UDC 519.3

Shtovba S.D., Kozachko O.M.
Vinnitsa National Technical University, Vinnitsa
Dounias G.D.
University of the Aegean, Chios

A FAST GENETIC ALGORITHM FOR OPTIMIZING THE CHECKING – RETROFIT PROCEDURES IN MULTIDIMENSIONAL TECHNOLOGICAL PROCESSES

ABSTRACT

We propose an improvement on a genetic algorithm used for the optimization of checking – retrofit procedures in multidimensional technological processes. The improvement allows the genetic algorithm to find out the optimal solutions very quickly. The speed up of the genetic algorithm is achieved with the aid of (a) a smart initialization, (b) a fast calculation of technological process reliability, and (c) a specific adaptive fitness function.

INTRODUCTION

A technological process (TP) is considered multidimensional, if a number of defects of diverse types occur, are being detected and corrected simultaneously, within the process execution [3]. The TP is estimated by the probability of output zero-defects, as well as by the probabilities of zero-defect for each of the defect types. The tasks of TP-optimization, involve the choice of such a process structure that will provide the necessary output level of product quality given some certain cost limits [3]. Typical examples of such optimization tasks are the optimal allocation of control procedures in a TP, and the optimal choice of multiplicity of control-retrofit procedures. The first attempt of using a genetic algorithm (GM) in order to solve these nonlinear optimization problems was proposed in [2]. In that work, a GA manages to find out the optimal solutions, but the whole process proves very time-consuming. A fast greedy method (GM) for solving these problems is proposed in [1], but the GM is not always able to achieve the global optimal solutions. In this paper we propose an improvement of GA [2, 1]. The improved GA quickly produces the solutions, performing better than the GM.

PROBLEM STATEMENT

Let us introduce the following notations: m is number of diverse types of defects; n is the number of working technological operations; x_i is a number of checking-retrofit procedures after i-th operation, $i=\overline{1,n}$; $\mathbf{X}=(x_1,x_2,...,x_n)$ is a vector denoting controlled variables, used to determinate the TP structure; $p^1(\mathbf{X})$ is the probability of executing process $\mathbf{X}$ without any defect; $p_j^0(\mathbf{X})$ is the probability of executing process $\mathbf{X}$ with defect of type j; $C(\mathbf{X})$ is cost required for process $\mathbf{X}$. According to [1], the problem consists of finding that $\mathbf{X}$, for which

$$p^1(\mathbf{X})\to \min,\quad \text{subject to } C(\mathbf{X})\le C^* \text{ and } p_j^0(\mathbf{X})\le q_j,\ j=\overline{1,m}, \tag{1}$$

where q_j is the admissible probability threshold of the j-th type defect on process output,

$j=\overline{1,m}$; $C*$ is the admissible cost threshold of the TP. Let us call the task of optimal allocation of control-retrofit procedures as task A, and the optimal choice of the multiplicity of control-retrofit procedures as task B. For task A $x_i \in \{0,1\}$, and for task B $x_i \in \{0,1,2,...\}$.

MODELS OF MULTIDIMENSIONAL TP RELIABILITY

The relations connecting reliability figures with TP parameters [3] are described below. The reliability figures of the working operation A and the retrofit U are listed as follows:

$$\mathbf{P}_A = \begin{pmatrix} p_A^1 & p_A^{0_1} & \dots & p_A^{0_m} \\ 0 & 1 & \dots & 0 \\ \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \qquad \mathbf{P}_U = \begin{pmatrix} 1 & 0 & \dots & 0 \\ v_U^{1_1} & v_U^{0_1} & \dots & 0 \\ \dots \\ v_U^{1_m} & 0 & \dots & v_U^{0_m} \end{pmatrix},$$

where p_A^1 is the probability of zero-defect execution of the operation A; $p_A^{0_j}$ is the probability of execution of the operation A with j-th type of defect, $j=\overline{1,m}$; $v_U^{1_j}\left(v_U^{0_j}\right)$ is the probability of correcting the j-th type of defect during the execution of the retrofit U, $j=\overline{1,m}$.

