

УДК 681.51:681.3

І.Г. Кіріллов, К.В. Брага, В.Ю. Суходольський

Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків

УДОСКОНАЛЕНА МЕТОДИКА СИНТЕЗУ ПАРАЛЕЛЬНИХ ЧАСОВИХ МОДЕЛЕЙ ФУНКЦІОНАЛЬНОГО АЛГОРИТМУ З МІНІМАЛЬНИМ ЧАСОМ ЙОГО РЕАЛІЗАЦІЇ

При розробці апаратно-програмних паралельних обчислювальних засобів АСУ реального часу, виникає необхідність як обґрунтованого вибору чисельних методів вирішення прикладних завдань, відповідних алгоритмів, так і паралельних часових моделей їх реалізації за мінімальний час з мінімальними апаратними витратами. При цьому особливої актуальності набуває пошук методів (методик) оцінювання потенційних можливостей розпаралелювання відповідних алгоритмів та їх практичної реалізації за критерієм мінімуму (або не перевищення допустимого) часу виконання з одночасною мінімізацією апаратних витрат.

В статті розглядається можливість вирішити завдання одночасної мінімізації часу виконання функціонального алгоритму та апаратних витрат при його паралельній реалізації шляхом узагальненої постановки задачі планування організації обчислювального процесу.

Ключові слова: метод мережевого планування та управління, паралельна часова модель алгоритму, пріоритети і ранги операцій, потенційні можливості розпаралелювання алгоритму, резерви часу подій та операцій.

Вступ

Постановка проблеми. Центральною проблемою сучасної обчислювальної техніки, що є важливою складовою розвідувально-управлінських інформаційних систем різного призначення, на теперішній час вважається істотне підвищення ефективності реалізації паралельної обробки даних [1–2]. До числа основних факторів, що визначають можливості підвищення ефективності паралельних обчислювальних систем (ПОС), відноситься застосування прогресивних концепцій, принципів побудови та архітектури ПОС, забезпечення їх відповідності різним вимогам і обмеженням в залежності від призначення.

Необхідною умовою вирішення проблем створення перспективних ПОС слід вважати розробку і вдосконалення математичного апарату формалізованого синтезу та аналізу паралельних цифрових структур і програм з заданими показниками ефективності. Такі показники, наприклад, оперативність, своєчасність, зокрема забезпечуються обґрунтованим вибором чисельних методів реалізації функціональних завдань. Тому при розробці апаратно-програмних паралельних обчислювальних засобів АСУ реального часу, виникає необхідність як обґрунтованого вибору чисельних методів вирішення прикладних завдань, відповідних алгоритмів, так і паралельних часових моделей їх реалізації за мінімальний час з мінімальними апаратними витратами. При цьому особливої актуальності набуває пошук методів (методик) оцінювання потенційних можливостей розпаралелювання відповідних алгоритмів та їх практичної реалізації за критерієм мінімуму (або не перевищення допустимого) часу виконання з одночасною мінімізацією апаратних витрат.

Аналіз літератури. В даний час відомо багато публікацій, присвячених розробці паралельних структур спеціалізованих і паралельних програм реалізації

різних функціональних алгоритмів [3–6]. Методика синтезу часопараметризованих мультипаралельних цифрових апаратних та програмних засобів, що запропонована в [3–4], дозволяє отримати паралельні часові моделі алгоритму (ПЧМА), тобто часові паралельні орієнтовані графи, число яких визначається різноманітністю апаратних компонент (часом виконання елементарних операцій, функцій), поєднанням методів паралельної обробки та сукупністю обмежень. В умовах відсутності останніх реалізується процедура оцінювання потенційних можливостей розпаралелювання обчислень при реалізації функціональних алгоритмів, що базується на використанні методу суміщення незалежних операцій алгоритму (глобального паралелізму). Під «потенційними» зазвичай розуміються граничні (максимальні) можливості розпаралелювання обчислень, що обумовлюються мірою «незалежності» операцій алгоритму (насамперед мірою незалежності їх вхідних операндів або початкових даних), часом їх реалізації апаратними компонентами (процесорами, обчислювальними модулями тощо). Реалізація таких можливостей дозволяє мінімізувати час виконання функціональних алгоритмів (комплексу алгоритмів), тобто підвищити оперативність обробки інформації в системах реального часу.

