

札幌市における交通渋滞の発現特性と空間分布に関する研究

Data mining of congestion trends and congestion-point distributions
-A case of the Sapporo urban area-

室蘭工業大学工学研究科公共システム工学専攻
秋田工業高等専門学校環境都市工学科
(株) ドーコン交通部
室蘭工業大学建築社会基盤系学科

○学生員 木村 洋平 (Yohei Kimura)
正 員 長谷川 裕修 (Hironobu Hasegawa)
非会員 内藤 利幸 (Toshiyuki Naito)
フェロー 田村 亨 (Tohru Tamura)

1. はじめに

インフラの賢い使い方が議論される時代においては、道路容量の増強だけではなく、ソフトを含む適切な施策の選定、その実施位置やタイミング、効果的な対象範囲の設定等も重要な評価項目となる。細密かつ大量な ITS データは、これらのニーズに応えるものであるが、その解析については、ある仮説をもって集計し、それを統計的に検証するアプローチだけではなく、統計的に検証可能なデータ群を機械的に選定して、その背後の仮説を推測する、仮説発見型アプローチも重要な試みとなる。

そこで本研究では、札幌市内で観測された渋滞情報データにシーケンシャルマイニングを適用し、アソシエーション分析と頻出パターンマイニング分析を行い、渋滞発生頻度の少ない地点における渋滞発生ルールの抽出を試みた。

2. 既往研究レビューと本研究の位置づけ

これまでの交通情報データを利用した研究は、主に道路利用者への情報提供の質向上を目標として進められてきた。しかし、蓄積が進むデータの背後に潜んでいるマネジメント上有用なパターンや仮説を発見する視点から、研究を進めている事例はまだ多くない。

渋滞現象理解のために仮説発見型分析アプローチであるデータマイニングを試みた研究としては、塚井ら¹⁾がトラフィックカウンターから得られる広島空港の流入自動車交通量を対象にデータマイニングを試み、空港流入車両の到着時刻の分布特性を明らかにした研究がある。

道路ネットワーク上の複数の VICS データから、データマイニング的アプローチにより渋滞現象理解と施策意思決定支援を行った事例としては、筆者ら²⁾が札幌市内約 2,500 箇所に設置されたセンサーから大量に得られる渋滞情報データを用いて、通年・夏期冬期における時間的・空間的に特徴のある渋滞発現パターンの抽出を行い、渋滞現象理解と対策立案支援を試みた研究がある。しかし既往の研究成果では、年間総渋滞発生の多い地点を対象とした時間的・空間的な特徴の渋滞発現パターンを把握することはできたが、未知の情報の発掘までには至らなかった。

そこで本研究では、自明でなく、未知であり、かつ、役立つ可能性のある情報の発掘のために、シーケンシャルマイニングの適用を試みる。シーケンシャルマイニングとは、アイテムの発生順序を保ったままマイニングを

行う手法であり、時系列を考慮した相関ルールの抽出や頻出アイテム群の抽出が可能である。

時間・空間を変数としたシーケンシャルマイニングを行った研究としては、武智ら³⁾が 128 名のアンケートから得られた個人の行動をデータとし、モニターが訪問施設に訪れる行動の周期性や条件をマイニングによって評価した研究がある。しかし、大量の交通渋滞データベースにシーケンシャルマイニングを適用した事例はまだない。大量の渋滞情報データベースから未知で有用なルールやパターンを抽出することができれば、渋滞現象の総合的な理解と道路運営者への施策立案支援が可能となる。

3. データの前処理

3.1 渋滞情報データの観察

図-1 に 4 年次の渋滞発生回数と発生地点数の分布を示す。図の縦軸は一日に観測された渋滞回数の合計、横軸は渋滞が観測された地点数である。なお、4 月から 11 月を夏期、12 月から 3 月を冬期とした。夏季の渋滞発生地点数は 400 箇所程度で、渋滞が発生する地点はほぼ決まっている。しかし、冬期の地点数は 1,000 を超えているときもあり、降雪状況次第により渋滞発生箇所数が増加し、面的な広がりをみせることが分かる。

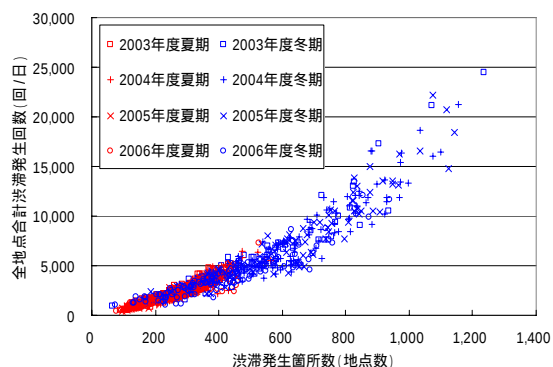


図-1 渋滞発生回数と地点数の分布

道路のサービスレベルは“渋滞の発生回数”だけに依存するものではない。時間・場所・気象に応じて発生する突発的な事象によって、「ドライバーが予想する移動時間やサービスレベル」と、「実際の移動時間や道路のサービスレベル」間の差が増大し、サービスレベルが低下するものと考えられる。このことから、既往研究内で棄却していた渋滞発生回数の少ない地点の中に、未知であり、かつ、有用な情報が隠れている可能性がある。

