行政院國家科學委員會專題研究計畫 期中進度報告

繪圖演算法及資訊視覺化(2/3)
期中進度報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 95-2221-E-002-074-
執 行 期 間 ：95 年 08 月 01 日至 96 年 07 月 31 日
執 行 單 位 ：國立臺灣大學電機工程學系暨研究所

計畫主持人：顏嗣鈞

報 告 附 件 ：出席國際會議研究心得報告及發表論文

處 理 方 式 ：期中報告不提供公開查詢

中 華 民 國 96 年 06 月 02 日

行政院國家科學委員會專題研究計劃成果報告

繪圖演算法及資訊視覺化(2/3)

計畫編號: 95-2221-E-002-074-
執行期間:95 年 8 月 1 日至 96 年 7 月 31 日

計畫主持人: 顏嗣鈞 教授 國立台灣大學電機工程系

一、摘要

本年度之研究計畫中，我們探討*多一邊界標示 (many-to-one boundary labeling)*之架構。在這個架構中，地圖中每一點被用與座標軸平行之多邊形線段連接到置於地圖四周的文字標示(text label)，且允許可能超過一個點連接到一共同的標示。在這種情況下，線段的交叉與總長度的最小化是使多對一邊界標示看起來較美觀的準則。因此，本研究計畫著重在處理如何最適化這些準則的各種議題。

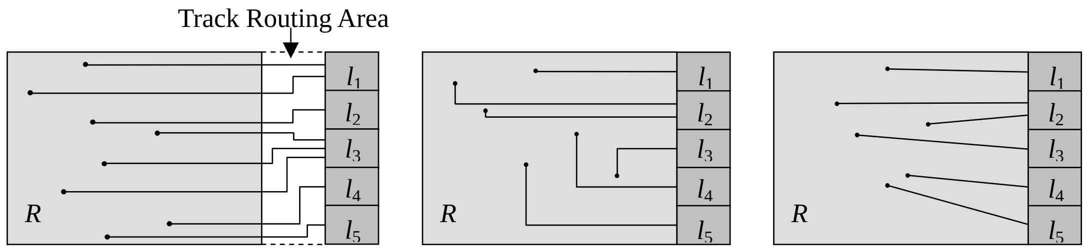
關鍵字：地圖標示，邊界標示

In this project, we investigate the framework of *many-to-one boundary labeling*, where every site in a map is connected to a text label placed on the boundary of map by an axis-parallel polygonal line segment, called *leader*; more than one site may be connected to a common label. In this case, the minimization of the number of line crossings and the total length of leaders is the main aesthetic criteria of many-to-one boundary labeling. This project focuses on coping with various issues in optimizing these aesthetic criteria.

Keyword: map labeling, boundary labeling

二、計畫緣由與目的

In *boundary labeling* [1][2][3], each point site is uniquely connected to a label placed on the boundary of an enclosing rectangle by a *leader*, which may be a rectilinear or straight line segment. To our knowledge, all the results reported in the literature for boundary labeling deal with the so-called one-to-one boundary labeling, i.e., different sites are labelled differently. In certain applications of boundary labeling, however, more than one site may be required to be connected to a common label. In this case, the presence of crossings among leaders often becomes inevitable. Minimizing the total number of crossings in boundary labeling becomes a critical design issue as crossing is often regarded as the main source of confusion in visualization. In this project, we consider the crossing minimization problem for multi-site-to-one-label boundary labeling, i.e., finding the placements of labels and leaders such that the total number of crossings among leaders is minimized. We show the crossing



(a) Type-*opo* leaders (b) Type-*po* leaders (c) Type-*s* leaders

Figure 1: Illustration of leaders

minimization problem to be NP-complete under certain one-side and two-side labeling schemes. Subsequently, approximation algorithms are derived for the above intractable problems. We also present an $O(n^2 \log^3 n)$-time algorithm for the problem of minimizing the total leader length for multi-site-to-one-label boundary labeling, where n is the number of labels.