Проте ця методика не дозволяє одночасно вирішити завдання й мінімізації в цих умовах апаратних витрат. Як показують дослідження, така мінімізація можлива і зводиться до мінімізації числа операторів алгоритму, одночасно розміщених на часових ярусах моделі. Крім того ця методика використовує множинний опис початкових (проміжних) даних та відповідну формалізацію кроків і етапів, який в більшості випадків є надлишковим в порівнянні з матричним описом.

Мета статті. Вирішити завдання одночасної мінімізації часу виконання функціонального алгори-

тму та апаратних витрат при його паралельної реалізації шляхом узагальненої постановки задачі планування організації обчислювального процесу.

Основна частина

До числа загальних методів рішення завдань планування складних проєктів відноситься метод мережевого планування і управління [7], який полягає в побудові (синтезі) мережних моделей (графів процесу), оцінюванні часових параметрів подій і операцій (робіт) обчислювального процесу (плану) з подальшою їх оптимізацією, а також пошуку критичного (підкритичного) шляху (шляхів) на графі. В якості вузлів такого графу зазвичай вибираються події, що полягають у початку або завершенні операцій алгоритму. Часові параметри подій (ранні і пізні терміни їх настання) являються часовими ярусами обчислювального процесу, який описується ПЧМА. Ребрами графа виступають операції алгоритму (в загальному випадку – роботи), що ілюструють інформаційні зв'язки в алгоритмі і задаються як правило часом їх реалізації, якій у свою чергу залежить від вибору апаратних компонент (їх швидкодії).

Для формалізації опису орієнтованого графа алгоритму, що є формою представлення ПЧМА, в відомій методиці використовуються так звані спряжені $S(P_j) = S_j$ і зовнішні $W(P_j) = W_j$ множини $P = \{P_j\}_{j=1}^{|P|}$ операцій P_j з номерами $j \in 1, |P|$, де $|P| = N$ – потужність (число N елементів) множини P .

Згідно визначення спряжена множина $S_j = \{P_\xi\}_{\xi=1}^N$ j -ї операції містить ξ -ті операції (чи їх номери), результати яких є її вхідними операндами. Помітимо, що для переважної більшості елементарних арифметичних (логічних) операцій (складання / віднімання, множення / ділення тощо) $|S_j| = 2$, а для деяких з них (витягання квадратного кореня, знаходження зворотного елементу) і більшості елементарних функцій ($\sin(x)$, $\cos(x)$, $\log_m(x)$ тощо) – $|S_j| = 1$. Зовнішня множина $W_j = \{P_k\}_{k=1}^N$ j -ї операції навпаки містить k -і операції (чи їх номери), що використовують в якості одного з вхідних операндів результат j -ї операції. При цьому число W_j може мати будь-яке, у тому числі і досить велике, значення. Зводячи постановку завдання до загальної термінології формального опису орієнтованого графа довільної інформаційної системи помітимо, що множини S_j і W_j є відповідно множинами лівих (G_j^{-1}) і правих (G_j) інцидентів j -го вузла графа [8].

Помітимо, що матричним аналогом множини $G_j = W_j$ правих інцидентів є слабозаповнена $(N \times N)$ – мірна матриця суміжності $A = \{a_{ij}\}_{i,j=1}^N$ 3

елементами a_{ij} , що набувають значення «1», якщо результат i -ї операції є вхідним операндом j -ї операції і «0» в інших випадках, а множини $G_j^{-1} = S_j$ лівих інцидентів – є матриця A^T («Т» – символ транспонування), тобто матриця A повністю описує інформаційні зв'язки функціонального алгоритму.

Удосконалена методика синтезу паралельних часових моделей функціонального алгоритму з мінімальним часом його реалізації зводиться до наступної послідовності дій.

