

## Раздел I. Эволюционное моделирование, генетические и бионические алгоритмы

УДК 321.3

В.М. Курейчик

### МОДИФИЦИРОВАННЫЕ ГЕНЕТИЧЕСКИЕ ОПЕРАТОРЫ\*

*В работе рассматриваются генетические алгоритмы (ГА). Основным блоком в ГА является блок генетических операторов. Он выполняет преобразование популяции альтернативных решений задач оптимизации и искусственного интеллекта. В работе анализируются новые генетические операторы. Это позволяет формализовать процесс поиска, сократить размерность задач, время получения новых альтернативных решений и качество результата.*

*Генетический алгоритм; генетические операторы; операторы кроссинговера; мутации; инверсии; сегрегации; операторы на основе «Решета Эратосфена», «треугольника Паскаля»; эффект Основателя.*

V.M. Kureichik

### MODIFIED GENETIC OPERATORS

*Genetic algorithms (GA) were considered in this paper. The main block in GA is the block of genetic operators. It makes changes in population of optimization and artificial intelligence problems alternative solutions. New genetic operators were analyzed. It allows formalizing the search process, reducing the problems dimension and the time of getting new alternative solutions and the result quality.*

*Genetic algorithm; genetic operator; crossover operator; mutation operator; segregation operator; "Eratosfen's strainer" operator; "Pascal's triangle"; "Effect of the Founder".*

**Введение.** В настоящее время во многих областях науки и техники эффективно применяются алгоритмы, инспирированные природными системами. К ним относятся генетические, эволюционные, «муравьиные», моделирования отжига, «пчелиных колоний» и т.п. [1-7]. В работе рассматриваются генетические алгоритмы (ГА) [1-4]. Основным блоком в ГА является блок генетических операторов. Он выполняет преобразование популяции альтернативных решений задач оптимизации и искусственного интеллекта. В результате моделирования эволюции получается набор локально-оптимальных решений задач оптимизации. В работе анализируются новые генетические операторы. Это позволяет формализовать процесс поиска, сократить размерность задач, время получения новых альтернативных решений и качество результата.

**1. Некоторые сведения из теории генетических алгоритмов.** Генетический алгоритм (ГА) моделирует эволюцию. При этом каждой особи (их часто называют хромосомами) ставится в соответствие одно альтернативное решение.

---

\* Работа выполнена при поддержке: РФФИ (гранты № 09-01-00492, № 10-01-90017), г/б № 2.1.2.1652, г/б № 2.1.2.4595.

В результате имеем популяцию хромосом, т.е. набор альтернативных решений рассматриваемой задачи. Для задач САПР, искусственного интеллекта, оптимизации, принятия решений существует огромное пространство поиска допустимых решений. Генетический алгоритм позволяет анализировать разные подобласти пространства поиска параллельно и независимо от локальных экстремумов целевой функции (ЦФ) задачи. Следует отметить, что увеличение количества точек, соответствующих подобластям поиска не гарантирует сходимости к глобальному оптимуму, хотя повышает вероятность такой сходимости.

Эволюционный процесс представляется как способность «лучших» хромосом оказывать большее влияние на состав новой популяции на основе длительного выживания из более многочисленного потомства. Основные этапы эволюционного поиска следующие:

1. Случайным, направленным или комбинированным методом конструируется начальная популяция. Вводится точка отсчета поколений  $t = 0$ . Вычисляется приспособленность (целевая функция) каждой хромосомы в популяции, а затем средняя приспособленность (средняя целевая функция) всей популяции.

2. Производится переход к новому поколению и устанавливается  $t = t + 1$ . Производится выбор двух или некоторого множества родителей (хромосом) для реализации генетических операторов. Первым обычно выполняется оператор кроссинговера (ОК). Существует большое число ОК. В настоящее время приоритет отдается ОК со знанием о структуре решаемой задачи. Простейшим является ОК, выполняемый случайным образом пропорционально приспособляемости родителей.

