ユーザの移動状況に適応したダミーによる位置曖昧化手法

林田 秀平[1] 水野 聖也[2] 天方 大地[1] 原 隆浩[1] Xing Xie[3]

概要：位置情報サービス利用時にユーザの位置プライバシを保護するための手法として，ダミーによる位置曖昧化手法がある．この手法では，ユーザが位置情報サービスを利用する際，自身の位置情報と同時に複数のダミーの位置情報をサービスプロバイダに送信することにより，ユーザの位置を曖昧化している．筆者らの研究グループでは，ユーザから訪問場所，訪問順序，移動経路，停止時間，および移動速度で構成される行動プランが事前に得られるという想定のもとでダミーを生成する手法を考案した．この手法は，移動経路，停止時間，および移動速度が行動プランと異なる場合においても，実際のユーザの移動状況を考慮してダミーの行動プランを修正することにより，ユーザの位置を十分に曖昧化できることが確認されている．しかし，この手法では，訪問場所の訪問順序が異なる場合に対応できず，性能の大幅な低下を招いていた．そこで本研究では，ユーザの訪問場所の訪問順序について，事前の入力を必要とせず，任意の訪問順序に対して柔軟に対応する位置曖昧化手法を提案する．提案手法では，ユーザからの事前の入力は訪問場所の集合のみとし，ユーザの実際の移動状況に応じてダミーの行動プランを動的に修正する．シミュレーションによる評価実験の結果，任意の順序で訪問場所を訪問するユーザに対して，効果的に位置を曖昧化できることを確認した．

1. 序論

近年，スマートフォンを初めとする GPS を搭載した携帯型端末の普及により，ユーザの現在地に対応した情報を提供する位置情報サービスが数多く展開されている．例えば，ぐるなび[*1]では，ユーザの自身の現在地情報を用いて，周辺にあるレストランの検索が行える．

一般的に，位置情報サービスでは，ユーザは自身の位置情報をサービスプロバイダに送信し，サービスプロバイダは受信した位置情報に対応したサービスを提供する．この際，サービスプロバイダにはユーザの位置情報が蓄積され，その管理はサービスプロバイダに一任されている．そのため，サービスプロバイダが不正アクセスを受けたり，サービスプロバイダ自体に悪意がある場合，ユーザの位置情報が悪用され，自宅や勤務地，訪問場所および行動パターン等が露見し，位置プライバシが侵害される可能性がある．文献 [1] によると，約 50 万人のトラジェクトリデータから，あるユーザの位置情報が 4 点特定されると，95 %の確率で個人との紐づけが可能であることが報告されており，わずかな位置情報の流出でもユーザの位置プライバシが侵害される危険性は高い．そのため，位置情報サービスを有用なものとするには，ユーザの位置プライバシを保護することが重要である．

これまでに，位置プライバシの保護を目的とした研究は数多く行われている．その一つに，偽の位置情報 (ダミー) を用いた位置曖昧化手法が存在する [**?**,2–5]. この手法では，ユーザがサービスプロバイダに自身の位置情報を送信する際，同時に複数のダミーの位置情報も送信する．サービスプロバイダは受信した位置情報それぞれに対応したサービスを提供し，ユーザはその中から，自身の位置情報に対応したサービスのみを利用する．これにより，ユーザは自身の位置プライバシを保護しつつ，サービスを利用することが可能である．

文献 [2] では，ユーザの施設利用に伴う停止時間を考慮し，いくつかの訪問場所で停止しながら移動するユーザを想定したダミー生成手法が提案されている．この手法では，ユーザから訪問場所，訪問順序，移動経路，停止時間，および移動速度で構成される行動プランが事前に得られ，ユーザがそのプランに完全に従って行動するという条件下で，ユーザの位置を曖昧化できる．しかし，ユーザが行動プランに完全に従って行動することは非現実的であり，状況に応じて訪問場所への到着時刻や訪問場所での停止時間に変更が生じることが一般的である．

また，文献 [3] では，入力された行動プランに一部変更が生じるような状況を想定したダミー生成手法が提案され

[1] 大阪大学 大学院情報科学研究科 マルチメディア工学専攻
[2] サイバーエージェント
[3] Microsoft Research Asia

[*1] http://r.gnavi.co.jp/

ている．この手法では，実際のサービス利用時のユーザの移動状況を監視し，それが入力された行動プランと異なる場合には，変更後のプランの予測を行い，それに合わせてダミーの行動プランを修正する．評価実験の結果より，移動経路，停止時間，および移動速度に関しては，入力された行動プランから変更が生じた場合においても，ダミーの行動プランの軽微な修正によってユーザの位置を十分に曖昧化できることが報告されている．そのため，これら 3 つの要素については，入力が省略された場合でも簡単な予測によって補完できると考えられる．しかし，この手法では，訪問順序に変更が生じないという想定に基づき行動プランの予測を行っているため，実際に変更が生じた場合にはユーザの位置を十分に曖昧化できない．つまり，位置を十分に曖昧化するためには，ユーザが事前に入力した訪問順序に従って行動する必要がある．しかし，実際のサービス利用では，訪問場所の混雑状況などにより，ユーザが訪問場所の訪問順序を事前に入力した行動プランから変更することが考えられる．そのため，このような訪問順序に関する制約は，実環境への適用性を困難とする．

