

В целом, ИМ позволяет получить точные оценки значений ρ_k для конкретного словаря, а также сопоставить их с результатами возможных попыток улучшения качества БС за счет исключения некоторых слов с определенными свойствами, например, слов с малой востребованностью [6].

Более точную информацию для решения упомянутой задачи оптимизации БС могли бы дать «точные» оценки конкретного вклада потенциально исключаемых слов с учетом не только их востребованности, но и взаимосвязей с «близкими» словами. Описанная ИМ может служить инструментальной основой для такого точечного исследования.

Литература

1. Литвинов В.А., Майстренко С. Я., Юденко О.П. Экспериментальная оценка эффективности автоматического обнаружения типовых ошибок пользователя по словарям русского и украинского языков // System Analysis and Information Technologies SAIT 2012, April 24, 2012 Kyiv, Ukraine. - P. 374.
2. Словари русского языка, <http://speakrus.ru/dict/>.
3. Словарь Лопатина, http://royallib.ru/book/lopatin_vladimir/russkiy_orfograficheskiy_slovar.html.
4. Словарь английского языка wordlist.txt, www-personal.umich.edu/~jlawler/wordlist.
5. Словарь английского языка brit-a-z.txt, <http://www.curlewcommunications.co.uk/resource/british.zip>.
6. Litvinov V.A, Maistrenko S.Y. About formation of the basic vocabulary in the system of automatic detection of user errors // System Analysis and Information Technologies SAIT 2014, May 26-30, 2014 Kyiv, Ukraine. - P. 322.

УДК 004.94

ДОСЛІДЖЕННЯ РОБОТИ ГЕНЕТИЧНИХ АЛГОРИТМІВ ВІДНОВЛЕННЯ ЧАСОВИХ РЯДІВ ЗАСОБАМИ ІМІТАЦІЙНОГО МОДЕЛЮВАННЯ

Ю.О. Олійник, Р.А. Шигида

Національний технічний університет України «КПІ», кафедра автоматизованих систем обробки інформації та управління, Україна

Постановка проблеми

У базовому варіанті робота генетичних алгоритмів (ГА) налічує такі операції:

1. Створення початкової популяції.
2. Репродукція особин за допомогою операцій схрещування та мутацій.
3. Проводиться оцінювання та селекція особин популяції.

4. Формування оновленої популяції шляхом заміни особин популяції.

Цикл 2-4 повторюється доки не буде досягнуто умову закінчення генетичного пошуку.

В роботі [1] досліджено застосування імітаційного моделювання для оптимальної конфігурації обчислювального програмно-апаратного комплексу прогнозування за методом адаптивного МГУА. Але не досліджено структуру обчислювальних процесів виконання ГА.

Цілі дослідження

В даній роботі необхідно засобами імітаційного моделювання дослідити вплив параметрів ГА на оптимізацію обчислювальних процесів прогнозування часових рядів.

Аналіз останніх досліджень та публікацій

У роботі [2] пропонується ГА, що здатний до самонавчання, та аналізується задача виявлення закономірностей з часових рядів, після чого отримані знання використовуються для прогнозування. У реалізованому алгоритмі виконується одноточкове схрещування шаблонів. Після етапу рекомбінації отримані шаблони перевіряються на наявність надлишкових тверджень за принципом пошуку схожих компонентів при нішуванні. Подібні компоненти вилучаються з шаблону. У роботі [3] запропонована реалізація ГА, що визначає оптимальні параметри моделі часового ряду на основі нечітких множин: кількість біт, та два дійсні числа для корегування юніверсумів. Такий алгоритм підходить для короткострокового прогнозування стаціонарних та не-стаціонарних рядів. У роботі [3] представлено дослідження розміщення завдань на обчислювальних ресурсах для виконання ГА. Задача зводиться до вибору таких параметрів запуску, що цільова функція оцінки ефективності плану стає мінімальна. В експеримент проводився з залученням обчислювального кластера та з урахуванням топології та властивостей обчислювальної мережі. В результаті було дослідження було доведено ефективність роботи алгоритму планування при високій конкуренції за ресурси обчислювальної мережі. В то же час в дослідженнях не розглянуто використання імітаційного моделювання для дослідження роботи генетичних алгоритмів.

Виклад основного матеріалу

Основною задачею ГА є знаходження максимуму функції пристосованості. Ціллю імітаційного моделювання є мінімізація T_{TOTAL} (часу завершення розрахунків). Основними обчислювальними процесами моделі є:

оцінювання та селекція;

репродукція;

заміна поточної популяції.

Параметрами ГА є:
величина хромосоми L ;
кількість ітерацій.

