ダミーの広範囲な分散を考慮したユーザ位置曖昧化手法

加藤 諒[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：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，停止しながら移動するダミーの位置情報を生成するユーザ位置曖昧化手法を提案した．しかし，この手法で生成されたダミーは，時間の経過にしたがって狭い範囲に密集してしまい，ユーザの位置を十分に曖昧化できていないことがわかった．そこで本稿では，定期的にダミーをユーザや他のダミーの存在数が少ない位置に移動させることで，時間に依らず，ダミーを広範囲に分散させるユーザの位置曖昧化手法を提案する．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，提案手法の方が十分にユーザの位置を曖昧化できていることを確認した．

A User Location Anonymization Method Achieving Widely Distributed Dummies

RYO KATO[1] MAYU IWATA[1] TAKAHIRO HARA[1] YUKI ARASE[2] XING XIE[2] SHOJIRO NISHIO[1]

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，ユーザの住居や勤務先，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．

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．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，ユーザの位置を曖昧化できる．しかし，既存の手法では，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など，実環境における制約を十分に考慮できていない．

そこで，これまでに筆者らの研究グループでは，実環境を考慮したダミー生成手法である DumGrid[9] および Dum-P[4] を提案した．両手法とも，地図情報を参照し，ユーザが存在できる場所にダミーを生成し，さらに，ユーザの移動速度も考慮してダミーを生成することで，ダミーをユーザと同様に移動させる．DumGrid では，ユーザの移動にその都度対応できるように，ユーザの周囲にダミーをグリッド状に配置し，ダミーをユーザに纏わりつかせるように移動させていた．しかし，DumGrid は，ユーザが停止せずに移動し続けるという行動モデルを想定していたため，ユーザがコンビニエンスストアに立ち寄ったり，休憩地点で休憩するなど，いくつかの地点で停止しながら移動する行動モデルを想定すると，ダミーが交差点の中で停止するなど不自然な行動をすることがある．そこで，Dum-P では，ユーザの停止地点や停止時間も含めた行動予測が可能な状況を想定し，予測されたユーザの行動に基づいて，停止地点や停止時間を含む自然なダミーの行動を事前に決定する．しかし，Dum-P では，ユーザとダミーの地理的分布を広範囲に広げるために設定する停止地点を，各ダミーの行動プランにおいて，早い時間帯に唯一つだけしか含ま

[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

ないため，ダミーは時間の経過に伴い狭い範囲に密集するように動いてしまう．ダミーが密集してしまうと，たとえユーザとダミーの区別ができなかったとしても，ユーザのおおよその位置が推測できてしまうため，ユーザの位置プライバシに関する要求を十分に満たしているとはいえない．

そこで，本研究では，ユーザの位置プライバシに関する要求を十分に満たすため，Dum-P を拡張し，時間の経過に依らず広範囲にダミーを分散する手法 (Dum-P-Cycle) を提案する．Dum-P-Cycle では，Dum-P 同様に，既知であるユーザの移動経路に基づいて，ダミーが停止すべき時間および地点を決定し，その時間にその地点を経由して移動を行うダミーの行動を決定する．この際，ダミーの総移動時間の中に周期を設け，周期ごとにユーザや他のダミーの存在数が少ない地点にダミーが停止する地点を決定することで，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，周期以外の時間帯で，ユーザや他のダミーと停止する地点を共有して交差をさせることで，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．これにより，Dum-P-Cycle では，時間が経過してもダミーの位置が狭い範囲に密集しないようにし，ユーザの総移動時間を通じて，ユーザの位置を曖昧にすることが可能になる．

以下では，2 章で関連研究を説明し，3 章で先行研究の問題点とダミーの広範囲な分散を考慮した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 つの手法について述べる．

文献 [3][7] では，ユーザが直接自身の位置情報をサービスプロバイダに送るのではなく，信頼された第三者サーバを利用する手法が提案されている．第三者サーバは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．これによりユーザの位置を $\frac{1}{k}$ 以上の確率で特定不可能になる．ただし，この手法では完全に信頼できる第三者サーバの存在を前提としており，実環境で用いるのは困難である．

