

УДК 004.052: 004.414.23

Преимущества, затраты и риски модификации реализаций методов дискретных особенностей с целью оптимизации

В. О. Мищенко, Б. В. Паточкин

Харьковский национальный университет имени В. Н. Каразина, Украина

Методы дискретных особенностей (МДО), группа молодых компьютер ориентированных вычислительных методов математического моделирования. Она не представлена в математических библиотеках и пакетах, но положение изменится, если реализовать её потенциал ускорения вычислений. Недавно были разработаны такие модификации представителя этой группы - метода дискретных токов (МДТ), реализация которых обладают повышенной скоростью выполнения расчётов на ПК в режиме многопоточности за счёт использования архитектурных особенностей ПК. Цель настоящей работы - оценка затрат и рисков создания и сопровождения такой программной реализации. Для этого для конкретной программной системы (ПС) оценивались энергетические метрики, исследована трудность модулей этой системы. Результаты помогут использовать изученную ПС в качестве основы реализации других МДО.

Ключевые слова: *методы дискретных особенностей, время вычислений, параллельные вычисления, персональный компьютер, надёжность, энергетический анализ программ.*

Методи дискретних особливостей (МДО), група молодих комп'ютер орієнтованих обчислювальних методів математичного моделювання. Вона не представлена в математичних бібліотеках і пакетах, але становище зміниться, якщо реалізувати її потенціал прискорення обчислень. Нещодавно були розроблені такі модифікації представника цієї групи - методу дискретних струмів (МДС), реалізація яких має підвищену швидкість виконання розрахунків на ПК в режимі багатопоточності за рахунок використання архітектурних особливостей ПК. Мета цієї роботи - оцінка витрат і ризиків створення і супроводу такої програмної реалізації. Для цього для конкретної програмної системи (ПС) оцінювалися енергетичні метрики, досліджена трудність модулів цієї системи. Результати допоможуть використовувати вивчену ПС в якості основи реалізації інших МДО.

Ключові слова: *методи дискретних особливостей, час обчислень, паралельні обчислення, персональний комп'ютер, надійність, енергетичний аналіз програм.*

The methods of discrete singularities (DSM) is a group of young computer oriented computational methods of mathematical modeling. It is not represented in mathematical libraries and packages, but the situation will change, should the group potential to speed up calculations be realized. Recently, accelerating modifications of software implementation of this group representative - the method of discrete currents (DCM) - have been developed. It had accelerated computing on PCs due to multi-thread mode of calculations developed using some PCs architecture features. The purpose of this work is assessment of the costs and risks of creating and maintaining such software. To come to this, the Energy Metrics of mentioned DCM implementation were estimated and the modules difficulty and system complexity were studied. The results will help to use this software system as a basis for other DSM implementations.

Key words: *discrete singularities methods, computing time, parallel computing, personal computer (PC), reliability, software energy analysis.*

1 Введение

Настоящая работа связана с реализацией методов дискретных особенностей (МДО). Это группа вычислительных методов современного математического моделирования, позволяющих решать на компьютере сложные задачи аэрогидродинамики, электродинамики и теории упругости [1-3]. Недавно с целью ускорения вычислений был разработан подход к модификации этих методов для наиболее полного использования резервов параллельной обработки данных на персональном компьютере (ПК) [4]. Он реализован для двумерных задач метода дискретных токов (МДТ), который дан в [5]. Целью настоящей работы является относительная оценка затрат на эту модификацию и обеспечение её надёжности. Методической основой послужили методы энергетического анализа программ [6]. Актуальность задачи объясняется тем, что выполненная модификация МДТ является прообразом модификации других вычислительных методов дискретных особенностей (МДО), существующих или разрабатываемых с прикладными целями для моделирования волновых и стационарных процессов аэрогидродинамики, электродинамики и других прикладных областей.

2 Обзор данного направления и задачи исследования

Известной реализацией МДО в электродинамике уровня продукта является [7]. Развивались традиционные подходы к распараллеливанию подобных реализаций, эффективные при большом числе параллельно работающих процессоров [8, 9]. Особенностью подхода [4] является комбинация оптимального использования кэш-памяти при многопоточных вычислениях [10] с использованием встроенных векторных (AVX) регистров. Это поднимает коэффициент ускорения вычислений существенно выше числа параллельно работающих ядер процессора ПК.

