

ラベル配置問題に対する実験的評価

亀田 貴之*

今井 桂子†

*中央大学大学院 理工学研究科 情報工学専攻

†中央大学 理工学部 情報工学科

概要

2次元上に描画されたグラフにラベルを配置する問題をここでは考える。ラベルをつける対象として、点、辺、領域の3種類がある。辺と点をまとめて Graphical Feature と呼び、[3]で Graphical Feature にラベルを配置する一般的手法が提案されているが、精細なアルゴリズムの決め方には自由度がある。本稿ではこれらの自由度を決定し、平面上のグラフの点と辺にラベルを配置する貪欲算法と交差集合を用いたアルゴリズムを提案する。また、東京23区内地下鉄とJRの路線図に駅名や路線名を配置する計算機実験によって、アルゴリズムの評価を行った。

Experimental Performances for Labeling Problems

KAMEDA Takayuki*

IMAI Keiko†

*Information and System Engineering Course,

Graduate School of Science and Engineering Chuo University

†Department of Information System and Engineering, Chuo University

Abstract

The automatic placement of text or symbol labels corresponding to graphical objects is critical in several application areas such as Cartography and Geographical Information Systems. The problem of placing such labels are called the labeling problem, and has been studied intensively. Recently, a unified approach for placing labels for edges and nodes simultaneously is proposed [3]. In this paper, we propose generalized algorithms based on greedy methods and overlap sets for train maps and show some computational results.

1 はじめに

ラベル配置問題は主に地図制作などに利用されており、自動描画などの分野とも関連して、研究が行われている([5]に主な論文リストがある)。地図制作においては、ラベルを付ける対象によって大きく次の3つに分けられる。

- 点(都市, 建物)にラベルを配置する NLP(Node Label Placement)問題。
- 辺(川, 路線)にラベルを配置する ELP(Edge Label Placement)問題。
- 地域(国, 海)にラベルを配置する多角形配置問題(polygon containment problem)。

NLP問題において、地図上の点をラベルを表す長方形のどこと一致させるかによっていろいろ

なモデルが考えられている。平面上の点の集合とラベルの大きさが与えられた時、ラベルの3点以上の点と平面上の点が一致するようにして、重ならないように配置することのできるラベルの最大数を求める問題は NP 困難であることが知られている[4]。

ELP問題は辺上の点にラベルを付けると考え、NLP問題に帰着して解く手法が取られていたが、最近では、辺と点をまとめて Graphical Feature として扱い、これらを同時に配置する問題(Graphical Feature Label Placement, 略して GFLP 問題)が考えられた。それに対する一般的な手法が[3]によって提案されている。この手法をもとに、他の辺や点に重なっても配置する試みに関する初期段階の実験を行ない、その結果を報

告した [1]. また, [1] では領域にラベルをつける問題も扱っている.

本稿では点と辺において, なるべく多くの良い位置にラベルを配置することを目的として [3] をもとにしたアルゴリズムを提案し, 計算機実験を行なった. また, [2] で開発した連続した辺の 1 つにラベルを配置する方法を用いて, 東京都 23 区内の地下鉄と JR の路線図に路線名と駅名を配置する実験も行なったのでそれを報告する.

2 Graphical Feature へのラベル配置

まず, [3] で提案されている一般的手法を述べる. そこでは, 配置の条件として次のようなものを考えている.

- (1) 他の Graphical Feature やラベルに重ならない.
- (2) ラベルは対応する Graphical Feature が一意的に定まる.
- (3) ラベルは最良の位置に配置する.

ここで (3) は配置させるグラフによって異なり, どの位置にラベルを付けると見やすいかを考えることとなり, 表現するのが難しい条件である.

GFLP 問題の一般的なアルゴリズムは次の 3 つのステップからなる.

- (1) 各 Graphical Feature に対して, ラベル配置可能位置を離散的に定義する.
- (2) 各ラベル配置可能位置にコストを定める.
- (3) コストの総和が最小になるように, ラベル配置可能位置の中からラベル位置を決める.

実際にこの枠組で計算機実験を行なうためには, ラベル配置可能位置の定め方, コストの決め方などに自由度がある. 次の節では, 今回の計算機実験で行なった手法を述べる.

