

В. В., ФРОЛОВ

ОПРЕДЕЛЕНИЕ РАЗМЕРА ПОПУЛЯЦИИ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ ЗАДАЧ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ В САПР

Харьковский национальный университет имени В. Н. Каразина

Предлагается метод определения размера популяции. Общий подход для определения размера популяции вытекает из утверждения, что хромосомы популяции должны содержать максимальное количество различных значений, которые покрывают большую часть области поиска. В основе метода лежит регрессионная модель, которая позволяет определить размер популяции в зависимости от допустимого количества значений независимой переменной. Регрессионная модель получена в результате обработки данных имитационного моделирования при формировании популяции для целевой функции с одной переменной. Задача решается для генетического алгоритма, где генотип представлен хромосомой в двоичном коде, а фенотип десятичным целочисленным кодом значений независимых переменных. Это позволяет моделировать формирование популяции без привязки к конкретным значениям переменных. Модель была получена для диапазона мощностей опорных множеств от 12 до 52, и позволяет прогнозировать размер популяции за пределами этого диапазона. Основная область использования данного метода это задачи дискретной оптимизации с целевыми функциями нескольких переменных, где области допустимых значений конечны и имеют небольшую размерность.

Ключевые слова. Популяция, генетический алгоритм, САПР, генотип, фенотип, хромосома.

Введение

Размер популяции – это один из основных параметров генетического алгоритма (ГА), который определяет эффективность поиска оптимального значения функции приспособленности. Общий подход для определения размера популяции вытекает из утверждения, что хромосомы популяции должны содержать максимальное количество различных значений, которые покрывают большую часть области поиска.

В работе [1] отмечается, что для ГА с одноточечным кроссовером и мутацией, размер популяции N_p зависит от длины «выживших» шаблонов $L(s)$ и может быть определен по формуле $N_p = 2^{L(s)/2}$. Цой Ю. Р. в [2] выделяет три основных подхода к определению размера популяции: определение фиксированного размера популяции по экспериментальным данным анализа работы ГА на тестовых функциях; определение размера популяции на основе анализа функции приспособленности, где выделяется наиболее значимый параметр, влияющий на значение функции; адаптивный – под-

ход, который предполагает изменение размера популяции в процессе работы ГА на основе анализа динамики изменений средней функции приспособленности в популяции.

В работе [3] предлагается определять минимальный размер популяции в зависимости от L размера хромосомы в битах и вероятности когда сумма выделенного бита хромосомы по всем хромосомам популяции N меньше размера этой популяции по формуле

$$N = \log_{0,5} \left(\frac{1 - \sqrt{P(S_L^N)}}{2} \right),$$

где $P(S_L^N)$ – вероятность события, попадания суммы по битам в диапазон.

В работах [4–10] проводится анализ влияния размера популяции на работу ГА в конкретной предметной области. Например, в [4] такой анализ выполняется из условия, что оптимальный размер должен лежать в диапазоне от 100 до 500 особей. Здесь авторы указывают, что исходный размер популяции соответствует линейному закону и для каждого дополнитель-

ного параметра модели увеличивается на 30 особей. В результате, для анализа моделей нефтяных месторождений он колеблется от 140 до 250 особей, по данным авторов.

Чтобы эффективно использовать ГА в САПР различного назначения необходимо учитывать, что инженер проектировщик должен пользоваться простыми рекомендациями для назначения параметров ГА. Решением, в данном случае, может быть максимальная формализация выбора размера популяции для настройки ГА до начала работы. Основой формализации может быть оценка вероятности появления максимального количества значений из всей допустимой области. Ограничения такого подхода заключаются в размерах опорных множеств независимых переменных, используемых для решения задач дискретной оптимизации.

Рассмотрим это противоречие на простом примере. Зададим целевую функцию от 22 переменных, каждая из которых имеет по 22 допустимых значений. Область поиска, в данном случае, определяется из прямого произведения этих множеств, и будет иметь $22^{22} \approx 3 \cdot 10^{29}$ сочетаний значений переменных. Тогда, в популяции необходимо иметь особи, которые будут содержать в себе все допустимые значения переменных из опорных множеств. Если убрать фактор вероятности появления повторяющихся значений в сочетаниях при формировании популяции, то имеем популяцию 22 особи, для которых множество значений каждой переменной, сформированное из данного множества особей, эквивалентно опорному множеству значений этой переменной. Популяция генерируется случайным образом, поэтому ее размер должен быть больше 22 особей, чтобы исключить повторяющиеся значения.

