



ТЕХНОЛОГИЯ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ В ЗАДАЧАХ АНАЛИЗА И ОПТИМИЗАЦИИ¹

О.В. Абрамов, Я.В. Катуева

Институт автоматики и процессов управления, г. Владивосток

Рассмотрена задача оптимального выбора параметров аналоговых технических устройств и систем по стохастическим критериям. Предложены параллельные алгоритмы многовариантного анализа и оптимизации, ориентированные на компьютерные системы с массовым параллелизмом. Уделено внимание влиянию особенностей архитектуры параллельной вычислительной системы, механизма передачи данных и организации вычислительного процесса на эффективность решения задач.

ВВЕДЕНИЕ

Создание технических систем связано с решением задач синтеза либо анализа. Любой технический объект можно представить в виде двойки $W = \langle X, S \rangle$, где $X = (x_1, x_2, \dots, x_n)$ – параметры объекта, а S – его структура. Критерий качества функционирования объекта можно задать в виде функционала $\Phi = \Phi(W, Q)$, где Q – условия эксплуатации объекта.

Задача синтеза сводится к решению оптимизационной задачи

$$\Phi(W, Q) \rightarrow \operatorname{extr}_{W \in \Omega_W} W_Q^0,$$

где Ω_W – допустимое множество, в рамках которого может варьироваться объект, W_Q^0 – оптимальный в условиях Q объект. Задача синтеза допускает естественную декомпозицию на параметрическую и структурную:

$$\Phi(X, S, Q) \rightarrow \operatorname{extr}_{X \in \Omega_X} \operatorname{extr}_{S \in \Omega_S} X_Q^0 S_Q^0,$$

где Ω_X – множество допустимых параметров, Ω_S – множество допустимых структур.

¹ Работа выполнена в рамках проекта «Разработка методов и алгоритмов параметрического синтеза стохастических систем» программы фундаментальных исследований Отделения ЭММПУ РАН № 14 «Проблемы анализа и синтеза интегрированных технических, социальных и медико-биологических систем управления».

Таким образом, синтез технических объектов включает в себя две основные части: формирование структуры объекта (структурный синтез) и выбор значений внутренних параметров (параметрический синтез) [1].

Разработка объекта начинается с формирования структуры. Для выбранного варианта структуры строится модель объекта (математическая или физическая) и выбираются исходные (начальные) значения параметров, при которых проверяются условия работоспособности. В случае их невыполнения значения внутренних параметров корректируются; при этом естественно попытаться найти такие их значения, при которых условия работоспособности выполняются с наибольшей вероятностью.

Проектирование технических объектов с учетом случайности процессов изменения их параметров связано с необходимостью решения ряда сложных и трудоемких задач. К их числу относится и задача оптимального выбора номинальных значений параметров проектируемых объектов, основные трудности решения которой обусловлены вероятностным характером критерия оптимальности и дефицитом информации о случайных закономерностях процессов изменения параметров проектируемых систем.

Высокая вычислительная трудоемкость решения оптимизационных задач со стохастическим критерием заставляет искать способы достаточно быстрого получения желаемых результатов. В работах [2, 3] предложены методы и алгоритмы, убывающие в известной мере процедуру поиска оп-

тимальных решений в задаче параметрического синтеза.

Сравнительный анализ и эксперименты показывают, что они позволяют сократить время решения задачи в 3...5 раз по сравнению с классическими методами, основанными на прямом методе статистического моделирования.

Вместе с тем, в последние годы активно развивается достаточно радикальный путь сокращения трудоемкости решения сложных вычислительных задач, в основе которого лежит идея распараллеливания процессов поиска конечного результата и который приводит к самым разнообразным вариантам архитектуры высокопроизводительных ЭВМ. В качестве параллельных вычислителей широко применяются открытые системы массового параллелизма, состоящие из стандартных компонентов, в том числе массовых серийных микропроцессоров. Для создания подобных кластерных компьютеров сформировался как рынок аппаратных средств, так и требуемый для распределенной обработки набор программных компонентов, состоящий из некоммерческого свободно распространяемого программного обеспечения. Естественная проблема для эффективной реализации требующих высокой производительности алгоритмов и приложений заключается в том, как наилучшим способом использовать имеющиеся аппаратные средства для конкретной вычислительной задачи.