そこで本研究では，ユーザの訪問場所の訪問順序について，事前の入力やユーザがそれに必ず従うといった強い制約を排除した位置曖昧化手法を提案する．提案手法では，ユーザからの入力を訪問場所の集合のみとし，ユーザの実際の移動状況に適応してダミーの行動プランを修正する．具体的には，まず，ユーザの初期位置にもとづいて，各訪問場所への移動距離を考慮して訪問順序を予測する．そして，予測によって得られた行動プランに対して従来手法 [2] を適用し，ダミーの行動プランを生成する．その後，サービス利用時のユーザの移動状況を監視し，実際のユーザの行動が予測した行動プランと異なる場合には，行動プランの予測を再び行い，それに合わせてダミーの行動プランを修正する．シミュレーション実験の結果から，任意の順序で訪問場所を訪問するユーザに対して，ユーザの訪問順序に応じたダミーの行動プランの修正を行わない手法よりも効果的に位置を曖昧化できることを確認した．

以下では，第 2 章で関連研究について説明し，第 3 章で本研究の想定環境および位置プライバシ保護に関する要求について述べる．第 4 章で提案手法の詳細を説明し，第 5 章でシミュレーションによる評価実験について述べる．最後に第 6 章で本研究のまとめについて述べる．

2. 関連研究

2.1 Cloaking Area を用いた手法

文献 [6] では，ユーザが自身の位置情報を直接サービスプロバイダに送信するのではなく，匿名化サーバに対して位置情報を送信する．匿名化サーバは管理するユーザの位置情報集合から，k 人以上のユーザを包含するような領域 (Cloaking Area) を選択し，その領域をサービスプロバイダ

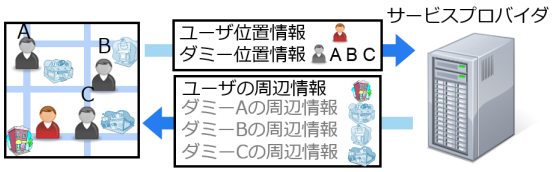


図 1　ダミーを用いた位置曖昧化手法

に送信する．サービスプロバイダは送信された領域に対応する情報を匿名化サーバに返信し，匿名化サーバはその情報をユーザに送信する．これにより，サービスプロバイダが得られる情報は Cloaking Area の範囲のみとなるため，ユーザの位置が k 匿名化される．しかしこの手法は，匿名化サーバが完全に信頼できると想定しているため，実環境への適用は困難である．

2.2 Mix Zone を用いた手法

文献 [7] では，ユーザの位置の長期的な追跡を防止するための手法が提案されている．ユーザはまず，サービス利用を行う前に自身がサービス利用を行いたい領域 (Application Zone) を第三者サーバに登録する．第三者サーバは，この領域以外の部分に Mix Zone と呼ばれる領域を生成し，その領域に同時に入ったユーザ同士で ID を入れ替える．これによりサービスプロバイダは，受信したリクエストで同一ユーザのサービス利用を対応付けることが困難になり，ユーザの連続的なサービス利用の追跡が防止される．この際，Mix Zone 内でのサービス利用を許可すると，ID を変更した場合も，移動可能な範囲の制約から入替え前と入れ替え後の ID との対応付けが可能となってしまうため，Mix Zone 内でのサービス利用は禁止される．この手法もまた，Cloaking Area を用いた手法と同様に，Mix Zone を管理する第三者サーバを完全に信頼できるという想定が必要である．

2.3 ダミーを用いた位置曖昧化手法

文献 [2–5] では，ダミーを用いた位置曖昧化手法が提案されている．この手法では，ユーザが位置情報サービスを利用する際，図 1 のようにダミーの位置情報を複数生成し，ユーザの位置情報とともにサービスプロバイダに送信する．サービスプロバイダは受信した位置情報それぞれに対応したサービスを提供し，ユーザは自身の位置情報に対応したサービスのみを利用する．サービスプロバイダはユーザとダミーを区別できないため，ユーザの位置情報を識別することが困難になり，ユーザの位置が曖昧化される．この手法では，サービスプロバイダに送信する位置情報の中に必ず自身の位置情報が含まれているため，サービスの質を低下させることはない．また，ダミーの生成は全てユーザの端末上で行えるため，第三者サーバを必要としない．以上の理由から，本研究ではダミーを用いた手法を採用する．

文献 [5] では，ユーザの各サービス利用時点までの位置情報に基づいて，ダミーを生成する手法を提案している．し

かし，この手法では，ユーザは停止することなく移動し続けるという行動モデルを想定をしているため，ユーザが停止する場合に対応できない．文献 [2] では，ユーザは事前に自身の行動プランを入力し，複数の訪問場所で停止しながら移動するという想定のもと，ダミーを生成する手法を提案している．しかし，ユーザが入力した行動プランに完全に従うという前提があるため，非現実的である．文献 [3] では，ユーザが，入力した行動プランから変更して行動する場合を考慮しているが，訪問場所の訪問順序が異なる場合には，ユーザの位置を十分に曖昧化できない．提案手法では，実際の訪問順序が予測と異なると判断した際に，入力された各訪問場所を訪問するために必要な移動距離を考慮して行動プランを再び予測し，ダミーの行動プランを修正する．これにより，任意の訪問順序への対応を図る．

3. 想定環境およびダミー生成時の制約と要求

3.1 想定環境

本研究では，Google Map*2のように，サービスプロバイダに頻繁にユーザの位置情報を送信する位置情報サービスを想定する．また，ユーザは，いくつかの訪問場所を訪問しながら移動する状況を想定する．ユーザは訪問場所に到着した際，施設の利用を行うために最小で pt_{min}[sec]，最大で pt_{max}[sec] の間停止し，訪問場所に移動する際，最短路もしくは目的地に近づく経路を通って v_{min}[m/sec] から v_{max}[m/sec] の範囲の一定の速度で移動する．