The probabilities of type I (false alarm) and type II (defect loss) when a checking process ω is carried out, are represented by the following matrices:

$$\mathbf{K}_\omega^1 = \begin{pmatrix} k_\omega^{11} & 0 & \dots & 0 \\ 0 & k_\omega^{01_1} & \dots & 0 \\ \dots \\ 0 & 0 & \dots & k_\omega^{01_m} \end{pmatrix}, \qquad \mathbf{K}_\omega^0 = \begin{pmatrix} k_\omega^{10} & 0 & \dots & 0 \\ 0 & k_\omega^{00_1} & \dots & 0 \\ \dots \\ 0 & 0 & \dots & k_\omega^{00_m} \end{pmatrix},$$

where $k_\omega^{11}\left(k_\omega^{10}\right)$ is the probability of right (wrong) decision made about the absence of defects during the checking process ω; $k_\omega^{01_j}\left(k_\omega^{00_j}\right)$ is the probability of on detecting defects of type j during the checking process ω, $j=\overline{1,m}$.

The costs (or other resources) required for the execution of the working operation A, the checking procedure ω, and the retrofit U, are denoted by c_A, c_ω, and c_U, respectively.

A model of a potentially checked working operation is employed for task A. This operation is defined as follows: potentially checked working operation A is an working operation, which may be carried out with checking ω and retrofit U (if $x=1$) or without them (if $x=0$). The output reliability figures on this operation are calculated as follows [3]:

$$\mathbf{P}_A(x) = \mathbf{P}_A \cdot \left(\mathbf{K}_\omega^1 + \mathbf{K}_\omega^0 \cdot \mathbf{P}_U\right)^x, \tag{2}$$

$$c_A(x) = c_A + x \cdot \left(c_\omega + c_U \cdot \sum_{j=1}^{m} (p_A^1 \cdot k_\omega^{01} + p_A^{0_j} \cdot k_\omega^{00_j})\right). \tag{3}$$

A model of a working operation with x-multiple checking is employed for task B. This operation is defined as follows: an working operation with x-multiple checking is an working operation A in which, it is carried out an x times checking ω and retrofit U. The output reliability figures on this operation are calculated by (2)-(3) in an iterative scheme.

The reliability figures of the whole TP denoted as $\mathbf{X}$ are calculated as follows:

$$\mathbf{P}(\mathbf{X}) = \prod_{i=\overline{1,n}} \mathbf{P}_{A_i}(x_i), \qquad C(\mathbf{X}) = \sum_{i=\overline{1,n}} c_{A_i}(x_i). \tag{4}$$

GA FOR OPTIMIZATION THE CHECKING – RETROFIT PROCEDURES

To speed up GA, described in [1,2], we created: a) a smart procedure for the generation of the proper initial population; b) a fast algorithm calculating the reliability figures of the whole TP; c) a specific adaptive fitness function. The features of the proposed GA are listed below.

Genetic coding of variants. A TP-variant can be represented by a chromosome that contains n-genes: $\mathbf{X} = (x_1, x_2, ..., x_n)$, where genes correspond to the controlled variables for task (1).

Generation the initial population. The proposed initialization allows the creation of the population with high quality chromosomes. They satisfy all constrains of (1) and they have a low value of C. For generating the good chromosomes we use the so-called gradient of the checking-retrofit procedure from GM [1]. The gradient $\boldsymbol{\gamma}_i$ of the checking-retrofit procedure with number i indicates a relatively efficient factor of impacting this procedure into the TP [1]. For the task A, the value of x_i for any chromosome is generated randomly from the set $\{\underline{x}_i, 1\}$, where $\underline{x}_i = \begin{cases} 1, & \text{if } \exists j:\ p^{0j}_{A_i} > q_j \\ 0, & \text{otherwise} \end{cases}$. Additionally, we set up in 1 the value of x_i, if $\boldsymbol{\gamma}_i$ is a large number. For the task B the value of x_i is generated as a random integer from $[\underline{x}_i, \overline{x}_i \cdot \gamma_i^{\xi_i}]$, where $\overline{x}_i$ is the a-priori defined upper bound of checking multiplicity, and ξ_i is a random number from $[0,1]$.

