

Інфокомунікаційні системи

УДК 004.9

А.Ю. Дяченко

Черкаський національний університет імені Богдана Хмельницького, Черкаси

РОЗПАРАЛЕЛЮВАННЯ ПРОЦЕСУ СИНТЕЗУ МОДЕЛЕЙ В СИСТЕМАХ БАГАТОРІВНЕВОГО ПЕРЕТВОРЕННЯ ІНФОРМАЦІЇ

Представлені результати досліджень методу розпаралелювання процесу синтезу моделей для формування структури глобальної функції перетворення інформації при багаторівневому моделюванні. Розпаралелювання здійснюється в локальній мережі на базі операційної системи Windows. На основі проведених досліджень встановлено, що розроблений метод дозволяє значно прискорити синтез моделей в системах багаторівневого перетворення інформації.

Ключові слова: розпаралелювання, індуктивна модель, МГУА, багаторівнева модель.

Вступ

Загальновідомо, що значна частина обчислювальних задач вимагає для свого розв'язку на типовому персональному комп'ютері занадто багато часу. Прикладами таких задач є побудова багаторівневих моделей [1], криптоаналіз, аналіз зображень з телескопів для пошуку нових об'єктів в астрономії та ін. Для вирішення цієї проблеми використовується одночасна робота декількох комп'ютерів, що об'єднані в локальну мережу. Такий тип одночасної роботи називають паралельним. Крім того, майже всі сучасні мікропроцесори складаються з декількох ядер, і, для того, щоб задіяти всі можливості мікропроцесора, необхідно завантажити всі його ядра. Для цього необхідно розпаралелити існуючі алгоритми. Цілком очевидно, що в першу чергу слід розпаралелювати ті програми, алгоритми яких є вимогливими до швидкодії мікропроцесора. Саме до таких відносяться алгоритми побудови багаторівневих моделей.

Аналіз досліджень

У роботі [2] здійснено аналіз існуючих підходів до паралельних обчислень у методі групового урахування аргументів (МГУА) [3] і запропоновані підходи для розпаралелювання комбінаторного алгоритму МГУА на базі системи з декількох комп'ютерів, що об'єднані в обчислювальний кластер. Також цікавий підхід для прискорення аналізу криптосистем з використання штучних нейронних мереж (НМ) шляхом розпаралелювання обчислень на кластері запропоновано у роботі [4]. Суттєвим недоліком підходів [2] і [4] є необхідність адаптації існуючих програм для роботи на кластері, на практиці досить часто це призводить до переписування більшої частини початкових кодів програм. Крім того, існує багато різних типів кластерних систем, що в основній своїй масі не сумісні між собою.

Постановка завдання

Синтез багаторівневих моделей вимагає суттєво більших обчислювальних ресурсів в порівнянні зі звичайними моделями. В переважній більшості випадків на 1-2 порядки більше ресурсів, в першу чергу ресурсів мікропроцесора. Відповідно, розв'язок деяких задач з використанням багаторівневого моделювання займає неприйнятно багато часу. Саме тому актуальною є задача прискорення процесу побудови багаторівневих моделей шляхом розпаралелювання їх синтезу на комп'ютерах локальної мережі. При багаторівневому моделюванні перетворення інформації здійснюється за допомогою ієрархічного поєднання об'єктів моніторингу. Моделі синтезуються за багаторядним алгоритмом МГУА. Різноманітність алгоритмів забезпечує побудову моделей на всіх рівнях перетворення інформації. Метою роботи є розробка нового методу розпаралелювання процесу синтезу моделей для формування структури глобальної функції перетворення інформації при багаторівневому моделюванні. Для цього необхідно спроектувати та реалізувати відповідну програмну систему і провести дослідження з використанням даної системи. Також при переході до паралельних обчислень ставиться додаткова вимога мінімізувати зміни початкових кодів існуючої програмної системи.

Виклад основного матеріалу дослідження

Запропонована програмна система складається з двох частин: серверної та клієнтської. Серверна формує та розподіляє обчислювальні задачі між клієнтами, а також обробляє результати розрахунків з клієнтських комп'ютерів. На рис. 1 наведено діаграму варіантів використання (прецедентів) технології розпаралелювання процесу синтезу моделей.

