

УДК 004.94

ПАРАЛЕЛЬНА РЕАЛІЗАЦІЯ АЛГОРИТМУ МОНТЕ-КАРЛО ДЛЯ МОДЕЛЮВАННЯ ПРОЦЕСУ УСЕРЕДНЕННЯ

Щукін Владислав

Національний технічний університет України "Київський політехнічний інститут" Україна

Анотація

В даній роботі представлена математична модель задачі усереднення в термінах теорії взаємодійних процесів; виконано моделювання її реалізації в паралельних архітектурах SIMD та MIMD. На основі результатів моделювання виконано аналіз щодо подальшого вибору архітектури.

This paper presents the mathematical model of the homogenization task in terms of theory of Communicating Sequential Processes; here is performed modeling its implementation in parallel architectures SIMD and MIMD. Based on modeling analysis performed for the further choice of architecture.

Вступ

Готова продукція виробничих комплексів зазвичай є результатом багатоетапної обробки сировини, яка постачається з різних джерел. На кожному етапі може відбуватись змішування декількох продуктів з різними параметрами з метою отримання нової суміші, яка б задовольнила технічні умови виробництва. Таку технологічну операцію називають усередненням. Математично її можна описати як сукупність випадкових процесів, які можна моделювати методом Монте-Карло [1]. Проте для забезпечення потрібної точності необхідно виконувати велику кількість статистичних випробувань, що унеможливило використання для цієї мети звичайного персонального комп'ютера [2]. Тому виникає задача побудови комп'ютерної системи з паралельною архітектурою, яка б дозволила отримати результат відповідної точності за прийнятний час.

1. Математична модель задачі усереднення

Як вже було зазначено вище, для моделювання використовується метод Монте-Карло. Його використання часто потребує генерування великої кількості випадкових чисел для досягнення необхідної точності. При цьому процес генерування послідовності випадкових чисел для одного процесу не може виконуватись паралельно декількома процесорами, оскільки існує можливість отримати неякісну випадкову послідовність. Для подолання зазначеної проблеми використовують декілька підходів. В даній роботі використовується метод, запропонований в [3]. Він полягає в тому, щоб згенерувати всю послідовність випадкових чисел на одному процесорі, а потім вже виконувати моделювання одного випадкового процесу паралельно на різних обчислювальних вузлах системи.

В рамках задачі усереднення промисловий комплекс можна представити як набір постачальників продуктів та змішувачів. Передача продуктів (потоки продуктів) від постачальників до змішувачів відбувається порціями певної маси або об'єму. Видобування первинної сировини можна описати генератором випадкових чисел (ГВЧ) з певним законом розподілу. Саме тому доцільно визначити множину процесів $G = \{G(Z_0), G(Z_1), \dots, G(Z_{k-1})\}$, які описують поведінку коливання параметрів продуктів за законом $(Z_i)_{i=0,1,\dots,k-1}$ в кожному із k потоків продуктів.

Процес усереднення потоків виконується в два етапи: 1) накопичення і усереднення параметрів в кожному потоці; 2) змішування потоків і знаходження параметрів суміші. Для реалізації першого етапу вводиться множина процесів $GA = \{GA_0, GA_1, \dots, GA_{k-1}\}$, які виконують функцію накопичення і усереднення в потоці, а для реалізації другого — процес A , який виконує змішування потоків продуктів.

Поведінку процесу змішування $PA = F(G, GA, A)$ можна описати в термінах теорії взаємодійних процесів (Communicating Sequential Processes, CSP).

Як було зазначено вище, розпаралелювання виконується тільки в межах одного потоку. Коли моделювання першого потоку завершилося, виконується моделювання другого. Процес повторюється для всіх k потоків продуктів:

$$\begin{aligned} PA_0 &= G(Z_0) \rightarrow (GA_{00} \parallel GA_{01} \parallel \dots \parallel GA_{0(C-1)}) \rightarrow PA_1; \\ PA_1 &= G(Z_1) \rightarrow (GA_{10} \parallel GA_{11} \parallel \dots \parallel GA_{1(C-1)}) \rightarrow PA_2; \\ &\dots \\ PA_{k-1} &= G(Z_{k-1}) \rightarrow (GA_{(k-1)0} \parallel GA_{(k-1)1} \parallel \dots \parallel GA_{(k-1)(C-1)}) \rightarrow A; \end{aligned} \quad (1)$$

де Z_i — закон розподілу параметру продукту в i -му потоці ($i = 0, 1, \dots, k-1$); C — кількість паралельних розгалужень така, що $(GA_{i0} \parallel GA_{i1} \parallel \dots \parallel GA_{i(C-1)}) = GA_i$; $i = 0, 1, \dots, k-1$.

