

МЕТОДИ ОРГАНІЗАЦІЇ ПРОГРАМНИХ ЗАСОБІВ, ЩО РЕАЛІЗУЮТЬ ГЕНЕТИЧНІ АЛГОРИТМИ

Розглянемо загальну організацію програмних засобів управління системою зв'язку (рис.1):

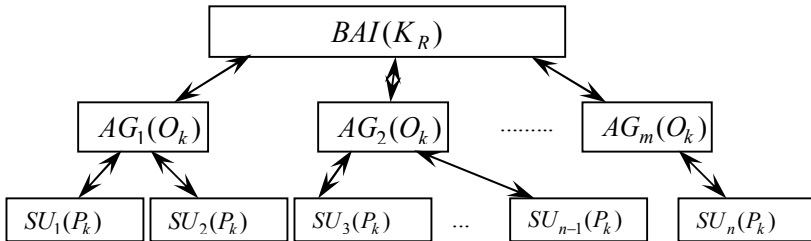


Рис.1. Загальна схема організації програмних засобів управління системою зв'язку

На рисунку прийнято наступні позначення:

BAI – блок алгоритмів ідентифікації,

$AG_1(O_k) \div AG_m(O_k)$ – генетичні алгоритми,

$SU_1(P_k) \div SU_n(P_k)$ – системи локального управління,

O_i – образи стану каналу на інтервалі Δt ,

P_k – параметри каналу, з яким працює система управління SU_i ,

K_k – критерії еволюційного розвитку каналів та системи зв'язку в цілому.

В даному випадку локальні системи управління не досліджуються, оскільки вони будуються на основі класичних підходів, в яких досить широко використовуються стохастичні моделі. Тому розглянемо алгоритм реалізації генетичної схеми аналізу функціонування і розвитку каналу зв'язку. Такий алгоритм представляє собою один з можливих варіантів реалізації генетичної схеми. Очевидно, що конкретна реалізація окремих фрагментів генетичної схеми може бути різною при різних реалізаціях конкретного алгоритму. Різні алгоритми можуть користуватись різними способами кодування генів, функціями, що визначають міру пристосування окремих хромосом при реалізації базової операції репродукції, та іншими компонентами, які є необхідними для реалізації базових фрагментів генетичної схеми перетворень відповідного опису об'єкта [1]. Наведемо опис повного циклу генетичного алгоритму для випадку, коли використовується бінарне кодування в рамках окремих хромосом. При цьому поточні значення параметрів, що характеризують функціонування каналу, розміщуються в бінарній формі

запису в рамках однієї хромосоми. Початкову популяцію значень параметрів можна вибирати на основі використання генератора випадкових чисел, що описується рівномірною функцією розподілу у відповідності з співвідношенням:

$$V = [v_1, v_2, \dots, v_N],$$

де $v_i = [v_{1i}, v_{2i}, \dots, v_{ni}]$, $i = 1, 2, \dots, N$ – i -та одинарна хромосома в популяції, v_{ki} – випадкова бінарна послідовність, що має встановлену довжину, є одним з відрізків хромосоми та кодує величину k -того параметра каналу.

Наступним кроком відповідної процедури є декодування генотипу у форму фенотипу:

$$x_i = [x_{1i}, x_{2i}, \dots, x_{ni}]^T,$$

де $x_{ki} = x_k + [(x'_k - x''_k) / (2^{m_k} - 1)] \sum_{j=0}^{m_k} v_{ki} 2^j$, $k = 1, 2, \dots, n$, m_k визначає довжину

бінарної послідовності, що кодує параметр x_k , x_k – k -тий фенотип, що визначається в межах $[x''_k, x'_k]$.

