УДК 004.8

І.В. Калініна, О.І. Лісовиченко

ВИКОРИСТАННЯ ГЕНЕТИЧНИХ АЛГОРИТМІВ В ЗАДАЧАХ ОПТИМІЗАЦІЇ

Анотація: Розглянуто та проаналізовано найбільш поширені моделі генетичних алгоритмів. Наведено послідовність та приклад вирішення класичної задачі використовуючи генетичні алгоритми.

Ключові слова: Генетичні алгоритми, генітор, кросовер,мутація, схрещування, функція пристосованності.

Вступ

Еволюція в природі показала себе як потужний механізм розвитку і пристосування живих організмів до навколишнього середовища і дивує різноманіттям і ефективністю рішень. В пошуках нових підходів, дослідники в області IT звернулися до запозичених у природи моделей та алгоритмів вирішення різнопланових задач оптимізації.

Генетичні алгоритми (ГА) працюють за аналогією з природою. Вони оперують з сукупністю “особин”, що представляють собою рядки, кожна з яких кодує одне з рішень задачі [1–3]. Пристосованість особини оцінюється за допомогою спеціальної функції. Найбільш пристосовані отримують шанс схрещуватися і давати потомство. Найгірші особини видаляються і не беруть участі в подальшому розвитку (еволюції). Методи з комбінованим використанням ГА, мають більшу ефективність вирішення певного класу задач оптимізації, ніж просто використання класичних методів.

Постановка задачі

В даний час ГА використовуються для вирішення наступних задач:

- пошук глобального екстремуму багатопараметричної функції;
- оптимізація функцій;
- оптимізація динамічних систем;
- задачі оптимізації параметрів;
- настройка штучної нейронної мережі;
- завдання на графах (задача комівояжера);
- навчання інтелектуальних моделей;
- комбінаторні задачі;
- знаходження коренів систем нелінійних рівнянь.

ГА в своїх моделях використовують еволюційні принципи спадковості, мінливості і природного відбору [4].

Розглянемо та проаналізуємо три найбільш поширених моделі:

1. класична (канонічна) модель ГА;
2. генітор (Genitor);
3. гібридна модель ГА.

Класичний генетичний алгоритм

Дана модель алгоритму була запропонована в 1975р. Джоном Холландом [5]. За допомогою пропорційного відбору формується проміжний масив, з якого випадковим чином вибираються двоє батьків.

Далі виробляється одноточковий кросовер, і створені два нащадка мутують (одноточкова мутація) із заданою вірогідністю. Після мутації, нащадки займають місця своїх батьків. Процес продовжується до тих пір, поки не буде досягнутий критерій закінчення алгоритму. Модель канонічного ГА має такі характеристики:

- фіксований розмір популяції;
- фіксована розрядність генів;
- пропорційний відбір;
- особини для схрещування вибираються випадковим чином;
- одноточковий кросовер і одноточкова мутація;
- наступне покоління формується з нащадків поточного покоління без "елітизму", нащадки займають місця своїх батьків;
- бінарне кодування змінних.

Генітор (Genitor)

У даній реалізації генетичного алгоритму використовується специфічна стратегія відбору. Спочатку, популяція ініціюється і її особи оцінюються. Потім вибираються випадковим чином дві особи, схрещуються, внаслідок чого, виходить тільки один нащадок, який оцінюється і займає місце найменш пристосованої особи. Далі, необхідно знову випадковим чином вибрати 2 особи, схрестити і їх нащадок займає місце особи з найнижчою пристосованістю. Таким чином на кожному кроці в популяції оновлюється тільки одна особа. Дана модель ГА має наступні характерні особливості:

- фіксований розмір популяції;
- фіксована розрядність генів;
- особини для схрещування вибираються випадковим чином;
- обмежень на тип кросовера і мутації немає;
- внаслідок схрещування особин виходить один нащадок, який займає місце найменш пристосованої особи;

- дійсне та бінарне кодування змінних.

Необхідно зазначити, що в Genitor-моделі пошук гіперплоскостей відбувається набагато краще, ніж в інших моделях, а також, збіжність при пошуку швидша, ніжу класичного ГА.

