ユーザの移動先候補の推定に基づく逐次的なダミー生成による位置曖昧化手法

林田 秀平[1] 天方 大地[1] 原 隆浩[1] Xing Xie[2]

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．これまでに，ダミーを用いた手法に関する研究が数多くなされている．しかし，既存研究は，ユーザは停止することなく移動し続ける，もしくはユーザの移動軌跡が事前に既知であるという非現実的な想定をしているため，実環境への適用は困難である．そこで本研究では，これらの想定を緩和し，ユーザの訪問場所の集合のみを入力とした際に位置曖昧化が可能なダミー生成手法 (**E**stimation-based **d**ummy sub-trajectory **ge**neration: Edge) を提案する．Edge は，ユーザの行動プランに関する詳細な入力を必要としないため，ユーザの負担を軽減できる．また，入力された訪問場所の集合から，ユーザの移動先の候補を推定し，逐次的にダミーのトラジェクトリを生成することで，ユーザの移動状況に柔軟に対応する．シミュレーションによる評価実験の結果，Edge は効果的にユーザの位置を曖昧化できることを確認した．

1. 序論

近年，スマートフォンやタブレット端末などの GPS を搭載したモバイル端末の普及により，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，GoogleMap*1では，現在地を送信すると，周辺に存在する人気のレストランや観光スポットの検索が行える．

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，勤務地，訪問場所，および行動パターン等が露見し，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．文献 [2] では，「第三者に自身の現在地や過去の位置情報履歴が露見することを防ぐ能力」を位置プライバシと定義しており，本稿で考える位置プライバシは，この定義に従う．

これまでに，位置プライバシの保護を目的とした研究が数多く行われている [3–10]．文献 [3] では，Cloaking Region を利用した手法が提案されている．この手法では，ユーザは自身の位置情報をサービスプロバイダに直接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，管理してい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 Region) を選択し，その領域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．文献 [4] では，Mix Zone を用いた手法が提案されている．この手法では，ユーザは第三者サーバに自身がサービス利用を行いたい領域を登録し，この領域以外の部分にサービス利用を禁止する領域 (Mix Zone) を生成する．その後，Mix Zone 内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，受信した位置情報群の中から，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．これらの手法で用いられている匿名化サーバおよび第三者サーバは，完全に信頼できるものと想定されているが，実環境でこの想定が保証されるとは限らない．

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 Microsoft Research Asia
*1 https://www.google.co.jp/maps

Copyright (c) 2017 by the Information Processing Society of Japan

この問題を解決するために，文献 [5–9] では，偽の位置情報 (ダミー) を用いることで，第三者サーバを使用せずにユーザの位置を曖昧化する手法を提案している．この手法では，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．また，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，サービスの質は低下しない．しかし，これらの文献では，海上など，歩行するユーザが存在し得ない位置にダミーを生成する可能性があったり [5]，ユーザは停止することなく移動し続ける [6]，もしくはユーザの移動軌跡が事前に既知である [7, 8] という非現実的な想定をしているため，実環境への適用が困難である．

そこで本研究では，ユーザの行動に関する入力を訪問場所の集合のみとした際に位置曖昧化が可能なダミーを生成する手法 (**E**stimation-based **d**ummy sub-trajectory **g**eneration: Edge) を提案する．ユーザの行動に関する情報が訪問場所の集合のみという条件下では，訪問場所の訪問順序が未知であるため，ユーザの位置を効果的に曖昧化するためのダミーを生成することはチャレンジングである．Edge は，ユーザの移動先の候補を巡回セールスマン問題を用いて推定し，その推定に基づいてエンティティ (ユーザおよびダミー) 同士のトラジェクトリを交差させたり，エンティティが広い範囲に分布するようにダミーを配置する．また，ユーザが訪問場所に到着する際に，逐次的にダミーのトラジェクトリを生成することにより，ユーザの移動状況に柔軟に対応する．シミュレーション実験の結果から，Edge は効果的にユーザの位置を曖昧化できることを確認した．

以下では，第 2 章でダミーを用いた位置曖昧化についての関連研究を紹介し，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

文献 [6] では，実環境における制約を考慮し，地図情報に基いて，ユーザが移動可能な範囲に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想定しており，ユーザが停止する場合に対応できない．文献 [7, 8] 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，ユーザがその行動プランに完全に従って行動すると想定している．つまり，ユーザの移動軌跡を事前に把握できるため，実際のユーザの移動軌跡を基に，定期的にエンティティが広い範囲に分布したり，エンティティ同士の移動軌跡が交差するようにダミーのトラジェクトリを設計する．しかし，実際は，訪問場所での停止時間や訪問場所の訪問順序などに変更が生じることが一般的であるため，ユーザが事前に入力した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る．文献 [9] では，実用性を向上させるため，ユーザの訪問場所の集合のみを入力とした際に位置曖昧化が可能なダミーを生成する手法を提案している．この手法では，入力された訪問場所の集合からユーザの行動プランを予測し，一度ダミーのトラジェクトリを生成する．その後，毎サービス利用時に，予測したユーザの行動プランと実際の行動プランの差異を判断し，その変化に応じて，ダミーのトラジェクトリを動的に修正する．しかし，一度生成したトラジェクトリをなるべく保全するように修正を行うため，予測したユーザの行動プランと実際の行動プランの差異が大きいと，ユーザの位置を十分に曖昧化できない．

