

МАТЕМАТИЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ ІНТЕЛЕКТУАЛЬНОЇ ПІДТРИМКИ АНАЛІЗУ ТА ТЕСТУВАННЯ ПРОГРАМ

Мархоцький Н.Ю.¹⁾, Олійник А.П.²⁾, Судейченко Д.В.³⁾, Костик Б.П.⁴⁾

*Західноукраїнський національний університет
1)магістрант;2)аспірант;3)магістрант;4)аспірант*

І. Постановка проблеми

Тестування відносять до динамічних методів верифікації, який дозволяє перевірити програму у процесі її виконання. Підходи до тестування та розроблення тестів використовуються для чіткого визначення мети тестування програми чи її компонента/модуля. Тестові набори для перевірки програми формуються на основі застосування методів верифікації або надаються після використання альтернативного підходу (білої скриньки)[1,2].

У той же час поточні наукові дослідження показують, що існують підходи, розробка та застосування яких може значно поліпшити якість тестів, що генеруються, виражається в ступені покриття ними коду, що тестується. Особливий інтерес представляють підходи на основі аналізу коду програм (тестування білої скриньки), оскільки витягнута з тексту програми інформація може сприяти подальшому збільшенню показника якості покриття коду [2,3].

Тому наразі більш реалістичним і ефективним для практичного використання у компаніях, що займаються виробництвом ПЗ, є динамічний підхід, заснований на фактичному виконанні програми при деяких значеннях вхідних змінних та наступному аналізі потоків даних [3].

Еволюційна основа, що лежить в основі генетичного алгоритму (ГА), використовує велику кількість випадкових тестових даних, згенерованих на початковому етапі, після чого проводиться послідовна "еволюція" даних з метою поліпшення якості покриття коду, що тестується. У зв'язку з вищезазначеним виникає припущення про можливість адаптації генетичного алгоритму для реалізації ідеї еволюційного покращення тестових даних з точки зору максимізації покриття ними коду, що тестується. У зв'язку з вищезазначеним виникає припущення про можливість адаптації генетичного алгоритму для реалізації ідеї еволюційного покращення тестових даних з точки зору максимізації покриття тестованого коду.

II. Мета роботи

Метою дослідження є розробка математичного та програмного забезпечення для інтелектуальної підтримки аналізу та тестування програм.

III. Застосування основних еволюційних операцій в задачах генерації тестових даних

Першим етапом роботи ГА, як було сказано вище, є формування початкової популяції. Стандартно у ГА початкова популяція визначається випадковим чином, тобто перша популяція наповнюється хромосомами з рівномірно розподіленими випадковими значеннями тестових змінних в межах заздалегідь заданих областей визначення. Незважаючи на те, що існують і інші методи ініціалізації, саме випадковий метод формування початкової популяції використовується у цій роботі, оскільки він дозволяє забезпечити достатню різноманітність отримуваних тестових варіантів.

Після того, як було сформовано початкову популяцію, починається послідовне виконання основних еволюційних операцій відповідно до алгоритму ГА. Далі розглянемо особливості їх застосування у завданні генерації тестових даних, використаних у цій роботі.

Відбір (селекція). Хромосоми у популяції сортуються відповідно до значення функції пристосованості. Сортування дозволяє розділити хромосоми на найкращі (з більшим значенням функції) та гірші (з меншим значенням функції), що є необхідним критерієм для відбору.

Перед етапом відбору частина найкращих хромосом переноситься до нового покоління без змін, тим самим формуючи пул найкращих хромосом. Інші хромосоми нового покоління будуть отримані в результаті схрещування. Саме для визначення пулу придатних для схрещування хромосом проводиться операція відбору. Пул кращих хромосом (P_{best}) та пул хромосом для схрещування (P_{cross}) можуть як збігатися, так і не збігатися.

Методи відбору мають свої переваги та недоліки. Найбільший інтерес представляють методи турнірного відбору та метод випадкового визначення пар. Перший метод дозволяє розділити популяцію на окремі підгрупи, з яких відбувається відбір батьків, а другий – досягти більшого

розмаїття, що у завданні по генерації тестових даних відіграє важливу роль. Тому у цій роботі пропонується гібридний метод відбору на основі цих методів.

