EXPONENTIAL SCALING OF INDIVIDUAL SUITABILITY IN EVOLUTION ALGORITHMS

Anton Moiseenko

Saint-Petersburg State University of Aerospace Instrumentation,
Saint-Petersburg, Russia
tel: +7-911-292-3902
canggu@mail.ru

Abstract

There is a rife problem of premature convergence to local optimum in genetic algorithms. One of feasible solutions lies in using scaling methods. The developed scaling method based on exponential function allows to transform individual weights relatively to the generation average fitness-function value. The suitability of weakly-fitness individuals decreases divisible by the average value before scaling. For individuals with average-fitness degree the suitability is hardly changed after scaling. The problem of evolution algorithm convergence to local extremum is solved by changing weights of the most suitable individuals in comparison with new weights of weakly-fitness ones.

I. INTRODUCTION

There are two reasons of scaling fitness-function for evolution algorithms as given in [1–3]:

At-first, to prevent premature convergence of genetic algorithms to local optimum. It appears when good(not best) chromosomes begin to dominate. For example, roulette selection [4] could result after some generations to new one including only copies of the best individs from initial population. The cause is in selection probability proportionate to fitness-function value. As initial generation is usually chosen accidentally from whole retrieval area so it will hardly ever include the most viable individums.

Secondly, when generation has sizeable heterogeneity but average value of fitness differs from maximum for a little [1]. As result average and the best individums form almost the same offspring amount in following generations.

The scaling is an appropriate transformation of fitness-function. Accordingly to [1–3] *the main idea of the approach is to decrease suitability of weakly fitness individums multiply to the average. The suitability of individums with average fitness degree should not change after scaling. The congestion of well-fitness individs also has to decrease due to new smaller fitness degree.* There are three main scaling methods in sources [1–3]: linear, sigma-trancation and power law scaling. Lets investigate it.

II. LINEAR SCALING

As shown in [1, 2] new value of fitness-function is defined by classical formula.

$$g(f) = k \times f + b\,, \qquad (1)$$

where k and b are constants, that should be defined empirically. As is well known repetition factor k sets the slope ratio of straight line to absciss and the scaling factor accordingly.

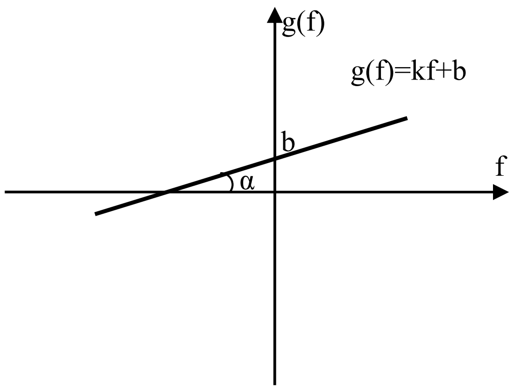

Fig.1-a Linear function

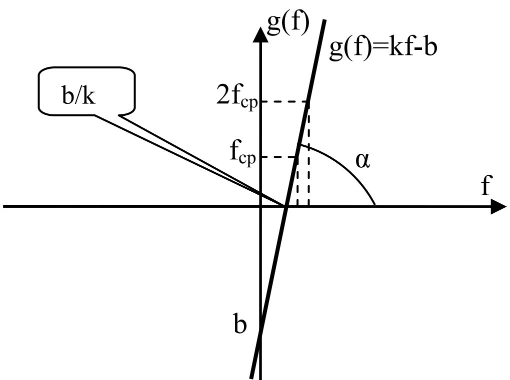

Fig.1-b Linear scaling
(the straight line crosses the abscissa in b/k point)

The advantage of this approach is in possibility of cutting off idividuals exceeding generation average fitness degree

It is necessary to keep up with function value to not take on a negative value. For example, after scaling new fitness degree less then b/k could be changed to zero.

The disadvantage is that such individuals will have the same weight and therefore the same selection probability.