Crossover and mutation. A uniform crossover with one cutting point and single-gene mutation [4] are employed. After the mutation, the value of gene x_i should not be below $\underline{x}_i$.

Fitness function: The following adaptive fitness function is proposed:

$$F(\mathbf{X}) = \begin{cases} 1/C(\mathbf{X}), & \text{if } \mathbf{X} \text{ is a feasible solution} \\ 1/(C(\mathbf{X}) \cdot D(\mathbf{X})), & \text{otherwise} \end{cases},$$

where $D(\mathbf{X}) = \dfrac{1}{m+1}\left(\max\left(0, \dfrac{C(\mathbf{X}) - C^*}{C^*}\right) + \sum_{j=\overline{1,m}} \left(\dfrac{\Delta b_j(\mathbf{X})}{\Delta b_j^{\max}} \right)^{\alpha} \right)$ is a penalty function;

α is an factor of importance of avoidance the defects;

$\Delta b_j(\mathbf{X}) = \max\left(0,\ p_j^0(\mathbf{X}) - q_j\right)$ means violation of j-th constraint by chromosome $\mathbf{X}$, $j = \overline{1,m}$;

$\Delta b_j^{\max} = \max\limits_{p=1,\, pop\_size} (\Delta b_j(\mathbf{X}_p))$ is max-violation of j-th constraint in the current population.

Fast calculation of the reliability figures. A profile of the GA shows that the most time-consuming part is the set of calculations by formulae (2)-(3). To speed up the process, we calculate in advance and store in memory the quantities $\mathbf{P}_{A_i}(x_i)$ and $c_{A_i}(x_i)$ for all possible values of x_i, $i=\overline{1,n}$. Accordingly, for the reliability figures of whole TP we simply apply formulae (4) for the corresponding (already computed) values of $\mathbf{P}_{A_i}(x_i)$ and $c_{A_i}(x_i)$.

Selection. We propose the implementation of a selection procedure based on the following elitist strategy: 1) find out the chromosome with the highest fitness and the chromosome with the highest fitness among the feasible ones and include them into the new population; 2) add the remaining chromosomes via the roulette wheel process [4].

COMPUTATIONAL EXPERIMENTS

The GA was tested for the tasks A and B, on two data sets. Data sets correspond to a TP with 7 diverse types of defects. The number of potential checking operations varied from 20 to 120. In the first data set, avoidance of two types of defects is the most important task, according to the very low levels of the admissible probabilities threshold for them. All the types of defects had approximately the same importance in the second data set. The data sets are available at www.ksu.vstu.vinnica.ua/shtovba/benchmark. Test results are shown in Table. As an alternative optimization routine we used the GM [1] with fast gradient computing. The GA provided better solutions than the GM, especially for the first data set. For large-scale tasks the GA finds out the optimal solutions quicker than the GM (Fig. 1).

