

交通機関の自動料金徴収のための分散データ配置方式

田 中 幹 夫[†]

鉄道旅客輸送や有料道路交通等の交通機関の利用料金徴収を利用者のユーザ ID 情報のみを用いて自動的に行うシステムは、従来の、利用者側に各種情報を携帯させる方式に比べていくつかの利点を有している。このシステムの実現を目的とした分散データ配置方式に関して論ずる。一般に利用料金徴収の自動化のためには、利用開始時点の情報と利用終了時点の情報とから利用区間の確定、および料金計算等を行う必要がある。利用終了時点の処理を実用的な処理時間内に収めるためには、その処理に必要とされるデータの複製を予想される利用終了地点に分散して配置することによって処理のレスポンスタイムを一定値以内に収める方式が考えられる。この方式の現実化のためには、様々な交通ネットワークの形態に応じた、利用者移動に関する複雑な解析が必要となる。ここではその基礎検討のために、単純な線形の交通ネットワークと利用者移動モデルを対象とした場合の実現方式としての分散データ配置方式を提案し、その性能とコストの評価を行った結果を示す。そして、システムの適用可能性、課題を述べる。

Decentralized Data Allocation for Automatic Toll Collection Systems for Transportation Services

MIKIO TANAKA[†]

For the purpose of an automatic toll collection systems for linear transportation services, a decentralized data allocation method in distributed communication systems is discussed. A simple model to formulate the movement of users (passengers or vehicles) along a linear transportation network is proposed as a first step. Then the merits of user-ID based toll collection system is introduced. For the purpose of this user-ID based toll collection system, the allocation of replicated data over a transportation network is required in order to reduce processing time at exit nodes. Some methods to implement this system and methods to evaluate their costs are studied.

1. はじめに

交通機関（鉄道輸送、道路交通等）の料金徴収業務は、従来、紙による切符類を使用して行われてきた。この業務の自動化のために各種のデータキャリア（その内部に必要情報を登録し、外部から情報のリード/ライトが可能であり、利用者側で携帯可能な小型の装置）を用いたシステムの開発、試行、一部実用化が進められている。現在試行されている方式は利用者（乗客、自動車等）に携帯させるリードライト可能のデータキャリア（IC カード等）を使用したものが主流である。本論文では、それらの方式とは異なり、利用者に関する ID 情報のみを用いるシステムに関して、そ

の意義、利点を示し、単純な線形の交通ネットワークと利用者移動モデルを対象とした場合の実現方式としての分散データ配置方式を提案する。そしてその性能とコストの評価を行った結果を示す。

2. システムの対象と目的

2.1 交通機関の料金徴収自動化システム

交通機関（鉄道輸送、道路交通等）の料金徴収自動化のためには、その利用者の利用区間を特定しなければならない。利用区間を特定することによって料金表等の情報から利用者に課すべき使用料金を計算し、それを利用者のストアードフェア（利用者が、あらかじめ交通機関の提供者に対して払い込んでおく金額）の残額情報や銀行口座等から差し引かなければならない。また、定期券を設定している利用者であれば、その利用者の定期券情報（設定区間、設定期間）を参照して利用の可否や追加料金徴収の必要性有無を判断し

[†] 財団法人鉄道総合技術研究所 輸送システム開発推進部情報システム担当グループ
Information Systems Group, Railway Technical Research Institute

なければならない。これらの一連の処理を利用区間が確定する時点、すなわち利用終了地点（鉄道輸送ならば駅の改札出口、有料道路ならば出口側ゲート等）にて行い、その結果を利用者に示すことが必要である。駅の改札出口では、定期券設定期間切れやストアードフェア残高不足の場合に、改札出口に設けられたドア（小扉）を閉めて利用者を物理的に停止させる必要があるかもしれない[☆]。

このような料金徴収自動化のために、近年では各種のデータキャリアを利用したシステムの開発、試行、そして一部実用化が進められている。現在試みられている方式は利用者（乗客、自動車等）に携帯または装着させるリードライト型のデータキャリアを使用したものが主流である¹⁾。現在は接触読取りを行う磁気カード（たとえば、JR 東日本のイオカードや有料道路のハイウェイカード等）が広く普及しているが、最近はより大きなメモリ容量を持ち、処理能力を備えた IC カードも使用され、さらには電波等の通信媒体を使用した非接触読取りを行う IC カードやデータキャリアを使用することによって、利用者の使い勝手を向上しようという試みがなされている^{2)~5)}。この場合には、現行の接触式カードのように利用者は切符類やカードを取り出して改札口の機械に挿入する必要がなく、読取り装置（小型のアンテナ）の近傍にかざせばよい。また、自動車の場合には前記のデータキャリアを自動車のフロントガラス等に取り付けることにより、出入口ゲートのアンテナの近傍をノンストップで通過できるといった、利用者にとっての大きな利便性の向上がある^{6),7)}。

2.2 料金徴収自動化システムの各種方式

料金徴収自動化システムのために必要な情報を保持/参照する方式は、大きく分けて、以下の2種類の形態が考えられる。

①利用者自身が情報を携帯する方式（図1）

②交通機関側が情報を保持する方式（図2）

①の形態では図1に示すように、利用者に関する情報（定期券情報、ストアードフェア情報等）が利用者の携帯または装着するデータキャリアの内部に蓄えられている。そして入口にて、交通機関側は利用開始に

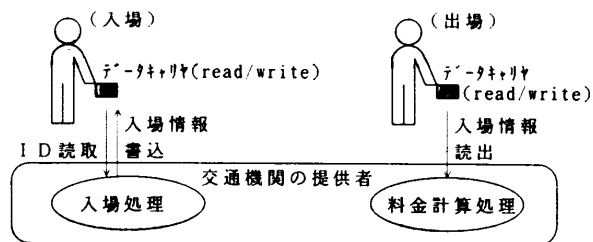


図1 利用者自身が情報を携帯する方式
Fig.1 Method of user carrying information.

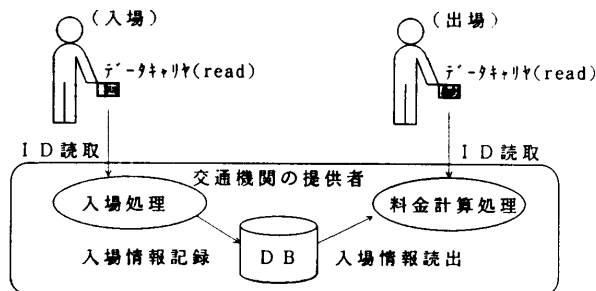


図2 交通機関の提供者が情報をデータベース内に保持する方式
Fig.2 Method of transporter storing information.

