

鉄道におけるユーザ ID ベースの料金徴収方式の適用と評価

田 中 幹 夫†

鉄道旅客輸送や有料道路交通等の交通機関の自動料金徴収を利用者のユーザ ID 情報のみを用いて行うシステムは、利用者側で保持するデータキャリア等に関連情報を負担させる既存方式と比較して種々の特徴を備え優れたサービスを実現する可能性を持っている。しかし実現するシステム側の負荷が大きく、その構築方法や性能、コストの解析・評価が課題となる。当論文では、交通機関内の旅客移動を反映するためのモデルを提案し、そのモデルに基づいたシステム実現方法として、交通機関利用者の予想される移動に合わせて必要情報の複製を分散配置する方式、および、必要情報を移動させる方式に関して、性能やコストの解析を行う。そして、現実の鉄道路線の旅客輸送データを適用した評価を試み、対象となる交通機関の輸送上の特性との関係について考察する。

An Analysis, Application and Evaluation of User-ID Based Automatic Toll Collection Systems for Transportation Services

MIKIO TANAKA†

For the purpose of an automatic toll collection system for transportation services, a user-ID based system, which has some valuable merits compared with existing systems, is proposed. A model to formulate the movement of users (passengers/vehicles) along a transportation network is proposed. In order to realize such a system efficiently, especially to reduce processing time at exit nodes, a method of data replication and a method of data re-allocation of data over a transportation network according to related entity in distributed communication systems are introduced. An analysis, application and evaluation of those methods are shown.

1. ま え が き

交通機関の料金徴収業務自動化のためにデータキャリア等を適用したシステムの開発、試行、実用化が進められている^{1)~6)}。現在、読取書込可能のデータキャリア(磁気カード、ICカード等)を利用者(乗客、自動車等)に携帯させ、必要情報をその中に書き込み保存する方式が主流である。当論文ではデータキャリア等から得る利用者に関する ID 情報のみを用いる「ユーザ ID ベースの料金徴収システム」⁷⁾に関して論ずる。

文献 7)では、利用者移動の(基礎的)モデル化の方法と予想される利用者移動に合わせ必要データの複製を作成して分散配置する方式に関して論じている。文献 8)では、同様モデルに基づき、予想される利用者移動に合わせ必要データを適切に移動させる方式に関して論じている。当論文では、現実への適応性を高め、輸送特性を反映した解析が可能な(拡張)モデル

を導入し、それに基づいた解析・評価方法を示す。そして、そのモデルに基づくデータ複製配置方式、データ移動方式に関して論じ、現実の鉄道旅客輸送データを適用した評価を試み、輸送上の特性との関係について考察する。

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)の自動料金徴収のためには、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残額情報や銀行口座等から差し引く。また、定期券利用者であれば、その情報(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の有無を判断する。これら一連の処理を利用区間確定時点、すなわち利用終了地点(改札口、ゲート等)で行い、その結果を利用者に示すことが必要である¹⁾。

このような自動料金徴収のために、近年では各種データキャリアを利用した方式の開発試行、実用化が

† 財団法人鉄道総合技術研究所技術開発事業本部情報システム担当グループ
Information Systems Group, Railway Technical Research Institute

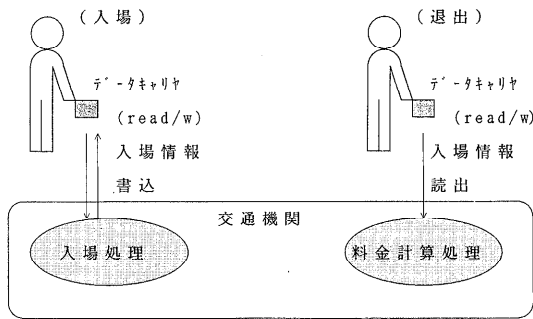


図1 利用者自身が情報を携帯する方式
Fig.1 Method of user carrying information.

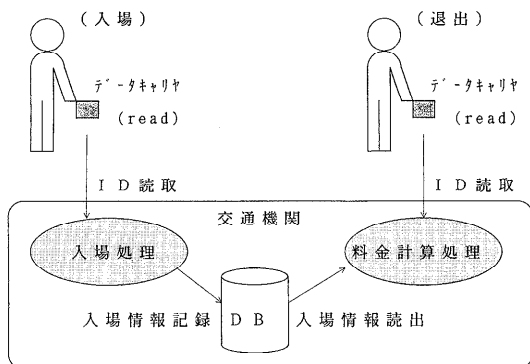


図2 交通機関の提供者が情報をデータベース内に保持する方式
Fig.2 Method of transporter storing information.

進められている。現在は利用者（乗客、自動車等）に携帯または装着させる読取書込型データキャリアを使用したものが主流である¹⁾。最近はより大きなメモリ容量、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を行うICカードやデータキャリアを使用することにより、利用者の利便性を向上しようという試みもなされている^{2)~6)}。

2.2 自動料金徴収システムの各種方式

自動料金徴収のための必要な情報（最低限の情報としては利用開始箇所を示す情報）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⁷⁾。

- (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が利用者の携帯または装着するデータキャリア内部に蓄えられる。入口で、交通機関側は利用開始に関する情報をシステム内部には特に取り込まず、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カードに必要な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、前記必要情報を交通機関側にデータとして保存し参照するもので、利用者に関しては、そのID（ユーザを特定する固定情報）だけを使用する。何等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等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小規模な場合に限定されている。

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ID情報のみで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。

- (A) データキャリア、および読取装置は利用者のIDを読み取るのみで十分であり、技術的、構造的に単純化され、経済性、安全性が向上する。
- (B) 入出場（改札等）時点でのトランザクションが単純となり、処理が高速化される。
- (C) 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。
- (D) バーコードのように安価であるが光学的方式で書き込み困難なデータキャリアも使用できる可能性がある。
- (E) 技術的課題は大きいが将来的にはデータキャリアを使用せず人間や自動車自体の認識結果（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、利用者に携帯させるものはまったく不要となり、いわゆるキップレスシステムが実現する。

(2)のような方式を、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。この方式は、上記のような特長を備えているが、前節で述べたように交通機関側システムに対し処理速度や容量の面で大きな能力を要求することが予想されるため、現状で実用化されているケースはほとんどない。また、定量的分析のために、利用者移動に関する複雑な解析が必要となる。当論文では、この「ユーザIDベースの料金徴収方式」を実現する方式の提案、解析、評価を主題とする。

3. 利用者の移動モデル

3.1 利用者と交通ネットワーク

4章で述べる実現方式に関して、その性能やコスト

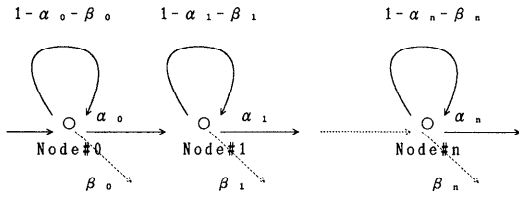


図3 交通ネットワーク上の利用者移動(拡張)モデル

Fig. 3 Extended model of transportation network and movement of a user.

