

ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

УДК 519.874

В.С. Власов

ЗАДАЧИ УПОРЯДОЧЕНИЯ И РАСПРЕДЕЛЕНИЯ РЕСУРСОВ ПРИ ИЗГОТОВЛЕНИИ ИЗДЕЛИЙ МИКРОЭЛЕКТРОННОГО ПРОИЗВОДСТВА

Нижегородский государственный университет им. Н.И. Лобачевского

Рассматриваются задачи построения расписаний выполнения комплекса операций при планировании производства микросэлектронных изделий. Для решения поставленных задач предлагаются вычислительные процедуры метода ветвей и границ с использованием эвристических схем нахождения верхних оценок.

Ключевые слова: комбинирование алгоритмов, оптимальное расписание, стохастические и детерминированные алгоритмы

Содержательное описание

Рассматривается задача построения расписания выполнения комплекса операций при планировании производства микросэлектронных изделий [1]. Микросэлектронное производство изготавливает большие интегральные схемы по различным технологиям. Используется позаказная система планирования. Каждый заказ включает в себя наборы партий пластин, из которых изготавливаются интегральные схемы. Каждая партия пластин определяет количество кристаллов, которые поступят на операцию сборки. Технология изготовления кристаллов, определяемых одной партией пластин, до операции резки одинакова. Для каждой партии пластин задан перечень технологических операций, которые выполняются на оборудовании, причем для каждой технологической операции однозначно определено оборудование, на котором эта технологическая операция должна выполняться. Для каждой технологической операции определено время ее выполнения на оборудовании. Время выполнения технологической операции любого изделия одинаково и определяется оборудованием, на котором эта технологическая операция выполняется. Некоторые технологические операции объединяются в группы (групповые операции), которые должны выполняться последовательно и без перерывов. Для операций, являющихся завершающимися для изделий, заданы директивные сроки (директивные операции). Для части технологических операций (операции с пролеживанием) задано время пролеживания – минимально возможный интервал времени до начала выполнения следующей технологической операции. Для некоторых технологических операций (операции с межоперационным временем) задано межоперационное время – максимально возможный интервал времени до начала следующей технологической операции. Требуется построить расписание выполнения комплекса операций таким образом, чтобы выбранная мера эффективности была минимальной. Особенности рассматриваемого технологического процесса производства микросэлектронных изделий позволяют отдельно рассматривать две взаимозависимые задачи – конвейерную задачу, позволяющую оптимизировать функционирование системы от момента запуска партии пластин до стадии резки, и сетевую задачу – оптимизирующую процесс производства от стадии резки до выпуска готовой продукции.

Математические модели

Конвейерная модель

До стадии резки процесс изготовления кристаллов можно представить с помощью модели построения конвейерного расписания [2, 3]. На содержательном уровне конвейерная модель может быть описана следующим образом. На нескольких станках должны пройти обработку несколько операций. Для каждой операции задан порядок ее обработки на станках. Каждый станок одновременно не может выполнять более одной операции, а операции обрабатываются на станках без перерывов. Порядок выполнения операций на каждом станке одинаков.

Пусть i - номер станка, j - номер операции, $T = \|t_{ij}\|$ - $m \times n$ действительная матрица, элемент $t_{ij} \geq 0$ которой определяет время выполнения операции j на станке i , $i = \overline{1, m}$, $j = \overline{1, n}$. Допустимое решение для конвейерной схемы производства определяется матрицей $X = \|x_{ij}\|$ - размерами $m \times n$, элемент которой x_{ij} - определяет момент начала выполнения операции j на станке i , $i = \overline{1, m}$, $j = \overline{1, n}$, для которой выполняются ограничения:

$$x_{ij} \geq x_{i-1j} + t_{i-1j}, \quad i = \overline{2, m}, \quad j = \overline{1, n},$$

(обработка операции на станке может начаться не раньше, чем эта операция завершит обработку на предыдущем станке),

$$x_{ij} \geq x_{ik} + t_{ik}, \quad \text{или} \quad x_{ik} \geq x_{ij} + t_{ij}, \quad i = \overline{1, m}, \quad j = \overline{1, n},$$

(на станке одновременно не может обрабатываться более одной операции),

$$\text{если } x_{ij} \geq x_{ik}, \text{ то } x_{sj} \geq x_{sk}, \quad i = \overline{1, m}, \quad s = \overline{1, m}, \quad j = \overline{1, n},$$

(порядок обработки операций одинаков на всех станках - перестановочность расписания),

$$x_{ij} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n},$$

(естественные условия на переменные).