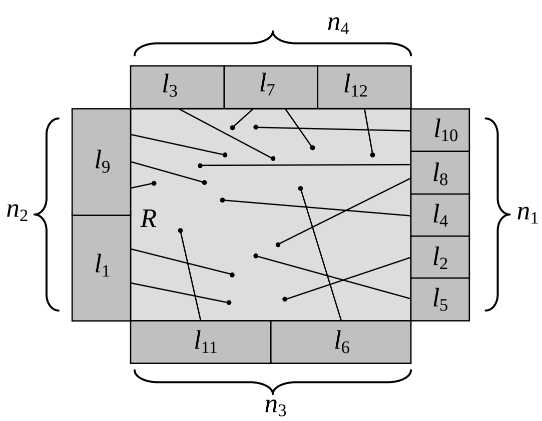


Figure 2: A 4-side many-to-one labeling with type-*s* leaders.

三、問題定義與研究結果

A *k-side type-t many-to-one boundary labelled map* (or *k-side type-t map*, for short), where $k \in \{1, 2, 4\}$ and $t \in \{opo, po, s\}$, is $\mathsf{M} = (P, L, n_1, n_2, n_3, n_4, f)$ where

- $P = \{p_1, p_2, ..., p_N\}$, $p_i \in \mathrm{R}^2$, $1 \le i \le N$, is the set of sites (points) on the plane enclosed in a rectangle,
- L is the set of *labels* with $|L| = n_1 + n_2 + n_3 + n_4$,
- $n_1, n_2, n_3, n_4 \in N$ are the numbers of *labels* to the *East*, West, South, and *North*, respectively, of the axis-parallel rectangle enclosing all sites in P,
- $f: P \to L$ is an onto function which assigns each site in P to a label in L. Note that f is a many-one function in general.

W.l.o.g., we assume that for $k = 1$ (resp., 2), labels are only placed on the *east* side (resp., *east* and *west* sides) of the enclosing rectangle of P. Hence, $n_2 + n_3 + n_4 = 0$ for $k = 1$ and $n_3 + n_4 = 0$ for $k = 2$. On the other hand, labels can be placed on the four sides of the rectangle when $k = 4$. The parameter t, $t \in \{opo, po, s\}$, refers to the type of *leaders* allowed for connecting sites to labels. The *opo*, *po*, and *s* stand for *orthogonal-parallel-orthogonal*, *parallel-orthogonal* and *straight-line*, respectively. See Figure 1 for these three types of leaders. For notational convenience, we refer to the *East*, *West*, *South*, and *North* sides to be the 1st, 2nd, 3rd, and 4th sides throughout the rest of this project. One should also note that every label *l* is connected with $|f^{-1}(l)|$ sites; hence, *l* has to have $|f^{-1}(l)|$ ports to which the sites are connected. Here we assume that the locations of ports of each label are predefined (see label l_3 with three ports in Figure 1 (a)). In addition, the leader connecting site u to label *l* is denoted by u_l.

A *boundary labeling* of a map M is a sequence of labels ($l_{1,1}$, ..., $l_{1,n1}$, $l_{2,1}$, ..., $l_{2,n2}$, $l_{3,1}$, ..., $l_{3,n3}$, $l_{4,1}$, ..., $l_{4,n4}$) such that $\forall\ 1 \le i \le 4$, $0 \le j \le n_i$, $l_{i,j} \in L$. Intuitively speaking, $l_{i,1}$, ..., $l_{i,ni}$ is the sequence of labels along the *i*-th side. W.l.o.g., for *i* = 1 and 2 (i.e., east and west sides, resp.) a top-down ordering is assumed; for *i* = 3 and 4 (i.e., south and north sides, resp.) a left-to-right ordering is assumed. Figure 2 illustrates a 4-side type-*s* boundary labeling. Note that we assume labels (drawn as rectangles) along the same side to be of uniform size; hence, the ordering of labels along each side is sufficient to determine the exact positions of the labels. As *f* is a many-one function in general, there might be several sites connecting to the same label. For example, three sites are connected to label l_3 in Figure 1(a). It is easy to observe that to minimize the number of crossings (or the total leader length), the ordering of ports at which the three leaders touch label l_3 (drawn as a rectangle) must respect the ordering (in increasing order) of the *y*-coordinates of the three sides. The *crossing number* is the number of crossings among leaders in a drawing.