Решение задач анализа и оптимизации подразумевает отображение всей вычислительной схемы, включая методы многовариантного детерминированного и статистического анализа, а также оптимизации по стохастическим критериям, на параллельную архитектуру ЭВМ, учитывая топологию межпроцессорных связей и обеспечивая правильность взаимодействия множества параллельно выполняющихся независимо друг от друга процессов.

ЗАДАЧА ОПТИМАЛЬНОГО ПАРАМЕТРИЧЕСКОГО СИНТЕЗА

Данная задача состоит в выборе номинальных значений внутренних параметров исследуемого устройства $x_{\text{ном}} = (x_{1\text{ном}}, \dots, x_{n\text{ном}})$, обеспечивающих максимум вероятности его безотказной работы в течение заданного времени:

$$x_{\text{ном}} = \arg \max P\{X(x_{\text{ном}}, t) \in D_x, \forall t \in [0, T]\}, \quad (1)$$

где $X(x_{\text{ном}}, t)$ – случайный процесс изменения параметров; D_x – область работоспособности (область допустимых изменений параметров); T – заданное время эксплуатации устройства.

Область допустимых вариаций внутренних параметров D_x , как правило, неизвестна, поэтому условия работоспособности обычно задаются систе-

мой неравенств $a_j \leq y(x)_j \leq b_j, j = \overline{1, m}$, где $y = \{y_j\}_{j=1}^m$ – вектор выходных параметров устройства, причем $y_j = F_j(x_1, \dots, x_n)$, а $F_j(\cdot)$ – известный оператор, зависящий от топологии исследуемого устройства.

Обычная процедура получения оценок (1) вероятности безотказной работы связана со статистическим моделированием процессов изменения параметров $X(t)$ и вычислением методом Монте-Карло вероятности $P_r(t) = P\{X(t) \in D_x, \forall t \in [0, T]\}$.

Основные трудности при параметрическом синтезе технических систем с учетом случайных вариаций их параметров связаны с высокой вычислительной трудоемкостью решения возникающих при этом задач оптимизации по стохастическим критериям. На каждом шаге оптимизации необходим статистический анализ для получения оценки критерия оптимальности. При этом на основе метода статистических испытаний (Монте-Карло) многократно рассчитывается исследуемая система (устройство) при разных случайных значениях параметров элементов. Число расчетов N , равное числу реализаций случайного вектора параметров, определяется из условия обеспечения необходимой точности оценки критерия. Обычно для обеспечения требуемой точности оценки на каждом шаге поиска необходимо полностью рассчитывать систему от нескольких сотен до тысячи раз.

Традиционно алгоритм параметрического синтеза, основанный на методе прямого статистического моделирования и технологии последовательных вычислений, выглядит следующим образом.

1. Задается исходный вектор номиналов параметров $x_{\text{ном}}^{(1)}$.

2. Исходя из заданных законов распределения параметров x_1, \dots, x_n , генерируется реализация случайного вектора параметров $x^{(k)}$.

3. Для заданной реализации значений параметров вычисляются выходные параметры $y_j = F(x^{(k)})$, $j = \overline{1, m}$.

4. Проверяются условия работоспособности $y \in D_y$, где D_y – известная область допустимых значений выходных параметров y .

Наиболее трудоемкий п. 3, поскольку вычисление выходных параметров связано нередко с решением систем дифференциальных (и не всегда линейных) уравнений.

Выполнение п. 4 позволяет отнести данную реализацию $x^{(k)}$ к числу «хороших» (обеспечивающих работоспособное состояние системы) или «плохих». Этой операцией завершается первый внутренний цикл и происходит возврат к п. 2. Генерируется следующая реализация $x^{(k+1)}$, выпол-



няются пп. 3 и 4 и т. д. Общее число итераций N определяется необходимой точностью оценки вероятности нахождения параметров в области работоспособности $\hat{P} = n_x/N$, где n_x – число «хороших», принадлежащих области работоспособности, реализаций из общего числа N испытаний.

Не останавливаясь на известных процедурах определения необходимого числа испытаний N , отметим только, что обычно $N > 100$. Для вычисления оценки вероятности безотказной работы в течение определенного времени число выполнений п. 2–4 определяется необходимым числом временных сечений случайного процесса $X(t)$, так называемых, t -сечений [1].