Сетевая модель

После стадии резки технология изготовления кристаллов в общем случае задается сетевой канонической структурой, которая предполагает (в отличие от конвейерной схемы), что у операции могут быть несколько непосредственно предшествующих и несколько непосредственно следующих операций. Каноничность сетевой структуры означает, что никакая операция не может быть активизирована до тех пор, пока не завершится выполнение всех ей предшествующих операций. Связи между операциями (технология изготовления изделий) задаются ориентированным графом без петель и контуров, элементам которого поставлены в соответствие некоторые характеристики. Пусть $T = \{0, 1, 2, \dots, T_0\}$ - множество тактов планирования; $J = \{1, 2, \dots, n\}$ - множество всех операций, подлежащих выполнению; $I = \{1, 2, \dots, m\}$ - множество различных видов оборудования; $K(j)$ - множество операций, непосредственно предшествующих операции с номером j , $j \in J$. Обозначим через t_i - время выполнения операции на i -м оборудовании, $i \in I$; $\varphi(j)$ - функцию, определенную на множестве J ($j \in J$) со значениями из множества I ($\varphi(j) \in I$), которая задает номер оборудования, на котором должна выполняться соответствующая операция.

Пусть $G(j_s)$ - множество групповых операций, начинающихся с операции j_s , $G(j_s) = \{j_s, j_k, \dots, j_t\}$, $j_s \in J$. Введем множество операций, являющихся начальными для соответствующих им групп - $G = \{\lambda_1, \dots, \lambda_k\}$, $\lambda_1, \dots, \lambda_k \in J$. Обозначим через t_j^{\min} - время пролеживания j -й операции, $j \in J^{\min}$, где J^{\min} - множество операций, для которых определено время пролеживания; t_j^{\max} - межоперационное время j -й операции, $j \in J^{\max}$, где J^{\max} - множество операций, для которых определено межоперационное время. Пусть D_j - директивный срок окончания выполнения операции j , $j \in J^D$, где J^D - множество директивных операций, $J^D \subseteq J$.

В качестве варьируемых параметров математической модели выступают n -мерные целочисленные векторы \vec{x} и \vec{y} , компоненты которых определяют такты начала и окончания выполнения операций.

В качестве ограничений системы выступают условия каноничности, условия непрерывности выполнения операций, условия совместности и ограничения для групповых операций, операций с пролеживанием и операций с межоперационным временем.

Постановка оптимизационных задач

В рамках построенных математических моделей ставятся различные оптимизационные задачи. Постановка задачи определяется двумя факторами: множеством допустимых решений и критерием оптимальности. Множество допустимых решений задается системой ограничений математической модели.

Для конвейерной схемы критерий оптимальности может быть формально представлен в виде функционала $F(X) = \max_{j=1,n} (x_{mj} + t_{mj}) \rightarrow \min$, который определяет минимизацию времени завершения работ от момента запуска партий пластин до стадии резки.

Для сетевой модели рассмотрим два типа оптимизационных задач. Первый тип связан с возможными *нарушениями директивных сроков*. При этом на систему накладываются штрафные санкции. Для каждой директивной операции определим функцию штрафа, связанную с возможным нарушением директивного срока. Тогда группа частных критериев, связанных с возможными нарушениями директивных сроков, определяется частными критериями оптимальности: $f_j(y_j, D_j) \rightarrow \min, j \in J^D$. В качестве обобщенного критерия оптимальности можно рассмотреть аддитивную свертку частных критериев $F^D(\vec{y}) = \sum_{j \in J^D} f_j(y_j, D_j) \rightarrow \min$. Функционал $F^D(\vec{y})$ определяет количество тактов нарушения

сроков выполнения директивных операций. Минимизация функционала отражает стремление своевременно завершать выполнение работ.