На наступному кроці визначається міра пристосування параметрів, яка обчислюється за допомогою відповідних функцій пристосування або

$$f(x_i) = [f_1(x_i) f_2(x_i) \dots f_m(x_i)].$$

Якщо відповідна функція інтерпретується як функція вартості, то використовується наступне перетворення:

$$f_e(x_i) = C_{\max} - f_e(x_i),$$

де C_{\max} – максимальна вартість $f_e(x_i)$ в поточній популяції. Якщо функція $f_e(x_i)$ є міра типу виграшу, то використовується наступне перетворення:

$$f_e(x_i) = f_e(x_i) - C_{\min},$$

де C_{\min} – мінімальна вартість в даній популяції.

Якщо кількість поколінь, які отримані протягом генетичних перетворень, дорівнює встановленій кількості для даного перетворення, то алгоритм записується на даному етапі функціонування, а отриманий індивид приймається як оптимальний на даний момент. В протилежному випадку алгоритм переходить до виконання наступного кроку.

На наступному кроці виконується змішування сукупності параметрів, які для зручності будемо називати вектором міри пристосування кожного індивіду у відповідності з формулою:

$$f'(x_i) = f(x_i) / \sum_{j=1}^N \delta_{ij},$$

де $\delta_{ij} = 1 - \|x_i - x_j\|_p$, якщо $0 \leq \|x_i - x_j\|_p < 1$, та $\delta_{ij} = 0$, якщо $\|x_i - x_j\|_p \geq 1$,

$\|x_i - x_j\|_p = \sqrt[p]{(x_i - x_j)^{T p-2} (x_i - x_j)}$, $p = \text{diag}\{\varphi_1, \varphi_2, \dots, \varphi_n\}$, де φ_k ($k = 1, 2, \dots, n$) –

k -та діагональ i -го індивіду і визначається із співвідношення:

$$\varphi_k = \frac{\Delta_k}{\varepsilon},$$

де Δ_k – ненульовий діапазон значень шуканого значення k -того параметру, ε – число, що визначає одиницю масштабу або частину, на які поділено відповідний діапазон.

Кожному індивіду призначається скалярна величина $\rho(x_i)$, яку будемо називати її рангом, для чого можна використовувати мажоритарний метод, наприклад, метод, що визначається уявленнями про Парето-оптимізацію:

$$\rho(x_i) = \mu_{\max} - \mu(x_i) + 1,$$

де $\mu_{\max} = \max \mu(x_i)$, $\mu(x_i)$ – міра переваги або кількість індивідуумів, що домінують над ним в сенсі Парето або μ_{\max} має максимальну вартість серед всіх $\mu(x)$.

На наступному кроці батьківську популяцію V_p вибираємо з популяції індивідуумів V , наприклад, використовуючи метод пропорційного залишкового вибору. Для цього використовуємо наступні дії:

- визначаємо очікувану кількість копій індивідуумів з популяції у відповідності з співвідношенням:

$$e(x_i) = [\rho(x_i) / \sum_{i=1}^N \rho(x_i)] N, \quad i = 1, 2, \dots, N,$$

- копіюємо $N_{\text{int}} = \sum_{i=1}^N \lfloor e(x_i) \rfloor$ індивідуумів до популяції, яку приймемо батьківською, використовуючи тільки цілу частину відповідного числа $e(x_i)$, а також визначимо кількість невистаючих індивідуумів для заданої популяції з співвідношення $N' = N - N_{\text{int}}$,
- визначаємо розподіл індивідуумів відповідно до залишку від числа $e(x_i)$, по співвідношенню

$$q(x_i) = \sum_{j=1}^i [e(x_j) - \lfloor e(x_j) \rfloor],$$

- N' кратно розділяємо інтервали генератора, наприклад, поле рулетки:
- генеруємо випадкове число $r \in [0, 1]$,
- визначаємо першого випавшого індивідуума в популяції, для якого виконується умова $r \leq q(x_i) / q(x_N)$,
- копіюємо вибраною індивідуума до батьківської популяції.

