ユーザの訪問場所の傾向を考慮したダミーによるユーザ位置曖昧化手法

加藤 諒[1]　松野 有弥[2]　原 隆浩[1]　荒瀬 由紀[3]　Xing Xie[3]　西尾 章治郎[1]

近年注目を集めている位置情報サービスでは，サービス利用時にユーザの位置情報をサービスプロバイダへ送信する必要があり，位置情報が第三者に流出することで，ユーザの個人情報が漏洩する可能性がある．このようなユーザの位置情報に関するプライバシを保護するために，筆者らの研究グループでは，先行研究において，実環境における制約条件を考慮し，いくつかの場所を訪れながら移動するダミーの位置情報を生成することで，ユーザの位置を曖昧化する手法を提案した．しかし，先行研究では飲食店や映画館といった訪問場所の意味的情報が考慮されていないため，ダミーの訪問場所の傾向から容易にダミーであると判断されてしまう可能性がある．そこで本稿では，ユーザの訪問場所の履歴から抽出した訪問場所の傾向に従いダミーを移動させることで，よりユーザらしい動きを実現するダミーを生成する手法を提案する．訪問場所が密集した環境を想定した地図上でユーザの動きをシミュレートし，先行研究の手法と比較した結果，提案手法の方がユーザの位置プライバシ保護により有効であることを確認した．

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数多く展開されている．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，ユーザの訪問箇所や行動パターンなどの重要なプライバシが第三者に把握される可能性が指摘されている [2].

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研究の一つとして，ダミーの位置情報を用いたユーザの位置曖昧化手法がある [6][7][11]．この手法では，ユーザが位置情報サービスを利用する際，同時に複数のダミーの位置情報も送信する．これにより，送信された位置情報の中から，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．筆者らの研究グループでは，これまでに先行研究において，いくつかの場所を訪れながら移動するというユーザの行動を想定し，既知であるユーザの行動に基づいて，複数の場所を訪れながら移動するダミー生成手法を提案した [5]．しかし，この先行研究では飲食店や映画館といった訪問場所の意味的情報が考慮されていないため，生成されたダミーの訪問場所を調べると，利用目的が同じ場所に連続して立ち寄るような不自然な動きを示すものがダミーであると識別可能な可能性が高い．また，ユーザの訪問場所の傾向が第三者に特定されている場合，ユーザと異なる訪問場所の傾向を示すものがダミーであると容易に判断されてしまう．したがって，ユーザの位置プライバシを十分に保護するためには，訪問場所の意味的情報や，ユーザの訪問場所の傾向が第三者に特定されている場合でも，ユーザの位置情報は把握されないようにすることが求められる．

そこで，本研究では，ユーザがある場所を訪問した際に位置情報サービスを利用し，次の訪問場所に向けて移動するという移動モデルを想定して，ユーザの訪問場所の履歴から抽出した訪問場所の傾向に従い，よりユーザらしく移動するダミーを生成する手法を提案する．提案手法では，十分な量の行動履歴を保持したユーザ自身の端末上において，ユーザの訪問場所や停止時間も含めた行動予測が完全に可能な状況を想定し，予測されたユーザの行動，および，実環境における制約条件を考慮してダミーの行動を決定する．具体的には，提案手法はまず，ユーザの訪問履歴に基づいて，ユーザの訪問場所の傾向を抽出する．そして，得られた訪問場所の傾向に基づいて，ユーザとダミーが同期してサービス利用しながら移動するように，ダミーが訪問

[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] 東京大学 大学院情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo
[3] マイクロソフトリサーチアジア
Microsoft Research Asia

すべき場所およびその時間を決定する．この際，ユーザと他のダミーを包含する最小凸多角形の面積を考慮しながら，ユーザおよびダミーを広範囲に分散させ，ユーザの位置が広範囲に曖昧になるようにする．また，ユーザの訪問場所の傾向を考慮し，ユーザが訪問する可能性が高い場所にダミーを配置することで，ユーザとダミーの識別をより困難にしつつ，互いの交差を促す．これにより，ユーザの位置が一時的に特定された場合でも，その曖昧性を短時間で回復できるようにする．

以下では，2 章で関連研究を説明し，3 章でユーザの訪問場所の傾向を考慮したダミーによるユーザ位置曖昧化手法について述べる．4 章で評価実験の結果を示し，最後に 5 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

本章では，ユーザの位置プライバシの保護を目的とした代表的な 3 つのアプローチについて述べる．

文献 [4][8] 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るのではなく，信頼できる第三者サーバが自身の管理するユーザの位置情報から，あらかじめ決められた k 人以上のユーザを含むような領域を選択し，その領域に対するクエリをサービスプロバイダに送信する手法が提案されている．これにより，ユーザの位置は $\frac{1}{k}$ 以上の確率で特定不可能になる．ただし，この手法は第三者サーバを完全に信頼できるという想定のもとに成り立っているため，実環境で用いるのは困難である．

文献 [1][3] では，サービス要求時に送信する位置情報として，ユーザ自身の位置ではなく，ユーザ付近の交差点や建物などのあらかじめ決められた地点を送信する手法が提案されている．これにより，プロバイダはユーザの正確な位置を知ることができなくなり，ユーザの位置が曖昧化される．しかし，近隣に適当な地点が存在しない場合，ユーザ位置との乖離が大きくなるため，サービスの質が低下してしまう．

