

- 14 ETC データを利用した高速道路の交通量分析 に関する基礎研究

A Study for Analyzing of Traffic Situation of Expressway Using ETC Data

吉田博哉¹・田中成典²・古田均²・

杉江功³・山本昌孝³・神野裕昭⁴・榎山武浩⁵

Yoshida Hiroya, Tanaka Shigenori, Furuta Hitoshi,

Sugie Isao, Yamamoto Masataka, Jinno Hiroaki and Kashiyama Takehiro

抄録：近年、交通量分析に関する研究が盛んに行われており、橋梁や道路の整備計画や道路沿道地域への商業活動の戦略策定において活用されている。しかし、既存の交通量分析では、道路交通センサスや起終点(OD:Origin Destination)調査により収集した交通情報を利用しているため、多大なコストを要する。そこで、本研究では、全国の高速道路で導入が進められる自動料金収受システム(ETC:Electronic Toll Collection System)において記録される情報からOD情報を生成する手法を考案し、そのシステム化を図った。本システムでは、利用者毎にOD情報を生成し、継続的に蓄積することにより、これまで不可能であった移動履歴の分析を可能とした。さらに、データマイニングを用いて、移動履歴から交通流動を分析する事例を示す。

Abstract: Recently, the research on the traffic analysis is actively done, and it is used to decide the maintenance plan of the bridge and strategy of the commercial operation to the road coast region. However, a large cost is required in an existing analysis, because traffic information collected in a road traffic census and the Origin Destination Survey is used. In this research, we design the method for generating OD information based on the Electronic Toll Collection System, which maintained in Japanese expressway. This system continuously accumulated OD information generated to each user, and the analysis of impossible movement history became possible. In addition, the case that traffic flow analysis is done from the movement history is shown by using data mining.

キーワード： ETC, OD 情報, データマイニング, 交通流動分析

Keywords : Electronic Toll Collection System, Origin Destination Information, Data Mining, Traffic Flow Analysis

1. まえがき

阪神高速道路は、現在、総延長 233.8km、日平均交通量は約 91 万台と、関西都市圏の大動脈となっている。しかし、朝夕の混雑時をはじめ、慢性的な渋滞が問題となっており、合流部、特に環状線手前の本線合流および本線、ランプ合流で著しい渋滞を生じている。このような渋滞に対して、旧阪神高速道路公団では、流出促進ランプを建設し、また、平成 17 年 10 月の民営化後は、渋滞を回避する新規乗り継ぎルートの設定を検討している。これらの対策の効果予測等に用いるデータは、従来、道路交通セン

サスにあわせてほぼ 5 年毎に実施している阪神高速道路の起終点(OD:Origin Destination)調査のみに頼ってきた。しかし、OD 調査は、ある一日のアンケートでデータ収集しており、予算や回答率の制限から、一日あたり数万件から十数万件のデータを集めることが限界となっていた。また、全交通量を記録しているトラフィックカウンターのデータ¹⁾やブロープデータ²⁾では、交通量を把握できたとしても、利用者の行動パターンを分析し、将来予測することが困難と言える。

一方、平成 13 年度から本格的な運用を開始した有料道路料金自動収受システム³⁾(ETC:Electronic

1 : 学生員 情修 関西大学大学院 総合情報学研究科

(〒569-1095 大阪府高槻市霊仙寺町 2-1-1, Tel : 072-690-3213, E-mail : yoshida@kansai-labo.co.jp)

2 : 正会員 工博 関西大学 教授 総合情報学部

3 : 正会員 工修 阪神高速道路株式会社 (〒569-1095 大阪府大阪市中央区久太郎町 4-1-3, Tel : 06-6252-8121)

4 : 正会員 工修 株式会社建設技術研究所 (〒810-0041 福岡県福岡市中央区大名 2-4-12)

