

特集 軽量な渋滞予測エンジンの中国での実現性評価*

Feasibility study of a light weight traffic prediction system in China

増谷 修
Osamu MASUTANI

劉 正
Zheng LIU

Traffic prediction is a key technology of recent traffic information systems. Especially in China, traffic services become more popular and crucial for many people by the spread of personal nomadic devices or smart phones. We introduce a traffic prediction engine based on combination of 3 different (short, middle, long term) light-weight and accurate prediction methods into a traffic information system in China. We evaluated that its prediction accuracy outperforms baseline prediction methods such as AR model, by using floating car based traffic data. Moreover, the developed prediction system can process data from 150 thousand links in Beijing in a short period of time using a portable PC, with a reasonable cost, and without heavy investment in road related infrastructure.

Key words : Traffic prediction, Travel time

1. まえがき

渋滞予測は交通情報システム (TIS) やカーナビゲーション機器などでよく用いられる技術である。VICS¹⁾などの通常のTISは現状の交通状況を配信するサービスであるが、現況のみでは、目的地到着までに変化する渋滞を捉えることができず、最短経路の探索や、目的地までの正確な所要時間の算出に失敗することがある。この場合は、将来の渋滞を予測するための予測情報が必要となる。渋滞予測は、天気予報と同様、人々の日々の行動を左右するサービスといえる。一部のカーナビゲーションシステム、テレマティクスサービスで、すでに様々な形態の予測サービスが提供されているが、そのほとんどは、統計的渋滞予測と呼ばれる手法を用いている。これは交通の一日ごとの周期的な変化を利用して、過去の(曜日ごとなどの)統計情報(最も簡単な方法は平均値)を元に将来の特定の日時の渋滞情報を予測するものである。この手法は任意の予測ホライズン(何期先を予測するか)について予測可能となるため、特に長期(予測ホライズンが1時間以上先)の予測手法として用いられることが多い。

一方、短期(数分~1時間先等)の予測については、統計情報を用いるよりも、直前の状況をもとに予測した方が良い結果が得られることが多いため、短期予測について特別な予測を導入しているシステムもある。ただし、短期予測は、その性質上、刻一刻と変化する現状値の更新1ステップごとに予測をアップデートする必要がある。そのため短期予測を行うためには、比

較的高い計算コストが必要となる。このような短期予測の高い計算コストは、特に多数の道路リンクを持つ大都市への、予測システム導入には課題となる。

一方中国の都市のような、道路の拡張が速く交通情報インフラの整備が追い付かない状況では、交通情報の収集に、道路に設置される交通量感知器よりも、プローブカーシステムがその導入コストから好んで用いられる。プローブカーシステムでは、サービス提供リンクはその情報品質を問わなければ事実上都市のリンク全体となりしばしば膨大となる。

我々は、このようにプローブカーシステムを前提とし、大規模な都市へのサービス提供を行うTIS事業者に対する渋滞予測エンジンを提供する。そのために、計算コストにおいて軽量なシステムを目指した。我々は、このような要件を満たしながら、全予測ホライズンについて予測情報を提供しつつ、すべてのホライズンでベースライン手法の精度を超える予測手法を提案する。この手法を利用すれば、現状の交通情報システムに追加インフラ(サーバー等)なしに投入可能なアドオンエンジンとして利用することができ、カーナビ等でのオンボードの渋滞予測も可能となる。

2. 交通情報データの発展と予測エンジンへの要求

われわれの渋滞予測エンジンのターゲットは、中国の世紀高通社²⁾のTISである。北京において実に15万道路リンクの交通情報を提供している。これらはカーナビで用いられるGISリンクをもとにしており、細街路まで含んだものである(Fig. 1(a))。一方、中国の交通

*2012年ITS世界会議で発表論文を和訳し掲載

情報配信サービスでは日本と同様にFM電波によるリアルタイム交通情報の配信が主流である。そのため、実際に交通情報が配信されるリンクは、放送波の容量の制限から、一部の主要道路のみとなる。世紀高通社ではこれをRTIC (Realtime Traffic Information of China) リンクとして定義している (Fig. 1(b))。ただし、配信可能なリンク数は今後携帯網などWAN (Wide Area Network) の普及に連れて拡大すると思われる。