文献 [6][7][11] では，自身の位置情報と一緒に架空の情報であるダミーの位置情報をクエリに付加して，図 1 のようにサービスプロバイダにサービスを要求する手法が提案されている．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．返信された情報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することで，自身の位置情報に関連する情報のみを取得可能となる．サービスプロバイダは受信した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が低くなる．このようなダミーを用いた手法では，ダミーの位置情報はユーザの端末上で生成されるため，第三者サーバを必要とせず，サービスの質が低下することもない．そのため，本研究ではダミー手法を採用する．

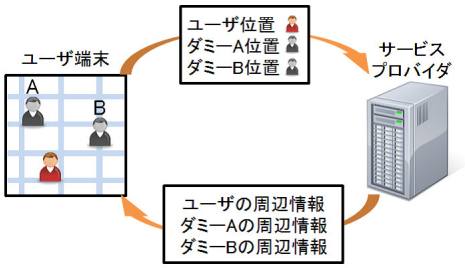

図 1 ダミーを用いた位置情報サービスの利用例

筆者らの研究グループでは，いくつかの場所を訪れながら移動するというユーザの移動モデルを想定し，ユーザと同様に行動するダミーを生成し，ユーザの位置を曖昧化する手法を提案した [5]．しかし，この先行研究の手法では，ダミーの訪問場所を決める際に，単純に地理的に到達可能であれば訪問可能であるとしており，ダミーが連続して複数回にわたり飲食店を訪れる可能性があるなど，訪問場所の意味的な情報に対する考慮が不足している．そのため，ダミーがユーザと異なる傾向で訪問場所を訪問し続けると，ダミーとユーザは容易に区別されてしまう．一方，提案手法では，文献 [12] の手法を適用し，ユーザの訪問履歴を基に，ユーザの訪問場所の傾向を訪問パターンとして抽出する．抽出された訪問パターンに従ってダミーの訪問場所を決定することで，よりユーザらしく行動をするダミーを生成できる．

3. ユーザの訪問場所の傾向を考慮したダミーの訪問経路生成手法

本章では，まず，想定環境を説明した後，実環境でダミーを生成する際に考慮すべき制約とユーザの位置プライバシに関する要求について述べる．次に，ユーザの訪問場所の傾向の抽出手法を説明し，最後に，それらの条件，要求とユーザの訪問場所の傾向を考慮したダミー生成手法の詳細について述べる．

3.1 想定環境

本研究では，ユーザがある場所を訪問した際に位置情報サービスを利用し，次の訪問場所に向けて移動するという環境を想定する．位置情報サービス利用のために，ユーザは自身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，現在位置に関連した情報を取得できる．ユーザが位置情報を送る際には，位置プライバシ保護のために，ユーザのモバイル端末上でダミーの位置情報を複数生成し，それらを自身の位置情報と共にサービスプロバイダに送信する．

具体的には，ユーザはある目的地をもって移動を開始し，移動中にいくつかの訪問場所で停止し，位置情報サービスを利用するという行動モデルを想定する．そして，それぞれの訪問場所で，ユーザやダミーが最小 T_m 秒から最大 T_M

秒までの範囲で停止する．訪問場所間は，最短路を通って移動するものとする．さらに，本研究では，ユーザの訪問場所，訪問場所に到着する時間といったユーザの行動がすべて事前に予測できるものと想定する．このような想定は実環境では必ずしも妥当ではないが，ユーザの過去の行動履歴から予測したりするなど，ある程度の精度で予測できる場合も多い．このような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．

ユーザが所有するモバイル端末上のシステムは，地図情報とユーザの訪問履歴を保持しており，ユーザやダミーが通っても不自然ではない道路，訪問可能な場所，訪問可能な場所の意味的情報，ユーザの訪問傾向をすべて把握しているものとする．

3.2 実環境でダミーを生成する際に考慮すべき条件

短時間にサービス要求が繰り返される場合，前後のサービス利用におけるダミー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を要求してから，三分後に新たにサービス要求した場合を考える．この際，新しいサービス要求において，直前のサービス利用時のどのダミー位置からも三分間で到達不可能な位置にダミーが存在する場合，その位置情報はユーザではないと容易に推測可能である．

そこで本研究では，実際の地図情報を用いてダミーの移動距離を計算することで，直前のダミー位置から移動可能な距離内にダミーが生成されることを保証する．

3.3 位置プライバシ保護に関する要求

3.3.1 匿名領域

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，図 2(a) のように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，ユーザの位置を正確には特定できないが，このようなダミーの配置は，ダミーの存在範囲が小さいため，ユーザが存在する可能性のある領域が絞り込めてしまい，ユーザのおおよその位置の予測が可能になってしまう．

そこで本稿では，Lu ら [7] の定義に基づき，ユーザとすべてのダミーを包括する凸多角形を匿名領域と定義し，そ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 2 の場合は，(b) の方が匿名領域が大きいので，ユーザの位置曖昧性は大きい．

