УДК 519.681.5

Informatics and Mathematical Methods in Simulation
Vol. 7 (2017), No. 3, pp. 234-239

ПАРАЛЕЛЬНІ АЛГОРИТМИ РОЗВ'ЯЗАННЯ ПРОСТОРОВОЇ ЗАДАЧІ ЦИФРОВОЇ ФІЛЬТРАЦІЇ ДАНИХ

М. С. Яджак

Інститут прикладних проблем механіки і математики ім. Я.С. Підстригача НАН України,
вул. Наукова, 3-Б, м. Львів, 79060, Україна; e-mail: yadzhak_ms@ukr.net

У роботі запропоновано паралельні алгоритми з автономними гілками для розв'язання просторової задачі цифрової фільтрації. Наведено оцінки прискорення, які підтверджують високу ефективність цих алгоритмів. Запропоновані паралельні алгоритми зорієнтовані для реалізації на сучасних широкодоступних високопродуктивних обчислювальних засобах – багатоядерних комп'ютерах, кластерах та системах з гібридною архітектурою. Одержані в роботі результати можуть бути використані під час дослідження в режимі реального часу стану та якості функціонування складних динамічних систем, зокрема транспортних.

Ключові слова: паралельний алгоритм, цифрова фільтрація, складна система, обмежений паралелізм, прискорення обчислень, кластер.

Вступ

У роботах [1–4] описано методику комплексного детермінованого оцінювання складних систем (СС) із ієрархічно-мережевою структурою. Цю методику використано для оцінювання стану та якості функціонування біомеханічних і робототехнічних систем локомоційного типу [5], а також для оцінювання окремих підсистем залізничної транспортної системи України [6]. У більшості випадків оцінювання СС необхідно здійснювати у режимі реального часу. Однією із важливих складових згаданої методики є попереднє оброблення великих масивів даних (Big Data) [7, 8], які використовуються під час дослідження системи. Такі вхідні дані можуть бути різних типів: числові, текстові, звукові, графічні (плоскі або просторові зображення), відеодані тощо. У [9, 10] запропоновано оптимальні за швидкодією (у заданих класах) паралельно-конвеєрні алгоритми розв'язання задач фільтрації різної вимірності. Ці алгоритми зорієнтовані на реалізацію на відповідних спеціалізованих обчислювальних засобах – квазісистолічних структурах. Тому актуальною є проблема розробки ефективних паралельних алгоритмів цифрової фільтрації даних на сучасних широкодоступних високопродуктивних обчислювальних системах – багатоядерних комп'ютерах та кластерах.

Метою цієї роботи є розробка та дослідження паралельних алгоритмів з автономними гілками для числового розв'язання просторової задачі цифрової фільтрації (ЗЦФ) даних.

Задача цифрової фільтрації та послідовний алгоритм її розв'язання

Загалом розглядувана просторова задача фільтрації полягає у виконанні C переобчислень згладжування масиву значень N змінних $x_{i,j,k}$ $(i=\overline{1,l_1};\ j=\overline{1,l_2}\ ;\ k=\overline{1,l_3})$ через рухоме вікно розміром M [9]. Для розв'язання цієї задачі можна скористатися послідовним алгоритмом, який має наступний вигляд:

$FOR\ \ l=1,\ C\ \ \ DO$

$\{\ FOR\ \ ALL\ \ (i,j,k)\in\{(i',j',k'):\ 1\le i'\le l_1;1\le j'\le l_2;1\le k'\le l_3\}\ \ \ DO$

$\{\ \ p=0$

$$FOR\ \ ALL\ \ (s_1,s_2,s_3)\in\{(s_1',s_2',s_3'):-m_1\le s_1'\le m_1;-m_2\le s_2'\le m_2;-m_3\le s_3'\le m_3\}\ DO \qquad (1)$$

$\{\ \ \ p=p+x^{(l-1)}_{i+s_1,j+s_2,k+s_3} * f_{s_1,s_2,s_3}\}$

$x^{(l)}_{i,j,k}=p\ \ \}\ \}.$

У наведеному алгоритмі $f_{-m_1,-m_2,-m_3},...,f_{0,0,0},...,f_{m_1,m_2,m_3}$ – задані вагові коефіцієнти; $x^{(l)}_{i,\ j,\ k}$ – l-те переобчислення значення змінної $x_{i,\ j,k}$. Згідно з (1) для переобчислення на l-му кроці значень змінних використовуються виключно значення, переобчислені на (l-1)-му кроці, при цьому $N=l_1l_2l_3$; $M=(2m_1+1)(2m_2+1)(2m_3+1)$. Зазначимо, що величини

