

**М. Д. ГОДЛЕВСЬКИЙ**, д-р техн. наук, проф., зав.каф. АСУ НТУ «ХПІ»;  
**О. О. АБАБІЛОВ**, студент, НТУ «ХПІ»

## **РОЗРОБКА ТА НАЛАШТУВАННЯ ПАРАЛЕЛЬНИХ ГЕНЕТИЧНИХ АЛГОРИТМІВ ДЛЯ РОЗВ'ЯЗАННЯ ЗАДАЧІ СТВОРЕННЯ РОЗКЛАДУ ЗАНЯТЬ ВУЗУ НА ОСНОВІ GRID- СИСТЕМИ**

У статті розглядаються питання автоматизації процесу побудови розкладу занять вузу. Обґрунтовується доцільність застосування генетичних алгоритмів і обчислень Grid. Розроблено структуру особини, описані три типи паралельних генетичних алгоритмів, розроблені їх процесні представлення, виділені критерії оцінки їхньої якості й налаштування.

В статье рассматриваются вопросы автоматизации процесса построения расписания занятий вуза. Обосновывается целесообразность применения генетических алгоритмов и вычислений Grid. Разработана структура особи, описаны три типа параллельных генетических алгоритмов, разработаны их процессные представления, выделены критерии оценки их качества и настройки.

In article, questions of automatised process of high school schedule development are given. The expediency of genetic algorithms and Grid computations application is proved. The structure of an individual is developed, three types of parallel genetic algorithms are described, their process views are developed, and criteria of their quality estimation and adjustment are allocated.

**Вступ.** Задача складання розкладу занять відноситься до класу NP-повних [1]. Це означає, що для її розв'язання досі не знайдено поліноміального алгоритму й імовірно, що його взагалі не існує. Тому задача складання розкладу потребує значних ресурсів: процесорного часу, а також оперативної та енергонезалежної пам'яті, проте від якості розкладу занять залежить і якість та економічна ефективність навчання, комфортність навчання студентів і роботи професорсько-викладацького складу і т.д.

Генетичні алгоритми є потужним інструментом оптимізації загального призначення [2]. Вони моделюють принципи розвитку біологічних систем згідно концепції виживання найбільш прилаштованого та природного відбору, описаного в Дарвіна [3]. Генетичні алгоритми добре підходять для розв'язання задачі складання розкладу, який має задовольняти великій кількості нетривіальних умов, що можуть бути враховані під час розрахунку функції придатності. Крім того, генетичні алгоритми за своєю природою працюють із сукупністю індивідів, виконуючи паралельний пошук, тому вони легко можуть бути реалізовані на паралельних архітектурах, що дає змогу значно підвищити швидкість пошуку та якість отриманих рішень. Останні десятиліття активно розвиваються розподілені широкомасштабні (глобальні) обчислення – обчислення Grid [4]. Вони дозволяють вирішувати задачі великої розмірності на доступних у великій кількості комп'ютерах по всій земній кулі.

**Вимоги до алгоритму.** Вибір підходящого алгоритму викликає певні складнощі [2]. Алгоритм має задовольняти двом вимогам, що протирічать одна одній:

- простота для розуміння, кодування, налагодження тощо;
- ефективне використання ресурсів комп'ютера та швидкість виконання.

В інформатиці термін «ефективність» використовується, щоб описати властивості алгоритму, які стосуються того, наскільки великі вимоги до ресурсів він висуває. Існує багато метрик, за якими вимірюють ефективність алгоритму. Найбільш часто зустрічаються наступні:

- швидкість або час виконання – час, який потрібний алгоритму, щоб завершитися;
- «простір» – пам'ять, що використовується алгоритмом під час його роботи;
- необхідна пропускна здатність мережі під час нормального функціонування;
- розмір зовнішньої пам'яті – тимчасовий дисковий простір, який потребує алгоритм.

Нарешті, бажано, щоб алгоритм був придатний для рішення цілого класу задач – це скоротить питому собівартість його реалізації й дослідження. Генетичні алгоритми чудово відповідають цій вимозі.

При реалізації генетичного алгоритму потрібно розробити функцію оцінювання придатності, структуру хромосом особини. До останньої висуваються наступні вимоги:

- особина несе всю необхідну інформацію рішення задачі, тобто про варіант розкладу;
- за особиною можна відносно легко розрахувати функцію придатності, тобто зважений показник якості розкладу занять;
- особина має якомога меншу кількість генів.

Остання вимога знаходиться у протиріччі до перших двох, але вона дуже важлива з огляду на розмір пам'яті, яка необхідна генетичному алгоритму.

