

デジタル道路地図のリンクの沿道状況の自動判別*

Automatic Distinction of Terrain Types of Digital Road Map Links*

鷹尾和享**・東徹***

By Kazutaka TAKAO**・Tohru HIGASHI***

1. はじめに

(1) 交通量配分のネットワークデータ

交通量配分は対象となるネットワークのデータを必要とする。したがって、ネットワークの全リンクのリンク・パフォーマンス関数を用意する必要がある。リンク・パフォーマンス関数は、リンク交通量が増えるに従ってリンク走行時間が増加する様子を表す関数であり、いくつかのパラメータを必要とする。具体的には、BPR 関数のように、各リンクの自由走行時間とリンク容量を必要とするのが一般的である。

リンク容量（基準交通容量）とは、道路条件および交通条件が基本的な条件を満たしている場合に単位断面を1時間に通過しうる乗用車の台数である。交通量配分で必要とするリンク容量は、各リンクの実際の条件下での容量であり、理想的な条件下での容量である基準交通容量に対して、交通容量を低下させる種々の要因（たとえば、車道幅員、側方余裕、沿道状況、二輪車類）による補正を行って求める¹⁾²⁾³⁾。

交通量配分の実務では、多くの場合、これらの補正率や補正後の容量が表で与えられている場合が多く、各リンクの属性に基づいて表を参照するだけでよい。たとえば、土木学会(2003)⁴⁾は道路種別と沿道状況の別による補正率を表の形で示している。また、交通工学研究会(1973)⁵⁾は種級・沿道状況・車線数によって容量を求められるような表を示している。

交通量配分で用いられるネットワークデータは、通常、手作業で作成され、維持管理されている。そのうえ、実務の交通量配分の対象範囲は、通常、広範囲である。したがって、数多くのリンクの属性を設定する必要がある。

(2) 沿道状況とは

沿道状況は、リンク容量を補正するためのリンク属性の1つである。Highway Capacity Manual (HCM) では、容量の算定は都市部と地方部とで異なる枠組みで行われる²⁾。地方部では、大型車の速度に着目することで、平地(level)・丘陵地(rolling)・山地(mountainous)の3つに分類される。

我が国では、HCM とは異なり、都市部も地方部も同じ枠組みで扱われ、補正率の値が両者で異なるだけである。我が国の沿道状況は、沿道施設や無信号交差点からの出入りや路上駐車によって交通流が妨げられ、交通容量が低減することを表現するための補正である⁶⁾。その結果、観測データの分析によって、沿道状況は次の3つに分類される⁷⁾。

平地：人家が連担しておらず、道路の勾配が緩やかな地域

山地：山地、丘陵及び山麓地域

都市部：DID（人口集中地区）およびその他市街部（道路の両側に人家が連担し、市街部を形成している地域）

(3) 道路データの管理

道路データは通常、国や道路管理者等の機関で管理される。たとえば、我が国では、主要道路のデータは、道路交通センサス⁸⁾によって何年かごとに調査される。センサスでは交通量・旅行速度・道路状況等が調査され、沿道状況も含まれている。米国でも同様に、Highway Performance Monitoring System (HPMS) という国家レベルの道路情報システムが構築されており⁷⁾、延長・交通条件・性能・運用特性等のデータが管理されていて、沿道状況も含まれている。

(4) 問題点

多くの場合、交通量配分で用いられるリンクデータは、センサスや地勢図を参照して人手で構築されてきた。しかし、この方法には次のような問題点がある。

第1に、センサスの該当箇所のデータを手作業で調べるという作業をネットワークの全リンクについて行うのは、非常に煩雑な作業であり、ミスも混入しやすい。しかも、この作業は1度きりで済むものではなく、ネット

* Keywords: 交通量配分、ネットワーク、地図、機械学習、決定リスト

** 正員 博士(工)

(社)システム科学研究所 (〒604-8223 京都市中京区新町通四条上ル小結棚町428 新町アイエスビル TEL: 075-221-3022, FAX: 075-231-4404)

*** 正員 工修

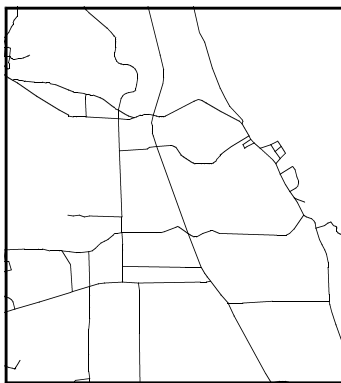
(社)システム科学研究所 (同上)