The rate of change for function g(f) is equal in any point f as the derivative represents the following value:

$$\frac{dg}{df} = k \qquad (2)$$

III. SIGMA-TRUNCATION

This scaling method looks like the *linear*. It is based on transformation of fitness-function f to g(f) by the following expression [1, 2]:

$$g(f) = f + (f_{avg} - c \times \sigma), \qquad (3)$$

where f_{avg} – is the average fitness-function value for generation; c – a little natural number (usually from 1 to 5), seeked in concordance with investigated problem; σ – standard deviation for generation that could be defined by the following expression:

$$\sigma = \sqrt{\sum_{i=1}^{N} \frac{(f_i - f_{avg})^2}{N}}, \qquad (4)$$

N – generation volume.

If design data F' is negative then it puts to zero like under linear scaling. The diagram of derivative for this function is also parallel to abscissa straight line.

$$\frac{dg}{df} = 1 \qquad (5)$$

Hence in this case the rate of change is like in the previous method.

Lets try a factor c and investigate the changes of the fitness degree after scaling on a specific example. By way of example look at the problem of audience time-table optimization. Let there is an initial generation consists of 7 casual individuals, which are full-fledged schedule variants for a half-year in an Institute of Higher Education [5]. There is a penalty system to evaluate fitness value for any schedule: the penalty for the "gap" in a group or teacher schedule, for moving between buildings of educational institution during a day, for one lesson is out of timetable's grid, for same lessons during a day in a group schedule and so on. Fitness-function (or total schedule weight) is a sum of penalties for all groups and teachers during both workweeks (even/odd). The suitability of any individum (schedule variant) is inversely to fitness-function value. So the heavier total weight, the smaller suitability. New generation will be formed by means of individum selection from previous generation with successive execution crossingover and mutation genetic operators [6].

Try to rate the suitability for 6 schedule generations being part of the era [7] (a set of generations), formed using roulette selection [4] (Fig. 2)

When c=1 the individum total after-scaling weight becames heavier then the average before scaling. But it is unacceptablly.

When c=2 the average of fitness-function increase on 25-30%. So the suitability decrease on the same value. But it is not suit us.

When c=4 the average of fitness-function increase on 60-100%. But it is not also right for the definition.

When c=5 total weight of many individum is put to zero (the highest suitability) because of the negative fitness-function value. So the selection probability of such individums becames equal, but for others it becames infinitesimal.

In such a way it is better to choose c equal to 3 or 4. Although in all variants individual density stays the same (diagram shapes are equal). Hence the selection probability is not changed.

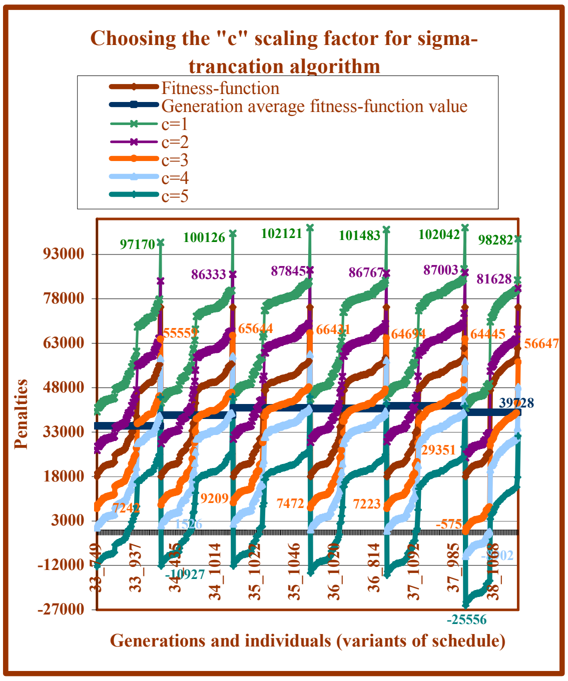

Fig 2. Sutability scaling for 6 generations of schedule variants (individums)

IV. POWER LAW SCALING

It is the method transforms fitness-function by the following expression:

$$g(f) = f^k \qquad (6)$$

where k – is a number near to 1, as mentioned in source [1,2]. Usually it is fit empirically depending on specific of solving problem. For example, k=1,005 could be used.