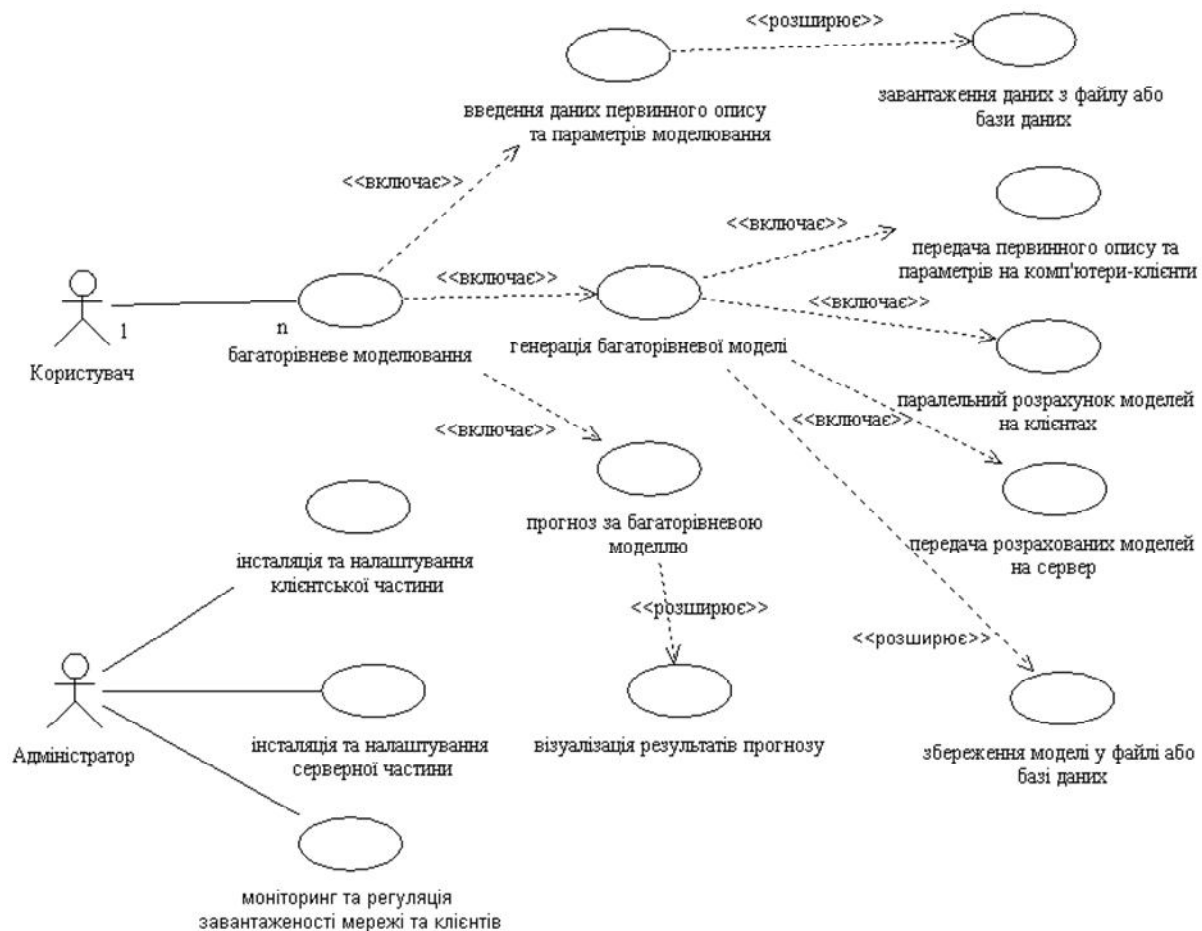


Рис. 1. Діаграма варіантів використання (прецедентів) технології розпаралелювання процесу синтезу моделей

Серверною частиною керує людина-оператор. При цьому від оператора необхідно лише задати первинний опису об'єкта моделювання та вхідні параметри багаторівневого моделювання, а весь подальший паралельний процес побудови багаторівневих моделей повністю автоматизований. Робота клієнтської частини взагалі повністю автоматизована.