Перш ніж приступати безпосередньо до відбору, необхідно визначити, яка частина нового покоління буде отримана перенесенням найкращих хромосом, а яка - за допомогою схрещування. Емпірично для завдання генерації тестових даних було визначено такі параметри. В якості кращих хромосом для перенесення в нове покоління без змін вибираються 20% хромосом з найбільшою функцією пристосованості, що становлять пул кращих P_{best} . Інші 80% хромосом нового покоління будуть відібрані з використанням гібридного методу:

- 50% хромосом для схрещування вибираються тільки з пулу P_{best} ;
- 50% хромосом відбиратимуться з популяції загалом, тобто значення функції пристосованості не враховується.

Схрещування серед усіх хромосом необхідне для досягнення більшого різноманіття популяції. Таким чином, гібридний метод дозволяє краще досліджувати тестовий код та відбирати хромосоми з метою досягнення більшої різноманітності. Після відбору пари батьківських хромосом вони формують одного нащадка з використанням еволюційної операції схрещування.

Схрещування. Схрещування дозволяє отримувати нові тестові набори перемішуванням значень вже існуючих. У термінах генетичного алгоритму, дві батьківські хромосоми обмінюються генами для отримання нащадка. Основна мета використання схрещування полягає в ітераційному покращенні покоління послідовним схрещуванням найкращих хромосом для отримання найкращих рішень. У цій роботі схрещування використовується також для збільшення різноманітності популяції гібридним методом відбору.

Існують різні методи схрещування. При їх описі, для зручності, визначимо одну з батьківських хромосом як материнську (x_{mother}), а другу – як батьківську (x_{father}). Схрещування за однією позицією. Для отримання нащадків у батьківських хромосом вибирається одна позиція, за якою відбувається обмін генами. Стандартно використовується одна статична позиція для схрещування, більш просунутим методом є визначення позиції випадковим чином. Графічно схрещування двох хромосом з 6 вхідними змінними та позицією схрещування $r = 4$ показано на рисунку 1.

Застосовуючи схрещування по одній позиції, можна отримати дві нові хромосоми, але вони не дуже відрізняються від батьківських. У разі якщо одна з частин хромосоми робить набагато більший вплив на результат, то нові тестові набори, отримані таким чином, не будуть вносити жодного різноманітності в популяцію, що негативно позначається на роботі алгоритму загалом. Одноточкове схрещування є простим для розуміння та реалізації, але має надто суттєві недоліки.

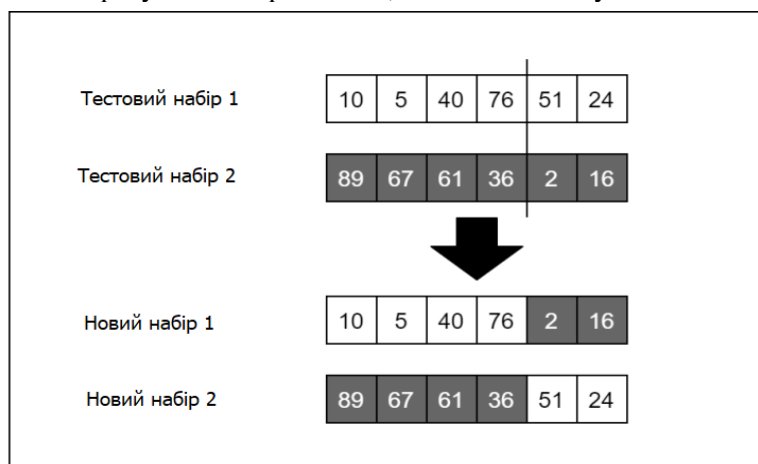


Рисунок 1– Приклад одноточкового схрещування тестових наборів

Так, збільшення кількості позицій для схрещування може сприяти отриманню більш відмінних тестових наборів, оскільки дозволяє більш глибоко і різноманітно комбінувати гени в батьківських хромосомах. Уніфіковане схрещування є одним з типів схрещування, яке відрізняється від інших у підході до отримання нащадків. У цьому методі кількість генів та кількість позицій для схрещування збігаються. Кожен ген нащадка випадково успадковується від одного з батьків. Оскільки батьки у генетичному алгоритмі рівнозначні, кожен ген нащадка має однакову ймовірність вибору від одного або іншого батька. Це призводить до того, що кожен ген нащадка може успадкувати значення або від одного, або від іншого батька з ймовірністю 0.5.