1. Формування на основі аналізу інформаційних зв'язків алгоритму матриці суміжності A виду [8]:

$$A = \{a_{ij}\}_{i,j=1}^N; (a_{ij} = 0, 1), \quad (1)$$

де $i, j \in 1, N$ – номери операції алгоритму.

При необхідності із матриці A легко можуть бути сформовані відповідні множини правих $G_\ell = W_\ell$ і лівих $G_\ell^{-1} = S_\ell$ інцидентів ℓ -ї ($\ell \in 1, N$) операції алгоритму, які містять зокрема номери операцій. Так множина W_ℓ ($\ell = i$) містить номери j -х ($j \in 1, N$) стовбців матриці A , де є ненульові елементи $a_{ij} = 1$ в i -му ($i \in 1, N$) рядку, а множина S_ℓ ($\ell = j$) – номери i -х рядків A , де є ненульові елементи $a_{ij} = 1$ в j -му стовпцю матриці суміжності A .

2. Побудова (синтез) мережевої моделі (орієнтованого графа) алгоритму, доповненої $j = 0$ -ю та $j = N + 1$ -ю подіями «початку» та «кінця» алгоритму відповідно.

3. Оцінювання «раннього» $t_j^{(np)}$ і «пізнього» $t_j^{(mn)}$ термінів настання [7] j -ї події ($j \in 1, N$), в якості якої виступає початок виконання в обчислювальному модулі (процесорі) j -ї операції алгоритму, резерву R_j часу j -ї події [9]:

$$R_j = t_j^{(mn)} - t_j^{(np)}, \quad (2)$$

а також відшукування критичного (підкритичного) шляху (шляхів) і сумарного часу T_k реалізації операцій «критичного шляху» алгоритму.

Час T_k являється мінімально можливим (граничним) часом $T_{ПА_{\min}}$ паралельної реалізації алгоритму в апаратних компонентах з часом $t_j^{(0)}$ виконання елементарної j -ї операції (функції) алгоритму (час виконання однотипних операцій вважається однаковим).

4. Синтез паралельних часових моделей алгоритму, в якості «граничних» варіантів яких виступають ПЧМА, відповідні «раннім» ($t_j^{(np)}$) і «пізнім» ($t_j^{(mn)}$) термінам початку виконання j -ї ($j \in 1, N$) операції алгоритму. При цьому часовими ярусами, на яких «розміщуються» операції алгоритму являються значення $t_j^{(np)}$ або $t_j^{(mn)}$, що не повторюються. Кількість варіан-

тів ПЧМА $M = \{M(t^{(n)})\}$ ($t^{(n)} = \{t_j^{(n)}\}_{j=1}^N$) визначається діапазоном можливих значень часу $t_j^{(n)}$ ($t_j^{(np)} \leq t_j^{(n)} \leq t_j^{(mn)}$) початку виконання j -ї операції в відповідності з наявними резервами R_j .

Звернемо увагу, що задача синтезу варіантів ПЧМА, які задовольняють умовам $T_k = T_{ПА_{\min}} = \text{const}$, вирішується методом розподілення резервів часу по фронту [9] з попереднім визначенням повного ($R_{ij}^{(n)}$), вільного ($R_{ij}^{(c)}$) і незалежного ($R_{ij}^{(h)}$) резервів часу i -ї операції, результати якої використовує j -та операція алгоритму.

5. Вибір з множини ПЧМА оптимальної моделі за критерієм мінімуму обчислювальних модулів (процесорів), необхідних для реалізації алгоритму за час $T_k = T_{ПА_{\min}} = \text{const}$. Обрана модель може бути використана в якості початкових даних для синтезу структури спеціалізованого процесора, що реалізує функціональний алгоритм за мінімально можливий час $T_{ПА_{\min}}$ з мінімальними апаратними витратами.