中国など交通渋滞の激しい国においては、渋滞を迂回する経路を探索することがますます重要であり、その迂回路は必然的に細街路が多く含まれることになる。そういった意味でも、細街路までの渋滞情報を提供することが今後重要度を増してくることは明らかである。それゆえ、我々の予測エンジンはGISリンクのような、膨大な数のリンクに対応する必要がある。



Fig. 1 (a) GIS Links

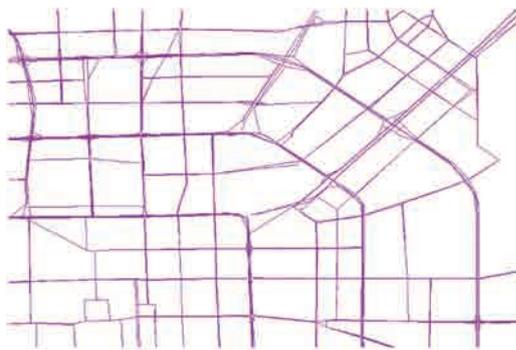


Fig. 1 (b) RTIC Links

Fig. 1 Example of Map of GIS Links and Telematics Links in Beijing

我々の渋滞予測システムでは、5分から60分までの5分ごとの12ホライズン予測について、渋滞情報の配信間隔 (5分) ごとにアップデート (再計算) を行っている。これらを北京15万リンクに対して、リンク情報の圧縮等を行わずに1~2分で完了をしている (後述)。リンクの圧縮などの処理を行っていないため、圧縮時の誤差などが出ない、並列処理化がしやすい、所要時間の予測がしやすい、などのメリットがある。これら

の特徴は、サービスエリアの拡張、地図の更新時の対応、オンデマンドの予測 (必要な期間だけ特定のリンクの予測を行うなど) に対する柔軟性を高めている。

3. 渋滞予測手法

我々のシステムでは短期、中期、長期の各予測ホライズンにて、それぞれ最適な予測手法を用いた。渋滞の計測方法としては、ここでは道路リンクのある時間間隔 (ここでは1ステップ5分間隔) に通過した車両の平均リンク旅行時間 (リンクの通過時間) である。

短期予測 (5分~10分先) の手法として、フェロモンモデルを用いている³⁾。この手法は群知能研究で用いられている生物のメカニズムにヒントを得た手法である⁴⁾。蟻などの昆虫が仲間とのコミュニケーションのために場に残していく化学物質 (フェロモン) の特性を模して、様々な動的な探索行動を効率よく行うモデルである。我々は仮想フェロモンを道路リンクに配置される渋滞度に対応したものとして定義し、そのダイナミクスを用いて、短期先の渋滞を予測する。具体的には下式で表せるような結合写像格子 (Coupled Map Lattice) の形式の漸化式である。

$$r(t, p) = \frac{1}{|C(t, p)|} \sum_{i \in C(t, p)} \frac{1}{v_i(t, p)}$$

$$s(t+1, p) = E_p \times s(t, p) + \alpha_p \times r(t, p) + q(t, p)$$

$$q(t, p) = \sum_{p' \in N(p)} \frac{F_p}{|N(p)|} r(t, p')$$

ここで、 t :時刻 p :リンクID s :フェロモン量 r :発生量 q :拡散量

v_i :リンク速度 C :隣接リンク a, E, F : [発生, 蒸発, 拡散]パラメータ

自リンクと隣接リンクの1ステップ前の状況から次の自リンクの状況を予測する。これは、フェロモンの発生、蒸発、拡散のオープンなパラメータを内包するが、そのパラメータを実際の交通データを利用してGA (遺伝的アルゴリズム) により最適化している。我々は、以前の研究で、このモデルの有効性を、日本の交通調査データ (三鷹) やIICプローブカーデータ (名古屋) にて実証している³⁾。

このモデルは多項式の単純な計算式であり、予測自体は非常に高速である。なお、以前の研究では、信号

サイクルや車両一台一台の加減速などの詳細なモデルを実装していたが、これらは1分間隔という短いサイクルでの現象に依拠しており、今回のような5分間隔程度の交通情報では有効でないと思われ採用していない。

