

## ETCを利用した交通管理用データの抽出方法について

狩野 雄一

日本道路公団東京管理局西局 〒192-8648 東京都八王子市宇津木町231

E-mail: Yuuichi.Karino@jhnet.go.jp

あらまし 本論文は、ETCシステム（Electronic Toll Collection System）で収集される履歴データを用いた交通管理への活用の可能性を探るにあたり、収集データから有効な統計データを抽出するための方法について記述したものである。代表的な出口料金所におけるサンプルデータから料金所区間所要時間を算出し、その統計データとしての特性をデータ分散、データ変動、料金所区間距離の観点から考察した。さらに履歴データの中の車両行動パターンに影響するような属性項目に着目した抽出を行い、抽出前後の結果を比較した。

**キーワード** ETC、交通管理、所要時間、データ抽出、フィルタリング

## The extraction method of the data for traffic management used ETC

Yuuichi Karino

Japan Highway Public Corporation Tokyo Operation Bureau

231,Utsuki-cho,Hachioji, Tokyo, 192-8648,Japan

E-mail: Yuuichi.Karino@jhnet.go.jp

**Abstract** This paper describes the method for extracting effective statistics data from collection data of ETC (Electronic Toll Collection System). Its aims at exploring the possibility of the practical use to the traffic management which used history data. The passing time required was computed from the data in a typical tollgate, and it considered from a viewpoint of data distribution, data change, and tollgate section distance. And its attention was paid to the attribute item related to the vehicles action pattern contained in history data.

**Keyword** ETC, traffic management, passing time required, data extraction, Data filtering

### 1.はじめに

ITS (Intelligent Transport Systems) サービスの先駆けである自動料金収受システム（Electronic Toll Collection System、以下 ETC という。）は平成 13 年 3 月末の運用開始から 2 年半近くが経った。その間 ETC 利用車両は増加を続け、平成 15 年にはさらなる利用可能料金所の拡大と社会実験に伴う割引策が実施され、より一層の普及が見込まれている。

ETC は料金所に設置された路側アンテナと ETC 利用車両に搭載された車載器との間で、予め登録された車載器を特定するための ID 情報、課金種別を区分するための車種情報等、入口料金所の通過情報（対距離料金収受の場合）を相互通信することにより実現されるものである。そのとき路側アンテナが受信するデータ（以下、ETC データという。）は料金収受の利用ばかりではなく、様々な利活用の可能性を秘めている。

第一に路側アンテナにおいて車載機との通信履歴の存在検知や ID 情報の収集することで、従来の車両感知器や AVI(Automatic Vehicle Identification)システムの延長にある情報収集装置としての利用ができる。第二に車載器または IC カードに格納された過去の通過情報を参照することにより走行経路、通過区間の所要時間などの通行履歴を収集することで、車両の走行状況を把握することができる。この第二の方法は、従来の磁気式通行券を利用した料金収受システムにおける通行履歴（通行券に書き込まれた入口データ）の集計と同じ概念である。しかし、ETC データは磁気式通行券より多様な属性情報を含んでおり、この属性情報によるフィルタリングを実施することにより、目的に合致した有効データを演算処理することなく抽出することが可能である。また、予めフィルタリングを行うことはデータが必要最小限となるため、データ処理性能の

確保の観点からも有利である。

これらのことから ETC データを利用した新たな統計が期待されているところであり、本論文は実際の ETC データを用いてこれらの統計に有効となるデータ抽出方法の可能性について検討したものである。

## 2. E T C データ収集

### 2.1. 収集方法

ある車両の通行履歴を把握するためには、その車両の過去の通過地点ごとのデータを集めて、トレースしなくてはならない。しかし、ETC データは図 1 に示すように、対距離出口料金所において、直前の入口料金所の通過情報が把握できる。この方法は入口料金所の通過情報が、出口料金所に至るまで把握できないためにデータ集計の遅延するという問題はあるものの、一箇所のアンテナで広域の通過情報を把握できるため、効率良くかつ経済的な方法である。本論文においてはこの方法により収集したデータを用いている。

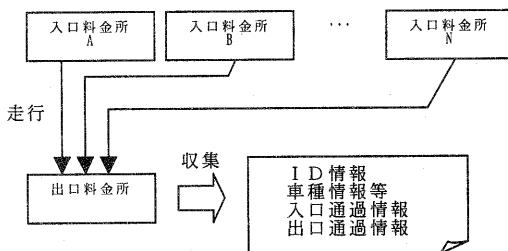


図 1 E T C データの収集概念

### 2.2. 利用データ

料金所で通信される E T C データは大きく表 1 のように分類できる。

- ① 車載機等を特定するためのユニークな識別情報である「一意性データ」
- ② E T C を提供するための車載機や契約に関わる区分するための「属性データ」
- ③ E T C 利用を確実に行うための機器動作や処理記録のための「証跡データ」

