

Модифікований генетичний алгоритм і його застосування для оптимізації розподілення завантаження робітників

Теличко Г.О.

Донецький національний технічний університет

E-mail: selevkoanna@mail.ru,

Abstract

Telichko G.O. Modified genetic algorithm and its using for optimization of workers load allocation. The article introduces into the issue of modern optimization methods used in the tasks of optimal decision search in case of compound dynamic objects managing, in particular, the work load allocation of employers. The brief description of the author version of the modified genetic algorithm for the optimization tasks (the MGA) is given. The MGA includes itself the block of prevention of the preliminary algorithm convergence in the local optimum. The results of testing of the MGA and its comparing with the performance of the Matlab tool of the genetic algorithm designing are also provided.

Вступ

Прикладна проблема ефективного розподілу персоналу між роботами підприємства, по проектах і проектних ролях (ресурсне планування) регулярно виникає перед виробничими керівниками і керівниками проектів. Задача пошуку оптимального розподілення робітників з урахуванням великої чисельності співробітників, множинності і різноманіття робіт і проектів є трудомісткою і такою, що трудно розв'язати. Це обумовлено багатофакторністю і нелінійністю у критеріях оптимізації. Постановка задачі пошуку оптимального розподілення завантаження робітників відрізняється від стандартної задачі про призначення робітників на роботи [1]. Головні відмінності наступні:

- 1) кількість робіт одного робітника не має обов'язково дорівнювати 1;
- 2) цільова функція задачі оптимізації є багатофакторною і нелінійною.

Тому для розв'язання поставленої задачі не можливо застосувати стандартні методи розв'язання задачі про призначення – угорський і ξ -метод.

Не можливо застосовувати і методи перебору, тому що варіації реального розподілення завантаження персоналу безперервні. Недоліком методу повного перебору є велика обчислювальна вартість. Потужність розглянутої множини можливих рішень можна оцінити як $(kol_p)^{kol_w}$, де kol_w , kol_p - відповідно число робіт і робітників. Якщо число робіт і

співробітників наближається до 100, то рішення задачі звичайними методами практично неможливо. Однак, якщо перебір усіх варіантів за розумний час можливий, то можна бути абсолютно впевненим у тім, що знайдене рішення дійсно оптимально. Найбільш широко використовуваний прийом скорочення перебору є прийом цілеспрямованого перебору, заснований на методі гілок і меж. Цей прийом полягає в побудові приватних рішень, наведених у вигляді дерева пошуку, і застосуванні методів побудови оцінок, що дозволяє пізнавати безперспективні часткові розв'язання, внаслідок чого від дерева пошуку відсікається ціла гілка. Тут для алгоритму подібного скороченого перегляду варіантів характерна організація пошуку з поверненням. Виконуючи ходу по дереву з певного вузла в деякому напрямі, після «дослідження» даного напрямку, здійснюється повернення в початкову точку і рух в раніше «не досліджених» напрямках. Використання методу гілок і меж дозволяє в тому або іншому ступені прискорити пошук точного розв'язання при збереженні в загальному випадку експоненціального закону залежності часу виконання алгоритму (трудомісткості) від початкових даних.

Для розв'язання задач з одним екстремумом існує набір градієнтних і числових методів. Застосування цих методів для задач з двома і більше параметрами формує неоднозначні рішення. У кожному випадку результати залежать від обраних початкових координат. З цього випливає висновок про багатоекстремальність подібних задач, і для їхнього рішення потрібні методи глобальної оптимізації [2,3].

В даний час найбільш переважними методами багатоекстремальної оптимізації є генетичні алгоритми (ГА), які реалізують постулати теорії еволюції і досвіду селекції рослин і тварин [2,4-6]. Стратегія пошуку оптимального розв'язання в генетичних алгоритмах спирається на гіпотезу селекції: чим вище пристосованість особня, тим вище імовірність того, що у нащадків, отриманих з її участю, ознаки, що визначають пристосованість, будуть виявлені ще сильніше [6].