Щоб нівелювати недоліки безперервного ГА, пропонується використовувати еволюційну операцію змішування, яка дозволяє вводити у популяцію нові значення в залежності від значень генів батьків. Як було сказано раніше, змішування необхідне для отримання нових значень генів у популяції, що є істотним для генерації нових значень у безперервному варіанті ГА.

Змішування є механізмом, що дозволяє отримати значення генів у нащадків, відмінне від батьків. Можна виділити такі методи змішування:

- Пряме змішування значень. Нове значення гена обирається випадковим чином у проміжку між батьківськими значеннями.

- Середні значення. Значення нового гена отримується усередненням значень батьківських генів. Очевидно, що такий метод погано підходить для збільшення різноманітності, але може використовуватися, наприклад, для пошуку відповідних меж вхідних змінних.

Мутація є еволюційною операцією, яка дозволяє вносити в популяцію нові значення за допомогою випадкової зміни генів. Мутація застосовується тільки до нащадків, тому найкращі тестові набори, що передаються у нове покоління безпосередньо (без операції схрещування), не змінюватимуться. Дуже важливо задати вірну ймовірність мутації, оскільки занадто невеликий шанс може нівелювати переваги мутації в цілому, і алгоритм може не підібрати тестові набори для виходу більш складними шляхами. Занадто високий шанс, навпаки, непередбачувано змінюватиме нащадків, через що нівелюються переваги використання ГА порівняно з випадковим пошуком.

Мутація застосовується до кожного гена окремо. Залежно від реалізації випадкового шансу мутації, вона може бути не застосована до хромосом зовсім або змінити тільки один ген (рисунок 2) або змінити кілька.



Рисунок 2 – Приклад одноточкової мутації тестового набору, в якому значення другої змінної змінилося

Хоча мутація є досить потужним інструментом для забезпечення різноманітності популяції, необхідно користуватися нею обережно. Мутація може як допомогти у знаходженні складніших шляхів, так і ускладнити пошук. Емпірично для завдання генерації тестових даних було визначено оптимальну ймовірність мутації на рівні 5% (0.05), яка, в сукупності з змішуванням, дозволила забезпечити достатню різноманітність популяції. Таким чином, ітераційно виконуючи раніше розглянуті еволюційні операції алгоритм забезпечує пошук тестових даних.

Висновок

Основні поняття генетичного алгоритму, такі як ген, хромосома і популяція, були адаптовані для вирішення задачі генерації тестових даних. Показано застосування безперервного варіанта ГА, визначено поняття невід'ємних хромосом, описано особливості реалізації еволюційних операцій для розв'язання даної задачі. Описано основні етапи генетичного алгоритму. Пропонується новий метод відбору, що ґрунтується на виборі кращих хромосом, випадковому формуванні пар та турнірному відборі. Для забезпечення більшої різноманітності популяції використовується еволюційна операція змішування, яка дозволяє вводити нові значення у популяцію.

Список використаних джерел

1. Jamil, M.A.; Nour, M.K.; Alhindi, A.; Awang Abhubakar, N.S.; Arif, M.; Aljabri, T.F. Towards Software Product Lines Optimization Using Evolutionary Algorithms. *Procedia Comput. Sci.* 2019, 163, 527–537.
2. Alsewari, A.A.; Zamli, K.Z.; Al-Kazemi, B. Generating t-way test suite in the presence of constraints. *J. Eng. Technol. (JET)* 2015, 6, 52–66.
3. Малихіна М. П., Частікова В. А., Власов К. А. Дослідження ефективності роботи модифікованого генетичного алгоритму в задачах комбінаторики // *Сучасні проблеми науки та освіти.* – 2013. – No 3.