

но большом количестве векторов (в данном случае – это миллионы векторов). В то же время, при малом количестве векторов из-за накладных рас-

ходов на пересылку данных между узлами видно значительное отставание от алгоритмов, выполнявшихся на одной машине.

СПИСОК ЛИТЕРАТУРЫ

1. Quinlan J. Ross. C4.5: Programs for Machine learning. Morgan Kaufmann Publishers, 1993.
2. Интеллектуальный анализ данных в распределенных системах / М. С. Куприянов, И. И. Холод, З. А. Каршиев, И. А. Голубев. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2012.
3. White T. Hadoop: The Definitive Guide. USA: O'Reilly Media, 2009.

I. V. Petukhov

DISTRIBUTED EXECUTION OF DECISION TREE BUILDING ALGORITHMS BASED ON A DATA MINING LIBRARY WITH MAP-REDUCE

This article describes the method of paralleling decision tree building algorithms based on a data mining library that uses units structure for algorithm building and execution. Also it describes the method of interconnection of a library and a map-reduce computing cluster. Finally it describes results of execution of that algorithm with different environments.

Decision trees, parallel algorithms, distributed computations

УДК 681.32

А. В. Бессонов, К. А. Кноп, Ю. Т. Лячек, Ю. И. Попов

Определение минимальной ширины канала между парой компонентов при топологической трассировке

Предложен метод расчета минимальной ширины канала между парой многоконтактных компонентов при трассировке соединений в произвольных направлениях.

Печатный монтаж, топологическая трассировка, размещение компонентов

Гибкая топологическая трассировка в произвольных направлениях [1] имеет целый ряд преимуществ по сравнению с традиционной (под 90 и 45°).

Уже только отказ от преимущественных направлений трассировки позволяет [2]:

- уменьшить суммарную длину проводников;
- сократить площадь, занимаемую проводниками;
- понизить уровень перекрестных электромагнитных помех как за счет уменьшения длины проводников, так и за счет снижения уровня их параллельности;
- уменьшить риск рассогласования задержек в группе сигналов или в дифференциальном сигнале, обусловленный неоднородностью материала печатной платы;
- снизить риск коробления платы при воздействии тепловых нагрузок.

При гибкой трассировке фиксируется только относительное расположение проводников, межслойные переходы с помощью специальной процедуры перемещаются в оптимальные положения, а форма проводников вычисляется автоматически уже по окончании трассировки.

Автоматический расчет оптимальной формы проводников и автоматическая подвижка объектов (межслойных переходов, точек ветвления проводников и, при необходимости, компонентов) позволяют решать трудноразрешимые оптимизационные задачи, например построение деревьев Штейнера, сжатие топологического рисунка.

Задача выравнивания задержек сигналов при нефиксированном положении проводников значительно упрощается. Главное, что ее можно решать не последовательно, а параллельно [3].

Однако при этом возникает ряд задач, не свойственных традиционному подходу, таких, например, как вычисление формы проводника и определение ширины канала между компонентами. Первая из указанных задач рассматривается в работе [4], а данная статья посвящена второй задаче.

Даны 2 параллельных друг другу многоконтактных компонента, все контакты которых расположены на равных расстояниях друг от друга. Требуется расположить их на минимальном расстоянии, позволяющем произвести «поконтактную» разводку проводников между ними с учетом заданной ширины проводника d_1 и минимального зазора d_2 .

В случае ортогональной разводки (рис. 1, а) минимальное расстояние подсчитывается легко: если число контактов равно $2N - 1$, то минимальное расстояние равно сумме N ширин проводников и N зазоров между ними. При трассировке в произвольных направлениях (рис. 1, б) с дугами все оказывается не так просто.