Благодаря универсальности схемы МДО программа EPolarizationMT, в которой нашли реализацию указанные выше методы ускорения вычислений и которая предназначена для решения задач дифракции E-поляризованных волн на цилиндрических идеально гладких проводящих экранах в 2D постановке, может быть использована в качестве образца для построения ускоренных реализаций других МДО. Мы ответим на вопросы: каковы затраты на модифицированную, «быструю» реализацию в сравнении с обычной реализацией МДО? трудно ли работать с такой модифицированной реализацией? можно ли сосредоточить в отдельных модулях все существенные изменения, связанные с ускорением, и насколько это рискованно? Свои прогнозы мы обоснуем аналогией с приложением EPolarizationMT, оценив его соответствующие характеристики в сравнении с однопоточной версией EPolarizationST. В качестве метрик (методов оценки) будут применены метрики энергетического анализа [6, часть 1]. Для их оценивания необходимо представить EPolarizationMT, написанную на языке C++, в форме схемы ПС (СПС), к которой предъявляются известные требования [6]. Данный язык не обеспечивает однозначного представления, оно обеспечивается UML моделированием системы. Для модулей СПС нужно отыскать N - длину в программных символах (ПСим), η - алфавит (точнее, мощность алфавита) использованных ПСим, а также разделение ПСим на

операторы, помечаемые *or*, и операнды – *od* ($N = N_{or} + N_{od}$, $\eta = \eta_{or} + \eta_{od}$). Это даёт холстедовские объёмы модулей $V = N \log_2 \eta$. Метрика оценки трудности

$$D^{\wedge} = N_2 \eta_1 / 2 \eta_2 \quad (1)$$

(по Холстеду) традиционно использовалась в программной индустрии как показатель понятности исходного текста модуля. Модули можно классифицировать по абсолютным значениям этой метрики [11]. Ещё необходима оценка числа формальных параметров η^* [6], после чего потенциальные объёмы корректно определяются как $V^* = (\eta^* + 2) \log_2 (\eta^* + 2)$. Для групп блоков и модулей эта величина получается суммированием по всем блокам. Добавляя к холстедовскому объёму модуля объёмы интерфейсных модулей, от которых он прямо зависит, и, делая поправки на другие отношения с учётом порядка разработки [6], получим метрику объёма разработки $W \geq V$. Она позволяет прогнозировать ошибки кодирования и оценивать трудность модулей:

$$B = W / V_{cr} \quad (V_{cr} = 3Hd, Hd = 10^3 \text{ bit} \cdot \text{sym}), \quad D = W / V^*, \quad (2)$$

реализуя идеи Холстеда и обобщая их для программ сложной архитектуры. По этой метрике модули имеет смысл ранжировать [11]. Произведение $A = D \cdot V$ даёт метрику работы программирования для модуля, а сумма по модулям распространяет её на всю систему. По данным об архитектуре системы определяется её спецификационная энергия [6] (во многих случаях для групп блоков, имеющих суммарный потенциальный объём V^* , эта энергия $E = (V^*)^3 / \lambda^2$, где λ – уровень, присвоенный языку программирования). Сравнение для ПС работы и энергии предоставляет метрика

$$q = (E - A) / \max(E, A), \quad (3)$$

- «относительное интеллектуальное тепло». Критерий энергетической сбалансированности состоит в том, что в десятичном представлении q первая значащая цифра $\neq 9$ [6, 11]. Устойчивая (от версии к версии) энергетическая сбалансированность свидетельствует о зрелости разрабатываемой системы.

3 Архитектура исследуемой системы

Общая логика компьютерного моделирования с помощью МДО стационарных и волновых процессов с линейной математической моделью вытекает из унифицированной схемы решения таких задач. Необходимы математическая модель на основе граничных уравнений (обычно интегральных, но понимаемых в смысле обобщенных функций в виду особенностей ядра), дискретная модель граничных уравнений, вычислительный метод для расчёта элементов матрицы дискретной модели, метод решения СЛАУ, метод расчёта требуемых прикладникам характеристик или полей. Соответственно этому выглядит на рис. 1 эскиз архитектуры ПС EPolarization, которая реализует МДТ для задач дифракции плоских Е-поляризованных волн на цилиндрических идеально проводящих экранах с любыми сечениями – гладкими несамопересекающимися кривыми. Эта ПС является базой двух приложений: EPolarizationST – однопоточная реализация (контрольная) и EPolarizationMT.