В связи с этим, цель данного исследования заключается в повышении эффективности работы ГА для САПР за счет определения размера популяции, обеспечивающего наибольшее разнообразие генетического материала, в зависимости от количества независимых переменных и мощностей их опорных множеств, при условии формирования начальной популяции случайным образом с равной вероятностью появления значений. Опорные множества значений переменных, конечны и имеют мощность не более 100, исходя из особенностей решения технологических задач.

Метод определения размера популяции для дискретной переменной

Решать задачу вычисления граничных размеров популяции будем для ГА, предложенного в работе [11] для проектирования технологических систем. Особенностью данного алгоритма является наличие нескольких этапов кодирования значений переменных из опорных множеств, которые вытекают из определений генотипа и фенотипа.

Генотип – состоит из хромосом, представленных двоичным кодом. С генотипом напрямую работают операторы генетического алгоритма. Фенотип – десятичный целочисленный код значений независимых переменных. Набор преобразований, при работе алгоритма, следующий (1): значение свойства реального объекта в фенотип; фенотип в генотип; применение операторов генетического алгоритма к генотипу; генотип в фенотип; фенотип в значение; вычисление функции приспособленности для реальных значений.

$$\left\{ \begin{array}{l} gf : B^n \rightarrow Z, Z = \\ \left\{ z : (\forall z \in \mathbb{Z}^+) \wedge (|Z| = |R|) \wedge (r \leftrightarrow z) \right\}, \\ fz : Z \rightarrow R, R = \{ r : P(r) = 1 \}, \\ fg : Z \rightarrow B^n, \\ zf : R \rightarrow Z, \end{array} \right. \quad (1)$$

где gf – преобразование генотип – фенотип; fz – преобразование фенотип – значение; Z – множество целочисленных кодов значений, которые удовлетворяют правилам кодирования P ; R – множество значений, которые определяют свойства объекта согласно правилу кодирования P ; B – исходное множество для формирования хромосомы $B = \{0, 1\}$, n – длина хромосомы в аллелях.

Такой подход при решении задач дискретной оптимизации позволяет работать с фенотипом для определения размера популяции, без привязки к конкретной предметной области, что выгодно для САПР. Решим задачу определения размеров популяции с помощью имитационного моделирования формирования популяции для генетического алгоритма, при следующих условиях: количество значений в опорных множествах $Mval = \{12, 22, 32, 42, 52\}$; количество переменных в целевой функции – одна.

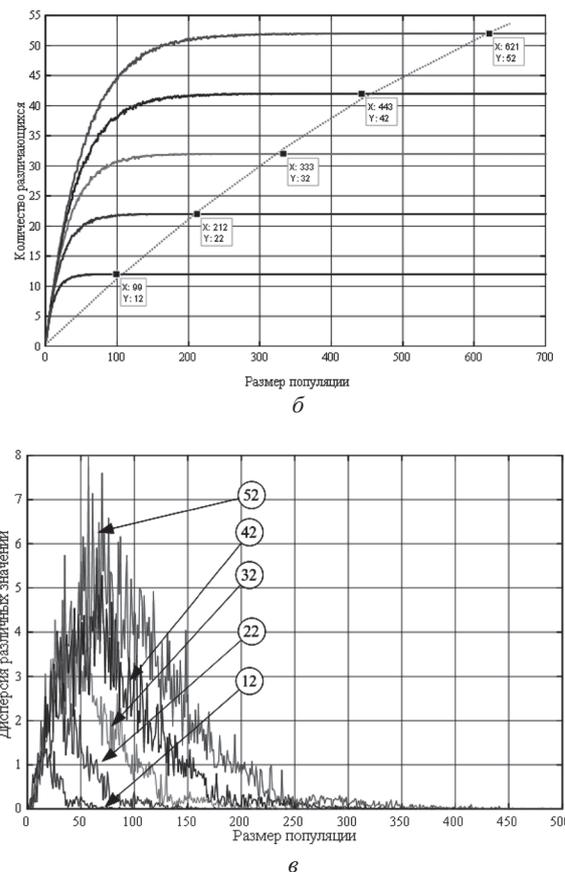
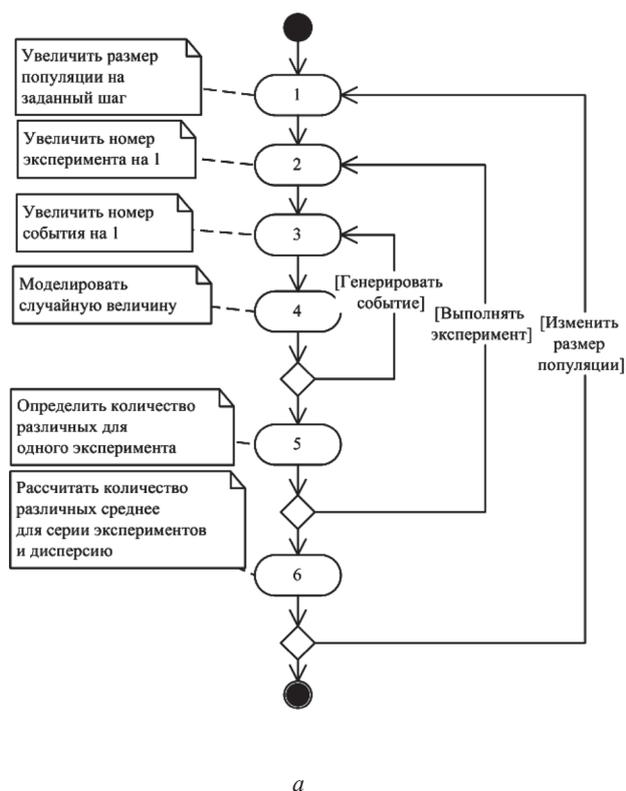