関する情報を特にシステム内部には取り込まず、利用者の保持するデータキャリア内に書き込みを行う。そして出口では、そのデータキャリア内に書かれた情報を読み出すことにより利用開始箇所等の情報を知り、料金表等を参照することで料金計算等、必要な処理を行う。前節で述べたような現在多く試みられている方式は、利用者の携帯する磁気カードや IC カードに利用開始に関する情報を書き込むものであり、この①の形態に該当する。

②の形態は図2に示すように利用者に関しては、その ID（ユーザを特定する固定情報）だけでシステムを実現でき、何ら、情報を利用者側に書き込む必要はない。しかしシステム側に、処理速度やデータベース容量の面で大きな能力を要求するため、実用化されているケースは、スキー場のリフト輸送での料金徴収等、小規模な場合に限られている。これらは利用終了点が1カ所、あるいは数カ所に限られており、データベースの参照が高速で容易な場合である。

これに含まれる、やや特殊な形態として「交通機関側が情報を利用者ととも移動させる方式」（図3）もあり、一部のバス等、小規模な交通機関に限って使用されている。これは、図3に示すように、たとえばバスに乗り込んだお客の情報（乗車停留所名等）を、そのバス内の記憶装置に保存しておき、降車時には、そのデータを参照して料金計算を行う方法である。これ

☆ 利用者の完全なクレジット（信用）形式の支払いであり、利用開始/終了の実績を適宜集めて後日にデータ処理することで区間を確定し、利用者に対して料金請求を行う方法もありうる。この場合には利用区間や料金の確定処理を利用終了時点に行う必要はない。しかし、プリペイド（ストアードフェア）形式の支払いへの対応や、クレジット利用者に対する料金の明示処理や与信チェック等のために、一般的に利用終了時点で何らかの処理が必要と考えられている。

表1 実現可能なシステム等に関する2方式の比較

Table 1 Comparison of 2 methods (user carrying information vs. transporter storing information).

方式	①利用者自身が情報を携帯する方式	②交通機関側が情報を保持する方式
概要	交通機関への入場情報等はユーザの所持するデータキャリア内に持つ。	交通機関への入場情報等は交通機関の保持するデータベースに持ち、利用者はユーザIDのみを与える。
実例	現在試みられている主流の方式 (接触式データキャリアは実用済)	実用例はわずか (小規模な輸送サービス)
実現可能なシステム	・リードライト型データキャリア (接触/非接触) を使用した自動料金徴収システム	・リードライト型データキャリア (接触/非接触) を使用した自動料金徴収システム ・リードオンリー型データキャリア (接触/非接触) を使用した自動料金徴収システム ・キップレス (利用者認識技術と組み合わせてキップ省略) システム
システム実現上の課題	・利用者の保持するデータキャリアおよびデータキャリア読取装置の負荷大 (機能複雑), コスト大	・交通機関側の情報処理システムの効率的経済的な実現

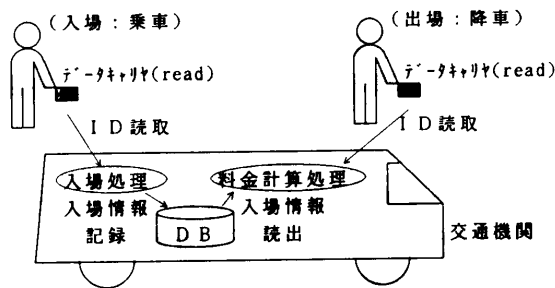


図3 交通機関の提供者が情報を利用者とともに移動させる方式
Fig.3 Method of transporter carrying information.

が①の形態を採用していない理由は、乗車時の処理時間 (データキャリアへの書き込み時間) を節約する狙いが大きい。この形態では利用終了の生じる箇所が決定される必要がある (この例では「乗ったバスから必ず降りる」ため、利用終了箇所が一意に決まる)、一般の鉄道輸送や自動車交通での採用は難しい。

2.3 ユーザIDベースの料金徴収方式

前節で述べた②の方式のように、利用者のユーザID、すなわちリードオンリー情報のみで (リードライト型のデータキャリアを使用せずに) 実現するシステムは、従来試みられてきた①の方式と比較して以下のような利点を持っている。

- (1) データキャリアおよび読取り装置は利用者のIDを読み取るのみで充分であり、技術的、構造的に単純化され、経済性、安全性が向上する。
- (2) 入出場 (改札等) 時点でのトランザクションが単純となり、一般に所要時間は減る (書き込みの場合、書き込み直後に再読出しを行ってチェックする等、一般に通信シーケンスが複雑である⁸⁾)。
- (3) データキャリアへの書き込み時のエラーによる情報不整合等のトラブルがない (無線等による非接触書き込みの際は、通信中断等に起因するエ

ラーの対策が必須で、既存の方式では、このために書き込みエリアの2重化等、複雑な対策を講じている場合が多い)。

- (4) バーコードのように、非常に安価ではあるが、光学的な方式で原理的に書き込みが困難なデータキャリアも使用できる可能性がある。
- (5) 現時点では技術的に困難であるが、将来的にはデータキャリアを使用せずに人間そのものの認識結果 (指紋、掌紋、声紋、容姿等) をユーザIDとして使用する方式が使える。この場合には利用者に携帯させるものはまったく不要となる。

現在試みられている自動料金徴収システム、特に利用者側の利便性に関して格段に優れている非接触データキャリアを使用したシステムの実現のためには、利用者が保持するデータキャリアのコストが大きな障害となっている。②の方式は、上記の(1)~(4)にあげた点から、このコスト面の障害を大きく軽減できる可能性を持っている。また、(5)のように、利用者の利便性を根本的に向上させる形態のシステムは、①の方式では不可能であり、②によって初めて可能となるものである。実現可能なシステム等に関する両者の比較を表1に示す。

このような②の方式を、利用者側に負担させる情報がユーザIDのみであることから、ここでは「ユーザIDベースの料金徴収方式」と呼ぶ。これに関する基礎的な検討が本論文の主題である。この「ユーザIDベースの料金徴収方式」は、上記のような優れた特徴を持っているが、前節で述べたように交通機関側のシステムに対して、処理速度やデータベース容量の面で大きな能力を要求することが定性的に予想されるため、