3 ラベル配置可能位置の決定とコスト

ここで問題を簡単にするために, 付けるラベルの大きさはすべて同じとし, 与えられた平面上のグラフの各辺は直線と仮定する. 辺に対するラベル配置可能位置は次のようにして決定した. ラベルの幅を W , 高さを H とし, 辺の傾き s と H/W を比較し,

$$-\frac{H}{W} \leq s \leq \frac{H}{W}$$

の場合には垂直線で, そうでない場合には水平線で辺を分割し, ラベルが辺に接する位置をラベル配置可能位置の候補とする (図 1). 点に対するラベル配置可能位置は, ラベルの各頂点が点に接する位置をラベル配置可能位置の候補とした (図 2). 他の辺や点と重なるようなラベル配置は行なわないので, 他の辺や点と重なるラベル位置をすべて削除し, 各辺と各点のラベル配置可能位置を決定する. 次にコストを定める.

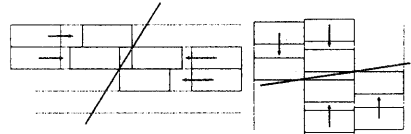


図 1. 分割線とラベル配置可能位置の関係

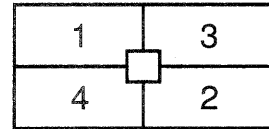


図 2. 点に対するラベル配置可能位置とコスト

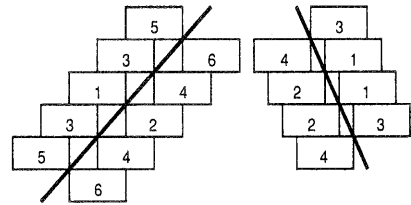


図 3. 辺に対するラベル配置可能位置のコスト

辺 e に対するラベル配置可能位置の中で, 付随する各ラベルの接点が, 辺 e の中央に近いラベル配置可能位置から小さなコストを割り当てる. また, 辺の上側のラベル配置可能位置の方がコストを小さくしてある (図 3). 次に点に対するラベル配置可能位置のコストの決定は, 左上, 右下, 右上, 左下の順にコストを割り当てた (図 2).

4 貪欲的解法

最も単純な方法として貪欲法を用いた算法がある. 同じ辺や点に対する 2 つのラベル配置可能位置は交わらないように定義したが, 異なる辺や点のラベル配置可能位置は一般に交わっている. 交わる 2 つのラベル配置可能位置を選ぶことはでき

ないので、どちらかを予め次の貪欲算法を用いて削除する。辺と各辺のラベル配置可能位置に順番をつけ、その順に処理を行っていく。ここで、 Λ を全ての Graphical Feature のラベル配置可能位置の集合とし、 Λ_e を辺 e のラベル配置可能位置の集合とする。また、ラベル l にコストを与える関数を $COST(l)$ とした。

[コスト優先法]

今、処理を行なう辺 e のラベル配置可能位置を $l \in \Lambda_e$ とする。もし、 l と交わるラベル配置可能位置 $l' \in \Lambda_{e'}$ がなければ、 Λ は更新せず、次のラベル配置可能位置の処理に移る。もし、 l と交わるラベル配置可能位置 $l' \in \Lambda_{e'}$ があれば、 $COST(l) < COST(l')$ の場合、 $\Lambda \leftarrow \Lambda - \{l\}$ とし、 e に対する処理を続ける。 $COST(l) \geq COST(l')$ の場合、 $\Lambda \leftarrow \Lambda - \{l\}$ とし、次のラベル配置可能位置に対して同様の処理をする。これを未処理のラベル配置可能位置がなくなるまで行なう。

[交差数優先法]

今、処理を行なう辺 e のラベル配置可能位置を $l \in \Lambda_e$ とする。もし、 l と交わるラベル配置可能位置 $l' \in \Lambda_{e'}$ がなければ、 Λ は更新せず、次のラベル配置可能位置の処理に移る。もし、 l と交わるラベル配置可能位置 $l' \in \Lambda_{e'}$ があれば、 l, l' それぞれと交わるラベル配置可能位置の数 (交差数) を数え、交差数の大きい方を Λ から削除する。交差数が同じである場合はコスト優先法のルールに従って削除する方を決める。

