

Реалізація генетичних
алгоритмів ранжування
текстової колекції
документів з
використанням
архітектури NVIDIA CUDA

Виконав:

Семилітко М.Ю.

Керівник:

Глибовець А.М.

План

- 1. Ранжування в пошукових системах
- 2. Еволюційні алгоритми
- 3. Архітектура NVIDIA CUDA
- 4. Опис та тестування алгоритму

Ранжування в пошукових системах

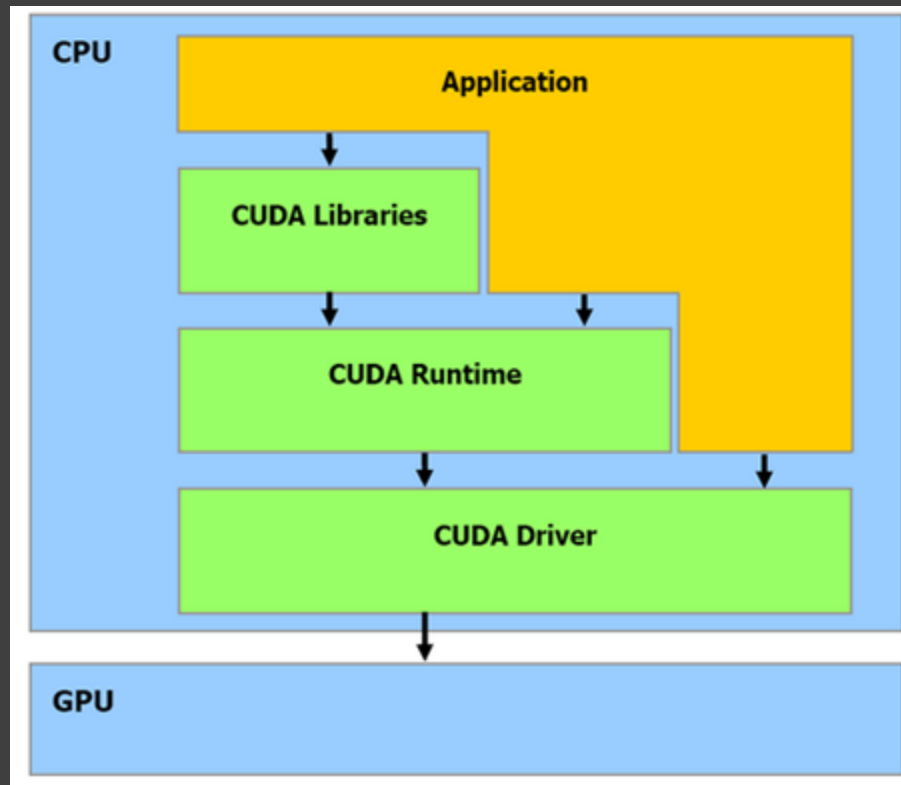


- Ранжування - це сортування джерел інформації (веб-сайтів), яке використовується пошуковими системами.
- Сьогодні основними критеріями ранжування є:
 - Посилання на домен
 - Кількість переходів за посиланням
 - Авторитет домену
 - Якість мобільної версії
 - Час завантаження сторінки
 - Загальна кількість зворотних посилань
 - Якість вмісту
 - Оптимізація контенту сторінки

