Vol.2012-DPS-153 No.12
2012/11/16

車両トレースデータを用いたダミー生成による位置プライバシ保護手法

山本 彩奈[1]　岩田 麻佑[1]　原 隆浩[1]　荒瀬 由紀[2]　Xing Xie[2]　西尾 章治郎[1]

概要：GPS 技術の発展に伴いユーザの位置情報を利用した位置情報サービスが数多く提供されている．しかし，位置情報サービスは，サービス利用時にユーザの位置情報を送信する必要があり，この情報をもとに住所などの個人情報が露見してしまう可能性がある．このようなプライバシを保護するために，筆者らの研究チームでは先行研究において，移動手段を徒歩と想定し，実環境を考慮したダミーを用いたユーザ位置曖昧化手法を提案した．本稿では，徒歩とは移動特性が異なる車で移動するユーザを想定した，ダミー生成による位置プライバシ保護手法を提案する．この手法では，車で移動するダミーを実環境に適した形で生成するために，車両トレースデータを使用する．

キーワード：位置情報サービス，位置プライバシ，車両トレースデータ

1. はじめに

GPS 技術の発展に伴い，ユーザの位置に対応した情報を提供する位置情報サービスが展開されている．しかし，位置情報サービスを利用する際には，ユーザは自身の位置をサービスプロバイダへ通知する必要があり，この位置情報が流出することにより，ユーザが訪問箇所が特定され，住居や勤務先，行動パターンなどを第三者に把握される可能性が指摘されている．

このようなユーザの位置情報（位置プライバシ）の保護を目的とした既存研究は多数行われている [1][3]．その一つとして，ダミーの位置情報を利用したユーザ位置曖昧化手法がある [5]．この手法では，図 1 のように，サービスプロバイダに位置情報を通知する際，同時に複数のダミーの位置情報も送信する．それにより，送信された位置情報のうち，ユーザの位置を一意に特定することが困難になり，ユーザの位置の曖昧化が可能になる．しかし，既存の手法では，サービスの対象領域としてユークリッド平面を想定しており，実環境においてはユーザが存在できない場所にダミーが生成される可能性がある．さらに，ユーザの移動速度を考慮しておらず，直前の問い合わせにおけるダミーとの位置関係によりダミーを特定される可能性がある．このように，既存の手法では実環境における制約を十分に考慮できていない．

そこで筆者らは先行研究において，実環境に適応したダミーの生成手法を提案した [6][7]．これらの先行研究では，ユーザが連続的にサービスを使用した場合に問題となる追跡可能性という観点に，着目している．追跡可能性とは，ある時点でユーザの位置が特定した場合に，位置情報の履歴を遡ったり，逆にその後のユーザの位置情報を追跡できる性質である．先行研究では，ダミーとユーザを交差させることで，追跡可能性を低下させた．

先行研究を含む従来研究の多くは，ユーザとして歩行者を想定しており，車で移動する場合を考慮していない．ユーザが車で移動する場合，徒歩と比較すると一定時間内に移動できる範囲が広く，また目的地を定めずにぶらぶらと歩く可能性のある歩行者と異なり，目的地まで最短の経路を辿る可能性が高い．さらに，信号や一方通行といった道路によって異なる交通規則による制限も大きくなり，徒歩よりも移動できる場所が限られる．そのため，歩行者を想定した手法をそのまま適用することは困難である．

そこで本稿では，車で移動するユーザの位置プライバシを保護するためのダミー生成手法を提案する．提案手法では，交通規則や道路の状況などの実環境に適したダミーを生成するために，車両トレースデータおよび道路データ[*1]を利用する．さらに，先行研究と同様に追跡可能性を低下させるために，ユーザとダミーを意図的に交差させ

[1] 大阪大学 大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University
[2] マイクロソフトリサーチアジア
Microsoft Research Asia