そこで提案手法では，ユーザの要求する匿名領域の大きさを達成できるように，ダミーを生成する．

3.3.2 追跡可能性

短期間の連続したサービス要求を想定すると，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，

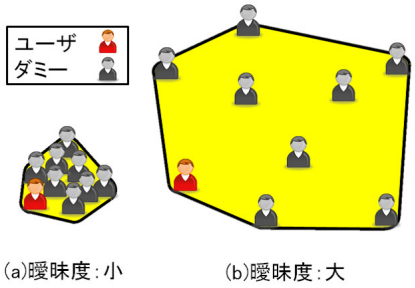

図 2 匿名領域

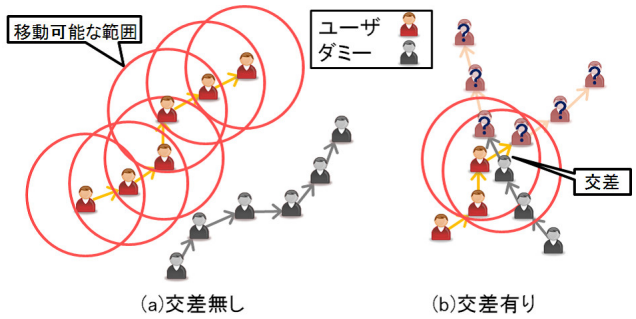

図 3 追跡可能性

複数の位置情報が短い時間間隔で与えられる場合に，それらを結ぶことにより，その軌跡を推測できてしまう性質を指す．これにより，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 3(a) のようにユーザの移動可能な範囲内をダミーが通過しない場合，ユーザの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．このような追跡を防ぐためには，図 3(b) のように，ユーザとダミーの移動経路を定期的に交差させることが有効である [11]．交差により，ユーザとダミーの移動可能な範囲に両者の位置が含まれ，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．

そこで提案手法では，訪問場所の意味的情報を考慮し，有効な交差となる場所にダミーを生成することにより，追跡可能性を低下させる．

3.4 ユーザの訪問場所の傾向の抽出

提案手法では，ダミーを生成する前に，まずユーザの訪問履歴からユーザの訪問場所の傾向を抽出する．具体的には，Ying[12] らの手法に基づき，ユーザの訪問履歴から訪問場所の意味的情報の系列を複数抽出し，系列パターンマイニングを施すことにより，ユーザの訪問場所の意味的情報に関して頻出する訪問パターンを求める．そして抽出した訪問パターンを木構造で表現し，ダミーの訪問場所間の遷移を決定する際に利用する．

具体的には，まず，ユーザの訪問履歴から表 1 のような訪問場所の意味的情報の系列を抽出する．得られたすべての

表 1　訪問履歴の意味的情報の系列

訪問場所
学校→病院→カフェ
学校→病院→カフェ→食事
カフェ→映画→食事

表 2　訪問パターンと支持度

訪問パターン	支持度
＜ 学校, 病院 ＞	0.667
＜ 学校, カフェ, 食事 ＞	0.333
＜ カフェ, 映画 ＞	0.333
⋮	⋮

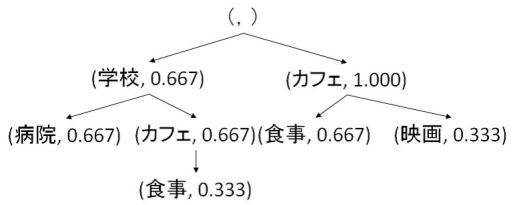

図 4　木構造で表現された訪問場所のパターンと支持度

系列に対して PrefixSpan[9] を適用することにより，頻出するユーザの訪問場所の傾向を，表 2 に示すような複数の訪問パターンとその支持度の組として取得する．PrefixSpan とは，頻出する部分系列パターンを抽出する系列パターンマイニングアルゴリズムであり，最小支持度以上のすべての系列パターンを出力する．系列パターンの支持度とは，全系列に対するその系列パターンを含む系列の割合のことであり，支持度が大きいほど，より頻出する系列パターンであるとみなすことができる．したがって，支持度が大きいほど，ユーザがより頻繁に訪問するパターンであると考えられる．

次に，得られた訪問場所の意味的情報に関する訪問パターンから，図 4 に示すような木構造を作成する．木構造を作成することで複数の訪問パターンをコンパクトに表現することができ，効率的に訪問パターンとその支持度を探索することができる．提案手法では，この木構造を利用してダミーの訪問場所を決定する．このように，ユーザの訪問場所のパターンを考慮することで，ダミーはユーザの訪問場所の傾向に従った遷移を行うことができる．

3.5　訪問場所の傾向を考慮したダミー生成手法

提案手法では，ユーザが要求したダミーの個数，匿名領域の大きさ，予測されたユーザの行動，および抽出されたユーザの訪問パターンに基づいて，ダミーが訪問すべき場所を決定する．そして，決定された訪問場所に向かって移動し，到着すべき時間にその訪問場所に到着して位置情報サービスを利用し，一定時間停止後に次の訪問場所に向けて移動するというダミーの行動スケジュールを生成する．