表 1 主なデータ項目

分類	データ項目例
一意性データ	車載器 ID、ナンバープレート等
属性データ	車両コード、利用契約種別、履歴等
証跡データ	処理結果、機器動作記録等

ETC データの利用にあたっては、個人情報保護の観点から個々を特定しないように加工されたデータを用

いることが必要である。「一意性データ」はその一意性から取扱いが難しい。「証跡データ」は ETC の運用管理的な侧面に使用するものであるから本論文の目的と性格が異なるものである。「属性データ」は個々のデータではなく類似グループを特定するものであるから、本論文においては容易に加工ができる本データを利用する可能性を検討した。

## 3. 区間所要時間の算出

### 3.1. 収集サンプル

以下のとおり、代表的な料金所区間ににおけるサンプルデータを収集した。

- ① 収集期間  
平成 15 年 6 月 5 日（木）及び 6 月 5 日（土）
- ② 収集対象  
0:00~24:00 に出口料金所を通過した ETC 車両
- ③ 料金所区間  
表 2 に示す比較的 E T C 利用台数が多く区間距離 100km 程度までの料金所区間を選定した。

表 2 収集料金所区間とその区間距離

入 口 料金所名	出 口 料金所名	区間距離 (Km)
所沢	新座	9.4
		39.4
	本庄児玉	69.6
	渋沢伊香保	103.4
	狭山日高	36.8
横浜青葉	東京	13.3
		35.1
	沼津	103.6
	柏	10.8
谷和原	三郷	19.1
	水戸	82.0

### 4. 区間所要時間の算出方法

#### 区間所要時間

$$= \text{出口料金所通過時刻} - \text{入口料金所通過時刻}$$

### 3.2. データ分布

収集データの代表例として東松山 - 新座間における収集データから算出された区間所要時間の分布の 6 月 5 日（平日）データを図 2、6 月 7 日（休日）データを図 3 に示す。グラフは収集された個体の全データデータに 1 時間毎の区間所要時間の平均値をプロットしたものである、図のパターンにより休日はデータ分散が大きいことが判る。これは、休日は渋滞、道路休憩施設（サービスエリア及びパーキングエリア）への立ち寄りなどの要因から所要時間にばらつきが生じているためと考えられる。

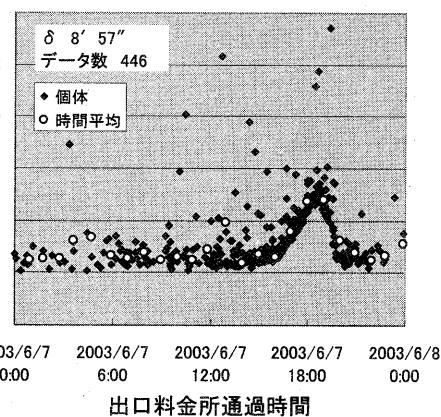
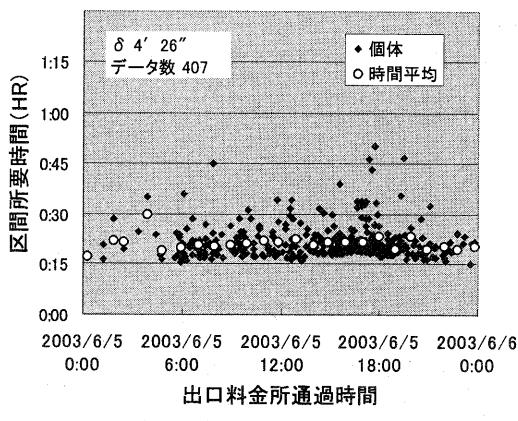


図3 東松山－新座間所要時間分布  
(2003年6月7日土曜日)

### 3.3. データ変動特性

データの分散を考察するにあたり、収集データの標準偏差  $\delta$  求め、式1に示す計算式により変動率  $\rho$  を求めた。また、データの分散の程度は区間距離に依存する関係にあると考えられることから、変動率と区間距離との関係を考察した。

$$\text{変動率 } \rho = \frac{\text{標準偏差 } \delta}{\text{日平均値 } A} \quad (\text{式 } 1)$$