の定量的解析を行うためには対象の移動のモデル化が必要である。またモデル化を行うことで対象交通機関の輸送特性が統計的な生データでなく幾種類かのパラメータで統一的に表現でき、システム性能との関係の議論が可能となる。ここでは文献7)で提案、使用した(基本)モデルを拡張する。各種交通機関において利用者は交通機関に沿って物理的移動を行う。ここではその移動する領域を交通ネットワークと呼ぶ。交通ネットワーク内の離散点(ノード: 駅, ゲート, 乗降口等)で、利用者の「利用開始」「利用終了」の事象が生じる。交通機関の利用料金は「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。

3.2 利用者移動モデル

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行先によって異なる^{9)~11)}。たとえば、鉄道の場合、高齢者は一般に駅構内移動や乗換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¹⁰⁾。また、利用駅が主要駅の場合には、急行や快速等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、「利用開始」や「利用終了」が生ずる確率は一般的には主要駅は他ノードに比べて高くなる。これらの現実的な分布を、ここでは理想的確率分布で代替する¹¹⁾(図3)。以下のような前提をおく。

- 利用者の状態(ノード間移動や「利用終了」事象発生)はマルコフ過程に従って離散の状態をとり、時間あたり一定確率で状態推移し過去履歴には依存しない。
- ノード n_i に存在する利用者の状態が、単位時間内に遷移する確率は、
次ノードへ移る確率(移動率): α_i
交通ネットワークから退出する確率(退出率): β_i
とする($0 \leq \alpha_i, \beta_i \leq 1$)。

同ノードにとどまる確率は $1 - \alpha_i - \beta_i$ となる。ここで単位時間内に2ノード(2駅)進む確率は無視する。

3.3 利用者移動モデルに基づく解析

対象 Entity(旅客, 自動車等)が、時刻 t , ノード n (これを、セル (t, n) と称す)で、exit する確率が、

複製データ配置ノードの決定、データ移動先ノードの決定に際して重要となる。以下の前提をおく。

- 時刻 $t = 0$ で、ノード 0 に Entity が存在する。
- 前記モデルに従い、この Entity の各ノード $\{n_i\}$ での移動率を $\{\alpha_i\}$, 退出率を $\{\beta_i\}$ とする。
- $1 - \alpha_i - \beta_i$ は、ノード n_i にて、Entity が単位時間中に退出も、次ノードへの移動の何れも起こさない(状態遷移上は“道草”する)確率である。これを道草率 $\{h_i\}$ とする。

すなわち、 $h_i = 1 - \alpha_i - \beta_i$

Entity のセル (t, n) での存在確率は、要約すれば、ノード: $0 \sim (n-1)$ の各移動率 $\{\alpha_i\}$ の積をとり $(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{n-1})$, さらに、道草数 $(t-n)$ を各ノード $(0 \sim (n-1))$ に割り振る各組合せごと (${}_t C_n$ 通り存在する)の道草率の積の総和を掛けることで求められる。以下に手順を示す。

まず、準備として、次のような「道草和(みちくさ)函数」を定義する。

- 道草和函数の定義

ノード $A \sim$ ノード B 間の道草率 $\{h_i\}$ の C 次元 (C 次元組合せ和)

A, B, C の函数になり、以下、 $H[A][B][C]$ と表現する(ただし、 $0 \leq A \leq B, 0 \leq C$)。

- 道草和函数: $H[A][B][C]$ を表す式

$$H[A][B][C] = \sum_{i=A}^B \left[h[i] \cdot \sum_{j=i}^B \left\{ h[j] \cdot \sum_{k=j}^B (h[k] \cdots) \right\} \right] \quad (1)$$

(C 重の入れ子となる)

- 求める手順

具体的には、この道草和函数は、以下の手順で再帰的に計算することで求める。

$$H[A][B][0] = 1 \quad (\text{任意の } A, B) \quad (2)$$

$$H[A][B][1] = \sum_{i=A}^B h[i] \quad (3)$$

$$H[A][B][2] = \sum_{i=A}^B \{h[i] \cdot H[i][B][1]\} \quad (4)$$

$$H[A][B][C+1] = \sum_{i=A}^B \{h[i] \cdot H[i][B][C]\} \quad (5)$$

この道草和函数: $H[A][B][C]$ を用い、該当 Entity の各セル (t, n) の存在率: $P(t, n)$ は下式で求められる。

$$P(n, n) = \alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{n-1} \quad (6)$$

$$P(t, n) = P(n, n) \cdot H[0][n][t - n] \quad (7)$$

(ただし, $n \leq t$)

セル (t, n) での Entity の exit 発生確率は以下となる。

$$\beta_n \cdot P(t, n) \quad (8)$$

Entity が, 時刻 t に交通ネットワーク内に存在する確率は以下となる。

$$E(t) = \sum_{n=0}^t P(t, n) \quad (9)$$

Entity の交通ネットワーク内平均滞在時間は以下となる。

$$T_{ave} = \sum_{t=0}^{\infty} \left\{ t \cdot \sum_{n=0}^t \beta_n \cdot P(t, n) \right\} \quad (10)$$

Entity の平均移動距離 (移動ノード数) は以下となる。

$$D_{ave} = \sum_{n=0}^{\infty} \left\{ n \cdot \beta_n \sum_{t=n}^{\infty} P(t, n) \right\} \quad (11)$$

ノード n に存在する Entity が, 最終的に ($t \rightarrow \infty$ の範囲で), そのノードで退出する確率は,

$$\begin{aligned} & \beta_n + (1 - \alpha_n - \beta_n)\beta_n + (1 - \alpha_n - \beta_n)^2\beta_n \\ & + (1 - \alpha_n - \beta_n)^3\beta_n + \dots \\ & = \sum_{i=0}^{\infty} (1 - \alpha_n - \beta_n)^i \beta_n \\ & = \beta_n / (\alpha_n + \beta_n) \end{aligned} \quad (12)$$