図 4 は、辺に対するラベル配置可能位置の例である。ラベルの中央の数字は、各辺によって決められたコストとなっている。また、辺に近い数字は、ラベルを捜査する順の番号である。図 5, 6 は、それぞれ図 4 におけるコスト優先法と交差数優先法のラベル配置可能位置の例である。図からも分かるように、ラベル配置可能位置数に違いがでていいる。コスト優先法ではラベル配置可能位置のない辺が存在するのに対して、交差数優先法ではコスト優先法でなかった辺にもラベル配置可能位置が存在している。

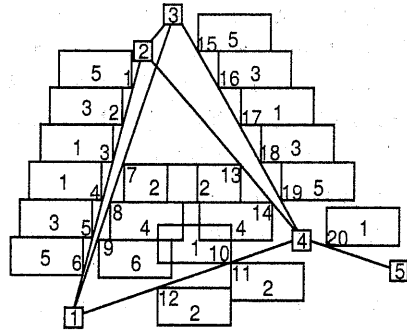


図 4. 辺に対するラベル配置可能位置の例

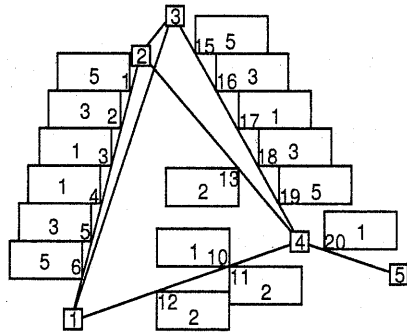


図 5. コスト優先法のラベル配置可能位置

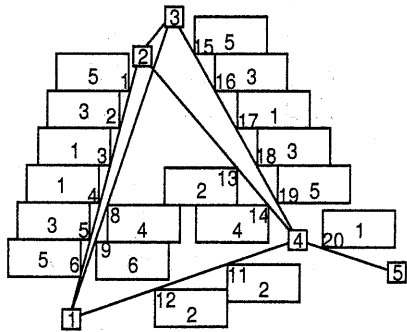


図 6. 交差数優先法のラベル配置可能位置

5 交差集合を用いた解法

ここではオーバーラップしているラベルどうしを1つのグループとして扱い、そのグループの中において各ラベルが他のどのラベルとも重ならないように分けた後に各 Graphical Feature についてラベル位置を選ぶ方法を考える。まずはじめにラベル位置を集め、2つ以上のラベルがオーバーラップしていたら同じグループにする。以後このグループを交差集合と呼ぶことにする。大まかなアルゴリズムを以下に記述する。

- (1) 交差集合を見つける。
- (2) 交差集合を下記アルゴリズムを用いて分ける。
- (3) 各 Graphical Feature において、コスト最小のラベルを選択する。

次に交差集合を分けるアルゴリズムを以下に記述する。

- 各交差集合の中で、次数最大のラベルを削除する。
- 次数が同じ場合、コストの大きいラベルを削除する。
- コストが同じ場合、そのラベルが対応する Graphical Feature のラベル総数の多いラベルを削除する。

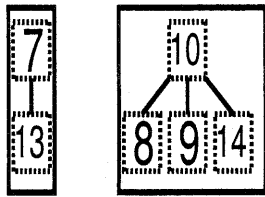


図 7. 図 4 に対する交差集合の例

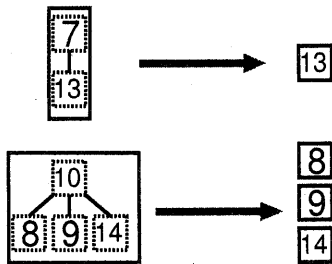


図 8. 交差集合を分ける

図 7 は図 4 に対する交差集合の例である。これを上記のアルゴリズムで分ける (図 8)。こうすることで、前節で述べたラベルに順番をつけて貪欲算法を用いる方法より、ラベル配置可能位置の数を多くできると思われる。