中期予測（10分～1時間先）の手法として、直前時系列のパターンをパターンマップという表ベースでクラスタリングし、予測値を決定する手法を用いている。Fig. 2のように学習時には、過去の直前時系列のデータ（historical pattern）を予測対象値ごとに分けてパターンマップ（clustered patterns）を構成し、クラスタリングを行う。予測時には、直前時系列（current pattern）から、このパターンマップからもっとも一致度の高いものを求め予測値を得る。

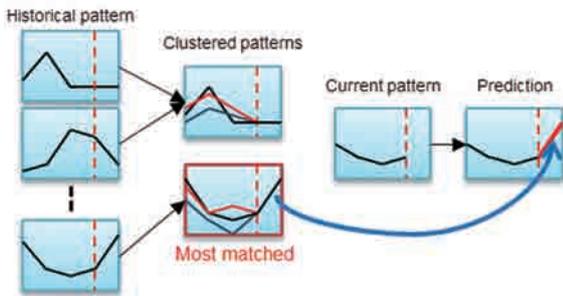


Fig. 2 Medium Term Prediction (clustered pattern map)

本手法は現在時刻の数ステップ前から現在までの直近交通量パターンを用いて数ステップ先までの予測を行うという点では、他の時系列手法と同じ形態をとる。ARモデルなどでは線形の予測しか行えないが、本手法はパターンを直接将来の値と対応付け、非線形な予測も可能となる。また、本手法は過去時系列をすべて保持して最近傍探索を用いるような手法よりも比較回数が圧倒的に少ないため高速に処理可能である。

具体的なパターンマップは、通常行われるようなクラスタごとの平均値を保持するのではなく、元データを時刻と量子化された値の2次元のマップに投票する形態で表現され、元データの形状をある程度記憶しながら1つのマップで複数のパターンを統合して保持できる。これにより、通常クラスタリングで行われるような平均化による過度の情報圧縮を避けることができ、精度の悪化を避けている。予測はマップ（確率表）の各エントリの値から予測値の条件付き確率を求める形態の単純ベイズ法にほぼ一致する。ただし、マップは時系列をそのまま表現するため、時間方向への減衰、交通量方向への値の補完などの操作などを行いやすい。本手法によると、モデルの学習は2次元マップへ

の投票のみ、予測は最適なマップの選択のみとなり、両フェーズとも加減算中心の処理となり、計算コストが非常に低い。

また、本手法ではマップの更新はオンライン（予測と同時に）で行えるため、変化していく都市の交通状況に適応しながら予測することができる。これにより、予測モデルの更新を別途行う必要がない。この特徴は中国のように発展の目覚ましい都市では必須の機能といえる。

長期予測（1時間以上先）の手法として、日々の交通データの時系列をクラスタ化した後に、決定木を用いて曜日など日付属性との関連づけを行うモデルを採用した（Fig. 3）。クラスタリングは、1日分の時系列同士の差分を用いた距離の逆数を類似度としたK-means手法を用いた。クラスタリング後の平均値時系列データを決定木の葉要素とし、日付属性としては、曜日、祝日、天候、気温などを用いた。

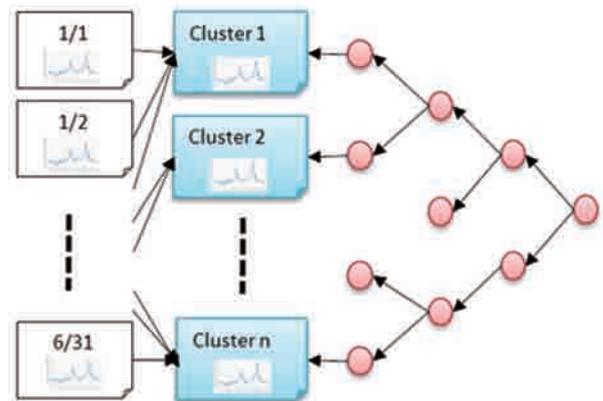


Fig. 3 Long Term Prediction (Decision tree model over clustered pattern)