ユーザは，事前に自身が訪問する予定である訪問場所の集合を入力し，本研究における提案手法を用いてダミーを生成することで自身の位置情報を保護する．ユーザは入力したすべての訪問場所に必ず 1 度ずつ訪問するものとし，ダミーの初期生成は，ユーザが最初の訪問場所に到着した際に行う．すなわち，ダミー生成時に最初の訪問場所は既知であるとする．

なお，ユーザが所有するモバイル端末は，地図情報を保持しており，ユーザが歩行可能な道路および訪問場所の情報を全て把握しているものとする．

3.2 実環境におけるダミー生成時の制約

実環境では海上や高速道路など，歩行するユーザが存在し得ない領域が存在する．このような領域にダミーを生成すると，それがダミーであると容易に特定されてしまう．そのため，ダミーは実際のユーザが存在し得る領域に生成する必要がある．また，短時間で連続的に位置情報サービスを利用する場合，直前のサービス利用を行った位置からの地理的な到達可能性を考慮する必要がある．よって，ダミーは，前回のユーザのサービス利用時に生成した位置から移動可能な範囲内に生成しなければならない．なお，移動可能な範囲は，マンハッタン距離に基づいて計算する必要がある．

*2 https://www.google.co.jp/maps

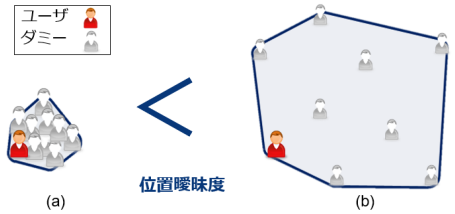


図 2 匿名領域

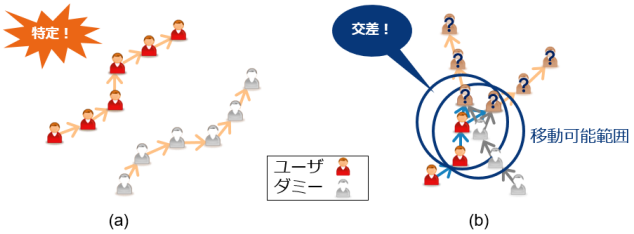


図 3 追跡可能性

3.3 ユーザの位置プライバシ保護に関する要求

- **匿名領域 [4]**：ユーザの位置プライバシを保護するためには，どの程度の大きさの領域にユーザの位置が曖昧化されているかが重要である．例えば図 2(a) のように，ユーザ付近に密集してダミーを配置した場合，ダミーの存在範囲が小さいため，おおよその位置が推測できてしまう．そのため，図 2(b) のように，ダミーを広範囲に配置させる必要がある．そこで本研究では，文献 [4] の定義に基づき，全てのエンティティ (ユーザもしくはダミー) で形成される凸包領域を匿名領域 [4] と定義し，この大きさが，ユーザの要求する大きさを満たせるようにダミーを配置する．
- **追跡可能性 [8]**：ユーザが短時間に連続して位置情報サービスを利用する場合，ユーザのサービス利用の追跡可能性も考慮しなければならない．例えば，図 3(a) ように，あるサービス利用時刻におけるエンティティの位置から次のサービス利用の時刻までに対応する位置情報が一つしかない場合，前後のサービス利用の対応関係が一意に定まることから特定のエンティティの移動軌跡が追跡できてしまう．この場合，ユーザが一度特定されてしまうと，その前後にサービスを要求した位置まで特定され，より多くの情報流出を招く．このような追跡を防ぐには，図 3(b) のように，ユーザとダミーの移動軌跡を交差させることが有効であり，追跡可能性を低下できる．

4. 提案手法

4.1 概要

図 4 に，提案手法によるダミーの行動プランの修正の流れを示す．提案手法ではまず，ユーザがサービス利用を開始する前に，(1) ユーザから入力される訪問場所の集合をもとに，(2) それらの訪問順序を予測する．(3) 予測によっ

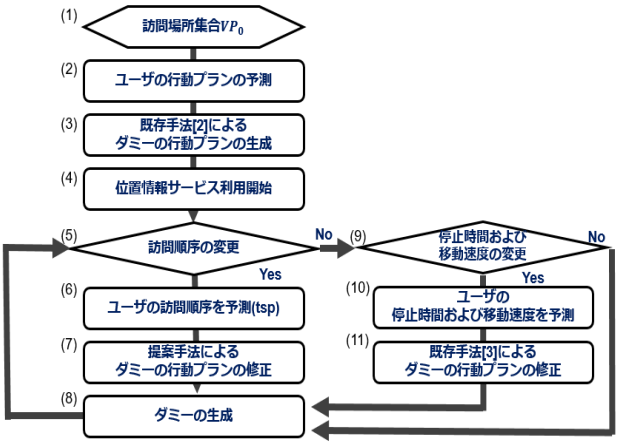


図 4　提案手法のフローチャート

て得られた行動プランに対して既存手法 [2] を適用し，ダミーの行動プランを生成する．(4) ユーザが位置情報サービスの利用を開始すると，位置情報を送信する各時刻でユーザの移動状況を確認し，(5) 予測した訪問順序に誤りが生じているかを判断する．もし予測した訪問順序に誤りがあると判断した場合，(6) 各訪問場所を訪問するために必要な移動距離を考慮して訪問順序を再び予測し，(7) それに基づいて，ダミーの行動プランを修正する．予測した通りの訪問順序で移動していると判断した場合，4.2.2 項で説明する既存手法 [3] に基づき，(9) 停止時間および移動速度の変更を確認し，変更が生じている場合，(10) それらを再予測し，(11) ダミーの行動プランを修正する．