5 : 学生員 情学 関西大学大学院 総合情報学研究科

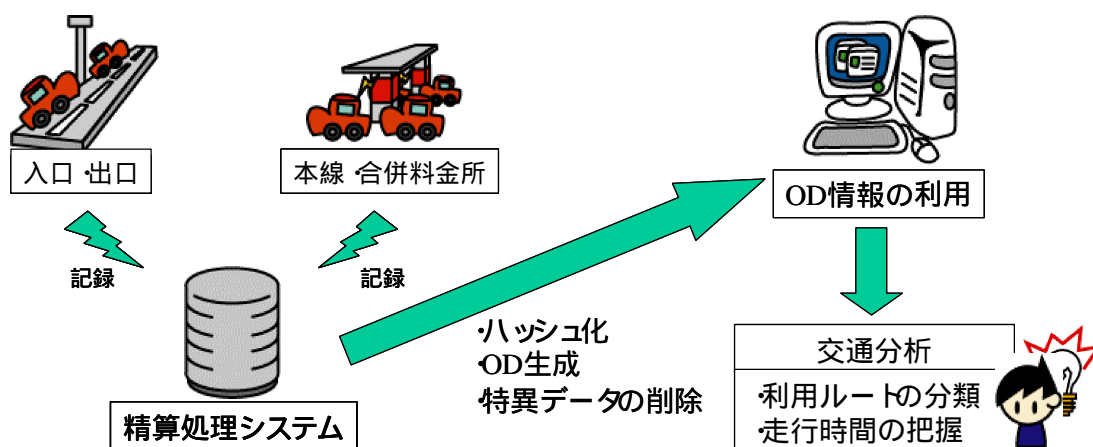


図 - 1 OD 情報分析システムのイメージ

Toll Collection system) は、阪神高速道路における利用率が約 60%に達しており、1ヶ月あたり 3,000 万件以上ものデータが記録される。そして、ETC によって記録される ETC データは、ETC カードを表す一意な値、料金所情報、通過時刻、車種といった情報で構成されており、各料金所における交通量として精算処理システムに保存される。ただし、精算処理システムに格納された ETC データは、料金所における通過情報であるため、利用者の行動パターンを分析するには至らない。

そこで、本研究では、図 - 1 に示すように各料金所から精算管理システムに蓄積された ETC データをもとに、OD 情報の生成を行った。そして、生成された OD 情報を利用して、交通量を分析するためのシステムを開発する。また、生成された OD 情報のデータを活用して、交通分析の一例として、OD 間交通の相関係数の算出、および OD 間の走行時間の分析を行う。

2. OD 情報分析システムの概要

(1) 個人情報の保護

本来、ETC によって得られた情報は、「有料道路自動料金収受システムを使用する料金徴収事務の取扱いに関する省令（平成 11 年建設省令第 38 号）」により、有料道路の料金徴収を行うことのみによる利用に限られている。また、個人情報を含む ETC データは厳重なセキュリティで守られた料金精算処理システム内にあり、直接、交通解析等に用いること

はできない。しかし、「有料道路自動料金収受システムにおける個人情報の保護に関する指針（平成 11 年建設省道有発第 19 号）」では、道路の交通量を把握するために、個人を識別できない情報を作成することは認められている。

そこで、各出入口および料金所で記録されている ETC データの取得時に、ETC カード ID の情報を暗号化して、個人を識別できないようにした。なお、ETC カード ID の暗号化は、Hash 関数を用いることで、32 文字の不可逆性の文字列に変換することとした。ただし、同一の ETC カードが使用された場合、それらの特定は可能で、当該カードを持つ利用者の基本的なデータを併せ持つ。

(2) 入力情報

本システムの入力情報は、ETC の路側装置が設置されている入口、出口、本線料金所、合併料金所において記録された ETC データを利用する。ETC データには、Hash 関数を用いて暗号化した ETC カード ID や通過した料金所コード、支払った金額、割引コードなど様々な情報が格納されている。そこで、精算処理システムに格納されている情報のうち、表 - 1 に示す情報を利用して OD 情報を生成する。

3. OD 情報の生成

(1) ETC カード ID の変換機能

OD 情報の生成の処理は、図 - 2 のフローチャートに示すように、まず ETC カード ID の変換を行う ETC

表 - 1 利用した ETC データの構成

項番	フィールド名	主キー	型名	サイズ	フィールドの説明
1	id		数値型	4	ETCデータのIDが格納
2	card_id	x	文字型	32	ETCカードID(暗号化)が格納
3	tollbooth_code	x	数値型	4	料金所コードが格納
4	antenna_kind	x	数値型	4	アンテナ種別が格納
5	pass_time	x	日付型	8	通過時刻が格納
6	pass_payment	x	数値型	4	通過料金が格納

カード ID の変換機能では、ETC データに含まれる暗号化された ETC カード ID の情報を数値に変換する。数値への変換は、データベースへ格納するデータ量を軽減するためである。暗号化された ETC カード ID の情報は、32 バイトの文字列で構成される。ETC データは 1 ヶ月間で 1,000 万件以上のデータ量となるため、ETC カード ID を格納するだけでも約 320MB 以上の領域が必要となる。

