \{x_{21}, x_{22} \ldots x_{2m}\} \tag{3.4}$$

Some of the different types of crossover and mutation operators are discussed below:

Uniform Mutation: Uniform mutation randomly selects one gene x_i and sets it equal to a uniform random number U(A_i , B_i) such that:

$$x_j' = \begin{cases} U(a_i, b_i) & \text{if } i = j \\ x_j, & \text{otherwise} \end{cases} \tag{3.5}$$

Gaussian Mutation: Gaussian mutation is commonly preferred in real-valued GAs. For example, when a gene is mutated, a random value from a Gaussian distribution is added to the gene value.

$x_i' = x_i + G(x)$ (3.6)

Where G(x) is the Gaussian function.

And another one is Gaussian with momentum, improves the convergence speed of GA on the problems involving real value alleles.

$x_i' = x_i + \acute{\alpha} M_i + G(x)$; $0 \leq \acute{\alpha} \leq 1$ (3.7)

Where M_i is the momentum and ά controls its amount

Arithmetic Crossover Arithmetic crossover produces two complimentary linear combinations of the parents based on the value of $\tau = U(0,1)$.This operator attempts to average or blend components across multiple parents to create an offspring as follow:

$X_1' = \tau X_1 + (1-\tau) X_2$ (3.8)

$X_2' = \tau X_2 + (1-\tau) X_1$ (3.9)

For example, if we assume τ = 0.2, then the newly generated offspring, obtained after applying arithmetic crossover operation on parents in Figure 1, are as follows:

GA terminates either a maximum number of generations elapsed or a desired level of fitness is reached.

Parent 1(X_1):

0	22	12	21	29	6	9	12	24	11	17	2	16	5	3	1

Parent 2(X_2):

10	2	15	2	9	13	7	1	9	17	2	4	27	4	30	7

(a) Before crossover

Offspring 1($X_1' = 0.2* X_1 + 0.8* X_2$):

8	6	14.4	5.8	13	11.6	7.4	3.2	12	15.8	5	3.6	24.8	4.2	24.6	5.8

Offspring 2($X_2' = 0.2* X_2 + 0.8* X_1$):

2	18	12.6	17.2	25	7.4	8.6	9.8	21	12.2	14	2.4	18.2	4.8	8.4	2.2

(b) After crossover

Fig. 1: Representation of Real-Valued Chromosomes A) before Arithmetic Crossover Operation and B) after Crossover Operation.

4. Experimental results

Due to the limited access of information about growth of cells, we have conducted our experiments on the dataset obtained from El-Gohary [1], illustrated in Table 1, under the set of parameters $a_1 = 2.134$, $a_2 = 15.319$, $a_3 = 0.425$, $a_4 = 24.124$, $a_5 = 4.472$, $a_6=17.005$. Moreover, the initial conditions in the present work are chosen as $[x_1(0) \quad x_2(0) \quad x_3(0)] = [0.528 \ 0.811 \ 1.741]$. To perform our experimental analysis we used available optimization toolbox in MATLAB with the following set of parameters mentioned in Table 2. The simulation results obtained through first experiment, presented in Figure 2, clearly shows the fitness value, best set of individuals, average distances between generations and the number of individuals being produced. GA process begins with roulette wheel selection for the next generation, followed by the real value assignment to each of the 9 parameters in the range [0...30]. The initial population size is 100, with 100 new individuals being produced from generation to generation. In order to produce 80 new individuals, arithmetic crossover operation is used while remaining 20 are being created by uniform mutation. An elitist approach is used, wherein the best 10 individuals from the previous population are used in next generation.