Гібридний генетичний алгоритм

Використання гібридного алгоритму дозволяє об'єднати переваги ГА з перевагами класичних методів. Справа в тому, що ГА є робастними алгоритмами, тобто вони дозволяють знаходити оптимальне рішення, але знаходження цього рішення найчастіше виявляється набагато більш ітераційно, важким завданням по причині стохастичності принципів роботи алгоритму. Внаслідок цього, для зменшення ресурсо-затрат, виникла ідея використовувати ГА на початковому етапі для ефективного звуження простору пошуку навколо глобального екстремуму, а потім взявши кращу особу (набір даних), застосувати один з "класичних" методів оптимізації, підвищивши, таким чином, ефективність розрахунку, в цілому. Після аналізу наведеної моделі ГА, можна виділити наступні, характеристики алгоритму:

- фіксований розмір популяції;
- фіксована розрядність генів;
- будь-які комбінації стратегій відбору та формування наступного покоління;
- обмежень на тип кросовера і мутації немає;
- використання ГА на початковій стадії вирішення задач оптимізації, а потім в роботу включаються класичні методи оптимізації;
- дійсне та бінарне кодування змінних.

Для більшої наглядності роботи ГА, запропонуємо приклад вирішення задачі пошуку максимального значення [6].

Приклад

Розглянемо спеціально спрощений і досить штучний приклад, що складається в знаходженні хромосоми (набору даних) з максимальною кількістю одиниць. Припустимо, що хромосоми складаються з 12 генів, а популяція налічує 8 хромосом. Це є класичний етап кодування, перехід до використання ГА. Зрозуміло, що найкращою буде хромосома, що складається з 12 одиниць. Спробуємо простежити та проаналізувати процес вирішення цієї досить тривіальної задачі за допомогою генетичного алгоритму, блок-схема, якої зображена на рис. 1.

Ініціалізація, або вибір вихідної популяції хромосом. Необхідно випадковим чином згенерувати 8 двійкових послідовностей довжиною 12 бітів.

Рис. 1 – Блок-схема генетичного алгоритму

Таблиця 1

Згенеровані хромосоми – 8 двійкових послідовностей довжиною 12 бітів

c_1	000100011001	c_5	101010111001
c_2	010010110011	c_6	110101101000
c_3	011011100100	c_7	010110110110
c_4	110101001101	c_8	101000100100

Це можна досягти, наприклад, підкиданням монети (96 разів, при випаданні "орла" приписується значення 1, а у разі "решки" — 0). Таким чином можна сформувати вихідну популяцію (табл. 1):

Оцінка пристосованості хромосом в популяції. У наведеному прикладі, для спрощення, розглядається задача знаходження хромосоми з найбільшим вмістом одиниць. Стосовно, функція пристосованості повинна визначати кількість одиниць у хромосомі. Значення функції пристосованості для кожної хромосоми з вихідної популяції будуть такі:

Таблиця 2

Значення функції пристосованості

$F(c_1)$	4	$F(c_5)$	7
$F(c_2)$	6	$F(c_6)$	6
$F(c_3)$	6	$F(c_7)$	7
$F(c_4)$	7	$F(c_8)$	4

Хромосоми c_4, c_5 і c_7 характеризуються найбільшими значеннями функції приналежності, яка дорівнює 7. У цій популяції вони вважаються найкращими кандидатами на рішення задачі.

Селекція хромосом. Селекція проводиться методом рулетки. На підставі формули: $V(c_i) = P(c_i) \cdot 100$ %, де $P(c_i) = \frac{F(c_i)}{\sum F(c_i)}$, причому $F(c_i)$ — значення функції пристосованості хромосоми $P(c_i)$ — вірогідність селекції хромосоми c_i.

Для кожної з 8 хромосом поточної популяції (у нашому випадку — вихідної популяції, для якої $n = 8$) отримуємо сектори (табл. 3) колеса рулетки, виражені у відсотках (рис. 1). Скажімо, для першої хромосоми, значення обчислюється так: $P(c_1) = \frac{4}{47} = 0,0851$, тому $V(c_i) = P(c_i) \cdot 100\ \% = 8,51\%$.