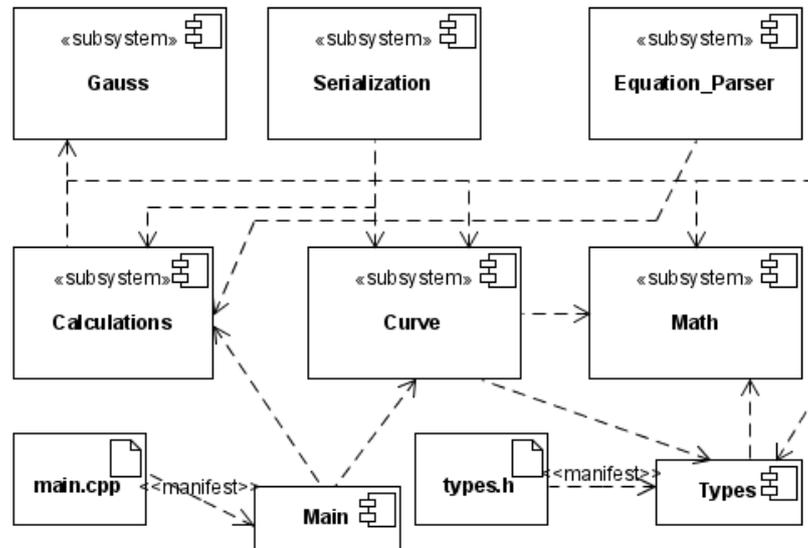


Рис.1 Эскизная схема ПС EPolarization (диаграмма компонент UML)

Раскрывая с помощью других диаграмм компонент архитектуру всех подсистем (пример на рис. 2) мы подойдем к получению схемы программной системы в смысле энергетического анализа. Для оценки СПС используется также следующая информация. Во-первых, характеристики интерфейсов, в частности, тех, на которые указывают отношения зависимости, показанные на диаграммах рис. 1-2 (и ещё на 5-ти, которые здесь не приводим). Во-вторых, внутренняя архитектура всех компонент, включая соответствующие интерфейсы (примеры на рис 3-5), и сведения об использовании операций ввода-вывода в телах интерфейсных модулей (в тех случаях, когда такие операции имеются).

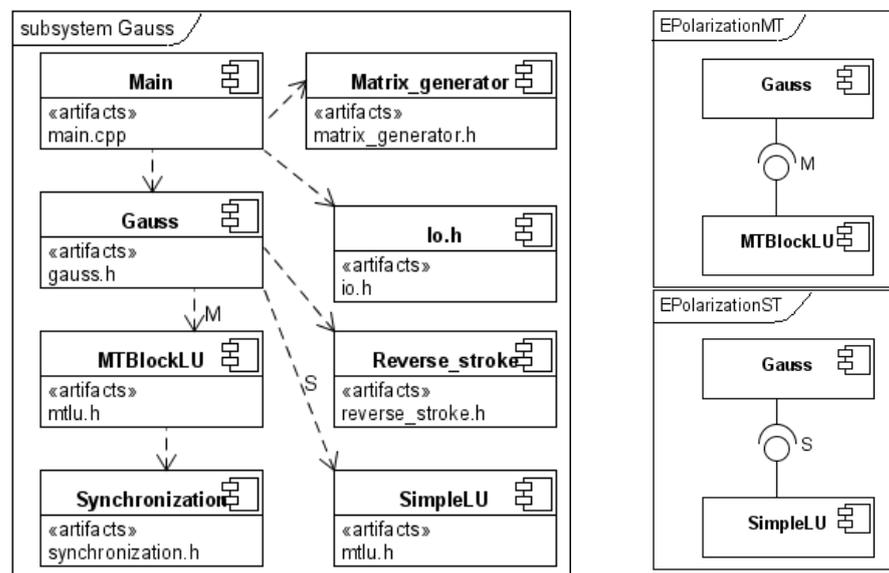


Рис.2 Компонент Gauss в однопоточной версии (EPolarizationST) имеет только зависимость S от SimpleLU, а в многопоточной зависимость M от MTBlockLU

В-третьих, отношение «предшества создания» между компонентами (которое в данном случае совпало с отношением стандартного умолчание для СПС).