На наступному кроці з батьківської популяції формуємо популяцію потомків V , виконуючи наступні дії:

- випадковим чином формуємо батьківські пари і виконуємо на них

операцію одноточкового схрещування з вибраною ймовірністю схрещування $p_c \in [0, 4; 1]$ наступним чином:

- генеруємо випадкове число $r \in [0, 1]$ з рівномірним розподілом для кожної батьківської пари,
- схрещуємо пари, якщо $r \leq p_c$, які є випадково вибраними точками схрещування j_c , що належать до дискретних одиниць $[1, m-1]$, де m – довжина хромосоми, і заміняємо місцями частини хромосом випадково вибраних батьківських пар, починаючи від біту $j_c + 1$ до біту m включно,
- проводимо бінарну мутацію з вибраного ймовірністю мутації наступним чином:
- генеруємо число $r \in [0, 1]$ з рівномірним розподілом для кожного гену,
- проводимо мутацію, якщо $r \leq p_m$, що полягає в заміні біту на обернене значення (звичайно p_m дорівнює оберненому числу індивідуумів популяції).

На останньому етапі заміняємо попередню популяцію новою популяцією потомків V' у відповідності з стратегією повної репродукції, що записується у вигляді $V \rightarrow V'$ і повертаємося до декодування кожного індивідуума до відповідних їм фенотипів.

Оскільки в процесі функціонування генетичних алгоритмів вибираються найбільш пристосовані індивідууми, то на черговому етапі кількість найбільш пристосованих може збільшитися до такої міри, що всі вибрані індивідууми будуть між собою дуже близькими по параметрах, що їх характеризують. Більш того, на деякому етапі такі індивідууми можуть стати однаковими або клонами, що приведе до повної відсутності еволюційного розвитку. Для уникнення цієї ситуації можливі наступні варіанти:

- збереження індивідуумів, які не є найбільш пристосованими до умов, що змінюються, в окремому місці, вибираючи останні на основі такого параметра як відстань між групою близьких індивідуумів та віддаль цих індивідуумів від фенотипу. Така функція близькості може описуватись співвідношенням $\delta_{ij} = 1 - \|x_i - x_j\|_p$, якщо $0 \leq \|x_i - x_j\| < 1$, та $\delta_{ij} = 0$, якщо $\|x_i - x_j\|_p \geq 1$;
- зміна умов, що визначають рівень пристосування індивідуумів, причому ці умови можуть бути різними по відношенню до різних індивідуумів;
- така модифікація функцій пристосування індивідуумів, яка дозволила б зберегти необхідний рівень різноманітності серед окремих хромосом.