Table – Testing rezults (infeasible solutions are in boldface)

Data set	Task	Algorithm	$p^1(\mathbf{X})$					
			n=20	n=40	n=60	n=80	n=100	n=120
First data set	Task A	GA	0.9681	0.8974	0.8399	0.7694	0.8031	0.7276
		GM	**0.9647**	**0.8952**	0.8369	0.7657	**0.7956**	**0.7245**
	Task B	GA	0.9745	0.9006	0.8375	0.7739	0.8057	0.7281
		GM	0.9701	**0.8975**	**0.8355**	0.7701	0.8040	**0.7261**
Second data set	Task A	GA	0.8520	0.6994	0.6130	0.4651	0.4061	0.2942
		GM	0.8493	0.6968	0.5680	0.4509	0.3671	0.2935
	Task B	GA	0.8652	0.7175	0.5895	0.4793	0.3974	0.3226
		GM	0.8643	0.7166	0.5838	0.4789	0.3972	0.3223

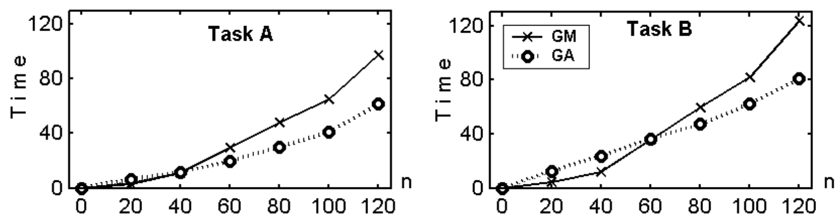

Figure 1 – Comparing GA and GM timing on the second data set (average for 20 runs)

Fig. 2 shows a satisfaction of the constraints by the solutions, founded by the GA and by the GM. In this figure y-axes correspond to the following factor of relative satisfaction:

$$\text{for } c(\mathbf{X}): \quad \psi(\mathbf{X}) = \frac{C^* - C(\mathbf{X})}{C^*} \cdot 100\%; \quad \text{for } p_j^0(\mathbf{X}): \quad \psi(\mathbf{X}) = \frac{q_j - p_j^0(\mathbf{X})}{q_j} \cdot 100\%, \quad j = \overline{1,m}.$$

A negative value of $\psi(\mathbf{X})$ means that solution $\mathbf{X}$ does not satisfy a constraint. There are 7 such solutions, founded by the GM-approach for the second data set.

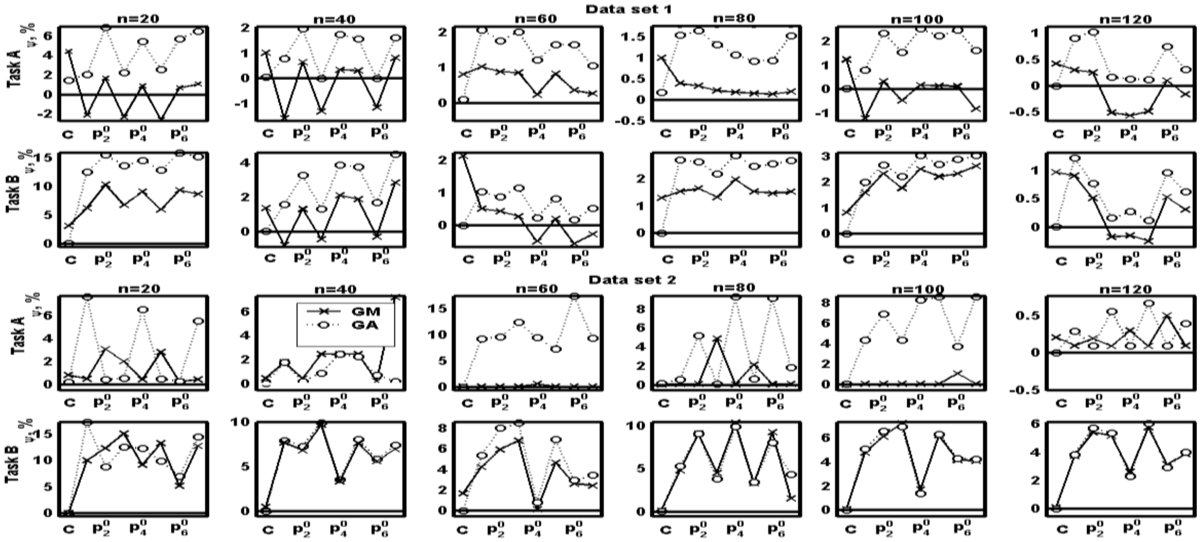

Figure 2 - A satisfaction the constrains by solutions, founded by the GA and by the GM