[*1] 日本デジタル道路地図協会：http://www.drm.jp/

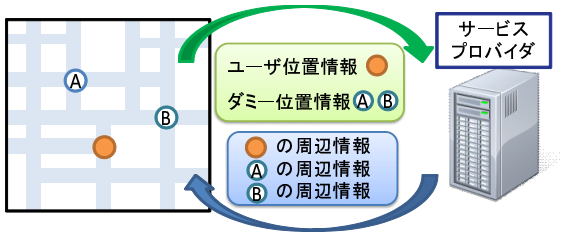

図 1 ダミーを用いた位置情報サービスの利用例

る．また，提案手法では，ユーザの移動経路および移動時間が予測可能な状況を想定し，既知であるユーザの移動経路に基づいて，実環境における制約条件を考慮したダミーの移動経路を決定する．このような想定は実環境では必ずしも妥当ではないが，ユーザが事前に移動経路をカーナビゲーションシステムで検索したり，ユーザの過去の行動履歴から予測したりなど，ある程度の精度で予測できる場合も多い．この予測の精度が低い場合の対応については，今後の課題と考え，本稿では対象としない．提案手法では，まず準備段階として，位置座標と時刻などの情報を含む車両トレースデータと，交差点などをノードとするグラフで道路網を表した道路データをマッチングし，各車両トレースデータがどの道路および交差点を通過するのかという情報を付加する．そして，道路情報を付加した車両トレースデータから，既知であるユーザの移動経路と交差するような車両トレースデータの軌跡を探し出し，ダミーの移動経路として決定する．

以下では，2 章で既存研究とその問題点について説明し，3 章でダミーを用いた位置プライバシ保護手法について述べる．最後に 4 章で本稿のまとめと今後の課題について述べる．

2. 関連研究

Lu らは，自身の位置情報と一緒に架空の位置情報であるダミー情報をクエリに付加して，サービスプロバイダにサービス要求をする位置プライバシ保護手法を提案している [5]．サービスプロバイダはクエリ中に含まれるすべての位置情報に関連する情報を返信する．返信を受け取ったユーザは自身の位置に対応する情報以外をフィルタリングし，自身の位置情報に関連する情報のみを取得できる．サービスプロバイダは位置情報群として送られてきた情報の一つ一つを区別できないため，ユーザの位置を正確に知られる可能性は小さくなる．しかし，この手法では，ダミーの生成位置に制約がなく，海などの通常ユーザが存在し得ない場所にもダミーを生成する可能性があるなど，実環境における考慮が不足していた．

筆者らは先行研究において，実環境を考慮し，ユーザのように自然に移動するダミーを生成する手法を提案している [2][6]．文献 [6] の手法では，まずダミーをユーザの周辺にグリッド状に配置することで十分な位置曖昧性を確保する．その上で道路などの実環境における制約を考慮して，各サービス利用間隔で移動可能な範囲を算出し，各ダミーの次の目的地を決定する．さらにダミー同士が定期的に目的地を交換することでグリッドを一時的に崩して交差を発生させ，追跡可能性を低下させる．さらに文献 [2] では，ユーザがいくつかの地点で停止しながら移動することを想定し，既知であるユーザの移動経路に基づいて，自然に停止しながら移動するダミーを生成する手法を提案している．これらの先行研究では，歩行するユーザを想定していた．そのため，グリッド状にダミーを配置させたり，ユーザやダミーが停止可能な地点が多く存在するという仮定をおくことが可能であった．しかし，歩行者よりも道路状況などの制約が大きい車での走行を想定した場合，これらの仮定やアプローチをそのまま適用することはできない．

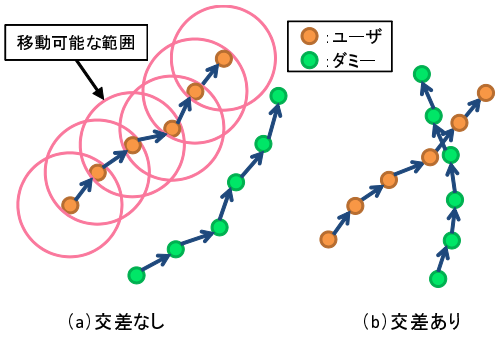

図 2 ユーザ追跡可能性

Krumm らは，車で移動するユーザを想定し，車両トレースデータを利用したダミー生成手法を提案している [4]．この手法では，収集した車両トレースデータを地図とマッチングさせ，交差点や行き止まり，道路の名前が変化する点などをノードとしたグラフを作成する．そして，各ノードにおける車両の平均移動速度や通過頻度を算出し，その情報を基にダミーが自然に移動するための経路の始点，終点，経由地，移動速度を決定し，実環境に適したダミーを生成する．しかしこの手法では，追跡可能性に関して考慮しておらず，ユーザが一度特定されてしまった場合，ユーザの移動軌跡全体を追跡できる可能性が高い．そこで本稿の提案手法では，車両トレースデータを利用することで自然な車の動きを再現しつつ，さらに追跡可能性を低下させるために，ユーザとダミーの交差を発生させる．