Таблиця 3

Значення у відсотках секторів колеса рулетки

$V(c_1)$	8,51	$V(c_5)$	14,89
$V(c_2)$	12,76	$V(c_6)$	12,76
$V(c_3)$	12,76	$V(c_7)$	14,89
$V(c_4)$	14,89	$V(c_8)$	8,51

Розіграш за допомогою колеса рулетки зводиться до випадкового вибору числа з інтервалу [0, 100], що вказує на відповідний сектор на колесі, тобто на конкретну хромосому (рис. 2). Припустимо, що випадково згенеровано наступні 8 чисел: 29, 62, 52, 54, 10, 4, 29, 99.

Дивлячись на рулетку, вибираємо сектори з цими значеннями, це означає вибір наступних хромосом (табл. 4).

Очевидно, що чим більше сектор, тим більше вірогідність "перемоги" відповідної хромосоми. Тому ймовірність вибору даної хромосоми виявляється пропорційною значенню її функції пристосованості. Якщо все коло рулетки представити у вигляді цифрового інтервалу $[0, 100]$, то вибір хромосоми можна ототожнити з вибором

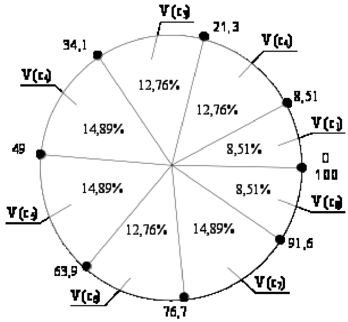

Рис. 2 – Використання колеса рулетки для вибору хромосом

Таблиця 4

Вибір хромосом методом рулетки

число	хромосома	число	хромосома
29	c_3	10	c_2
62	c_5	4	c_1
52	c_5	29	c_3
54	c_5	99	c_8

числа з інтервалу $[a, b]$, де a і b позначають відповідно початок і закінчення фрагмента кола, відповідного цьому сектору колеса; очевидно, що $0 < a < b < 100$. У цьому випадку вибір за допомогою колеса рулетки зводиться до вибору числа з інтервалу $[0, 100]$, яке відповідає конкретному сегменту на колі колеса.

Як видно, хромосома c_5 була обрана тричі, а хромосома c_3 — двічі. Зауважимо, що саме хромосома c_5 має найбільше значення функції пристосованості. Проте, паралельно була й обрана хромосома $c4$ з найменшим значенням функції пристосованості. Всі вибрані таким чином хромосоми включаються в так званий батьківський набір.

Застосування генетичних операторів. Припустимо, що ні одна з відібраних у процесі селекції хромосом не піддається мутації, і всі вони складають популяцію хромосом, призначених для схрещування. Це означає, що ймовірність схрещування $P_c = 1$, а ймовірність мутації $P_m = 0$. Далі, з цих хромосом випадковим чином сформовані пари батьків: $c_3 \leftrightarrow c_5$, $c_5 \leftrightarrow c_2$, $c_5 \leftrightarrow c_3$, $c_1 \leftrightarrow c_8$.

Для першої пари випадковим чином вибрана точка схрещування $l_k = 5$, для другої $l_k = 11$, для третьої $l_k = 3$, для четвертої $l_k = 1$. При цьому процес схрещування протікає так, як показано на рис. 3. В результаті виконання оператора схрещування виходять 4 пари нащадків.

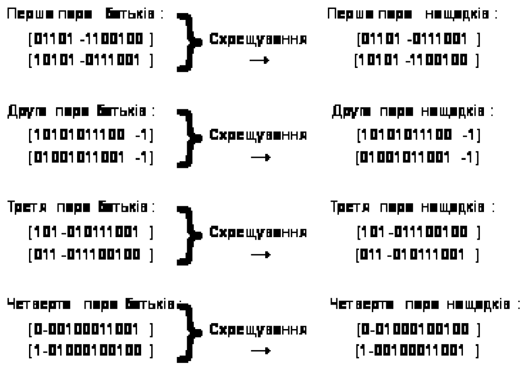

Рис. 3 – Процес схрещування хромосом

Слід зазначити, що в схрещувані однакових пар хромосом немає сенсу (наприклад, $c_3 \leftrightarrow c_3$).