同様に, 次ノードへ前進する確率は,

$$\alpha_n / (\alpha_n + \beta_n) \quad (13)$$

ノード n に存在する Entity が, ノード $(n+1)$ へ最終的に ($t \rightarrow \infty$ の範囲で) 移動する場合に, その移動に要する時間の平均値は以下となる。

$$\begin{aligned} & [\alpha_n \cdot 1 + (1 - \alpha_n - \beta_n)\alpha_n \cdot 2 \\ & + (1 - \alpha_n - \beta_n)^2\alpha_n \cdot 3 + \dots] / [\alpha_n / (\alpha_n + \beta_n)] \\ & = \left[\alpha_n \sum_{i=0}^{\infty} (1 - \alpha_n - \beta_n)^i (i+1) \right] / \\ & \quad [\alpha_n / (\alpha_n + \beta_n)] \\ & = [\alpha_n \{1 / (\alpha_n + \beta_n)^2\}] / [\alpha_n / (\alpha_n + \beta_n)] \\ & = 1 / (\alpha_n + \beta_n) = 1 / (1 - h_n) \end{aligned} \quad (14)$$

4. システムの実現方式

4.1 データ複製配置とデータ移動方式

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, Entity の Exit 処理の際に, 関連データ

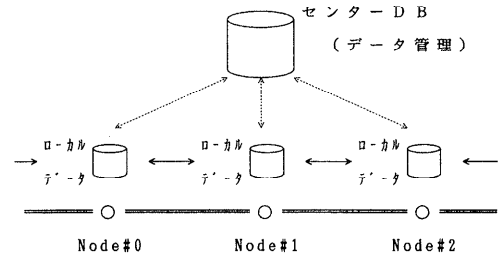


図4 システムの基本構成
Fig.4 Basic structure of the system.

(最低限, 入場箇所を示す情報) を高速参照して処理を完了することが必要となる。たとえば, 鉄道における改札処理の場合, 100 ms 以内に完了することが望ましいとされている³⁾。データベースを1カ所に集中して設置し, それに対して広範囲に散在する交通ネットワークの各ノードからアクセスを行う構成 (図4) は基本的なものであるが, この方法だけでは, システムの性能や容量面から不利な場合が多いと考えられる。これに代わる, あるいは補完するために, 必要データを要求箇所自身, あるいは物理的にその近接箇所に配置してデータ参照を高速化する手段が考えられる。以下の2種類の方式を提案する。ここでは Entity が物理的に移動する交通ネットワークに沿って, 情報を相互に伝送可能な通信ネットワークが存在し, これの利用が可能なることを前提としている。現実には交通事業者はこのような通信設備をすでに設置運用している場合が一般的である^{12),13)}。

(A) データ複製配置方式 (分散配置方式)⁷⁾

Exit 時の処理に必要なデータを極力要求ノードでローカルな DB アクセスによって得られるようデータの複製を作成し, Entity の動きに合わせて Exit 事象の発生確率の高い箇所に適切に配置する方式 (図5)。

(B) データ移動方式⁸⁾

Exit 時の処理に必要なデータを極力要求ノード, あるいは物理的に近いノードの DB アクセスで得られるよう, Entity の動きに合わせて Exit 事象の発生確率の高い箇所へデータを適切に移動する方式 (図6)。

データ複製配置方式は, データ参照が同一ノード内でローカルに行われる確率が高く, 通信ネットワークの性能にほとんど依存せずに高速なデータアクセスが期待できる。しかしデータ複製配置のための複雑な処理が必要であり, データのストレージコストも複製の分, 相当に大きくなることが予想される。またデータ更新処理は複製の同時更新の必要性から非常に複雑と

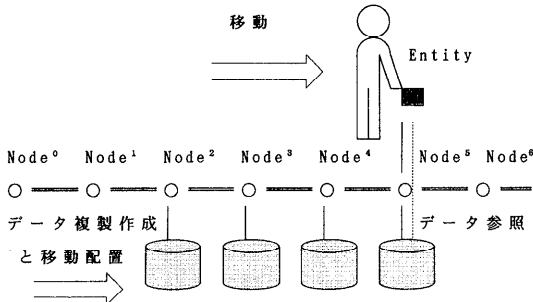


図5 データ複製配置方式
Fig. 5 Method of data replication.

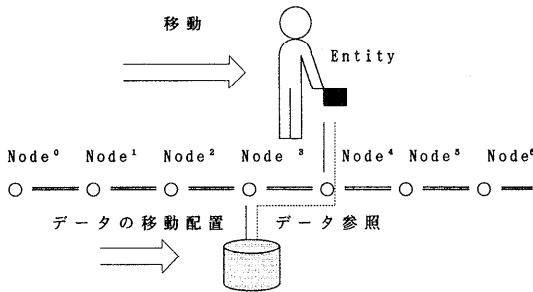


図6 データ移動方式
Fig. 6 Method of data re-allocation.

表1 複製配置方式と移動方式の主な比較
Table 1 Data replication vs. data re-allocation.

	データ複製配置方式	データ移動方式
長所	・データアクセス時間小 小	・データ配置コスト小 ・データ更新が可能
短所	・データ配置コスト大 ・データ更新が困難	・データアクセス時間大
主な課題	・データ配置コストの解析	・データアクセス時間の解析

なり^{14)~16)}現実的には困難である。一方、データ移動方式は、データ参照が同一ノード内でローカルに行われる確率が、ある程度は期待できるものの、他ノードに対して必要となる確率も高く、この際の通信ネットワーク経由でのデータアクセス時間が性能に大きく影響する。しかし、配置/移動の処理やデータの更新は複製配置方式に比較すれば単純、容易である。この両方式の主な特徴を比較すると、表1のようになる。

定性的には、データ複製配置の実現方式はデータベース技術に関連が深く、今後、この範疇の技術の進展、大容量化、コスト低下にともなって有利になっていくと考えられる。一方、データ移動の実現方式は通信ネットワークの技術に関連深く、この技術の進展、大