**Структура особини.** Введемо спочатку набір визначень.

Ресурс – це кожний з об'єктів, для яких складається розклад. Є чотири типи ресурсів: 1) викладач; 2) клас (академічна група, підгрупа чи потік); 3) аудиторія; 4) відрізок часу (перший чи другий тиждень, день та пара). Подія – це кортеж, що містить по одному об'єкту кожного з типів ресурсів. Для кожної події точно відомі клас та викладач, а час та місце повинні бути назначені в процесі пошуку розв'язку серед множин припустимих аудиторій та відрізків часу. Розклад – це множина подій. Об'єкти деяких подій тривіально залежать від об'єктів певної іншої події. Наприклад, якщо у розкладі є два тижні, то на кожне заняття, що проводиться в обидва тижні, треба завести дві пов'язані події, які відбуваються в одній аудиторії в один

день та пару, але в різні тижні. Заняття двох підгруп однієї групи мають проходити у один і той же час.

Нетривіальні вимоги до розкладу будемо задавати за допомоги часткових штрафних функцій. Кожна з них співвідносить розкладу ненегативне число. Будемо розрізняти обов'язкові та необов'язкові штрафні функції. Перші задають жорсткі вимоги, виконання яких обов'язкове для припустимих розкладів. Другі – це побажання. Зважена сума часткових штрафних функцій, взята з протилежним знаком, дасть значення функції придатності.

Найпростіша структура особи (див. рис. 1). Кожен індивід має 4 хромосоми. Кожна хромосома відповідає типу ресурсу і містить стільки генів, скільки є подій. Стрілки показують наявність залежностей ресурсів в подіях.

Очевидно, є пряма відповідність між розкладом і набором генів у хромосомах особи. Генетичний алгоритм змінює гени лише тих хромосом, що відповідають змінним ресурсам (тобто аудиторіям та академічним парам), тому перші дві хромосоми можна зберігати у одному екземплярі для всієї популяції. В цілому потрібно зберігати  $2n + 2np$  генів, де  $p$  – кількість особин в популяції.

Цю структуру можна поліпшити. Залежні ресурси (а таких серед хромосом часу та місця буде близько половини, тому що розклад складається для двох тижнів) можна не зберігати в особині, тому що вони однозначно розраховуються через незалежні ресурси. Нова структура включає одну особину, що зберігає лише незалежні гени (див. рис. 2, 3). В цілому буде потрібно зберігати  $2n + np$  генів.

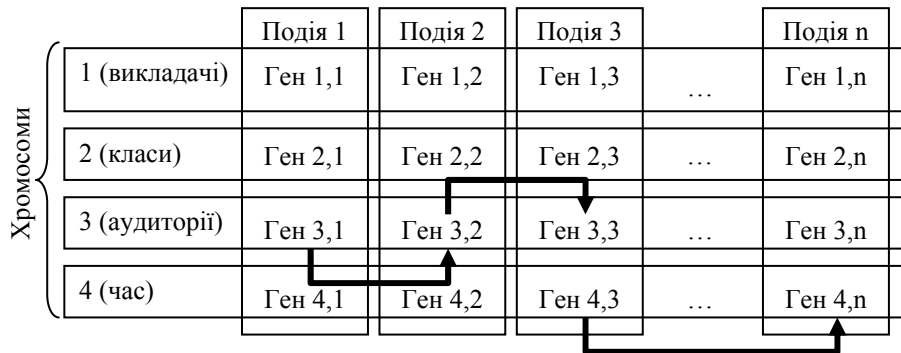


Рис. 1 – Найпростіша структура особи

Хромосома	Ген 3,1	Ген 3,n	Ген 4,1	Ген 4,2	Ген 4,3
-----------	---------	---------	---------	---------	---------

Рис. 2 – Структура особи, що враховує залежність подій

Хромосома	№ 3,1	№ 3,n	№ 4,1	№ 4,2	№ 4,3
	Ген 3,1	Ген 3,n	Ген 4,1	Ген 4,2	Ген 4,3

Рис. 3 – Структура особи для алгоритму з інверсією

**Алгоритмічне забезпечення.** Для розв'язання задачі пошуку розкладу пропонуються три різних паралельні генетичні алгоритми, процесні представлення яких наведено на рис. 4. На ефективність алгоритмів будуть впливати параметри:

- початкова популяція;
- стратегія уточнення цільової функції;
- рівень мутацій;
- рівень схрещувань;
- стратегія відбору;
- об'єм популяції;
- стратегія формування нової популяції;
- локальний пошук та специфічні оператори (наприклад, інверсія);
- кількість обчислювальних елементів.