Формування нової популяції. Після виконання операції схрещування ми отримуємо наступну популяцію нащадків:

Таблиця 5

Популяція нащадків

c_1	011010111001	c_5	101011100100
c_2	101011100100	c_6	011010111001
c_3	101010111001	c_7	001000100100
c_4	010010110011	c_8	100100011001

Значення функцій пристосованості хромосом цієї популяції складають:

Таблиця 6

Значення функції пристосованості

$F(c_1)$	7	$F(c_5)$	6
$F(c_2)$	6	$F(c_6)$	7
$F(c_3)$	7	$F(c_7)$	3
$F(c_4)$	6	$F(c_8)$	5

Далі, проводимо оцінку хромосом на критерій досягнення результату - знаходження хромосоми (набору даних) з максимальною кількістю одиниць. Так, як умову не виконано, згідно блок-схемі (рис. 1) проводиться повернення до другого етапу, тобто до оцінки пристосованості хромосом з новосформованої популяції, яка стає поточною (ітерація 2).

 ISSN 1560-8956

Таблиця 7

Значення у відсотках секторів колеса рулетки

$V(c_1)$	14,89	$V(c_5)$	12,77
$V(c_2)$	12,77	$V(c_6)$	14,89
$V(c_3)$	14,89	$V(c_7)$	6,38
$V(c_4)$	12,77	$V(c_8)$	10,64

Випадково згенеровано наступні 8 чисел: 66, 6, 70, 4, 32, 35, 71, 30.

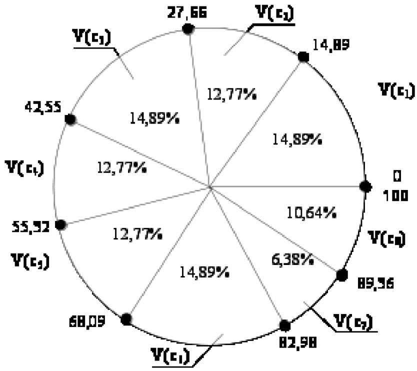

Рис. 4 – Використання колеса рулетки для вибору хромосом

Таблиця 8

Вибір хромосом методом рулетки

число	хромосома	число	хромосома
66	c_5	32	c_3
6	c_1	35	c_3
70	c_6	71	c_6
4	c_1	30	c_3

Випадковим чином сформовані пари батьків: $c_5 \leftrightarrow c_3$, $c_6 \leftrightarrow c_1$, $c_1 \leftrightarrow c_3$, $c_3 \leftrightarrow c_6$. Для першої пари випадковим чином вибрана точка схрещування $l_k = 6$, для другої $l_k = 11$, для третьої $l_k = 1$, для четвертої $l_k = 4$. При цьому процес схрещування протікає так, як показано на рис. 5.

Знову, повторюючи виконання операції схрещування, ми отримуємо наступну популяцію нащадків:

Знову проводимо оцінку хромосом, умову досягнення максимального значення не виконано, повертаємося до другого етапу.

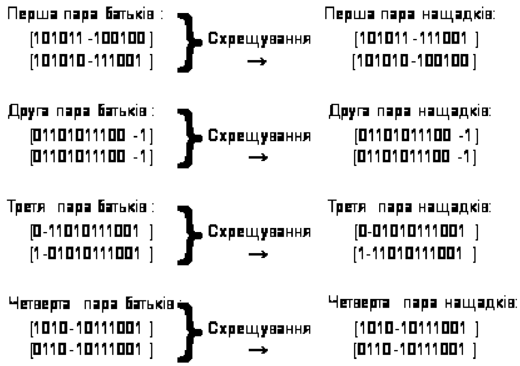

Рис. 5 – Процес схрещування хромосом

Таблиця 9

Популяція нащадків

c_1	101011111001	c_5	001010111001
c_2	101010100100	c_6	111010111001
c_3	011010111001	c_7	101010111001
c_4	011010111001	c_8	011010111001

Таблиця 10

Значення функції пристосованості

$F(c_1)$	8	$F(c_5)$	6
$F(c_2)$	5	$F(c_6)$	8
$F(c_3)$	7	$F(c_7)$	7
$F(c_4)$	7	$F(c_8)$	7

Наведемо кінцеву ітерацію навчання (ітерація 8) з новосформованої популяції.

Таблиця 11

Значення у відсотках секторів колеса рулетки

$V(c_1)$	13,33	$V(c_5)$	15,00
$V(c_2)$	11,67	$V(c_6)$	13,33
$V(c_3)$	11,67	$V(c_7)$	11,67
$V(c_4)$	10,00	$V(c_8)$	13,33

Знову вибираємо методом рулетки хромосоми, які увійдуть в наступний батьківський набір (табл. 12), генеруємо значення для рулетки: 30, 47, 68, 2, 97, 60, 83, 45.