☆ ただしバーコードのようなデータキャリアを使ったシステムの場合、その捏造や複写による不正利用が容易といった弱点に関する対策が重要である。

現状で実用化されているケースはほとんどない。また、定量的な分析を行うためには、様々な交通ネットワークの形態に応じた、利用者移動に関する複雑な解析が必要となる。本論文ではその基礎的な検討のために、線形の交通ネットワークと単純な利用者移動モデルを対象とした場合の分散データ配置方式を提案し、その性能とコストの評価を行う。

3. システムの対象のモデル化

3.1 利用者の移動モデル

4章以降のシステム提案や評価を行うために、ここでは対象の単純なモデル化を行う。各種の交通機関において、利用者は交通機関に沿って物理的な移動を行う。ここでは、その移動する領域を交通ネットワークと呼ぶことにする。一般に、交通ネットワーク内の離散的な点（ノードとする：駅、ゲート、乗降口等）において、利用者の「利用開始」、「利用終了」の事象が生じる。「利用開始」によって利用者は交通ネットワーク内に入り、「利用終了」によって利用者は交通ネットワークから脱退する。交通ネットワークの利用料金は「利用開始」、「利用終了」の事象が発生した場所（ノードID）で決定する（図4）。

交通ネットワーク内での利用者の移動速度は利用者の特性や先行によって異なる^{9),10)}。たとえば、鉄道による移動の場合、高齢者は一般に駅構内の移動や乗換えに要する時間が大きく、移動速度は相対的に低下するであろう¹¹⁾。また、利用する駅が主要駅の場合には、急行や快速等の優等列車を利用できることから移動速度は相対的に高くなる。また、交通ネットワーク内で「利用開始」や「利用終了」が生ずるノードの分布は一般的には主要駅は他のノードに比べて高くなる。

これらの現実的な分布を、ここでは理想的な確率分布で代替する。以下のような前提をおく。

- 利用者の状態（ノード間の移動や「利用終了」事象の発生）はマルコフ過程に従って遷移する（離

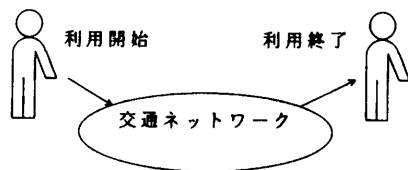


図4 交通ネットワークと利用開始/終了

Fig. 4 Transportation network and the entrance/exit of a user.

散の状態をとり、時間あたり一定確率で状態推し過去履歴には依存しない。

- あるノードに存在する利用者が、単位時間内に次ノードへ移る確率（移動率）：

$$\alpha \quad (0 \leq \alpha \leq 1)$$

交通ネットワークから退出する確率（退出率）：

$$\beta \quad (0 \leq \beta \leq 1)$$

とする。

同じノードに留まる確率は $1 - \alpha - \beta$ となる。ここでは単位時間内に2ノード（2駅）進む確率は無視する（あるいは2ノード進む確率を無視できる時間を単位時間としたともいえる）。

また、交通ネットワークの形状は線状（1次元）、格子状、環状、さらにそれらの組合せ等、様々な形態が考えられるが、本論文の解析では

- 線形状は線形（単純な直線状）であり、終端はない

とする。

上記のような前提で考えた交通ネットワーク上の利用者移動のモデルは図5のようになる。

3.2 利用者移動モデルの性質

前節で述べたモデルは定性的には、鉄道の例で表現すれば、以下のように解釈できる。ある時点である駅の管内に存在する利用者が、次の単位時間内にその駅の改札出口を出る事象の確率が β である。また次の単位時間内に、次の駅の管内へ進む事象の確率は α である。次の駅へ進んだり、改札へ向かって出る移動速度は利用者によって差があり一様ではない。それゆえ、次の単位時間内に、どちらの事象にも進まない確率が存在し、それが $(1 - \alpha - \beta)$ である。

上記のようなモデルの場合、定量的には以下のような基本的な性質が導かれる。以下、「利用者」を Entity、「利用開始」を Entry、「利用終了」を Exit とする。

いま、時刻 $t = 0$ において、Node#0 に Entity が存在しているとする。

- その Entity が、時間 t 後に交通ネットワーク内に存在する確率：

$$E(\beta, t) = (1 - \beta)^t \quad (1)$$

- その Entity が、時間 t 後に Node# n に存在する確率（ただし $t \geq n$ ）：

$$P(\alpha, \beta, t, n) = {}_t C_n \alpha^n (1 - \alpha - \beta)^{t-n} \quad (2)$$

- その Entity が、時間 $t \sim (t+1)$ 間に Node# n で Exit する確率（ただし $t \geq n$ ）：

$$\begin{aligned} W(\alpha, \beta, t, n) &= \beta P(\alpha, \beta, t, n) \\ &= \beta {}_t C_n \alpha^n (1 - \alpha - \beta)^{t-n} \end{aligned} \quad (3)$$

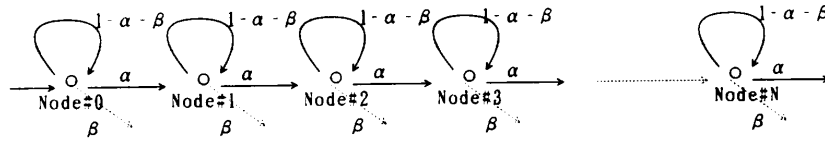


図5 交通ネットワーク上の利用者移動のモデル

Fig. 5 Model of transportation network and movement of a user.

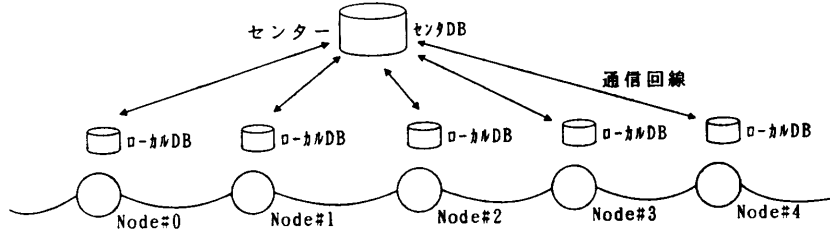


図6 システムの基本構成

Fig. 6 Basic structure of the system.