One of the optimization problems considered in this project is as follows:

The *Crossing Problem for k-Side Many-to-One Labeling with Type-t* (*CPkML-t*, for short): Given a *k*-side type-*t* map M, determine a boundary labeling for M so that the crossing number is minimized.

Our main results are correlated with two NP-complete problems, namely, the *Decision Crossing Problem* (DCP) [4] and the *Max-Bisection Problem* [5]. The DCP is stated as follows:

The *Decision Crossing Problem* (*DCP*): Given a two-layered network $G = (L_0, L_1, E)$, an ordering y_0 of L_0, and an integer *M*, is there an ordering y_1 of L_1, so that $cross(G, y_0, y_1) \le M$?

The Max-Bisection problem is stated as follows. So far the best approximation ratio for the Max-Bisection problem is 1.431 [5].

The *Max-Bisection Problem*: Given an undirected graph $G = (V, E)$ with non-negative weights w_{ij} for each edge in *E* (and $w_{ij} = 0$ if $(i,j) \notin E$), partition the nodes in *V* into two sets *S* and $V \setminus S$ of equal cardinality so that $w(S) := \sum_{i \in S, j \in V \setminus S} w_{ij}$ is maximized.

Reduced from the DCP, we can show the following theorem:

Theorem *CP1ML-opo is NP- complete*.

In addition, we show that the *median heuristic* introduced in [4] is an approximation algorithm with approximation ratio three.

Reduced from CP1ML-opo, we can show the follow theorem:

Theorem *CP2ML-opo is NP-complete* even when $n_1 = n_2$.

In addition, we modify the 1.431-approximation algorithm [5] for the MAX-Bisection problem to devise an approximation algorithm for the CP2ML with an approximation ratio about three.

As for the case of type-*po* leaders, we have the following results:

Theorem *DCP1ML-po is NP-complete*.

We investigate an alternative objective to minimize the total leader length, called the leader length problem for many-to-one labeling (*LLPML*). Following Bekos et al.'s result [2], we can show the following:

Theorem *LLPML can be solved in* $O(n^2 \log^3 n)$ *time*.

四、結論與未來展望

The main contributions of this project include: (1) Crossing minimization problems for both one-side and two-side many-to-one labeling with type-*opo* leaders are proved to be NP-complete. We also design approximation algorithms to cope with such intractable problems. (2) Crossing minimization problem for one-side many-to-one labeling with type-*po* leaders is proved to be NP-complete. A heuristic algorithm with satisfactory experimental results is also given for this problem. (3) We show the many-to-one labeling with the objective to minimize the total leader length to be solvable in polynomial time.

五、參考文獻

[1] Bekos, M., van Kreveld, M., Potika, K., and Symvonis, A., Boundary labeling of optimal total leader length, *Proc. of Panhellenic Conference on Informatics*, LNCS 3746, 80-89, 2005.

[2] Bekos, M., van Kreveld, M., Potika, K., and Symvonis, A., Polygons labeling of minimum leader length, *Proc. of APVIS 2006*, pp. 15—21, 2006.

[3] Bekos, M., van Kreveld, M., Symvonis, A., and Woff, A., Polygons labeling of: Models and efficient algorithms for rectangular maps, *Computational Geometry: theory and Applications*, 17:643-644, 1974.

[4] Eades, P. and Wormald, N. C., Edge crossings in drawing of bipartite graphs, *Algorithmica*, 11:379—403, 1994.

[5] Ye, Y., A .699-approximation algorithm for max-bisection, Manuscript, 1999.