文献 [1][2] では，自身の位置情報として，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はできなくなるが，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 [5][6][10] では，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，**図 1** のようにサービスプロバイダにサービス要求を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は小さくなる．このようなダミーを用いた手法は，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．しかし，文献 [5][6][10] の手法ではダミーの生成位置に制約がなく，高速道路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．

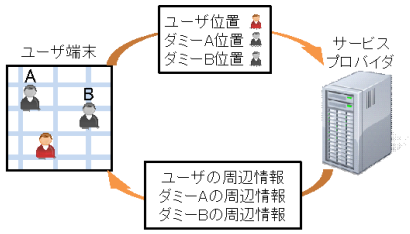

図 1 ダミーを用いた位置情報サービスの利用例

筆者らは先行研究において，これらの実環境における制約を考慮した 2 種類のダミーによるユーザ位置曖昧化手法 (DumGrid[9] および Dum-P[4]) を提案した．しかし，DumGrid では，ユーザが停止することなく移動し続けることを想定しており，ユーザが停止する状況を考慮していない．そのため Dum-P では，予測されたユーザの行動をもとに，ユーザと同様に停止しながら移動するダミーの行動を決定する．しかし，Dum-P によって生成されたダミーは，時間が経過すると徐々に密集するように動いてしまう．ダミーが密集してしまうと，たとえユーザとダミーの区別ができなかったとしても，ユーザのおおよその位置が推測できてしまう．一方，提案手法である Dum-P-Cycle では，周期的にダミーが広範囲に分散するようにすることで，ユーザの位置を広範囲に曖昧化することができる．

3. ユーザの行動に基づいたダミーの移動経路生成手法

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の保護に関する要求について述べる．次に，先行研究において提案した手法 (Dum-P)[4] を説明し，最後に，Dum-P における問題を解決する提案手法 (Dum-P-Cycle) の詳細について述べる．

3.1 想定環境

本研究では，ユーザは短時間に連続して位置情報サービスを利用する状況を想定する．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．

ユーザはある目的地を持って移動を開始し，移動中にいくつかの停止地点で停止しながら，最終的に目的地に到達するという行動モデルを想定する．そして，それぞれの停止地点で，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し，停止地点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，ユーザの停止地点，停止時間，移動経路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，ユーザが事前に行動プランを登録したり，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．

3.2 実環境でダミーを生成する際に考慮すべき条件

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を要求してから，3 分後に新たにサービス要求した場合を考える．この際，新しいクエリにおいて，直前のクエリのどのダミー位置からも 3 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できてしまう．

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いため，ユーザが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の予測が可能になってしまう．

そこで本稿では，Lu ら [6] 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．

Dum-P-Cycle では，生成するダミーの行動の総移動時間

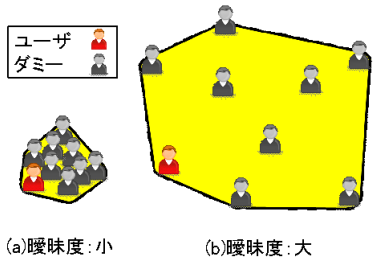

図 2 匿名領域

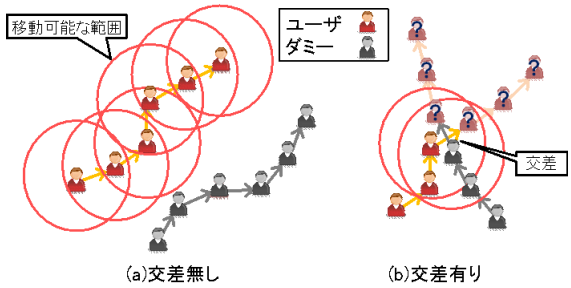

図 3 追跡可能性

を周期で区切り，各周期ごとにユーザおよびダミーの存在する数が少ない領域を探し出し，その領域にダミーを移動させることにより，要求される匿名領域の保証を試みる．

3.3.2 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，ユーザとダミー経路を定期的に交差することが有効である [10]．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．