容量化、コスト低下にともなって有利となるといえる。

一般的にデータアクセス時間短縮のための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^{17)~24)}がある。通常ファイル配置問題の場合、レプリカ更新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、その更新コスト増大に起因して不利になることが知られている¹⁵⁾。たとえば ARPANET の例で更新比率が 40% 以下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²⁰⁾。しかし当システムの場合は特殊性があり、データは参照のみで更新不要 (使用後のデータ削除のみ) というのが基本である。また、後者のデータ移動方式は複製を持たないことからデータ更新が比較的实现しやすく、当システムに限らず移動対象の動きに追従してデータを準備し参照更新を行う一般的なアプリケーションへの適用可能性が高いと考えられる。

4.2 データ複製配置方式の解析

各時刻ごとに対象 Entity の Entry に関するデータの複製を配置するノード、すなわち複製配置セル (t, n) は、3.3 節で示したようにして求められる各セルの存在確率: $P(t, n)$ を用いて以下に示す方法で決定する。セルの選択方法はいくつか考えられ、システム上の制約により、Entity の利用開始時にセル配置をステティックに決定し時間経過によって変更はしない方式、時間経過により配置箇所を変更するダイナミックな方式等いくつか考えられ、システム側機能によって採用できる方式の制約もありうるが、一般には以下に示すダイナミックな方式がコスト上、最も有利と考えられる。

前章で述べた道草和函数: $H[A][B][C]$ を用いて、該当 Entity の、各セル (t, n) の存在率: $P(t, n)$ は式 (7) で求められる。各セルの Exit 率は式 (8) である。これが、該当セルに複製を配置した際に得られる、該当 Entity に関するデータ参照のヒット率 (そのノード内でデータ参照が可能な確率) となる。

時刻 t で、Entity が交通ネットワーク内に存在する確率: $E(t)$ は以下の式 (9') となる。

$$E(t) = \sum_{i=0}^t P(t, i) \tag{9'}$$

1セル (1ノード、1単位時間) に複製を配置するのに要するコストを "1" とすれば、セル (t, n) に複製配置を予定した際に、その実効的なコスト: $C(t, n)$ は以下となる (配置を予定しても、それ以前に Entity が

Exit することで配置が不要になる場合があるために、コストは 1 以下となる⁷⁾。

$$C(t, n) = E(t) \tag{15}$$

各セルの配置効果率は以下の式で計算される。これは、該当セルに複製を配置する場合の効果の大きさ(得られるヒット率/コスト)を示している。

$$\beta_n \cdot P(t, n) / E(t) \tag{16}$$

上記の配置効果率の大きいセルを順に選定して複製の配置箇所を決定する。選定されたセルの Exit 率の総和が得られるヒット率である。

システムに対する要求がレスポンスタイムとの関係上、ヒット率下限： h_L で与えられるとするならば、そのヒット率が得られるまで、配置効果率の高い順に選定したセル群のコストの総和が、配置コストとなる。

すなわち、 $\{t_j, n_j\}$: セル $\{t, n\}$ を配置効果率 $[= \beta_n \cdot P(t, n) / E(t)]$ の大きい順にソートした配列として、

$$h_L \leq \sum_{j=0}^J [\beta_j \cdot P(t_j, n_j)] \tag{17}$$

となるような、最小の J を決定する。

この $\{t_j, n_j\}$ ($j = 0 \sim J$) が、複製配置予定セルとなり、そのセル数は、 $J + 1$ となる。複製配置の実効コスト(実際に複製データ配置を行うセル数の期待値)は、以下となる。

$$\text{cost} = \sum_{j=0}^J E(t_j) \tag{18}$$

4.3 データ移動方式の解析

この方式でも、前章で述べた、各セル (t, n) の存在率： $P(t, n)$ を使用する(式(6), (7))。

時刻 t での、データ移動先 node： $N(t)$ は以下のよう決定する。

$$D(n) = \sum_{i=0}^t \beta_i \cdot P(t, i) \cdot |n - i| \tag{19}$$

の値を最小にするような n 。

すなわち、Exit 処理時のデータアクセスの際、経路リンク数 n (絶対値) の平均値： $D(n)$ が最小となるようなノードをデータ移動先ノードとする。

Entity が、時間 $t \sim (t + 1)$ 間にノード $\#i$ で Exit した際、その Exit 処理にともなうデータアクセスの経路リンク数は上記のように決定する N を用いて以下となる。理想的にデータアクセスが行われた場合、すなわち対象データが自ノードに存在した場合、当値は 0 である。

$$|i - N(t)| \tag{20}$$

Entity が、時間 $t \sim (t + 1)$ 間で Exit したと想定した場合、その経路リンク数の期待値： r は、時刻 t 、ノード $\#n$ での存在率： P を用いて以下となる。

$$r(t) = \sum_{i=0}^t \{\beta_i \cdot P(t, i) \cdot |i - N(t)|\} \tag{21}$$

この値が、時間 $t \sim (t + 1)$ 間での Exit 事象に対する実効的な経路リンク数と考えられる。

Entity が、(時間 $0 \sim \infty$ 間で) Exit した際の経路リンク数の期待値： R は上記の値を総和した、以下となる。

$$R = \sum_{t=0}^{\infty} \left[\sum_{i=0}^t \{\beta_i \cdot P(t, i) \cdot |i - N(t)|\} \right] \tag{22}$$

上式を計算することで、データ移動方式の際の経路リンク数の期待値が、 $\{\alpha_i\}$ 、 $\{\beta_i\}$ の関数として求まる。

5. システムの適用評価

5.1 移動モデルのパラメータ推定

前章で示した方式を現実の鉄道路線に適用し評価した例を示す。対象は表 2 にその下車率等を示した首都圏の線状の鉄道路線であり、両端の駅を含めて 19 駅からなっている。3 章で示した旅客移動モデルを、この線区に適用する。

各ノードの移動率 $\{\alpha_i\}$ 、退出率 $\{\beta_i\}$ は、下車数等から各駅ごとに表 3 のように推定される。最

表 2 首都圏 X 線における始発駅 (A 駅) 発下り利用者の各駅での下車数, 下車率 (下車人数/到着人数) 等

Table 2 Number/ratio of alighting passengers at X-line in Tokyo metropolitan area.