Якщо прийняти, що кожний особень популяції є точкою в координатному просторі задачі оптимізації $X_i[x_{1i}, x_{2i}, \dots, x_{ji}]$, а пристосованість особня відповідним значенням функції цілі $f(x_i)$, тоді популяцію особнів можна розглядати як безліч координатних точок в просторі, а процес еволюції - як рух цих точок у бік оптимальних значень цільової функції (ЦФ).

Особливістю ГА є необхідність для кожної задачі визначати хромосому (генотип) і будувати свою ЦФ [5]. Від адекватності побудованої ЦФ залежить ефективність результату роботи ГА. При проектуванні параметрів і роботи ГА для конкретної задачі виникає також необхідність вибору і модифікації ГО. Ці аспекти необхідно враховувати при використанні ГА для задач пошуку оптимального розподілення завантаження робітників підприємства.

Слід зазначити, що класичний генетичний алгоритм знаходить глобальний екстремум у ймовірному розумінні. Тобто, ГА характерна попередня збіжність рішення у локальному оптимумі [6].

Перераховані особливості стримують широке застосування ГА в інженерній практиці. Проте потреба в таких алгоритмах для вирішення прикладних задач порівняно невеликої розмірності і складності постійно росте.

Метою статті є, по-перше, запропонувати модифікацію існуючого ГА, яка здатна з необхідною імовірністю досягти глобального екстремуму. Далі описати застосування МГА для пошуку оптимального завантаження робітників. В статті приводяться результати тестування блоку запобігання попередньої збіжності (БЗПЗ) і використання авторської версії модифікації генетичного алгоритму (МГА). Автор вважає, що такий алгоритм знайде застосування в задачах оптимального розподілення ресурсів, а також в інших дослідницьких задачах.

1. Загальний опис МГА

При реалізації МГА важливою є процедура збереження максимальних значень ЦФ кожної генерації. Робота БЗПЗ базується на аналізі найкращого значення ЦФ поточної генерації і найкращих ЦФ попередніх поколінь та прийнятті відповідних дій, якщо виконуються умови, що наведено у таблиці 2.2. При виконанні перерахованих вище умов відбувається зміна поточної популяції. Хромосоми поточної популяції замінюються відповідно алгоритму хромосомами з АОР, хромосомами сформованими випадково чи хромосомами сформованими на базі хромосом АОР з мутованими генами. Дії, що приймаються алгоритмом при спрацьовуванні умов БЗПЗ, описано в таблиці 2.2. В умові 3 пропонується $eq_hr = 20\%$ (достатня кількість для подальшого розвитку за принципом Парето), а дії (1-3) виконуються послідовно, поки не будуть замінені всі обрані хромосоми популяції.

Для підвищення імовірності пошуку глобального екстремуму в МГА реалізовано механізм БЗПЗ, який виконує функцію аналізу оптимального рішення поточної генерації і прийняття заходів для виходу з локальних „ям” за допомогою сформованого в процесі виконання алгоритму архіву оптимальних (квазіоптимальних) рішень (АОР). У МГА пропонується застосувати прийом повернення в попередню точку, який вдало використовується у методі гілок і меж. Цей прийом відкриває додаткові напрямки розвитку популяції для допомоги ГА вийти з локального оптимуму.

При реалізації МГА важливою є процедура збереження максимальних значень ЦФ кожної генерації. Робота БЗПЗ базується на аналізі найкращого значення ЦФ поточної генерації і найкращих ЦФ попередніх поколінь та прийнятті відповідних дій, якщо виконуються

умови, що наведено у таблиці 2.2. При виконанні перерахованих вище умов відбувається зміна поточної популяції. Хромосоми поточної популяції замінюються відповідно алгоритму хромосомами з АОР, хромосомами сформованими випадково чи хромосомами сформованими на базі хромосом АОР з мутованими генами. Дії, що приймаються алгоритмом при спрацьовуванні умов БЗПЗ, описано в таблиці 1. В умові 3 пропонується $eq_hr = 20\%$ (достатня кількість для подальшого розвитку за принципом Парето), а дії (1-3) виконуються послідовно, поки не будуть замінені всі обрані хромосоми популяції.

Таблиця 1.