Еволюційні алгоритми

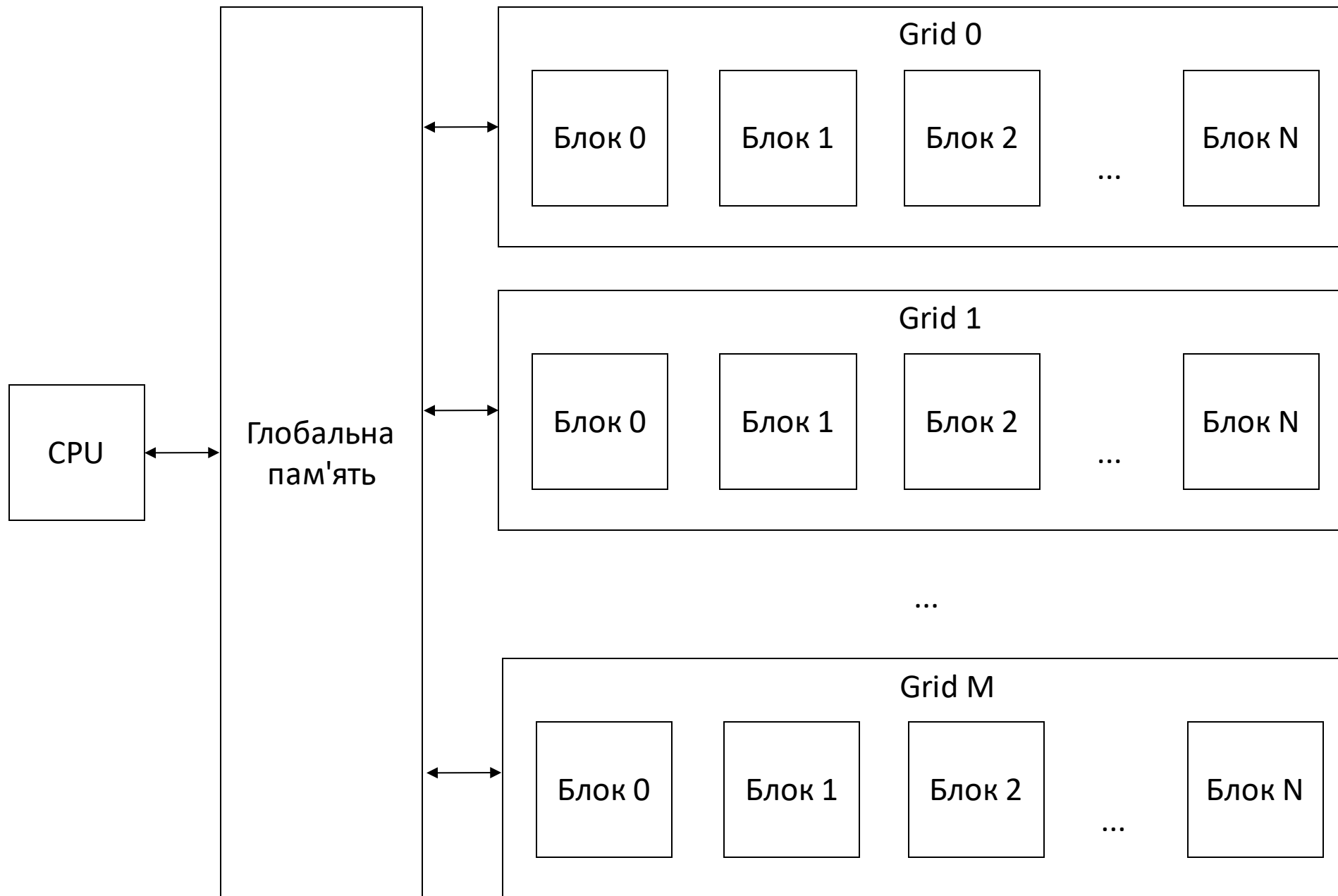
- Еволюційні алгоритми - це адаптивний пошуковий механізм, в якому моделюється процес біологічної еволюції.
- Генетичні алгоритми
- Генетичне програмування
- Еволюційне програмування
- Еволюційні стратегії





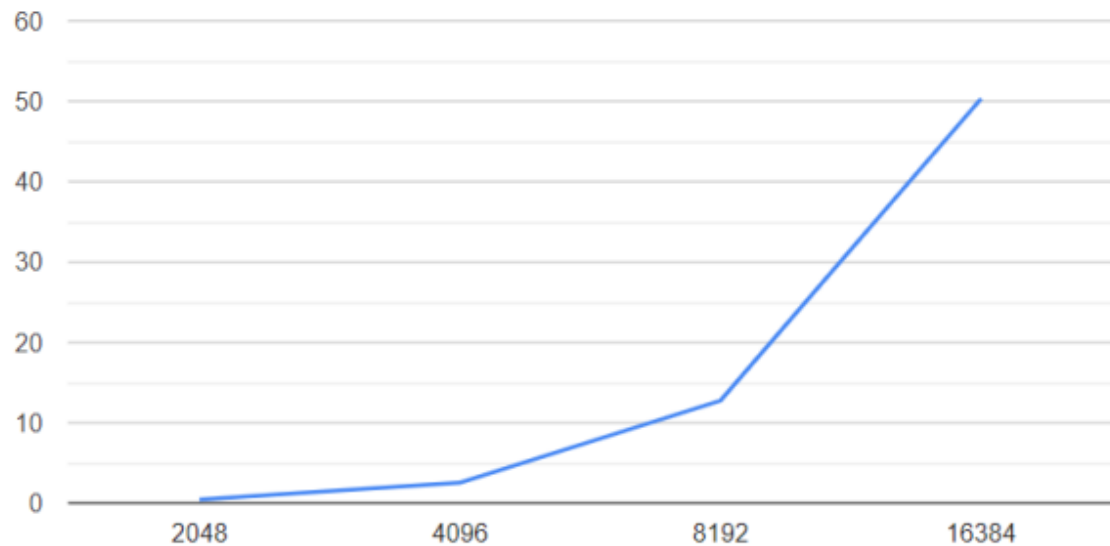
Архітектура NVIDIA CUDA

- Програмно-апаратна технологія
- В архітектурі CUDA є 3 основні абстракції:
 - Ієрархія блоків потоків
 - Спільна пам'ять
 - Синхронізація з використанням бар'єрів