どのような日付属性パターンに対して、これらのクラスタのうちどのパターンがよく発生するかを少ない条件分岐で精度良く対応づけるために決定木を用いる。実際に得られた決定木の例をFig. 4に示す。この例では、予測対象日が金曜日かどうかということが、対応する時系列クラスタを割り当てるうえで最重要ということになっている。

決定木による予測は単純な場合分けとなるため、予測時の負荷は最小限であり、ニューラルネットワークやベイジアンネットワークなどに比べて、通常のプログラミング言語での表現も容易である。また得られた予測モデルについて、人手による解析やメンテナンスをしやすい点も決定木の好ましい特徴である。



Fig. 4 Example of Induced Decision Tree

4. 評価結果

本予測エンジンを、実際に世紀高通社で提供している中国の交通情報を元に評価した。評価指標としてはリンク旅行時間の予測値の実測値に対する誤差をRMSE (Root Mean Square Error, 平均2乗誤差) で比較した。評価データはTable 1の仕様である。

Table 1 Specifications of Evaluation Data

	Short-term	Mid-long term
Location	3 rd ring in Beijing	Outer ring highway in Shanghai
# of links	397	55
Length	94 km	26km
Test duration	1 week	2 week

短期予測については、ベースライン手法であるAR (自己回帰) モデルよりも、予測誤差が6%ほど低減されている (Fig. 5)。中期予測については、ベースライン手法であるARモデルよりも、予測誤差が25%ほど低減されている。長期予測については、ベースライン手法である平均値予測よりも、予測誤差が4%ほど低減されている。

Table 2 Results Comparison between RTIC and the Fixed Sensor Data in RMSE (sec/km)

	baseline	Proposed method	Error ratio
RTIC data	10.3	10.1	-2%
Fixed sensor data	6.37	5.87	-8%

上記の評価はGISリンクを元に行っている。これに加えRTICリンクを用いた評価を、短期、中期に関して行った (Table 2)。またプローブカーデータではなく交通量感知器によるデータを利用した評価も行った。両データにおいて、ベースライン手法に対し良い性能を確認した。このように我々の手法は交通データ種類に依存しないモデルとなっており、様々なデータサービスへの統合が容易になっている。

Table 3 Effect of Using Arrival Time Prediction in RMSE (sec/km)

	Current data	Prediction	Error ratio
Mid-term (pattern map)	5.54	5.29	-5%
Long-term (decision tree)	9.04	6.30	-31%

さらに前記のリンク旅行時間をもとに到着時刻予測の評価を行った。到着時刻予測の精度はリンク旅行時間予測の単純な累積とはならないため、より実利用に即した評価を行うために本評価は不可欠である。

評価対象道路として、北京第三環状道路の15.9kmの区間 (リンク数102) を用いた。1週間分のデータを用い、5分おきに出発した場合の到着時刻予測の精度を2016回測定してその平均を求めた結果、Table 3のように現状値だけを用いた場合よりも大きな改善を見た。

プローブカーシステムではより広いカバレッジを得られる反面、プローブカーの往来の少ないリンクではデータが十分に得られない問題がよく発生する。データが得られない時間帯の交通量データを補完するためには直前のデータをそのまま用いることもできるが、予測を行うことで、この補完をより精度よく行えない