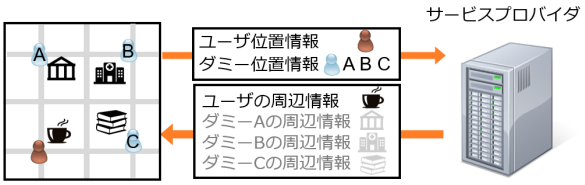


図 1: ダミーを用いた位置曖昧化手法

そこで Edge は，入力された訪問場所の集合から，ユーザの移動先の候補を推定し，逐次的にダミーのトラジェクトリを生成することにより，ユーザの移動状況に柔軟に対応する．

3. 事前準備

3.1 想定環境

本研究では，ユーザがサービスプロバイダに頻繁に位置情報を送信する位置情報サービス (例えば GoogleMap) を利用することを想定する．また，ユーザは，いくつかの訪問場所を訪問しながら徒歩で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．ユーザは入力した全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．なお，ユーザが所有するモバイル端末は，地図

Copyright (c) 2017 by the Information Processing Society of Japan

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 ユーザの位置プライバシ保護に関する要求

まず，実環境における制約について考える．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．つまり，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．

また，ユーザの位置プライバシ保護に関する要求として，追跡可能性と匿名領域の 2 点を考慮する．

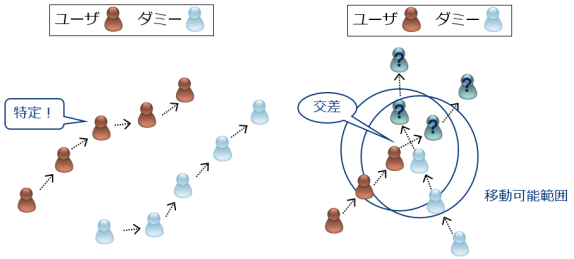
追跡可能性. ユーザが短時間に連続して位置情報サービスを利用する場合，ユーザのサービス利用の追跡可能性を考慮しなければならない．例えば，図 2(a) ように，あるサービス利用時刻におけるエンティティの位置から次のサービス利用の時刻までに対応する位置情報が一つしかない場合，前後のサービス利用の対応関係が一意に定まることから，特定のエンティティの移動軌跡が追跡できてしまう．この場合，ユーザが一度特定されてしまうと，その前後にサービスを要求した位置まで特定され，より多くの情報流出を招く．このような追跡を防ぐには，図 2(b) のように，ユーザとダミーの移動軌跡を交差させることが有効であり，これにより追跡可能性を低下できる．具体的には，以下で定義される指標の低下を図る．

定義 1 (Mean Time to Confusion) [10]. Mean Time to Confusion（MTC）は，ユーザが特定された時点から各エンティティ e_i のユーザ確率 $p(i)$ のエントロピー $H = -\sum p(i)\log_2 p(i)$ が 1 を超えるまでにかかる平均時間と定義される．すなわち，サービス利用の総数を l，時刻 t_j 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 $TC(t_j)$，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_j の数を $n_{achieve}$ とすると，MTC は，式 (1) で計算される．

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

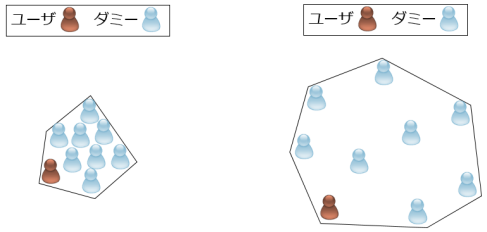
ただし，$TC(t_j)$ は，時刻 t_l までに $H > 1$ が満たせない場合は 0 と見なし，MTC の計算には考慮しない．

なお，ユーザ確率とは，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示す．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 x のエンティティとユーザ確率 y 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 $\frac{x+y}{2}$ となる．MTC 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．



(a) 連続的なサービス利用による移動軌跡の特定　(b) エンティティ同士の交差による移動軌跡の曖昧化

図 2: 追跡可能性



(a) 匿名領域が小さい場合　(b) 匿名領域が大きい場合

図 3: 匿名領域

匿名領域. 文献 [5] の定義に基づき，全てのエンティティで形成される凸包領域を匿名領域と定義する．ユーザの位置プライバシを保護するためには，ユーザの位置がどの程度の大きさの領域に曖昧化されているのかも重要である．例えば，図 3(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，おおよそのユーザの位置が推測できてしまう．そのため，図 3(b) のように，ダミーを広範囲に配置させる必要がある．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の配置を試みる．