6 計算機実験

今まで述べてきた手法によって、東京 23 区内の地下鉄などの路線図にラベルを配置する計算機実験を行なった。ここでは、駅の位置は地図データから取り、駅間は直線で結んで得られるグラフを入力としている。駅を点とし、それを結ぶ線路を辺と考え、路線名のラベルは同じ路線線上にある辺のどこかに配置すれば良い。[2] で提案された手法を用いて連続した辺へのラベル配置を行った。そこでは、次の 2 ステップを行ない配置を求めている。

- (1) 同一視する各辺にコストを割り当てる。
- (2) 割り当てられたコストの小さい順に辺を調べ、その辺にラベル配置可能位置が存在する場合、その中からコスト最小の位置を選ぶ。

ステップ 1 の辺に割り当てるコストの決め方は、連続した辺の中央に配置したい場合は次のようにする。

[辺コストの決定]

- (1) 同一視する辺が一本のパスの場合、始点から終点に向かって順に番号を振る。枝分かれがある場合、本線を上のようにして番号を振り、副線には分岐した辺の番号から降順に番号づけを行なう。図 9 では、上側の数字が辺の番号を表している。
- (2) 辺につけた番号の最大値を N とし、 $M = \lceil \frac{N}{2} \rceil$ とおく。
- (3) 辺番号と M の差の絶対値をコストとする。

次に、東京 23 区内の地下鉄と JR の路線図に対して駅名、路線名を配置した実験について述べる。

図 10, 図 11, 図 12 は東京 23 区内の JR と地下鉄の路線図をそれぞれコスト優先, 交差数優先, 交差集合を用いて路線名と駅名のラベルを配置した図である。各図とも最終的な位置ではそれほど変化がなかったが、ラベル配置可能位置数に大きな違いがある (表 1, 2)。表 1 と表 2 はそれぞれ

表 1. JR 路線図の点, 辺に対するラベル配置可能位置数とラベル配置数の比較

	コスト優先 (図 10(a))	交差数優先 (図 11(a))	交差集合 (図 12(a))
辺ラベル配置可能位置数	55	58	59
点ラベル配置可能位置数	173	180	181
ラベル配置数/路線数	7 / 7	7 / 7	7 / 7
ラベル配置数/駅数	77 / 81	79 / 81	72 / 81

表 2. 地下鉄路線図の点, 辺に対するラベル配置可能位置数とラベル配置数の比較

	コスト優先 (図 10(b))	交差数優先 (図 11(b))	交差集合 (図 12(b))
辺ラベル配置可能位置数	51	51	55
点ラベル配置可能位置数	216	229	235
ラベル配置数/路線数	9 / 12	9 / 12	10 / 12
ラベル配置数/駅数	132 / 210	132 / 210	128 / 210

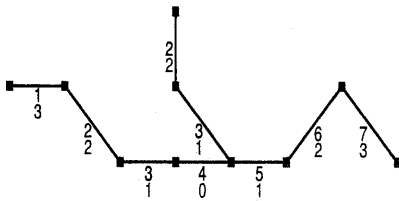


図 9. 同一視した辺の番号とコスト

れ JR (a) と地下鉄 (b) の辺に対するラベル配置可能位置の数と点に対するラベル配置可能位置の数を表している。また、実際に配置できたラベル数も表している。コスト優先, 交差数優先, 交差集合の順にラベル配置可能位置数が多くなっている。この例では, 交差集合を用いた方法は辺に対するラベルを多く配置できるが, 点に対するラベル数は貪欲算法より少なくなった。配置できるラベル数は, 入力するグラフやラベルの大きさに依存する。

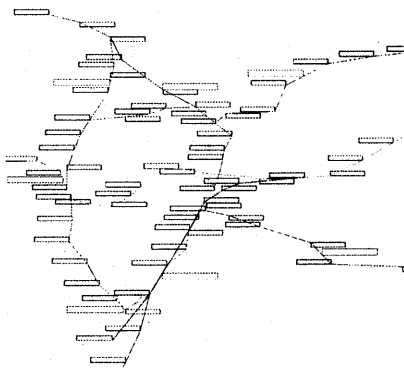
7 おわりに

本稿では, 平面に描画されたグラフの点と辺にラベル配置する手法について議論し, 今回用いたコスト関数やラベル選択の方法の有効性を計算機実験によって示した。また [3] の GFLP 問題についてのアルゴリズムを拡張し, 路線図への応用を考察した。アルゴリズムの詳細の決め方に自由度があるため, さまざまなものが考えられるが, それらのアルゴリズムは一長一短がある。具体的な状況に応じて使い分ければ, 要求にあう結果を得