以降では，提案手法で用いる既存手法 [2,3] の概要について説明し，その後，任意の訪問順序の変更に対応するためのダミーの行動プランの修正方法について述べる．

4.2　既存研究 [2,3] の概要

4.2.1　文献 [2] におけるダミー生成手法

既存研究 [2] では，要求ダミー数 d，要求匿名領域の大きさ ra[m^2]，訪問場所，訪問順序，移動経路，停止時間，および移動速度から構成されるユーザの行動プランに基づき，d 個分のダミーの行動プランを生成する．ここで，エンティティ e_j の行動プラン MP_j は，i 回目の位置情報送信時の位置情報 p_i，およびその時刻 t_i の組 $< p_i, t_i >$ で構成される位置情報系列 P_j，各時刻における移動速度の系列 V_j，および停止時間の系列 PT_j で表される．

$$
\begin{aligned}
MP_j &= \{P_j, V_j, PT_j\},\\
P_j &= \{< p_1, t_1 >, < p_2, t_2 >, \cdots, < p_l, t_l >\},\\
V_j &= \{v_{1,2}, v_{2,3}, \cdots, v_{l-1,l}\}, \quad v_{min} \leq \forall v \in V_j \leq v_{max},\\
PT_j &= \{pt_1, pt_2, \cdots, pt_l\}, \quad pt_{min} \leq \forall pt \in PT_j \leq pt_{max}
\end{aligned}
$$

本研究では，ユーザの訪問順序は任意であると想定しているため，ユーザからは訪問場所の集合 VP_0 のみが得られる．そのため，既存手法 [2] を適用するには，訪問順序，移動経路，停止時間，および移動速度を補完する必要がある．提案手法では 4.3.1 項で説明する方法に基づいてこれらを予測し，その結果得られる行動プランを入力として用いる．

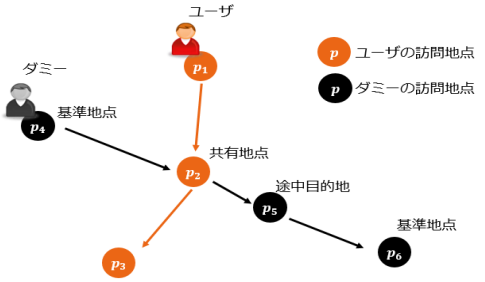


図 5　文献 [2] のダミー生成手法

既存手法 [2] では，各ダミーの行動プランを 1 つずつ確定させていくことによりダミーの生成を行う．この際，k 番目のダミー e_k の生成には，ユーザおよび生成済みダミーの集合 $E = \{e_i | 0 \leq i < k\}$ の行動プランを考慮する．具体的には，各ダミーの行動プランは，以下の手順で決定される．

(1) 匿名領域確保のための基準地点およびその到着時刻の決定
(2) 追跡可能性低下のための，エンティティと共有する地点およびその到着時刻の決定
(3) 自然な移動経路生成のための途中目的地およびその到着時刻の決定

手順 (1) では，要求された匿名領域を確保するため，ダミーが広範囲に分布するように訪問場所を定める．まず，ユーザの総移動時間を，周期 C[sec] で分割する．各周期において，生成済みエンティティの平均位置を計算し，その平均位置を中心とし，一片の長さが $\sqrt{ra}$[m] である 3×3 のグリッド領域を生成する．各時刻についてグリッド中の各セルに含まれる生成済みエンティティの数を計上し，これが最も少ないセル領域とその時刻に対し，基準地点を設定する．手順 (2) では，追跡可能性を低下させるための訪問場所を設定する．手順 (1) で生成した点から，生成済みエンティティの各訪問場所に対する到達可能性を計算し，到達可能な場所を可能な限り多く共有地点に設定する．

上記 2 つの手順によって設定される点を単純に結ぶだけでは，ダミーの移動速度が極端に遅くなるなど，ダミーの行動が不自然になる場合が生じる．これを防止するため，手順 (3) において，基準地点および共有地点の間に途中目的地を適宜設定し，ダミーの移動が自然なものとなるように行動プランを調整する．以上の手順に基づいて決定した訪問場所を図 5 のように接続し，ダミーの行動プランとする．

4.2.2　文献 [3] におけるダミー行動プランの修正手法

ユーザの行動プランの変更には，訪問順序の変更以外にも，移動経路，停止時間，および移動速度の変更が考えられる．この場合，提案手法では，既存手法 [3] に従ってダミーの行動プランを修正する．以下で，既存手法 [3] における行動プラン変更の判断方法と，その際のダミーの行動

プランの修正方法をそれぞれ説明する．

ユーザの行動プラン変更の判断方法. 既存手法 [3] では，実際のユーザの移動状況をもとに，行動プランにどのような変更が生じているかを判断し，次の訪問場所への到着時刻を予測する．以下に，移動経路，停止時間，および移動速度それぞれについて，行動プランからの変更の判断方法と，次の訪問場所の到着時刻の予測方法を示す．

- 移動経路. 行動プランと異なる経路上にユーザが位置した場合，移動経路が変更されたとみなせる．この場合，ユーザが現在地から次の訪問場所までの最短路を通って移動する時間を移動時間とみなし，到着時刻を予測する．
- 停止時間. 行動プランでは移動を開始している予定であるにも関わらず，実際のユーザが訪問場所に位置している場合，停止時間が増加したとみなせる．この場合，停止時間の増分だけ次の訪問場所への到着時間が遅れると予測する．
- 移動速度. 行動プラン通りの経路上に位置しているが，その位置が行動プランと異なる場合，移動速度を変更したと判断する．その場合，直前まで停止していた訪問場所を行動プラン通りに出発したと仮定し，直前まで停止していた訪問地点と現在の位置から移動速度を求める．次の訪問場所までその移動速度で移動するとみなし，到着時刻を予測する．