Приведемо співвідношення для формалізації основних кроків удосконаленої методики синтезу паралельних часових моделей функціонального алгоритму з одночасною мінімізацією часу його реалізації та апаратних витрат. Визначення «раннього» і «пізнього» термінів [9] початку виконання j -ї ($j \in 1, N$) операції алгоритму, тобто й відповідних часових ярусів ПЧМА, здійснюється за наступними співвідношеннями:

– «раннього» терміну $t_j^{(np)}$ настання події (початку виконання j -ї операції):

$$t_j^{(np)} = \begin{cases} 0; & a_{ij} = 0, i \in 1, N; \\ \max_i(t_i^{(np)} + t_i^{(0)}); & a_{ij} \neq 0, \end{cases} \quad (3)$$

де $t_i^{(0)}$ – час виконання i -ї операції алгоритму, обумовлений швидкодією апаратних компонент;

– «пізнього» терміну $t_j^{(mn)} = t_i^{(mn)}$ настання події (початку виконання i -ї операції):

$$t_i^{(mn)} = \begin{cases} T_{ПА_{\min}} = T_k; & a_{ij} = 0, j \in 1, N; \\ \min_j(t_j^{(mn)} - t_i^{(0)}); & a_{ij} \neq 0, \end{cases} \quad (4)$$

Слід зауважити, що в умовах $t_j^{(mn)} = T_{ПА_{\min}} - r_j$ ($j \in 1, N$), де r_j – так званий «ранг» [3; 5] j -ї операції алгоритму, (4) ідентично співвідношенню для розрахунку рангів, а саме

$$r_j = \begin{cases} t_j^{(0)}; & W_j = 0; \\ \max_{P_{\xi} \in W_j}(r_{\xi} + t_j^{(0)}); & W_j \neq 0. \end{cases} \quad (5)$$

З (5) видно, що «ранги» операцій є їх часовими параметрами, але позбавлені будь-якого фізичного сенсу, що ускладнює методику. Їх розрахунки необхідні для оцінювання пріоритетів b_j операцій алго-

риту, знання яких потрібно в умовах апаратних обмежень. В якості додаткового позитивного ефекту процедури визначення рангів операцій алгоритму відмітимо, що значення максимального рангу дорівнює мінімально можливому (граничному) часу $T_{ПА_{\min}}$ паралельної реалізації алгоритму ($T_{ПА_{\min}} = \max_j(r_j)$) при заданих часових параметрах $t_j^{(0)}$. Ця особливість дозволяє використати (5) для визначення потенційних можливостей розпаралелювання обчислень при реалізації функціонального алгоритму (визначення часу $T_{ПА_{\min}}$) без синтезу його паралельної часової моделі.

Для визначення $T_{ПА_{\min}}$ також можна використовувати співвідношення

$$T_{ПА_{\min}} = \max_j(t_j^{(np)} + t_j^{(0)}) = \max_j(t_j^{(p)}), \quad (6)$$

де $t_j^{(p)} = t_j^{(np)} + t_j^{(0)}$ – «ранній» час завершення j -ї операції алгоритму.

Відшукування $t_j^{(p)}$ незалежно від (3), також можливо по співвідношенню:

$$t_j^{(p)} = \begin{cases} t_j^{(0)}; & a_{ij} = 0, i \in 1, N; \\ \max_i(t_i^{(p)} + t_j^{(0)}); & a_{ij} \neq 0. \end{cases} \quad (7)$$

При необхідності можливо визначити і «пізній» час $t_j^{(n)}$ завершення операцій алгоритму:

$$t_j^{(n)} = t_j^{(mn)} + t_j^{(0)} \quad (j \in 1, N). \quad (8)$$

Знання «ранніх» і «пізніх» термінів початку виконання операцій алгоритму дозволяє розрахувати резерви R_j часу подій початку виконання j -ї операції. Нерівність нулю цих резервів свідчить про можливість синтезу замість однієї – множини ПЧМА, які задовольняють $T_{ПА_{\min}}$. Це дозволяє, наприклад, методом послідовного перебору, визначити модель, яка вимагає мінімум апаратних компонентів. Першим кроком такої мінімізації може виступати порівняння «граничних» ПЧМА, відповідних випадкам «ранніх» і «пізніх» термінів початку виконання операцій алгоритму.