Робота БЗПЗ

| № | Умова | Опис дій |
|---|---|---|
| 1 | максимальне значення ЦФ поточної генерації не змінилося щодо попередніх генерацій | При кількості генерацій більше 2, 20% популяції замінюється випадковими хромосомами |
| 2 | максимальне значення ЦФ зменшується протягом декількох генерацій | При кількості генерацій рівним 2 замінюється 20% популяції з АОР (не більш 3 ітерацій). При кількості генерацій > 3 , 20% популяції замінюється випадковими хромосомами |
| 3 | кількість однакових хромосом з максимальної ЦФ поточної генерації перевищує 20% (достатня кількість для подальшого розвитку за принципом Парето). | Заміна 80% однакових хромосом. Способи заміни: 1) хромосоми з АОР (перезапис не більш 3 ітерацій); 2) випадково сформована хромосома; 3) мутирована хромосома з АОР. |

2. Опис методики та результатів тестування МГА

Відомою особливістю ГА є імовірний характер визначення глобального екстремуму. Ступінь досконалості алгоритмів залежить як від запрограмованого механізму генетичної обробки інформації (закладені в програмі генетичні операції і стратегія роботи алгоритму), так і від призначених для користувача установок початкових умов, обсягу використовуваної при пошуку інформації (розмір популяції pop , кількість генерацій $generation$, область пошуку $rang$).

Для запропонованої модифікації алгоритму виконано тестування впливу заданих значень pop , $generation$, $rang$ на імовірність p визначення глобального екстремуму. Для порівняння паралельно з МГА проведено тестування інструментарію ГА в Matlab [7].

Тестування проведено для трьох двомірних функцій цілі. Перша має назву функція Растрігіна [8,9] другу і третю взято з джерел [10] і [11] відповідно. Функції містять множину локальних мінімумів. Проте мають тільки один глобальний мінімум. Аналітичний вигляд функцій наведено у формулах (7-9).

$$Ras(x) = 20 + x_1^2 + x_2^2 - 10 \cdot (\cos 2\pi x_1 + \cos 2\pi x_2) \quad (1)$$

$$f(x_1, x_2) = 3 \cdot (1 - x_1)^2 \cdot e^{[-x_1^2 - (x_2 + 1)^2]} - 10 \cdot (0,2 \cdot x_1 - x_1^3 - x_2^5) \cdot e^{(-x_1^2 - x_2^2)} - e^{[-(x_1 + 1)^2 - x_2^2]} / 3 \quad (2)$$

$$f(x_1, x_2) = [\sin(\pi \cdot x_1) + \sin(\pi \cdot x_2)] \cdot 0,2 + 0,01 \cdot [0,4 \cdot (x_1 - 5,5)^2 + 0,5 \cdot (x_2 - 5,5)^2] + 0,4 \quad (3)$$

Функція Растрігіна (рис. 1) часто використовується для тестування генетичного алгоритму, оскільки наявність великої кількості локальних мінімумів створює труднощі для використання стандартних методів пошуку глобального мінімуму на основі використання градієнтних методів.