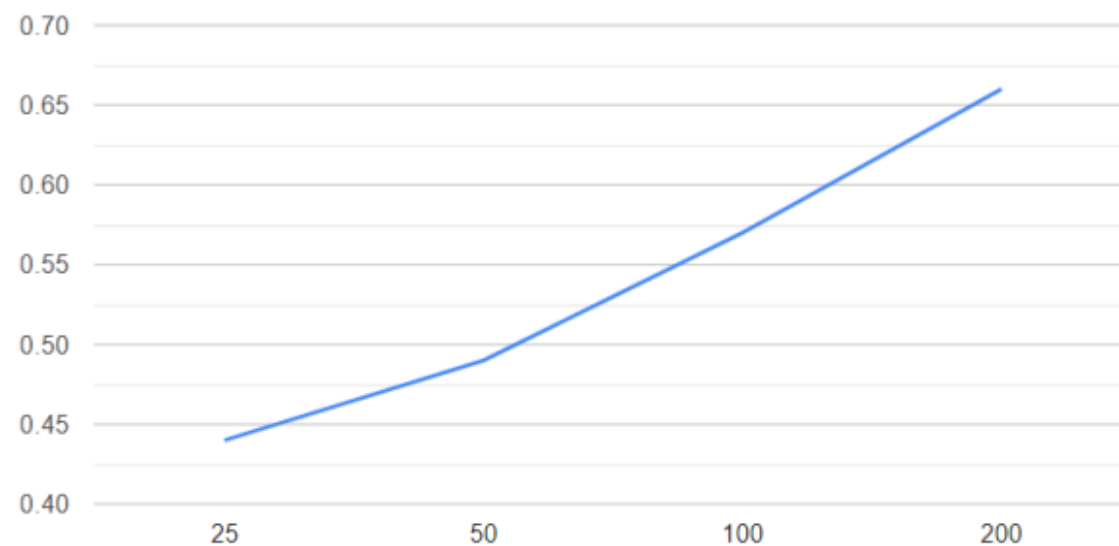


Опис та тестування алгоритму

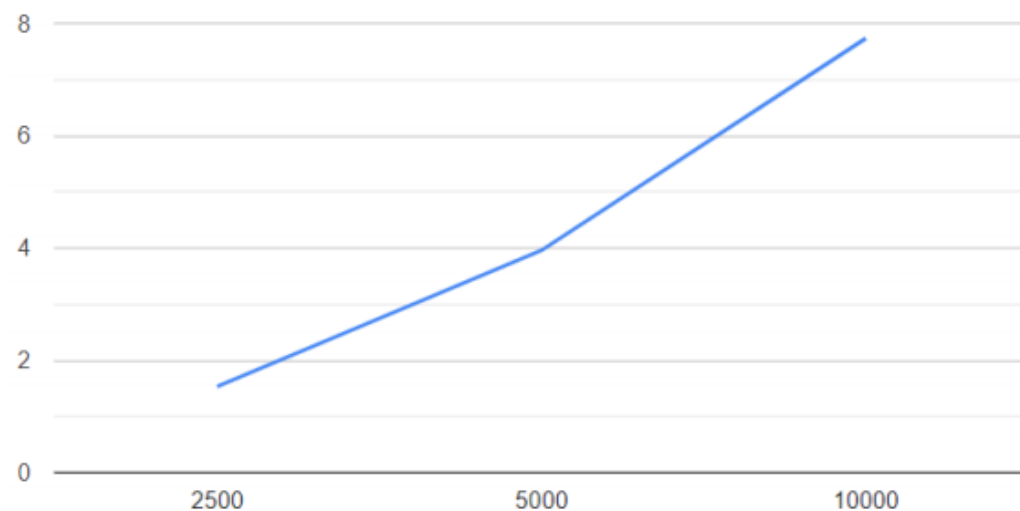
- Ініціалізація даних для оцінки популяції
 - Виділення необхідної пам'яті
 - Генерація популяції
 - Оцінка популяції фітнес-функцією
 - Відбір найкращих особин з найкращими генами
 - Схрещення відібраних екземплярів
 - Мутація нащадків
 - Оцінка нового покоління фітнес-функцією
 - Селекція
- Відбір - турнірний
 - Схрещення - лінійне
 - Мутація - випадкова
 - Селекція - нове покоління повністю замінює старе



Час 1 ітерації алгоритму для різних розмірів популяції



Час 1 ітерації алгоритму для різних розмірів генів



Час ітерації алгоритму для різних розмірів тестових даних

Висновки

Переваги

- Багато обчислень виконується паралельно
- Відсутні жорсткі прив'язки по вхідним параметрам
- Оптимальна робота з пам'яттю
- З реальними тестовими наборами має забезпечити результати близькі до найкращих
- Є можливість модифікувати кожен етап алгоритму
- Кількість генів та фітнес функція не сильно сповільнюють роботу алгоритму.

Недоліки

- Дуже повільний алгоритм схрещення
- Залежить від якості вхідних даних
- Не гарантує знаходження найкращого розв'язку

Дякую за увагу