$$x^{(l)}_{1-m_1,j^*,k^*},x^{(l)}_{2-m_1,j^*,k^*},...,x^{(l)}_{0,j^*,k^*};\qquad x^{(l)}_{l_1+1,j^*,k^*},x^{(l)}_{l_1+2,j^*,k^*},...,x^{(l)}_{l_1+m_1,j^*,k^*};$$

$$x^{(l)}_{i,\ 1-m_2,\ k^*},x^{(l)}_{i,\ 2-m_2,\ k^*},...,x^{(l)}_{i,\ 0,\ k^*};\qquad x^{(l)}_{i,\ l_2+1,\ k^*},x^{(l)}_{i,\ l_2+2,\ k^*},...,x^{(l)}_{i,\ l_2+m_2,\ k^*};$$

$$x^{(l)}_{i,\ j,\ 1-m_3},x^{(l)}_{i,\ j,\ 2-m_3},...,x^{(l)}_{i,\ j,\ 0};\qquad x^{(l)}_{i,\ j,\ l_3+1},x^{(l)}_{i,\ j,\ l_3+2},...,x^{(l)}_{i,\ j,\ l_3+m_3};$$

$$(j^*=\overline{1-m_2,\,l_2+m_2};\ k^*=\overline{1-m_3,\,l_3+m_3};\ i=\overline{1,\,l_1};\ j=\overline{1,\,l_2};\ l=\overline{0,\,C-1})$$

є заданими константами.

Розробка паралельних алгоритмів

Для (1) розглянемо простір ітерацій I [11], що є наступною множиною цілочисельних векторів:

$$I=\{(i,j,k,l):\ 1\le i\le l_1;1\le j\le l_2;1\le k\le l_3;1\le l\le C\}.$$

Результуючими у наведеному просторі будуть N ітерацій, що утворюють його деяку підмножину I_r:

$$I_r=\{(i,j,k,C):1\le i\le l_1;1\le j\le l_2;1\le k\le l_3\}.$$

Зазначимо, що на кожній ітерації виконуються обчислення, задані в (1) фрагментом із чотирьох останніх операторів.

На підставі використання методу пірамід [12] для розпаралелювання циклів із (1) одержуємо паралельний алгоритм з автономними гілками для розв’язання просторової ЗЦФ, який можна описати конструкцією:

$$
\begin{aligned}
&FOR\ \ ALL\ \ (k_1,k_2,k_3)\in\{(k_1^{'},k_2^{'},k_3^{'}):1\le k_1^{'}\le l_1;\,1\le k_2^{'}\le l_2;\,1\le k_3^{'}\le l_3\}\ \ DO\ \ PAR\\
&\qquad\{\,FOR\ \ l=1,\ C\ \ DO\\
&\qquad\{\,FOR\ ALL\ (i,j,k)\in\{(i_1,i_2,i_3):\max\{1,(l-C)m_t+k_t\}\le i_t\le\min\{l_t,(C-l)m_t+k_t\};\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 1\le t\le 3\ \}\ \ DO\\
&\qquad\{\ \ p=0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2)\\
&FOR\ \ ALL\ (s_1,s_2,s_3)\in\{(s_1^{'},s_2^{'},s_3^{'}):-m_1\le s_1^{'}\le m_1;\,-m_2\le s_2^{'}\le m_2;\,-m_3\le s_3^{'}\le m_3\}\ \ DO\\
&\qquad\{\ \ p=p+x_{i+s_1,\,j+s_2,\,k+s_3}^{(l-1)}*f_{s_1,\ s_2,\ s_3}\ \ \}\\
&\qquad x_{i,j,k}^{(l)}=p\ \ \}\ \}\ \}.
\end{aligned}
$$

У наведеному алгоритмі тип паралельності PAR є $AUTON$ [13] і задає паралельне виконання N гілок, які не взаємодіють між собою. Очевидно, що на практиці це є можливим, якщо можна залучити N процесорів (ядер) або обчислювальних вузлів. Однак, кожен багатоядерний комп’ютер або кластер має наперед визначену кількість обчислювальних елементів, яка зазвичай є значно меншою, ніж N. Тому для ефективного розв’язання просторової ЗЦФ нами запропоновано алгоритм з обмеженим паралелізмом [14], який дозволяє враховувати обсяг наявних ресурсів обчислювального засобу:

$$
\begin{aligned}
&FOR\ \ ALL\ (r_1,r_2,r_3)\in\{(p,r,s):\ 1\le p\le P_1;\,1\le r\le P_2;\,1\le s\le P_3\}\ \ DO\ PAR\\
&\{\,FOR\ \ l=1,\ C\ \ DO\\
&\{\,FOR\ ALL\ \ (i,j,k)\in\{(i_1,i_2,i_3):\ \ \max\{1,\ (l-C)m_t+(r_t-1)l_t/P_t+1\}\le i_t\le\\
&\qquad\qquad\qquad\qquad\qquad\le\min\{l_t\,,\ (C-l)m_t+r_tl_t/P_t\};\,1\le t\le 3\ \}\,DO\\
&\{\ \ p=0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3)\\
&FOR\ ALL\,(s_1,s_2,s_3)\in\{(s_1^{'},s_2^{'},s_3^{'}):\ -m_1\le s_1^{'}\le m_1;\ -m_2\le s_2^{'}\le m_2;\ -m_3\le s_3^{'}\le m_3\}\,DO\\
&\{\ \ p=p+x_{i+s_1,\ j+s_2,\ k+s_3}^{(l-1)}*f_{s_1,\ s_2,\ s_3}\ \ \}\\
&\quad x_{i,\ j,\ k}^{(l)}=p\ \ \ \}\ \ \}\ \ \}.
\end{aligned}
$$

Тут P – кількість паралельних гілок, які можна реально одночасно виконати на кластері або багатоядерному комп’ютері; $P=P_1P_2P_3,\ P<N$ і $l_1,\ l_2,\ l_3$ є кратними відповідно до $P_1,\ P_2,\ P_3$. Тип паралельності PAR в (3) є $AUTON$.

Твердження. Алгоритми (1), (2) та (3) є еквівалентними за інформаційним графом.

Доведення цього твердження випливає із структури терм-історій (дерев) алгоритмів (1), (2) та (3) обчислення значення деякої змінної $x_{i,\ j,k}$ на заданому кроці [9].

Аналогічно, як і в [15], нами встановлено, що згідно із цими деревами обчислюється один і той же вираз.

Зазначимо, що у (2) та (3) процес згладжування можна дещо покращити, якщо останні чотири рядки в них замінити фрагментом:

$\{\ p_{i,j,k} = 0$

$FOR\ \ ALL\ (s_1, s_2, s_3) \in \{(s_1', s_2', s_3') : -m_1 \le s_1' \le m_1; -m_2 \le s_2' \le m_2;\ -m_3 \le s_3' \le m_3\}\ DO$

$\{\ p_{i,j,k} = p_{i,j,k} + x_{i+s_1,\ j+s_2,\ k+s_3} * f_{s_1,\ s_2,\ s_3}\ \}$

$x_{i,\ j,\ k} = p_{i,\ j,\ k}\ \ \}\ \}\ \}$.

Одержані унаслідок цього паралельні алгоритми дозволять у кожній автономній гілці під час переобчислення значення деякої змінної на заданому кроці використовувати значення, які є вже переобчисленими на цьому ж кроці.

Оцінювання алгоритмів фільтрації

Використовуючи сформульоване вище твердження, одержимо прискорення для запропонованих паралельних алгоритмів фільтрації (2) та (3). Припустимо, що t_{op} – це час виконання подвійної операції додавання-множення. Тоді для виконання алгоритмів (1)–(3) потрібен відповідно такий час:

$$T_1 = t_{op} MNC\,;$$

$$T_2 = t_{op} MC[1 + (2C(C-1)m_1 m_2 m_3 + (2/3)(2C-1)(m_2 m_3 + m_1 m_3 + m_1 m_2) + m_1 + m_2 + m_3)(C-1)];$$

$$T_3 = t_{op} MC\left[\frac{N}{P} + \left(m_1 \frac{l_2}{P_2}\frac{l_3}{P_3} + m_2 \frac{l_1}{P_1}\frac{l_3}{P_3} + m_3 \frac{l_1}{P_1}\frac{l_2}{P_2} + (2/3)\left(m_1 m_2 \frac{l_3}{P_3} + m_1 m_3 \frac{l_2}{P_2} + \right.\right.\right.$$

$$\left.\left.\left. + m_2 m_3 \frac{l_1}{P_1}\right)(2C-1) + 2 m_1 m_2 m_3 (C-1) C\right)(C-1)\right].$$