出席國際學術會議心得報告

計畫編號	95-2221-E-002-074
計畫名稱	繪圖演算法及資訊視覺化(2/3)
出國人員姓名 服務機關及職稱	顏嗣鈞　台大電機系 教授
會議時間地點	5-7 February, 2007,　澳洲 雪梨
會議名稱	2007年亞太區視覺化會議(Asia Pacific Symposium on Visualization 2007)
發表論文題目	Many-to-One Boundary Labeling

一、參加會議經過

亞太區資訊視覺化會議的前身是已舉辦過三屆的澳洲資訊視覺化會議 (Australasian Symposium on Information Visualisation (Invis.au))，自 2005 年後，此會議擴大為在亞太地區舉辦的國際會議，更名為亞太區資訊視覺化會議 (Asia-Pacific Symposium on Information Visualisation)，是每一年都會舉行的視覺化領域的研討會。今年，除了資訊視覺化 (Information Visualization) 的範疇，更進一步把會議的領域擴大為視覺化 (Visualization) 的範疇，因此會議更名為亞太區視覺化會議 (Asia-Pacific Symposium on Visualisation)。並且由 IEEE VGTC 所贊助，且與 ACM SIGGRAPH 合作。且將在明年正式由 IEEE 所主辦，更名為 IEEE Pacific Symposium on Visualisation。今年此會議是在澳洲雪梨的 National ICT Australia (NICTA)和雪梨大學所舉辦，澳洲必須搭乘十個小時的飛機才能到達。會議的主題包含所有與視覺化相關的議題，包含：

- Display and interaction devices
- Theoretic foundations of visualization
- Usability and human-factors
- Visual models and representations
- Visual knowledge discovery
- Visual data mining
- Visual analytics
- Interaction techniques
- Scientific visualisation
- Information visualisation
- Graph drawing algorithms
- Visualisation of large data sets

- Visualisation of time-varying data
- Visualisation in bioinformatics
- Visualisation in social network analysis
- Software visualisation
- Case studies
- Real-world systems

與會的人士包含了論文發表者與相關領域研究人員，而會議進行包含了一般的邀請演講、口頭及壁報論文的發表。 此次會議包含了社交晚宴，目的在於讓世界各國的參會者先熟悉環境，並且使前輩與後進可以互相交流、討論。其餘正常日安排了正式的研討會議，從早上的九點到下午的四、五點。 一整天的會議分成好幾個段，每段會議都會有一個主題，分配在各段的演講者的題目會和這個主題有所相關。除此之外，大會還邀請到美國 National Visualization and Analytics Center (NVAC) 的主任 Jim Thomas 教授來演講。本人擔任此會議的議程委員。會議期間，除了發表論文外(發表論文題目為「Many-to-One Boundary Labeling」），並擔任 Session Chair。

二、與會心得

APVIS 2007 會議對論文具有相當嚴謹之審查。本會議同一時段只有一 Session，因此與會學者均能熱烈參與討論，交換研究心得。與會中有多位筆者認識多年的國際知名學者，筆者有機會與他們交換研究心得、成果，穫益良多。本人擔任此會議的議程委員。會議期間，除了發表論文外，並擔任 Session Chair。本屆年會共有約五十多位各國學者專家出席。筆者的論文為國內唯一的論文。本人也參訪雪梨大學資訊科學系， 參觀其研究設施與研究成果，並與相關研究學者討論有關 Information Visualization 的研發。雪梨大學資訊科學系中資訊視覺化的研究團隊，陣容相當堅強，在國際上為最頂尖的研究團隊之一。主持人與其中多位研究人員(包括團隊主持人 Prof. Peter Eades 以及核心研究人員 Prof. Seokhee Hong) 熟識，此次也充分和他們交換研究心得，收益良多。

近幾年深深感覺到，Information visualization 是一項不論是在理論以及應用層面，均是目前國際上在計算機科學以及工程方面，相當重要的研究方向，值得國內研究人員積極從事這方面的研究，參與國際活動，與國際上頂尖研究機構從事學術交流，才能迎頭趕上國際水準。