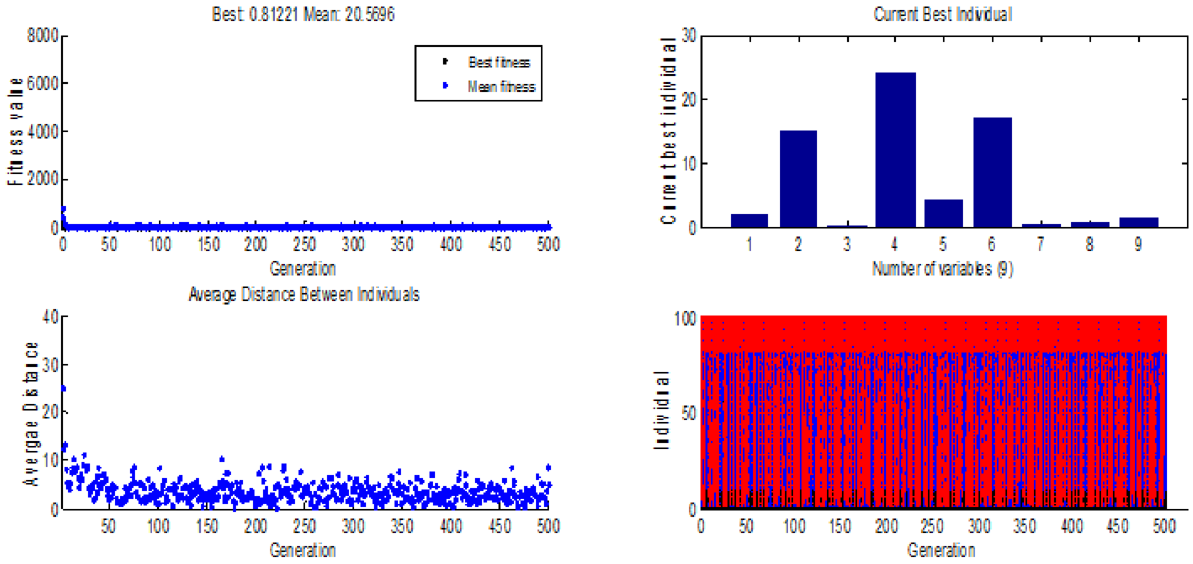

Fig. 2: The Representation of Different Stages of Each Generation during Evolution Process

Following is the set of parameters and initial conditions generated via real-valued GA using arithmetic crossover and uniform mutation operators, where we have performed 10 iterations for the experiment to verify their sustainability.
Set 1: a_1 = 2.134, a_2 = 15.319, a_3 = 0.425, a_4 = 24.124, a_5 = 4.472, a_6=17.005 and $x_1(0)$ =0.528, $x_2(0)$=0.811, $x_3(0)$=1.741.

Table 1: Experimental data for El-Gohary [1]

t	x_1	x_2	x_3		T	x_1	x_2	x_3
0		0.5000	1.0000	1.5000	2.4500	0.8096	1.4156	0.0079
0.0082		0.5090	1.0506	1.4539	3.0000	0.9296	1.0432	0.0132
0.0163		0.5178	1.1020	1.4099	3.4945	1.0135	0.8092	0.0276
0.1327		0.6065	1.8771	0.9277	4.0055	1.0744	0.6604	0.0689
0.3327		0.6147	2.9953	0.4295	4.5305	1.1039	0.6308	0.1718
0.4077		0.5928	3.2437	0.3147	5.5084	0.9999	1.2467	0.3405
0.9077		0.4998	3.2486	0.0453	6.2030	0.8181	1.9334	0.1931
1.0000		0.5025	3.1242	0.0340	7.1249	0.7590	1.8490	0.0797
1.8427		0.6579	1.9967	0.0079	8.0519	0.8492	1.4241	0.0719
1.9919		0.6954	1.8351	0.0074	8.5769	0.9023	1.2561	0.0912
2.0000		0.6975	1.8267	0.0074	9.4519	0.9450	1.1908	0.1519
2.2500		0.7606	1.5852	0.0073	10.0000	0.9314	1.3010	0.1824

Table 2: GA Parameter Values Used in Experiments

Parameter name	Parameter Value	Description
Population Size	100	The size of individuals in the population at each generation
Termination threshold	0.005	When the fitness value of the best individual is either less than or equal to threshold a good solution is found
Maximum number of Generation	500	If the number of generation reaches to this value and solution has not been found, then the best individual of this generation is used as a final set of parameters
Crossover fraction	0.8	It specifies the fraction of the next generation that crossover produces
Mutation rate	0.05	It defines the probability of an individual in population is being mutated

Second, to observe the change in growth analysis of x_1, x_2 and x_3 using the one set of parameters we generated through real-valued GA with the set used in Gohary [1].

Table 3: Comparison of Estimated Parameters with El-Gohary [1] Model

Parameters	Parameters with Real-valued GA	Parameters of El-Gohary [1] model
a_1	2.134	2.5
a_2	15.319	4.5
a_3	0.425	0.6
a_4	24.124	3.5
a_5	4.472	2
a_6	17.005	0.1
x_1	0.528	0.5
x_2	0.811	1
x_3	1.741	1.5

5. Conclusions

The parameter estimation using genetic algorithms with a predator-prey type tumor model has a favorable fit but it requires a large number of iterations (generations) to find the best solution. The estimation can be improved by allowing the algorithm to run for longer. This is not a surprise because of the number of parameters and the difficult of measuring them.

Although the parameter estimation with a model that includes inputs was made in a simulation context, it is an interesting possibility for future researches, where the biological system can be manipulated with persistently excited inputs or controlled to keep a stable system. An interesting application of estimation with a model with inputs is the estimation of pulse parameters (amplitude and width), which can help to identify the time and magnitude of a change in population due to an external event; that is, identify input shape.