- その Entity が, Node# n で時間 t 以降に Exit する確率 (ただし $t \geq n$):

$$\begin{aligned} Z(\alpha, \beta, t, n) &= \sum_{i=t}^{\infty} \{W(\alpha, \beta, i, n)\} \\ &= \beta \alpha^n \sum_{i=t}^{\infty} \{i C_n (1 - \alpha - \beta)^{i-n}\} \quad (4) \end{aligned}$$

- その Entity が, Node# n で Exit する確率:

$$\begin{aligned} Z(\alpha, \beta, 0, n) &= \beta \alpha^n \sum_{i=n}^{\infty} \{i C_n (1 - \alpha - \beta)^{i-n}\} \quad (5) \end{aligned}$$

- Entity の平均システム滞在 (Entry ~ Exit) 時間:

$$\begin{aligned} &\sum_{i=1}^{\infty} [i \beta (1 - \beta)^i] \\ &= \beta \sum_{i=1}^{\infty} \left[\sum_{j=1}^{\infty} (1 - \beta)^j \right] \\ &= \beta \sum_{i=1}^{\infty} \left[\frac{(1 - \beta)^i}{\beta} \right] = \frac{1 - \beta}{\beta} \quad (6) \end{aligned}$$

- Entity の平均移動距離 (移動ノード数):

$$= \frac{1 - \beta}{\beta} \alpha \quad (7)$$

4. 複製データ配置方式

4.1 システムの基本構成

2章で述べた「ユーザ ID ベースの方式」を実現するためには利用者の利用開始時点でシステム側にて利用開始に関する情報 (たとえば, 鉄道なら乗車駅名等)

とユーザ ID との組合せの情報を内部データとして登録する. そして利用者の利用終了地点で, そのデータを用いて料金表等を参照して料金計算等必要処理のある処理時間内で行うことが必要である. この処理時間に関する要求の程度はアプリケーションによって異なるが, ある例では 100 ms 以内という目標値が示されている³⁾.

ここではセンターに存在するデータベースに利用者に関する情報 (定期券やストアードフェア等の情報) が登録されており, 利用開始時点で利用開始に関する情報が通信回線経由でそこに登録されるとする. そして利用終了時には, それを同じく通信回線経由により参照可能とする. 利用終了時の処理時間を一定値以下に収めるという制約からノード自体にもある程度, そのデータの複製 (以下, レプリカ) を配置し利用終了事象の発生時に, その利用者情報がそのノードに存在していれば (レプリカヒット時), その自ノード上データを参照し, そうでなければ (ヒットミス時), センターのデータベース本体へアクセスして処理を行うこととする^{12)~14)}. システムの基本構成は図6のようになる.

4.2 複製データ配置方式の評価方法

前節で述べたような基本構成を持つシステムにおいて, データ分散配置は複雑な問題であり, ファイル配置問題 (FAP: File Allocation Problem) として, 多くの研究¹⁵⁾がある. 通常ファイル配置問題の場合, レプリカ更新時のコストの算出が重要となる. 一般的には参照頻度と比較して更新頻度が相対的に高い場合, 多数のレプリカを配置することは, その更新コストの増大に起因して, 不利になることが知られている¹³⁾. たとえば ARPANET の例で, 更新の比率が 40% 以下

表 2 W の実際の値の例
Table 2 An example of the value of W .

t/n	0	1	2	3	4	5	6	7
000	0.100							
001	0.070	0.020						
002	0.049	0.028	0.004					
003	0.034	0.029	0.008					
004	0.024	0.027	0.012	0.002				
005	0.017	0.024	0.014	0.004				
006	0.012	0.020	0.014	0.005	0.001			
007	0.008	0.016	0.014	0.007	0.002			
008	0.006	0.013	0.013	0.008	0.003			
009	0.004	0.010	0.012	0.008	0.003			
010	0.003	0.008	0.010	0.008	0.004	0.001		
011	0.002	0.006	0.009	0.008	0.004	0.002		
012	0.001	0.005	0.007	0.007	0.005	0.002		
013		0.004	0.006	0.006	0.005	0.002		
014		0.003	0.005	0.006	0.005	0.003	0.001	
015		0.002	0.004	0.005	0.004	0.003	0.001	
016		0.002	0.003	0.004	0.004	0.003	0.001	
017		0.001	0.003	0.004	0.004	0.003	0.002	
018			0.002	0.003	0.003	0.003	0.002	
019			0.002	0.003	0.003	0.003	0.002	
020			0.001	0.002	0.003	0.002	0.002	

の場合に限って、レプリカ配置が有利になるという報告がある¹⁸⁾。しかし今回の「利用料金徴収のための分散データ配置方式」の場合、レプリカ配置に要するコストは、レプリカ配置数によって決まるストレージコストが支配的と考える。そして得られるメリット（平均レスポンスタイムの向上）と、レプリカ配置数との関係の分析に焦点をあてる。その理由は、今回の場合はアプリケーションの性質上、以下のように、一般的なファイル配置問題とは異なった特殊性、単純性があると考えられるからである。

複製されたデータの更新手続きは一般に不要である。交通ネットワークに入場した Entity に関して、その Exit はただ 1 回、いずれかのノードにて発生する。その際に Entity に関するデータが参照され、それ以降はその複製データは不要となる。データストレージの管理上、Exit 事象後には複製データが、全交通ネットワーク上で速やかに削除されることが有利であるが、その削除はリアルタイムで必要になるのではなく、ストレージ管理上必要な範囲で遅滞なく行えばよい。それゆえ、データ複製更新に関するコストはストレージコストと比べた場合に相対的に無視できるものとする。

すなわち、本論文では以下を前提とする。

- 要求条件は Entity の Exit 時の処理時間の最大値で与えられるが、直接的にはレプリカヒット率の下限が具体的目標になる。
- レプリカ配置のコストは、その配置数で決まる

ストレージコストが支配的であり、直接的には 1 Entity あたりのレプリカ配置数の上限が具体的目標になる。

- レプリカの再配置はセンタが主導して随時、ダイナミックに行えるものとする。
- Entity の Exit 事象は、Exit 後、速やかにセンタへ通知されるものとする。
- 対象とする Entity に関して、その移動率： α 、退出率： β は Entry 時点には与えられており、その Entity が退出するまでは不変とする。

上記の前提と 3 章で提案したモデルに従い、いくつかの種類レプリカ配置、評価方法の検討を行った。ここでは具体的方法として、以下のようなダイナミックな配置方式に関して述べる。