На уровне исходных кодов все существенные отличия многопоточного варианта приложения сосредоточены в подсистеме Gauss (рис. 2) и состоят в замене компонента SimpleLU (рис. 3) на компонент MTBlockLU (рис. 4), что повлекло появление дополнительного модуля Synchronization и его тела (рис. 5).

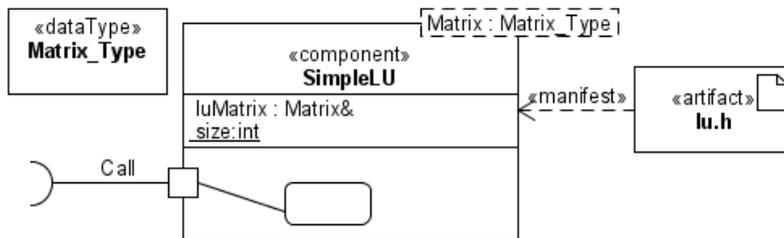


Рис.3 Альтернативная компонента подсистемы Gauss однопоточного приложения

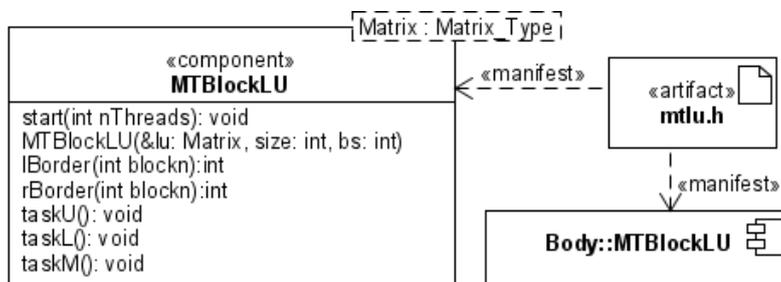


Рис.4 Альтернативная компонента подсистемы Gauss для приложения EPolarizationMT

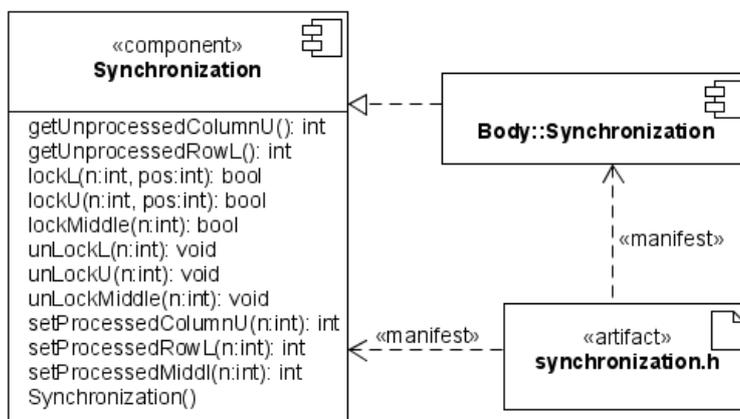


Рис.5 Дополнительные компоненты подсистемы Gauss в многопоточном приложении EPolarizationMT

4 Характеристики реализации рассматриваемой системы

Распределение модулей по классам трудности D^{\wedge} (1) дано в табл. 1. Перечень наиболее трудных модулей по сумме номеров классов трудности D^{\wedge} и D (2) - в табл. 2, результаты энергетического анализа по системе в целом – в табл. 3.