Рис. 1. *a* – обобщенный алгоритм имитационного моделирования; *б* – тенденции изменения средних различающихся значений для множества $Mval$; *в* – тенденции изменения дисперсий для множества $Mval$

Обобщенный алгоритм имитационного моделирования, представленный на рис. 1, *a*, предполагает, для каждого нового размера популяции, генерацию случайных значений из опорного множества значений, которое согласно (1), будет изменяться от 1 до $Mval_i$, где $i = \overline{1,5}$. События в блоке 4 на рис. 1, *a* генерируются с равной вероятностью. Для каждого размера популяции рассчитывается среднее число различных значений, отображенных на рис. 1, *б*, их дисперсия, представленная на рис. 1, *в*, и среднее квадратическое отклонение. Согласно рис. 1, *б*, максимальный размер популяции может быть определен из условия приближения среднего различающихся значений к $Mval_i$. Точка, после которой увеличение размера популяции не приводит к изменению среднего будет максимально допустимым размером популяции для каждого $Mval_i$.

На рис. 1, *б* выделены указанные точки для каждого значения из множества $Mval$ и тренд, определяющий границу максимальных размеров популяций. На основе этих данных в системе Matlab построена модель полинома тре-

тьей степени, для определения максимального размера популяции, изображенная на рис. 2. На рис. 2 по оси абсцисс откладывается мощность опорного множества, а по оси ординат максимальное значение размера популяции. Множество x_k определяется на основе анализа графика изменения дисперсии на рис. 1, *в*.

В этом случае граничные точки – это первые точки, в которых дисперсия равна нулю. В результате, происходит разделение координатной плоскости на две области, где $D > 0$ и $D = 0$, на рис. 3, *a*.

Разделение на две области выполняется с помощью решающего правила, исходя из следующего:

1. Средняя дисперсия справа от границы на рис. 3, *a* равна нулю;

2. Перемещая разделительную линию в пределах изменения размера популяции и перемножая средние дисперсии слева и справа от границы, получим функцию, которая гарантированно будет определять максимальную границу размера популяции (см. рис. 3, *б*) в момент достижения оси абсцисс.