3. 提案手法

本章では，想定する環境とダミーを生成する際に考慮すべき制約，車両トレースデータをダミーとして使用するための前準備について述べた後，それらを考慮したダミー生成手法について説明する．

3.1 想定環境

本研究では，ユーザがある目的地を持ち，位置情報サー

ビスを連続的に使用しながら車で移動する状況を想定する．車で移動する場合，歩行者と比較すると移動可能な場所が制限され，移動経路の選択肢も減少する．例えば，一方通行の道路は逆方向から進入できず，一方通行でなくとも道路の幅によって通行不可能な場所も存在する．一方で，自動車専用道路を通行することができたり，移動速度が大きいために一定時間内に移動できる範囲が広いといった点も歩行者と大きく異なる．

また，本研究では，ユーザの目的地および移動する経路があらかじめ予測できるものと想定する．例えば，ある目的地までの経路をカーナビゲーションで検索し，その検索結果通りに移動するといった状況である．既知であるユーザの移動経路に基づいてダミーの移動経路を決定することで，自然なダミーを作成可能になる．

また，ダミーの個数についてはユーザの要求に基づいて決定するものとする．

3.2 ダミー生成の際に考慮すべき制約条件と評価指標

- **移動可能性**

サービス要求が頻発する場合，前後のクエリにおけるダミーとの位置関係を考慮する必要がある．例えば，あるユーザが一度サービス要求してから，３分後に新しくサービス要求した場合を考える．この際，３分後のクエリ中に，直前のクエリにおけるどのダミーおよびユーザの位置からも３分間で到達不可能な位置にダミーもしくはユーザが存在する場合，その位置情報はユーザではないと容易に推測できてしまう．また，本研究では，車での移動を想定しているため，ダミーが存在できる場所が道路上に限られる．さらに，進行方向によって通行できない一方通行の道路や道路の幅によって通行できない道路も存在するなど，移動できる場所も制限される．

そこで，本研究では，実際に車の移動を記録した車両トレースデータを利用することで，直前のダミーの位置から移動可能な範囲，かつ，車が通行可能な場所にダミーが生成されることを保証する．

- **追跡可能性**

短期間の連続したサービス要求の際には，ユーザの追跡可能性も考慮しなければならない．追跡可能性とは，短い時間間隔で複数の位置情報が与えられた際に，それらを結合することにより，その軌跡を推測出来てしまう性質を指す．ある特定の経路の通過など何らかの理由でユーザの位置が一旦特定された時，その前後のサービス要求時のユーザ位置まで特定されてしまう可能性がある．例えば，図２（a）のように，ユーザの移動可能範囲内をダミーが通過しない場合，ユーザ位置を一旦特定できると，ユーザの行動軌跡（前後の位置情報）を完全に追跡できてしまう．このような推測を防ぐためには，ユーザとダミーの経路が定期的に交差する方法が有効と考えられる．図２（b）のような交差により，サービスプロバイダはユーザに対応する軌跡と交差したダミーの軌跡の区別が困難になる．

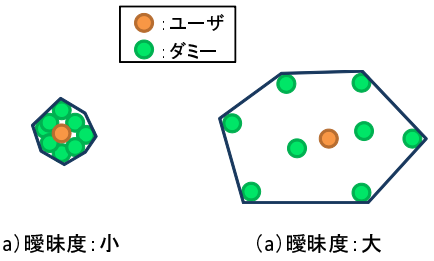

図 3 位置曖昧度

そこで，提案手法では，ユーザの経路と同じ地点を同時刻に通過する経路を持つ車両トレースデータの軌跡をダミーの移動経路とすることで，ユーザとダミーを交差させる．

- **位置曖昧度**

ユーザの位置プライバシを保護するためには，複数の位置情報から一意に特定できないだけでなく，どの程度の大きさの領域に位置情報が曖昧化されているかも重要である．例えば，ユーザ付近にダミーを配置した場合，複数の位置情報の中から，ユーザに対応する位置情報を容易に特定できない．しかし，このようなダミーの配置は，ダミーの存在範囲が小さく，ユーザの存在する可能性のある領域が小さく絞り込めてしまい，ユーザのおおよその位置が予測可能になってしまう．

そこで本研究では，ユーザとすべてのダミーを包含する凸多角形の大きさをユーザ位置の曖昧度の評価値として用いる．例えば，図 3(b) の方が多角形が大きいので，ユーザの位置曖昧性は大きい．