図4に今回収集した料金所区間における変動率と区間距離の関係を示す。ただし、図3に示すような交通集中時の分布においては、変動率の原因要素が平常時と異なると思われることから算出の除外とした。

図4より以下のことが言える。

- ① 平日より休日において変動率が高い

### ② 区間距離が大きくなると変動率が高くなる

これは一般的に休日の道路利用者はレジャー等のマイカーによる自由行動傾向の強いばらついたデータ傾向であり、一方、平日は通勤、流通、営業活動などによる定時性、反復性が強いデータ傾向であるからと言える。また、区間距離の増加と共に変動係数が増大するのは、道路休憩施設などで休憩時間が多く含まれるようになるからと推測される。

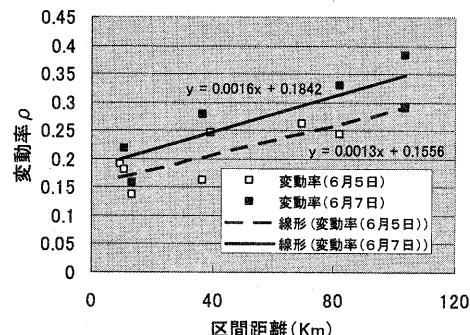


図4 変動率と区間距離の関係

### 3.4. 属性データによる抽出

収集されるデータの精度を高めるためには、前述のようなデータの変動率が少ないような有効性の高いデータを抽出することが必要になる。今回、データの有効性を高めるひとつの方法として以下の仮定を立てて、その仮定に基づく有効データの抽出に効果的と思われる属性データを抽出条件とした。

#### 3.4.1. 利用頻度による抽出

##### ① 仮定

利用頻度の高い車両は通勤、流通、営業活動などによる定時性、反復性が強い変動要因の少ないデータである

##### ② 抽出条件

利用頻度の指標としてETCデータの履歴情報に含まれる「処理一連番号」を使用する。なお、サンプル数を多く確保する目的からに処理一連番号100以上とした(抽出条件1)

図3における全データに、この条件により抽出したデータをプロットしたものを図5に示す。

この例において偏差は全データの場合が8分57秒なのにに対し、抽出条件1による抽出データは8分30秒と若干改善されている。

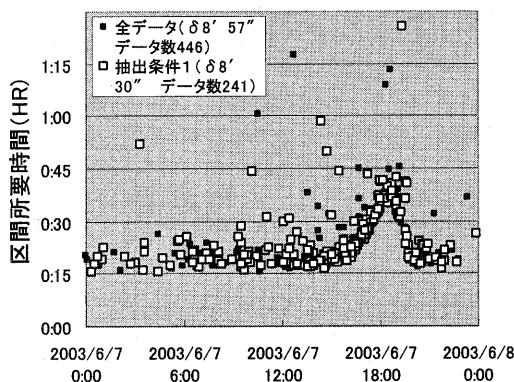


図 5 東松山－新座間所要時間分布（抽出条件1）  
(2003年6月7日土曜日)

次に抽出前と抽出後のデータ変動を比較するために、3.3と同様に式1による変動率を求め、区間距離との関係を考察した。6月5日及び7日における抽出前後のデータ変動の比較をそれぞれ図6、7に示す。

図6及び図7により以下のことが推測できる。

- ① 平日、休日共に利用頻度による抽出にはデータ変動の低減効果がある
- ② 平日においては、もともとマイカーユーザが少ないことから利用頻度による抽出による効果は休日よりも少ない
- ③ 休日において、区間距離が短い場合には交通集中のため走行時間に差が出にくく、区間距離が大きい場合はマイカーユーザによる道路休憩施設への立ち寄り率が高くなることから、区間距離が大きくなればなるほど利用頻度による抽出によるデータ変動の低減効果が大きく現れている。

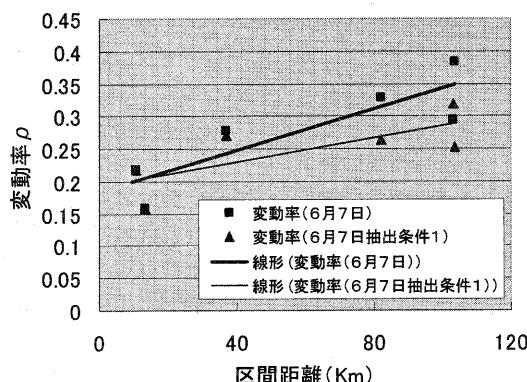


図 6 変動率と区間距離の関係〔抽出前後比較〕  
(2003年6月5日木曜日)