The derivative for this function is:

$$\frac{dg}{df} = k \times f^{k-1} \qquad (7)$$

When k is close to one the function increases extremely slow (look at Fig.1). So after scaling her shape is changed very weakly (Fig. 4). Hereinafter we will use k=1,005 for our task.

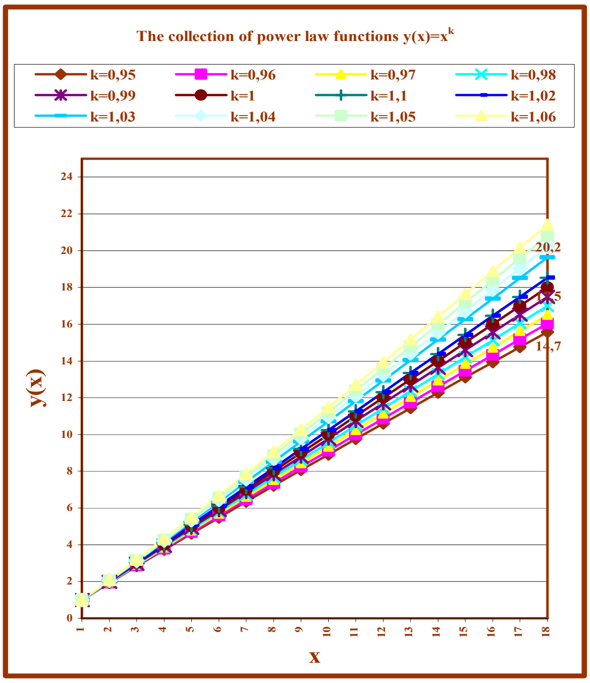

Fig.3 Collection of functions y=x^k for $k \ni (0.95;1.06)$

V. EXPONENTIAL SCALING

The problem of search optimum schedule is lead to search for fitness-function minimum as individum suitability is inversely to its total weight. Since scaling function should be monotonic and increasing for any initial fitness-function value so try to use exponent besides upper mentioned functions. It also complies with definition.

Lets try to fit scaling factors:

$$g(f) = e^{f+\beta}; \qquad \beta = \ln A - A; \qquad (8)$$

where A – is an average of fitness-function value for generation (always positive).

After differentiation we have:

$$\frac{dg}{df} = e^{f+\ln A - A} \qquad (9)$$

i. e. the rate of scaling-function change increases exponentially.

Now it is necessary to verify how it meets the definition of scaling-function from [1–3]:

1. *To prevent evolution algorithm convergence exponentially increase suitability of best individums decreasing their "density" in that way.*

If we use formaula (8) this condition will be met:

$$\lim_{A->0} g(A) = \lim_{A->0} e^{\ln A - A} \approx 0 \qquad (10)$$

2. *As on definition individual suitability equal to the average of generation fitness-function value should not be change after scaling.* This condition is also met.

$$g(A) = e^{A+\ln A - A} \Rightarrow g(A) = A \qquad (11)$$

3. Also by definition after-scaling suitability of individums with lowest suitability should became divisible by before-scaling average value. I.e., for example, the following inequality should be true:

$$g(2A)>2A, \qquad g(2A) = e^{A+\ln A} \qquad (12)$$

So it is necessary to prove, that $e^{A+\ln A} > 2A$:

$$e > 1 \Rightarrow e^{A+\ln A} > e^{\ln 2A}, \quad A + \ln A > \ln 2A, \quad A > \ln 2A - \ln A, \quad A > \ln 2 \qquad (13)$$

In such a way the function will work for all generations with fitness-function average value A higher than $\ln 2 \approx 0.7$. If 0<A<ln2 then all more suitable individums should be put to zero after scaling.