Зауважимо, що наведені оцінки для часу виконання алгоритмів (1)–(3) враховують лише час виконання арифметичних операцій додавання та множення. На підставі цих оцінок одержуємо прискорення паралельних алгоритмів (2) та (3) порівняно з послідовною реалізацією:

$$S_2 = \frac{T_1}{T_2} = \frac{N}{1 + (2C(C-1)m_1 m_2 m_3 + (2/3)(2C-1)(m_2 m_3 + m_1 m_3 + m_1 m_2) + m_1 + m_2 + m_3)(C-1)}\,;$$

$$S_3 = \frac{T_1}{T_3} = \left[\frac{1}{P} + \left(\frac{m_1}{l_1 P_2 P_3} + \frac{m_2}{l_2 P_1 P_3} + \frac{m_3}{l_3 P_1 P_2} + \frac{2}{3}\left(\frac{m_1 m_2}{l_1 l_2 P_3} + \frac{m_1 m_3}{l_1 l_3 P_2} + \frac{m_2 m_3}{l_2 l_3 P_1}\right)(2C-1) + \right.\right.$$

$$\left.\left. + \frac{2 m_1 m_2 m_3}{N}(C-1)C\right)(C-1)\right]^{-1}.$$

Зазвичай на практиці $l_i >> m_i$, $l_i >> C$ $\forall i: \ i = \overline{1,3}$, а також l_i відрізняється ($l_i > m_i C^3$) від $m_i C^3$ $\forall i: \ i = \overline{1,3}$ не менш, як на декілька порядків, тому прискорення S_2 буде значним, а $S_3 \approx P$. Отже, S_3 за зроблених припущень є близьким до свого оптимального значення. У разі $l_1 = l_2 = l_3 = 1000$; $C = 10$; $m_1 = m_2 = m_3 = 1$ одержуємо, що $S_2 \approx 502512.56$, а для параметрів задачі фільтрації $l_1 = 100$; $l_2 = 500$; $l_3 = 200$; $C = 5$; $m_1 = 2$; $m_2 = m_3 = 3$ маємо: $S_2 \approx 1425.11$.

Розроблені паралельні алгоритми (2) та (3) зорієнтовані для ефективної реалізації на багатоядерних комп'ютерах (засоби зі спільною пам'яттю), кластерах (засоби з розподіленою пам'яттю) та обчислювальних засобах з гібридною архітектурою (кластери, що залучають для обчислень блоки графічних процесорів) [16, 17].

Висновки

У роботі розроблено та досліджено паралельні алгоритми з автономними гілками для розв'язання просторової ЗЦФ на сучасних високопродуктивних обчислювальних засобах. Наведено оцінки прискорення, які підтверджують високу ефективність запропонованих алгоритмів фільтрації. Одержані результати можуть бути використані для попереднього оброблення в режимі реального часу просторових зображень та експериментальних даних для дослідження багатовимірних об'єктів СС.

Подальші дослідження автор вбачає у виявленні особливостей застосування одержаних результатів на практиці для конкретних предметних областей, зокрема і під час оцінювання якості функціонування таких складних динамічних систем, як транспортні системи, системи газо-, водо-, енергопостачання тощо.

Список літератури

1. Поліщук, Д. О. Комплексне детерміноване оцінювання складних ієрархічно-мережевих систем. Частина I. Опис методики / Д. О. Поліщук, О. Д. Поліщук, М. С. Яджак // Системні дослідження та інформаційні технології, 2015. – № 1. – С. 21 - 31.
2. Polishchuk, D. About Evaluation of Complex Dynamical Systems / D. Polishchuk, O. Polishchuk // Journal of Complex Systems, 2013. – Article ID 204304, 6 p. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1155/2013/204304 (дата звернення: 5.08.2017).
3. Polishchuk, D. Complex Evaluation of Hierarchically-Network Systems / D. Polishchuk, O. Polishchuk, M. Yadzhak // Automatic Control and Information Sciences, 2014. – 2, № 2. – Pp. 32 - 44.
4. Поліщук, Д. О. Інформаційна технологія комплексного детермінованого оцінювання складних ієрархічно-мережевих систем / Д. О. Поліщук. – Автореф. дис. к. т. н., спеціальність 05.13.06 – інформаційні технології. – Львів: Українська академія друкарства, 2017. – 21 с.
5. Поліщук, О. Д. Оптимізація оцінки функціонування опорно-рухового апарату людини / О. Д. Поліщук // Комп'ютерна математика. Оптимізація обчислень: Зб. наук. праць. – Київ: Інститут кібернетики ім. В. М. Глушкова НАН України, 2001. – Т. 2. – С. 360 - 367.
6. Поліщук, Д. О. Оцінювання стану колійного господарства Укрзалізниці / Д. О. Поліщук // Вісник Дніпропетровського національного університету залізничного транспорту імені акад. В. Лазаряна, 2012. – Вип. 41. – С. 203 - 211.
7. Polishchuk, O. Big Data Processing in Complex Hierarchical Network Systems I: Structures and Information Flows / O. Polishchuk, D. Polishchuk, M. Tyutyunnyk, M. Yadzhak // AASCIT Communications, 2016. – 3, № 3. – Pp. 112 - 118.
8. Polishchuk, O. Big Data Processing in Complex Hierarchical Network Systems II: Computer Environments and Parallelization / O. Polishchuk, D. Polishchuk, M. Tyutyunnyk, M. Yadzhak // AASCIT Communications, 2016. – 3, № 3. – Pp. 119 - 124.
9. Анисимов, А. В. Построение оптимальных алгоритмов массовых вычислений в задачах цифровой фильтрации / А. В. Анисимов, М. С. Яджак // Кибернетика и системный анализ, 2008. – № 4. – С. 3 - 14.