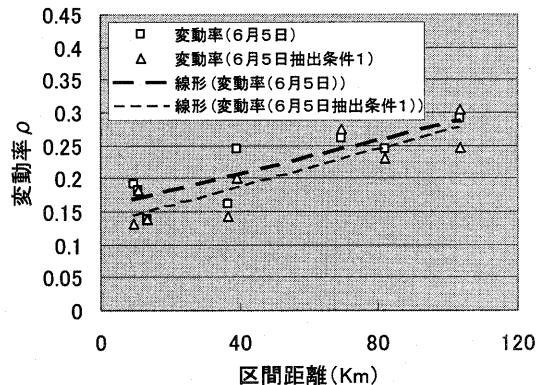


図 7 変動率と区間距離の関係〔抽出前後比較〕  
(2003年6月7日土曜日)

なお、図4、図6及び図7において除外した交通集中時におけるデータについても同じグラフを作成してみると図8のようになる。

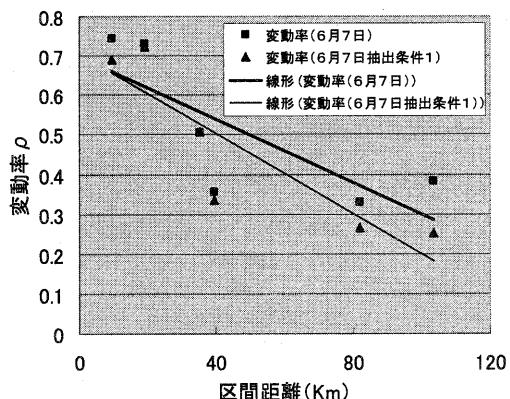


図 8 変動率と区間距離の関係〔交通集中時の抽出前後比較〕(2003年6月7日土曜日)

交通集中時の場合のデータは前述したように走行時間に差が出にくく、区間距離が大きくなればなるほど変動率が小さくなる傾向にある。図6及び図7と図8の傾向の相違から言えることは、データ分散の扱いにおいて平常時と交通集中時を区別して考えるべきことを意味している。その区別すべきデータ傾向の分岐点は変動率0.3程度であった。

### 3.4.2. 車両の用途による抽出

#### ① 仮定

利用頻度の高い車両は通勤、流通、営業活動などによる定時性、反復性が強い変動要因の少ないデータである

## ②抽出条件

利用頻度の指標として車種コード等の車両用途が営業用途のものとする（抽出条件 2）  
この条件により抽出したデータを図 9 に示す。

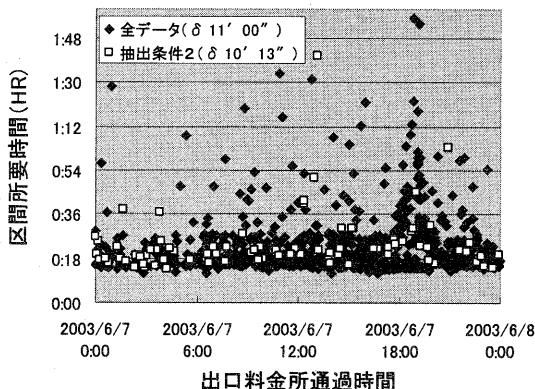


図 9 厚木-東京間所要時間分布(抽出条件 2)  
(2003 年 6 月 7 日土曜日)

抽出条件 2 による抽出の良いところは抽出条件が極めて簡単なところである。図 9 に示すように、この例において偏差は全データの場合 11 分に対し、抽出条件 2 よる抽出データは 10 分 30 秒であり、抽出条件 1 の場合と同様に若干改善されている。しかし、全車両に占める営業用途車両の割合が小さいことから、抽出条件 1 の抽出の場合より得られるデータ数が少なくなることが欠点である。

### 3.5. データ有効性の考え方

#### 3.5.1. ETC 利用の状況

日本道路公団の道路における平成 15 年 5 月の月平均 ETC 利用交通量は約 285,000 台／日、平均利用率は 6.2% である。この数字は全国の平均値であり、ETC 車両の通行が多い料金所では 10% を超える利用率となっており、また、朝夕の通勤、業務始業、終業時間においては、一時的に 1 時間交通量 780 台（5 分間交通量 65 台）（6 月末実績）程度の利用状況となっている。また、首都圏の通行頻度の高い料金所においては、特定の料金所区間の ETC 走行台数が全 ETC 流入台数の 20% 以上を占めるような区間もある。つまり、利用率の高い料金所区間においては 65 台の 20%、10 台強が走行する状況になっている。