そこで、データ管理の効率化を図るために、暗号化された ETC カード ID を 4 バイトの一意な数値に変換することで、データ量の軽減を図る。ただし、ETC データ内には、同一の ID を持つ情報が含まれている場合があるため、変換後においても、それらの特定を可能にする必要がある。本機能では、暗号化された ETC カード ID (32 バイト) と変換後の ETC カード ID (4 バイト) との対応表を作成する。なお、対応表には、変換前と変換後の ETC カード ID を対にして登録するものとする。ETC カード ID の変換では、まず、暗号化された ETC カード ID を変換し、対応表に既に ETC カード ID が登録されているかどうかの確認を行う。対応表に暗号化された ETC カード ID が存在しない場合は、ETC カード ID の一意な数値に変換し、変換前と変換後の ETC カード ID を対応表に新規に登録する。また、既に対応表に登録されている場合、対応表内に記載されている値を取得する。

ETC カードの ID 数は、1 ヶ月で 60 万件以上になる。さらに、複数月にわたり ETC カード ID が追加される対応表では、100 万件以上の膨大なデータ件数になり、この対応表から全検索を行うと、膨大な計算量が必要となる。本機能では、Hash 法を用いることで、ETC カード ID の検索速度の向上を図った。Hash 法とは、データ検索アルゴリズムの一種で、Hash 関数を用いて Hash 値と呼ばれる一意な値を生成し、その値

をもとに検索を行うものである。Hash 法では、データ件数が増加した場合においても最適なツリー構造を構成することで、全探索に比べ、探索回数を軽減することが可能である。本機能では、メモリ上で Hash の対応表を作成し、計算量の軽減を図った。

(2) OD 情報の生成機能

OD 情報の生成機能では、まず生成した OD 情報を保存するデータベースの構成について検討した。さらに、ETC カード ID を変換した ETC データを利用して、OD 情報を生成した。

a) データベースの構成

本システムのデータベースは、ETC データを一時的に格納するテーブルと、生成した OD 情報を格納するテーブルで構成される。

ETC データは、1 ヶ月間で 1,000 万件以上の膨大なデータ量になる。データベースから情報を取得する速度は、テーブルから取得するレコード容量に依存するため、1 ヶ月間の ETC データを 1 つのデータベースで管理すると、データ取得速度が低下する。一方、ETC データを管理するデータベースを分割すると、手動でデータを確認する手間が増大する問題が生じる。本システムでは、手動でデータを確認する箇所をシステム化し、ETC データを複数のテーブルに分割して格納することで、手動でデータを確認する手間の増大を押さえると同時に、膨大な ETC データの処理時間短縮を図る。

また、別の問題として、データベースには、データを格納する上での制約がある。例えば、本システムで利用する Microsoft Access の場合には、1 つのデータベースファイルとして 2GB 以上のデータを保持できないという制約がある。そこで、上記を考慮して本システムでは、ETC データを一時的に保存するテーブル及び生成した OD 情報を格納するテーブルを半月毎に作成し、カード ID の先頭番号で区分し 16 分割した。

b) OD 情報の生成

OD 情報の生成は、阪神高速道路内の料金所によって記録された ETC データを利用する。本システムで利用する料金所種別は、「FF 入口」、「一般入口」、「FF 出口」、「一般出口」、「本線料金所」、「合併料金所」の 6 種類となっている。「FF 入口」と「FF 出口」は、料金所がないため料金徴収を行わず FF (フリーフロー) の状態で、ETC 車両の通過情報を取得する出入口である。また、「一般入口」と「一般出口」は、料金所があり料金徴収を行う出入口の料金所を示す。「本線料金所」は、高速道路の本線上にある料金所を示す。「合併料金所」とは、阪神高速と西日本高速道路株式会社の管理する高速道路の料金を一括して徴収する料金所で、ETC を西日本高速道路