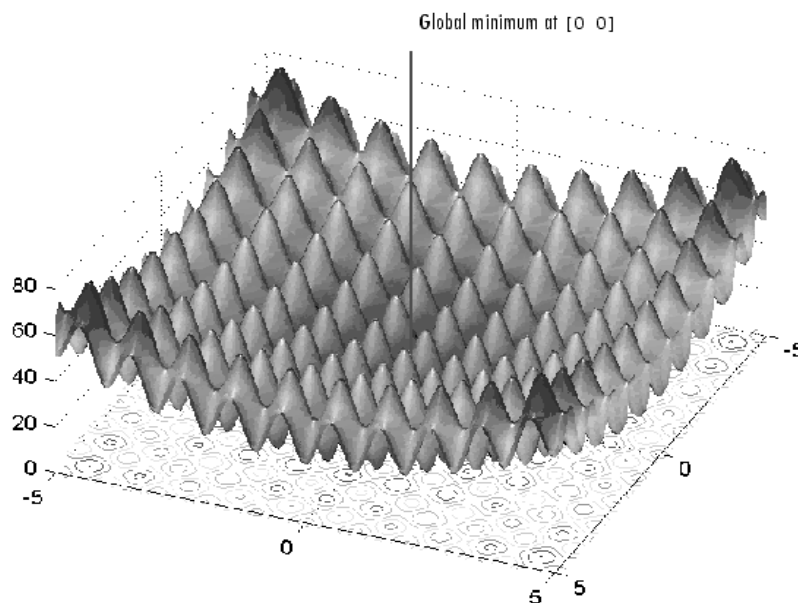


Рисунок 1 – Функція Растрігіна

Для визначення оцінки імовірності p для заданих значень pop , $generation$, $rang$ було виконано по 100 розв'язань (такий розмір вибірки відповідає 95%-му довірчому рівню [12]) МГА і ГА. Потім з кожних 100 результатів відбиралися правильні розв'язання і обчислювався відсоток їхньої появи. Знайдена таким чином величина приймалася за оцінку імовірності появи правильного розв'язання p .

Така процедура проводилась для двох діапазонів пошуку: $[-1, 1]$, $[-5, 5]$. Проводилось моделювання розміру популяції (20, 100, 200) і кількості генерацій (20, 100, 500). Отримані дані оброблялися за допомогою електронних таблиць Excel.

В таблиці 2 зведено результати тестування функцій (1) – (3) двома алгоритмами. З таблиці 2 видно, що імовірність пошуку оптимального екстремуму МГА вище порівняно з інструментарієм ГА в Matlab.

Таблиця 2.

Результати тестування

| Імовірності пошуку для функцій (1), (2), (3) | | | | | | Діапазон пошуку | | Параметри пошуку | |
|--|------|------|------|------|------|-----------------|---|------------------|-----|
| МГА | | | ГА | | | Rang | | generation | pop |
| (1) | (2) | (3) | (1) | (2) | (3) | | | | |
| 0,1 | 0,11 | 0,2 | 0 | 0 | 0 | -1 | 1 | 20 | 20 |
| 0,4 | 0,45 | 0,5 | 0 | 0 | 0 | -1 | 1 | 100 | 20 |
| 0,82 | 0,8 | 0,9 | 0 | 0 | 0 | -1 | 1 | 100 | 100 |
| 0,98 | 0,98 | 1 | 0,02 | 0,03 | 0,1 | -1 | 1 | 200 | 100 |
| 0,1 | 0,07 | 0,15 | 0 | 0 | 0 | -5 | 5 | 100 | 100 |
| 0,3 | 0,4 | 0,4 | 0 | 0 | 0 | -5 | 5 | 100 | 200 |
| 0,75 | 0,6 | 0,85 | 0 | 0 | 0 | -5 | 5 | 200 | 200 |
| 0,95 | 0,95 | 0,98 | 0,01 | 0,02 | 0,05 | -5 | 5 | 500 | 200 |

На рисунку 2 наведено графік залежностей імовірності визначення глобального екстремуму від числа особин в популяції для функцій (1)-(3).