3. Формируется генотип потомков. Для этого с заданной вероятностью производится оператор кроссинговера над генотипами выбранных хромосом. Далее с вероятностью 0,5 выбирается один из потомков  $P_i(t)$  и сохраняется как член новой популяции. После этого к  $P_i(t)$  последовательно применяется оператор инверсии, а затем мутации, сегрегации, транслокации и т.п. с заданными вероятностями. Полученный генотип потомка сохраняется как  $P_k(t)$ .

4. Определяется количество хромосом для исключения их из популяции, чтобы ее размер оставался постоянным. В необходимом случае ЛППР задает размер изменения популяции в каждом поколении. Текущая популяция обновляется заменой отобранных хромосом на потомков  $P_k(t)$ .

5. Производится определение приспособленности (целевой функции) и пересчет средней приспособленности всей полученной популяции.

6. Если  $t = t_{\text{заданному}}$ , то переход к 7, если нет, то переход к 2.

7. Конец работы.

Данный алгоритм является упрощенным модифицированным «репродуктивным планом Д. Холланда» [3].

В каждой генерации генетического алгоритма хромосомы являются результатом применения некоторых генетических операторов. Оператор – это языковая конструкция, представляющая один шаг из последовательности действий или набора описаний алгоритма [2,7]. Генетический оператор по аналогии с оператором алгоритма – это средство отображения одного множества на другое.

Рассмотрим основные операторы генетических алгоритмов [8].

**Оператор репродукции (селекция) (ОР)** – это процесс, посредством которого выбираются хромосомы (альтернативные решения), имеющие более высокое значение ЦФ. Элементы, выбранные для репродукции, обмениваются генетическим материалом, создавая аналогичных или различных потомков. Оператор репродукции считается эффективным, если он создает возможность перехода из одной по-

добласти альтернативных решений области поиска в другую. Это повышает вероятность нахождения глобального оптимума целевой функции.

**Оператор кроссинговера (ОК)** – это языковая конструкция, позволяющая на основе преобразования (скрещивания) хромосом родителей (или их частей) создавать хромосомы потомков. Существует огромное число ОК, так как их структура в основном и определяет эффективность ГА.

**Оператор мутации (ОМ)** – это языковая конструкция, позволяющая на основе преобразования родительской хромосомы (или ее части) создавать хромосому потомка.

**Оператор инверсии (ОИ)** – это языковая конструкция, позволяющая на основе инвертирования родительской хромосомы (или ее части) создавать хромосому потомка. При его реализации случайным образом определяется одна или несколько точек разреза (инверсии), внутри которых элементы инвертируются.

**Оператор транслокации (ОТ)** – это языковая конструкция, позволяющая на основе скрещивания и инвертирования из пары родительских хромосом (или их частей) создавать две хромосомы потомков. Существуют большое число других видов оператора транслокации.

**Оператор транспозиции (ОТР)** – это языковая конструкция, позволяющая на основе преобразования и инвертирования выделяемой части родительской хромосомы создавать хромосому потомка. Отметим, что существует большое количество модификаций оператора транспозиции.

**Оператор сегрегации** – это языковая конструкция, позволяющая на основе выбора строительных блоков из хромосом родителей (или их частей) создавать хромосомы потомков.

**Оператор удаления** – это языковая конструкция, позволяющая на основе удаления строительных блоков из хромосом родителей (или их частей) создавать хромосомы потомков.

**Оператор вставки** – это языковая конструкция, позволяющая на основе вставки строительных блоков в хромосомы родителей создавать хромосомы потомков. При его реализации направленным или случайным образом создается хромосома (донор), состоящая из строительных блоков, которые желательно разместить в другие хромосомы популяции. После этого направленным или случайным образом определяется хромосома для реализации оператора вставки. Далее производится пробная вставка строительных блоков с вычислением изменения значения ЦФ и получением реальных решений. Отметим, что оператор удаления и оператор вставки могут изменять размер хромосом. Для сохранения постоянным размера хромосом эти операторы можно применять совместно.