	A 駅	B 駅	C 駅	D 駅	E 駅	F 駅
下車数	0	765	456	9	1881	8709
---%	0	0.817	0.487	0.01	2.01	9.3
下車率%	0	0.008	0.005	0.000	0.020	0.096

G 駅	H 駅	I 駅	J 駅	K 駅	L 駅	M 駅
5715	6536	6500	10542	960	27998	809
6.1	6.98	6.94	11.25	1.025	29.89	0.864
0.070	0.086	0.093	0.167	0.018	0.543	0.034

N 駅	O 駅	P 駅	Q 駅	R 駅	S 駅	合計
691	16339	226	1787	268	3477	93668
0.74	17.44	0.24	1.91	0.29	3.71	100
0.030	0.739	0.039	0.323	0.072	1.000	

表3 実際の鉄道路線での移動率, 退出率の例

Table 3 An example of $\{\alpha, \beta\}$ in existing railways.

駅	A 駅	B 駅	C 駅	D 駅	E 駅	F 駅	G 駅
α_i	0.750	0.744	0.746	0.750	0.735	0.678	0.698
β_i	0.000	0.006	0.004	0.000	0.015	0.072	0.052

駅	H 駅	I 駅	J 駅	K 駅	L 駅	M 駅	N 駅
α_i	0.686	0.680	0.625	0.736	0.343	0.724	0.727
β_i	0.064	0.070	0.125	0.014	0.407	0.026	0.023

駅	O 駅	P 駅	Q 駅	R 駅	S 駅
α_i	0.195	0.721	0.508	0.696	0.000
β_i	0.555	0.029	0.242	0.054	0.750

終駅 (α_{18}, β_{18}) と中間駅 (α_9, β_9), (α_{11}, β_{11}), (α_{14}, β_{14}), (α_{16}, β_{16}) で相対的に退出率が大きい. 一般的に現実の鉄道路線で, よく見られるパターンと考えられる.

なお前述のように性能上の重要な要件は Entity の Exit 時の処理時間である. その評価指針となる値は対象により様々であるが, ここでは鉄道駅の改札口での処理時間: 0.1s (100ms) 以内という値³⁾を採用する.

5.2 データ複製配置方式の適用評価

データ複製配置方式の際の, 性能 (レプリカヒット率), コスト (配置レプリカ数) の解析評価の例を示す.

Entity の Exit 時のデータ処理の際にデータ参照がヒットした場合 (データの複製がその Node に存在した場合) の処理時間は, ローカルな DB の検索処理の時間である. ここでは以下とする.

$$0.05s \quad (50ms)$$

そして, ヒットしない場合の処理時間は, 中央 DB アクセスによりデータを得る時間 (ローカルでの処理時間, 中央での処理時間, 通信時間等の総和) である. これに関しては多くの要素があり定量的な決定は難しいが (中央 DB のアクセス処理の性能, 通信回線の伝送容量, 発生トランザクション数等に依存), ここでは以下とする*.

$$0.3s \quad (300ms)$$

以上のような前提の場合, 平均処理時間は達成ヒット率を h_L とすれば,

$$h_L \times 0.05 + (1 - h_L) \times 0.3 = 0.3 - 0.25h_L \tag{23}$$

これが, 先ほどの指標の 0.1s 以下であることから,

* たとえば, クレジットカード式公衆電話のクレジットカードの中央装置での受付処理に要する時間がこの程度であり, 妥当な水準と考えられる.

表4 データ複製配置方式でのセル順位の例

Table 4 An example of ranked cells in data-replication mechanism.

順位	セル(t, n)	効果率	ヒット率(累計)	実効コスト(累計)
1	(30, 18)	0.359	0.001 (0.001)	0.002 (0.002)
2	(29, 18)	0.330	0.001 (0.002)	0.004 (0.006)
3	(28, 18)	0.297	0.002 (0.004)	0.007 (0.013)
4	(27, 18)	0.261	0.003 (0.007)	0.011 (0.024)
5	(26, 18)	0.221	0.004 (0.011)	0.017 (0.040)
6	(25, 18)	0.179	0.005 (0.015)	0.026 (0.066)
7	(21, 14)	0.162	0.017 (0.032)	0.102 (0.169)
8	(20, 14)	0.161	0.022 (0.054)	0.137 (0.306)
9	(22, 14)	0.152	0.011 (0.065)	0.075 (0.381)
10	(19, 14)	0.147	0.026 (0.092)	0.180 (0.561)
⋮	⋮	⋮	⋮	⋮
67	(9, 8)	0.014	0.012 (0.804)	0.827 (21.031)
⋮	⋮	⋮	⋮	⋮

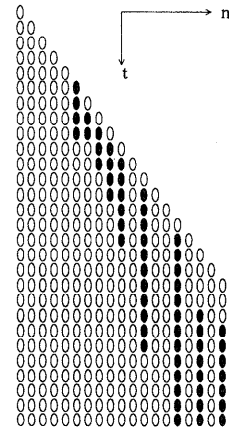


図7 レプリカ配置セルのレイアウト (1)

Fig. 7 Layout of cells where replica is distributed (1).

$$0.3 - 0.25h_L \leq 0.1 \quad h_L \geq 0.8 \tag{24}$$

ヒット率としては 0.8 以上が必要ということになる. それゆえ, 目標ヒット率 ≥ 0.8 , 時間範囲は $t = 0 \sim 30$, として解析する. 前記のような, 配置効果率順にソートしたセルと, そのヒット率, 実効コスト等を表4に示す. 効果率の大きいセルは, ある時間経過以降の最終到着駅である.

表4が示すように, 結果は以下となる.

レプリカ配置予定セル数: 67

レプリカ配置実効コスト: 21.031

達成ヒット率 : 0.804

図7にレプリカ配置予定セルのレイアウトを示す. 横軸方向がノードの軸 ($n \rightarrow$ 大) で縦軸方向が時間経過の軸 ($t \downarrow$ 大) である. 図中で ● は配置が予定されるセルを, ○ は Entity は存在し Exit は起こりうるが配置予定セルには選定されなかった箇所を示す. 大きな退出率 (β) を持つ, 最終駅といくつかの中間駅とが複製配置予定駅として選択される傾向の強いことが分かる.

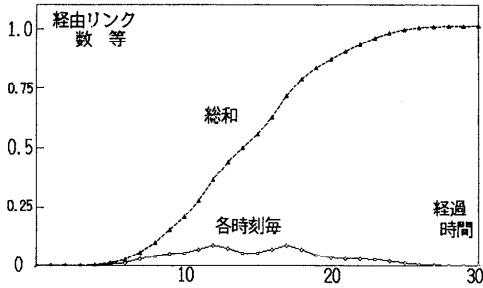


図8 実路線での経路リンク数等の例

Fig. 8 A variation of number of passing links in existing railway line.