提案手法では，交差する経路および時間帯を散在させることで，ダミーの経路が狭い場所に集中することを回避する．

3.3 前準備

本研究では，車両トレースデータに含まれる軌跡をそのままダミーの経路として利用する．しかし，車両トレースデータは，位置座標と時刻からなる GPS データの集合体であるため，同じ道路を通過している車両トレースデータであっても，経路の座標は異なることがあり，座標ベースのまま，ユーザの移動経路に交差するような車両トレースデータを探すのは困難である．そこで，車両トレースデータをグラフ形式で表す道路データをマッチングすることにより，各車両トレースデータがどの道路および交差点を通過するのかという情報を一律的に付加する．付加した情報によって，ある交差点を通過する車両トレースデータを検索することが容易になる．

3.4 ダミーの移動経路決定方法

3.1 節で述べた想定環境下で，提案手法は，ユーザが要

求したダミーの個数，予測されたユーザの移動経路，および道路情報を付加した車両トレースデータに基づいて，既知であるユーザの移動経路と交差する車両トレースデータを探し，ダミーの移動経路とする．追跡可能性を低下させるため，交差させる経路はユーザだけでなく，ダミー同士の交差も発生させる．この際，ユーザや特定のダミー間で交差が偏らないように，ユーザや各ダミーの交差回数を記録しておき，ユーザ，ダミーの全てが同程度の交差回数となるように，ダミーの移動経路を順次決定していく．

既知であるユーザの移動経路に対して，最初に追加するダミーは以下の手順で決定する．

(1) ユーザの移動経路から交差点をランダムに選択．
(2) (1) で選択した交差点を通過する車両トレースデータを列挙．
(3) (2) で列挙した中から，各車両トレースデータの始点から (1) の交差点までの走行時間がユーザより長いもののみ抽出．
(4) 交差点までの走行時間がユーザと同程度になるよう，交差点までの経路長を調整．
(5) 調整してもユーザよりも総走行時間が長いものをダミーとして決定（複数存在する場合は，最初に見つかったものを採用）．
(6) 条件にあうトレースデータが存在しない場合，(1) ～ (4) を反復．

ユーザよりも総走行時間が長い車両トレースデータのみをダミーの軌跡として利用するのは，ユーザが位置情報サービスを利用している間，常にダミーを存在させるためである．サービス利用途中で存在しなくなるものはダミーであることが明白であり，ユーザが特定される可能性が高くなる．

２つ目以降のダミーは，ユーザおよびそれまでに決定したダミーの移動経路と交差回数を考慮し，移動経路を決定する．具体的には，まず，ユーザの移動経路における始点から目的地までの経過時間を一定時間ごとに区切り，どの時間帯で交差が起きるかを時間帯ごとの交差発生回数として記録する．この情報を基に，交差が起きる時間帯を分散させるようにダミーに対応する車両トレースデータを選択する．これにより，生成した経路が集中した場所に存在するのを防ぐことが可能となる．さらに，他の経路との交差回数（経路交差回数）をユーザおよびダミーの各経路ごとに保持する．例えば，図４のようにユーザにダミー１とダミー３がそれぞれ交差する場合，ユーザの経路交差回数は２，ダミー１とダミー３の経路交差回数は１となる．

以下で，ダミーが他の経路と交差する回数を一回とする場合，および二回とする場合のダミー決定方法について，詳しく説明する．

3.4.1　１つの経路と交差する場合

まず，前述の手順でユーザに交差するダミーを一つ決定し，どの時間帯で交差が起きたか，および経路交差回数（ここでは１回）を保持する．２つ目以降のダミーは，以下の手順で決定する．

(1) ユーザと既に決定しているダミーの中から，経路交差回数が一番少ない経路 R を選択．
(2) 経路 R から，交差が起きる回数が最も少ない時間帯に通過する交差点 C を一つ選択し，始点から交差点 C までの走行時間 T を算出．
(3) 交差点 C を通過する車両トレースデータを列挙．
(4) (3) で絞り込んだ車両トレースデータのうち，始点から交差点 C までの走行時間が T よりも長いものを選択．
(5) 選んだ経路を始点から交差点 C までの走行時間が T となるように調整．
(6) 調整された車両トレースデータから，調整後の総走行時間がユーザよりも長いものの一つをダミーとして決定．
(7) 経路 R および選択されたトレースデータにおいて，交差の起きる時間帯と経路交差回数の情報を更新．