4. 提案手法

4.1 概要

短時間で連続的な位置情報サービスの利用を想定すると，ユーザの移動軌跡が追跡されないことが重要である．そこで Edge は，まず MTC の低下を行うためのダミーのトラジェクトリを生成し，その後，匿名領域を大きくするためのダミーのトラジェクトリを生成する．ここで，ユーザの行動プランに関する入力が訪問場所の集合のみという条件

Copyright (c) 2017 by the Information Processing Society of Japan

下では，ユーザの移動先が既知でない．そのため Edge は巡回セールスマン問題を用いてユーザの移動先の候補を推定し，ユーザとダミーの移動軌跡を交差させる．具体的には，ユーザが訪問場所に到着した際に，次回以降の訪問場所の訪問順序のうち，移動距離が短い上位 k 個をユーザの移動先と推定し，k 個の次の訪問場所でユーザと交差を行うダミーのトラジェクトリを生成する．また，MTC を低下させるためには，交差が生じる間隔が短いことも重要である．そこで，ユーザの次の訪問場所の候補 (k 個) のそれぞれの次の訪問場所で交差が発生するようにダミーのトラジェクトリを生成する．すなわち，ユーザと交差を行うダミーのトラジェクトリを $2k$ 個生成する．これにより，ユーザの複数の訪問順序への対応および交差の間隔の短縮を図る．なお，k はユーザが入力するものとし，k が増えるほど，訪問場所での交差が発生する確率が高くなる．また，交差の対象となるエンティティがユーザのみである場合，ユーザの特定のヒントになってしまう．これを防ぎつつ匿名領域を大きくするため，β 個のダミーを用いて，他のダミーと交差を行うダミーのトラジェクトリおよび匿名領域の拡大を行うダミーのトラジェクトリを生成する．（つまり，生成するダミーのトラジェクトリの数 d は，$2k+\beta$ 個となる．）なお，β もユーザが入力するものとし，β が増えるほど，ダミー同士の移動軌跡の交差が増えたり，エンティティが広い範囲に分布する．以上の方法を用いて，ユーザが訪問場所に到着する際に，ダミーのトラジェクトリを逐次的に生成することで，ユーザの移動状況に柔軟に対応する．

Edge の入出力. 入力は，$\{k, \beta, r, VP\}$で表される．r は，要求匿名領域を面積 $r^2[\mathrm{m}^2]$ の正方形とした際の一辺の長さであり，$VP = \{vp_1, vp_2, \cdots, vp_q\}$は，ユーザの訪問場所 vp_i の集合である．出力は，$\{T_1, T_2, \cdots, T_d\}$ で表される．T_j は，ダミー e_j のトラジェクトリであり，i 回目の位置情報送信時の位置 p_i およびその時刻 t_i の組 $\langle p_i, t_i\rangle$ 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 を用いて表される．すなわち，

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

である．

フレームワーク. アルゴリズム 1 に，Edge のフレームワークを示す．ユーザは，k，β，r，および VP を入力し，位置情報サービスの利用を開始する．ユーザが訪問場所 vp_i に到着する (初回のサービス利用を含む) と (2 行)，VP に基いて，ユーザの移動先および行動プランを k 個推定し (3 行)，ユーザとの交差を行うダミーのトラジェクトリを決定する (4 行)．次に，各ダミーのエンティティとの交差回数を基

Algorithm 1: Edge

Input: VP, k , β, r

1 **for** $time = 0$ **to** l **do**
2 　**if** *User has arrived at* vp_i **then**
3 　　Estimate-user-movement-plan(k, *time*)
4 　　Intersect-with-user(k, *time*)
5 　　$\langle D_1, D_2\rangle \leftarrow$ Assign-dummies(β, *time*) //D_1 (D_2) is a list of a dummy's trajectory for intersection with other dummies (for enlarging anonymous area)
6 　　Intersect-with-dummy(D_1, *time*)
7 　　Enlarge-anonymous-area(D_2, *time*)
8 　**else**
9 　　Check difference between actual and estimated user movement plan
10 　　Adjust arrival time of dummies for intersection with user

に，β 個のダミーを，他のダミーと交差させる，または匿名領域を大きくするダミーとする (5-7 行)．その後，ユーザとダミーの交差を発生させるために，ダミーの訪問場所への到着時刻の調整を行う．まず，推定したユーザの行動プランと実際の行動プランが異なるか否かを判断し，もし異なる場合は，各移動先の候補への到着時刻を再び推定する (9 行)．その後，ユーザと交差を行うダミーのトラジェクトリを，推定した到着時刻に合わせて調整する (10 行)．これらの操作をユーザのサービス利用が終了するまで繰り返し行う．以降で各フローの詳細を述べる．