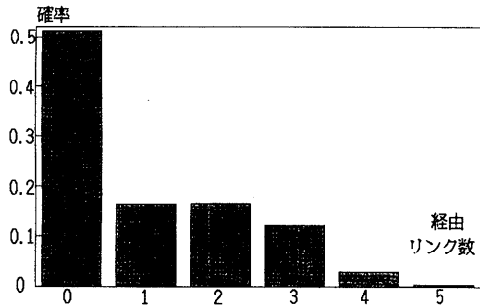


図9 経路リンク数の確率分布の例

Fig. 9 Probability of numbers of passing links.

5.3 データ移動方式での適用評価

先に示した実路線にデータ移動方式を適用評価した結果を示す。{ α_i }, { β_i } は表3のものが使用される。

各時刻の平均経路リンク数、およびその総和は、図8のように推移する。また、経路リンク数の確率分布は図9のようになる。平均経路リンク数を、4.3節のプロセスにより解析すると、1.012となった。

この場合のデータ配置先セルの、レイアウトを図10に示す。○印は Entity が存在し Exit の起こりうるセルで、●印が各時刻でのデータ移動先セルである。

このようにして求められた経路リンク数からデータ処理時間を解析する。ここで、各 Node を接続する通信機構としては、文献8)で示した、各要素の挙動が待行列モデルで近似できるような通信伝送モデルとする。文献8)に示されている手順をベースとして、データ処理時間を求める^{25),26)}。システム側の容量等は、比較的一般的と考えられる下記の前提とする。

1 パケットの正味の実伝送時間 = 0.004 [s]

(たとえば、32 Byte のパケットを 64 kbps の回線で伝送する場合は、 $T = 32 \times 8 \div (64 \times 10^3) = 0.004$)

データベース検索時間 = 0.05 [s]

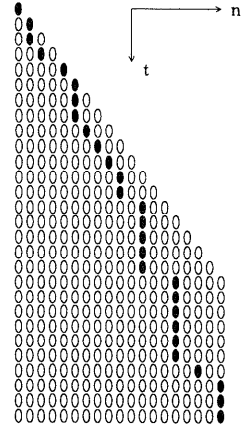


図10 データ移動先セルのレイアウト (1)

Fig. 10 Layout of cells where data is located (1).

Exit 処理のトランザクション数は、1日のピーク時利用者数 = 56,100 人/時 (15.6 人/秒) というデータから決定する。これらの条件と、先ほどの平均経路リンク数 = 1.012 とから以下が導かれる。

回線利用率:

$$15.6 \times 1.012 \times 0.004 \div 19 = 0.0033$$

リンク上の伝送に要される時間 (s):

$$1.012 \times 0.004 \div (1 - 0.0033) = 0.0041$$

データアクセス所要時間 (s):

$$0.05 + (0.05 + 0.0041 \times 2) \times 1.012 = 0.11 \text{ (110 ms)}$$

前記の、鉄道の Exit 処理時間は 100 ms 以下という指標³⁾を採用すれば、上記前提の下での試算結果は、この処理要件の、ほぼ最低水準にあるといえる。

6. 輸送特性との関連

6.1 データ複製配置方式

現実の交通ネットワークでは各 Node における Exit 率、あるいは下車率の相対的な差 (バラツキ) は様々である。鉄道路線でいえば各駅で下車数が比較的均一であるような線区と、バラツキの非常に大きい線区とが存在する。このような特性に対して、まずデータ複製配置のコスト (実効的レプリカ配置数) が、どのように影響されるか、に関していくつかの例を対象に解析を行う。たとえば、路線上各駅の利用客数が均一的な状況、すなわち各駅の下車数がほとんど均等の場合を考える。以下を前提とする。

- 線区は直線状で、両端の駅 (始発駅、終着駅) を含めて N 駅 (駅₀ ~ 駅_{N-1})、存在する。
- 始発駅で Z 人の旅客が乗車したとする。
- 各駅 (始発駅を除く) での下車数が等しいという条件から、始発駅以外下車数は以下となる。

表5 各駅の下車数が等しい場合の移動率, 退出率
Table 5 $\{\alpha_i\} \{\beta_i\}$ in case of equalized alighting passengers.

駅	A 駅	B 駅	C 駅	D 駅	E 駅	F 駅	G 駅
α_i	0.750	0.708	0.706	0.703	0.700	0.696	0.692
β_i	0.000	0.042	0.044	0.047	0.050	0.054	0.058

駅	H 駅	I 駅	J 駅	K 駅	L 駅	M 駅	N 駅
α_i	0.688	0.682	0.675	0.667	0.656	0.643	0.625
β_i	0.063	0.068	0.075	0.083	0.094	0.107	0.125

駅	O 駅	P 駅	Q 駅	R 駅	S 駅
α_i	0.600	0.563	0.500	0.375	0.000
β_i	0.150	0.188	0.250	0.375	0.750

$$Z/(N-1) \tag{25}$$

- 駅 n の下車率 (下車数/到着数) は, 以下となり, 式 (12) から, これが駅 n の $\beta_i/(\alpha_i + \beta_i)$ となる (ここで, $n = 1 \sim N-1$).

$$\begin{aligned} Z/(N-1)/\{Z-(n-1)Z/(N-1)\} \\ = 1/(N-n) \end{aligned} \tag{26}$$

- 駅間の平均所要時間は駅間ですべて等しいと仮定し, これを Tm とする. 式 (14) から,

$$Tm = 1/(\alpha_i + \beta_i) \tag{27}$$

- 上記の 2 式から,

$$\begin{aligned} \beta_n = 1/\{Tm(N-n)\} \quad (n = 1 \sim N-1) \\ \text{なお, } \beta_0 = 0 \end{aligned} \tag{28}$$

先ほどの実路線で, このように各駅の下車数が等しくなる場合の, $\{\alpha_i\}, \{\beta_i\}$ を設定する. α, β の値は表 5 のようになる. このケースでは, α_i は単調減少, β_i は単調増加となる.

当条件下でデータ複製配置方式のヒット率, コストを解析すると以下の結果が得られる. ここで目標ヒット率 ≥ 0.8 , 時間範囲は $t = 0 \sim 30$, としている.