Приведені вище міркування дозволяють удосконалити співвідношення, що наведені в [3; 5], для розрахунку пріоритетів b_j операцій алгоритму, відмовившись від відшукування їх рангів r_j , що не мають явного фізичного сенсу.

Пропонується для визначення b_j використовувати «пізні» терміни $t_j^{(mn)}$ ($j \in 1, N$) початку виконання операцій, що дає можливість додатково врахувати важливі умови, наприклад, надати більш високий пріоритет операціям, які мають менший резерв часу. Тоді уточнений алгоритм визначення пріоритетів b_j , необхідних при синтезі ПЧМА в умовах апаратних обмежень, має вигляд:

$$\left\{ \begin{array}{l} \text{а) } b_j = 1, \text{ при } t_j^{(nn)} = \max_j t_j^{(nn)} \text{ та } t_j^{(nn)} < t_\xi^{(nn)} (\xi \neq j, j, \xi \in 1, N) \text{ (вищий пріоритет);} \\ \text{б) } b_j < b_\xi, \text{ при } t_j^{(nn)} < t_\xi^{(nn)}; \\ \text{в) } b_j < b_\xi, \text{ при } t_j^{(nn)} = t_\xi^{(nn)}, \text{ операція } j \text{ на критичному шляху, } \xi - \text{ні;} \\ \text{г) } b_j < b_\xi, \text{ при } t_j^{(nn)} = t_\xi^{(nn)}, \text{ (однакові умови), } t_j^{(0)} < t_\xi^{(0)}; \\ \text{д) } b_j < b_\xi, \text{ при } t_j^{(nn)} = t_\xi^{(nn)}; t_j^{(0)} = t_\xi^{(0)}; |W_j| > |W_\xi|; (|W_\ell| = \sum_j a_{\ell j}); \\ \text{е) } b_j < b_\xi, \text{ при } t_j^{(nn)} = t_\xi^{(nn)}; t_j^{(0)} = t_\xi^{(0)}; |W_j| = |W_\xi|; R_j < R_\xi. \end{array} \right. \quad (9)$$

Аналіз показав, що використання запропонованої методики синтезу паралельних моделей навіть у випадку обчислювально простих алгоритмів (декілька десятків елементарних операцій), наприклад, для алгоритму розрахунку дальньої та ближньої зони пуску зенітних керованих ракет [1], дозволяє на 20% скоротити апаратні витрати та приблизно на 25% підвищити коефіцієнт завантаження обладнання (процесорів) [3].

Висновки

Запропонована у статті методика може бути покладена в основу формального синтезу проблемно- і архітектурно орієнтованих часопараметризованих мультипаралельних моделей функціональних завдань для програмної або апаратної реалізації, що забезпечують врахування особливостей архітектури паралельних процесорів і мультипроцесорних ЕОМ (супер-ЕОМ) різних класів і розподілених обчислювальних мереж, оптимізацію складу використовуваних методів паралелізму і задовольняють заданій системі вимог і обмежень.

Список літератури

1. Ярош С.П. Теоретичні основи побудови та застосування розвідувально-управлінських інформаційних систем ППО: моногр. / С.П. Ярош; ред. І.О. Кириченко. – Х.: ХУПС, 2012. – 512 с.
2. Шпаковский Г.И. Реализация параллельных вычислений: MPI, OpenMP, кластеры, грид, многоядерные процессоры, графические процессоры, квантовые компьютеры / Г.И. Шпаковский. – Минск: БГУ, 2011. – 176 с.

3. Синтез и анализ параллельных процессов в адаптивных времяпараметризованных вычислительных системах: монография / [Г.А. Поляков, С.И. Шматков, Е.Г. Толстолужская, Д.А. Толстолужский]; Харьк. нац. ун-т им. В.Н. Каразина, Акад. наук прикладной радиоэлектроники (АН ПРЭ). – Х.: ХНУ, 2012. – 670 с.