具体的には，以下の手順を要求されたダミー数だけ繰り返し，ユーザの訪問場所の傾向に従い移動するダミーの行動を一つずつ順に決定する．初めのダミーの行動は，ユーザの行動のみを考慮して決定し，二番目以降のダミーの行動は，ユーザと生成済みダミーの行動を考慮して決定する．

(1) 追跡可能性低下のための訪問場所の決定

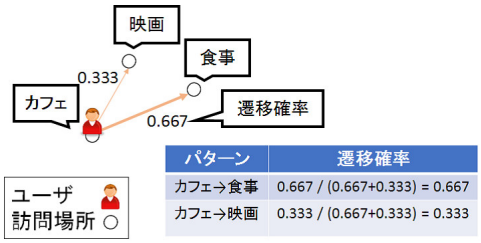

図 5　訪問場所間の遷移確率

(2) 匿名領域確保のための訪問場所の決定

まず，手順 (1) でユーザの訪問場所の傾向を考慮した効果的な交差を実現可能な訪問場所を一つ決定し，手順 (2) で，手順 (1) で得られた訪問場所からユーザのサービス利用開始時間および利用終了時間までのすべてのサービス利用時間に対して，サービス利用間隔内に到達可能で，かつ，ユーザが要求する匿名領域の面積を満たすことが可能な訪問場所を探索し，決定する．

3.5.1　追跡可能性低下のための訪問場所の決定

3.4 節で抽出した各訪問パターンの支持度を基に，追跡可能性を低下させるため，訪問傾向を考慮した効果的な交差を実現可能な訪問場所を決定する．ここで，効果的な交差とは，単にユーザとダミーの移動可能な範囲内にユーザとダミーの訪問場所が位置しているだけでなく，訪問場所の意味的情報も考慮した上で，ユーザが訪問する可能性が高い場所にダミーを配置することで，ユーザとダミーの識別をより困難にするような交差である．

まず，効果的な交差を実現可能な場所を探すために，図 4 で示した各訪問パターンの支持度の比から，到達可能な訪問場所間の遷移確率を計算する．図 5 に，図 4 から求められる遷移確率の例を示す．⟨ カフェ，食事 ⟩ の支持度は 0.667 で，⟨ カフェ，映画 ⟩ の支持度は 0.333 であるため，図 5 におけるカフェから食事への遷移確率は $\frac{0.667}{(0.667+0.333)} = 0.667$ と求められる．位置情報サービスの利用間隔から，時間的に複数の訪問場所へ到達可能である場合，この遷移確率が高い訪問場所ほど，ユーザが訪問する可能性が高くなる．ユーザの遷移確率を考慮したとき，効果的な交差が発生する条件として以下の 2 つの場合に分けられる．

(i)　ユーザが遷移確率の低い場所を訪問する場合 (図 6)

(ii)　ユーザが遷移確率の高い場所を訪問する場合 (図 7)

図 6(a) は (i) の場合であり，サービス利用間隔を考えると，遷移確率が高い場所 (訪問場所 C) へユーザが訪問可能であるにも関わらず，実際には，遷移確率の低い場所 (訪問場所 B) へ訪問する場合を表している．この場合には，図 6(b) のように，次のサービス利用時に，前のユーザの訪問場所 (訪問場所 A) から遷移確率が高い場所 (訪問場所 C) にダミーを配置させる．このようにすることで，訪問場所 A から，ダミーが訪問した場所 (訪問場所 C) へのユーザの遷移確率が高いため，訪問場所 C に訪問したダミーがユーザであると第三者に誤った認識をさせる可能性が高く

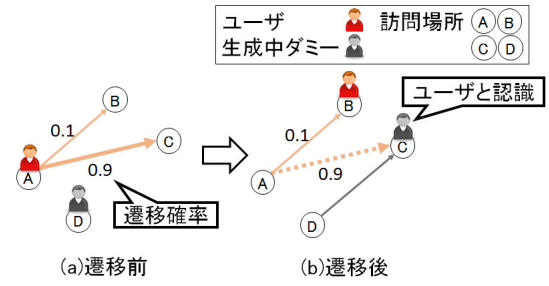

図 6 ユーザが遷移確率の低い場所を訪問する場合

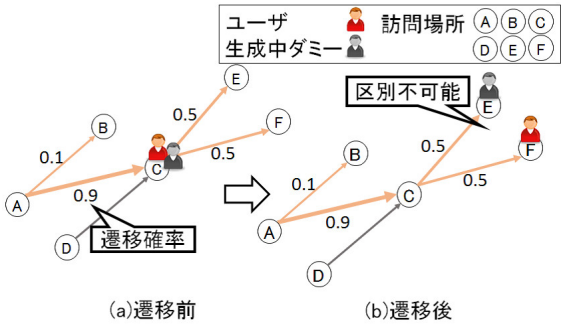

図 7 ユーザが遷移確率の高い場所を訪問する場合

なり，効果的な交差となる．

図 7(a) は (ii) の場合であり，ユーザが次の訪問場所として，遷移確率が高い場所 (訪問場所 C) に訪問する場合を表している．このとき，その場所 (訪問場所 C) へダミーも訪問させることにより，一時的に訪問場所を共有させる．さらに, 図 7(b) のように次の訪問場所を異なる場所 (訪問場所 E もしくは F) に設定することにより，第三者にはどちらにユーザが訪問するかを判断することが困難となり，効果的な交差となる．