ダミーの行動プランの修正方法. ユーザの実際の行動が行動プランと異なると判断された場合，ユーザの次の訪問場所への到着時刻と，元の行動プランにおけるその到着時刻の差分 Δ[sec] を求め，ダミーの行動プランも Δ だけ変化するように行動プランを修正する．具体的には，変化量の総和が Δ となるように，MP_j 中に含まれる訪問場所の内，現在時刻に近い訪問場所から順に停止時間，移動経路，および移動速度を変更する．この際，停止時間および移動速度の変化量が極端に大きいと，動作が不自然になり，ダミーが特定されてしまう可能性がある．そのため，これらの変化量には上限を設け，その範囲内で修正を行う．

まず，図 6 のように，現在位置から最初に訪問する訪問場所 p_i を取得し，この訪問場所での停止時間を変化させることで，到着時刻を修正する．この変更による到着時刻の変化量が Δ に満たない場合，移動経路の変更を行う．p_i から次の訪問場所 p_{i+1} に近づく移動経路を探索し，到着時刻の変化量が Δ に最も近づく移動経路に変更することで，行動プランを修正する．それでも到着時刻の変化量が Δ に満たない場合，移動速度を変更する．この修正を，変化量が Δ になるまで，そのダミーが訪問する全ての訪問場所について繰り返す．もし全ての訪問場所に対して修正を行っても変化量が Δ に満たない場合，訪問場所の数を増減させる．$\Delta > 0$ の場合，p_i から p_{i+1} に近づく範囲に存在する訪問場所から，無作為に訪問場所を設定する．一方，$\Delta < 0$ の場合，訪問場所 p_{i+1} を VP_j から削除し，直接 p_{i+2} に向かうように行動プランを修正する．

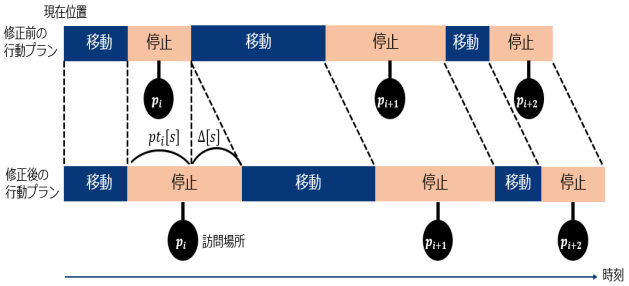


図 6　文献 [3] の手法によるダミーの行動プランの修正方法

4.3　提案手法におけるダミー行動プラン修正手法

本節では，任意の訪問順序に対応可能なダミーの行動プランの修正方法 (提案手法) について詳述する．

4.3.1　ユーザ行動プランの予測方法

本研究では，ユーザが，行動プランの入力として訪問場所の集合 VP_0 のみを入力する状況を想定する．そのため，既存手法 [2] を適用するには，訪問順序，停止時間，および移動速度を予測により補完する必要がある．そこで，これらの値を以下の方法に基づいて補完する．

- 訪問順序．ユーザは自身の移動距離がなるべく小さくなる訪問順序で訪問場所を移動すると仮定し，始点を p_0，終点を未定とした巡回セールスマン問題を解くことによりユーザの訪問場所の訪問順序を予測する．すなわち，訪問順序は総移動距離が最小となる訪問順序と推定する．
- 停止時間．各訪問場所での停止時間は，$[pt_{min}, pt_{max}]$ 内の一様分布に基づき任意の値を一つずつ定める．
- 移動速度．各訪問場所間の移動の速さは，停止時間と同様に，$[v_{min}, v_{max}]$ における一様分布に基づき任意の値を一つ定める．

4.3.2　ユーザの訪問順序の予測方法

位置情報を送信する各時刻において，ユーザが現在向かっている訪問場所を予測し，その訪問場所が予測した行動プランにおける訪問場所と異なる場合，訪問順序が予測と異なると判断する．各サービス利用時のユーザの向かっている訪問場所 vp_c は，VP で未訪問の訪問場所のうち，前回のサービス利用位置からの距離が最も短い訪問場所とする．つまり，$dist(a, b)$ を地点 a から地点 b までのマンハッタン距離および i 回目のサービス利用時の位置情報を p_i とすると，式 (1) のように表せる．

$$vp_c = \min\{vp_j \in VP_j | dist(p_i, vp_j) - dist(p_{i-1}, vp_j)\} \tag{1}$$

なお，その後の各訪問場所の予想訪問順序は，vp_c を始点とした巡回セールスマン問題の解とする．

4.3.3　ダミー行動プランの修正手法

方針. 提案手法では，既存手法 [2] に基づいてダミーの行

動プランを生成し，実際のユーザの移動状況に応じて，訪問順序を予測する．その後，ダミーの行動プランを修正することにより，訪問順序が予測と異なることにより生じるユーザの交差回数の減少分の補填を試み，追跡可能性の増加の防止を図る．この際，修正の対象であるダミーは，初めに匿名領域や追跡可能性を考慮して生成した行動プランを可能な限り維持するため，行動プランを変更する際，基準地点と共有地点に影響する数が最も小さくなるダミーを選ぶ．

ダミーの行動プラン修正手順. ダミーの行動プランを修正する手順の詳細をアルゴリズム 1 に示す．まず最初に，既存手法 [2] で設定されたユーザの交差回数を取得する (1 行)．そして，予測したユーザの行動プランにおける交差回数が取得した回数と等しくなるまで，以下の手順で繰り返しダミーの行動プランの修正を試みる．(2-12 行)