Таблиця 1. Кількість модулів в класах складності D^\wedge прикладження EPolarizationMT

Класи по [7]:	0 клас	1 клас	2 клас	3 клас	4 клас	5 клас
модулів в класі	31	4	4	2	2	2

Таблиця 2. Найбільш складні модулі прикладження EPolarizationMT

Модуль	Підсистема	Кл. D^\wedge	D^\wedge	Клас D	D
json_ui.cpp	Serialization	5	987	4	293
calc_manager.cpp	Calculations	5	435	4	318
grammar.h	Equation_Parser	3	263	5	668

Таблиця 3. Характеристики EPolarizationMT, отримані при енергетичному аналізі

$\sum B_i$	$\max D_0^*$	$\min D_5^*$	$\max D_5^*$	A (kHd)	E (kHd)	q (3)
56	70,8	368	814	12,52	0,983	-0,92

5 Результати енергетичної оцінки, їх аналіз і висновки

Пусть B_i - оцінка потенціального числа помилок в i -м модулі на момент переходу від написання коду до систематичному тестуванню, оцінена по об'єму розробки (2). Ця оцінка служить орієнтиром, реалістичним звичайно по першій значущій цифрі [6]. Суммування по всім модулям, в принципі, законно в силу використання при оцінках саме об'ємів розробки, а не Холстеда. Воно дає орієнтир 56 помилок (табл. 3), а, значить, приблизно стільки ж первинних дефектів цієї розробки. Їх успішне (без внесення суттєвого числа вторинних дефектів) виправлення во багато залежало від складності модулів системи. Використовуючи для висвітлення питання про складність метрику її оцінки М. Холстеда по методу [11], ми на основі табл. 1 можемо утвердити компонентну зрілість системи EPolarizationMT. Дійсно, більше половини (69%) модулів належить до нульового класу складності ($D^\wedge \leq 115$), а в 5-й клас максимальної складності ($D^\wedge > 430$) потрапляє не більше 10% (4,4% на самому справі). Тому велика ймовірність, що за період розробки і налагодки даного програмного засобу наукоємного моделювання (около 2-х років) і його використання в експериментах (около півроку) число залишкових дефектів було зменшено практично до нуля. Відзначимо, що згідно з пропозиціями в [11] нормативам, на обидва модуля, що складають п'ятий клас, вказується, як на складні, «безумовно вимагають додаткового інтенсивного тестування зі сторони розробника» також і D (табл. 2). Таке тестування було проведено. Оцінки по метриці складності D , взагалі, в цілому свідчать про помірну складність системи, прогножують, в частині, низьку ймовірність внесення помилок при модифікаціях. Во-перше, верхня межа 0* класу, до якої належить половина модулів з найменшим оціненим значенням D , це всього лише ≈ 71 (проти орієнтира 115 – верхньої межі 0 класу по D^\wedge). Во-друге, з 4-х модулів класу 5*, до якого слід відносити 10-ю частину найбільш складних в сенсі D модулів, три належать до нульового класу D^\wedge (они мають відносно слабо проработану архітектуру і, тим самим, менш прозорі інших, але, тим не менше, невеликі за розміром). Нарешті, хоча система в цілому виглядає несбалансованою, $q = -0,92$, вона досягне зрілості в наступній версії («лишні» 0,02-0,03 зберуться за рахунок декомпозиції 1-2 модулів).

Относительный уровень затрат на разработку, который характеризует метрика A , в данном случае (табл. 3), для реализации, форсированной по скорости выполнения нашими методами, на порядок выше обычного для реализаций МДО, основанных на одномерных граничных уравнениях (напр., в [12] $A = 1,2 \text{ kNd}$). Однако он вдвое ниже, чем для традиционных реализаций МДО на базе двумерных граничных уравнений (напр., для основных подсистем в [13] суммарно $A = 26,4 \text{ kNd}$). Какую долю в 10 кратном возрастании играет собственно модификация, необходимая для ускорения? Во-первых, нужно иметь в виду, что и EPolarizationMT, и EPolarizationST включают нетривиальный интерфейс для ввода данных об экранах произвольной формы, тогда, как в [12] общий интерфейс отсутствует, а в данных из [13] он не учтён. Реализация собственно МДТ у нас поддерживается модулями с суммарной работой программирования $A_0 = 5,1 \text{ kNd}$. Модификация, обеспечивающая ускорение, сводится (рис. 3-5) к замене 1 базового компонента (с объёмом разработки 0,68 Nd) на 3 других (суммарный объём 18,9 Nd). Прирост работы программирования $\Delta A = 3,47$, и

$$\Delta A/A_0 = 0,68. \quad (4)$$

Следовательно, подобная модификация основных подсистем реализации МДО должна вызывать 5-8 дополнительных ошибок кодирования (около 9% от общего числа для всей системы) и дополнительные затраты времени на разработку, которые составят ориентировочно 70% от времени реализации непосредственно МДО. При этом для дополнительных модулей СПС характерно попадание в 4-е классы трудности по обоим метрикам трудности.