提案手法では，この効果的な交差をユーザのどの訪問場所間で実現可能であるかを順次探索し，訪問可能な場所を一つ選択することでダミーの訪問場所を決定する．生成済みダミーとの交差を行う際には，生成済みダミーをユーザとみなし，同様の方法で訪問可能な場所を探索し，ダミーの訪問場所を決定する．交差場所を探索する際には，ユーザおよび全てのダミーの交差回数と交差が起きる時間帯に大きな差が生じないように，ユーザおよびダミーの交差回数と交差が起きる時間帯をシステム内で記録しておき，ユーザと生成済みダミーの中で交差回数の少ないものから交差があまり起きていない時間帯を見つけ出し，その時間帯で交差を実現可能な場所がないかを調べる．ここで，ユーザとダミーの交差回数を均等化するのは，頻繁に交差しているものがユーザであると容易に判断されないようにするためである．また，ある時間帯に交差が頻発すると，その時間帯はユーザとダミー間およびダミー同士の距離が極端に短くなり，匿名領域が非常に小さくなる可能性があるため，交差の時間帯も均等化を図る．

3.5.2 匿名領域確保のための訪問場所の決定

効果的な交差が実現可能な訪問場所の決定後，匿名領域を確保するために，その前後のユーザのサービス利用時にあわせてダミーが訪問する場所を決定する．訪問場所を決定する際には，ユーザの要求する匿名領域の面積を満たす場所に，訪問すべき場所を決定する．

初めに，3.5.1 項で決定した訪問場所の前後のサービス利用時におけるユーザと，生成済みダミーによって形成される匿名領域の面積を計算する．そして，形成された匿名領域と，ユーザの要求する匿名領域の面積を比較する．形成された匿名領域がユーザの要求する匿名領域よりも小さい場合，ユーザの要求する匿名領域を満たすことができる場所を訪問場所として決定する．ユーザの要求する匿名領域を満たす場所が存在しない場合，訪問可能な場所から，ユーザの要求する匿名領域に最も近い大きさの匿名領域を形成可能な場所を訪問場所として決定する．

一方，形成された匿名領域がユーザの要求する匿名領域より大きい場合，訪問可能な場所から，ユーザや生成済みダミーと成す匿名領域の大きさが最も小さい場所を訪問場所として決定する．

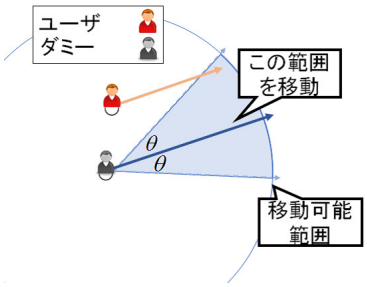

図 8 ユーザの進行方向を考慮したダミーの移動範囲

また，図 8 のようにユーザの進行方向を考慮し，ユーザの訪問場所間の進行方向のベクトルを取得し，それを基にダミーの訪問場所を決定する．具体的には，ユーザの進行方向から左右 θ 度以内の場所を探索し，ユーザの訪問傾向を考慮して訪問場所を決定する．これは，ユーザのおおよその進行方向にダミーを移動させることで，ユーザが孤立して移動し，ユーザであると特定され易くなることを防ぐためである．

そして，3.5.1 項で決定した訪問場所から，ユーザのサービス利用開始時間および利用終了時間までこの操作を繰り返し，訪問場所を順次決定する．決定されたすべての訪問場所をつなぐことで最終的なダミーの訪問経路を作成する．

4. 評価実験

提案手法の有効性を確認するために，地図上でユーザの動きをシミュレーションできるネットワークシミュレータ MobiREAL[*1]を用いて，京都の街を再現し，評価実験を行った．ユーザの動きは，ある訪問傾向を示しながら移動するモデルを利用した．シミュレーションにおける各パラメータは表のように定めた．地理情報のソーシャルネット

*1 http://www.mobireal.net

表 3　パラメータ

パラメータ	範囲
シミュレーション時間 [s]	36000
サービス利用間隔 [s]	[500,1000]
歩行速度 [m/s]	[1.0,2.0]
領域 [m^2]	15200^2
ダミー数 [個]	9,16
停止時間 [s]	[60,300]
要求匿名領域 [m^2]	$1000^2, 1200^2, \cdots, 2000^2$
訪問場所のカテゴリ数 [個]	10

ワーキングサービスである foursquare[*2]から取得した訪問場所のカテゴリ情報を，実験環境における地図上の訪問場所の意味的情報として割り当てた．取得したカテゴリの構造は，カテゴリ数 10，階層数 1 である．提案手法において，匿名領域確保のための訪問場所を決定する際に用いる，ユーザの進行方向からの角度 θ は 45° に設定した．