(1) 予測したユーザの行動軌跡と交差させるダミーおよび交差場所を決定．

(2) 手順 (1) で決定したダミーの行動プランを交差場所を含むように修正．

手順 (1) では，交差を発生させるためのダミーの行動プランの修正をシミュレートし，行動プランを修正する際に消去される基準地点と共有地点の個数が少ないダミーを求める．具体的に図 7 を用いて説明する．

図 7 のように，ダミーの行動プランを，あるユーザの予想行動プラン中に含まれる訪問場所である $< p_i, t_i >$ を含むように修正する状況を考える．まず，ユーザと交差させる時間およびその場所 $< p_i, t_i >$ を取得する (6 行)．その後，ダミーの行動プラン中に含まれる時刻 t_i の直前の訪問場所である $< p_h, t_h >$ を取得する．もし，p_h から p_i への移動時間が $t_i - t_h$ 以内である場合，訪問場所を減少させることなく，$< p_i, t_i >$ を含むように修正可能であるが，図 7 のように $t_i - t_h$ より大きい場合，時刻 t_i に訪問場所 p_i には到達できない．そのため，行動プラン中に $< p_i, t_i >$ を含むには，$< p_h, t_h >$ を消去しなければならない．この際，その訪問場所が基準地点か共有地点である場合，消去される数としてカウントする．この動作を，t_i 以内に p_i に到達できる訪問場所が見つかるまで逆上り，消去される訪問場所の数を求める (7 行)．ただし，4.2.1 項で述べた自然な移動経路生成のための途中目的地が消去される場合，匿名領域や追跡可能性に対する直接的な影響は少ないと考えられるため，消去される数には含まないものとする．同様の操作を，時刻 t_i 直後の訪問場所 p_j に対しても行う (8 行)．なお，交差場所では少なくとも最低停止時間 pt_{min} は停止しなければならないため，時間内に到着できるかを計算する際は，pt_{min} も考慮する．このようにして，交差場所 $< p_i, t_i >$ を含むために消去しなければならない基準地点と共有地点の総数を取得する．この操作をユーザの予想行動プラン中に含まれる全ての訪問場所および全てのダミーに対して行い，最終的に，消去数が最も少ないダミーの行動プランを，その際求めた交差場所を含むように修正する．

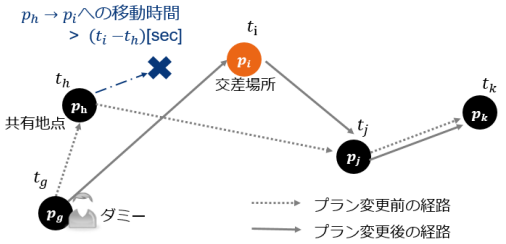


図 7　提案手法におけるダミーの行動プラン修正方法
$(0 < t_g < t_h < t_i < t_j < t_k)$

Algorithm 1: Proposed Method

input : MP_0:User trajectory, d: Number of dummy

1 $original\_cross\_times \leftarrow$ getOriginalCrossTime(MP_0)
2 **for** $j = 0$ **to** $original\_cross\_times$ **do**
3 　$ST \leftarrow$ all time when the user visits POIs
4 　**for** $\forall st_j \in ST$ **do**
5 　　**for** $dummy\_id = 1$ **to** d **do**
6 　　　$< p_i, t_i > = < MP_0[st_j], st_j >$
7 　　　$num_{dummy\_id} =$ getPrevDeletedPosNum($MP_{dummy\_id}, < p_i, t_i >$)
8 　　　$num_{dummy\_id} +=$ getAfterDeletedPosNum($MP_{dummy\_id}, < p_i, t_i >$)
9 　　　**if** $num_{dummy\_id} \leq num_{min}$ **then**
10 　　　　$num_{min} = num_{dummy\_id}$
11 　　　　$dummy\_id_{min} \leftarrow dummy\_id$
12 　　　　$< p_s, t_s > \leftarrow < p_i, t_i >$
13 　ReviseDummyPlan($MP_{dummy\_id_{min}}, < p_s, t_s >$)

手順 (2) では，手順 (1) で決定したダミーに対し，$< p_i, t_i >$ を含むように行動プランを修正する (12 行)．この際，交差場所での最大停止時間 pt_{max} を考慮しなければならない．これは，交差場所で pt_{max} 以上の停止を行った場合，それがダミーであると容易に推測できるためである．もし，交差場所を含むようにダミーの行動プランの修正を行う際，交差場所での停止時間が pt_{max} を超える場合，交差場所と交差場所の直前の訪問場所の間の近づく範囲内で，自然な経路生成のための途中目的地を設ける．これにより，交差場所での停止時間が pt_{max} を超えないように停止時間を設定する．

5. 評価実験

5.1 実験環境

東京 23 区の地図上で，ユーザのサービス利用をシミュレートした評価実験を行った．シミュレーションにおける各パラメータは表 1 のように定めた．訪問場所には，FourSquare API[*3]を用いて取得した Venue を割り当て，実際の訪問場所の分布を地図上に再現した．ユーザデータに

[*3] https://developer.foursquare.com

表 1　シミュレーションにおけるパラメータ

パラメータ	値
地図領域 [km^2]	76.644
訪問場所の数	22354
交差点の数	26102
ユーザの訪問場所の数	7
サービス利用間隔 [sec]	180
歩行速度 [m/sec]	[1.05, 1.55]
停止時間 [sec]	[600, 1800]
ダミー数	9
要求匿名領域 [km^2]	[1000, 2000]