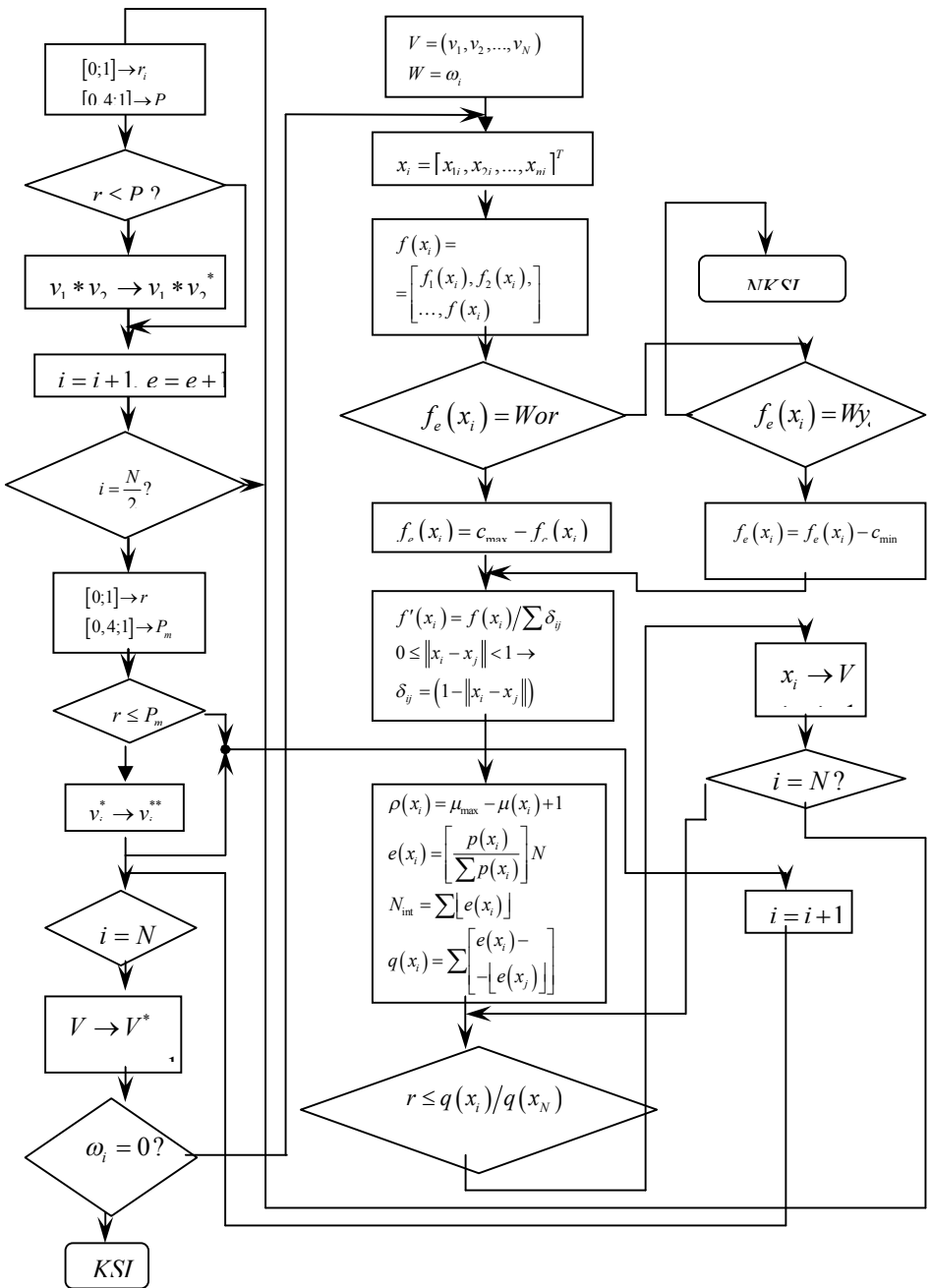


Рис. 2. Блок-схема алгоритму $AG_i(O_k)$.

Найбільш поширеним є перший варіант, який передбачає створення ним для відкладання не найбільш пристосованих хромосом [1]. Цей підхід реалізується в рамках описаної вище процедури.

Зміна умов визначення міри пристосування є більш складним варіантом реалізації протидії виникненню клонів, оскільки процес виявлення найбільш пристосованих хромосом є в більшості випадків імовірнісним і при реалізації алгоритмів для конкретних задач імовірнісні процеси вибору ґрунтуються на предметній області задачі, яка розв'язується.

Модифікація функцій пристосування приводить до необхідності створення цілого ряду функцій, які по певних ознаках необхідно вибирати на одному кроці реалізації процедури репродукції. На рис. 2 наведена блок-схема алгоритму, що реалізує генетичну схему AG_i .

На блок-схемі додатково до описаних вище позначень використовуються наступні позначення:

- KSI – завершення роботи алгоритму AG_i ,
- $NKSI$ – несанкціоноване завершення роботи алгоритму AG_i ,
- W_{ar} – вартісна характеристика,
- W_{yq} – характеристика, що означає перевагу.

Наведені вище підходи до модифікації генетичних алгоритмів з метою уникнення можливості виникнення клонів використовуються в тій чи іншій формі в рамках реалізації самих генетичних схем. Такий підхід приводить до збільшення складності алгоритмічної реалізації відповідних генетичних схем та до складностей, пов'язаних з алгоритмами модифікації генетичних схем, які в цьому випадку повинні розміщуватися в рамках генетичних алгоритмів [2].