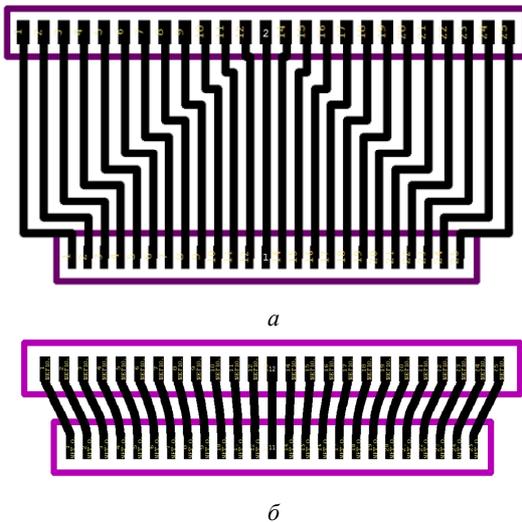


Рис. 1

Введем обозначения:

- h – расстояние между компонентами (считаем их горизонтальными);
- $(0, 0)$ – координаты самой левой точки O нижнего компонента;
- P_k – правая точка того контакта верхнего компонента, для которого отрезок OP_k пересекает ровно k проложенных проводников (это k -й слева контакт). Координаты этой точки равны $(A + kB, h)$, где коэффициенты A и B ищутся из системы уравнений, которые можно получить, зная координаты правых точек любых двух контактов верхнего компонента;

- L_k – левая точка того же контакта;
- d_1 – заданная ширина проводника;
- d_2 – значение минимально допустимого зазора между проводниками.

Вычисление A и B для стандартной конфигурации компонентов. Стандартной будем называть такую конфигурацию, в которой:

- 1) количество контактов каждого компонента нечетно (равно $2N - 1$), при этом центральные (N -е) контакты находятся точно друг под другом;
- 2) расстояния между центрами двух соседних контактов нижнего компонента минимальны, т. е. равны $d_1 + d_2$, а расстояния между центрами соседних контактов верхнего компонента равны $d_3 + d_4$, где d_3 – ширина контактов для верхнего компонента, а d_4 – зазор между соседними контактами верхнего компонента.

При этих предположениях расстояние между точками P_k и P_{k+1} равно $d_3 + d_4$, откуда

$$B = d_3 + d_4. \tag{1}$$

Абсцисса точки L_k равна $A - d_3 + kB$, а для обеих средних точек центральных (N -х) контактов абсцисса равна $(d_1 + d_2)(N - 1) + 0.5d_1$, откуда получаем второе уравнение системы:

$$A + NB - 0.5d_3 = (d_1 + d_2)(N - 1) + 0.5d_1.$$

После подстановки B из (1) вычисляем A :

$$A = 0.5(d_1 - d_3) - d_4 - (N - 1)(d_3 + d_4 - d_1 - d_2). \tag{2}$$

Условие допустимости разводки (выполнение конструктивно-технологических ограничений):

$$|OP_k| \geq d_1k + d_2(k - 1) \text{ для любого } k, \tag{3}$$

$$|OL_k| \geq d_1(k - 1) + d_2(k - 1) \text{ для любого } k. \tag{4}$$

Смысл этих неравенств в том, что длины отрезков OP_k и OL_k , пересекающих несколько проводников, должны быть больше, чем суммарная ширина всех этих проводников и минимальных зазоров между ними. Если условия (3) и (4) выполнены, то аналогичные неравенства выполнены и для любого отрезка между концами каких-либо контактов нижнего и верхнего компонентов (поскольку $d_3 \geq d_1$ и $d_4 \geq d_2$), а это является необходимым и достаточным условием допустимости прокладки проводников.

Отрезки OL_k или OP_k назовем критическими, если для данного k в (3) или в (4) соответственно достигается равенство. Если критических отрезков нет, можно немного уменьшить h , при этом левые части неравенств (3) и (4) уменьшатся (для каж-

дого k), а правые не изменятся. Поэтому будем считать, что критичный отрезок существует и идет от O к правой или левой точке с номером j .