Технологія містить наступні етапи:

1. Клієнти знаходяться в режимі очікування команд від сервера. Якщо з'єднання з сервером розривається, то через певний проміжок часу клієнт автоматично спробує встановити нове з'єднання з сервером. Відразу після запуску клієнтської частини вона також намагається встановити з'єднання з сервером. В режимі очікування клієнти практично не витрачаються ресурсів мікропроцесора. При частотному використанні клієнтського комп'ютера для паралельних обчислень рекомендується внести відповідну програму в список програм, що автоматично завантажуються з операційною системою. В такому випадку немає необхідності «ручного» запуску програми-клієнта.

2. Сервер формує черговий набір вхідних даних для моделювання. При цьому для моделей першого рівня вхідними даними є дані з первинного опису

об'єкта моделювання, а для моделей інших рівнів в якості вхідних даних виступають вихідні дані моделей попередніх рівнів або частина даних з первинного опису.

3. На базі чергового набору вхідних даних та параметрів моделювання сервер формує і передає на кожний з під'єднаних клієнтів поточне завдання для обчислення моделі з поточного рівня моделювання. Якщо клієнтський комп'ютер є багатоядерним то йому одночасно надсилається декілька завдань для обчислення моделей, при цьому кількість завдань прямо пропорційна кількості ядер мікропроцесора. Якщо кількість моделей на поточному рівні багаторівневого моделювання більша, ніж сумарна кількість ядер на підключених комп'ютерах-клієнтах, то процес обчислень розбивається на декілька етапів. Наприклад, якщо в поточній обчислювальній мережі сумарна кількість ядер 14 (3 2-ядерні комп'ютера та 2 4-ядерних), а нам необхідно розрахувати 26 моделей, то на першому етапі будуть розраховані перші 14 моделей з 26, а на другому останні 12 моделей. Досить часто на практиці комп'ютери, що задіяні в паралельних обчисленнях, мають однакову кількість ядер мікропроцесора але різну потужність. Тому сервер намагається передати задачі з великим обся-

гом обчислень більш потужним клієнтам, а з меншим обсягом - менш потужним.

4. Клієнти здійснюють обчислення в відповідності з раніш отриманим завданням, а сервер чекає, доки всі клієнти не завершать розрахунок заданих моделей з поточного рівня багаторівневого моделювання.

5. Після завершення розрахунку клієнти передають розраховані моделі на сервер, а сервер додає шойно розраховані моделі до багаторівневої моделі.

6. На сервері здійснюється перевірка умови завершення процесу багаторівневого моделювання: якщо моделювання завершено то відбувається перехід на крок 7, якщо ні то процес повторюється з кроку 2. Є дві умови завершення процесу багаторівневого моделювання: до досягнення заданої точності моделі або наперед фіксована кількість рівнів моделювання.

7. Клієнти переходять в режим очікування, а сервер формує остаточний вигляд багаторівневої моделі і видає результат моделювання оператору.

Фактично розпаралелювання здійснюється на двох рівнях: на рівні комп'ютерів локальної мережі та на рівні ядер мікропроцесора кожного з комп'ютерів.

Програмна система розроблено з використанням середовища розробки Embarcadero C++Builder XE6 на мові програмування C++, що надає можли-

вість її використання практично на будь якій версії операційної системи Windows, зокрема 98, Me, 2000, XP, Vista, 7 та 8. Для перенесення серверної та клієнтської частини достатньо просто скопіювати відповідні файли в одну з папок комп'ютера і запустити програму на виконання. Єдиним налаштуванням є введення на клієнті мережевого імені або локальної IP-адреси сервера. Тобто відсутня необхідність в процесі інсталяції та спеціальних налаштуваннях комп'ютерів. Що є суттєвою перевагою в порівнянні з кластерним підходом, оскільки при для налаштування кластеру необхідно інстальювати та налаштувати спеціальне програмне забезпечення. Для більшості типів кластерів це займає багато часу та вимагає високої класифікації адміністратора.