Другой тип оптимизационных задач связан с *простоем оборудования*. Определим для каждого оборудования функцию простоя, связанную с неиспользованием оборудования на некотором временном участке. Время занятости оборудования на произвольном интервале $[0, t]$ при $t \in T$, находится как сумма времен выполнения операций, времена начала и окончания которых принадлежат рассматриваемому интервалу, а также времен выполнения операций, начавшихся на интервале $[0, t]$, но не завершённых к моменту времени t . Тогда время занятости i -го оборудования на интервале $[0, t]$ можно рассчитать как:

$$f_i^{\text{занят}}(\vec{x}, \vec{y}) = f_i^{\text{полн}}(\vec{y}) + f_i^{\text{част}}(\vec{x}, \vec{y}),$$

$$\text{где } f_i^{\text{полн}}(\vec{y}) = \sum_{j: y_j \leq t} t_{(j)}, j \in J, (j) = i, i \in I,$$

$$f_i^{\text{част}}(\vec{x}, \vec{y}) = t - x_k, x_k < t < y_k, k \in J, (k) = i, i \in I.$$

Тогда функция, определяющая время простоя i -го оборудования на интервале $[0, t]$ будет определяться как $f_i(\vec{x}, \vec{y}) = t - f_i^{\text{занят}}(\vec{x}, \vec{y}), i \in I$. Определим частные критерии оптимальности, связанные с минимизацией времени простоя конкретного оборудования, как $f_i(\vec{x}, \vec{y}) \rightarrow \min, i \in I$.

В качестве обобщенного критерия оптимальности будем рассматривать аддитивную свертку частных критериев по каждому виду оборудования при условии, что $t = \max_{j \in J} y_j$:

$$\Phi(\vec{x}, \vec{y}) = \sum_{i \in I} f_i(\vec{x}, \vec{y}) \rightarrow \min.$$

Функционал $\Phi(\bar{x}, \bar{y})$ определяет количество тактов суммарного простоя оборудования на интервале от такта начала планирования до такта, соответствующего окончанию выполнения всех запланированных операций. Минимизация функционала отображает стремление максимизировать загрузку оборудования.

Поставленные задачи относятся к классу NP-трудных, и для их решения в работе предлагается использовать вычислительные процедуры метода ветвей и границ с определением оценок на основе стохастических и детерминированных алгоритмов.

При этом изложенные процедуры метода ветвей и границ применимы для решения как общей задачи построения расписания микроэлектронного производства, так и задачи построения конвейерного расписания. Это позволило произвести декомпозицию решения задачи на построение решения конвейерной задачи до стадии резки и решение сетевой задачи с меньшим числом операций для следующих этапов производства.

Алгоритм решения задачи методом ветвей и границ

Различают четыре основные вычислительные процедуры метода ветвей и границ, которые делятся на индивидуальные, зависящие от специфики задачи, и универсальные, которые являются общими для любых решаемых задач. К индивидуальным процедурам относятся процедура оценок (в общем случае нахождение верхней и нижней оценок) и процедура ветвления. К универсальным процедурам относятся процедура отсева (отбрасывания неперспективных направлений) и процедура останова (определение оптимальности найденного решения).

Индивидуальные процедуры метода ветвей и границ

Процедура оценок включает в себя определение верхней (достижимой) оценки V и нижней оценки H . Рассмотрим перестановку $\rho = (\rho_1, \rho_2, \dots, \rho_n)$, $\rho_j \in J$, $j = \overline{1, n}$. Если перестановка ρ соответствует условиям математической модели, то такую перестановку будем называть “допустимой”. Очевидно, что допустимая перестановка однозначно определяет решение системы, а тем самым соответствует некоторому расписанию, для которого можно найти значение критерия решаемой задачи, которое и определяет верхнюю (достижимую) оценку. В качестве верхней оценки выбирается минимальное значение критерия для допустимых перестановок, сгенерированных различными стохастическими и детерминированными алгоритмами.