ワークデータを最新の状態に更新し続けなければならない。

第2に、交通量配分のネットワークデータとセンサスとの同期を取る必要があるため、データを更新するのに次のセンサスのリリースを待つ必要がある。したがって、ネットワークデータの最新性はセンサスの制約を受けることになる。

(5) 目的

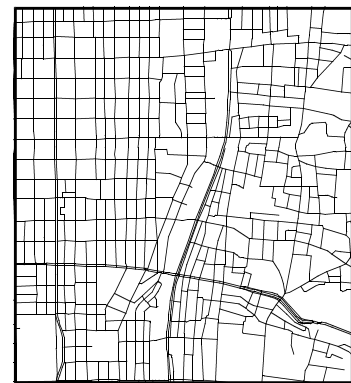
問題の解決方法の1つは、デジタル道路地図(Digital Road Map, DRM)から配分用のネットワークデータを作成する方法である。DRMは全国の道路管理者からの情報に基づき、毎年データの更新が行われている。DRMは交通量配分のリンクデータに利用できる情報を多く含んでいる。しかし、沿道状況は含まれておらず、別途与える必要がある。⁸



(a)平地



(b)山地



(c)都市部

図-1 沿道状況によるネットワーク形状の違い

図-1に例を示す。これはDRMのデータ⁸⁾を筆者が図化したものである。したがって、次のようなif-thenルールをうまく作成できれば、沿道状況の3区分を自動的に判別することができる：

もし(if)、リンクが曲がりくねっていれば、(then)沿道状況は山地である。

次に、このようなif-thenルールを作成する方法について述べる。

(2) 決定リスト

機械学習とは、人間が自然に行っている学習能力と同様の機能をコンピュータで実現させるための技術・手法のことであり、情報処理の分野で盛んに研究されている。これは、ある程度の数のサンプルデータ集合を対象に解析を行い、そのデータから有用な規則、ルール、知識表現、判断基準などを抽出する。このサンプルデータ集合のことを「学習セット」という。学習セットのリンクに正しい沿道状況を与えて解析を行い、有用な規則を

そこで、本稿では、DRMのリンクの沿道状況をDRM自身を用いて自動的に判別する方法を述べる。DRM自身から読み取る情報を用いることで、センサスとの同期を取る必要がなくなり、また、判別を自動的に行うことができる。

2. 沿道状況の判別方法

(1) 基本的な考え方

3つの沿道状況区分には視覚的な形状の違いがある：

- (a) 平地はネットワークが疎であり、リンクは長くまっすぐな傾向がある。
- (b) 山地はネットワークが疎であり、リンクは長く曲がりくねっている傾向がある。
- (c) 都市部はネットワークが密であり、リンクは短くまっすぐな傾向がある。⁹

抽出すれば、対象ネットワークのリンクすべてに規則を適用し、沿道状況を自動的に判別することができる。

機械学習は自然言語処理の分野で広く用いられており、いくつかの手法がある。決定リストはその1つで、Rivest(1987)⁹⁾によって提案され、Yarowsky(1994)¹⁰⁾によって自然言語処理に導入された。これは、対象データの何らかの素性に着目して、信頼度による優先順位付きの規則の集合を得る方法である。

これを沿道状況の判別に適用するには、対象ネットワークから一部のリンクをランダムに抽出し、それを学習セットとして用いて規則の信頼度を求め、優先順位を定めることになる。具体的には次のような手順となる。

Step 1: 素性の選定

沿道状況の違いの判別に利用できるリンクの素性を選定する。素性とは、リンクに関する情報の種類のことであり、沿道状況の違いをよく説明でき、かつ、DRMから容易に読み取れるものでなければならない。本稿では

次の素性を選定した。A は当該リンクが位置する付近の場所についての情報であり、B は当該リンク自身についての情報である。

(A-1) ノード密度：リンクの周囲{500, 1000, 1500}m 以内のノード数

都市部のネットワークは密であるから、都市部でこの値が大きくなり、それ以外では小さくなるのが期待できる。周囲何メートルが効果的なのか前もって判断できないので3種類の距離を用意した。