10. Яджак, М. С. Високопаралельні алгоритми та засоби для розв'язання задач масових арифметичних і логічних обчислень / М. С. Яджак. – Автореф. дис. д. ф.-м. н., спеціальність 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем. – Київ: КНУ імені Тараса Шевченка, 2009. – 33 с.
11. Lamport, L. The parallel Execution of DO Loops / L. Lamport // Communications of ACM, 1974. – 17, № 2. – Pp. 83 - 93.
12. Вальковський, В. О. Проблеми подальшого розвитку та модифікації методу пірамід для розпаралелювання циклів / В. О. Вальковський, М. С. Яджак // Мат. методи та фіз.-мех. поля, 2000. – 43, № 1. – С. 68 - 75.
13. Вальковский, В. А. Распараллеливание алгоритмов и программ. Структурный подход / В. А. Вальковский. – М.: Радио и связь, 1989. – 176 с.
14. Яджак, М. С. Алгоритмы с ограниченным параллелизмом для решения одной задачи цифровой фильтрации / М. С. Яджак // Проблемы управления и информатики, 2001. – № 6. – С. 109 - 118.
15. Яджак, М. С. Об оптимальном в одном классе алгоритме решения трёхмерной задачи цифровой фильтрации / М. С. Яджак // Проблемы управления и информатики, 2000. – № 6. – С. 66 - 81.
16. Информационно-аналитический центр по параллельным вычислениям [Електронний ресурс]. Режим доступу: www.parallel.ru (дата звернення: 10.08.2017).
17. Український національний грід. Базовий координаційний центр [Електронний ресурс]. Режим доступу: www.ung.in.ua (дата звернення: 12.08.2017).

ПАРАЛЛЕЛЬНЫЕ АЛГОРИТМЫ РЕШЕНИЯ ПРОСТРАНСТВЕННОЙ ЗАДАЧИ ЦИФРОВОЙ ФИЛЬТРАЦИИ ДАННЫХ

М. С. Яджак

Институт прикладных проблем механики и математики им. Я. С. Пидстригача НАН Украины, ул. Научная, 3-Б, г. Львов, 79060, Украина; e-mail: yadzhak_ms@ukr.net

В работе предложены параллельные алгоритмы с автономными ветвями для решения пространственной задачи цифровой фильтрации. Приведены оценки ускорения, подтверждающие высокую эффективность этих алгоритмов. Предложенные параллельные алгоритмы ориентированы для реализации на современных широкодоступных вычислительных средствах – многоядерных компьютерах, кластерах и средствах гибридной архитектуры. Полученные в работе результаты могут быть использованы для исследования состояния и качества функционирования сложных динамических систем, в том числе и транспортных.

Ключевые слова: параллельный алгоритм, цифровая фильтрация, сложная система, ограниченный параллелизм, ускорение вычислений, кластер.

PARALLEL ALGORITHMS FOR SOLVING THE SPATIAL PROBLEM OF DATA DIGITAL FILTERING

M. S. Yadzhak

Ya. S. Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, NAS Ukraine, 3B, Naukova Str., Lviv, 79060, Ukraine; e-mail: yadzhak_ms@ukr.net

In the work the parallel algorithms with autonomous branches for solving the spatial of digital filtering problem are proposed. The speed up estimates, which confirm high effectiveness of these algorithms, are pointed. The proposed parallel algorithms are oriented for implementation on modern widely-accessible computation means such as multicore computers, clusters and means with the hybrid architecture. The results of this work can be used for research of the state and operation quality of complex dynamical systems, including transport ones.

Keywords: parallel algorithm, digital filtering, complex system, limited parallelism, speed up of computations, cluster.