は，50 個の人工的に作成した軌跡を用いた．各軌跡は，地図領域中から訪問場所を 7 つランダムに選択し，各訪問場所をランダムな順序で訪れるものとし，4.3.1 項と同様の方法で，移動速度および停止時間を決定した．

5.2　評価指標

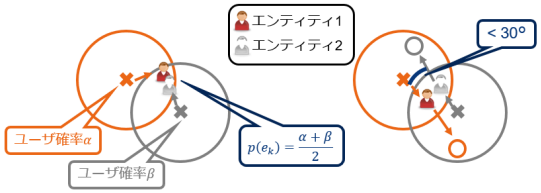
本評価では以下の 3 つの性能指標を用いた．

- **AR-Size(Anonymous area achieving Ratio - Size)**．要求匿名領域に対し，ダミーの配置によって実際に確保できた匿名領域の平均面積の割合を AR-Size と定義する．AR-Size が 1 より大きければ，平均的にユーザの要求以上に位置を曖昧化できたと見なせる．
- **MTC(Mean Time to Confusiton) [8]**．攻撃者から見たある位置情報がユーザのものである確率をユーザ確率と呼ぶ．ある時点において，目撃等によりユーザが特定された場合，ユーザ自身のユーザ確率は 1 となり，他のダミーのユーザ確率は 0 となる．ここで，エンティティ間で交差が発生した場合，交差前後の遷移の対応関係が一意に定まらなくなるため，それらのエンティティ間でユーザ確率が等配分される．すなわち，ユーザ確率 α のエンティティとユーザ確率 β のエンティティが交差した場合，交差後の両エンティティのユーザ確率は $\frac{\alpha+\beta}{2}$ となる．ただしこの際，図 8 のように，交差の角度が 30° 以下となるような交差の場合，進行方向に基づく追跡が可能であると考えられるため，ユーザ確率の配分は行わない．

  この条件下で MTC を，ユーザが特定された時点から各エンティティ e_k のユーザ確率 $p(k)$ のエントロピー $H = -\sum p(k)\log_2 p(k)$ が 1 を超えるまでにかかる平均時間と定義する．すなわち，サービス利用の総数を L，時刻 t_i でユーザが特定されてから $H > 1$ となるまでにかかる時間を $TC(t_i)$ とすると，MTC は式 (2) で計算される．

$$MTC = \frac{1}{n_{achieve}} \sum_{i=1}^{L} TC(t_i) \qquad (2)$$

  ただし，$TC(t_i)$ は，時刻 t_L までに $H > 1$ が満たせない場合は 0 と見なし，MTC の計算には考慮しない．



(a)交差に判定する場合　　(b)交差に判定しない場合

図 8　交差の判定方法

  ここで，$n_{achieve}$ は，最後のサービス利用時刻 t_L までに $H > 1$ を達成できたユーザ特定時刻 t_i の数である．この指標は，ユーザが特定されてから再び曖昧化されるまでにかかる平均時間であるため，この値が小さければ追跡可能性が低いことを表す．

- **CR(Confusion achieving Ratio)**

  式 (2) における L，$n_{achieve}$ を用いて，追跡可能性を示す指標として CR を式 (3) で定義する．

$$CR = \frac{n_{achieve}}{L} \qquad (3)$$

  CR は，ユーザの行動プラン全体のうち，最初のサービス利用からどの程度の割合のサービス利用までのユーザ特定に対し，最終サービス利用時刻までにユーザの曖昧性を回復することができたかを表す．そのため，この値が高いほど追跡可能性が低いことを意味する．

5.3　評価手法

本実験では，以下に示す 3 つの手法の性能を比較した．

- **No-Revise**：巡回セールスマン問題の解として予測したユーザの行動プランに既存手法 [2] を適用してダミーを生成する手法．ただし，ユーザの訪問順序に応じたダミーの行動プランの修正は行わない．
- **Proposed**：4 章で述べた，ユーザの移動状況に適応したダミーの行動プランの修正を行う提案手法．
- **Given-Actual**：ユーザの行動を既知とし，そのユーザに既存手法 [2] を適用してダミーを生成する手法

5.4　評価結果

5.4.1　匿名領域について

要求匿名領域に対する AR-Size を図 9 に示す．横軸は要求匿名領域であり，縦軸は AR-Size である．この結果から，Proposed，No-Revise，Given-Actual の順で AR-Size が大きくなることがわかる．これは，Proposed では修正を行ったダミーが，No-Revise ではユーザが，他のエンティティ群から離れて行動するためである．図 10 はダミーのみで形成される匿名領域を用いた際の要求匿名領域に対する AR-Size の変化を表す．図 10 より，Given-Actual および No-Revise よりも Proposed の値の方が大きいことから，ダミーの行動プランの修正によって，ダミーは広範囲に分布することがわかる．また，図 9 と図 10 における Proposed と No-Revise の AR-Size の値の差を比較すると，図 10 における差分の

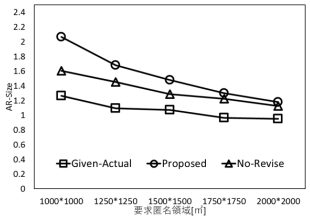


図 9 要求匿名領域に対する AR-Size

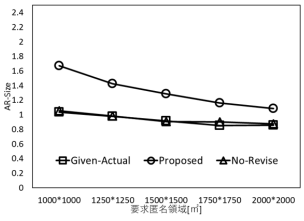


図 10 要求匿名領域に対するダミーのみの AR-Size

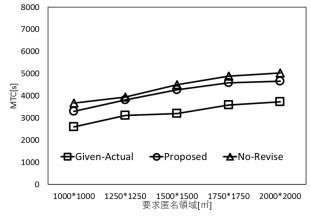


図 11 要求匿名領域に対する MTC

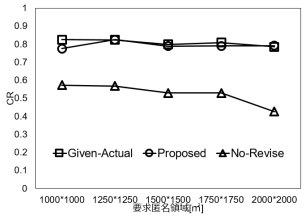


図 12 要求匿名領域に対する CR

方が大きい．この結果より，Proposed はダミーの行動プランをユーザと交差するように修正することで，ユーザが，全体のエンティティ群から離れて行動することを抑制していることがわかる．