レプリカ配置予定セル数 = 112

レプリカ配置実効コスト = 37.379

達成ヒット率 = 0.807

先ほどの例 (現実のデータに即してパラメータを設定し各駅の下車数に相当のバラツキのある場合) と比較すると, 同等ヒット率を達成するためのコストは約 2 倍に上昇している. セル配置のレイアウトを図 11 に示す. 今回の場合は, 先ほどの例 (図 7) とは異なり, 特定駅がセル配置箇所として選択される傾向はない.

次に, これとは逆に先ほどの線区で, 駅集中の度合いを, 現実以上に大きくした場合を考える. 19 の駅の下車数の割合は相対的に以下 (表 6) とする (中間

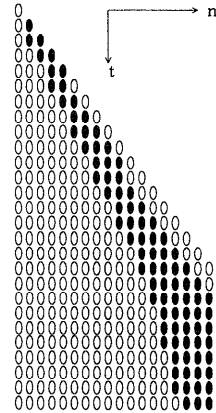


図 11 レプリカ配置セルのレイアウト (2)

Fig. 11 Layout of cells where replica is distributed (2).

表 6 バラツキを大きくした下車数分布

Table 6 An example of biased number of alighting passengers.

駅	A 駅	B 駅	C 駅	D 駅	E 駅	F 駅	G 駅	H 駅	I 駅
下車数の比	0	10	10	10	10	200	10	10	10

J 駅	K 駅	L 駅	M 駅	N 駅	O 駅	P 駅	Q 駅	R 駅	S 駅
10	250	10	10	10	10	10	10	10	147

表 7 バラツキを大きくした際の移動率, 退出率

Table 7 An example of $\{\alpha\} \{\beta\}$ when number of alighting passengers is biased.

駅	A 駅	B 駅	C 駅	D 駅	E 駅	F 駅	G 駅
α_i	0.750	0.740	0.740	0.740	0.740	0.538	0.735
β_i	0.000	0.010	0.010	0.010	0.010	0.212	0.015

駅	H 駅	I 駅	J 駅	K 駅	L 駅	M 駅	N 駅
α_i	0.735	0.735	0.734	0.349	0.715	0.714	0.712
β_i	0.015	0.015	0.016	0.401	0.035	0.036	0.038

駅	O 駅	P 駅	Q 駅	R 駅	S 駅
α_i	0.710	0.708	0.705	0.702	0.000
β_i	0.040	0.042	0.045	0.048	0.750

の 5, 10, および終端の 18 での下車数を大きく設定した).

この場合, 各駅の α, β は表 7 のようになる. 駅 No. 5, 10, 18 で, β が大きくなる.

この場合, 前記と同様に目標ヒット率 ≥ 0.8 , 時間範囲は $t = 0 \sim 30$ として解析すると以下となる.

レプリカ配置予定セル数 = 45

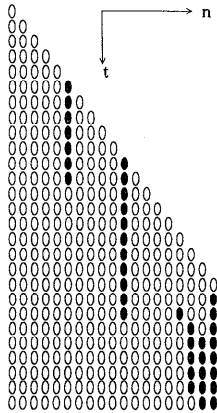


図 12 レプリカ配置セルのレイアウト (3)

Fig. 12 Layout of cells where replica is distributed (3).

レプリカ配置実効コスト = 13.372

達成ヒット率 = 0.805

前記までに示した、現実のデータに即した結果、および、各駅の下車数を均一化した結果の両者よりも、同様のヒット率を達成するためのコストは減少している。先と同様にセル配置のレイアウトを図 12 に示す。今回の場合は下車数を大きく設定した駅 (No.5, 10, 18) がセル配置予定駅として選択される傾向が強く出ている。

6.2 データ移動方式

次に、線区の輸送特性とデータ移動方式での性能、コスト (経路リンク数) との関係性を調べる。先ほどの実路線データによって設定する $\{\alpha_i\}$, $\{\beta_i\}$ の場合、結果は 5.3 節で示したように以下であった。

平均経路リンク数: 1.012

次に下車数を均等化し、表 5 で示したような α , β を適用した場合は、結果は以下となり、平均経路リンク数は 5% 程度の増加となる。

平均経路リンク数: 1.059

この場合のデータ移動セルレイアウトを図 13 に示す。

次に表 6, 表 7 で示したようなデータを適用して、下車数を現実以上に偏らせた場合 (中間 2 駅と終端駅 5, 10, 18), 結果は以下となり、前の 2 例と比較すると、平均経路リンク数は約 50% となる。

平均経路リンク数: 0.548

この場合のデータ移動セルレイアウトを図 14 に示す。

7. む す び

7.1 ユーザ ID ベースの料金徴収方式に関して

交通機関における「ユーザ ID ベースの料金徴収方式」は、種々の特徴を備え優れたサービスを実現する

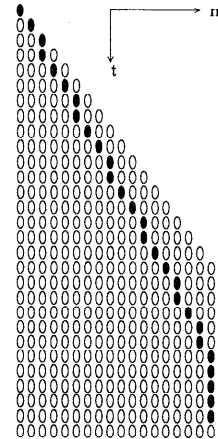


図 13 データ移動先セルのレイアウト (2)

Fig. 13 Layout of cells where data is located (2).

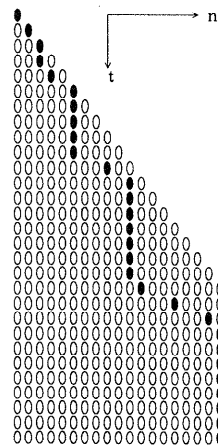


図 14 データ移動先セルのレイアウト (3)

Fig. 14 Layout of cells where data is located (3).