3.2 データの前処理

そこで、渋滞発生回数の少ない地点を抽出するために、4年次の地点別渋滞発生回数を多い順に並べ替え、全地点の総渋滞回数の総和に占める累積渋滞回数の比率が80%以上となる渋滞箇所を抽出した。

例として、2006年度の渋滞発生回数の分布を図-2に示す。図中の点線の右側にプロットしてある地点が、累積率80%以上の渋滞箇所となっている。この累積率80%以上の地点を対象に分析を試みる。

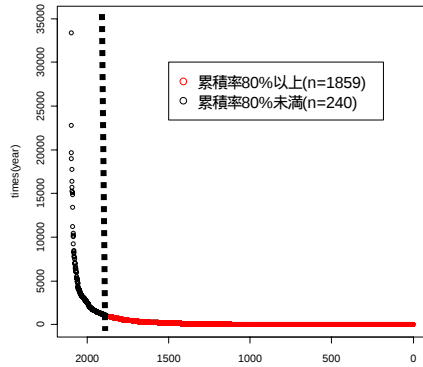


図-2 2006年度渋滞発生回数の分布

4. 関連・頻出パターンルール抽出

4.1 アソシエーション分析について

アソシエーション分析とは、巨大なデータベースの中から価値のある相関関係（ルール）を見つけ出す分析のことである。マーケットバスケット分析とも呼ばれ、POSデータなどに適用されている。POSデータをイメージした架空のデータを表-1に示す。表-1のようなデータをマーケットバスケットトランザクションと呼び、各行をそれぞれ一つのトランザクション、あるいはバスケットと呼ぶ。

表-1 POSデータのイメージ例

TID (Transaction ID)	アイテム集合
1	{ハム、果物、パン、弁当}
2	{オムツ、ビール、パン}
3	{ハム、牛乳、パン、ビール}
4	{ビール、タバコ、オムツ}
5	{弁当、オムツ、ビール、ジュース、}

大量のデータベースの中から関連ルールを抽出する際には、何らかの評価指標が必要となる。関連ルールの代表的な評価指標として、support（支持度）と confidence（確信度）が挙げられる。

supportとは、全トランザクションのうち、前提部（X）と帰結部（Y）に登場する全てのアイテムを含むトランザクションの割合であり、(1)式で定義される。confidenceとは、条件部を含むトランザクションのうち、帰結部も同時に含むトランザクションの割合であり、(2)式で定義される。

$$supp(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{M} \quad (1)$$

$$conf(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} = \frac{supp(X \Rightarrow Y)}{supp(X)} \quad (2)$$

本研究では、年月日をトランザクション、路線・方向・住所・時刻を一つにまとめたものをアイテムとし、相関ルール抽出を行う。アソシエーション分析の出力例を表-2に示す。

表-2 相関ルール出力例

```
conf=1.0    supp=0.0737
道5号_上り_手稲区手稲本町 2-2_08:05
<= 道5号_上り_手稲区富丘 3-7_08:05
```

このルールでは、発生頻度は少ないが“道5号_上り_手稲区富丘 3-7_08:05”で渋滞が発生した場合には、必ず“道5号_上り_手稲区手稲本町 2-2_08:05”で渋滞が発生するというを示している。両交差点は近接しており、高い相関関係が見受けられた。上記ルールは既往の分析では発見できなかった未知の情報であるといえ、これらの情報を積み上げていくことで、渋滞発生頻度の少ない地点同士の隠れた相関関係を見出すことが可能である。

4.2 頻出パターンマイニングについて

頻出パターンマイニングとは、データベースの中で、一定頻度以上で現れるパターンを列挙・抽出する手法である。この手法を用いることで、渋滞発生頻度が少ない地点集合における「複数の渋滞が同時に発生する地点群」や、渋滞発生頻度が少ないある一地点における「連続して渋滞が発生する時間」といった頻出パターンの抽出が可能である。分析結果は都合により口頭での発表とする。

5. おわりに

本研究では、札幌市内で観測された渋滞情報データにシーケンシャルマイニングを適用し、アソシエーション分析と頻出パターンマイニング分析を行い、渋滞発生頻度の少ない地点における渋滞発生ルールの抽出を試みた。発生頻度の少ない地点においても、頻出アイテムの時間は8時台に集中していた。そのため、他の時間帯における“発生頻度は少ないが確信度の高いルールの抽出”が難しいということがわかった。全時間帯でマイニングを行うのではなく、ある時間帯で区切って分析を進めていく必要があることがわかった。時間帯で区切った分析結果は、発表時に公表することとする。

謝辞：本研究の遂行にあたっては、北海道開発局札幌開発建設部から渋滞情報データをご提供頂いた。ここに記して感謝の意を表します。

参考文献

- 塚井誠人、井上真一、奥村誠：データマイニングによる広島空港流入車両の時間分布パターンの分析、第28回交通工学研究発表会論文報告集、（2008）
- 木村洋平、郎芮漪、有村幹治、長谷川裕修、田村亨：データマイニング法による交通渋滞の季節変動把握、第39回土木計画学研究発表会（春大会）
- 武智環、羽藤英二、渡部眞幸、中山沖彦：個人の行動パターンのシーケンシャルマイニング、第39回土木計画学研究発表会（春大会）