Сначала выпишем условия для правых точек:

$$|OP_j| = d_1j + d_2(j-1), \quad (5)$$

$$|OP_{j+1}| \geq d_1(j+1) + d_2j, \quad (6)$$

$$|OP_{j-1}| \geq d_1(j-1) + d_2(j-2). \quad (7)$$

Нетрудно убедиться в том, что одновременное выполнение равенств в (5)–(7) невозможно, поэтому неравенство (6) положим строгим, т. е. будем считать, что j – номер *последнего* критичного отрезка к правой точке. Вычитая из (6) равенство (5), а из (5) (7), получим:

$$\begin{aligned} & |OP_{j+1}|^2 - |OP_j|^2 > \\ & > (d_1 + d_2)[d_1(2j+1) + d_2(2j-1)] \end{aligned} \quad (8)$$

и

$$\begin{aligned} & |OP_j|^2 - |OP_{j-1}|^2 \leq \\ & \leq (d_1 + d_2)[d_1(2j-1) + d_2(2j-3)]. \end{aligned} \quad (9)$$

Выразим координаты левых частей (8) и (9) по теореме Пифагора:

$$|OP_{j+1}|^2 - |OP_j|^2 = (A + B + jB)^2 + h^2 - (A + jB)^2 - h^2 = B[2A + (2j+1)B], \quad (10)$$

$$|OP_j|^2 - |OP_{j-1}|^2 = (A + jB)^2 + h^2 - (A - B + jB)^2 - h^2 = B[2A + (2j-1)B]. \quad (11)$$

Подставив (10) и (11) в (8) и (9) соответственно, получим после преобразований:

$$2j-1 \leq \frac{-2AB - 2d_2(d_1 + d_2)}{B^2 - (d_1 + d_2)^2} < 2j+1, \quad (12)$$

$$2j \leq \frac{B(B-2A) - (3d_2 + d_1)(d_1 + d_2)}{B^2 - (d_1 + d_2)^2} < 2j+2, \quad (13)$$

$$j = \left\lceil 0.5 \frac{B(B-2A) - (3d_2 + d_1)(d_1 + d_2)}{B^2 - (d_1 + d_2)^2} \right\rceil. \quad (14)$$

Таким образом, получили явную формулу для номера j последнего критичного отрезка к правой

точке. Теперь из (5) можно найти минимальное значение h :

$$(A + Bj)^2 + h^2 = (d_1j + d_2(j-1))^2, \quad (15)$$

$$h = \sqrt{[d_1j + d_2(j-1)]^2 - (A + Bj)^2}, \quad (16)$$

где значение j найдено по формуле (14), а A и B – по формулам (2) и (1).

Аналогичный подсчет критичных значений для левых точек дает такие результаты:

$$j = \left\lceil 0.5 \frac{B(B-2A-2d_3) - 3(d_1 + d_2)^2}{B^2 - (d_1 + d_2)^2} \right\rceil, \quad (17)$$

$$h = \sqrt{[(d_1 + d_2)(j-1)]^2 - (A - d_3 + Bj)^2}, \quad (18)$$

где значение j найдено по формуле (17), а A и B – по формулам (2) и (1).

Из двух значений h , получаемых по формулам (16) и (18), следует выбирать наибольшее, потому что в противном случае какие-то из условий допустимости разводки (3), (4) будут нарушаться.

Вычислительный эксперимент. Для компонентов на рис. 1: $d_1 = d_2 = d_4 = 0.2$ мм, $d_3 = 0.25$ мм, $2N - 1 = 25$, откуда $N = 13$.

В случае ортогональной разводки будет 12 горизонтальных участков различной длины, поэтому расстояние между компонентами должно быть таким, чтобы уместить 12 проводников и 12 зазоров между ними. Отсюда $h_{\min} = 4.8$ мм.

В случае разводки в произвольных направлениях последовательно находим:

$$B = 0.45, A = -0.825, A - d_3 = -1.075.$$

Для правых точек: $j = [7.353] = 7$, $h^2 = 6.76 - 5.405625 = 1.354375$, откуда $h \approx 1.17$ мм.