Система ідентифікації представляє собою сукупність алгоритмів, які аналізують результати роботи генетичних алгоритмів $AG_i(O_k)$ і на основі такого аналізу здійснюють управління. Очевидно, що, як мінімум, повинні існувати критерії, відносно яких може проводитися аналіз результатів роботи генетичних алгоритмів.

Деякі з таких критеріїв можуть полягати у наступних обставинах та величинах:

- міра близькості окремих хромосом після визначеної кількості циклів роботи генетичних алгоритмів, яку будемо позначати критерієм типу ρ або K_ρ ,
- характеристика динаміки зміни, що визначають той чи інший спосіб роботи генетичних алгоритмів, яку будемо називати критерієм динамічних змін в генетичних алгоритмах K_z ,
- величина впливу зміни результатів роботи генетичних алгоритмів на послідовних етапах його роботи на величини змін споживчих або

інтегральних параметрів об'єкта управління, яку будемо називати критерієм інтегрального впливу K_v ,

- міра взаємозв'язку між змінами результатів роботи окремих генетичних алгоритмів, яку будемо називати критерієм взаємозв'язку K_G ,
- величина інтервалу часу функціонування генетичних алгоритмів, протягом якого не відбувалася модифікація генетичних алгоритмів K_m .

Розглянемо більш детально перший критерій K_p . Суть роботи найпростішого генетичного алгоритму полягає в оптимізації деякого об'єкту чи деякого процесу, який може бути представлений в рамках структури хромосом, генів та, відповідно, популяцій хромосом. Очевидно, що будь-яка оптимізація передбачає існування оптимізаційних принципів, які дозволяють стверджувати факт існування чи відсутності оптимізації. Перш за все відмітимо, що такі критерії можуть ґрунтуватися на структурах алгоритмів чи на аналізі динаміки їх функціонування, оскільки ці два фактори є найбільш загальними з точки зору характеристики генетичних алгоритмів. Цей тип критеріїв будемо називати внутрішніми критеріями об'єкту, який передбачається оптимізувати.

Одним з основних параметрів або однією з основних компонент, стосовно яких може бути сформульований критерій, що може використовуватися для аналізу роботи генетичних алгоритмів, є міра подібності хромосом між собою в останній популяції за певний період часу роботи алгоритму. Слід зауважити, що аналіз роботи генетичних алгоритмів може проводитись не тільки з метою оптимізації їх роботи. Для того, щоб можна було ефективно використовувати цей критерій, необхідно прийняти метод обчислення міри близькості двох хромосом між собою. Найбільш просто ця задача розв'язується в тому випадку, коли хромосома представляється в бітовому вигляді, а окремий ген відповідає одному біту. Тоді відстань між двома хромосомами x_1 і x_2 буде вимірюватись у вигляді відстані Лемінга між двома бінарними кодами. В цьому випадку відстань між хромосомами вимірюється кількістю бітових позицій, в яких коди двох хромосом різні або

$$d_H(x_i, x_j) = \sum_{k=1}^n (\alpha_{ik} + \alpha_{jk}) .$$

У випадку, коли хромосоми складаються з окремих бітових послідовностей, які описують окремі параметри, останні перетворюються у відповідні фенотипи і між фенотипами обчислюється різниця у відповідності з співвідношенням:

$$d(x_i, x_j) = [x_{ik} + (x'_{ik} - x''_{ik}) / (2^{m_k} - 1)] \sum_{j=1}^{m_k} v_{ki} 2^j - [x_{jk} + (x'_{jk} - x''_{jk}) / (2^{m_k} - 1)] \sum_{j=1}^{m_k} v_{kj} 2^j$$

1. *Батищев Д.А.* Генетические алгоритмы решения экстремальных задач. – Воронеж: Изд-во ВГТУ, 1995.

2. *Олешко Т.І.* Дослідження взаємозв'язку між інформаційною моделлю та базами даних, що входять в склад інформаційної технології. Захист інформації – 2005. Спецвипуск, с.30-35.