У випадку, якщо виникає потреба додати існуючий програмний код, що реалізує новий алгоритм побудови моделей в систему паралельних обчислень, то немає необхідності внесення суттєвих змін в існуючий програмний код. Достатньо лише додати можливість конвертації вхідних даних алгоритму та результатів його роботи в формати, що прийняті в системі. При кластерному підході у випадку необхідності додавання нового алгоритму слід переписати більшу частину програмного коду для її адаптації під конкретну кластерну технологію.

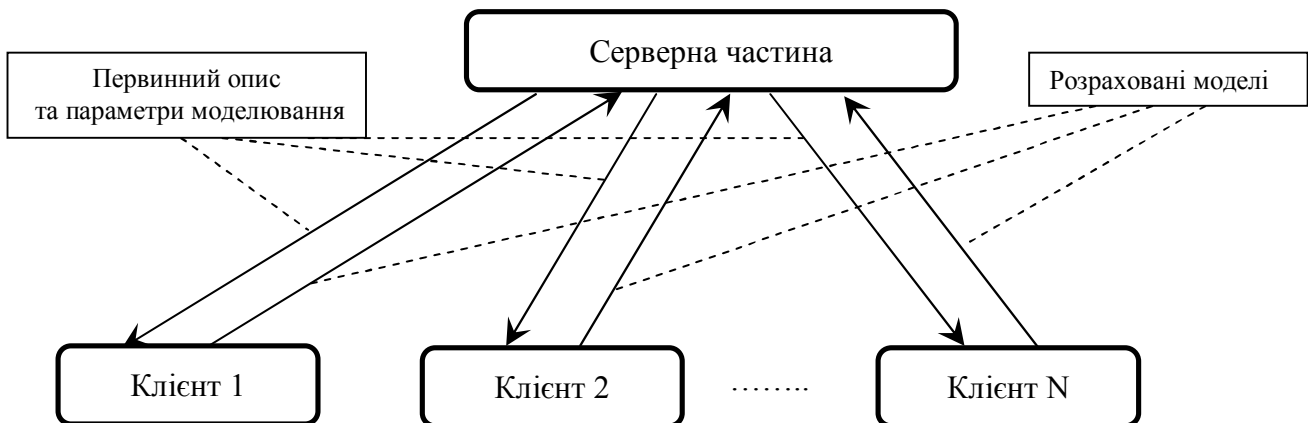


Рис. 2. Структурна схема технології розпаралелювання процесу синтезу моделей

Структурна схема, яка відображає нову технологію розпаралелювання процесів синтезу моделей подана на рис. 2. Для обміну даними між клієнтом на сервером використовується технологія Sockets, яка є загальноприйнятим стандартом для передачі даних у локальних та глобальних комп'ютерних мережах. Вказана технологія є крос-платформенною, тобто дозволяє обмінюватися даними між комп'ютерами під керуванням різних операційних систем.

Слід зазначити, що клієнтська частина працює з мінімальним пріоритетом, а тому не заважає користувачу вирішувати своє практичні задачі одночасно з розрахунком моделей, що виконується в фоновому режимі. Особливо якщо задачі користувача не вимагають багато ресурсів мікропроцесора.

В таблиці 1 наведені результати досліджень часу побудови багаторівневих моделей різноманітних предметних областей в залежності від сумарної кількості обчислювальних ядер мікропроцесорів (ОЯМ), що задіяні в паралельних обчисленнях. Первинні дані отримані с сайтів www.kdnuggets.com та www.ics.uci.edu мережі Internet.

За результатами досліджень видно, що розпаралелювання процесу формування структури глобальної функції при багаторівневому моделюванні суттєво зменшує час, що необхідний для побудови моделей. При цьому зменшення часу побудови моделей практично прямо пропорційне кількості ядер мікропроцесорів, що задіяні в паралельних обчисленнях.