Dum-P-Cycle では，ユーザとダミー，またはダミー間で停止する地点を共有させることで交差を発生させ，追跡可能性を低下させる．

3.4 先行研究 (Dum-P)

3.4.1 Dum-P のダミー生成の手順

Dum-P では，予測されたユーザの行動に基づいて，ダミーが停止すべき地点 (停止地点)，その地点に到着すべき時間 (停止地点到着時間) を決定する．そして，停止地点に向かって移動し，停止地点到着時間にそこに到着し，その後停止するというダミーの行動プランを生成する．

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，ダミーの行動を 1 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，2 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) 匿名領域確保のための停止地点 (基準地点) および停止地点到着時間 (基準地点到着時間) の決定：
ユーザおよび生成済みダミーの存在数が連続的に最も少ない領域と時間帯を探し出し，その領域内に基準地点を決定し，その時間帯の最も早い時間として基準地点到着時間を決定する．

(2) 追跡可能性低下のための停止地点 (共有地点) および停止地点到着時間 (共有地点到着時間) の決定：
基準地点から，移動可能な範囲を考慮して，ユーザおよび生成済みダミーと停止地点を共有できる地点に，共有地点を決定し，同じ時間に停止地点を共有するように，共有地点到着時間を決定する．

(3) 自然なダミーの移動経路作成のための停止地点 (途中停止地点) および (途中停止地点到着時間) の決定：
基準地点および共有地点を単純に結ぶだけでは，ダミーの移動が不自然になることがあるため，そのような場合には途中停止地点および途中停止時間を決定する．

以上のように，Dum-P は，基準地点，共有地点，途中停止地点という 3 種類の停止地点で停止しながら移動するダミーを生成する．

3.4.2 Dum-P の問題点

Dum-P の問題点を明らかにするため，ユーザとダミーの移動をシミュレートし，統計量による評価実験を行った．評価実験では，シミュレーション時間を 50000[s]，サービス利用間隔を 180[s]，ダミー数を 16，要求匿名領域を 1600^2[m^2] とした．Dum-P の性能を測るため，以下に示す評価指標を用いた．

- **AR-Count（Anonymous area achieving Ratio - Count）**
  要求された匿名領域を，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 (要求) 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．
- **MTC（Mean Time to Confusion）**
  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．ここで，何らかの原因によりユーザ位置が特定された時，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遷移を以下の条件により求める．ある時点において，ユーザ確率が α であるダミー (もしくはユーザ) とユーザ確率 β のダミーが，次の時点でお互いの移動可能範囲に入った場合，二つのダミーは区別不可能となる．このとき，両ダミーのユーザ確率を $\frac{\alpha+\beta}{2}$ と計算する．

  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 [8] 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 $H = -\sum p_i \log p_i$ が閾値を越えるまでの時間である．本稿では，閾値を 1 とし，ユーザ位置がサービスプロバイダに特定され，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．

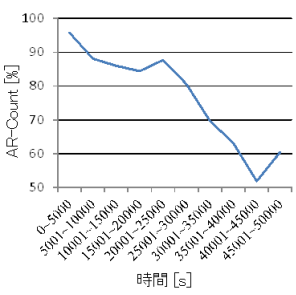

図 4 Dum-P における AR-Count の時間経過

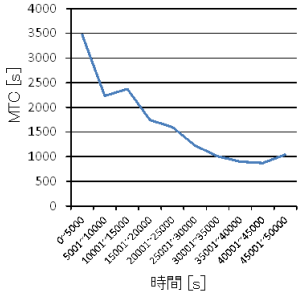

図 5 Dum-P における MTC の時間経過

Dum-P における AR-Count および MTC の時間経過をそれぞれ，図 4 および図 5 に示す．この結果より，時間が経過するに従い，AR-Count および MTC の値は減少している．つまり，Dum-P によって生成されたダミーは，移動開始直後は広範囲に位置しているが，時間が経つにつれて，狭い範囲に密集するように動くことがわかる．Dum-P では，ユーザとダミーの存在数が最も少ない時間帯のなかで，最も早い時間に基準地点到着時間を決定するため，多くのダミーは総移動時間のなかで，比較的早い時間に唯一つの基準地点に移動し，その後は匿名領域を確保する動きを行わない．その結果，時間経過とともに匿名領域は小さくなる．