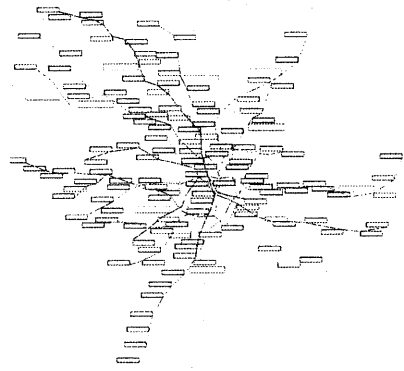
られると思われる。計算機実験から分かるように, 辺に対するラベルを配置することで点に対するラベルが配置できなくなる場合もある。このように, 密集した Graphical Feature に大量の情報が存在する場合には, ラベルを配置できない Graphical Feature が存在する。このような場合には, これを解決する方法を考えなければならない。また, ラベルを離散的に定義するのではなく, ラベルを移動して, 空いている空間を捜し出し, 各ラベルとの間隔が同じような長さになる位置に配置する方法を考察中である。

文献

- [1] 今井桂子, 中岡北巳, 林恵意: “地図におけるラベル配置問題,” 第 3 回 ‘統合型地理情報システム’ シンポジウム予稿集, pp.11-18, 1999.
- [2] 今井桂子, 亀田貴之: “ラベル配置問題とその路線図への応用,” 第 5 回 ‘統合型地理情報システム’ シンポジウム予稿集, pp.17-24, 2000.
- [3] K. G. Kakoulis and I. G. Tollis: “A Unified Approach to Labeling Graphical Features,” *Proc. 14th Annual ACM Symposium on Computational Geometry*, pp.347-356, 1998.
- [4] M. Formann and F. Wagner: “A Packing Problem with Applications to Lettering of Maps,” *Proc. 7th Annual ACM Symposium on Computational Geometry*, pp.281-288, 1991.
- [5] A. Wolff: “A Map Labeling Bibliography,” 2000. <http://www.inf.fu-berlin.de/map-labeling/bibliography.html>

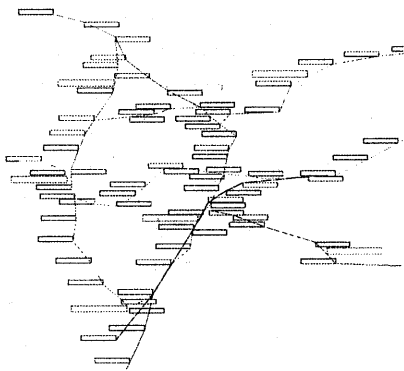


(a) JR

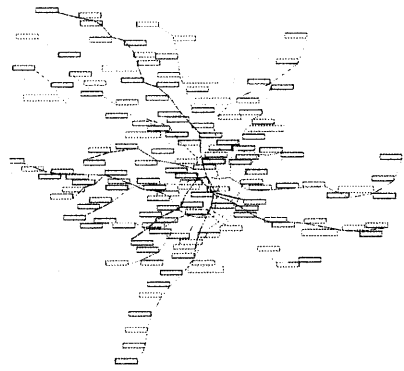


(b) 地下鉄

図 10. コスト優先によるラベル配置

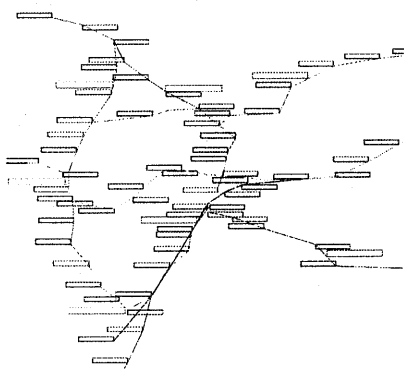


(a) JR



(b) 地下鉄

図 11. 交差数優先によるラベル配置



(a) JR



(b) 地下鉄

図 12. 交差集合によるラベル配置