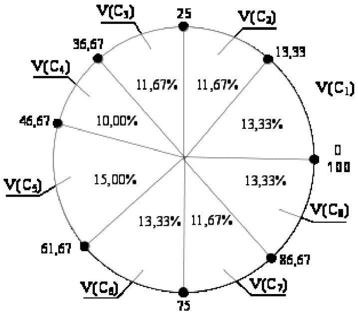

Рис. 6 – Використання колеса рулетки для вибору хромосом

Таблиця 12

Вибір хромосом методом рулетки

число	хромосома	число	хромосома
30	c_3	97	c_2
47	c_5	60	c_1
68	c_5	83	c_3
2	c_5	45	c_8

Для першої пари випадковим чином вибрана точка схрещування $l_k = 8$, для другої $l_k = 6$, для третьої $l_k = 2$, для четвертої $l_k = 5$.

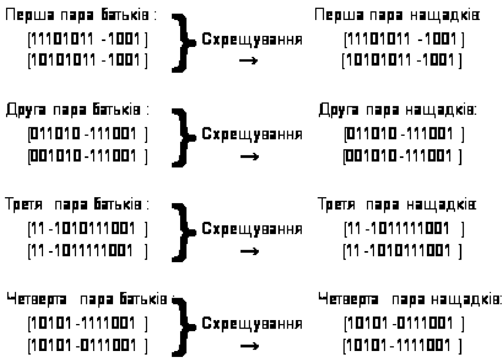

Рис. 7 – Процес схрещування хромосом

Значення функції пристосованості для кожної хромосоми з вихідної популяції будуть такі:

Таблиця 13

Популяція нащадків

c_1	111010111001	c_5	111011111001
c_2	101010111001	c_6	111010111001
c_3	011010111001	c_7	101010111001
c_4	001010111001	c_8	101011111001

Таблиця 14

Значення функції пристосованості

$F(c_1)$	8	$F(c_5)$	9
$F(c_2)$	7	$F(c_6)$	8
$F(c_3)$	7	$F(c_7)$	7
$F(c_4)$	6	$F(c_8)$	8

Помітно, що популяція нащадків характеризується набагато більш високим середнім значенням функції пристосованості, ніж популяція батьків. Проте так і не було досягнуто максимальної кількості одиниць, тому використаємо оператор мутації - інвертування значень відібраних генів з 0 на 1 та навпаки. (ітерація 9). Повторюємо цикл, починаючи з другого етапу (рис. 1).

Таблиця 15

Значення у відсотках секторів колеса рулетки

$V(c_1)$	12,9	$V(c_5)$	12,9
$V(c_2)$	12,9	$V(c_6)$	9,68
$V(c_3)$	12,9	$V(c_7)$	12,9
$V(c_4)$	11,29	$V(c_8)$	14,51

Знову випадково згенеровано наступні 8 чисел: 27, 97, 10, 26, 41, 91, 11, 23.

Таблиця 16

Вибір хромосом методом рулетки

число	хромосома	число	хромосома
27	c_3	41	c_4
97	c_8	91	c_8
10	c_1	11	c_1
26	c_3	23	c_2

Випадковим чином сформовані пари батьків: $c_3 \leftrightarrow c_8$, $c_1 \leftrightarrow c_3$, $c_4 \leftrightarrow c_8$, $c_1 \leftrightarrow c_2$.

Для першої пари випадковим чином вибрана точка мутації $l_k = 10$, для другої $l_k = 7$, для третьої $l_k = 10$, для четвертої $l_k = 9$.