5.4.2 追跡可能性について

要求匿名領域に対する MTC および CR をそれぞれ図 11 および図 12 に示す．Proposed と No-Revise を比較すると，図 11 より，Proposed の方が，より短い時間でユーザの位置を曖昧化できることがわかる．また，図 12 から，Proposed は No-Revise よりも，サービス利用全体においてユーザの位置を曖昧化できることがわかる．一方，Proposed と Given-Actual を比較すると，サービス利用全体においてユーザの位置を曖昧化できる回数の割合はあまり変わらないが，ユーザの位置を曖昧化できるまでの平均時間は Proposed の方が大きいことがわかる．これは，両手法において，交差が生じるタイミングが異なるためである．

図 13 および図 14 は，シミュレーション経過時間の前半および後半における，ユーザとダミーが訪問場所を共有した数 (交差した数) を示すグラフである．横軸は要求匿名領域，縦軸はユーザとダミーが訪問場所を共有した数を表す．図 13 および図 14 より，Given-Actual では，経過時間の前半と後半の交差回数の差のばらつきは小さいが，Proposed では，交差が後半に集中していることが確認できる．Proposed では，ユーザの訪問順序を予測することでダミーの行動プランを修正しているが，ユーザが完全にランダムな順序で移動する場合，予測がほぼ当たらず，前半部分での交差が設定できないためであると考えられる．

6. おわりに

本研究では，ユーザの任意の訪問順序に対してユーザの位置を曖昧化できるように，ユーザの移動状況に応じてダミーの行動プランを修正する手法を提案した．FourSquare のデータセットを用いたシミュレーションによる評価実験

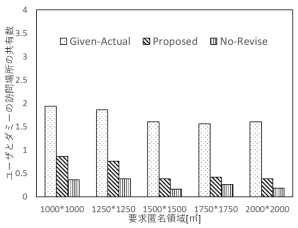


図 13 経過時間の前半におけるユーザとダミーの訪問場所の共有数

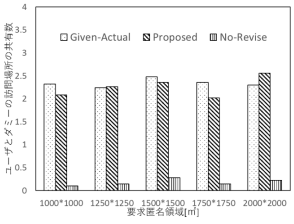


図 14 経過時間の後半におけるユーザとダミーの訪問場所の共有数

を行った結果，提案手法は，ユーザの移動状況に応じたダミーの行動プランの修正を行わない手法よりも，効果的にユーザの位置を曖昧化できることを確認した．しかし，今回の評価環境では，交差の分布がサービス利用の後半に集中したため，ユーザの位置を曖昧化できるまでの平均時間 (MTC) を十分に小さくできなかった．今後は，MTC をより小さくできるように，手法の改善を行う予定である．

7. 謝　辞

本研究の一部は，文部科学省科学研究費補助金・基盤研究 (A)(26240013)，挑戦的萌芽研究 (16K12429)，および日立財団研究助成「倉田奨励金」の研究助成によるものである．評価で利用した東京 23 区の地図データは一般財団法人日本デジタル道路地図協会より貸与されたものである．ここに記して謝意を示す．

参考文献

[1] De Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M. and Blondel, V. D.: Unique in the crowd: The privacy bounds of human mobility, Scientific Reports, Vol. 3 (2013).

[2] Kato, R., Iwata, M., Hara, T., Arase, Y., Xie, X. and Nishio, S.: User location anonymization method for wide distribution of dummies, Proc. SLGSPATIAL Int'l Conf. on Database and Expert Systems Applications, pp. 259–273 (2013).

[3] 加藤 諒，原 隆浩，Xie, X.，岩田麻佑，西尾章治郎：ユーザの行動プランの変更を考慮したダミーによるユーザ位置曖昧化手法，データ工学と情報マネジメントに関するフォーラム (2015).

[4] Lu, H., Jensen, C. S. and Yiu, M. L.: PAD: privacy-area aware, dummy-based location privacy in mobile services, Proc. ACM Int'l Workshop on Data Engineering for Wireless and Mobile Access, pp. 16–23 (2008).

[5] Suzuki, A., Iwata, M., Arase, Y., Hara, T., Xie, X. and Nishio, S.: A user location anonymization method for location based services in a real environment, Proc. SIGSPATIAL Int'l Conf. on Advances in Geographic Information Systems, pp. 398–401 (2010).

[6] Gkoulalas-Divanis, A. and Verykios, V. S.: A privacy-aware trajectory tracking query engine, ACM SIGKDD Explorations Newsletter, Vol. 10, No. 1, pp. 40–49 (2008).

[7] Palanisamy, B. and Liu, L.: MobiMix: Protecting location privacy with mix-zones over road networks, Proc. IEEE Int'l Conf. on Data Engineering, pp. 494–505 (2011).

[8] Shokri, R., Freudiger, J., Jadliwala, M. and Hubaux, J.-P.: A distortion-based metric for location privacy, Proc. ACM Workshop on Privacy in the Electronic Society, pp. 21–30 (2009).