可能性を持っている。当論文ではこのサービスの実現を目標として、輸送特性を反映して、この解析・評価を行うための旅客移動モデルを導入し、それに基づく、「データ複製配置方式」、および、「データ移動方式」の、性能・コスト等の解析方法、鉄道の実路線のデータを使用した適用評価例を示した。前者のコストはデータベース保持、管理上のコストに大きく依存し、後者は通信ネットワークを通じたデータ通信の性能、コストに依存する特徴を持っている。

7.2 交通機関の輸送特性との関係

今回の方式では、基本的に (時間 t) \times (ノード数 n) 個存在する配置単位 (セル) を対象として、Exit 率等を解析し、データ複製配置、移動配置を予定することが有利なセルを決定している。6 章の結果でも示したように、定性的には、

- 各セルの Exit 率の値のバラツキが大きい
- Exit 率の大きい箇所では配置や移動効果率が大きい場合に、より効果的にデータ複製配置/移動セルを決定することが可能で、相対的にコストが低減できる。逆に、上記の値が各セルで均一な場合にはコスト上、不利になる。鉄道の路線でいえば、各駅の下車数のバラツキが大きい、すなわち特定の駅での乗降頻度が大きいような線区の場合にコスト上有利になる。

7.3 今後の課題

当論文では線状の交通機関を直接の対象として解析を行った。鉄道の場合、都市圏では複雑な網状となっている場合があるが、総体的には線状の独立したものも多い。また物理的形狀は一部枝状となるが、旅客の主要な流れが線状で、線状の解析で近似できる場合も多い。鉄道以外の交通機関、たとえば有料道路等に関しては線状のものが多数を占める。これらに関しては今回の結果は基本的に適用可能である。

複雑な網状のものを対象とする場合、旅客移動モデル化や解析の手法は準用でき、今回の結果がベースとなる。たとえば、移動モデルに関しては分岐ノードで方向別移動率 ($\alpha_1, \alpha_2, \dots, \alpha_n$) を実績から決定し、パラメータが増加することで計算過程は複雑化するが、各ノードでの旅客存在確率、退出確率が基本的に同様の方法で推定可能である。データ複製配置セル選択やデータ移動先ノード決定の論理についても同様の考えが適用できる。網状の場合の問題点として、旅客存在確率が広範囲に拡散することからシステム負荷が増大し、当論文で示した方式そのままでは現実性が薄れることが考えられる。より負荷を削減し、最適化を実現する方式の検討が今後の課題である。この方法としては、たとえば、

- 個人の移動特性の利用
利用者個人ごとの利用履歴情報から、個人ごとの移動特性を把握してその移動先確率の推定に利用する。
- 交通機関側の特性の利用
たとえば、鉄道輸送は列車ダイヤに基づいて提供されており、旅客移動はこれに制約される。この列車ダイヤ情報を移動先確率の推定に利用する。等が考えられる。

「ユーザ ID ベースの料金徴収方式」に関しては技術上の課題のほかに、利用者情報に関するセキュリティ確保等、別次元で重要な課題も存在する。これに関しては、たとえば文献 6), 27) 等で暗号技術の適用により利用者の ID を隠し匿名性を保つ方式が提案されている。

謝辞 交通機関の利用者へのサービス向上を実現するシステムに関して、貴重なご意見や示唆をいただいております。東京大学曾根悟教授に感謝いたします。また当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) 日本国有鉄道規格：JRS76321-5B-15AR5M(自動改札装置)，日本国有鉄道車両局(1985.10).
- 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 (1991.4).
- 5) 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収受システムについて，交通工学，Vol.26, No.4, pp.29-35, 交通工学研究会 (1991).
- 6) Tetsusaki, Y.: 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).
- 7) 田中幹夫：交通機関の自動料金徴収のための分散データ配置方式，情報処理学会論文誌，Vol.37, No.11, pp.2127-2137 (1996).
- 8) 田中幹夫：交通機関の自動料金徴収のためのデータ移動方式，電気学会論文誌 (D)，Vol.117-D, No.9, pp.1106-1114 (1997).
- 9) 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, 交通工学研究会 (1994).
- 10) 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, 交通工学研究会 (1991).
- 11) 西井和夫，古屋秀樹，坂井 努：時間軸を考慮したマルコフ連鎖モデルによる観光周遊行動分析，交通工学，Vol.31, No.5, pp.21-30, 交通工学研究会 (1996).
- 12) 宮地正和，松本憲二郎，松木孝仁：新幹線 ATC の一方向，鉄道と電気技術，Vol.6, No.7, pp.51-56 (1995).
- 13) 石塚幸治：新しい情報発信のあり方を提案する：京浜急行のインタラクティブなマルチメディアサービス，鉄道と電気技術，Vol.7, No.1, pp.41-45 (1996).

- 14) Mullender, S.: *Distributed Systems*, ACM Press (1989).
- 15) 前川 守 (編): 分散オペレーティングシステム, 共立出版 (1991).
- 16) Coulouris, G.F. and Dollimore, J.: *Distributed Systems, Concept and Design*, 分散システム, 電気書院 (1991).
- 17) Dowdy, L.W. and Foster, D.V.: Comparative Models of the File Assignment Problem, *Computing Surveys*, Vol.14, No.2, IEEE (1982.6).
- 18) Smith, A.J.: Long Term File Migration: Development and Evaluation of Algorithms, *Comm. ACM*, Vol.24, No.8, pp.521-532 (1981).
- 19) 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 (1990.5).
- 20) Fisher, M. and Hochbaum, D.: Database Location in Computer Networks, *JACM*, Vol.27, No.4, pp.718-735 (1980).
- 21) Casey, R.G.: Allocation of Copies of a file in an information network, *Spring Joint Computer Conference*, pp.617-625, AFIPS Press (1972).
- 22) 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 (1994.7).
- 23) Wolfson, O. and Milo, A.: The Multicast Policy and Its Relationship to Replicated Data Placement, *ACM Trans. Database Syst.*, Vol.16, No.1, pp.181-205 (1991).
- 24) Kumar, A. and Segev, A.: Cost and Availability Tradeoffs in Replicated Data Concurrency Control, *ACM Trans. Database Syst.*, Vol.18, No.1, pp.102-120 (1993).
- 25) Tanenbaum, A.S.: *Computer Networks*, Second Edition. タンネンバウムコンピュータネットワーク, pp.737-747, 丸善 (1992).
- 26) 吉岡良雄: ネットワークの基礎, pp.84-106, オーム社 (1991).
- 27) Chaum, D.: A New Paradigm for Individuals in the Information Age, *Proc. Symposium on Security and Privacy*, IEEE Computer Press (1984).

(平成 10 年 9 月 24 日受付)

(平成 11 年 5 月 7 日採録)



田中 幹夫 (正会員)

1953年生。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、鹿児島鉄道管理局等にて信号通信、情報設備の計画、技術開発、工事、保守に従事。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に従事。現在、同所主幹技師(情報システム担当)。電気学会会員。