Нахождение значений нижних оценок определяется с учетом, как длительностей выполнения операций, так и каноничности сетевой модели.

Процедура ветвления рассматривает все допустимые варианты построения допустимых перестановок - на каждом шаге ветвления выбор осуществляется только из тех операций, непосредственно предшествующие для которых уже выполнены.

Универсальные процедуры метода ветвей и границ

Процедура отсева предполагает, что если значение верхней (достижимой) оценки в одной из вершин дерева ветвлений не больше значения нижней оценки в другой вершине, то вторая вершина исключается из рассмотрения не в ущерб оптимальности.

Процедура останова определяет окончание процесса вычислений. Если осталась неотброшенной лишь одна вершина, в которой значения оценок совпадают, то найдено оптимальное решение задачи, которое определяется перестановкой, соответствующей верхней (достижимой) оценке.

Основным достоинством метода ветвей и границ является то, что, остановив вычисления в любой момент времени, лучшее значение верхней оценки (рекорд) может быть принято за эвристическое решение задачи. Основным недостатком метода является необходимость определения оценок в каждой вершине дерева ветвления, а при большом числе исходных параметров число вершин становится очень большим, что не позволяет провести рассмотрение всех вершин дерева ветвлений.

Получение совокупности верхних оценок Алгоритм-построитель расписания

Алгоритм-построитель расписания формирует из произвольной перестановки ρ допустимую перестановку с учетом особенностей каноничности системы и специфики микроэлектронного производства (групповые операции, операции с пролеживанием и межоперационным временем). Схематично работу алгоритма можно представить в следующем виде:

$$\rho \rightarrow \rho^{\text{кан}} \rightarrow \rho^{\text{доп}},$$

где ρ – исходная перестановка; $\rho^{\text{кан}}$ – перестановка с учетом сетевой модели; $\rho^{\text{доп}}$ – допустимая перестановка, удовлетворяющая всем ограничениям системы.

Использование алгоритма-построителя расписания позволяет производить формирование перестановок, как для конвейерной задачи, так и общей задачи микроэлектронного производства с использованием одних и тех же разработанных алгоритмов.

Стохастические алгоритмы

Генетический алгоритм. В основу эволюционно-генетических алгоритмов заложена идея наследственности в биологических популяциях. Структура данных алгоритма состоит из набора хромосом. Работа алгоритма продолжается до тех пор, пока не будет выполнен критерий останова. В рассмотренной постановке расписание определяется допустимой перестановкой из n элементов. Представление допустимой перестановки формально является генетическим и задает хромосому в виде упорядоченной последовательности из n генов (n – мерного вектора). Ген, соответствующий j -й компоненте, определяет номер операции, которая будет выполняться j -й по порядку, $j = \overline{1, n}$. Работа генетического алгоритма основана на использовании ряда операторов кроссовера и мутации, которые, используя допустимые генотипы “родительских” особей, сохраняют допустимость генотипов потомков.

Алгоритм simulated annealing. Используется аналогия между процессом нахождения решения задачи и моделью охлаждения термодинамической системы. Считается, что процесс протекает при постоянно понижающейся температуре. На каждой итерации при температуре T система с некоторой вероятностью может перейти из состояния с энергией E_1 в состояние с энергией E_2 . Эта вероятность рассчитывается как $P(\Delta E) = e^{-\Delta E/kT}$, где $\Delta E = E_2 - E_1$ и k – постоянная Больцмана. Работа алгоритма начинается с выбора случайного состояния π_1 , удовлетворяющего ограничениям системы, для которого находится значение критерия E_1 . Применяется оператор перехода из состояния π_1 в состояние π_2 и находится значение E_2 для этого состояния. Если $\Delta E \leq 0$ или $P(\Delta E) > \xi$, где ξ – случайное число, равномерно распределенное в диапазоне $[0, 1]$, то π_2 принимается за текущее состояние системы. После этого происходит ее охлаждение. Условием останова алгоритма является охлаждение системы до заданной температуры.