3.5 提案手法 (Dum-P-Cycle)

3.4.2 項で言及した Dum-P の問題を解決するため，Dum-P-Cycle を提案する．Dum-P-Cyle では，Dum-P と同様に，ユーザの移動経路に基づいて，ダミーの停止地点と停止地点到着時間を決定し，停止地点到着時間に停止地点で停止するというダミーの移動経路を作成する．Dum-P では，基準地点はダミーの移動経路の中にただ 1 つだけ含まれるように設計したが，それだけでは，時間が経過するに伴い，ダミーの位置が狭い範囲に密集するようになってしまった．そこで，Dum-P-Cycle では，ダミーの移動経路のなかに複数の基準地点を決定することで，ダミーが匿名領域を確保する動きを複数回にする．

3.5.1 Dum-P-Cycle の方針

Dum-P-Cycle では，匿名領域をユーザの総移動時間を通して確保できるように，総移動時間の中に周期を設け，その周期ごとに基準地点および基準地点到着時間を決定する．周期的に基準地点を決定することで，その周期の前後には匿名領域を確保するようにダミーを移動させる．ここで，基準地点を設定するための周期は全てのダミーで統一する．各ダミーに異なる周期を与えた場合，あるダミーは広範囲に移動するが，別のダミーは交差をしようと移動する．その結果，交差をするために移動しなければならないダミーの移動距離が長くなってしまい，十分に追跡可能性を低下できなくなってしまう．ダミーに同一の周期を与えることで，全てのダミーは周期ごとにユーザの位置の曖昧性を高めるために広範囲に位置し，周期以外の時間帯は交差をするためにある程度密集するように位置する．これにより，匿名領域の確保と，追跡可能性の低下が効率的に行われることが期待できる．

3.5.2 Dum-P-Cycle の手順

Dum-P-Cycle では，ユーザの移動経路および生成済みダミーの移動経路に基づいて新たなダミーを作成する．具体的には，ユーザの要求したダミー数分だけ以下の手順を繰り返す．

(1) 周期に分割：新たなダミーの総移動時間を周期 T[s] ごとに区切る．

(2) 各周期の基準地点および基準地点到着時間の決定：周期ごとにダミーの基準地点および基準地点到着時間を決定する．最初の周期から順に基準地点と基準地点到着時間を決定し，全ての周期における基準地点および基準地点到着時間を決定する．

(3) 共有地点および共有地点到着時間の決定：各周期の基準地点間に可能な限り多くの共有地点および共有地点到着時間を決定する．

(4) 途中停止地点および途中停止地点到着時間の決定：自然なダミーの移動経路になるように適宜途中停止地点および途中停止地点を決定する．

このように Dum-P-Cycle では，各ダミーの基準地点をユーザの総移動時間/周期 T 個決定し，複数の基準地点を決定した後は，文献 [4] の Dum-P における方法と同様に共有地点と途中停止地点を順に決定する．

ダミーの生成の手順 2 にあたる周期における基準地点および基準地点到着時間を決定するアルゴリズムを Algorithm 1 に示す．具体的には，以下の手順で基準地点および基準地点到着時間を決定する．