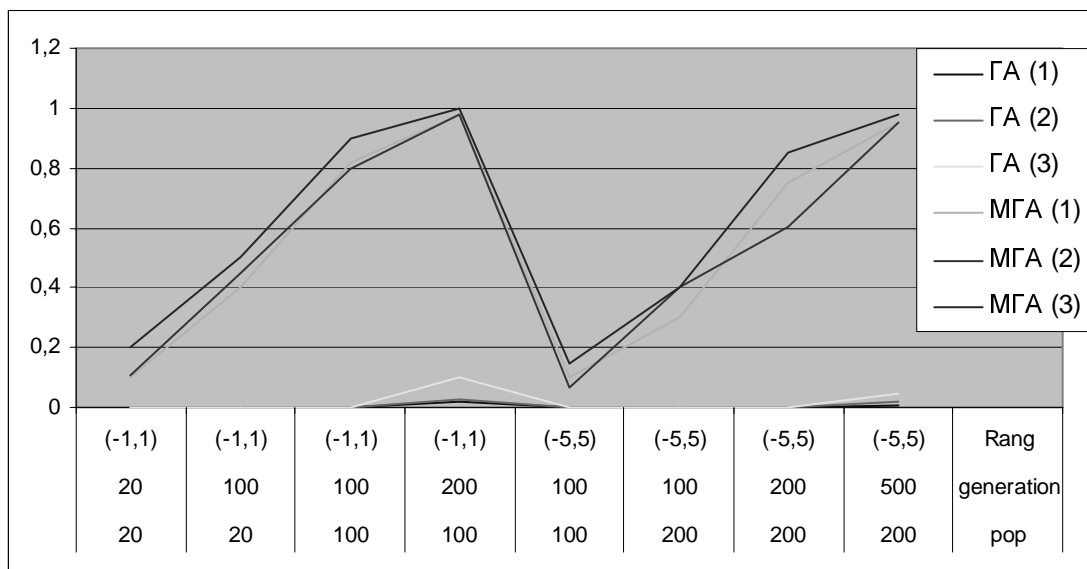


Рисунок 2 – Залежність пошуку глобального екстремуму від розміру популяції і кількості генерацій

3. Застосування МГА для пошуку оптимального розподілу завантаження робітників

При проектуванні роботи ГА головною є задача визначення генотипу, побудови ЦФ, визначення ГО. Відповідно математичної моделі задачі [13], в якій визначено елементи фенотипу задачі необхідно сформулювати генотип задачі пошуку і визначити формулу розрахунку пристосованості хромосом.

3.1. Генотип задачі

Формування генотипу умов задачі є одним з складних блоків для проектувальника, тому що потребує від нього об'єктивності і чіткості при формуванні умов задачі. Від достовірності і однозначності сформованих структур залежить ефективність пошуку. Відповідно математичної моделі [13] множина змінних x_{ij} формує структуру хромосоми. Отже, хромосома є бінарною структурою розміром $kol_i \times kol_j$. Змінна x_{ijh} зберігає значення ij -ого гена хромосоми h . Значення гена може дорівнювати або 0, або 1. Відповідно обмеженням задачі значення гена x_{ijh} може дорівнювати 1 тільки один раз для кожної i для $\forall j$. Це потрібно врахувати при програмуванні алгоритму щоб будувати допустимі рішення. На рис. 3 наведено загальний вигляд хромосоми популяції за номером h . Кількість хромосом популяції pop є змінною величиною, визначається проектувальником від потрібної імовірності пошуку глобального оптимуму. При формуванні початкової популяції значення генів хромосоми визначається випадковим чином.

| | 1 | 2 | 3 | ... | j | ... | kol_j |
|-------|-----|-----|-----|-----|-----|-----|-------|
| 1 | 0 | 1 | 0 | ... | 0 | ... | 0 |
| 2 | 0 | 0 | 0 | ... | 0 | ... | 1 |
| 3 | 0 | 0 | 1 | ... | 0 | ... | 0 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| i | 1 | 0 | 0 | ... | 0 | ... | 0 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| kol_i | 1 | 0 | 0 | ... | 0 | ... | 0 |

ген x_{ijh}

Рисунок 3 - Загальний вигляд хромосоми h

3.2. Розрахунок цільової функції

Відповідно математичної моделі задачі оптимізації [13] необхідно максимізувати ефективність виконання робіт і мінімізувати витрати на оплату праці. При цьому необхідно врахувати обмеження на унікальність роботи (кожну роботу виконує тільки один робітник) і тривалість робочого дня робітника (сумарний час роботи на добу не повинен перевищувати 8 годин). Пристосованість або ЦФ хромосоми h необхідно сформувати з врахуванням описаних вище функцій і обмежень. Значення пристосованості хромосоми h визначає корисність хромосоми для подальшого розмноження популяції у напрямку оптимального рішення.