Алгоритм ant colonies. Идея алгоритма основана на моделировании поведения муравьев, связанного с их способностью быстро находить кратчайший путь от муравейника к источнику пищи и адаптироваться к изменяющимся условиям, находя новый кратчайший путь. При движении муравей метит свой путь феромоном, и эта информация используется другими муравьями для выбора пути. Дойдя до преграды, муравьи с равной вероятностью будут обходить ее справа и слева. Поскольку движение муравьев определяется концентрацией феромона, то следующие муравьи будут предпочитать более насыщенный феромоном путь, продолжая обогащать его феромоном. Моделирование испарения феромона гарантирует нам, что найденное локально оптимальное решение не будет единственным – муравьи будут искать и другие пути. В основе предлагаемого алгоритма лежит n -мерная квадратная матрица $P(t) = \|p_{ij}(t)\|$, определяющая состояние системы на такт работы алгоритма t . Элемент матрицы $p_{ij}(t)$ в такт t определяет вероятность того, что операция с номером i будет входить в искомую перестановку j -й по порядку. При этом в силу канонической особенности задачи,

вероятность выбора операции, для которой не выполнены все предыдущие, приравнивается нулю. В основу стратегии построения перестановки заложены как элементы стохастичности, так и «предыстория» жизни системы, которая отображается матрицей $P(t)$. Каждая допустимая перестановка однозначно определяет расписание, а тем самым и значение критерия задачи, соответствующее этому расписанию.

Детерминированные алгоритмы

Фронтальный алгоритм. Предлагаемый алгоритм основан на идеологии «жадных алгоритмов»: включенная в строящееся расписание на каком-то шаге работы алгоритма операция в дальнейшем из расписания не исключается, причем для ее выполнения используются все доступные к данному моменту времени ресурсы в максимально возможном объеме.

На подготовительном этапе для каждой операции j производится расчет временных характеристик: находятся моменты раннего начала выполнения t_j^{PH} , раннего окончания выполнения t_j^{PK} , позднего начала t_j^{PN} , позднего окончания t_j^{PK} выполнения операций, и резервы времени операций $r_j = t_j^{PN} - t_j^{PH} = t_j^{PK} - t_j^{PK}$, $j \in J$. Пусть t – произвольный такт планирования, $t \in T$. Назовем «фронтом операций» множество $F(t)$ – множество операций, любая из которых может начать выполняться с такта t , $t \in T$. Установим соответствие между фронтом операций и допустимой перестановкой, для чего перейдем от множества $F(t)$ к вектору $\rho(t)$, компоненты которого и будут определять порядок выполнения операций. Определим вектор $\rho(t)$ следующим образом. Если $F(t) = \{j_1, j_2, \dots, j_k\}$, то $\rho(t) = (\rho_1, \rho_2, \dots, \rho_k)$, $\rho_1 \neq \rho_q$, $\rho_1, \rho_q \in F(t)$, где ρ_q – номер операции, выполняемой q -й по порядку. Для чего упорядочим операции из множества $F(t)$ по неубыванию резервов времени. Далее последовательно просматриваем операции построенного вектора и включаем в строящееся расписание те из них, для которых свободно соответствующее оборудование. После этого изменяем значение такта планирования и заново формируем фронт операций, соответствующий этому такту.