4.2 Edge の各フローの詳細

4.2.1 ユーザの行動プランの推定方法

ユーザが訪問場所 vp_i に到着した際に，入力された訪問場所の集合 VP を用いて，ユーザの移動先および行動プランを k 個推定する. 移動先 (訪問順序) の候補および行動プランの各要素 (移動経路，停止時間，および移動速度) を以下の通りに推定する．

- 訪問順序. ユーザは，移動距離がなるべく小さくなる順序で訪問場所を移動すると仮定し，始点を停止中の訪問場所，終点を未定とした巡回セールスマン問題を解くことで訪問順序を推定する．すなわち，訪問順序は移動距離が小さい上位 k 件を選択する．ただし，ある順位の訪問順序において，次に向かう訪問場所が上位のものと同一の場合，その順位のものは候補から外す．
- 移動経路. 各訪問場所 vp_i 間を最短経路で移動するものとする．
- 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，$[v_{min}, v_{max}]$ における一様分布に基づき任意の値を一つ定める．

Copyright (c) 2017 by the Information Processing Society of Japan

Algorithm 2: Intersect-with-User

Input: k, *time*

1 $D \leftarrow \{T_1, T_2, \cdots, T_d\}$
2 Sort D in ascending order of times of intersection with user
3 **for** $n = 1$ **to** 2 **do**
4 　**for** $i = 1$ **to** k **do**
5 　　$\langle p_s, t_s \rangle \leftarrow$ a tuple of estimated vp and arrival time of user
6 　　**for** $j = 1$ **to** d **do**
7 　　$\langle p_f, t_f \rangle \leftarrow$ a tuple of the last generated position and arrival time of $D[j]$
8 　　**if** reachable($\langle p_f, t_f \rangle, \langle p_s, t_s \rangle$) **then**
9 　　Decide-Trajectory($D[j]$, $\langle p_s, t_s \rangle$) //decide elements of $D[j]$
10 　　**break**

Algorithm 3: Intersect-with-dummy

Input: D_1(D_1 is a list of dummy's trajectory for interesction with dummy), *time*

1 **for** $i = 0$ **to** D_1.size()-1 **do**
2 　$d_1 \leftarrow D_1[i]$
3 　$d_2 \leftarrow D_1[i+1]$
4 　$\langle p_{d_1}, t_{d_1} \rangle \leftarrow$ a tuple of the last generated position and arrival time of d_1
5 　$\langle p_{d_2}, t_{d_2} \rangle \leftarrow$ a tuple of the last generated position and arrival time of d_2
6 　$p_g \leftarrow$ a random point on the segment between p_{d_1} and p_{d_2}
7 　$\zeta \leftarrow$ a random value $\in [0, 1]$
8 　$\overrightarrow{w_h} \leftarrow \frac{\zeta \cdot range}{|\overrightarrow{w_{d_1}} + \overrightarrow{w_{d_2}}|} (\overrightarrow{w_{d_1}} + \overrightarrow{w_{d_2}})$ //$\overrightarrow{w_{d_1}}$ and $\overrightarrow{w_{d_2}}$ are vectors composed of moving direction of d_1 and d_2 starting at p_g, and $range$ is movable range of d_1 and d_2
9 　$p_s \leftarrow$ the nearest PoI of $\overrightarrow{w_h}$
10 　$t_s \leftarrow$ estimated arrival time to the next vp of user
11 　**if** reachable($\langle p_{d_1}, t_{d_1} \rangle, \langle p_s, t_s \rangle$) $\wedge$ reachable($\langle p_{d_2}, t_{d_2} \rangle, \langle p_s, t_s \rangle$) **then**
12 　　Decide-Trajectory(d_1, $\langle p_s, t_s \rangle$) //decide elements of d_1
13 　　Decide-Trajectory(d_2, $\langle p_s, t_s \rangle$) //decide elements of d_2

4.2.2 ダミーのトラジェクトリの生成方法

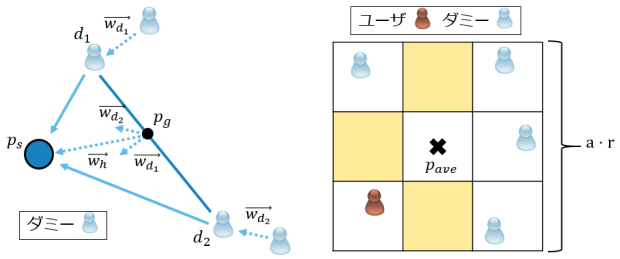
d 個のダミーのトラジェクトリの生成方法について述べる．d 個のうち，$2k$ 個のダミーをユーザとの交差のために用い，β 個のダミーを他のダミーとの交差または匿名領域の拡大のためのダミーに用いる．以下で詳細を述べる．