i. 基準地点到着時間の決定 (Algorithm 1 の 6 および 7 行目) : 各周期の前後にダミーごとに異なる基準地点到着時間を設定する．

ii. グリッド領域の作成 (Algorithm 1 の 8 行目) : 各周期でユーザと生成済みダミーの平均位置を中心とするグリッド領域を作成する．

Algorithm 1：生成中ダミー (n 番目) の基準地点および基準地点到着時間の決定

```
1: input: a list of a user and generated dummies' movements D = {D_0, ..., D_{n-1}} (sets of their pause positions, pause start times and pause durations), possible pause positions based on the map information P = {P_0, ..., P_m}, cycle lengh T, simulation end time t_end
2: output: sets of pause positions and pause start times PP of nth dummy
3:
4: k ← 0
5: repeat
6:    t ← kT
7:    t_base ← t+ random(−a,+a) // determine kth base pause start time
8:    generate a grid with 3 × 3 cells G = {G_0, ..., G_8} around the center of positions of D at t
9:    for i = 0 to 8 do
10:      G_i.exist_t ← the number of D within G_i
11:   end for
12:   if k = 0 then
13:      //determine first base pause position
14:      G_base ← G_i with min(G_0.exist_t, ..., G_8.exist_t)
15:      p_base ← random(P in G_base)
16:   else
17:      //determine subsequent base pause position
18:      p_base ← NULL
19:      repeat
20:         if p_base ≠ NULL then
21:            append p_base to p_false  // p_false is a list of unreachable pause positions
22:            if the number of P in G_base = the number of p_false then
23:               append G_base to G_false  // G_false is a list of grid cell where all pause positons are unreachable
24:               p_false ← 0
25:            end if
26:         end if
27:         G_base ← G_i with min(G_0.exist_t, ..., G_8.exist_t; except G_false)
28:         p_base ← random(P in G_base; except p_false)
29:      until (p_base is within the reachable area from a pause position with the latest pause start time in PP of nth dummy) OR (the number of cells in G_false = 9)
30:   end if
31:   if the number of cells in G_false ≠ 9 then
32:      PP ← <p_base, t_base>
33:   end if
34:   k ← k + 1
35: until kT > t_end
36: return PP
```

iii. 各グリッドに存在するユーザおよびダミー数のカウント (Algorithm 1 の 9〜11 行目) : 各周期において，ユーザおよび生成済みダミーがグリッド領域内のどのセルにどれだけ存在するのかを算出する．

iv. 基準地点の決定 (Algorithm 1 の 12〜30 行目) : ユーザや生成済みダミーの存在する数が少ないセル内に基準地点を設定する．

手順 i において，周期に基づき基準地点到着時間を決定する．ここで，周期の時間に基準地点到着時間を決定すると，全てのダミーが同時に停止してしまうため，周期の時間を中心とした一定時間の中からランダムに基準地点到着時間を決定する．手順 ii では，図 6 に示すように，各周期において，ユーザおよび生成済みダミーが存在する平均位置を中心としたグリッド領域を作成する．グリッド領域は 3×3 のセルを持ち，ユーザの要求する匿名領域の大きさ S を満たすように，一辺の長さを $\sqrt{S}$ とする．そして，手順 iii において，配置したグリッド領域に対してユーザおよびダミーが各セルにどれだけ存在するのかを算出する．これにより，周期ごとにユーザおよび生成済みダミーの存在数が最も少ない領域 (セル) を特定することができる．手順

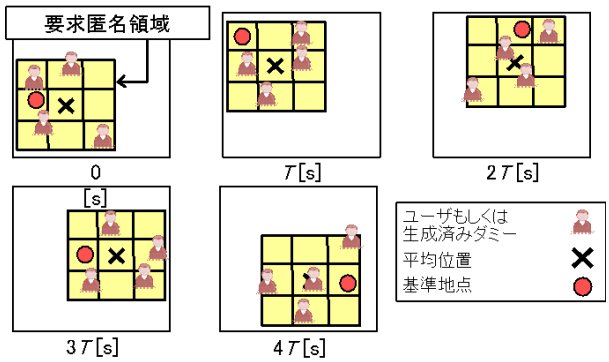

図 6　各周期 T におけるグリッド領域の利用

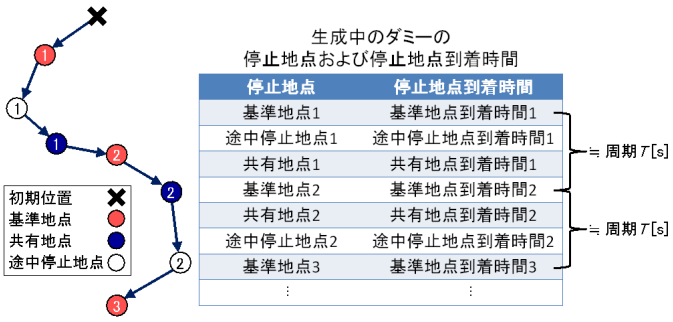

図 7　Dum-P-Cycle で生成されたダミーの移動経路の例 (周期 T[s])