Алгоритм, основанный на решении задачи о назначениях. Тот факт, что перестановки из n натуральных чисел можно представить в виде квадратных бистохастических булевых матриц, позволил для определения верхних оценок использовать хорошо разработанный аппарат решения задач о назначениях. Будем предполагать, что с использованием понятия предшества “ \prec ” между операциями, которое задается совокупностью множеств $K(j)$, $j \in J$, на множестве операций определяются линейный порядок π . Пусть $Y = \|y_{ij}\|$ – квадратная матрица порядка n , где $y_{ij} = 1$, если $i \pi j$, и $y_{ij} = 0$ в противном случае, $i = \overline{1, n}$, $j = \overline{1, n}$. Рассмотрим задачу о назначениях с матрицей неизвестных Y . Пусть P – множество различных перестановок из n натуральных чисел, а B – множество различных n -мерных бистохастических булевых матриц. Рассмотрим соответствие $\Gamma = (Q, P, B)$, где $Q = P \times B$. Соответствие Γ является функциональным, инъективным, всюду определенным и сюръективным, т.е. взаимнооднозначным (биективным). Отсюда генерировать различные перестановки можно посредством решения задач о назначениях, для чего необходимо связать значения критерия, зависящие от коэффициентов функционала, со значениями критерия решаемой исходной задачи. Это можно сделать, например, используя в качестве матрицы, определяющей коэффициенты критерия задачи о назначениях, найденную в результате работы алгоритма *ant colonies* матрицу $P(t) = \|p_{ij}(t)\|$. Кроме того, так как у задачи о назначениях в общем случае существует множество оптимальных решений, то имеет смысл рассматривать все оптимальные решения задачи о назначениях, для каждого решения находить соответствующую перестановку, строить по ней допустимую перестановку, и выбирать из построенных допустимых перестановок наилучшую, с точки зрения критерия исходной решаемой задачи.

Метод комбинирования алгоритмов

Так как предложенные алгоритмы используют перестановку (или набор перестановок) в качестве исходных данных, а результатом работы алгоритмов также являются перестановки, то для наиболее эффективного решения задачи предлагается последовательное использование алгоритмов. Перестановка, соответствующая лучшему значению критерия рассматриваемой задачи, полученная одним из алгоритмов, используется как начальная для другого алгоритма. Тем самым с помощью предложенных алгоритмов можно генерировать различные схемы построения допустимых расписаний, а тем самым определять верхнюю (достижимую) оценку в методе ветвей и границ.

Заключение

Таким образом, не смотря на то, что рассматриваемые задачи относятся к классу NP-трудных, и то, что реальные производственные задачи имеют большие размерности, применение метода ветвей и границ с использованием эвристических алгоритмов для нахождения верхних оценок позволяет в реальном времени находить допустимые решения задач планирования. Это связано с тем, что, даже рассмотрев не все вершины дерева ветвления, в процессе решения задачи многократно применяются различные алгоритмы нахождения верхних оценок. А так как верхние оценки для рассматриваемого метода ветвей и границ являются достижимыми, то они определяют допустимые решения исходной задачи, лучшее решение (с точки зрения критерия) может быть принято за эвристическое решение поставленной задачи.

Полученные в работе результаты легли в основу созданных диалоговых программных средств решения задач распределения ресурсов при планировании процесса изготовления изделий микроэлектронного производства. Проведенный вычислительный эксперимент показал высокую степень качества разработанных алгоритмов оптимизации. Так, для конвейерной задачи для тестовых примеров (число операций до 1000) отклонения значений критериев на найденных решениях относительно значений нижних оценок не превысили 6%.

Библиографический список

1. Прилуцкий, М.Х. Оптимизационные задачи распределения ресурсов при планировании производства микроэлектронных изделий / М.Х. Прилуцкий, В.С. Власов // Системы управления и информационные технологии. 2009. №1(35). С. 38–43.
2. Прилуцкий, М.Х. Метод ветвей и границ с эвристическими оценками для конвейерной задачи теории расписаний / М.Х. Прилуцкий, В.С. Власов // Вестник Нижегородского государственного университета. Математическое моделирование и оптимальное управление / Изд-во ННГУ. Нижний Новгород. 2008. Вып. 3. С. 147–153.
3. Прилуцкий, М.Х. Метод комбинирования эвристических алгоритмов для конвейерных задач теории расписаний / М.Х. Прилуцкий, В.С. Власов // Исследовано в России. 2007. С. 901–905. <http://zhurnal.ape.relarn.ru/articles/2007/086.pdf>

*Дата поступления
в редакцию 02.02.2010*

V.S. Vlasov

SEQUENCING AND ALLOCATION PROBLEM AT MICROELECTRONIC PRODUCTION

Scheduling problem of complex operation at microelectronic production is considered. Branch and bound method with heuristic schemes for finding the top estimations are offered.

Key words: algorithms combination, the optimum schedule, stochastic and deterministic algorithms