ユーザと交差を行うダミーのトラジェクトリの生成方法. アルゴリズム 2 に，ユーザと交差を行うダミーのトラジェクトリの生成方法を示す．追跡可能性を低下させるために，4.2.1 項で説明した方法で推定したユーザの次の訪問場所の候補，および，それぞれの次の訪問場所で交差が発生するようにトラジェクトリを決定する．まず，各ダミーを，その時点までに生じたユーザとの交差回数の少ない順にソートする (1-2 行)．その後，ソート順に各ダミーを走査し，もし各ダミーの最後に生成された場所および時刻の組 $\langle p_f, t_f \rangle$ から，ユーザとの交差場所および時刻の組 $\langle p_s, t_s \rangle$ に到達可能である場合，そのダミーのトラジェクトリをユーザと交差するように生成する (5-10 行)．この操作を $2k$ 回繰り返し行うことで，ユーザとの交差の発生を図る．なお，到達可能性の定義は，以下の通りである．

定義 2 (到達可能性). $\langle p_i, t_i \rangle$ から $\langle p_j, t_j \rangle$ へ到達可能であるとは，$dist(p_i, p_j)$ を p_i と p_j との地図上での距離とすると，式 (2) を満たす状態である．

$$dist(p_i, p_j) \leq v_{max} \cdot (t_j - t_i),\ t_i < t_j \qquad (2)$$

β 個のダミーの割り当て方法. β 個のダミーを，他のダミーとの交差または匿名領域の拡大を行うように割り当てる．この際，ダミーの識別を困難にするため，交差するダミーおよび匿名領域の拡大を行うダミーの数に偏りが生じないようにする．具体的には，まず，β 個のダミーをエンティティとの交差回数の少ない順にソートする．その後，各ダミーのエンティティとの交差回数の順位が $\frac{d}{2}$ 以下のダミーを，ダミー同士の交差に割り当て，$\frac{d}{2}$ より大きいダミーを匿名領域の拡大に割り当てる．



(a) ダミー同士の交差場所　(b) 匿名領域拡大のためのダミーの移動先の決定方法

図 4: ダミーの移動先の決定方法

ダミー同士の交差のためのダミーのトラジェクトリの生成方法. 交差の対象となるエンティティがユーザのみである場合，ユーザの特定のヒントになってしまう．これを防ぐため，ダミー同士の交差を発生させる．アルゴリズム 3 に，ダミー同士の交差を行うダミーのトラジェクトリの生成方法を示す．また，図 4(a) に交差場所を示す．まず，ダミー同士の交差に割り当てられたダミー群 D_1 から 2 つのダミーを選択し，一方のダミー d_1 の最後に生成された位置 p_{d_1} から，もう一方のダミー d_2 の最後に生成された位置 p_{d_2} で形成される線分の間からランダムに選択した点を p_g とする (2-6 行)．次に，p_g を始点とした d_1 および d_2 の進行方向ベクトルを $\overrightarrow{w_{d_1}}$ および $\overrightarrow{w_{d_2}}$，d_1 および d_2 が移動可能な距離を *range*，0 から 1 の間のランダムな値を ζ と

Copyright (c) 2017 by the Information Processing Society of Japan

Algorithm 4: Enlarging-anonymous-area

Input: D_2(D_2 is a list of dummy's trajectory for enlarging anonymous area), *time*

1 $p_{ave} \leftarrow$ average position of generated entities
2 Generate a grid G of 3×3 *cells* $\{c_1, c_2, \cdots, c_9\}$ centered at p_{ave}
3 Count the number of entities in $c \in G$ for each c
4 Sort *cells* in ascending order of the number of entities
5 **for** $\forall T \in D_2$ **do**
6 $\quad \langle p_f, t_f \rangle \leftarrow$ a tuple of the last generated position and arrival time of T
7 $\quad$ **for** $\forall c \in G$ **do**
8 $\quad\quad p_s \leftarrow$ a PoI in c
9 $\quad\quad t_s \leftarrow$ estimated arrival time to the next vp of user
10 $\quad\quad$ **if** reachable($\langle p_f, t_f \rangle, \langle p_s, t_s \rangle$) **then**
11 $\quad\quad\quad$ Decide-Trajectory($T, \langle p_s, t_s \rangle$) //decide elements of T
12 $\quad\quad\quad$ **break**

し，大きさが $\zeta \cdot$ *range*，向きが $\overrightarrow{w_{d_1}} + \overrightarrow{w_{d_2}}$ のベクトル $\overrightarrow{w_h}$ を求める (7-8 行)．その後，$\overrightarrow{w_h}$ の最近傍の場所を交差場所 p_s とし (9 行)，p_{d_1} および時刻 t_{d_1} の組 $\langle p_{d_1}, t_{d_1} \rangle$，および p_{d_2} および時刻 t_{d_2} の組 $\langle p_{d_2}, t_{d_2} \rangle$ から，p_s および推定したユーザの次の訪問場所への到着時刻 t_s の組 $\langle p_s, t_s \rangle$ に到達可能であれば，p_s を d_1 および d_2 の交差場所として決定する (11-13 行)．この操作を，割り当てられた全てのダミーのペアで繰り返すことで，ダミー同士の交差を発生させる．