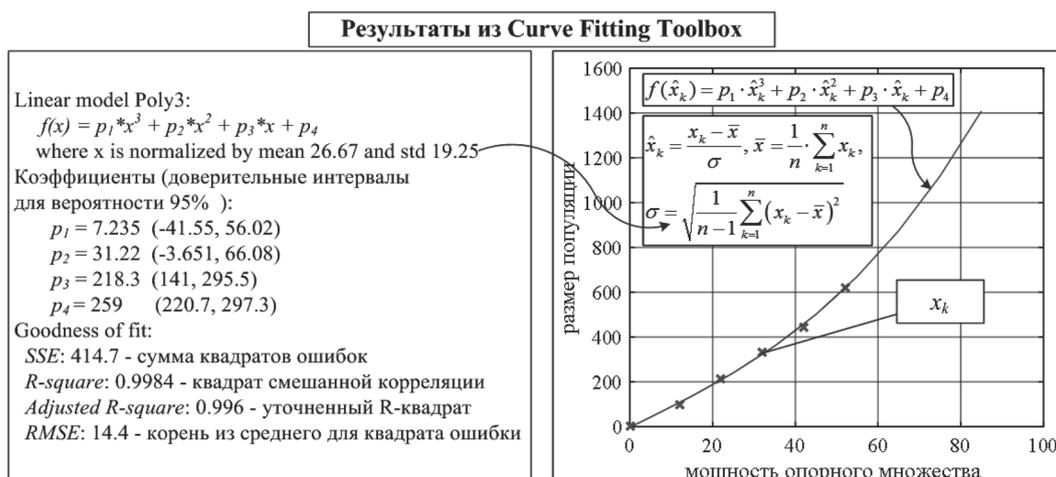


Рис. 2. Модель для определения размера популяции

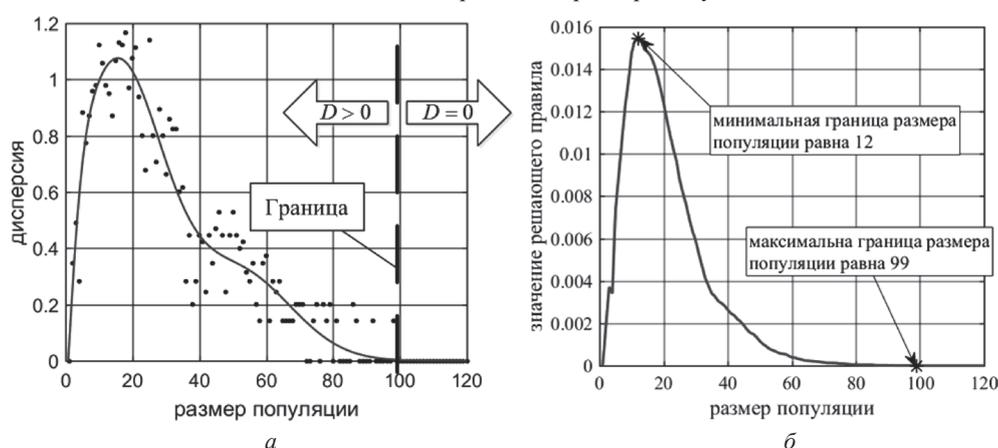


Рис. 3. а – разделение на две области, б – график изменения решающего правила

Формула решающего правила, тогда будет:

$$y(k) = \left(\frac{1}{k} \sum_{j=1}^k D_j \right) \left(\frac{1}{N - (k+1)} \sum_{j=k+1}^N D_j \right) = \frac{1}{k(N-1) - k^2} \left(\sum_{j=1}^k D_j \right) \left(\sum_{j=k+1}^N D_j \right),$$

где k – размер популяции; D_j – дисперсия повторяющихся значений при заданном размере популяции; N – предельное значение размера популяции для эксперимента. График изменения решающего правила изображен на рис.3б, здесь можно выделить две граничные точки: первая – максимум дисперсии, после которого значение дисперсии начинает убывать, его можно использовать для определения нижней минимально необходимой границы размера популяции; вторая – точка, после которой дисперсия равна нулю, которая используется для определения максимальной границы размера популяции.

Например, для множества из 12 значений на рис. 3, б этот диапазон будет от 12 до 99,

а для 22 значений этот диапазон будет от 27 до 212. Результаты моделирования для большего количества переменных в таблице 1 показывают, что определять размер популяции можно по лимитирующему опорному множеству, согласно модели рис. 2.

Таблица 1. Максимальный размер популяции для функций многих переменных

Значения/Переменные	12	22	32	42	52
12	113	108	98	116	135
22	235	207	235	212	242
32	254	332	373	313	300
42	402	531	447	424	551
52	546	590	581	589	614

Теоретическое обоснование метода

Эта задача относится к классу задач теории вероятности о повторных независимых испытаниях. Здесь проводится n независимых испытаний, при которых появляется хотя бы одно число из опорного множества M . Тогда вероят-