4.1　評価指標

- **AR-Count（Anonymous area achieving Ratio - Count）**

  要求された匿名領域のサイズを，ダミー配置により実際に達成できた回数のサービス利用の総数に対する割合を AR-Count と定義する．実際に確保できる匿名領域の大きさは，要求された匿名領域の大きさよりも大きくなる場合も，小さくなる場合も存在する．AR-Count は，要求匿名領域をどの程度の頻度で達成できたかを示しており，常時達成できた場合には 100%となる．

- **AR-Size（Anonymous area achieving Ratio - Size）**

  要求された匿名領域のサイズに対する，ダミー配置により実際に達成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 の値が 100%よりも大きければ，平均的にユーザの要求以上にユーザ位置を曖昧化できているとみなすことができる．

- **MTC（Mean Time to Confusion）**

  第三者から見たある位置情報がユーザのものである確率を，ユーザ確率とよぶ．ここで，何らかの原因によりユーザ位置が特定されたとき，ユーザ確率は 1 となる．その後の，各々の位置情報のユーザ確率の変化を，3.5.1 項で説明した訪問場所間の遷移確率を利用して，以下の条件により求める．図 9(a) のように，ある時点において，ユーザ確率が α であるダミー a(もしくはユーザ) とユーザ確率 β のダミー b が，次の時点で図 9(b) のようにユーザの訪問パターンを満たす訪問可能な場所 (訪問場所 B，C) に訪問する場合，二つのダミーのユーザ確率が分配される．図 9(b) における交差後のダミー b のユーザ確率を γ とすると，$\gamma = \frac{0.1}{(0.1+0.9)} \times \alpha + \frac{0.2}{(0.2+0.8)} \times \beta$ と求められる．また，図 10(a) のように，ダミー同士が訪問場所を共有した場合，図 10(b) のように次のサービス利用時に別々の場所へ訪問すると，ユーザ確率が分配される．図 10(b) における交差後のダミー b のユーザ確率 δ は，$\delta = (\alpha + \beta) \times \frac{0.2}{(0.2+0.8)}$ と求められる．

  このように求めた，各々のダミーのユーザ確率に，既存研究 [10] で提案されている MTC を適用し，ユーザの追跡可能性を評価する．MTC はダミーのユーザ確率 p_i としたときのエントロピー $H = -\sum p_i \log p_i$ が閾値を超えるまでの時間である．本報告では，閾値を 1 とし，ユーザ位置が第三者に特定され，エントロピーが 0 になった時点から，エントロピーが 1 を超えるまでにかかる時間の平均を MTC とする．この指標は，ユーザ位置が特定されてから再び曖昧化させるまでの平均時間であるため，この値が小さければ追跡可能性が小さいということを表している．

*2　https://ja.foursquare.com/

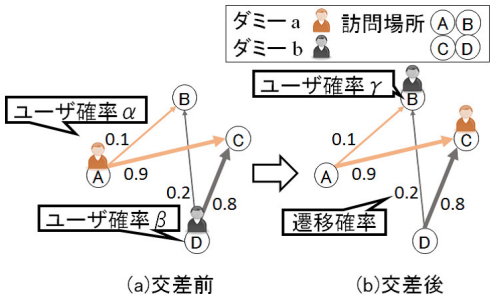

図 9　交差によるユーザ確率分配

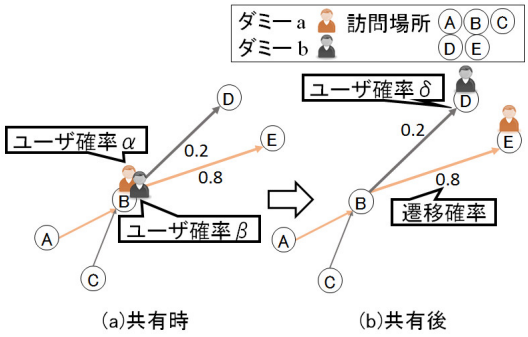

図 10　訪問場所を共有する場合のユーザ確率分配

4.2　評価手法

本実験では，以下の二つの手法の性能を比較する．

(1) 比較手法

ユーザの訪問傾向を考慮しない先行研究の手法 [5]．この手法は，ある場所を訪問し，一定時間停止した後，再び次の訪問場所に移動するというユーザの移動を想定し，ユーザの行動が既知として，それを基にユーザと同様に訪問場所を訪れながら移動するダミーを生成する．この手法では，ユーザおよびダミーの存在が少ない領域を周期的に探しだし，その領域内に匿名領域確保のための訪問場所を決

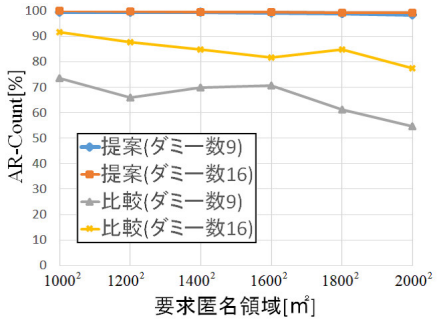

図 11 AR-Count

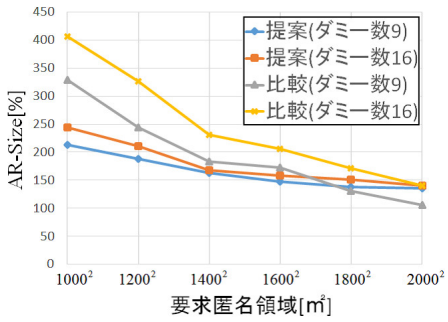

図 12 AR-Size

定する．それ以外の場所では，可能な限りユーザおよび生成済みダミーと訪問場所を共有することで追跡可能性を低下させる．比較手法は，ユーザの訪問場所の傾向を考慮していないため，ダミーがユーザと同様の訪問傾向をもつことが難しく，ダミーの訪問場所の傾向がユーザと異なる場合，ダミーであると特定されてしまう．本実験では，この問題自体は無視しているが，MTC への影響があることが予想される．