(1) では，交差される経路が偏ることを防ぐために，経路交差回数が少ない経路に交差するように新たなダミーを追加する．また，(5) において始点から交差点 C までの走行時間が T と同程度の経路が存在する場合，(6) は行わずにダミーとして決定する．ここで，(4) および (5) で条件に合う車両トレースデータが存在しなかった場合，(2) において交差点 C を変更し，(3) ～ (6) の操作を繰り返す．経路 R の交差点すべてにおいて条件に合う車両トレースデータが存在しなかった場合，(1) において経路 R を次に交差回数が少ない経路に変更する．

以上の操作を要求されたダミー数に達するまで繰り返す．

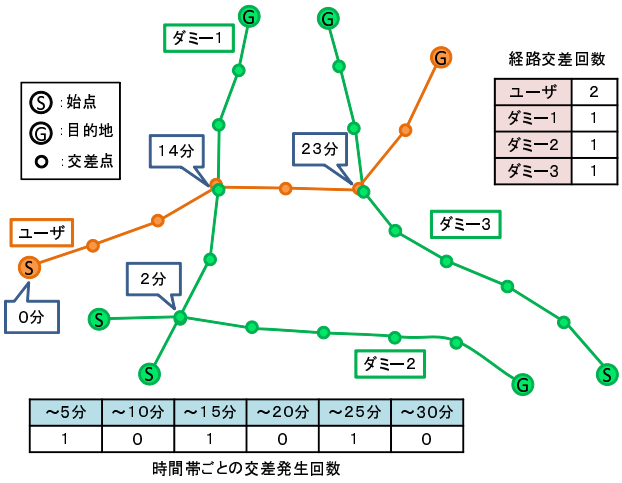

図 4　１つの経路と交差する場合

3.4.2　２つの経路と交差する場合

3.4.1 と同様に，まずユーザに交差するダミーを１つ決定し，どの時間帯で交差が起きたか，および経路交差回数

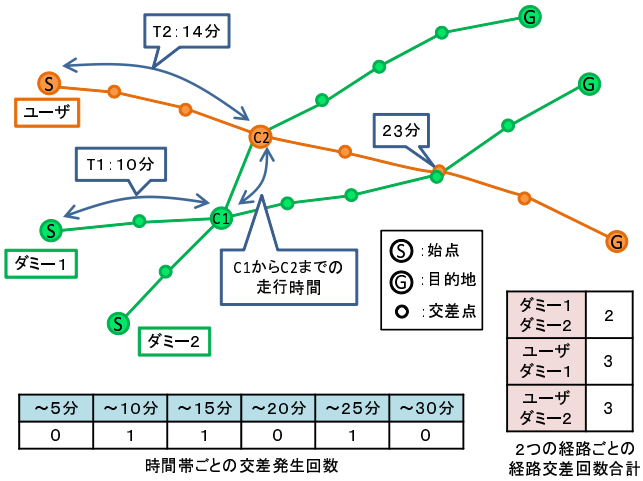

図 5　2つの経路と交差する場合

を保持する．ここで，１つ目に決定したダミーをダミー１とする．２つ目のダミーは，図５のように，以下の手順でユーザと既に決定したダミー１の両方に交差するように選ぶ．

(1) ユーザの経路に存在する任意の交差点とダミー１の経路に存在する任意の交差点の両方を通過する車両トレースデータを列挙．

(2) （1）で列挙された各車両トレースデータにおいて，ユーザおよびダミー１と交わる交差点のうち，先に通過するものを C1，後で通過するものを C2 とし，ユーザおよびダミー１の始点から C1 もしくは C2 までの走行時間 T1，T2 をそれぞれ算出．

(3) 各車両トレースデータの始点から C1 までの走行時間を算出し，T1 よりも長い場合は候補として残留．T1 よりも短い場合は候補から削除．

(4) （3）で候補に残った各車両トレースデータにおいて，C1 から C2 までの走行時間を算出し，それと T1 との合計が T2 と同程度のものを選択．

(5) （4）で選択した車両トレースデータの始点から C1 までの走行時間が T1 と同程度になるように調整．

(6) （5）で調整したものから，調整後の総走行時間がユーザよりも長い車両トレースデータの一つをダミーとして決定．

(7) 各経路において，交差の起きる時間帯と経路交差回数の情報を更新

（1）において，各車両トレースデータが交差するのはユーザおよびダミー１のどちらが先でも構わない．（4）～（5）において，条件にあうトレースデータが存在しない場合は，ユーザおよびダミー１のどちらか片方の経路に交差する経路を 3.4.1 と同様に選び出し，新たなダミーとする．