Паралельний генетичний алгоритм, що використовує модель ведучого-веденого. Він працює з єдиною популяцією, але розрахунок функції придатності виконується паралельно на множині процесів-ведених (SLAVE). Процес-ведучий (MASTER) виконує всі інші задачі.

Паралельний генетичний алгоритм з багатьма підпопуляціями, що використовує острівну модель. Цей метод вимагає підрозділу популяції на деяке число підпопуляцій (демів), які розташовуються на множині процесів-ведених (SLAVE). Деми розділено один від іншого (географічна ізоляція), і індивіди конкурують тільки в межах дему. Час від часу спрацьовує додатковий оператор – міграція: деякі індивіди переміщуються (копіюються) від одного дему до іншого. Процес-ведучий (MASTER) займається координацією.

На ефективність алгоритму впливають такі специфічні параметри:

- число мігрантів;
- стратегія відбору мігрантів;
- стратегія заміщення мігрантів;
- довжина епохи (задається фіксовано або розраховується динамічно після реєстрації застою).

Паралельний генетичний алгоритм з динамічними демами. Є множина процесів-ведучих (MASTER), відповідальних за вибір і спарювання. Інші операції (мутація, схрещування, оцінка придатності) виконуються на множині процесів-ведених (SLAVE). Є тільки одна популяція, але процеси-ведучі фактично поділяють її на динамічні демы, які перекриваються.

На ефективність алгоритму впливають такі специфічні параметри:

- кількість процесів-ведучих;
- кількість процесів-ведених;
- фактор перекриття.

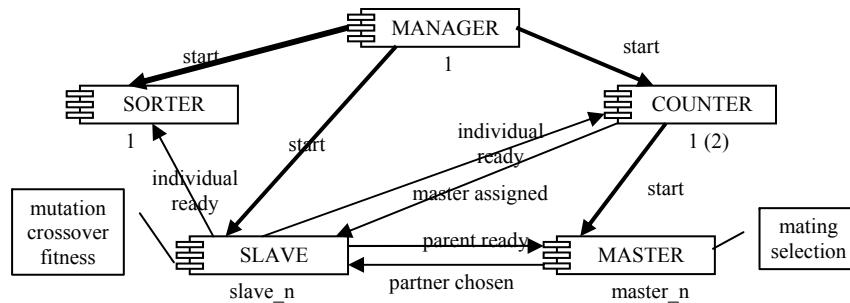
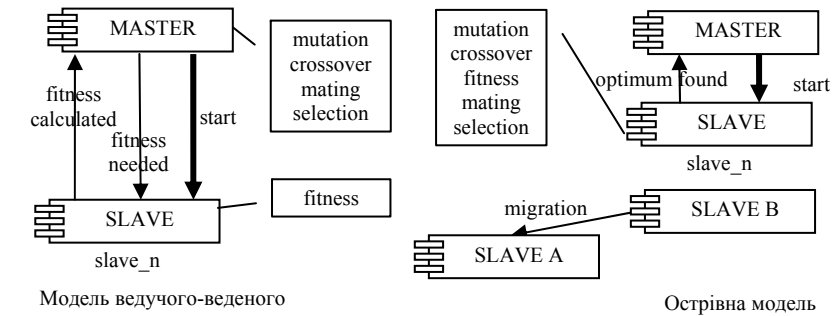


Рис. 4 – Процесні представлення паралельних генетичних алгоритмів

**Висновки.** Побудовано математичну модель для задачі складання розкладу факультету. Обрано та описано паралельні генетичні алгоритми та розроблено їхні процесні представлення.

**Список літератури:** 1. Т. Кормен, Ч. Лейзерсон, Р. Ривест. Алгоритмы : построение и анализ / Пер. с англ. под ред. А. Шеня. – М. : МЦНМО, 2002. – 960 с. 2. Holland, John H. Adaptation in Natural and Artificial Systems. Ann Arbor, MI : University of Michigan Press, 1975. – 228 p. 3. Darwin, Charles. On the Origin of Species by Means of Natural Selection. John Murray, 1859. – 502 p. 4. The Grid: Blueprint for a New Computing Infrastructure / Eds. I. Foster, C. Kesselman. – San Francisco : Morgan Kaufmann Publishers. – 1999. – 675 p. 5. Ахо, А. В., Хонкрофт, Дж., Ульман, Дж. Д. Структуры данных и алгоритмы. Пер. с англ. : М. : Издательский дом «Вильямс», 2001. – 384 с.

Надійшла до редколегії 13.11.2010