4. Поляков Г.А. Функциональный синтез параллельных неперестраиваемых специпроцессоров с использованием аппарата структур семантико-числовой спецификации / Г.А. Поляков, В.В. Лысых, В.В. Толстолужская // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. – Белгород, 2012. – № 13-1 (132). – С. 142-149.

5. Кириллов И.Г. Методика оценки потенциальной распараллеливаемости циклических задач / И.Г. Кириллов, Е.Г. Толстолужская // Системы обработки информации / ХУПС. – Х., 2006. – Вып. 5(54). – С. 44-52.

6. Кириллов И.Г. Анализ эффективности мультипараллельных вычислений в системах первичной и вторичной обработки радиолокационной информации / И.Г. Кириллов, А.П. Сахаров, В.В. Лашин // Системы обработки информации. – Х.: ХУПС, 2016. – Вып.2 (139). – С. 27-30.

7. Гвоздева Т.В. Проектирование информационных систем / Т.В. Гвоздева, Б.А. Баллод. – Ростов н/Д: Феникс, 2009. – 508 с.

8. Меньков А.В. Теоретические основы автоматизированного управления / А.В. Меньков, В.А. Острейковский. – М.: Оникс, 2005. – 640 с.

9. Организация управления в военно-технических системах / Ю.П. Пятков, С.А. Войтович, І.О. Романенко, І.О. Борозенець [за ред. Ю.П. Пяткова]. – Х.: ХУПС, 2006. – 239 с.

Надійшла до редколегії 4.10.2016

Рецензент: д-р техн. наук, проф. М.А. Павленко, Харківський національний університет Повітряних Сил ім. І. Кожедуба, Харків.

УСОВЕРШЕНСТВОВАНАЯ МЕТОДИКА СИНТЕЗА ПАРАЛЛЕЛЬНЫХ ВРЕМЕННЫХ МОДЕЛЕЙ ФУНКЦИОНАЛЬНОГО АЛГОРИТМА С МИНИМАЛЬНЫМ ВРЕМЕНЕМ ЕГО РЕАЛИЗАЦИИ

И.Г. Кириллов, Е.В. Брага, В.Ю. Суходольский

В статье обсуждаются особенности совершенствования методики синтеза параллельных часовых моделей функционального алгоритма с минимальным временем его реализации. При этом задание одновременной минимизации времени выполнения алгоритма и аппаратных расходов при его параллельной реализации возможно решить путем обобщенной постановки задачи планирования организации вычислительного процесса. Показано, что использование предложенной методики синтеза параллельных моделей на основе метода сетевого планирования и управления даже в случае вычислительных простых алгоритмов (несколько десятков элементарных операций) позволяет сокращать аппаратные расходы и повышать коэффициент загрузки оборудования (процессоров).

Ключевые слова: метод сетевого планирования и управления, параллельная часовая модель алгоритма, приоритеты и ранги операций, потенциальные возможности распараллеливания алгоритма, резервы времени событий и операций.

IMPROVED METHODS OF SYNTHESIS OF PARALLEL TEMPORAL MODELS OF FUNCTIONAL ALGORITHM WITH MINIMAL TIME FOR ITS REALIZATION

I.G. Kirillov, K.V. Braga, V.Yu. Sukhodolskiy

In the article are discussed the features of improving the method of synthesis of parallel temporal models of functional algorithm with minimal time for its realization. Wherein the task of simultaneous minimization of the algorithm execution time and hardware costs while its parallel implementation can be solved by the generalized formulation of the problem of computing process organization planning. It is shown that using of the proposed method of synthesis of parallel models on the basis of network planning and management even in case of computationally simple algorithms (several tens of elementary operations) can reduce hardware costs and increase the equipment (processors) load index.

Keywords: method of the network planning and management, parallel sentinel model of algorithm, priorities and grades of operations, potential possibilities of распараллеливания of algorithm, float of events and operations times.