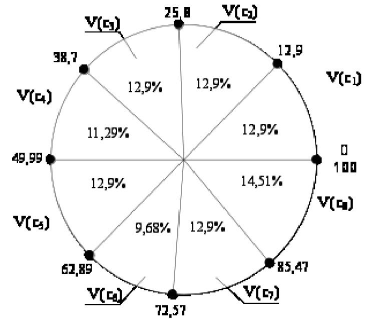

Рис. 8 – Використання колеса рулетки для вибору хромосом

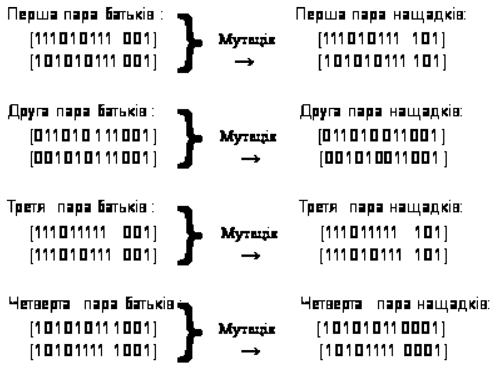

Рис. 9 – Процес мутації хромосом

Після виконання операції мутації, ми отримуємо наступну популяцію нащадків:

Таблиця 17

Популяція нащадків

c_1	111010111101	c_5	111011111101
c_2	101010111101	c_6	111010111101
c_3	011010011001	c_7	101010110001
c_4	001010011001	c_8	101011110001

Обчислюємо значення функції (табл. 18).

Отже, після використання оператора мутації було досягнуто результат - знаходження гена з найбільшою кількістю одиниць, який задовольняє поставленій задачі.

Таблиця 18

Значення функції пристосованості

$F(c_1)$	9	$F(c_5)$	10
$F(c_2)$	8	$F(c_6)$	9
$F(c_3)$	6	$F(c_7)$	6
$F(c_4)$	5	$F(c_8)$	7

Насамперед, по-перше, приклад показує, що результат можна отримати, як максимальне значення, так паралельно і мінімальне. По-друге, метод ГА слід використовувати при вирішення вузького класу задач, де немає оптимальних класичних методів. Потрете, ГА оптимально дає результати, при комбінованих методах ГА з іншими класичними методами, використовуючи ГА як проміжну ланку.

В таблиці наведено систематизовані характеристики розглянутих алгоритмів та підходів до вирішення певного класу задач.

Таблиця 19

Систематизовані характеристики розглянутих алгоритмів

Назва ГА	Тип відбору	Тип кросовера	Кодування змінних
Класичний ГА	$P(i) = \frac{F(i)}{\sum_{i=1}^{n} F(i)}$,	Вибирається одна з $l-1$ точок розриву	Бінарне
Генітор ГА	Турнірний,$k = 2$ використовує n турнірів, щоб вибрати n особин.	Може вибиратися як одна, так і дві точки розриву	Дійсні та бінарні
Гібридний ГА	Елітний	Може вибиратися декілька точок розриву	Дійсні та бінарні

Висновок

На підставі проведеного аналізу можна зробити висновки про те, що найкращим методом генетичного алгоритму є гібридний алгоритм, тому як він складається з двох алгоритмів дозволяє використовувати переваги обох, внаслідок чого, здатний вирішувати більш широкий спектр завдань. Але у деяких випадках класичний (чистий) генетичний алгоритм виявляється більш ефективним а це слід пам'ятати. Пов'язано це з рельєфом оптимізації функцій, коли класичні методи оптимізації потребують велику кількість проміжних і не результативних ітерацій.

Для застосування певного генетичного алгоритму необхідно лише в тих випадках, коли для вирішення заданої задачі немає відповідного спеціального алгоритму рішення, і що для кожної розв'язуваної задачі застосовувати лише ту модель ГА, яка дає найбільшу ефективність вирішення поставленої задачі.

Перелік використаних джерел

1. Панченко Т.В. Генетическиеалгоритмы: учебно-методическое пособие / под ред. Ю. Ю. Тарасевича. — Астрахань : Издательский дом"Астраханский университет", 2007. — 87 [3] с.
2. Darrel Whitley "A Genetic Algorithm Tutorial", 1993.
3. Soraya Rana "Examining the Role of Local Optima and Schema Processing in Genetic Search", 1999.
4. Дмитриев С. В., Тененев В. А. Оптимизация многоэкстремальных функций с помощью гибридных генетических алгоритмов: Известия Института математики и информатики. Ижевск, 2006 №2(36)
5. Holland J. H. Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, 1975.
6. http://www.znannya.org/?view=ga_general - Портал знань: Генетичні алгоритми. Ключові поняття і методи реалізації.

Отримано 16.04.2015 р.