Для левых точек: $j = [8.11] = 8$, $h^2 = 7.84 - 6.375625 = 1.464375$, откуда $h \approx 1.2101$ мм.

Итого $h_{\min} = 1.2101$ мм.

В случае разводки в произвольных направлениях расстояние между параллельными 25-контактными компонентами можно уменьшить почти в 4 раза.

СПИСОК ЛИТЕРАТУРЫ

1. Модели и алгоритмы автоматизированного проектирования радиоэлектронной аппаратуры / С. Ю. Лузин, Ю. Т. Лячек, Г. С. Петросян, О. Б. Полубасов // СПб.: БХВ-Петербург, 2010.

2. Лузин С. Ю., Попов С. И., Попов Ю. И. Гибкая топологическая трассировка в произвольных направлениях. Особенности и преимущества // Электроника: НТБ. 2013. Вып. 1. С. 96–104.

3. Лысенко А. А., Лячек Ю. Т., Полубасов О. Б. Автоматическое формирование линий задержки в топологии печатного монтажа // Изв. СПбГЭТУ "ЛЭТИ". 2011. Вып. 9. С. 61–65.

4. Попов Ю. И., Попов С. И. Вычисление минимального по длине пути проводника в топологической трассировке печатного монтажа // Науч.-техн. вестн. информационных технологий, механики и оптики. 2012. № 6 (82). С. 117–122.

A. V. Bessonov, K. A. Knop, Yu. T. Lyachek, Yu. I. Popov

THE DETERMINATION OF MINIMUM DISTANCE BETWEEN TWO COMPONENTS IN THE TOPOLOGICAL ROUTING

A method is proposed for calculating the minimum width of a channel between a pair of multipin components in any-angle routing.

PCB, topological routing, components placement

УДК 378.1:167

Ю. В. Воронов, Е. А. Законников, Д. Ю. Куликов, В. А. Мейев

Система мониторинга научного и инновационного потенциала Санкт-Петербурга

Рассматриваются особенности подхода к разработке системы мониторинга научного и инновационного потенциала Санкт-Петербурга, выбора системы показателей, организационное и программное обеспечение мониторинга научных учреждений государственных академий наук, высших учебных заведений и отраслевых организаций различных форм собственности.

Система мониторинга, научная и инновационная инфраструктура, программное обеспечение

Основной целью создания системы мониторинга научного и инновационного потенциала Санкт-Петербурга является обеспечение правительства города эффективным средством информационной и аналитической поддержки стратегического и оперативного планирования в сфере науки для повышения эффективности управления потенциалом одного из ведущих научных и производственных центров страны.

Особенностью системы мониторинга научного и инновационного потенциала организаций и учреждений Санкт-Петербурга является необходимость ввода пользователями значительных объемов числовой и текстовой информации в достаточно компактной форме и увязывания вводимых отдельных параметров с разными отчетными формами, представляемыми в государственные органы статистического наблюдения.

Материалы о научном и инновационном потенциале Санкт-Петербурга агрегированы в следующие разделы: основные сведения об органи-

зации; кадровый потенциал; подготовка специалистов и кадров высшей квалификации; финансирование научных исследований и разработок; результативность научных исследований и разработок; инновационная активность организации; экспортный потенциал.

В первом разделе описывается тип организации, приоритетное направление деятельности, научная и инновационная инфраструктура, материально-техническая база.

Второй раздел отражает кадровый состав организации, в том числе сотрудников с учеными степенями, а также разделение кадров по отраслям наук.

Третий раздел охватывает подготовку кадров (студентов, аспирантов и докторантов) и наличие диссертационных советов.

В четвертом разделе собираются сведения об источниках финансирования научных исследований и разработок (федеральный бюджет, средства коммерческих организаций, зарубежные источники,