(2) 提案手法

既知であるユーザの行動に基づいて，ユーザの訪問場所の傾向を考慮しながら移動するダミーを生成する 3.5 節で提案した手法．

4.3 実験結果

4.3.1 AR-Count, AR-Size

さまざまな要求匿名領域のサイズに対する，AR-Count および AR-Size を調べた．その結果をそれぞれ図 11，図 12 に示す．要求される匿名領域が大きくなるに従い，提案手法，比較手法ともに AR-Count，AR-Size の値が低くなっている．これは，交差を発生させることで追跡可能性を低下させるため，ユーザとダミー間やダミー同士の距離が短くなり，匿名領域が小さくなってしまう傾向があることに起因する．そのため，要求される匿名領域が大きくなると，それを十分に満たすことがより困難になる．

ダミー数が 9 および 16 の場合共に，提案手法では，比較手法より高い AR-Count を実現できているにも関わらず，AR-Size は比較手法よりも低い値をとる．ダミー数が 9 のとき，提案手法は比較手法に比べて平均 33.5[%] ほど高

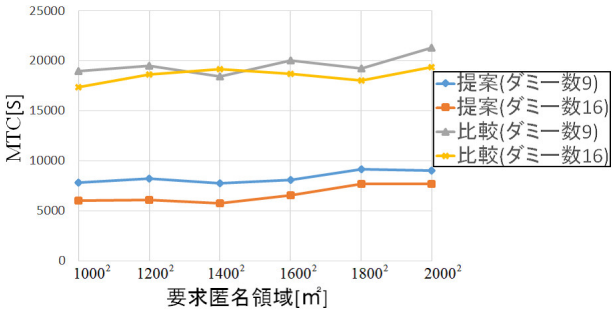

図 13 MTC

い AR-Count となり，AR-Size は平均 15.3[%] ほど低くなる．ダミー数が 16 のとき，提案手法は比較手法に比べて平均 14.9[%] ほど高い AR-Count となり，AR-Size は平均 27.6[%] ほど低くなる．これは，提案手法では，ユーザおよび生成済みダミーが形成する匿名領域の面積を，ユーザの要求する匿名領域の面積と比較し訪問場所を決定することで，ユーザとダミーが形成する匿名領域をむやみに広げることなく，要求する匿名領域の面積を確保する場所，もしくは，維持する場所を探索して訪問場所として決定するためである．一方，比較手法では，匿名領域確保のための訪問場所を決定するとき以外，地理的に到達可能である場所があれば形成されている匿名領域の面積を考慮することなく訪問場所を決定するため，極端に大きな匿名領域を形成したり，極端に小さな匿名領域を形成したりすることがある．そのため，比較手法では，ダミー数 9，要求匿名領域サイズ $1000^2[m^2]$ の場合のように，AR-Size は 328%と高い値になるが，AR-Count は 73.7%と低い値になるといった，匿名領域の面積のばらつきが生じる．提案手法は，ダミー数 9，ダミー数 16 共に，すべての要求匿名領域サイズにおいて AR-Size が 100%以上となり，かつ，比較手法に比べて高い AR-Count を維持できている．

提案手法において，ダミー数が 9 と 16 の場合を比較すると，すべての要求匿名領域サイズにおいて，ダミー数が 9 の場合に比べ，ダミー数が 16 の場合のほうが AR-Count，AR-Size 共に大きな値となる．ダミー数が 9 の場合とダミー数が 16 の場合の AR-Count の差の平均は 0.3[%] であり，AR-Size の差の平均は 7.9[%] である．これは，ダミー数が多くなると，ユーザの要求する匿名領域の大きさを満たす訪問場所が見つけられない可能性を減らすことができるからである．

4.3.2 MTC

さまざまな要求匿名領域に対する，MTC を調べた．その結果を図 13 に示す．ダミー数が 9 の場合，提案手法と比較手法を比較すると，提案手法は，全ての要求匿名領域サイズにおいて，比較手法よりも MTC を低減できている．ダミー数が 9 のとき，提案手法は比較手法に比べて，MTC は平均 11,216[s] ほど小さくなる．これは，比較手法はユーザの訪問場所の傾向を考慮することなく訪問場所を