匿名領域拡大のためのダミーのトラジェクトリの生成方法. アルゴリズム 4 に，匿名領域の拡大のためのダミーのトラジェクトリの生成方法を示す．エンティティが密集し，ユーザの位置のおおよその推測が可能になることを防ぐために，ダミーを広い範囲に分布させる．まず，既にトラジェクトリを生成済みのエンティティの平均の位置 p_{ave} を計算し，3×3 のセルで構成され，中心を p_{ave}，一片の長さを $a \cdot r$ とするグリッドを生成する (1-2 行)．次に，各セルに含まれるトラジェクトリを生成済みのエンティティの数を計上する (3 行)．その後，エンティティの数が少ない順にセルをソートする (4 行)．ソート順にセルを走査し，あるダミーの最後に生成された場所および時刻の組 $\langle p_f, t_f \rangle$ から，セルの中に含まれる場所 p_s および推定したユーザの次の訪問場所への到着時刻 t_s の組 $\langle p_s, t_s \rangle$ への到達可能性を調べる．もし到達可能であれば，p_s を匿名領域拡大のための移動先とする (5-12 行)．図 4(b) では，色のついたセルがエンティティが最も少ない領域となるため，このセルの中に存在する場所から移動先を選択する．

4.2.3 ダミーの訪問場所への到着時刻の調整方法

ユーザとダミーを交差させるには，訪問場所への到着時刻を合わせる必要がある．そこで，推定したユーザの行動プランと実際の行動プランとの差異を判定し，差異がある場合は，ユーザと交差を行うダミーの訪問場所への到着時刻の調整を行う．以下で詳細を述べる．

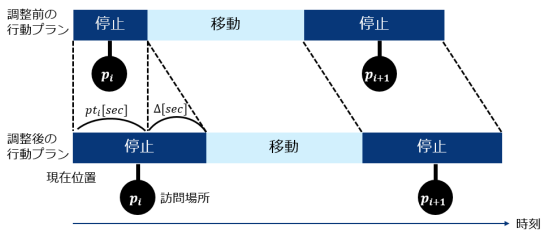


図 5: ダミーの訪問場所への到着時刻の調整方法

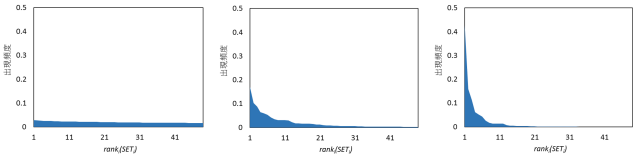
推定したユーザの行動プランと実際の行動プランの差異の判定方法. 実際のユーザの移動状況をもとに，推定したユーザの行動プランとの差異を判定し，各訪問場所への到着時刻を推定する．以下に，移動経路，停止時間，および移動速度それぞれについて，差異の判断方法および各訪問場所の到着時刻の推定方法を示す．

- 移動経路. 推定した行動プランと異なる経路上にユーザが位置した場合，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，到着時刻を推定する．
- 停止時間. 推定した行動プランでは移動を開始している，または停止中の予定であるにも関わらず，実際は訪問場所に停止中，または移動を開始している場合，停止時間が増加または減少したとみなす．この場合，停止時間の変化分だけ各訪問場所への到着時間が遅れる，または早まると推定する．
- 移動速度. 推定した行動プラン通りの経路上に位置しているが，その位置が異なる場合，移動速度に差異があると判断する．その場合，直前まで停止していた訪問場所を推定した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および現在の位置から移動速度を求める．また，各訪問場所までその移動速度で移動するとみなし，到着時刻を推定する．

ユーザと交差を行うダミーの訪問場所への到着時刻の調整方法. 推定したユーザの行動プランと実際の行動プランに差異があ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到着時刻の差分 Δ[sec] を求め，ダミーの訪問場所への到着時刻も Δ だけ変化するようにトラジェクトリを調整する．具体的には，生成済みのダミーのトラジェクトリにおいて，変化量の総和が Δ となるように，停止時間，移動速度，および移動経路を調整する．

図 5 のように，あるダミーが p_i に停止中の場合，p_i での停止時間を変化させることで，到着時刻を調整する．この変更による到着時刻の変化量が Δ に満たない場合，p_i から次に訪問する場所 p_{i+1} までの距離を基に，到着時刻が Δ だけ変化するように移動速度の調整を行う．それでも到