**Оператор редукции** – это языковая конструкция, позволяющая на основе анализа популяции после одной или нескольких поколений генетического алгоритма уменьшать ее размер до заданной величины. Он выполняется для устранения неудачных решений. В некоторых генетических алгоритмах этот оператор применяется для сохранения размера начальной популяции постоянным.

**Оператор рекомбинации.** Оператор рекомбинации это языковая конструкция, которая определяет, как новая генерация хромосом будет построена из родителей и потомков. Другими словами, оператор рекомбинации – это технология анализа и преобразования популяции при переходе из одной генерации в другую.

В инженерных задачах используются различные механизмы и модели этого процесса. Приведем несколько из них:

- ♦ M1 – вытеснение (crowding factor). Этот механизм определяет способ и порядок замены родительских хромосом из генерации  $t$  хромосомами по-

томками после генерации  $t+1$ . Механизм реализован таким образом, что стремится удалять «похожие» хромосомы из популяции и оставлять отличающиеся.

- ◆ M2 – разделение (sharing). Этот механизм вводит зависимость значения ЦФ хромосомы от их распределения в популяции и поисковом пространстве. Это позволяет копиям родительских хромосом или близких к ним не появляться в популяциях.
- ◆ M3 – введение идентификаторов (tagging). Специальным хромосомам присваиваются метки. Операторы ГА применяются только к помеченным хромосомам.
- ◆ M4 – сжатие. Специальные хромосомы объединяются в строительные блоки и считаются, как единичные элементы.
- ◆ M5 – реализации «грубой» и «точной» моделей. Грубая модель позволяет анализировать алгоритм на уровне популяций. Точная модель позволяет анализировать алгоритм на уровне альтернативных решений (хромосом).

## 2. Модифицированные генетические операторы

**2.1. Оператор мутации (ОМ).** Он реализуется обычно на одном альтернативном решении. Например, имеем родительскую хромосому  $P_1=1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9$ .

Выделим три части в исходной хромосоме на основе точек разреза.

$$P_1=1\ 2\ 3\ | \ 4\ 5\ 6\ | \ 7\ 8\ 9.$$

Опишем несколько вариантов оператора мутации.

- ◆ **Замена частей генов в исходной хромосоме по цепочке.** Механизм такого ОМ, например, имеет вид:

$$(123) \rightarrow (456) \rightarrow (789) \rightarrow (123).$$

При этом получим хромосому потомка:

$$P_1: 7\ 8\ 9\ | \ 1\ 2\ 3\ | \ 4\ 5\ 6.$$

На следующей генерации применение ОМ, например, введем 6 точек разреза:

$$P_1: 7\ | \ 8\ | \ 9\ | \ 1\ 2\ 3\ | \ 4\ | \ 5\ | \ 6.$$

После использования механизма замены  $8 \rightarrow 5 \rightarrow 8$ , получим преобразованное значение хромосомы потомка:

$$P_2: 7\ 5\ 9\ | \ 1\ 2\ 3\ | \ 4\ 8\ 6.$$

Вероятность применения такого оператора зависит от количества точек разреза. Например, в  $P_1$  при двух точках разреза  $P_1(\text{ОМ}) = 2/L$ , где  $L$  – длина хромосомы ( $L=9$ ).

- ◆ **Перестановка генов, расположенных между точками разрыва**

Пусть задана хромосома А. В хромосоме А заданы 4 точки разреза.

$$A: 1\ | \ 2\ | \ 4\ 3\ 5\ 7\ 6\ | \ 8\ | \ 9.$$

Применим следующий механизм перестановки генов.

Здесь  $1 \rightarrow 2$  и  $4 \rightarrow 1$ ;  $6 \rightarrow 8$  и  $9 \rightarrow 6$ .