ここで，（7）における経路交差回数は，図５にあるように，ユーザと既に決定されたダミーのすべての組合せについて，経路交差回数の合計が少ない順に並べて保持しておく．

３つ目以降のダミーは，上記の経路交差回数のランキングを基に，（1）において経路交差回数の合計が少ない２つの経路の組合せから選び，（2）～（6）の手順を繰り返す．（5）までにおいて条件に合うトレースデータが存在しない場合は，新たなダミーが交差する２つの経路の組合せを変更する．以上の操作を，要求されたダミー数に達するまで繰り返す．

このように，車両トレースデータからユーザおよび他のダミーと交差するものを探し，新たなダミーとすることで，自然な車での移動を再現し，かつユーザと交差するダミーを生成することができる．

4. おわりに

本稿では，車で移動するユーザが連続的に位置情報サービスを利用することを想定した，ダミー生成による位置プライバシ保護手法を提案した．提案手法では，実環境を考慮し，自然な動きのダミーを生成するために車両トレースデータをダミーの移動経路として利用する．この際，追跡可能性を低下させるため，既知であるユーザの経路と交差を行う車両トレースデータを利用する．

本稿では，車両トレースデータをそのまま，もしくは長さのみ変更して使用する手法を検討した．しかし，車両トレースデータの長さには偏りが存在するため，全てのユーザの経路に対応して，十分な数のダミーを生成できるとは限らない．そのため，今後の拡張として，車両トレースデータを継ぎ接ぎすることで，自然な動きを保ちつつ，交差の条件に合致する十分な数のダミー経路を生成する手法を検討している．

さらに，提案手法を統計的観点，視認性観点から評価する．統計的評価では，ユーザとダミーがまばらに生成されているかという位置曖昧度，および，ユーザの追跡可能性を評価する．視認性評価では，被験者がどの程度ユーザを識別することが可能であるかを評価する予定である．評価結果から，車両トレースデータをそのまま使用する手法の有効性を検証する．

本研究では同じ交差点を通るという条件のみで交差と判断し，ダミーを決定しているため，実際は同じ交差点内で反対車線を逆方向に走行しているだけであったり，交差してもどちらがユーザなのか簡単に判断できるようなものが含まれている可能性が高い．より自然な交差を実現するためには，交差時の二つの経路の角度を考慮したり，二つの経路が一時的に並走するようにするなど，工夫すべき点が多数ある．車両トレースデータを継ぎ接ぎする際も，自然な車での移動を再現するためには，道路同士の接続の角度や進行方向を考慮する必要がある．今後はこのような点についても考慮して，提案手法の拡張を行う予定である．

謝辞

本研究の一部は, マイクロソフトリサーチ CORE 連携研究プログラムおよび文部科学省研究費補助金・萌芽研究（24650038）の研究助成によるものである．ここに記して謝意を表す.

参考文献

[1] B. Gedik and L. Liu: LocationPrivacy in Mobile Systems: A Personalized Anonymization Model, *In Proc. ICDCS*, pp. 620-629, 2005.
[2] R. Kato, M. Iwata, T. Hara, A. Suzuki, Y. Arase, X. Xie, S. Nishio: A Dummy-based Anonymization Method Based on User Trajectory with Pauses, *In Proc. ACM-GIS*, 2012.
[3] H. Kido, Y. Yanagisawa, and T. Satoh: An Anonymous Communication Technique using Dummies for Location-based Service, *In Proc. IEEE Int'l Conf' on Pervasive Services*, pp. 88–97, 2005.
[4] J. Krumm: Realistic Driving Trips For Location Privacy, *In Proc. Int'l Conf' on Pervasive*, pp. 25–41, 2009.
[5] H. Lu, C. S. Jensen, and M. L. Yiu: PAD: Privacy-Area Aware, Dummy-Based Location Privacy in Mobile Services, *In Proc. MobiDE*, pp. 16–23, 2008.
[6] A. Suzuki, M. Iwata, T. Hara, X. Xie, S. Nishio: A User Location Anonymization Method for Location Based Services in a Real Environment, *In Proc. ACM-GIS*, pp. 308–401, 2010.
[7] 鈴木晃祥, 岩田麻佑, 荒瀬由紀, 原隆浩, Xing Xie, 西尾章治郎: ダミーを用いた位置曖昧化手法の評価, 情報処理学会マルチメディア通信と分散処理ワークショップ論文集, pp. 194–199, 2011.