2. Аналіз продуктивності виконання задачі усереднення в паралельних комп'ютерних системах із архітектурами SIMD та MIMD

Для оцінки продуктивності використовуються аналітичні моделі паралельних архітектур, які представлені в роботі [4]. Розрахунок довжини програмного рішення виконувався на основі алгоритмів роботи [2]. При цьому об'єм послідовної ділянки алгоритму складає 15% від загального об'єму обчислень. Також зроблено припущення, що комп'ютерна система має рівно стільки обчислювальних модулів, скільки існує готових до виконання паралельних гілок алгоритму. Моделювання проводилося для одного потоку продукту, тобто $k = 1$. На рисунку 1 наведено залежності часу виконання алгоритму в комп'ютерній системі від кількості паралельних гілок алгоритму.

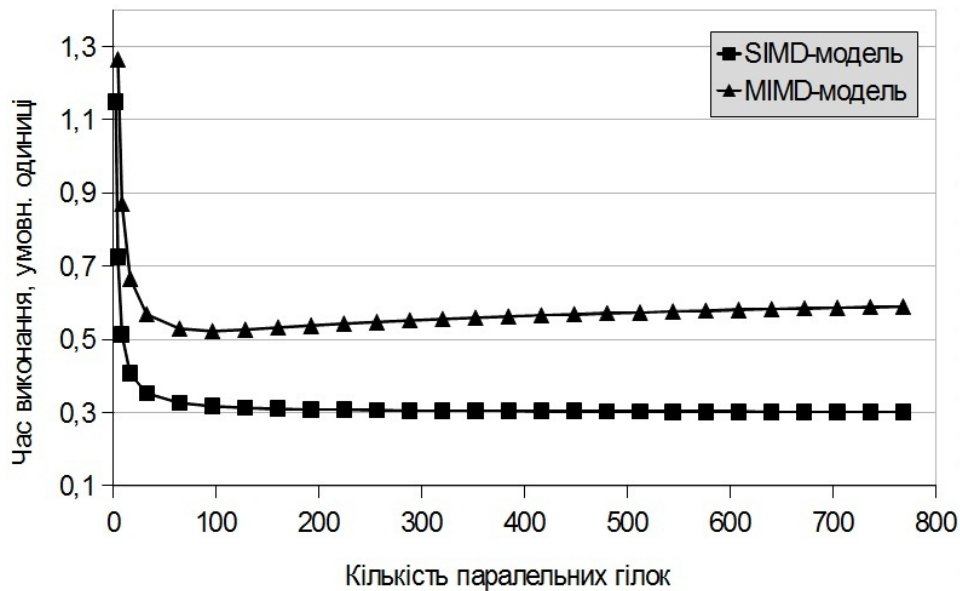


Рисунок 1 – Результати моделювання

Отримані результати дають підстави стверджувати, що для даного способу розпаралелювання алгоритму більш кращим варіантом є архітектура SIMD. При цьому слід зазначити, що кожне зі статистичних випробувань, які виконуються при моделюванні методом Монте-Карло, може бути оброблено незалежно одне від одного. Оскільки таких випробувань необхідно дуже багато, то реально кількість обчислювальних модулів є набагато меншою, ніж кількість можливих паралельних гілок алгоритму. Кожен обчислювальний модуль виконує обробку набору таких статистичних випробувань.

Висновки

В результаті виконання даної роботи була представлена модель задачі усереднення в термінах теорії CSP, в якій показано один із способів розпаралелювання методу Монте-Карло. Оцінка подальшої реалізації алгоритмів в комп'ютерних системах з паралельними архітектурами різних типів проводилася на аналітичних моделях. Результати моделювання показали, що SIMD-архітектура є більш продуктивною для даної форми алгоритму ніж MIMD-архітектура. Також слід зазначити, що були проведені додаткові дослідження, які показують, що час виконання паралельного алгоритму в архітектурі MIMD сильніше залежить від об'єму послідовної ділянки алгоритму. При описаному підході розпаралелювання алгоритму Монте-Карло необхідно на початку сформувати всю послідовність випадкових чисел (без розпаралелювання). Перспективою даної роботи є дослідження інших способів паралельної реалізації алгоритму Монте-Карло на предмет виявлення оптимальної паралельної архітектури.

Список використаних джерел:

1. Шупов Л.П. Математические модели усреднения. Справочное пособие. / Шупов Л.П. - М.: "Недр", 1978. - 287 с.
2. Купін А.І. Застосування методу Монте-Карло для комп'ютерного прогнозування якості сировини в умовах збагачувального виробництва / А.І. Купін, Д.С. Замятін, В.Б. Щукін, А.В. Петрашенко // Науковий вісник чернівецького університету. Збірник наукових праць. Комп'ютерні системи та компоненти. - 2010. - Том 1, випуск 2. - С. 130 - 137
3. Сайїд Реза, Кен Стрэндберг Прирост производительности с помощью распараллеливания приложения, использующего генератор случайных чисел [Электронный ресурс]: материалы Intel Software Network. - Режим доступа: <http://software.intel.com/ru-ru/articles/gaining-speedup-in-a-parallelizable-workload-with-dependence-on-random-number-generation/>
4. Параллельная обработка структур данных / [Шпаковский Г. И., Липницкий А. С., Черников Г. Н. и др.]; под ред. В. А. Мищенко. - Мн.: Университетское, 1988. - 272 с.