Тогда хромосома В примет вид:

$$B: 2\ 4\ 1\ 3\ 5\ 7\ 8\ 9\ 6.$$

- ◆ **Перестановка генов, ансамблей генов и строительных блоков между собой.**

Пусть заданы 2 хромосомы С и D. В них заданы 4 точки разреза.

$$C: 4\ | \ 5\ | \ 6\ 1\ 2\ 3\ 7\ | \ 8\ | \ 9.$$

$$D: 1\ | \ 2\ | \ 3\ 7\ 8\ 9\ 4\ | \ 5\ | \ 6.$$

Для хромосомы С применим механизм перестановки  $5 \leftrightarrow 8$ , а для хромосомы D механизм перестановки  $2 \leftrightarrow 5$ . Тогда получим хромосомы потомки С1 и D1:

$$C_1: 4\ 8\ 6\ 1\ 2\ 3\ 7\ 5\ 9.$$

$$D_1: 1\ 5\ 3\ 7\ 8\ 9\ 4\ 2\ 6.$$

Отметим, что, в принципе, можно построить множество вариантов ОМ. Выбор того или иного ОМ зависит от вероятности применения и других параметров генетического поиска. Вероятность выживания лучших решений зависит от выбора места разреза и знаний о решаемой задаче. В лучшем случае, вероятность выживания равна  $L-1/L$ , в худшем случае, вероятность выживания лучшего решения равна  $1/L$ , а в среднем –  $L/2/L$ .

**2.2. Совместный оператор скрещивание-мутация (СМО).**

Пусть задано 2 родительских хромосомы (альтернативных решений).

P1=1 2 3 | 4 5 6 | 7 8 9.

P2=2 4 6 | 8 9 7 | 5 3 1.

Кроме того, задан ансамбль генов (8,9,7), который представляет строительный блок, закрепляемый в решении. Предлагается ансамбль (8 9 7) из P2 вставить в P1 после первой точки разрыва. Далее в новой хромосоме необходимо устранить повторяющиеся гены. В результате получим хромосому потомка F1:

F1: 1 2 3 8 9 7 4 5 6 7 8 9 → 1 2 3 8 9 7 4 5 6.

Аналогичные преобразования выполним для построения хромосомы потомка F2:

F2: 2 4 6 4 5 6 8 9 7 5 3 1 → 2 4 6 5 8 9 7 3 1.

Отметим, что важным вопросом при разработке генетических операторов является определение числа и места точек разреза. Автор предлагает для нахождения точек разреза использовать треугольник Паскаля. Он основан на определении биномиальных коэффициентов в формуле бинома Ньютона:

$$(x + y)^m = \sum_{n=0}^m C_m^n x^n y^{m-n} .$$

На основе биномиальных коэффициентов строится список чисел. Он определяет количество точек разреза в различных генетических операторах. Размер списка зависит от длины (L) наибольшей хромосомы в популяции.

Для рекуррентного вычисления значений биномиальных коэффициентов используется треугольник Паскаля, представленный в табл. 1.

Таблица 1

<b>1</b>									
1	<b>2</b>	1							
1	3	<b>3</b>	1						
1	4	<b>6</b>	4	1					
1	5	10	<b>10</b>	5	1				
1	6	15	<b>20</b>	15	6	1			
1	7	21	35	<b>35</b>	21	7	1		
1	8	28	56	<b>70</b>	56	28	8	1	
				...					

Из таблицы выбираются числа, отмеченные жирным шрифтом. Они составят следующую последовательность для разрезов в генетических операторах:

TP = 1, 2, 3, 6, 10, 20, 35, 70, ....

Эта последовательность справедлива при длине хромосомы  $L \leq 70$ .

Для построения аналогичной последовательности на основе простых чисел автор предлагает использовать Решето Эратосфена. Идея показана в виде одной строки:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	...
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	-----

Механизм получения последовательности чисел для разрезов следующий. Сначала в строке вычеркиваются все числа, делящиеся на 2 (исключая само число 2). Потом берем первое из оставшихся (число 3). Это число простое. Далее вычеркиваются все идущие после него числа, делящиеся на 3. Первым оставшимся числом будет 5. Затем процесс продолжается аналогично. Числа, остающиеся после всех вычеркиваний, являются простыми. Например, для рассматриваемой строки имеем следующую последовательность простых чисел: 1, 2, 3, 5, 7, 11, 13, 17, 19. Она используется, когда  $L \leq 19$ .