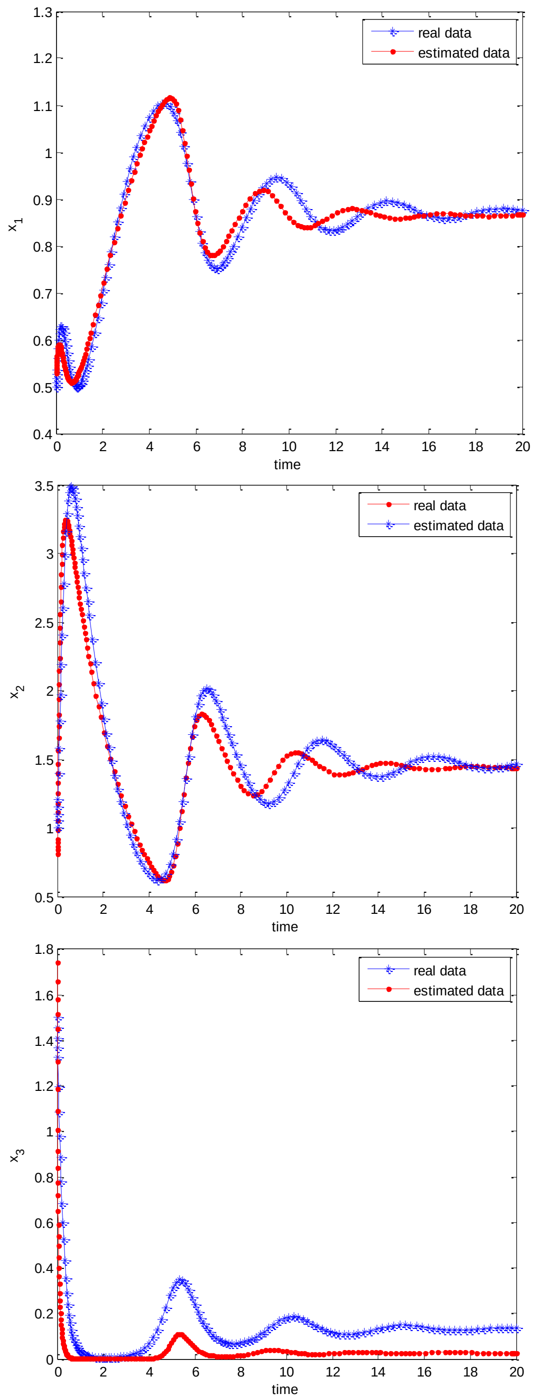

Fig. 3: The Variation of Densities of Tumor Cells, Hunting Cells and Resting Predator Cells with Parameters and Initial Conditions Mentioned in Table 3.

References

[1] El-Gohary A, Chaos and optimal control of cancer self-remission and tumor system steady states. Chaos, Solitons & Fractals, 37(5); 1305-16, (2008).

[2] Abbasi B, Jahromi AHE, Arkat J, Hosseinkouchack M, Estimating the parameters of Weibull distribution using simulated annealing algorithm. Applied Mathematics and Computation, 183(1): 85–93, (2006).

[3] Bayram M., Yildiz B, Parameter estimation of an enzyme kinetic system using computer algebra techniques. Applied Mathematics and Computation, 183; 93–98, (1999).

[4] Yildirim N, Akcay F, Okur H, Yildirim D, Parameter estimation of nonlinear models in biochemistry: A comparative study on optimization methods. Applied Mathematics and Computation, 140(1): 29–36, (2003).

[5] Coley DA, Introduction to genetic algorithms for scientists and engineers. London: World Scientific Publishing, (2003).

[6] Gen M, Cheng R, Genetic algorithms and engineering optimization. New York: John Wiley & Sons, (2000).

[7] Goldberg DE, Genetic algorithms in search, optimization and machine learning. Reading: Addison-Wesley, (2003).

[8] Tutkun N, Parameter estimation in mathematical models using the real coded genetic algorithms. Expert Systems with Applications 36: 3342–3345, (2009).

[9] Jost C, Arditi R, from pattern to process: Identifying predatorprey models from time-series data, Population Ecology, Vol. 43: 229-243, (2001).

[10] Mitchell M, Genetic Algorithms: An Overview, Complexity, 1(1): 31-39, (1995).

[11] Jeffries C, Mathematical modeling in Ecology. Boston: Birkhäuser, (1989).

[12] Katz C, A Nonequilibrium Marine Predator-Prey Interaction, Ecology, 66(5): 1426-1438, (1985).

[13] Bharadwaj KK and Al-Shamri MYH, Fuzzy-Genetic Approach to Recommender Systems Based on a Novel Hybrid User Model, Expert Systems with Applications, Elsevier, 35: 1386-1399, (2007).

[14] Mitchell M, An introduction to genetic algorithms, Cambridge, MA: MIT Press, (1998).

[15] Corcoran AL and Sen S, Using Real-valued Genetic Algorithm to Evolve Rule Sets for Classification, In: IEEE-CEC for Complex Social Networks, (1994).

[16] Michalewicz Z, Genetic Algorithms + Data Structures = Evolution Programs, AI Series, New York, Springer-Verlag, (1992).