В економічній теорії функція корисності є моделлю для визначення переваги економічних суб'єктів. Основоположною умовою концепту функції корисності є раціональна поведінка споживача, яка розкривається у виборі з великої кількості альтернатив саме ті, які дають йому найбільший рівень корисності. В мікроекономіці концепт функції корисності служить для пояснення поведінки споживачем та виробників, в

той час як у макроекономіці їм користуються для зображення переваги державних інтересів. У генетичних алгоритмах відповідно функції корисності вибираються хромосоми (рішення) для подальшого розвитку. Тому, значення ЦФ хромосоми h пропонується розраховувати за формулою (4), яка побудована відповідно неокласичній функції корисності з урахуванням штрафної функції.

$$F_h = a \cdot f l_h^\alpha \cdot (1/f 2_h)^\beta - \gamma \cdot g(X_h) \rightarrow \max, \quad h = \overline{1, pop} \quad (4)$$

Тут припускається, що $\alpha + \beta = 1$. Коефіцієнти функції корисності a , штрафної функції γ і показники ступеня α , β визначається відповідно потреб задачі і є базою для моделювання. За формулою (4) розраховується показник сумарної ефективності хромосоми h . Показник сумарної ефективності залежить від ефективності запропонованого рішення $f l$ (формула (5)) і показника витрат рішення $f 2$ (формула (6)) з урахування величини відхилення від обмежень задачі $g(X_h)$ (формули (7-9)).

$$f l_h = \sum_{i=1}^{kol_i} \sum_{j=1}^{kol_j} x_{ijh} \cdot \frac{E_i}{c_{ij} + 1}, \quad h = \overline{1, pop} \quad (5)$$

$$f 2_h = \sum_{i=1}^{kol_i} \sum_{j=1}^{kol_j} x_{ijh} \cdot z_i \cdot k_j, \quad h = \overline{1, pop} \quad (6)$$

При формуванні штрафної функції використовується коефіцієнт sh , який визначає розмір штрафу. Значення коефіцієнту sh визначається відповідно параметрам ефективності і витрат задачі.

$$g(X_h) = \sum_{j=1}^{kol_j} sum_n_j \quad (7)$$

$$sum_n_j = \begin{cases} 0, & sum_kol_j = 0, sum_kol_j = Norm_kol_j; \\ sh / (sum_kol_j + 1), & 0 < sum_kol_j < Norm_kol_j; \end{cases} \quad (8)$$

$$\begin{cases} sh \cdot sum_kol_j, & sum_kol_j > Norm_kol_j. \\ sum_kol_j = \sum_{i=1}^{kol_i} x_{ijh} \cdot norm_i, & j = \overline{1, kol_j} \end{cases} \quad (9)$$

3.3. Опис генетичних операторів

При формуванні нової популяції вибирається один із ГО (селекція, кросинговер, мутація) в залежності від заданої імовірності. Пропонуються наступні значення імовірностей: $P_{селекція} = 0,399$; $P_{кросинговер} = 0,6$; $P_{мутація} = 0,001$ [6]. Сума імовірностей генетичних операторів повинна дорівнювати 1. Пропонується застосовувати такі оператори як селекція, кросинговер,

мутація. Але можливе використання і інверсії, транслокації, транспозиції та інших, а також і можливі модифікації ГО. Вибір додаткових ГО залежить від умов задачі, напрямків пошуку і потреб проектування. На рисунках 4-5 наведено виконання ГО кросинговер і мутація відповідно заданої структури хромосоми.