4.3 複製データの配置方式

対象 Entity の各時刻/各ノードでの Exit 率の分布： $W(\alpha, \beta, t, n)$ に注目する (式 (3))。この分布 W はその時刻 t /ノード n にレプリカを配置した場合に得られるヒット率を示すものといえる。参考として、 $\alpha = 0.2$ 、 $\beta = 0.1$ の場合の W の値を表 2 に示す。縦軸は経過時間、横軸はノードである (太字は同時刻での最大値を示す)。

上記の W の表の各セル (t と n との組合せで決定する表中の柀) に示された値は、そのセル (t, n) に新たにレプリカを配置することで得られるヒット率の増加値を表す。これらのヒット率増加値を、配置するセ

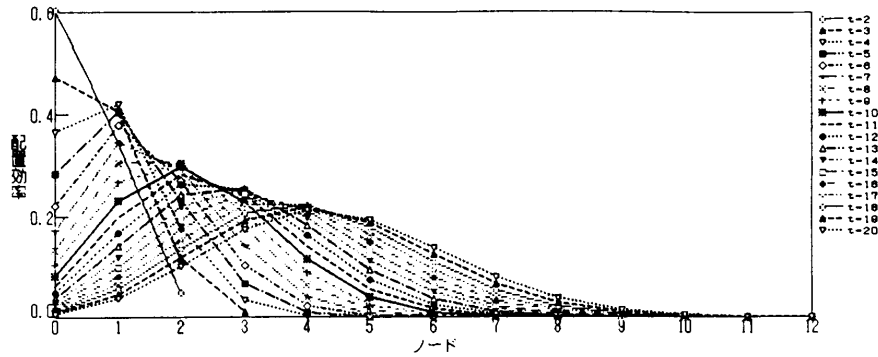


図7 レプリカ配置効果率の分布例
Fig. 7 An example of a distribution of R.

ル全体に関する総和をとれば全体のヒット率が計算される。すなわちヒット率は以下となる。

$$\sum_{t,n} W(\alpha, \beta, t, n) \tag{8}$$

一方、各時刻/ノードにレプリカを配置するためのコストを考える。ここでいうコストは、以下の式で与えられるようなストレージコストとする。

$$\int_{t=0}^{\infty} (\text{レプリカ個数}) \times (\text{レプリカ存在時間}) \times (\text{存在確率}) dt \tag{9}$$

1セルをカバーするレプリカ配置コスト(1つのノードに1単位時間1レプリカを配置するコスト)を基本的に1とする。しかし実効的なコストは実際に配置が必要となる確率との積をとった値と等しくなる。すなわち、セル (t, n) にレプリカ配置を予定する場合の実効コストは以下の式で計算される。

$$\begin{aligned} \text{実効コスト} &= (\text{配置コスト}) \times (\text{配置を実際に行う確率}) \\ &= (\text{配置コスト}) \times (\text{Entity 存在確率}) \\ &= (1 - \beta)^t \end{aligned} \tag{10}$$

この値が、そのセルにレプリカ配置を予定した場合のコストの期待値であり、これを以下「実効コスト」と定義することにする。

各セルごとに、そこにレプリカ配置を行うことにより得られるヒット率と、その実効コストが求められる。この2つの値の比率をレプリカ配置効果率 $R(\alpha, \beta, t, n)$ と呼ぶことにする。これは、いわゆるレプリカ配置のコストパフォーマンスを示す指標ともいえる。 $R(\alpha, \beta, t, n)$ は以下の式で計算される。

$$\begin{aligned} R(\alpha, \beta, t, n) &= (\text{得られるヒット率}) / (\text{実効コスト}) \\ &= [\beta_t C_n \alpha^n (1 - \alpha - \beta)^{t-n}] / (1 - \beta)^t \end{aligned} \tag{11}$$

上式は以下のようにも変形でき、一般的な2項分布の

形となることが分かる。

$$R(\alpha, \beta, t, n) = \beta_t C_n \lambda^n (1 - \lambda)^{t-n} \tag{12}$$

ただし、 $\lambda = \alpha / (1 - \beta)$

参考として、 $\alpha = 0.2, \beta = 0.1$ の場合の $R(\alpha, \beta, t, n)$ の値を図7に示す。最小コストで最大のレプリカヒット率を得るために、このレプリカ配置効果率の大きいセルから順にレプリカ配置を予定する。レプリカ配置予定セル $(t_1, n_1), (t_2, n_2), \dots, (t_m, n_m)$ を対象として、各時刻 t ごとに、対象 Entity が存在していれば (Exit していなければ)、 $t \sim (t+1)$ の時間に該当するセルに、実際にレプリカ配置を行う。この場合の実効コストは、決定された、 m 個の配置予定セル $(t_1, n_1), (t_2, n_2), \dots, (t_m, n_m)$ に関して、実効コストの総和をとることにより求められる。すなわち、以下となる。

$$\begin{aligned} \text{実効コスト} &= \sum (\text{配置予定セルの実効コスト}) \\ &= \sum_{i=1}^m (1 - \beta)^{t_i} \end{aligned} \tag{13}$$

このような手順に従ってレプリカ配置を決定する過程の例を表3に示す。これは、 $\alpha = 0.5, \beta = 0.3, t = 0 \sim 100$ の条件の際に、配置効果率 $R(\alpha, \beta, t, n)$ の順に上位10セルを並べたものである。ヒット率の項を上位から順に加えて行き、満たすべきヒット率に達するまでのセルがレプリカ配置予定セルとなる。そして、実効コストは、配置予定の各セルの実効コストの項の総和となる。

表3からも見られるように、配置効果率 R の最大のセルは、時間 $t = 0$ 、ノード $n = 0$ のものである。配置効果率の順は、概して時間の順に近いが、一部で逆転現象が起きる。この表3の内容を、上位80セルに関してグラフにしたものが図8である。

実際のレプリカ配置箇所の例を図9に示す。この図

表3 レプリカ配置セル決定のための計算過程の例
Table 3 An example of the process of computation.