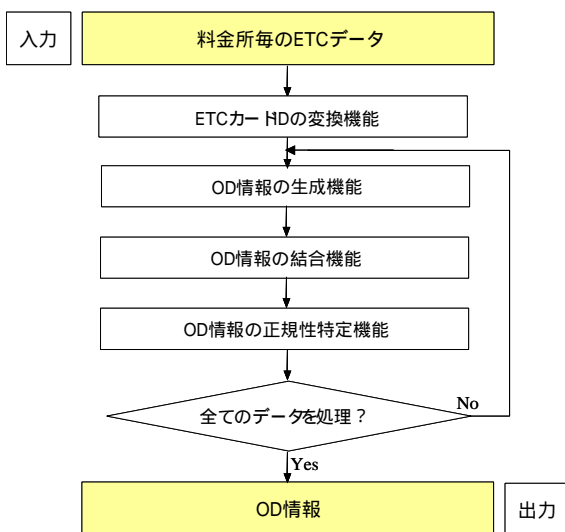


図 - 2 OD 情報の生成のフローチャート

表 - 2 生成したOD情報の構成

項番	フィールド名	主キー	型名	サイズ	フィールドの説明
1	od_id		数値型	4	ODデータのIDが格納(一意な値)
2	card_id	x	数値型	4	カードID(カードマスタを参照)が格納
3	od_date	x	日付型	8	OD情報の日付が格納
4	kind	x	数値型	4	データの正規性が格納
5	in_time	x	日付型	8	入口の通過時刻が格納
6	out_time	x	日付型	8	出口の通過時刻が格納
7	change_time	x	数値型	4	走行時間(秒)が格納
8	in_tollbooth	x	数値型	4	入口ランプIDを格納
9	out_tollbooth	x	数値型	4	出口ランプIDを格納

株式会社が管理しており、そのデータを西日本高速道路株式会社から入手する料金所を示す。

OD情報の生成処理は、カードID毎に行う。まず、ETCカードID毎に、ETCデータを並び替え、さらに料金所通過時の通過時刻に着目し、昇順で並び替える。この段階で、半月ごとに16分割された各テーブルに格納された利用者毎に、料金所利用履歴が時系列に並んだことになるが、この利用履歴には複数の利用に対応したOD情報が含まれている。これを下記に示す規則に従って、個別のOD情報に分割する。

- ETCデータの料金所種別が「FF入口」もしくは「一般入口」である場合、ETCデータに「入口フラグ」の情報を付加する。
- 「FF出口」もしくは「一般出口」の場合、ETCデータに「出口フラグ」の情報を付加する。
- その他の料金所種別の場合は、ETCデータに何も情報を付加しない。
- 時系列順に「入口フラグ」と「出口フラグ」の組合せを一つのOD情報として生成する。

ただし、「FF入口」や「FF出口」では料金徴収が行われないため、利用者がETC車載器にカードを挿入していない場合や、料金所通過時に挿入していたカードを抜く場合がある。この場合には、「入口フラグ」もしくは「出口フラグ」が付加されないため、前記ルールだけでは、生成に不備が発生する。そのため、「入口フラグ」のみのOD情報や「出口フラグ」のみのOD情報が含まれるように、OD情報生成の規則を定義した。なお、生成したOD情報の構成を表-2に示す。このOD情報を用いることで、例えば、入口の通過時刻、出口の通過時刻をもとに、OD毎の走行時間⁴⁾等を算出することができる。

(3) 乗り継ぎ区間におけるOD情報の結合

阪神高速では、特定の出口料金所から入口料金所まで、一定時間内に通過する場合、通行料金が無料となる乗り継ぎ区間がある。例えば、阪神高速を利用して、「5号湾岸線住吉浜出口」から出て、90分以内に「3号神戸線京橋(西行)入口」から再度入

表 - 3 OD情報における正規性の種別

ID	正規性の種類	説明
1	完全	入口 出口共に特定
2	入不明	入口が特定不可
3	出不明	出口が特定不可
4	出入不明	入口 出口共に特定不可

り、阪神高速に乗り継いだ場合は、「3号神戸線京橋(西行)入口」の通行料金が無料となる。

このような利用については、乗り継ぎ区間前のOD情報と乗り継ぎ区間後のOD情報を合わせて1つOD情報とする。ここで、前述(3)b)のアルゴリズムでOD情報を作成した場合、乗継区間前後でOD情報が2つに分割して生成される。そこで、前記アルゴリズムでOD情報を生成した後に、連続する2つのOD情報の間が乗継区間となる場合については、2つのOD情報を結合する。また、1回の利用で複数の乗り継ぎをおこなう場合も考えられることから、阪神高速道路の現行の乗り継ぎ区間の実態を勘案して最大3区間の乗り継ぎまでを結合した。

(4) OD情報の正規性の特定

本システムで作成するOD情報の中には、前述のように「入口フラグ」のみのOD情報や「出口フラグ」のみのOD情報が含まれる。この場合、出口、もしくは入口が特定出来ない。そこで、入口および出口が特定されているかどうかを判別するために、情報の正規性を区分する。表-3にOD情報の正規性を区分した種別を示す。入口・出口共に特定できた「完全」、入口が特定できなかった「入不明」、出口が特定できなかった「出不明」、入口・出口共に特定できなかった「出入不明」の4種類に区分している。ここで、例えば、出入口の情報を含むOD情報を分析する場合は正規性が「完全」のOD情報を利用し、利用頻度等特に入出口の情報が必要ない分析を実施する場合には、すべてのOD情報を対象に分析を行う。