決定しているため，効果的な交差の機会が少ないことに起因する．実際，シミュレーション内の比較手法のユーザおよびダミーの総交差回数の平均は，ダミー数が 9 の場合で約 4 回，ダミー数 16 の場合で約 6 回となり，多くの交差を発生させているが，ユーザの訪問傾向を考慮した効果的な交差にはなっていない．一方，提案手法では，ダミー数が 9，16 の場合共にユーザおよびダミーの交差回数の平均は約 2 回であるものの，ユーザの訪問傾向を考慮することにより，確実にユーザ確率を減少できる効果的な交差を発生できている．その結果，MTC を効果的に低下できている．

提案手法，比較手法共に，要求匿名領域サイズが変化しても MTC に大きな変化は見られない．これは，両手法とも要求匿名領域の大きさに関わらず，交差を決定しているためである．

また，提案手法において，ダミー数 9 と 16 の場合を比較すると， ダミー数 16 の方が MTC が小さくなり，その値の差の平均は 1740[s] となる．これは，むやみに匿名領域を広げようとしていないため，ダミー数が多くなるとユーザやダミーが互いの移動可能な範囲に入る可能性が高くなり，また，ダミーがユーザの訪問傾向を考慮しながら移動していることと相まって，意図しない効果的な交差が発生しているからである．

以上の結果から，MTC をより低下させるためには，効果的な交差を増やすことが有効であることが分かる．提案手法において現状以上に MTC を低減するためには，効果的な交差を実現可能な訪問場所を，各ダミーで複数箇所選定するような拡張が必要であると考えられる．

5. おわりに

本稿では，位置情報サービス利用におけるユーザの位置プライバシ保護を目的として，ユーザの訪問場所の傾向を考慮したダミー生成手法を提案した．提案手法では，ユーザがいくつかの場所を訪問しながら移動する際に，ユーザの訪問場所の傾向に合わせてユーザと同様にいくつかの場所に訪問するダミーの行動スケジュールを生成する．この際，各ダミーがユーザの要求する匿名領域の大きさを満たすように，次に訪問すべき場所を決定することで匿名領域を確保する．さらに，ユーザの訪問傾向を考慮して，効果的な交差を実現可能な場所にダミーを訪問させることにより，追跡可能性を低下させる．

評価実験の結果，提案手法では，先行研究で提案した手法に比べ，すべての要求匿名領域サイズにおいて，十分な匿名領域を確保できていることを確認した．また，ユーザの訪問場所の傾向を考慮することで追跡可能性を低減できていることを確認した．

提案手法では，効果的な交差を実現する訪問場所を各ダミーに対して一つだけ決定している．評価実験により，効果的な交差は追跡可能性の低下に非常に有効であることを確認したため，更なる追跡可能性の低下を目指して，効果的な交差を実現可能な訪問場所を複数箇所決定できるように手法の拡張を行う予定である．さらに，より現実的な状況に適用可能なように，ユーザの行動予測が外れてしまった場合でも，ユーザの行動に対応できるように提案手法を拡張することを検討している．

謝辞 本研究の一部は，マイクロソフトリサーチアジア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] C. A. Ardagna, M. Cremonini, E. Damiani, S. D. C. di Vimercati, and P. Samarati: Location Privacy Protection through Obfuscation-Based Techniques, *In Proc. DBSec*, pp. 88–97, 2005.

[2] A. R. Beresford and F. Stajano: Location privacy in pervasive computing, *In Proc. Pervasive Computing*, pp. 46-—55, 2003.

[3] M. Duckham and L. Kulik: Simulation of Obfuscation and Negotiation for Location Privacy, *In Proc. CONSIT*, pp. 31–48, 2005.

[4] M. Gruteser and D. Grunwald: Anonymous Usage of Location-Based Services through Spatial and Temporal Cloaking, *In Proc. ISENIX MobiSys*, 2003.

[5] R. Kato, M. Iwata, T. Hara, A. Suzuki, Y. Arase, X. Xie, and S. Nishio: A dummy-based anonymization method based on user trajectory with pauses, *In Proc. GIS*, pp. 249–258, 2012.

[6] H. Kido, Y. Yanagisawa, and T. Satoh: An Anonymous Communication Technique using Dummies for Location-based Service, *In Proc. ICPS*, pp. 88–97, 2005.

[7] H. Lu, C. S. Jensen, and M. L. Yiu: PAD : Privacy-Area Aware, Dummy-Based Location Privacy in Mobile Services, *In Proc. MobiDE*, pp. 47–60, 2007.

[8] M. Mano and Y. Ishikawa: Anonymizing User Location and Profile Information for Privacy-aware Mobile Services, *In Proc. GIS-LBSN*, p. 69–75, 2010.

[9] J. Pei, J. Han, B. Mortazavi-Asi, and H. Pinto: Prefixspan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth, *In Proc. ICDE*, pp. 215–224, 2001.

[10] R. Shokri, J. Freudiger, M. Jadliwala, and J. P. Hubaux: A Distortion-Based Metric for Location Privacy, *In Proc. WPES*, p. 6, 2009.

[11] Y. Yanagisawa, H. Kido, and T. Satoh: Location Traceability of Users in Location-Based Services, *In Proc. MobiQuitous*, 2006.

[12] J. J. C. Ying, W. C. Lee, T. C. Weng, and V. S. Tseng: Semantic Trajectory Mining for Location Prediction, *In Proc. GIS*, 2011.