Copyright (c) 2017 by the Information Processing Society of Japan



(a) Low (b) Mid (c) High

図 6: ユーザの行動モデル

着時刻の変化量が Δ に満たない場合，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，到着時刻を調整する．もし，移動中の場合は，移動速度の調整から行い，以後同様の操作を行う．

5. 評価実験

5.1 実験環境

東京 23 区の地図上 (地図領域: 67.2[km^2], 訪問場所の数: 26102, 交差点の数: 22724) 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．訪問場所には，FourSquare API[*2]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．ユーザデータには，図 6 のように，ある訪問場所の集合 *VP* の半順序集合 $SET_i(1 \leq i \leq (n(VP)-1)!)$ の出現頻度が異なる行動モデルを人工的に 3 つ (Low, Mid, および High) 生成し，各モデルから 50 個ずつ作成したトラジェクトリを用いた．各行動モデルは，SET_i を総移動距離が短い順にソートした際の順位 $rank_i(SET_i)$ を基に，SET_i の出現頻度を変化させており，Low, Mid, High の順に総移動距離が短い順序で移動する頻度が増加する．ユーザの訪問順序の推定には，総移動距離が短い上位 k 件を用いているため，直感的には，Low では推定の精度が低く，High では推定の精度が高いことを示す．*VP* の要素は，地図領域中から 7 つの訪問場所をランダムに選択した．なお，サービスの利用間隔を 180[sec]，v_{min} を 1.05[m/sec]，v_{max} を 1.55[m/sec]，pt_{min} を 600[sec]，pt_{max} を 1800[sec] とし，行動プランの各要素は 4.2.1 項と同様の方法で決定した．また，各パラメータは，表 1 のように定めた．$a = 1.5$ とし，各パラメータのデフォルト値は，$r = 1500$，$k = 2$，$\beta = 5$ とした．

表 1: 各パラメータの値

パラメータ	値
r[m]	1000, 1250, 1500, 1750, 2000
k	2, 3, 4
β	3, 4, 5, 6, 7

5.2 評価指標

本実験では以下の 2 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 Confusion) [10]. 式 (1) で定義される指標．ただし，交差の角度が 30° より大きい場合は，交差が生じたと見なし，ユーザ確率の配分を行うが，交差の角度が 30° 以下となる場合，進行方向に基づく追跡が可能であると考えられるため，ユーザ確率の配分を行わない．

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．

Dum-P-Cycle [8]. ユーザの行動プランを既知とし，ダミーを生成する手法

E-Dum-P-Cycle（Estimation-based-Dum-P-Cycle）[9]. ユーザの行動プランに関する既知の情報は *VP* のみとし，訪問順序を巡回セールスマン問題の最適解と予測し，予測したユーザに対して Dum-P-Cycle を適用する．その後，毎サービス利用時に，実際のユーザの移動状況と予測したユーザの差異を判断し，その変化に応じて，ダミーのトラジェクトリを動的に修正する手法．

Edge. 本稿における提案手法．

5.4 評価結果

要求匿名領域，k，β，およびユーザの行動モデルを変えて行った実験の結果を示す．なお，要求匿名領域，k，および β の実験には，図 6(b) の Mid の行動モデルを用いた．

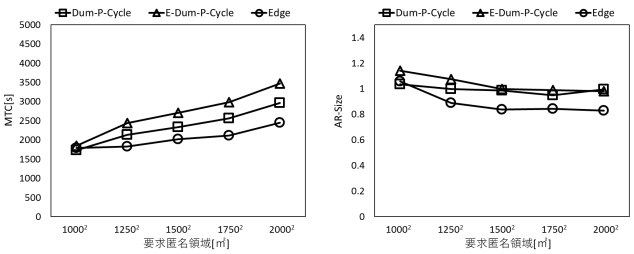
要求匿名領域の影響． 図 7(a) は，要求匿名領域に対する MTC の変化を示す．Edge は，ユーザの $2k$ 個の訪問場所で交差を発生させているため，ユーザとの交差が発生する確率が高く，また，交差の間隔が短いため，最も MTC を短くできている．図 7(b) は要求匿名領域に対する AR-Size の変化を示す．Edge は 0.8 程度の AR-Size は確保できているものの，1 には満たない．Edge は，追跡可能性の低下のために，エンティティ同士の移動軌跡の交差を最優先に行っており，ダミーを十分に広げることができなかったためである．

k の影響． 図 8(a) は，k に対する MTC の変化を示す．MTC は，k が増加するに従って短くなる傾向を示す．これは，ダミーの数が増加するにつれて，エンティティ同士の移動軌跡が交差しやすくなるためである．図 8(b) は，k に対する AR-Size の変化を示す．AR-Size は，k が増加するに従って増加傾向である．これは，ダミーの数が増加するにつれて，エンティティが広範囲に分布しやすくなるためである．