4. OD情報の分析手法

OD情報の分析手法では、交通量分析の一例として、OD間交通の相関関係、およびOD間の走行時間につ

いて分析を行った。

(1) OD 間交通の相関関係の算出

OD 間交通の相関関係は、生成した OD 情報の入口、出口の相関関係⁵⁾として求める。相関関係は、データマイニングの一種であるアプリアリアルゴリズムによって算出した。

a) アプリアリアルゴリズム

アプリアリアルゴリズムとは、アイテムの集合 I と、トランザクションテーブル D を利用する。式(1)にアイテムの集合と、トランザクションテーブルの関係を示す。

$$\begin{aligned} I &= \{i_1, i_2, i_3, \dots, i_m\} \\ D &= \{t_1, t_2, t_3, \dots, t_n\} \\ t_i &\subseteq I \end{aligned} \quad (1)$$

なお、各要素 t_i をアイテム集合と呼ぶ。そして、個々のトランザクションには、ユニークなトランザクション ID を割当てる。この時、導出される相関ルールは、支持度 (support) および確信度 (confidence) といった 2 つのパラメータによって算出される。これらのパラメータ値は、相関ルールの重要さを表す。相関ルールである $X \rightarrow Y$ の支持度を算出する support($X \rightarrow Y$) は、トランザクションテーブル D 全体に対し、 X, Y 共に含むトランザクションの割合によって定義される。また、確信度 confidence($X \rightarrow Y$) は、 X を含むトランザクションのうち、 Y も含むトランザクションの割合によって定義される。

b) 相関ルール抽出の流れ

相関ルールの抽出は、分析時に、指定された最小支持度 (minimum support) と最小確信度 (minimum confidence) を満足する全てのルールを導出する。なお、相関ルールは、次の 2 つの処理で抽出される。

- ・ 最小支持度を満足するアイテム集合を全て見つける。見つけたアイテム集合は、ラージアイテムと呼ぶ。
- ・ 上記ルールで算出したラージアイテム集合から、最小確信度を満たす相関ルールを抽出する。

相関ルール抽出処理のうち、まず、可能な全てのアイテム集合について、支持度を算出する。このため、アイテム数が増えると、組合せ論的にアイテム集合のバリエーションが増え、それに伴い、計算が膨大となる。そして、最小支持度を越えたラージアイテム集合だけを対象にして、相関ルールの生成を行

う。

(2) 走行時間の把握

OD 情報は、生成時に入口の通過時刻、出口の通過時刻から走行時間を算出する。そして、走行時間の平均値を時刻毎に算出し、利用者の走行時間を把握する。ただし、OD 情報の走行時間の値には、極端に長い走行時間のデータなどが存在する。そこで、生成された OD 情報のうち、極端に長い走行時間のデータや走行時間データの分布において、主な分布から外れているデータは、特異なデータとして扱う。そして、長時間データなどの特異データを除去した上で、走行時間を算出する。

a) クラスタ分析による特異データの除去

特異データは、サービスエリアやパーキングで休憩のため長期停車している車両や、事故や故障で長期停車している車両が相当すると考えられる。これらの特異データは、渋滞が発生したことによって長い走行時間となっている場合がある。そのため、データの分布状況によって除去するかどうかを判別する。特異データの除去判別は、クラスタ分析⁶⁾を使用し、統計学的な手法でグループ化を行う。

クラスタ分析では、クラスタの形成時に、オブジェクト間の非類似度や距離を定義して、凝集または結合ルールに則ってグループ化を行う。本研究では、走行時間、及び走行時刻を距離測定として、ウォード法を利用して各クラスタの凝集または結合を評価した。

ウォード法では、クラスタ (p) に含まれる i 番目の対象を考え、その変量に関する観測値を X_{ji} (p) と表せば、クラスタ (p) 内における分散の和の合計は、式(2)で表される。

$$S_p = \sum_{j=1}^m \sum_{i=1}^{n_p} (x_{ji}(p) - \bar{x}_i(p))^2 \quad (2)$$

今、(p) と (q) を融合してクラスタ (t) を生成する。この時、クラスタ内平方和における合計の増分 S_{pq} は、式(3)で表される。

$$\Delta S_{pq} = \frac{n_p n_q}{n_p + n_q} \sum_{i=1}^n (\bar{x}_i(p) - \bar{x}_i(q))^2 \quad (3)$$