	t	n	R	w	Σw	Cost	Σ Cost
1	0	0	1.0000	0.3000	0.3000	1.0000	1.0000
2	1	1	0.7143	0.1500	0.4500	0.7000	1.7000
3	2	2	0.5102	0.0750	0.5250	0.4900	2.1900
4	3	2	0.4373	0.0450	0.5700	0.3430	2.5330
5	4	3	0.4165	0.0300	0.6000	0.2401	2.7731
6	2	1	0.4082	0.0600	0.6600	0.4900	3.2631
7	5	4	0.3719	0.0188	0.6788	0.1681	3.4312
8	3	3	0.3644	0.0375	0.7162	0.3430	3.7742
9	6	4	0.3187	0.0113	0.7275	0.1176	3.8918
10	6	5	0.3187	0.0113	0.7387	0.1176	4.0095

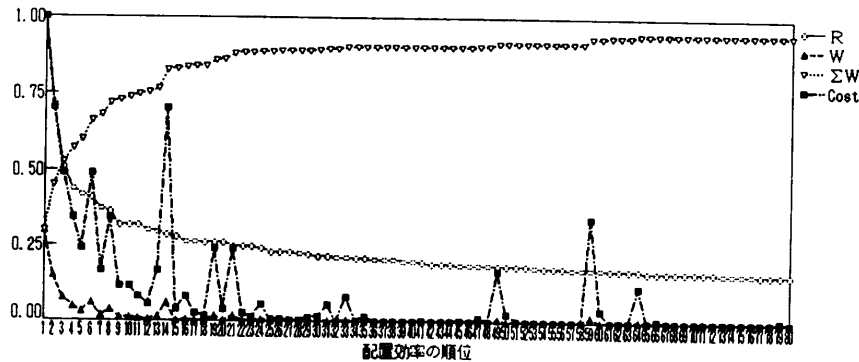


図8 ダイナミックなレプリカ配置方式でのグラフ

Fig. 8 An example of the value of hit-ratio and cost using dynamic method.

は、各セル (t, n) のレプリカ配置を示したものであり、横軸がノード n (左端が $n = 0$)、縦軸が時間 t (上端が $t = 0$) である。図中で 'o' はレプリカの配置セル、'_' はレプリカは配置されないが、Exit 率がゼロではないセルを示している。ここでは、 $\alpha = 0.45$ 、 $\beta = 0.1$ 、 $(\gamma = 1)$ 、 $h = 0.88$ 、 $t = 0 \sim 100$ としている。

上記のような手順に従ってレプリカ配置を行った場合の実効コストと得られるヒット率との関係を計算した結果の例を、図10のグラフで示す。ここで、

- 対象 Entity に関するパラメータ：
 $\alpha = 0.1 \sim 1.0$ 、 $\beta = 0.1 \sim 1.0$
- 満たすべきヒット率：
 $h = 0.5 \sim 0.9$

としている。

図10からも分かるように、 α と、 $1 - \alpha - \beta$ とを逆転した場合の結果は同様になる。これは、以下のように $W(\alpha, \beta, t, n)$ (式(3)) を分析すれば導かれる。

$$W(\alpha, \beta, t, n) = \beta P(\alpha, \beta, t, n) \\ = \beta_t C_n \alpha^n (1 - \alpha - \beta)^{t-n}$$

この $W(\alpha, \beta, t, n)$ の式を観察すると

$$W(1 - \alpha - \beta, \beta, t, n) \\ = \beta_t C_n (1 - \alpha - \beta)^n \alpha^{t-n} \\ = \beta_t C_{t-n} \alpha^{t-n} (1 - \alpha - \beta)^{t-(t-n)} \\ = W(\alpha, \beta, t, t-n) \quad (14)$$

すなわち、 α と $(1 - \alpha - \beta)$ とを交換した場合、分布は左右逆になるだけである。このため、このようなダイナミックなレプリカ配置方式の場合、 α と $(1 - \alpha - \beta)$ とを交換しても、各時間ごとのレプリカ配置を行うノードは異なるが、各時間ごとの分布形状が同様なことから、配置ノード数は同一であり、したがってコストも同一である。

ここで、拡散率 γ を以下のように定義する。

$$\gamma = \frac{\min(1 - \alpha - \beta, \alpha)}{\max(1 - \alpha - \beta, \alpha)} \quad (15) \\ (0 \geq \gamma \geq 1)$$

この方式の場合のレプリカ配置コストは、 β と上記の拡散率 γ で決定されるともいえる。すなわち、ある β を固定して考えた場合には、拡散率が等しければ、レプリカ配置コストは同様となる。定性的には、 γ は、3章で述べた今回の移動モデルで、対象エンティティの存在確率の拡散の速さを表す値といえることができる。

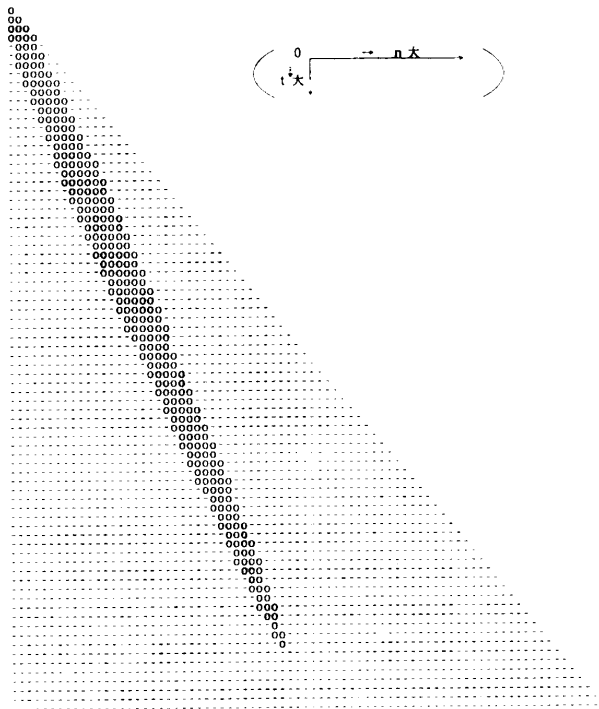


図9 ダイナミック方式のレプリカ配置セルのレイアウト

Fig.9 Layout of cells where replica is distributed using dynamic method.

拡散率 γ が大きい (1 に近い) 場合に 退出率分布のノード拡散が最大となり, 与えられたヒット率を満たすために必要なレプリカ数が最大となる. 拡散率が最大 ($\gamma = 1$) となる移動率 (α_0) は, $1 - \alpha_0 - \beta = \alpha_0$ から

$$\alpha_0 = \frac{1 - \beta}{2} \quad (16)$$

となる.