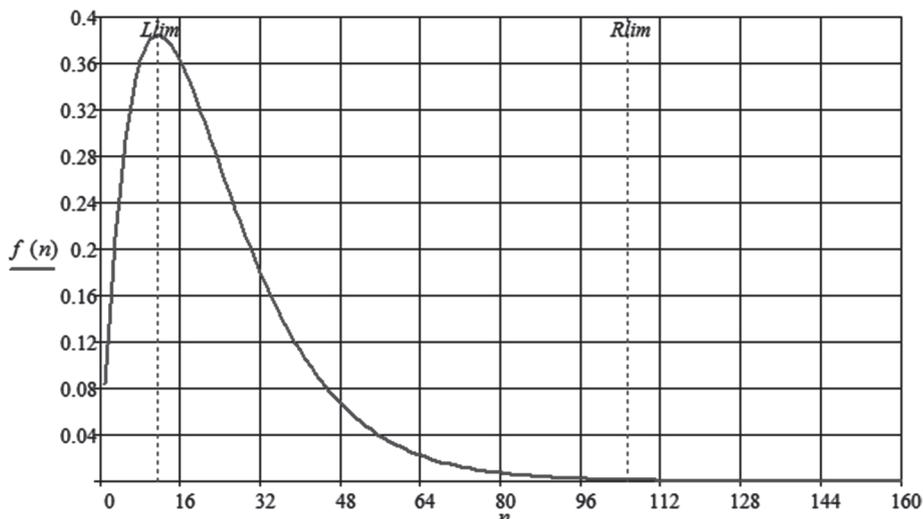


Рис. 4. Изменения вероятности от числа испытаний: L_{lim} – минимально необходимый размер популяции; R_{lim} – максимальный размер популяции

ность этого события зависит от мощности множества M и равна $p = 1/|M|$, где $|M|$ – мощность опорного множества. Вероятность, того, что появятся другие числа из этого множества $q = 1 - p$. По формуле Бернулли можно определить вероятность числа положительных исходов при n независимых испытаниях. В нашем случае, при заданном числе испытаний должен быть один положительный исход, т. е. вычисления по формуле Бернулли сводятся к частному случаю, где

$$P_n(1) = C_n^1 p q^{n-1} = n p q^{n-1} = n p (1-p)^{n-1}.$$

Учитывая, что вероятность зависит от мощности опорного множества, получим

$$P_n(1) = n \frac{1}{|M|} \left(1 - \frac{1}{|M|}\right)^{n-1}.$$

Рассмотрим характер изменения вероятности при заданном значении $|M| = M_{val_1}$ на рис. 4, для следующей функции

$$f(n) = n \frac{1}{|M|} \left(1 - \frac{1}{|M|}\right)^{n-1},$$

при диапазоне изменения испытаний от $n_{min} = 1$ до $n_{max} = 160$, поскольку из предыдущего известно, что максимальная граница для 12 равна 99.

По графику на рис. 4 можно определить минимальный размер популяции, это будет максимум функции $L_{lim} = 11,493$. Максимальный размер популяции, определяется из условия, что отклонение графика от оси абсцисс

будет не более 0,001. Решаем трансцендентное уравнение с заданной областью определения переменной

$$n \frac{1}{|M|} \left(1 - \frac{1}{|M|}\right)^{n-1} = 0,001, L_{lim} < n \leq n_{max}.$$

В результате, максимальный размер популяции $R_{lim} = 105,356$.

Вероятность появления всех значений при заданном размере популяции будет

$$p = 1 - P_{R_{lim}}(1) = 0,999.$$

Вероятность появления всех значений при максимальном размере популяции, полученном, в результате имитационного моделирования будет

$$p = 1 - P_{99}(1) = 0,998.$$

Выше приведенное показывает, что определение размера популяции на основе анализа дисперсии различающихся значений по сравнению с аналитическим методом, дает погрешность допустимую в инженерных расчетах.

Практическая реализация и сравнительный анализ методов

Предлагаемый метод используется для автоматического определения размера популяции в САПР расчета размерных цепей. САПР предназначена для выполнения проектного расчета размерных цепей, когда по допуску замыкающего звена необходимо подобрать допуски составляющих звеньев. Например, для связанной размерной цепи, состоящей из 11 звеньев,

задача подбора с помощью генетического алгоритма решается в области поиска, где хромосома состоит из 11 генов, каждый ген имеет опорное множество, состоящее из 12 значений допусков. Длина гена в этом случае составляет 4 бита, а общая длина хромосомы 44 бита. Область поиска имеет $12^{11} = 7,43 \cdot 10^{11}$ значений.