АНАЛИЗ ВЫЧИСЛИТЕЛЬНЫХ ОСОБЕННОСТЕЙ КОМПЛЕКСА МВС-1000/16

Прежде чем создавать параллельную программу, необходимо знать общую архитектуру параллельной машины и топологию межпроцессорных связей, которая существенна для программирования. Это связано с отсутствием средств автоматического распараллеливания, которое позволяло бы превращать любую последовательную программу в параллельную и обеспечивало бы ее высокую производительность [4]. Нужно в явном виде увязывать структуру алгоритма решаемой задачи со структурой вычислительной системы и обеспечивать правильность взаимодействия множества параллельно выполняющихся независимых друг от друга процессов.

Предлагаемые параллельные алгоритмы реализованы на имеющейся в Институте автоматизации и

процессов управления ДВО РАН многопроцессорной вычислительной системе МВС-1000/16, характеристики которой представлены в таблице. Она является типичным представителем параллельных систем с массовым параллелизмом (MPP – Massively Parallel Processing). Каждый процессор имеет только свою локальную память и обменивается информацией с соседями по сети. Доступ к памяти других процессоров возможен только с помощью механизма передачи сообщений.

Основное достоинство таких систем – высокая степень масштабируемости. Для достижения необходимой производительности требуется просто собрать систему с нужным числом узлов.

Главные ограничивающие факторы вычислительного процесса, характерные для систем с массовым параллелизмом и разделенной памятью – медленная скорость межпроцессорного обмена и высокая латентность.

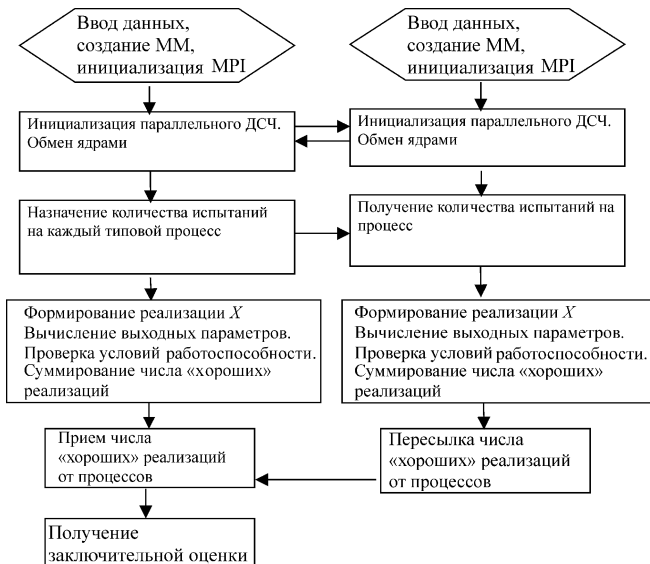
В параллельных системах данного класса особую роль играет закон Амдаля об уменьшении эффективности параллельных программ с ростом доли последовательных вычислений [4]. Кроме того, необходимо следить за общей балансировкой нагрузки процессоров и синхронизировать процессы на различных этапах выполнения алгоритмов.

ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ МЕТОДА СТАТИСТИЧЕСКИХ ИСПЫТАНИЙ

Параллельные вычисления в методе статистических испытаний (метод Монте-Карло) выглядят наиболее логичными, поскольку идея параллелизма – повторения некоторого типового процесса с различными наборами данных – заложена в самой

Характеристики вычислительной системы МВС-1000/16

Архитектура	Система состоит из однородных <i>вычислительных узлов</i> , включающих в себя: <ul style="list-style-type: none"> • один центральный процессор (Intel Pentium III с частотой 800 МГц и кэш-памятью второго уровня 256 кБ); • локальную/оперативную память 512 МБ (прямой доступ к памяти других узлов невозможен); • сетевой адаптер; • жесткий диск HDD 10 ГБ
Межпроцессорный обмен	Связь узлов осуществляется по одной из стандартных сетевых технологий (Fast Ethernet) на базе коммутатора
Латентность (MPI)	170 мкс
Пропускная способность (MPI)	10 МБ/с
Операционная система	На каждом узле работает OS Linux Red Hat 7.2
Модель программирования	Программирование в рамках модели передачи сообщений MPI (Message Passing Interface)
Программное обеспечение	Стандартные средства Linux (GCC, G77), могут быть привязаны специализированные компиляторы (Intel C, Fortran)