У якості досліджуваного ГА, взято алгоритм відновлення часових рядів. На вхід алгоритму надходить часовий ряд фіксованої довжини. Алгоритмом генеруються набір хромосом, довжина яких відповідають довжині вхідного ряду. Якість набору визначається схожістю сгенерованого ряду до вихідного, чим вища оцінка тим краща пристосованість. На кожні ітерації покращуючи якість набору. На кінцевому етапі отримуємо відновлений ряд. Для збору статистики було зроблено не менше ніж 3 прогони для довжини хромосоми від 4 до 15.

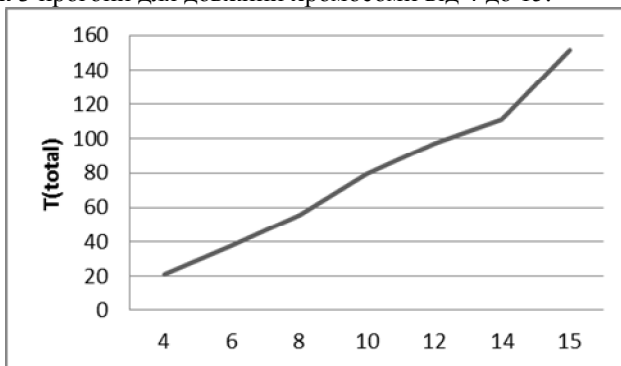


Рис. 1. Залежність T_{TOTAL} від довжини L .

Для моделювання вищенаведеного процесу створимо модель системи масового обслуговування (СМО). Параметри моделі отримуємо за допомогою тестових випробувань реалізації ГА: статистичну інформацію по обчислювальних процесах. Структура моделі наведена на рис.2.

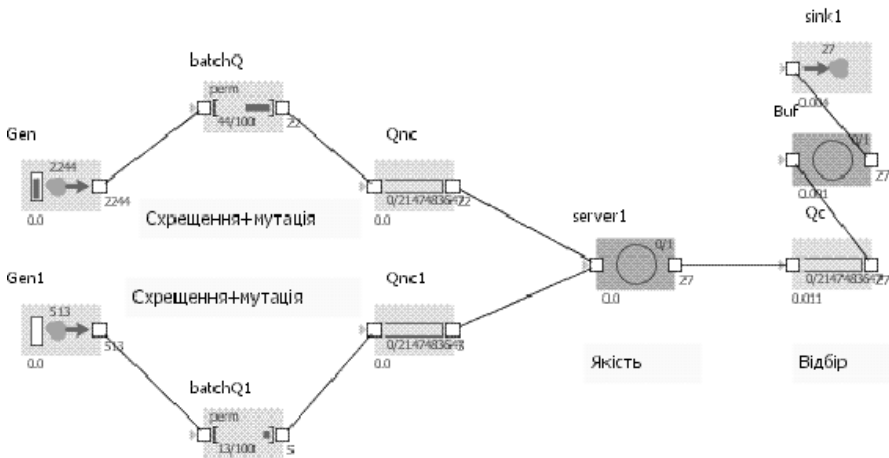


Рис.2. Структура моделі

Отримано наступні результати моделювання:

1. З рис.1 видно, що час розрахунку лінійно залежить від довжини хромосоми N .
2. Найбільш завантажений процес – схрещення та мутація. Даний процес може виконуватись паралельно на обчислювальних вузлах, за рахунок поділу множин хромосом.
3. Для зменшення втрат часу передачі даних від обчислювального вузла до іншого застосовано елемент Batch для групування особин перед оцінкою якості та селекцією.
4. Кількість ітерації практично не впливає на поведінку моделі.
5. Для мінімізації $T_{\text{ТОТАЛ}}$ необхідно використовувати десятки обчислювальних процесів схрещення та мутації.

Література

1. Олійник Ю.О., Виноградов О.М., Красовський К.М., Система розподілених обчислень прогнозних показників за методом адаптивного МГУА // Наукові вісті НТУУ “КПІ”. – 2007. – № 5.
2. Гуляницький Л.Ф., Павленко А.І. Розробка і дослідження еволюційних методів прогнозування. [Електронний ресурс] / Гуляницький Л.Ф., Павленко А.І. // System Analysis and Information Technologic. – 2013. – Режим доступу до ресурсу: <http://sait.kpi.ua/books/sait2013.ebook.pdf>.
3. А.Б. Новиков. Генетический алгоритм планирования конкурирующих за канал передачи данных пластичных заданий [Електронний ресурс] / А.Б. Новиков // Вычислительные методы и программирование. – 2012. – Режим доступу до ресурсу: <http://agora.guru.ru/abrau2012/pdf/652.pdf>.