Предлагаемые в работах [1][3] методы оценки, позволяют вычислить ориентировочный размер популяции для генетического алгоритма с одноточечным кроссовером, соответственно, получаем:

1. Согласно [1] для простого генетического алгоритма, если длина хромосомы 44 аллели минимально возможная длина шаблона будет $L(s) \approx 18$, тогда размер популяции $Np = 2^{\frac{18}{2}} = 512$ особей;

2. Согласно [3] с учетом вероятности 0,999 и длины хромосомы, получаем размер популяции

$$N = \log_{0,5} \left(\frac{1 - \sqrt[44]{0,999}}{2} \right) \approx 16.$$

Предлагаемый в данной статье метод дает следующие диапазоны размера популяции: теоретический расчет от 11,493 до 105,3564; моделирование от 12 до 99; упрощенный расчет с коэффициентом 10 до 120.

Серия из пяти экспериментов, при количестве тактов генетического алгоритма 1000 и одноточечном кроссовере, показывает следующие средние значения (в скобках среднеквадратическое отклонение по пяти экспериментам):

– размер популяции 16 особей – погрешность 51 мкм (40,44), время расчета 3,8 сек (0,44), смена популяций 9,4 (0,5);

– размер популяции 512 особей – погрешность 15 мкм (6,1), время расчета 47,2 сек (2,2), смена популяций 0;

– размер популяции 120 особей – погрешность 15 мкм (3,9), время 13,4 сек (0,5), смена популяций 2,4 (0,9).

Эксперименты показывают, что предлагаемый метод за приемлемое время дает результаты не хуже [1], поскольку средняя погрешность одинакова, а время вычислений меньше в 3,5 раза, и значительно лучше [3].

Заключение

Метод определения граничных размеров популяции, позволяет на основе анализа дисперсии различающихся значений определить максимальный размер популяции. При этом размере популяции с высокой вероятностью будет генерироваться весь набор значений из опорного множества, что обеспечивает максимальное разнообразие генетического материала в популяции. Метод не зависит от особенностей реализации операторов генетического алгоритма. Математическая модель для определения максимального размера популяции позволяет вычислить, с точностью достаточной для инженерных расчетов, по мощности опорного множества независимой переменной, необходимый размер популяции. Приближенно размер популяции можно определить, умножив мощность опорного множества на 10, согласно рис. 1, б. Для функции нескольких переменных определение размера популяции выполняется по лимитирующему опорному множеству, с точностью достаточной для инженерных расчетов.

ЛИТЕРАТУРА

1. Емельянов В. В., Курейчик В. В., Курейчик В. М. Теория и практика эволюционного моделирования. – М.: ФИЗМАТЛИТ, 2003. – 432 с.
2. Цой Ю. Р. Нейроэволюционный алгоритм и программные средства для обработки изображений. [Электронный ресурс]: дис. на соискание ученой степени канд. техн. наук: спец. 05.13.01 – Системный анализ, управление и обработка информации (отрасль: информация и информационные системы) / Ю. Р. Цой. – Томск, 2007. – 209 с. – Режим доступа: http://qai.narod.ru/Dissertation/tsoy_phd.thesis.zip (дата обращения: 27.12.2017).
3. Чипига А. Ф., Петров Ю. Ю. Оценка выбора размера популяции в простом генетическом алгоритме. [Электронный ресурс] // Информационное противодействие угрозам терроризма. – 2005г. – № 5. – С. 26–29. – Режим доступа: <https://elibrary.ru/item.asp?id=9572194> (дата обращения: 23.03.2018).
4. Семёнычев Е. В., Куркин Е. И., Данилова А. А. Выбор параметров генетических алгоритмов в задачах параметрической идентификации нелинейных моделей динамики // Вестник самарского муниципального института управления. – 2013г. – № 1(24). – С. 130–140. – Режим доступа: <https://elibrary.ru/item.asp?id=19024835> (дата обращения: 27.12.2017).
5. Жеребцова О. В., Зубьяк Д. Р. Генетический алгоритм и его свойства для решения задачи поиска пустой прямоугольной области максимальной площади // Вестник Волжского университета им. В. Н. Татищева. – 2011г. – № 17. – С. 160–164. – Режим доступа: <https://elibrary.ru/item.asp?id=16858033> (дата обращения: 23.03.2018).