Now compare the analyzed methods evaluating individual suitability by sigma-trancation (kind of linear scaling per se), power low scaling and exponential scaling.

After sigma-trancation and power low scaling the shape of fitness function is not changed (Fig. 4). So the selection probability is not changed as well.

After exponential scaling the shape of curve firstly looks like others, but fitness-function value is much lower for well-suitable individums (75% on average). However when the suitability is closer to the average the curve rises strongly. For individums with average suitability the fitness-function value is hardly changed. After that the curve rises extremely abruptly. In such a way the probabilty of weakly-suitable individums selection becomes extremely low and conversely for well-suitable ones.

Now create some new generations using developed exponential scaling method to evaluate individums. By way of initial set the 38-th population has been used (the final one from the era on Fig. 4). The selection of well-suitable individuals (with light-weight) rises after scaling, starting with 39-th generation. From 33-th to 38-th generation the count of "light-weight" and "heavy-weight" individuals is almost the same, but starting from 39-th the count of heavy-weight individums decrease sharply. Of course, part of heavy-weight individums migrate to new population as reproduction strategy is used as well. On the whole after scaling the generation suitability has increased on 26%.

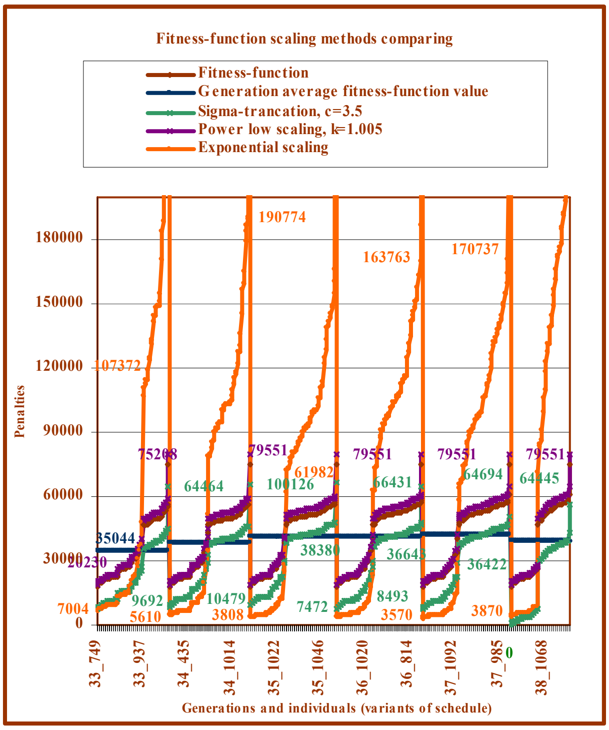

Fig 4. Scaling methods comparing
for 6 population of schedule variants

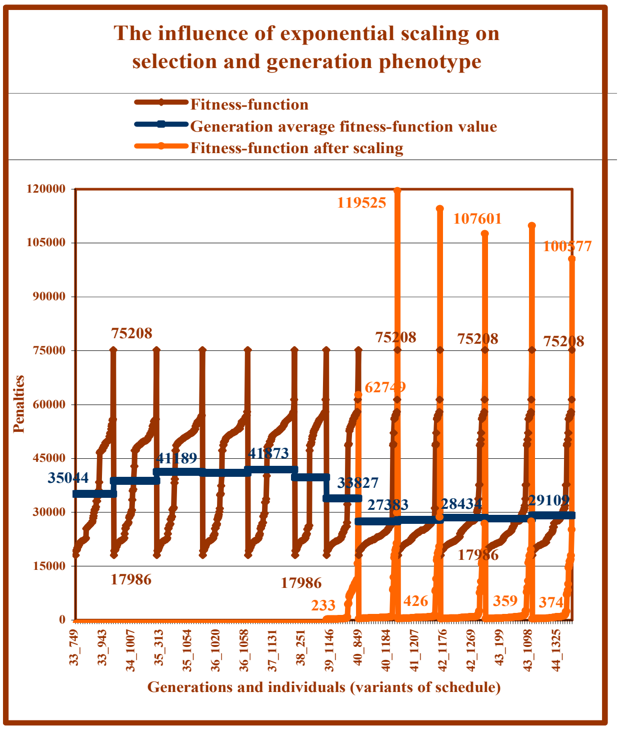

Fig 5. New era, created on basis of roulette method and exponential scaling (initial population - 38)

There is a class implementing exponential scaling on basis of derived formula (see fig.6)

Class TFitnessExpScale has only 4 private properties:

– generationAvgWeight – the generation average fitness-function value

– generationVolume – the volume of generation

– oldWeights – the massive of initial weights (fitness-function values for every individual)

– newWeights – the massive of weights after scaling

TFitnessExpScale
- int generationAvgWeight - int generationVolume - int[] oldWeights - int[] newWeights
+ TFitnessExpScale () + TFitnessExpScale (int[] weights, int volume) + ~TFitnessExpScale () + void setGenerationVolume(int volume) + void setOldWeights(int[] weights) + int getNewWeight(int chromosome) + int* getNewWeights(void)