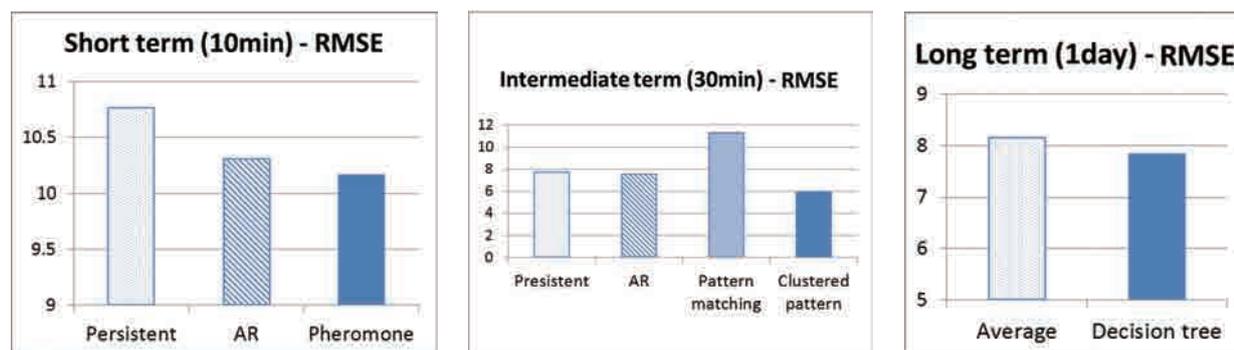


Fig. 5 Result of Short Term Prediction (The rightmost bar shows the proposed method in graphs)

かを評価した。データが得られない時間帯の時系列を同リンクのデータより予測することで補完する時間的補完と、周辺リンクのデータより予測することで補完する空間的な補完を検証した。評価は、時系列が実際には得られている時間帯のデータを仮想的に削除して、補完することで実測値との誤差を評価した。ベースラインの補完手法としては直前の現状値をそのまま用いる方法と比較した。

評価の結果、短期予測手法を補完に用いた場合、時間空間双方とも誤差を5%ほど低減させる効果があるという結果となった (Table 4)。

Table 4 Effect of Temporal / Spatial Complementation in RMSE (sec/km)

	Current data	Prediction	Error ratio
Temporal complementation	4.36	4.14	-5%
Spatial complementation	4.47	4.23	-5%

5. 実装方法と計算時間

我々は通常のWindows PC上で.NETベースの予測エンジンを実装した。現状の計算時間はTable 5に示すとおりである。モデルおよび予測結果はSQL Serverデータベースに保存されておりDBデータから逐次データのフェッチを行いながら処理している。処理時間にはこれらのオーバーヘッドも含まれ実際の計算時間はさらに短い。このように、数年前のリーズナブルなPCサーバ機一台にて十分に処理を行うことができる。またTable 5に示すように、昨今のマルチコアCPUを有効活用できるアーキテクチャとなっており、コア数に対して線形な処理時間低減を実現している。

Table 5 Hardware Specifications and Computation Time

	Short term	Middle term
PC spec	CPU: Xeon E5410 (Quad core 2.3GHz) Memory: 4GB OS: Windows Server 2008 R2 x64	
# of links	150,000 links (Beijing, GIS link)	
Processing time with single core	2.6 s	45.4 s
Processing time with quad core	0.7 s	11.3 s

6. むすび

我々は3つの比較的軽量な予測手法を用いて、全予測ホライズンに対して従来手法を凌駕する精度を持つ渋滞予測エンジンを構築した。中国世紀高通社の提供するプローブカーデータ、車両感知器データにおいて、評価を行い、GISリンク、RTICリンク上でのリンク旅行時間予測の精度、到着時刻予測の精度、リンク旅行

時間の補完の精度において、ベースライン手法を超える精度を達成した。

また、これらすべての予測を5分ごととアップデートするために十分な計算速度を達成し、中国のような大規模な都市の膨大な道路リンクへ対応するとともに、既存交通情報サービスへのアドイン予測エンジンとして扱いやすい製品展開を可能とした。

<参考文献>

- 1) VICS (Vehicle Information and Communication System) Center, <http://www.vics.or.jp/>
- 2) 北京世紀高通科技有限公司, <http://www.cennavi.com.cn/>
- 3) Masutani, O., Sasaki, H. Iwasaki, H., Ando, Y., Fukazawa, Y., Honiden, S. (2005). Traffic Prediction using Pheromone Model. In Proceedings 12th World Congress on ITS, San Francisco.
- 4) J. A.Sauter, R. Matthews, H.Van Dyke Parunak, S. Brueckner, Evolving Adaptive Pheromone Path Planning Mechanisms, Proc. First International Conference on Autonomous Agents and Multi-Agent System (AAMAS'2002), 2002

<著 者>



増谷 修
(ますたに おさむ)
株デンソーアイティーラボラトリ
研究開発グループ
ITS, 交通工学, GISに関する
研究開発に従事



劉 正 (りゅう せい)
株デンソー 情報通信システム
開発部
ITS, 交通工学, LBSに関する
研究開発に従事