式(2)と(3)の n_p および n_q は、クラスタの大きさである個数を表す。ウォード法は、クラスタ内平方和が、できるだけ小さいことを望ましいと考え、各段階でクラスタの融合による平方和の増分 S_{pq} が、最も小さい (p) と (q) を融合する。そのため、クラスタ (p) と (q) の非類似度 d_{pq} として S_{pq} を用いる。2 つのクラスタ (p)、(q) を融

合して生成されたクラスター (t) と、別のクラスター (r) を融合するときの非類似度 d_{tr} の更新式を式 (4) に示す。

$$d_{tr} = \frac{n_p + n_r}{n_t + n_r} \Delta s_{pr} + \frac{n_q + n_r}{n_t + n_r} \Delta s_{qr} - \frac{n_r}{n_t + n_r} \Delta s_{pq} \quad (4)$$

本手法では、クラスターリングを繰り返し、グループ化された領域に含まれる OD 情報の件数が閾値を下回る場合に除去する。なお、閾値の算出は、画像処理で利用される大津の閾値法⁷⁾を利用した。

b) 走行時間の算出

走行時間の算出は、特異データを除去するために生成したクラスター毎に平均値として算出する。次に、クラスター内に存在する OD 情報の数からクラスター毎の重みを算出する。最後に、クラスター毎の観測値、重みをもとに、式 (5) に示す加重平均で算出する。

$$\overline{X_w} = \frac{\sum_{i=1}^n W_i \cdot X_i}{\sum_{i=1}^n W_i} \quad (5)$$

W_i は、クラスターの重みを表す。また、 X_i は、クラスターの平均値を表す。

5. OD 情報の分析結果

(1) OD 情報の生成結果

本節では、OD 情報生成を行うために、ETC データ

を入力し、OD 情報を生成した。本実験で採用したデータは、2005 年 9 月の ETC データと 2005 年 11 月の ETC データである。入力した ETC データ件数を表 - 4 に示す。また、OD 情報として生成した件数を表 - 5 に示す。

表 - 4 に示すように入力した ETC のデータ件数は、9 月として 31,458,403 件、11 月として 36,057,802 件の合計 67,516,205 件を入力した。また、表 - 5 に示すように、特異データを削除する前段階において、生成された OD 情報において、正規性データの種類の構成比は「完全」が 83.77%、「入不明」が 1.94%、「出不明」が 14.28%、「出入不明」が 0.01% という結果となった。

(2) OD 間交通の相関関係の分析結果

OD 間交通の相関関係は、生成した OD 情報をもとに算出する。本分析では、ある入口情報、出口情報によって構成される OD₁ と、OD₁ とは別の入口情報、出口情報によって構成される OD₂ の相関関係を導出した。相関関係は、OD₁ と OD₂ の最小支持度と最小確信度がパラメータを超えている場合に相関関係として導出される。

最小支持度とは、全 ETC カード ID において、OD₁ と OD₂ が同時に利用されている割合を示す。また、最小確信度とは、OD₁ を利用している ETC カード ID において、OD₂ を利用している割合を示す。なお、周辺地区に住居している利用者は、OD の種類が特定さ

表 - 4 入力した ETC データ件数

入力年月	データ件数
2005年9月	31,458,403
2005年11月	36,057,802
合計	67,516,205

表 - 5 生成した OD 情報の件数

生成年月	OD情報の正規性					構成比			
	完全	入不明	出不明	出入不明	合計	完全	入不明	出不明	出入不明
2005年9月	10,077,007	290,824	1,796,703	1,653	12,166,187	82.83%	2.39%	14.77%	0.01%
2005年11月	11,186,980	195,620	1,821,793	1,508	13,205,901	84.71%	1.48%	13.80%	0.01%
					平均	83.77%	1.94%	14.28%	0.01%

表 - 6 OD 情報より導出した相関関係の結果 (最大確信度順)

順位	OD情報 1		OD情報 2		最小支持度	最大確信度
	入口情報	出口情報	入口情報	出口情報		
1	西宮山口合併	伊川谷 JC 出	伊川谷 JC 入	西宮山口東出	0.009	0.529
2	須磨合併	京橋 (東行) 出	京橋西行	神戸線 (本線) 出	0.012	0.486
3	安治川 (本線) 入	大正西出口	大正西入口	安治川 (本線) 出	0.006	0.484
4	安治川 (本線) 入	北津守出口	北津守入口	安治川 (本線) 出	0.012	0.472
5	松原 (本線) 入	神戸線 (本線) 出	須磨合併	松原 (本線) 出	0.008	0.462
6	松原 (本線) 入	文の里出	文の里	松原 (本線) 出	0.007	0.461
7	西宮入	神戸線 (本線) 出	須磨合併	西宮出	0.012	0.458
8	西宮 JCT 入	神戸線 (本線) 出	須磨合併	西宮 JCT 出	0.011	0.402
9	守口 (本線) 入	南森町出	扇町	守口 (本線) 出	0.004	0.359
10	摩耶西行	神戸線 (本線) 出	須磨合併	摩耶 (東行) 出	0.008	0.332