Fig. 6 Structure of class TFitnessExpScale

Also class TFitnessExpScale includes 6 public methods:

– TFitnessExpScale () – constructor by default;

– ~TFitnessExpScale () – destructor by default;

– void setGenerationVolume(int volume) – to set generation volume;

– void setOldWeights(int[] weights) – to set pointer for oldWeights;

– int getNewWeight(int individ) – return after-scaling weight for individual with chromosome number in oldWeights massive;

– int[] getNewWeights(void) – return pointer for newWeights massive.

VI. CONCLUSION

The research results allow to draw the following conclusions:

– the linear scaling and sigma-truncation methods proposed in source [1–3] don't affect the shape of generation individums fitness-function in contrast to power scaling method. Hence the individums selection probability is not changed.

– the rate of change of exponential scaling function is more higher than the rate of power scaling function taken from source[1–3]. The selection probability increases exponentially with the growth of individum suitability after exponential scaling.

The exponential scaling method allows:

– to prevent genetic algorithm convergence to local optimum due to increase of distance between individums congested near the most suitable one

– to increase the selection probability of well-suitable individums

– to decrease the selection probability of weakly-suitable individums

REFERENCES

[1] D.Rutkovskaya, M.Pilinsky, L.Rutkovsky “Neyronnie seti, geneticheskie algoritmi & netchetkie sistemi” –M.: Goryatchaya liniya – Telecom, 2006

[2] GoldBerg D.E., Algorithmy genetyczne I ich zastosowania, WNT, Warszawa, 1995

[3] Michaleviwicz Z., Genetic Algorithms + Data Structures = Evolution Programs, Springer – Verlag, 1992

[4] Brindle M. "Genetic Algorithms for Function Optimization", Ph.D.dissertation, University of Alberta

[5] Moiseenko A.S. Applying the genetic algorithm for the problem of the studies schedule optimization. // Challenges of the information age in the field of education. St-Petersburg, 2006 – p.53 -60

[6] Moiseenko A.S., Matyash V.A. «Realizatsiya otdelnikh metodov selectsii & reproductsii v genetitcheskih algoritmakh primenitelno k zadatche optimizatcii raspisaniy» - «Mekhatronika, avtomatizatciya, upravlenie», SPb, 2008

[7] Moiseenko A.S. «Epokhi v genetitcheskih algoritmakh. Skreshivanie epokh.» / Transactions of the 9-th International Odessa Scientific Industrial Conference SIET-2008 – Odessa, may 2008