5. 結論

交通機関の自動料金徴収システムに関して, 「ユーザ ID ベースの料金徴収方式」は 2 章で述べたような大きな利点をいくつか持っているが, 定性的に処理時間やデータベース容量の関係から困難と考えられてきた. この方式の実現のためには, 様々な交通ネットワークの形態に応じて, 利用者移動に関する複雑な解析が必要となる. 本論文ではその基礎検討のために, 単純な線形の交通ネットワークと, 3 章で述べたような単純なマルコフ過程に基づく利用者移動モデルを対象とした場合における分散データ配置方式, それらによるレスポンスタイムの短縮効果, 分散データ配置のためのコストの決定方法等の解析を行った. レプリカデータ配置箇所の決定方式はいくつかが考えられるが, どの

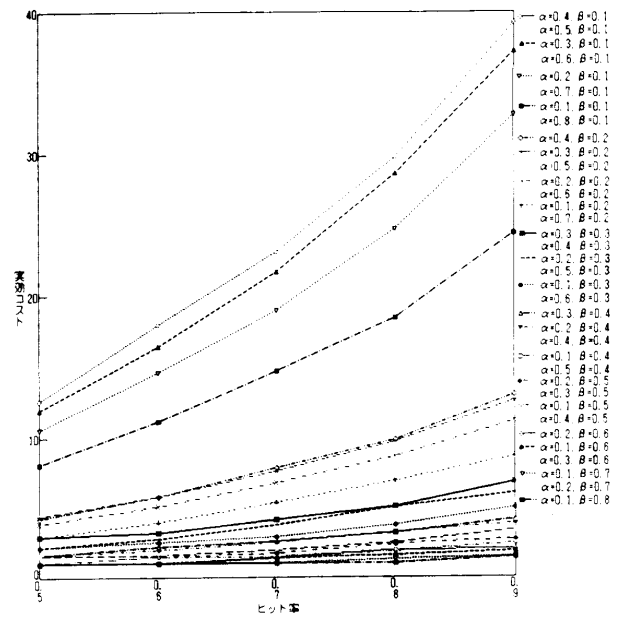


図10 ダイナミックなレプリカ配置方式でのヒット率とコスト計算例

Fig. 10 An example of the value of hit-ratio and cost using dynamic method.

のような方法が適切かは,

- システム機能上の制約事項
 - コスト上の制約事項
 - 対象モデルの性質 (移動率 α や退出率 β の値)
- 等によって左右されるが, 一般的には 4 章で述べたようなダイナミックな方式が最も有利と考えられる.

参考までにスタティックな配置方式を採用した場合のコスト計算の結果を図 11 に示す. これは Entity の Entry 時にレプリカ配置箇所をすべて決定し, 以降, Exit まで配置を固定するという比較的単純なものである. 4 章の方式と比べた場合, コストは上昇している.

6. 今後の課題

単純な条件に限定された範囲での「ユーザ ID ベースの料金徴収方式」の実現と評価の一方法を示してきた. この結果をベースとして現実との整合性評価や, より現実に近づけるためのパラメータ設定方法等を, 明らかにしていきたい. 以下のような項目を今後の課題として考えている.

6.1 パラメータの決定方法

3 章の利用者移動モデルをベースとして解析する場合, 移動率 α , 退出率 β の決定方法が重要である. 3 章ではこれらのパラメータを固定値として考えていたが, この拡張として, 以下のように, 時間に応じて, 位置 (ノード) に応じて, 個人に応じて等, 柔軟に変

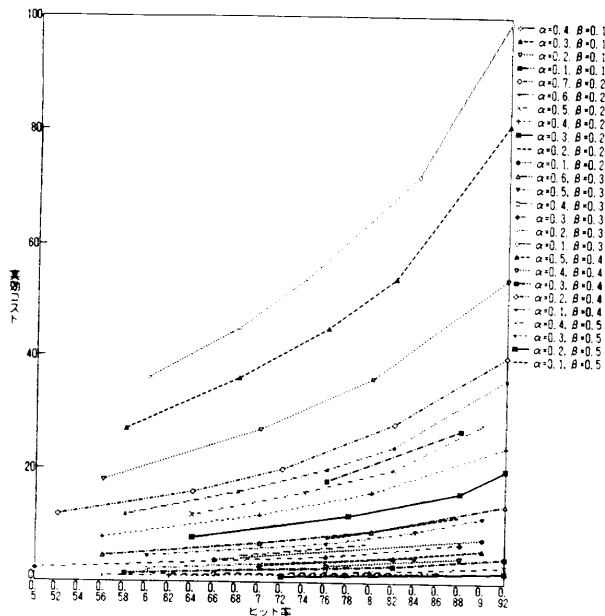


図 11 スタティックなレプリカ配置方式でのヒット率とコスト試算例

Fig. 11 An example of the value of hit-ratio and cost using static method.

更する方法が考えられる。

● 総体的な利用者移動特性の利用

現実には、利用者総体としての環境ごとの移動特性がある。たとえば、退出率の高いノード（駅）、低いノード（駅）が存在する。これらの特性を統計的に得て、ネットワーク（路線）別やノード（駅）別に移動率、退出率を設定する方法が考えられる。

● 個々の利用者移動特性の利用

個々の利用者の過去の移動の状況が把握できたとすれば、入場の際に得られるユーザ ID に従い、個人別に移動率、退出率を設定する方法が考えられる。具体的には、ある個人が特定の駅で降りる確率が高いことが把握できれば、その駅での退出率をそれに応じて高くする。現状では個々の利用者の履歴情報を交通機関側が獲得できている例はほとんどない。しかし今後は、IC カード等のデータキャリアを電子切符として使用し、そのデータ処理能力、蓄積能力を活用することで、個々の利用者の活動状況を把握して営業活動にフィードバックさせる方法も提唱されている²³⁾。この際に獲得できる、個人別の情報の利用が考えられる。

● 交通機関側の特性の利用

鉄道の場合には輸送サービスは列車ダイヤに基づいて提供されており、利用者の移動はこれに制約されている。交通機関側は、この列車ダイヤ情報を利用することで、時刻、ネットワーク（路線）、ノード（駅）別

に移動率の値を設定することが考えられる。道路交通の場合でも標準走行時間によって設定する移動率を渋滞情報によって変更する等が考えられる。

6.2 システム構成

3章のモデルは 1 Entity（利用者）を対象としたものであり、Entity の増加にともない解析結果は重畳されデータベース、トラフィックの量は増大する。それらが膨大となった場合のシステム構築上の課題の検討も必要である。システムが大規模となる場合のシステム構成に関しては、4章で述べたように、遠隔参照用のデータベースを単純に中央集中とするのではなく、ある程度分散し、あるいは階層化された構成とする検討も進めている。