Известно, что 32 числа от 1 до 100 не делятся ни на 2, ни на 3, ни на 5. Кроме них, останутся сами числа 2, 3, 5. Следовательно, всего 35 чисел в последовательности от 1 до 100. Заметим, что последовательность разрезов можно получить, используя зачеркнутые числа.

В нашем примере получим: 4, 6, 8, 9, 10, 12, 14, 15, 16, 18 при  $L \leq 19$ .

Отметим, что предложенный механизм получения точек разреза эффективно используется в простых и комбинированных генетических алгоритмах [9,10].

Леонардо да Винчи деление отрезка в отношении  $\Phi$  назвал «золотым сечением» [7]. Число  $\Phi$  связано с последовательностью Фибоначчи: 1, 1, 2, 3, 5, 8, 13, ..., где каждый член, начиная с третьего, равен сумме двух предыдущих

$$\Phi_n = \frac{1}{2}(a_n \sqrt{5} + b_n),$$

где  $a_n$  – n-е число Фибоначчи, а  $b_n$  – n-й член последовательности: 1, 3, 4, 7, 11, 18, 29, строящейся по тому же принципу, что и последовательность Фибоначчи.

Числа Фибоначчи, числа «золотого сечения», а также члены  $b_n$  – n-й последовательности могут использоваться или используются для определения точек разреза в генетических алгоритмах.

Леонардо да Винчи считал, что идеальные пропорции человеческого тела должны быть связаны с числом  $\Phi$ . В большинстве живописных пейзажей линия горизонта делит полотно в отношении, близком к  $\Phi$ .

Опишем теперь одну из возможных структур модифицированного генетического алгоритма, который использует предложенные идеи для создания генетических операторов.

#### 1. Конструирование популяции:

1.1. На основе создателя.

1.2. Дихотомии.

1.3. Фибоначчи.

1.4. Решета Эратосфена.

1.5. Фокусировки.

1.6. Одеяла.

1.7. Дробовика.

1.8. Конструирования строительных блоков на основе знаний о решаемой

задаче.

#### 2. Селекция:

2.1. Элитная.

2.2. Один с многими.

2.3. Многие с одним.

2.4. Лучшие с худшими.

2.5. На основе метолов ветвей и границ.

#### 3. Операторы:

3.1. Кроссинговера.

3.2. Мутации.

- 3.3. Транспозиции.
  - 3.4. Транслокации.
  - 3.5. Инверсии.
  - 3.6. Интегрированные операторы.
  4. Редукция популяции.
  5. Проверка окончания алгоритма. При необходимости выполнение следующих генерация алгоритмов.
  6. Конец работы алгоритма.
- Приведем упрощенную схему одной ветки генетического алгоритма на основе эффекта «Основателя» (рис. 1).