Общая схема параллельного метода статистического оценивания

структуре метода. Интуитивно понятно, что применение k независимых процессоров и распределение между ними независимых испытаний уменьшает трудоемкость статистического моделирования в k раз, поскольку затраты на заключительное суммирование и осреднение результатов практически несущественны. Заключительную оценку можно получить по формуле $\hat{P} = \sum_{i=1}^k n_{gi}/N$, где n_{gi} — количество «хороших» реализаций для каждого из процессоров, N — требуемое число испытаний.

Существенное требование при распределении испытаний по процессорам заключается в отсутствии межпроцессорных корреляций при формировании реализаций векторов варьируемых параметров на различных процессорах. Для его выполнения привлекаются специальные библиотеки параллельных датчиков случайных чисел, что позволяет осуществить распределенный подход в получении последовательностей псевдослучайных чисел [5]. Применение таких приемов параллельного программирования позволяет практически избежать накладных расходов на пересылку данных и высоких временных задержек, вызванных латентностью вычислительной системы, а также серьезной разбалансировки вычислительного процесса, возникающей в ходе анализа исследуемого устройства методом критических сечений [1]. Общая схема параллельного метода статистических испытаний приведена на рисунке, где ММ — математическая модель, ДСЧ — датчик случайных чисел.

Как показывают результаты численных экспериментов, применение параллельных датчиков

случайных чисел позволяет получить на этапе статистического оценивания и анализа исследуемых устройств ускорение, близкое к линейному (для 16-ти процессоров среднее ускорение составляет около 15,6 раз).

ПАРАЛЛЕЛЬНЫЕ АЛГОРИТМЫ ДИСКРЕТНОЙ ПОИСКОВОЙ ОПТИМИЗАЦИИ

Универсальным и достаточно эффективным средством решения задачи (1) может стать применение параллельного аналога метода сканирования (слепого поиска).

Сущность метода заключается в том, что вся допустимая область пространства параметров разбивается на элементарные ячейки, в каждой из которых по определенному алгоритму выбирается точка: в центре ячейки, на ребрах или вершинах, случайно. Для каждой ячейки последовательно просматриваются значения целевой функции и находится среди них экстремальное. Точность метода, естественно, определяется тем, насколько плотно располагаются выбранные точки в области поиска.

Основное достоинство метода сканирования — при достаточно густом расположении точек всегда гарантируется отыскание глобального экстремума. Однако требуется значительный объем вычислений, снизить который можно распараллеливанием алгоритма. Наиболее простой алгоритм поиска экстремума методом сканирования (поиска на сетке переменных) заключается в том, что по каждой независимой переменной задаются приращения в соответствующем порядке, обеспечивающем заполнение всей исследуемой области равномерной и достаточно густой сеткой.

Поскольку номинальные значения параметров схемных элементов x_n обычно должны принадлежать ряду стандартных значений, регламентированных техническими условиями или ГОСТ, иногда предпочтительнее искать оптимальный вектор номиналов параметров на дискретном множестве номиналов, соответствующем стандартным значениям и ограниченном областью допустимых значений D_x .

Пусть известен вектор номиналов внутренних параметров $x_n^{BH} \in D_x$. Тогда в каждой точке дискретного множества $D_n^{BH} = \{x_n^{BH}/x_n \in D_x\}$ необходимо найти оценку $\hat{P}_r(x_n^{BH})$. Искомый оптимальный вектор номиналов x_n^{OPT} находим, решая задачу $x_n^{OPT} = \max_{x_n^{BH}} \hat{P}_r(x_n^{BH})$.

В простейшем случае решение задачи сводится к полному перебору элементов множества D_n^{BH} ,



для каждого из которых ищется оценка вероятности $\hat{P}_r(x_H^{BH})$. Формирование множества D_H^{BH} можно осуществить как предварительную процедуру, и его элементы поместить в базу данных исходной информации.