6.3 セキュリティ

「ユーザ ID ベースの料金徴収方式」は将来的に利用者へのサービスを格段に向上させる可能性を持つものであるが、本論文で述べてきたようなシステム上の課題のほかに利用者のプライバシーの保護、たとえば、利用履歴情報のセキュリティ確保という問題も存在する。これは社会的な問題として別途の検討も必要となる。これに関しては、たとえば文献 24) 等で、暗号技術を利用して匿名性を保つ方式が検討されている。

7. おわりに

現実化には前章で述べたことをはじめ多くの課題がある。今回の検討結果が第一ステップとして役立つことを期待している。ここで述べたような方法は、交通機関の料金徴収という適用分野だけでなく、一般的には、移動する Entity を対象として、その行動に応じて関連データの準備/参照が必要な状況、たとえばセルラー通信システムでの移動体追跡等にも適用できる可能性があると考えている。

謝辞 本論文を作成するに際して、背景となる多くの情報を、東日本旅客鉄道（株）、西日本旅客鉄道（株）の関係者の方々から得たことを感謝いたします。

参考文献

- 1) 日本国有鉄道車両局：日本国有鉄道規格 JRS 76321-5B-15AR5M（自動改札装置）（1985）。
- 2) 三木彬生：非接触 IC カードによる乗車券システムの基本構想，鉄道総研報告，Vol.4, No.12, pp.53-60（1990）。
- 3) 後藤浩一，松原 広，佐々木健一，永井 昇：非接触 IC カードによる自動改札システム，鉄道総研報告，Vol.7, No.12, pp.1-8（1993）。
- 4) 曾根 悟，井上照久：乗客用非接触 IC カードによる鉄道利用システムの改善の可能性，電気学会

- 交通・電気鉄道/道路交通合同研究会, TER91-4, 電気学会 (1991).
- 5) Goto, K., Matsubara, H. and Sasaki, K.: New Railway Ticket System Using Contactless IC Cards, *Proc. IFAC-TS'94 International, Federation of Automatic Control* (1994).
 - 6) 矢野代二郎, 鎌田譲治, 駒田牧夫: ノルウェーの有料道路制と料金自動收受システムについて, *交通工学*, Vol.26, No.4, pp.29-35 (1991).
 - 7) Tetsusaki, K.: High Security Electronic Toll and Traffic Management and Road Pricing System Using Encrypted Messages and Personal Identity Number, *Vehicle Navigation & Information Systems Conference*, pp.695-698 (1994).
 - 8) 松川公一, 竹下孝徳, 青柳秀幸: 非接触 IC カードを使用した自動改札システム, *日本信号技報*, Vol.18, No.4, pp.1-12 (1994).
 - 9) 柴田正雄ほか: シミュレーションによる年間平均旅行速度の推定, *交通工学*, Vol.24, No.6, pp.11-20 (1989).
 - 10) 柴田正雄: 旅行時間推定式の提案, *交通工学*, Vol.29, No.1, pp.35-37 (1994).
 - 11) 清水浩志郎, 木村一裕, 吉岡靖弘: 道路横断施設における高齢者の歩行特性に関する考察, *交通工学*, Vol.26, No.2, pp.29-38 (1991).
 - 12) Mullender, S.: *Distributed Systems*, ACM Press (1989).
 - 13) 前川 守 (編): 分散オペレーティングシステム, 共立出版 (1991).
 - 14) Coulouris, G.F. and Dollimore, J.: *Distributed Systems, Concept and Design*, 邦訳: 分散システム, 電気書院 (1991).
 - 15) Dowdy, L.W. and Foster, D.V.: Comparative Models of the File Assignment Problem, *Computing Surveys*, Vol.14, No.2 (1982).
 - 16) Smith, A.J.: Long Term File Migration: Development and Evaluation of Algorithms, *Communication of the ACM*, Vol.24, No.8, pp.521-532 (1981).
 - 17) 重田和弘, 高野 誠, 斎藤 勲: 通信システムにおける分散データ配置方式, マルチメディア通信と分散処理研究会, 45-13, pp.93-100, 情報処理学会 (1990).
 - 18) Fisher, M. and Hochbaum, D.: Database Location in Computer Networks, *JACM*, Vol.27, No.4, pp.718-735 (1980).
 - 19) Casey, R.G.: Allocation of Copies of a File in an Information Network, *Spring Joint Computer Conference*, pp.617-625, AFIPS Press (1972).
 - 20) 高品智一, 宮西洋太郎, 渡辺 尚, 水野忠則: 分散環境における部分データベースの複製配置法, マルチメディア通信と分散処理研究会, 66-2, pp.7-12, 情報処理学会 (1994).
 - 21) Wolfson, O. and Milo, A.: The Multicast Policy and Its Relationship to Replicated Data Placement, *ACM Transaction on Database Systems*, Vol.16, No.1, pp.181-205 (1991).
 - 22) Kumar, A. and Segev A.: Cost and Availability Tradeoffs in Replicated Data Concurrency Control, *ACM Transactions on Database Systems*, Vol.18, No.1, pp.102-120 (1993).
 - 23) Takagi, R., Goto, K., Tanaka, M., Koseki, T. and Sone, S.: Management Data Acquisition System Using Smartcards Carried by Individual Passengers, *Proc. WCRR* (to be published) (1996).
 - 24) Chaum, D.: A New Paradigm for Individuals in the Information Age, *Proc. the Symposium on Security and Privacy*, IEEE Computer Press (1984).

(平成 8 年 3 月 1 日受付)

(平成 8 年 9 月 12 日採録)

田中 幹夫 (正会員)



1953年生。東京都出身。1976年、東京大学工学部電気工学科卒業。同年、日本国有鉄道入社。仙台電気工事局、盛岡鉄道管理局、鹿児島鉄道管理局、国鉄本社にて信号通信、コンピュータ関係設備の計画、工事、保守、技術開発等に従事。1987年から(財)鉄道総合技術研究所に勤務し、情報システム、ネットワークに関する研究開発、設計、コンサルティング等に従事。現在、輸送システム開発推進部主幹技師(情報システム担当)。IEEE-CS会員。