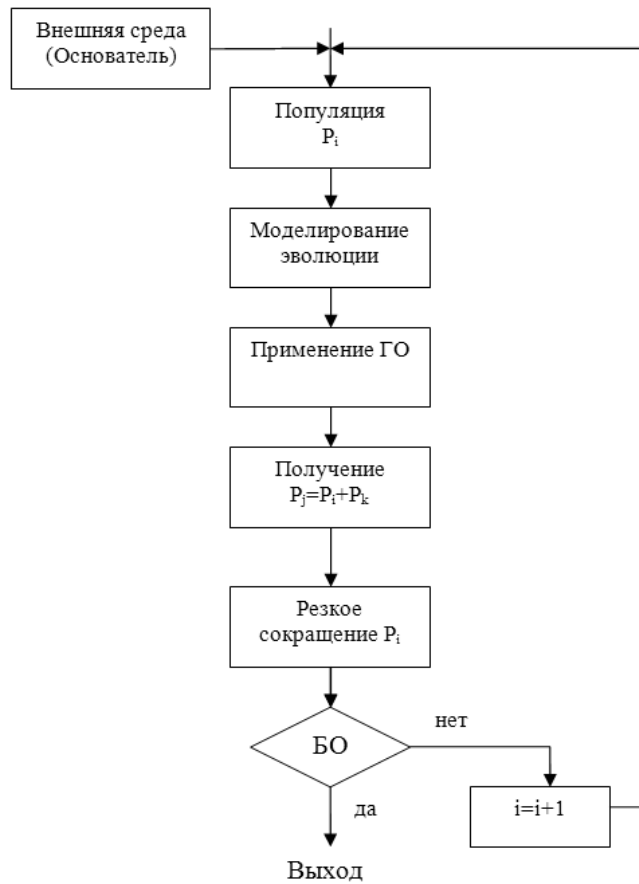


Рис. 1. Схема ГА на основе эффекта «Основателя»

Здесь популяция сначала состоит из одного элемента «Адама, созданного Основателем». Далее, на основе внешней среды из «Адама» создается «Ева». Теперь популяция состоит из двух элементов. На основе этих двух элементов реализуется одно поколение генетического алгоритма (см. рис. 1).

Затем получаем изменяемую на каждой итерации популяцию. Она анализируется и для нее реализуется модифицированный ГА. На рисунке (см. рис. 1) ГО – генетические операторы, БО – блок оценки, определяющей окончание эволюции и алгоритма. В алгоритме используется популяция изменяющегося размера

на каждой генерации. Это позволяет, в некоторых случаях, избегать преждевременной сходимости алгоритма.

**Заключение.** Отметим, что модифицированные генетические операторы выполняются в блоке эволюционной адаптации совместно с внешней средой и экспертной советующей подсистемой. Это позволяет эффективно подбирать такие параметры, как число точек разреза и вероятность применения генетического оператора. Новая структура генетических алгоритмов позволяет сократить время поиска и получить локально-оптимальные решения за полиномиальное время.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Holland John H.* Adaptation and Artificial Systems: An Introductory Analysis with Application to Biology, Control and Artificial Intelligence. USA: University of Michigan, 1975.
2. *Goldberg David E.* Genetic Algorithms in Search, Optimization and Machine Learning. USA: Addison-Wesley Publishing Company, Inc., 1989.
3. Handbook of Genetic Algorithms. Edited by Lawrence Davis. USA: Van Nostrand Reinhold, New York, 1991.
4. Practical Handbook of Genetic Algorithms. Editor I. Chambers. V.3, Washington, USA: CRC Press, 1999.
5. *Редько В.Г.* Эволюция, нейронные сети, интеллект: модели и концепции эволюционной кибернетики. – М.: Комкнига, 2005.
6. *Гладков Л.А., Курейчик В.В., Курейчик В.М.* Генетические алгоритмы / Под ред. В.М. Курейчика. Учебное пособие. – Ростов-на-Дону: Ростиздат, 2004.
7. *Емельянов В.В., Курейчик В.М., Курейчик В.В.* Теория и практика эволюционного моделирования. – М.: Физматлит, 2003.
8. *Гладков Л.А., Курейчик В.В., Курейчик В.М.* Генетические алгоритмы / Под ред. В.М. Курейчика. Учебное пособие. – М.: ФИЗМАТЛИТ, 2006.
9. *Курейчик В.В., Сороколетов П.В.* Концептуальная модель представления решений в генетических алгоритмах // Известия ЮФУ. Технические науки. – 2008. – № 9 (86). – С. 7-12.
10. *Курейчик В.М.* Об одной модели эволюции Шмальгаузена // Известия ЮФУ. Технические науки. – 2009. – № 4 (93). – С. 7-16.

#### **Курейчик Виктор Михайлович**

Технологический институт федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: kur@tsure.ru.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 88634393260.

Заместитель руководителя по научной работе; профессор.

#### **Kureychik Viktor Mihailovich**

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: kur@tsure.ru.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: 88634393260.

Deputy head on scientific work; professor.