β の影響． 図 9(a) は，β に対する MTC の変化を示す．Edge において，MTC は β の影響を受けづらい．これは，β 個のダミーの一部をダミー同士の交差に用いているが，MTC は，ユーザとダミーの交差のみ関係するため，ユーザと

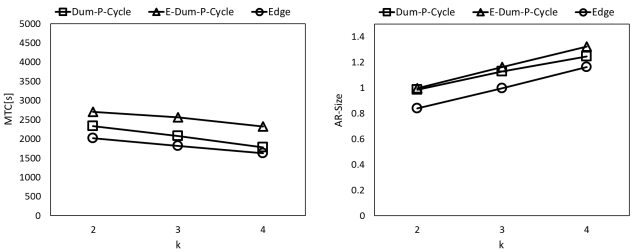
*2 https://developer.foursquare.com

Copyright (c) 2017 by the Information Processing Society of Japan



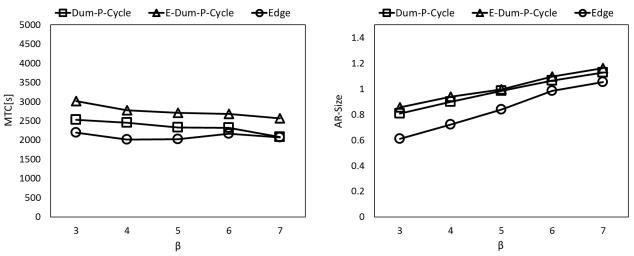
(a) MTC (b) AR-Size

図 7: 要求匿名領域の影響



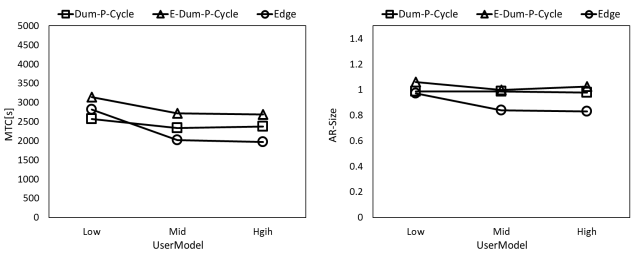
(a) MTC (b) AR-Size

図 8: k の影響



(a) MTC (b) AR-Size

図 9: β の影響



(a) MTC (b) AR-Size

図 10: ユーザの行動モデルの影響

の交差に関係しないダミー同士の交差の数を増やしても MTC には影響しないためである．図 9(b) は，β に対する AR-Size の変化を示す．β が増加するに従って，AR-Size は増加している．Edge では，β が増加すると，匿名領域の拡大を行うダミーが増えるためである．

ユーザの行動モデルの影響． 図 10(a) は，ユーザの行動モデルに対する MTC の変化を示す．Edge では，推定精度が高くなるに従って，MTC を短くできている．これは，推定精度が高いほど，ユーザとの交差が発生する確率が高くなるためである．図 10(b) は，ユーザの行動モデルに対する AR-Size の変化を示す．Edge は，訪問順序の推定精度が高くになるに従って，AR-Size は減少している．これは，推定精度が低いほど，ダミーが凸包領域を形成するエンティティになりやすいためである．

6. おわりに

本研究では，実用性を向上させるため，ユーザの行動に関する入力を訪問場所の集合のみとした際にも対応可能なダミーを生成する手法（Edge）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を行った結果，Edge は，効果的にユーザの位置を曖昧化できることを確認した．今後は，匿名領域をより確保できるように，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)(JP26240013) および挑戦的萌芽研究 (JP16K12429) 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] Beresford, A. R. and Stajano, F.: Location privacy in pervasive computing, IEEE Pervasive computing, Vol. 2, No. 1, pp. 46–55 (2003).

[3] Mokbel, M. F., Chow, C.-Y. and Aref, W. G.: The new casper: Query processing for location services without compromising privacy, Proc. VLDB, pp. 763–774 (2006).

[4] Palanisamy, B. and Liu, L.: MobiMix: Protecting location privacy with mix-zones over road networks, Proc. ICDE, pp. 494–505 (2011).

[5] 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).

[6] 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. GIS, pp. 398–401 (2010).

[7] Kato, R., Iwata, M., Hara, T., Suzuki, A., Arase, Y., Xie, X. and Nishio, S.: A Dummy-based Anonymization Method based on User Trajectory with Pauses, Proc. GIS, pp. 249–258 (2012).

[8] Kato, R., Iwata, M., Hara, T., Arase, Y., Xie, X. and Nishio, S.: User location anonymization method for wide distribution of dummies, Proc. DEXA, pp. 259–273 (2013).

[9] 林田秀平，水野聖也，天方大地，原　隆浩，Xie, X.：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，情報処理学会マルチメディア，分散，協調とモバイルシンポジウム論文集，pp. 328–335 (2016).

[10] 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).

Copyright (c) 2017 by the Information Processing Society of Japan