6. Мухамедиева Д. Т. Анализ особенностей генетических алгоритмов // Проблемы вычислительной и прикладной математики. – 2015 г. – № 1(1). – Режим доступа: <https://elibrary.ru/item.asp?id=25625477> (дата обращения: 23.03.2018).
7. Сапрыкин А. Н., Акиннина К. Д., Сапрыкина Е. Н. Нахождение оптимального числа полезных особей в популяции и конвергируемых поколений генетического алгоритма оптимизации простых многоэкстремальных функций // ACTUALSCIENCE. – 2016 г. – № 11. – Том 2. – С. 168–169. – Режим доступа: <https://elibrary.ru/item.asp?id=27542766> (дата обращения: 23.03.2018).
8. Сапрыкин А. Н., Акиннина К. Д. Основные характеристики генетического алгоритма компоновки блоков электронно-вычислительных средств // Информационные технологии, межвузовский сборник научных трудов. – Рязань: Рязанский государственный радиотехнический университет. – 2017. – С. 128–132. – Режим доступа: <https://elibrary.ru/item.asp?id=30470701> (дата обращения: 24.03.2018).
9. Сапрыкин А. Н., Акиннина К. Д., Сапрыкина Е. Н. Оптимальное число конвергируемых поколений и полезных особей при варьирующихся значениях вероятности мутации генетического алгоритма на малых пространствах поиска // Современные технологии в науке и образовании – СТНО-2017 Сборник трудов II международной научно-технической и научно-методической конференции Том. 4. – Рязань: Рязанский государственный радиотехнический университет. – 2017. – С. 248–252. – Режим доступа: <https://elibrary.ru/item.asp?id=29875353> (дата обращения: 24.03.2018).
10. Федоренко К. В., Оловяников А. Л. Исследование основных параметров генетического алгоритма применительно к задаче поиска оптимального маршрута // Вестник Государственного университета морского и речного флота имени адмирала С. О. Макарова. – 2017. – Т. 9. – № 4. – С. 714–723. – Режим доступа: <https://elibrary.ru/item.asp?id=29922476> (дата обращения: 24.03.2018).
11. Фролов В. В. Особенности реализации генетического алгоритма для проектирования технологических систем механической обработки // Восточно-Европейский журнал передовых технологий. – 2012. – № 3/9(57). – С. 60–64.