Оценим общую трудоемкость задачи дискретной оптимизации в случае полного перебора элементов множества D_H^{BH} . Если R – число элементов множества D_H^{BH} и для получения оценки методом Монте-Карло для каждого элемента необходимо провести N испытаний, то общее число испытаний равно RN . Соответствующий алгоритм представляет собой двухуровневый распределенный процесс, для реализации которого необходимо RN процессоров. Если $RN > k$, где k – число имеющихся процессоров, то можно представить алгоритм поиска экстремума в виде последовательно-параллельных процедур, число которых определяется количеством свободных процессоров $k_{CB} = [N/k]$, где скобки означают целую часть.

Множество D_H^{BH} разбивается на непересекающиеся подмножества $D_{Hi}^{BH} = \bigcup_{i=1}^k \{D_{Hi}^{BH}\}$, каждому

j -му процессору назначается своё подмножество D_{Hj}^{BH} исходных данных. Таким образом, каждый из процессов последовательно находит оптимальный вектор номиналов для своей подобласти, затем результаты передаются основному процессу и выбирается оптимальный вектор по заданному критерию.

ЛИТЕРАТУРА

1. *Абрамов О.В.* Параметрический синтез стохастических систем с учетом требований надежности. – М.: Наука, 1992.
2. *Abramov O.V., Katueva Y.V., Suponya A.A., Lazarev G.I.* Computer aided design system for optimal parametric synthesis // Proc. of the 3-rd Asian Control Conference (ASCC'2000), Shanghai, 2000. – P. 2047–2050.
3. *Абрамов О.В., Катueva Я.В., Супоня А.А.* Эффективные методы параметрической оптимизации по стохастическим критериям // Тр. междунар. конф. по проблемам управления. – М.: Фонд «Проблемы управления», 1999. – Т. 2. – С. 130–132.
4. *Воеводин В.В., Воеводин В.В.* Параллельные вычисления. – СПб.: БХВ-Петербург, 2002.
5. *Coddington P.* Random number generators for parallel computers // NHSE Review. – 1996. – № 2.

☎(4232) 31-02-02

E-mail: abramov@iacp.dvo.ru, gloria@iacp.dvo.ru □

Новая книга

Трахтенгерц Э.А., Шершаков В.М., Камаев Д.А. Компьютерная поддержка управления ликвидацией последствий радиационного воздействия. – М.: СИНТЕГ, 2004.

В первой части книги рассмотрены назначение и структура компьютерных систем поддержки управления ликвидацией последствий радиационного воздействия. Обсуждены особенности функционирования этого типа систем, в частности, необходимость принятия решений в максимально сжатые сроки при неполной информации, роль субъективных оценок и влияние фактора неопределенности. Вторая часть посвящена математическому и алгоритмическому обеспечению, в ней дано теоретическое и методологическое обоснование методов построения и эксплуатации таких систем. В третьей части дано описание конкретных алгоритмов, интерфейсов и процедур, реализованных в системе поддержки принятия решений по ликвидации последствий радиационного воздействия.

Книга предназначена для руководителей предприятий и организаций, специалистов по разработке компьютерных систем, преподавателей вузов, студентов и аспирантов, а также для широкого круга читателей, интересующихся проблемой принятия управленческих решений.

Э.А. Трахтенгерц – главный научный сотрудник Института проблем управления им. В.А. Трапезникова РАН, д-р техн. наук, заслуженный деятель науки РФ. Им опубликовано 11 монографий, свыше 200 печатных работ, под его руководством защищено свыше 30 диссертаций. Область научных интересов – теория и методы создания и использования компьютерных систем поддержки принятия управленческих решений.

В.М. Шершаков – заместитель генерального директора, начальник Федерального информационно-аналитического центра НПО «Тайфун». Доктор техн. наук, автор свыше 70 научных работ. Занимается созданием и внедрением автоматизированных систем формирования и ведения банков данных, усовершенствованием и развитием систем мониторинга радиационной обстановки на территории нашей страны. Под его руководством разработана компьютерная система поддержки принятия решений управления ликвидацией последствий ядерных аварий, которая в настоящее время функционирует в системе аварийного реагирования Минатома России и ряда других министерств и ведомств.

Д.А. Камаев – начальник лаборатории математических методов Федерального информационно-аналитического центра НПО «Тайфун», канд. физ.-мат. наук. Специалист в области теории динамических систем. Им опубликовано около 60 научных работ. В настоящее время занимается разработкой методов построения систем поддержки принятия решений.