Залежність часу (сек.) побудови багаторівневих моделей
в залежності від сумарної кількості обчислювальних ядер

Предметна область	1 ОЯМ	2 ОЯМ	5 ОЯМ	10 ОЯМ
1. Серцево-судинні захворювання	132,729	68,474	28,163	14,531
2. Економіка США	223,173	115,037	47,282	24,390
3. Економіка Німеччини	189,463	97,661	40,140	20,706
4. Забрудненість атмосфери	435,267	224,364	92,217	47,570
5. Захворювання на рак	384,923	198,413	81,551	42,068
6. Точність лазерного приладу	74,856	38,585	15,859	8,180
7. Хімічний реактор	284,349	146,571	60,243	31,076

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямку

Запропонований метод розпаралелювання процесу синтезу моделей дозволяє значно прискорити формування структури глобальної функції перетворення інформації при багаторівневому моделюванні, за рахунок паралельного використання обчислювальних потужностей декількох комп'ютерів-клієнтів, що об'єднані в комп'ютерну мережу. При цьому прискорення прямо пропорційне кількості та потужності під'єднаних клієнтів. Суттєвою перевагою даного методу, в порівнянні з локальними суперкомп'ютерами, є можливість практично необмеженого збільшення швидкодії за рахунок масштабування. Процес масштабування відносно простий і полягає в під'єднанні нового комп'ютера-клієнта до існуючої системи. Розпаралелювання процесу побудови моделей при багаторівневому моделюванні є важливим вдосконаленням, оскільки дозволяє значно розширити клас розв'язуваних задач. Ще однією перевагою даного підходу, в порівнянні з кластерними системами, є відносна простота під'єднання нового комп'ютера до існуючої системи. Для під'єднання нового комп'ютера з локальної мережі достатньо просто скопіювати на нього програму-клієнт.

В майбутньому планується дослідження, направлені на використання для розпаралелювання

комп'ютерів з локальної мережі під керуванням інших операційних систем: Linux, Unix, Max OS X та інших.

Список літератури

1. Голуб С.В. *Методологія створення автоматизованих систем багаторівневого соціоекологічного моделювання* / С.В. Голуб. – Автореферат дисертації на здобуття ступеня доктора техн. наук. – Черкаси: ЧНУ, 2008. – 36 с.
2. Кошулько О.А. *Технології паралельних обчислень в методі групового урахування аргументів* / О.А. Кошулько. – Автореферат дисертації на здобуття ступеня кандидата техн. наук. – К.: Інститут кібернетики ім. В.М. Глушкова НАН України, 2007. – 20 с.
3. Ивахненко А.Г. *Індуктивний метод самоорганізації моделей складних систем* / А.Г. Ивахненко. – К.: Наук. думка, 1981. – 296 с.
4. Тумоян Е.П. *Метод аналізу рюкзакових критосистем з використанням штучних нейронних мереж* / Е.П. Тумоян. – Таганрог: Технологічний інститут південного федерального університету, 2007. – С. 78-82.
5. Дяченко А.Ю. *Криптографічний захист інформаційних систем на основі індуктивних моделей* / А.Ю. Дяченко // Збірник «Вісник Інженерної академії України». – К., 2012. – Т. 2. – С. 65-58.

Надійшла до редколегії 12.06.2014

Рецензент: д-р техн. наук, проф. С.В. Голуб, Черкаський національний університет імені Богдана Хмельницького МОН України, Черкаси.

РАСПАРАЛЛЕЛИВАНИЕ ПРОЦЕССА СИНТЕЗА МОДЕЛЕЙ В СИСТЕМАХ МНОГОУРОВНЕВОГО ПРЕОБРАЗОВАНИЯ ИНФОРМАЦИИ

А.Ю. Дяченко

Представлено результати досліджень методу распараллеливания процесса синтеза моделей для формирования структуры глобальной функции преобразования информации при многоуровневом моделировании. Распараллеливание реализуется в локальной сети на базе операционной системы Windows. Установлено, что разработанный метод позволяет значительно ускорить синтез моделей в системах многоуровневого преобразования информации.

Ключевые слова: распараллеливание, индуктивная модель, МГУА, многоуровневая модель.

PARALLELIZATION OF PROCESS OF SYNTHESIS OF MODELS IN SYSTEMS OF MULTILEVEL TRANSFORMATION OF INFORMATION

A.U. Diachenko

The article presents the results of research of the parallelization method to build models for the formation of the structure of the global transformation function with multilevel modeling. Parallelization is performed in a local network under the operating system Windows. Found that the developed method can significantly accelerate the synthesis of multilevel models in systems transformation of information.

Keywords: parallelization, inductive models, GMDH, multilevel model