れる事と、ODのパターンが膨大となる関係から、最小支持度と最小確信度の値を小さくして、分析を行った。本実験では、相関関係を導出するパラメータとして、最小支持度に0.0025を最小確信度に0.25を与えて、OD情報の相関を導出した。

表-6にOD情報より導出した相関関係の結果を示す。これによると、導出した相関関係は、入口情報が「西宮山口合併」、出口情報が「伊川谷JCT出」と入口情報が「伊川谷JCT入」、出口情報が「西宮山口東出」のOD情報に対して、最も高い相関関係を得られた。また、入口情報が「須磨合併」、出口情報が「京橋(東行)出」と入口情報が「京橋西行」、出口情報が「神戸線(本線)出」のOD情報に対して、2番目に高い相関関係として得られた。

本分析におけるOD間交通の相関関係導出の結果は、関連性の高いOD₁とOD₂を導き出す事ができた。本結果として導き出されたOD₁とOD₂は、同時に利用されている事を表すため、OD₁における交通量増加に伴うOD₂の将来交通量の予測に必要な外的要因などに利用できると思われる。

(3) 走行時間の分析結果

走行時間把握の実証実験では、まず、生成したOD情報の正規性種別が「完全」のデータから走行時間を算出する。走行時間算出の前処理として、特異データの除去を行う。例えば、入口Aから出口BまでのOD情報のうち、特異データ除去前のOD情報の分布を図-3に示す。OD情報の分布を確認すると、データにばらつきがあることが確認できる。それらの情報は、走行時間算出時に雑音となってしまうため、クラスター分析を用いて特異データを除去(図-4)

する。生成したOD情報のうち、表-5で示す正規性が「完全」の情報において、特異データの除去を行った件数を表-7に示す。これによると、OD情報には、特異データが3.37%存在しており、それらを除去することで、生成したOD情報の80.49%を対象として、走行時間を分析した。

走行時間の分析は、時刻別平均走行時間、OD別平均走行時間を算出し、走行時間の差が大きい順に出

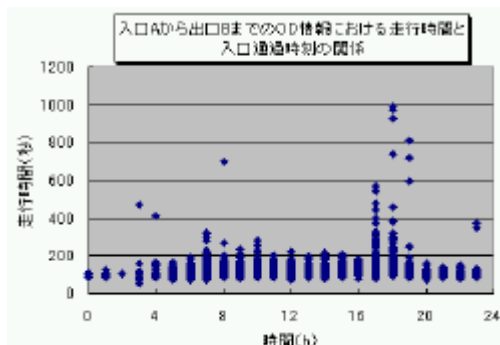


図-3 特異データ除去前のOD情報の分布

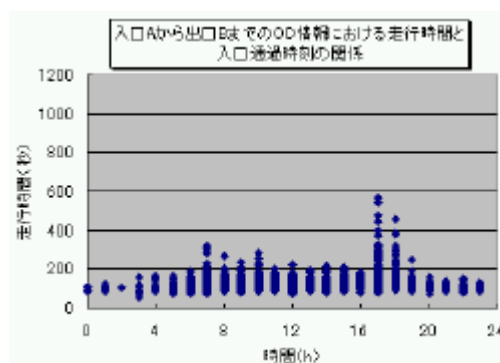


図-4 特異データ除去後のOD情報の分布

表-7 OD情報のデータ件数および特異データ件数

対象年月	OD情報のデータ件数			構成比	
	除去前	特異データ	除去後	特異データ	除去後
2005年9月	10,077,007	387,090	9,689,917	3.84%	96.16%
2005年11月	11,186,980	444,489	10,742,491	3.97%	96.03%
			平均	3.91%	96.09%

表-8 OD情報による平均走行時間(走行時間差順)