REFERENCES

1. Emel'janov V. V., Kurejchik V. V., Kurejchik V. M. Teorija i praktika jevolucionno-go modelirovanija. – M.: FIZMATLIT, 2003. – 432 s.
2. Соj Ju. R. Nejroevoljucionnyj algoritm i programmnye sredstva dlja obrabotki izobrazhenij. [Jelektronnyj resurs]: dis. na soiskanie uchenoj stepeni kand. tehn. nauk: spec. 05.13.01 – Sistemnyj analiz, upravlenie i obrabotka informacii (otrasl': informacija i informacionnye sistemy) / Ju. R. Соj. – Tomsk, 2007. – 209 s. – Rezhim dostupa: http://qai.narod.ru/Dissertation/tsoy_phd.thesis.zip (data obrashhenija: 27.12.2017).
3. Chipiga A. F., Petrov Ju. Ju. Ocenka vybora razmera populjacji v prostom geneti-cheskom algoritme. [Jelektronnyj resurs] // Informacionnoe protivodejstvie ugrozam terrorizma. – 2005 g. – № 5. – С. 26–29. – Rezhim dostupa: <https://elibrary.ru/item.asp?id=9572194> (data obrashhenija: 23.03.2018).
4. Semjonychev E. V., Kurkin E. I., Danilova A. A. Vybora parametrov geneticheskih algoritmov v zadachah parametriceskoj identifikacii nelinejnyh modelej dina-miki // Vestnik samarskogo municipal'nogo instituta upravlenija. – 2013 g. – № 1(24). – С. 130–140. – Rezhim dostupa: https://elibrary.ru/download/elibrary_19024835_95315377.pdf (data obrashhenija: 27.12.2017).
5. Zherebcova O. V., Zub'jak D. R. Geneticheskiy algoritm i ego svojstva dlja reshenija zadachi poiska pustoj prjamougol'noj oblasti maksimal'noj ploshhadi // Vestnik Volzh-skogo universiteta im. V. N. Tatishheva. – 2011 g. – № 17. – С. 160–164. – Rezhim dostupa: <https://elibrary.ru/item.asp?id=16858033> (data obrashhenija: 23.03.2018).
6. Muhamedieva D. T. Analiz osobennostej geneticheskih algoritmov // Problemy vychislitel'noj i prikladnoj matematiki. – 2015 g. – № 1(1). – Rezhim dostupa: <https://elibrary.ru/item.asp?id=25625477> (data obrashhenija: 23.03.2018).
7. Saprykin A. N., Akinina K. D., Saprykina E. N. Nahozhdenie optimal'nogo chisla poleznyh osobej v populjácii i konvergiruemyh pokolenij geneticeskogo algoritma op-timizacii prostyh mnogoekstremal'nyh funkcij // ACTUALSCIENCE. – 2016 g. – № 11. – Том 2. – С. 168–169. – Rezhim dostupa: <https://elibrary.ru/item.asp?id=27542766> (data obrashhenija: 23.03.2018).
8. Saprykin A. N., Akinina K. D. Osnovnye harakteristiki geneticeskogo algoritma komponovki blokov jelektronno-vychislitel'nyh sredstv // Informacionnye tehnologii, mezhvuzovskij sbornik nauchnyh trudov. – Rjazan': Rjazanskij gosudarstvennyj radio-tehniceskij universitet. – 2017. – С. 128–132. – Rezhim dostupa: <https://elibrary.ru/item.asp?id=30470701> (data obrashhenija: 24.03.2018).
9. Saprykin A. N., Akinina K. D., Saprykina E. N. Optimal'noe chislo konvergiruemyh pokolenij i poleznyh osobej pri var'irujushhijah znachenijah verojatnosti mutacii geneticeskogo algoritma na malyh prostranstvah poiska // Sovremennye tehnologii v nauke i obrazovanii – STNO-2017 Sbornik trudov II mezhdunarodnoj nauchno-tehniceskoj i nauchno-metodiceskoj konferencii Tom. 4. – Rjazan': Rjazanskij gosudarstvennyj radiotehniceskij universitet. – 2017. – С. 248–252. – Rezhim dostupa: <https://elibrary.ru/item.asp?id=29875353> (data obrashhenija: 24.03.2018).
10. Fedorenko K. V., Olovjannikov A. L. Issledovanie osnovnyh parametrov geneti-cheskogo algoritma primenitel'no k zadache poiska optimal'nogo marshruta // Vestnik Gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S. O. Makarova. – 2017. – Т. 9. – № 4. – С. 714–723. – Rezhim dostupa: <https://elibrary.ru/item.asp?id=29922476> (data obrashhenija: 24.03.2018).
11. Frolov V. V. Osobennosti realizacii geneticeskogo algoritma dlja proektirovanija tehnologiceskih sistem mehaniceskoj obrabotki // Vostochno-Evropskij zhurnal peredovyh tehnologij. – 2012. – № 3/9(57). – С. 60–64.

Поступила
06.02.2018

После доработки
10.07.2018

Принята к печати
31.08.2018

Frolov V. V., prof., D. Sc., Assoc.

DETERMINING THE SIZE OF THE POPULATION OF THE GENETIC ALGORITHM FOR THE PROBLEMS OF DISCRETE OPTIMIZATION IN CAD

vvicfrol@rambler.ru

A method for determining the size of a population is proposed. The general approach for determining the size of a population follows from the statement that the chromosomes of a population must contain the maximum number of different values that cover most of the search area. The method is based on the regression model, which allows you to determine the size of the population, depending on the permissible number of values of the independent variable. The regression model is obtained as a result of processing simulation data in the formation of a population for a single-variable objective function. The problem is solved for a genetic algorithm, where the genotype is represented by a chromosome in binary code, and the phenotype by a decimal integer code of values of independent variables. This allows you to model the formation of a population without reference to specific values of variables. The model was obtained for the power range of the reference sets from 12 to 52, and allows to predict the size of the population beyond the limits of this range. The main area of use of this method is discrete optimization problems with objective functions of several variables, where the ranges of admissible values are finite and have a small dimension.

Keywords. *Population, genetic algorithm, CAD, genotype, phenotype, chromosome.*



Фролов В. В., д.т.н., доцент, профессор кафедры теоретической и прикладной информатики Харьковского национального университета им. В. Н. Каразина. E-mail: vvicfrol@rambler.ru.

Frolov V. V., D. Sc., Associate Professor, Professor of the Department of Theoretical and Applied Informatics of the Kharkov National University. V. N. Karazin.

E-mail: vvicfrol@rambler.ru