#### 3.5.2. 有効なデータ数について

データを正規データとみなした場合に、データ信頼度 95% を得るために必要なデータ数  $n$  は次式で示される。

$$n \geq \left( \frac{1.96\sigma}{\varepsilon} \right)^2 \quad (\text{式 } 2) \quad \varepsilon : \text{信頼区間}$$

今回収集、抽出したデータの平均日交通量及び 5 分間交通量を区間距離別にまとめて表 3 に示す。

表 3 収集データの平均日及び 5 分間平均交通量  
(抽出条件 1 の場合)

区間距離 (Km)	区分	抽出前 (台)	抽出後 (台)
0~30	日交通量	1019.9	638.6
	5分換算交通量	3.5	2.2
30~100	日交通量	425.8	245.5
	5分換算交通量	1.478	0.852
100~	日交通量	134.3	66.0
	5分換算交通量	0.466	0.229

表 3 における値と  $\varepsilon$  を 5 分とし式 1 より計算した値を照らし合わせてみると、データ偏差を平常時は 4~5 分程度、交通集中時は 8~9 分程度と想定したときに、それぞれ 3~4 台、10~13 台程度の ETC 台数があれば十分であることになる。これは現在の利用台数の多い料金所区間においては既に実現または近い将来に実現すると考えられる数字である。

しかしながら、本データは全交通データの中から ETC に限定し、さらに作成的な抽出条件により抽出されたデータであるから、高速道路の全般のデータ特性を代表するものではない。結果の有意性について確認するために今後、ETC 以外を含む全交通量を考慮した十分な実データに基づくデータ検証が必要である。

#### 3.5.3. 抽出条件について

今回の結果から利用頻度及び車種用途による抽出のいずれの場合も偏差の減少が確認できた。しかし、それらの抽出条件がその車両の利用パターンを十分に表すものではないことから、より的確な利用パターンを抽出できる抽出方法を考慮する必要がある。例えば、今後蓄積される長期データから利用パターン因子を見つけ出して利用することや、場合によっては今回触れていない一意性データを統計により匿名化するような手法を用いて定期運行などの定時性を把握して利用するなどが考えられる。後者は路線バスや交通巡回車両を利用することも考えられる。

#### 3.5.4. 時間変動について

本論文では ETC データ特性の概略をつかむことを目的としたことと時間的な制約から一日単位のデータを考察した。本来、交通データは時々刻々と変化するものであり、その都度の交通状況の断面で切り出した

時間帯に分解して考察を行う必要がある。

参考に6月5日における東松山－新座間のデータから1時間毎に変動率を算出したところ図10のようになつた。これから交通が集中する朝夕の時間帯において変動率が高いことが判る。

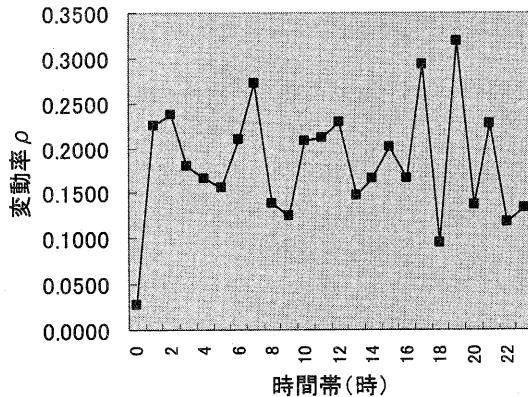


図10 東松山－新座間時間帯別変動率  
(2003年6月5日木曜日)

#### 4. おわりに

交通データに関する統計的研究は長い歴史により築き上げられ、既に様々な試みがされている。そのような背景であっても、ETCデータを用いて時間、環境、社会情勢により変化する挙動、また人が車を運転するという行動心理学的な挙動を把握することは非常に有意義なことと考える。また、ETCデータから利用傾向の想定できるような属性または行動のプロファイルとでも言うべき挙動パターンを見つけ出して、多様な目的に応じた統計を行うことは十分な価値な生むものであろう。さらに、それが少数データであってもプローブ情報として全体の交通状況を把握するための指標データとなる可能性もあると考えられる。

今回、ETCデータの交通管理データへの活用への可能性について十分な感触をつかめたと感じる。さらに、今後、ETC利用台数及び利用率が増加し、利用者層が広がることよりそれはさらに確実なものとなると考える。

#### 参考文献

- [1] 大場義和、上野秀樹、桑原雅夫：「磁気式通行券データを利用した高速道路旅行時間演算方法に関する研究」（東京大学生産技術研究所研究速報 1999年2月）