(A-2) 補間密度：リンクの周囲{500, 1000, 1500}m 以内のリンクあたり補間点数

DRM ではくねくね道は補間点を細かく配置することで折れ線として表される。山地はくねくね道が多いのでこの値が大きくなるのが期待できる。

(B-1) 延長(m)

都市部ではネットワークが密なため、結果としてリンク延長は短くなる傾向がある。

(B-2) 曲がり具合 (度/km)：(補間点で曲がる角度の総和 -180°) ÷ リンク延長

山地のくねくね道はこの値が大きくなる。なお、都市部でも宅地の周囲等に直角に曲がるリンクがあるので、-180 とした。

(B-3) 道路種別

指定市の一般市道は都市部であると期待できる。

(B-4) 車線数

4車線以上のリンクは高規格のリンクであり、山地ではないことが期待できる。

Step 2: 学習セットの選定

対象ネットワークの 63547 本のリンクの中から、1000 本をランダムに選び、それらのリンクに人手で正しい沿道状況を与えた。そのうち、750 本は学習セット、250 本はテストセットとして用いる。学習セットは規則の信頼性を推計するために用い、テストセットは判別実験に用いる。

Step 3: テンプレートの用意と規則の抽出

規則抽出のためのテンプレートを用意し、規則を抽出する。本稿では次の2つのテンプレートを用いた。[] はテンプレートパラメータを表す。

もし (条件) ならば沿道状況は[Type]である。

もし (条件) かつ (条件) ならば沿道状況は[Type]である。

(条件) の部分には、量的な素性の場合には次のパターンが入る。

[素性] ≥ [値]

[素性] < [値]

カテゴリカルな素性の場合には次のパターンが入る。

[素性] = [属性値]

[素性]の部分には Step 1 で選んだ素性が順に代入される。[値]の部分には、データの最大値と最小値の間でいくつかの値を設定した。この結果、この段階では信頼度とは無関係に、多くの規則のインスタンスが抽出される。

Step 4: 規則の信頼度による並べ替え

学習セットを用いて各規則の信頼度を求め、規則を信頼度の高い順に並べ替える。規則の信頼度は尤度によって求めた。尤度は次のように定義した。

$$L = \log ((\text{真のデータ数} + \alpha) / (\text{偽のデータ数} + \alpha))$$

α は 0.25 とした。尤度の絶対値は規則の信頼度を表す。ただし、尤度が正の場合は述語は「である」となり、負の場合は絶対値にするとともに述語の「である」を「でない」に反転させる。並べ替えの結果、本稿では表-1のような規則が得られた。

表-1 信頼度の順に並べ替えた規則

信頼度	規則
7.44833	もし[補間密度(1000m)] < [2]なら[山地]でない
7.42952	もし[補間密度(1500m)] < [2]なら[山地]でない
7.25912	もし[ノード密度(500m)] ≥ [33]かつ[補間密度(500m)] < [2]なら[山地]でない
7.21008	もし[ノード密度(500m)] ≥ [33]かつ[延長] < [321]なら[山地]でない
...	...

Step 5: 規則の適用

規則を対象のリンクに順に適用する (図-2)。

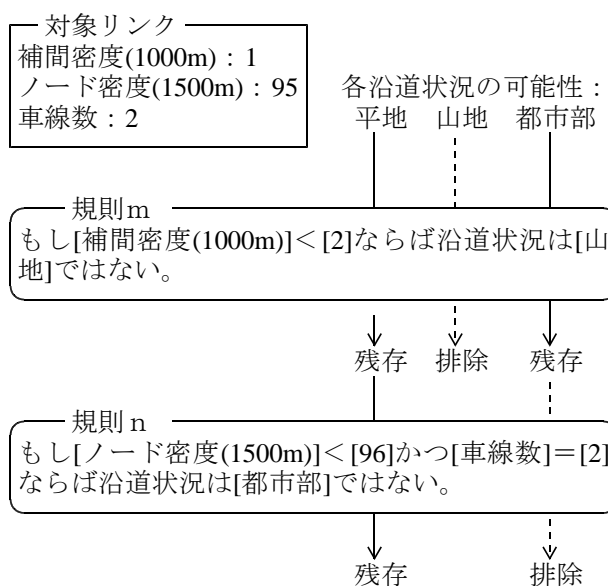


図-2 規則の適用

3. 判別実験

提案手法の性能をテストするため、得られた規則をテストセットの 250 本のリンクに対して適用し、判別実験を行った。本手法による判別結果が、予め人手で付与された正解の沿道状況と一致するかどうかを調べた。表-2 に結果を示す。「成功」は一致したもの、「失敗」は一致せず、判別結果が不適切であったものを表す。「許容範囲内」は、もともとはっきりと判別するのが困難なリンクであり、ベストアンサーには一致しなかったが、許容範囲内として挙げておいた沿道状況に一致したリンクである。これを見ると、かなり好成績で判別できていることがわかる。