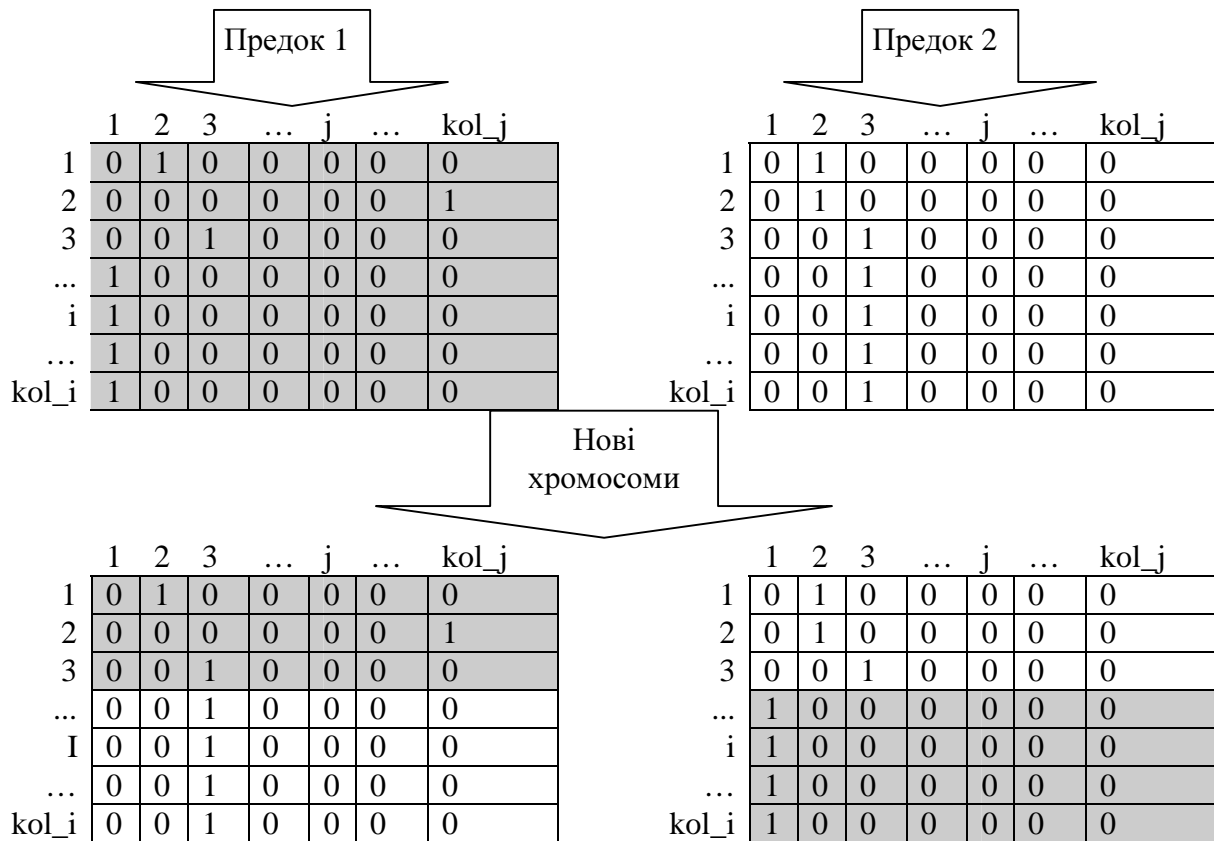


Рисунок 4 - Реалізація кросинговеру

Кросинговер виконується за наступним алгоритмом:

- 1) вибирається рядок i , за яким буде відбуватися розрив хромосом предків (наприклад, рядок 3);
- 2) будуються перша нова хромосома: отримує гени предку_1 з 1 по i і гени предку_2 з $i+1$ по kol_i ;
- 3) будуються друга нова хромосома: отримує гени предку_2 з 1 по i і гени предку_1 з $i+1$ по kol_i .

Мутація виконується за наступним алгоритмом:

- 1) вибирається рядок i , який буде мутовано (наприклад, рядок 3);
- 2) обнуління рядку i : всім значенням j рядку i ставиться у відповідність 0;
- 3) вбирається колонка j рядка i , за якою значення гена x_{ij} змінюється з 0 на 1;

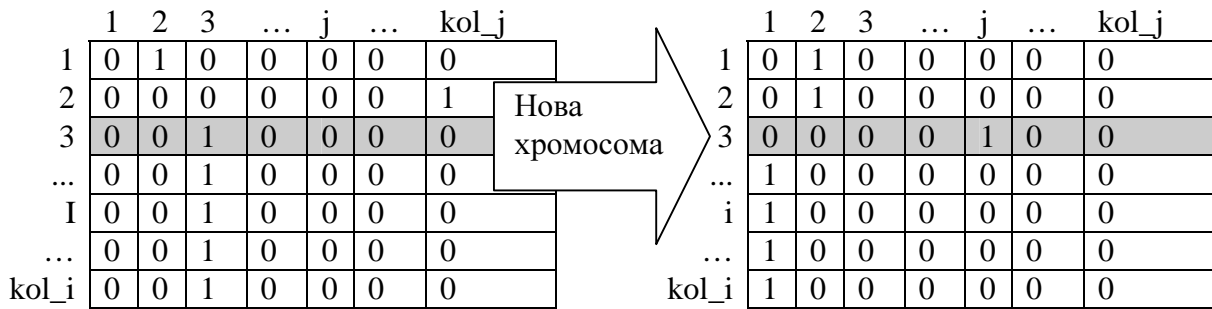


Рис. 2.5. Реалізація мутації

Результатом роботи алгоритму є хромосома, яка визначає значення змінної x_{ij} . Отже, результатом буде перелік робіт, які буде виконувати кожен з робітників.