6 Заключение

В статье на основе опыта разработки зрелой, как показывают метрические оценки, реализации ускоренного варианта МДО обоснована гипотеза об умеренных затратах на такие реализации и в будущем. Их разработка на основе методов оптимизации работы программы с кэш-памятью и векторными регистрами [4, 10] по затратам требуемой работы ожидается не более трудоёмкой, чем вдвое по сравнению с реализация базовых алгоритмов МДО. Риски, связанные с надёжностью программных реализаций не должны возрастать более, чем на 10%. По мере создания таких реализаций целесообразно было бы исследовать точность полученных априорных оценок.

ЛИТЕРАТУРА

1. Гандель Ю. В. Краевые задачи для уравнений Гельмгольца и их дискретные математические модели. / Ю. В. Гандель // Современная математика. Фундаментальные направления. – М.: РУДН, 2010. – Т. 36. – С. 36–49.
2. Дмитриев В. И. Метод интегральных уравнений в вычислительной электродинамике / Дмитриев В. И., Захаров Е. В. - М.: МАКС Пресс, 2008. - 316 с.
3. Довгий С. О. Метод сингулярних інтегральних рівнянь. Теорія та застосування / Довгий С. О., Ліфанов І. К. - К.: Наукова думка, 2004. - 510 с.

4. Mishchenko V. O. Accelerating the computation of the discrete currents method by modification takes into account the architectural features of a modern PCs / VO Mishchenko, BV Patochkin // Вісник Харківського національного університету серія МІА. – № 1056. – вып. 26. – С. 129–139.
5. Гандель Ю. В. Математические вопросы метода дискретных токов. Обоснование численного метода дискретных особенностей решения двумерных задач дифракции электромагнитных волн. / Ю. В. Гандель, С. В. Еременко, Т. С. Полянская // Учебное пособие. Ч. 2 – Харьков: ХГУ, 1992. – 145 с.
6. Мищенко В. О. CASE–оценка критических программных систем. Том 1. Оценка качества [Монография] / В. О. Мищенко, О. В. Поморова, Т. А. Говорушенко ; под ред. Харченко В. С. – Х : Нац. аэрокосмический ун–т «Харьк. авиац. ин–т», 2012. – 201 с.
7. Davydov A. Opportunities of program EDEM for development of devices of antenna techniques / A. Davydov, E. Zakharov // Antennas. – 2006.- № 10. - P. 52-57.
8. Медведик М. Ю. Параллельный алгоритм расчёта поверхностных токов в электромагнитной задаче дифракции на экране / М. Ю. Медведик, Ю. Г. Смирнов, С. И. Соболев // Вычислительные методы программирования. – 2005. - Т. 6. - С. 99–108.
9. Гахов А. В. Схема параллельной модификации систем компьютерного моделирования, использующих методы дискретных особенностей / А. В. Гахов, В. О. Мищенко // Радиоэлектронные и компьютерные системы. – 2010. – № 6 (47). – С. 143–147.
10. Б. В. Паточкин Минимизация времени компьютерного моделирования задачи дифракции на экранах методом оптимизации локальности памяти / Б. В. Паточкин // Вісник Кременчуцького національного університету імені Михайла Остроградського. – 2014. – Випуск 6 (89). С. 58-64.
11. Мищенко В. О. Метрики трудности в оценке надёжности инструментальных библиотек и фреймворков / В. О. Мищенко // Вісник Харківського національного університету серія МІА. – № 1131. – вып. 25. – С. 126–147.
12. Боровинский А. В. Компьютерное моделирование 3D дифракции в плоскопараллельной среде: опыт реализации МДО методом расширенных программ / А. В. Боровинский, А. В. Гахов, В. О. Мищенко // Вісник Харківського національного університету серія МІА. – № 863. – вып. 12. – С. 21–35.
13. Духопельников С. В. Математические модели для расчета излучения из продольных щелей в волноводе кругового сечения / С. В. Духопельников // Вісник Харківського національного університету серія МІА. – № 661. – вып. 4. – С. 104–113.