CONCLUSIONS

A fast GA of the optimization of the checking – retrofit procedures in multidimensional TP is proposed. The acceleration of the GA is achieved through (a) a smart procedure for the generation of a high-quality initial population, (b) a fast algorithm-calculation of the reliability figures of the whole TP, and (c) a specific adaptive fitness function proposed. Computational experiments carried out, show that the GA finds out better solutions, and proceeds faster than GM for large-scale instances. Future work is directed towards the comparison of the influence caused by the application of different competitive selection schemes on the speed of the optimization process for multidimensional TPs.

REFERENCES

1. Ротштейн О.П., Штовба С.Д., Дубіненко С.Б., Козачко О.М. Евристична оптимізація розстановки контрольних точок в технологічних процесах при багатовимірному просторі типів дефектів // Вісник ВПІ. - 2004.- №1.- С.54-62.
2. Ротштейн О.П., Штовба С.Д. Оптимізацiя багатовимірних технологічних процесів генетичними алгоритмами // Вимірювальна та обчислювальна техніка в технологічних процесах.- 1999.- №2.- С.7-13.
3. Ротштейн А.П., Кузнецов П.Д. Проектирование бездефектных человеко-машинных технологий. К.: Техніка. – 1992. – 180с.
4. Gen M., Cheng R. Genetic Algorithms and Engineering Design.- John Wiley &S ons.- 1997.-352 p.

Реферат (рус.)

БЫСТРЫЙ ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ДЛЯ ОПТИМИЗАЦИИ КОНТРОЛЯ В МНОГОМЕРНЫХ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССАХ
В статье предлагается улучшенный генетический алгоритм оптимизации контроля в многомерных технологических процессах. Предлагаемый генетический алгоритм находит оптимальные решения очень быстро за счет использования: а) процедуры инициализации хорошей исходной популяции; б) быстрого способа расчета показателей надежности технологического процесса; в) адаптивной фитнесс-функции.

Реферат (укр.)

ШВИДКИЙ ГЕНЕТИЧНИЙ АЛГОРИТМ ДЛЯ ОПТИМІЗАЦІЇ КОНТРОЛЮ В БАГАТОВИМІРНИХ ТЕХНОЛОГІЧНИХ ПРОЦЕССАХ
В статті запропоновано покращений генетичний алгоритм оптимізації контролю в багатовимірних технологічних процесах. Запропонований генетичний алгоритм знаходить оптимальні розв'язки дуже швидко за рахунок використання: а) процедури ініціювання доброї початкової популяції; б) швидкого способу розрахунку показників надійності; в) адаптивної фітнес-функії.

ВІДОМОСТІ ПРО АВТОРІВ:

Штовба Сергій Дмитрович,
к.т.н., доцент,
докторант кафедри комп'ютерних систем управління,
Вінницький національний технічний університет,
Хмельницьке шосе, 95, Вінниця, 21021
Тел. (0432)-440430, 440222, 519533.
Факс (0432)-465772
Ел. пошта: shtovba@ksu.vstu.vinnica.ua shtovba@svitonline.com

Козачко Олексій Миколайович,
аспірант кафедри комп'ютерних систем управління,
Вінницький національний технічний університет,
Хмельницьке шосе, 95, Вінниця, 21021
Тел. (0432)- 440222, 538466.
Факс (0432)-465772
Ел. пошта: okozachko@yahoo.com

Дуніас Георгіос,
доктор, профессор Егейського університету, Хіос, Греція.
Dounias Georgios,
Ph.D, Assoiate Professor,
Department of Financial and Management Engineering,
University of the Aegean, Chios, Greece,
g.dounias@aegean.gr