順位	OD情報		時刻 (h)	時刻別平均 走行時間	OD別平均 走行時間	走行時間 の差
	入口情報	出口情報				
1	堺(本線)入	扇町出	17	75.83	19.34	56.50
2	須磨合併	法円坂出	18	82.11	43.07	39.04
3	出島	鳴尾浜出	7	50.81	17.63	33.17
4	池田(本線)入	南港中出	17	54.35	28.74	25.61
5	須磨合併	長田出	16	71.83	47.32	24.51
6	松原(本線)入	豊中南(北行)出	8	43.71	23.96	19.74
7	西長堀	神戸線(本線)出	15	55.81	36.46	19.35
8	須磨合併	水走出	17	67.82	48.54	19.29
9	水走入	神戸線(本線)出	9	65.39	46.44	18.94
10	須磨合併	甲子園浜出	14	61.54	42.76	18.78

単位 分

力(表-8)した。これによると、入口情報が「堺(本線)入」、出口情報が「扇町出」の17時台の時間帯に、時刻別平均走行時間とOD別平均走行時間の差が最も大きい結果が得られた。また、入口情報が「須磨合併」、出口情報が「法円坂出」の18時台の時間帯に、2番目に大きい時間差が発生している結果が得られた。

本分析における走行時間の算出は、長時間停車している車両を特異データとして削除したため、精度の高い時刻別の走行時間が把握出来る結果となった。本結果として算出された走行時間は、時間別に算出しているため、通常よりも遅い走行時間となる原因を解明し、より快適に利用できる環境作りに役立てることができると考えられる。

6. おわりに

本研究では、精算処理システムから取得したデータをもとにOD情報を生成した。その方法として、OD情報分析システムを開発して、暗号化されたETCカードIDを数値に変換し、データのダウンサイジングを行い、データベースへの操作速度の向上を図った。そして、乗り継ぎパターンを考慮せずに余分に分割された情報に対して、OD情報の結合処理を行った。

本システムでは、生成したOD情報の中から、入口と出口ともに特定できる84%のデータを用いて、交通利用の相関性が高いODを抽出した。その結果、「西宮山口合併」～「伊川谷」CT出」を利用する交通は「伊川谷」CT入」～「西宮山口東出」を利用する割合が高いこと等の利用特性を明らかにできた。さらに、OD間の所要時間について、前記データから、長期停車している車両等をクラスター分析を用いて特異データとして判別し、それらを除去したデータにより、時間帯毎の所要時間分布を明らかにできた。

今後、ETCで得られるデータを用いて、応用システムを構築し、以下に示す分析を実施することも考えられる。

・各種車種特定による混入率の計測

ETCでは、車種によって利用料金が異なる。そのため、ETCデータ内には、車種を特定するデータが含まれている。そのため、それらの情報を利用し、大型車混入率や各種車両別混入率を算出することが可能となる。

・料金割引の効果予測

各高速道路では、ETCを利用することで、様々な割引サービスを提供している。しかし、サービスを提供したとしても、その効果を分析することは困難である。ETCでは、割引条件によって利用

料金が異なる。そのため、ETCデータ内には、割引コードを特定するデータが含まれており、それらの情報を利用し、車両の利用履歴から、割引料金を利用したかどうかを判別し、効果予測へと繋げることができる。

参考文献

- 1) 井坪慎二, 塚田幸広: 情報機器の道路交通調査への適用に関する検討, 土木技術資料, 土木研究センター, Vol.47, No.8, pp.56-61, 2005年.
- 2) 堀場庸介, 松本幸正, 松井寛, 高橋政稔: プローブデータに基づく推定経路交通量への観測誤差の影響分析と推定経路交通量の更新手法, 土木計画学研究・論文集, 土木学会, Vol.22, No.3, pp.495-506, 2005年.
- 3) 平出貴也, 森俊介, 堂脇清志: 首都高速道路におけるETC(自動料金収受システム)の導入効果の解析, システム・情報部門学術講演会講演論文集, 計測自動制御学会, Vol.2001, pp.409-412, 2001年.
- 4) 上野秀樹, 大場義和, 島田重人, 桑原雅夫: ETC応用の所要時間算出ロジックと応用システム, 電気学会道路交通研究会資料, 電気学会, Vol.RTA-00, No.11-20, pp.41-45, 2000年.
- 5) 岡田孝, 元田浩: データマイニング - 相関ルールとその周辺, オペレーションズ・リサーチ, Vol.47, No.9, pp.565-571, 2002年.
- 6) Michael.R.Anderberg, 西田英郎: クラスタ分析とその応用, 内田老鶴圃, 1988年.
- 7) 大津展之: 判別および最小2乗規準に基づく自動しきい値選定法, 電子通信学会論文誌, 電子通信学会, Vol.63, No.4, pp.349-356, 1980年.

(2006.5.19受付)