Проведені дослідження показали, що запропонований МГА дозволяє успішно побудувати задачу пошуку оптимального розподілу завантаження робітників підприємства. При цьому запропоновані елементи можуть ефективно використовуватися і для інших задач управління персоналом (добору, оцінки, тестування, аналізу).

Висновки

В статті наведено опис модифікованого ГА. Модифікація існуючого ГА зроблена з метою створення алгоритму, який за допомогою БЗПЗ здатний з необхідною імовірністю досягти глобального екстремуму.

Алгоритм випробуваний на тестових задачах. При тестуванні зроблено порівняння з результатами роботи інструментарію ГА Matlab. Алгоритм реалізовано у автоматизованій системі підтримки прийняття рішень добору і розподілу персоналу підприємства, а також робота БЗПЗ у вигляді m-файлу програми Matlab, який використовувався для тестування роботи алгоритму.

Як приклад, що ілюструє застосування МГА, розглянуто задачу пошуку оптимального розподілення завантаження робітників між роботами підприємства. Запропонований алгоритм може знайти застосування в задачах оптимального розподілу ресурсів, а також в інших дослідницьких і прикладних багатоекстремальних задачах.

Подальші дослідження направлені на застосування МГА для пошуку оптимального рішення різних класів задач, проведення аналізу і вибору оптимальних параметрів пошуку.

Література

1. Малыхин В.И. Математическое моделирование экономики. — М.: Изд-во УРАО, 1998. — 160 с.
2. Курейчик В.В. Эволюционные методы решения оптимизационных задач: Монография. Таганрог: Изд-во ТРТУ, 1999. — 95 с.
3. Селевко Г.О., Заболотникова В.С. Вибір методу оптимального розподілу кадрів на малому підприємстві // Наукові праці Донецького національного технічного університету. Випуск 70. Серія: „Інформатика, кібернетика та обчислювальна техніка”: - Донецьк: ДонГТУ. — 2003. — С. 47-53.
4. Goldberg D.E. Genetic Algorithms in Search Optimizations and Machine Learning.-Addison.Wesly, 1989.
5. Вороновский Г.К. и др. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности / Г.К. Вороновский, К.В. Махотило, С.Н. Петрашев, С.А. Сергеев. — Х.: ОСНОВА, 1997. — 112 с.
6. Курейчик В.М. Генетические алгоритмы и их применение. Таганрог: Изд-во ТРТУ, издание второе, дополненное.- 2002.- 242 с
7. Matlab & Toolboxes. Genetic Algorithm and Direct Search Toolbox — Математика <http://matlab.exponenta.ru/genalg/05.php>
8. Функція Растрігіна <http://www.cs.rtu.lv/dssg/ru/staff/rastrigin/rastr-function.html>
9. Matlab & Toolboxes. Функція Растрігіна http://matlab.exponenta.ru/genalg/06_02_03_01.php
10. Attia A.A., Horacek P. Adaptation of genetic algorithms for optimization problem solving // 7th International Conference on Computing MENDEL 2001. Brno, 2001. P. 36-41.
11. Сабанин В.Р., Смирнов Н.И., Репин А.И. Модифицированный генетический алгоритм для задач оптимизации и управления. // Exponenta Pro. Математика в приложениях. 2004. №3-4. С.78-85
12. Левин Д., Стефан Д., Кребиль Т., Беренсон М. Статистика для менеджеров с использованием Microsoft Excel. — М.: Издательский дом «Вильямс», 2004. — 1312 с.: ил.
13. Селевко А. А. Специфика и методы решения задачи управления персоналом малого предприятия // «Практика и перспективы развития партнерства в сфере высшей школы»: Материалы шестого научно-практического семинара.- Донецк.: ДонНТУ, 2005 — С . 383-394.