表-2 実験結果

区分	リンク数
成功	215 (86.0%)
許容範囲内	18 (7.2%)
失敗	17 (6.8%)

「許容範囲内」と「失敗」の原因を分類し、表-3 に示す。許容範囲内となったものは場所の境界部分のものが多く、たとえば、都市部から平地へは徐々に変化するため、その境目付近に位置するものは人間が見ても判別が悩ましい場合が多い。一方、失敗したものは特殊な場所に位置するものが多い。

表-3 不成功の原因

原因	許容範囲内	失敗
場所の境界部分		
異なる沿道状況の境目付近	11	5
程度の境界部分		
大きくない集落	3	2
険しくない丘陵地	1	1
特殊な場所		
山間部集落等、例外的に密	1	3
都市部海岸等、例外的に疎	0	2
山地だが線形は良い	2	1
その他の例外的な場所	0	3

4. 結論

本稿では、決定リストを用いて DRM リンクの沿道状況を自動的に判別する方法を述べた。本稿の成果は次のように要約される。

- (a) 提案手法は、少数のリンクに人手で沿道状況を与えることで、信頼度の高い判別規則を得ることができる。一旦規則が得られれば、大量のリンクの沿道状況を自動的に判別することができる。
- (b) 提案手法は DRM 自身から読み取れる情報を用い

る。したがって、センサス等の他のデータと同期を取る必要がなくなる。

- (c) 提案手法は DRM に含まれる全てのリンクの沿道状況を判別することができる。したがって、センサス対象外のマイナーなリンクについても自動的に沿道状況を得ることができる。

本稿の成果は、ネットワークデータの最新性を保つのに貢献できると思われる。残された問題は、特殊な場所の問題である。そのような特殊な場所はランダムに抽出した学習セットには多くは含まれず、対処方法を検討する必要がある。また、量的な素性を論理的な if-then 条件に組み込む方法も検討の余地がある。

本稿は沿道状況について述べたが、本手法は沿道状況以外のリンクの 2 次属性の判別にも簡単に応用できる。

参考文献

- 1) 近畿地方整備局: 平成 11 年度全国道路交通情勢調査 (道路交通センサス) 一般交通量調査箇所別基本表, 2001
- 2) 交通工学研究会: 道路の交通容量 1985, コロナ社, 1987
- 3) Luttinen, R.T., Innamaa, S.: *Nordic Highway Capacity - Uninterrupted Flow Facilities in Denmark, Finland, Norway, and Sweden*, Finnra Internal Publications 4/2000, Finnish National Road Administration, Traffic and Road Engineering, Helsinki 2000
- 4) 土木学会: 道路交通需要予測の理論と適用 (第 1 編) 利用者均衡配分の適用に向けて, 土木学会/丸善, 2003
- 5) 交通工学研究会: 交通工学ハンドブック, 技報堂出版, 1973
- 6) 塚田幸広, 桐山孝晴, 保久原均, 濱谷健太: 道路の交通容量における新しい設計法に関する検討, 国土技術政策総合研究所資料第 317 号, 2006
- 7) Highway Performance Monitoring System (HPMS), Federal Highway Administration (FHWA), <http://www.fhwa.dot.gov/policy/ohpi/hpms/index.htm>
- 8) 住友電工製 全国デジタル道路地図データ (DRM) 近畿地方福井県 2003 年度版 informatix
- 9) Rivest, R.L.: Learning Decision Lists, *Machine Learning*, 2, pp. 229-246, 1987
- 10) Yarowsky, D.: Decision Lists for Lexical Ambiguity Resolution: Application to Accent Restoration in Spanish and French, in *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, Las Cruces, NM, pp. 88-95, 1994