iv において，手順 iii で求めた領域内のランダムな位置に最初の周期の基準地点を決定する．そのような領域が複数存在する場合は，その中からランダムに 1 つの領域を選び出す．また，2 番目の周期以降の基準地点は，既に決定した基準地点から，1 周期の間に到達可能な地点であるかを考慮しなければならない．既に決定した基準地点から到達可能でない場合は，ダミーがユーザが移動可能なスピードよりも早く移動してしまい，ダミーと特定されてしまう．そのような場合は，ユーザおよびダミーの存在数が等しい別の領域，もしくはユーザおよびダミーの存在数が次に少ない領域に基準地点を決定する．このようにすることで，ダミーの移動が不自然にならないようにする．

Dum-P-Cycle によって生成されたあるダミーの移動経路の例を図 7 に示す．図 7 では，周期を T[s] としているため，前後 2 つの基準地点到着時間の間隔はこの周期に近い時間となる．このようにダミーの停止地点を決定して移動させることで，匿名領域の確保のための動きと追跡可能性の低下のための動きを周期的に繰り返すダミーを生成できる．

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，道路上をランダムに停止しながら移動するモデルを利用した．ユーザやダミーの停止可能地点は，交差点間の道路に 50[m] 間隔で位置するものとした．また，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，3.4.2 項において説明した二つの評価指標 (AR-Count および MTC) を用いた．

[*1] http://www.mobireal.net

表 1　パラメータ

パラメータ	範囲
シミュレーション時間 [s]	50000
サービス利用間隔 [s]	180
歩行速度 [m/s]	1.30
領域 [m^2]	15200^2
ダミー数 [個]	16, 25
最大停止時間 [s]	600
最小停止時間 [s]	60
要求匿名領域 [m^2]	$1000^2, 1200^2, \cdots, 2000^2$
基準地点を決定する周期 [s]	5000,10000,$\cdots$,30000

4.1　評価手法

本実験では，以下の三つの手法の性能を比較する．

(1) DumGrid

ユーザの行動を予測できない状況を想定した，先行研究の提案手法 [9]．この手法では，匿名領域を確保しつつ，ユーザの動きの変化に対応するため，ダミーをユーザの周りにグリッド状に配置し，移動させる．この手法はユーザの行動を予測できると想定していないため，ダミーの生成はユーザの行動にリアルタイムに対応して行われる．そのため，ユーザが停止しながら移動する際に，ダミーも適宜自然に停止しながら移動することが難しくなり，不自然な場所で停止してしまう．したがって，視覚的にユーザを特定されてしまう可能性があるが，本評価ではその欠点は無視するものとする．

(2) Dum-P[4]

(3) Dum-P-Cycle(提案手法)

4.2　実験結果

4.2.1　AR-Count

さまざまな要求匿名領域に対する，AR-Count を調べた．その結果を図 8 に示す．ここで，Dum-P-Cycle の基準地点を設定する周期は 20000[s] とした．要求匿名領域が大きくなるに従い，全ての手法で AR-Count の値が小さくなっている．これは，ユーザ，ダミー間で交差を発生させることで追跡可能性を低下させるため，ユーザやダミーとの距離が近くなり，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 16 において，Dum-P-Cycle と DumGrid を比較すると，すべての要求匿名領域において，Dum-P-Cycle は DumGrid に比べて AR-Count の値は大きくなり，その差の平均は 27.8%である．また，Dum-P-Cycle は Dum-P と比べても AR-Count は大きくなり，その差の平均は 31.2%であ

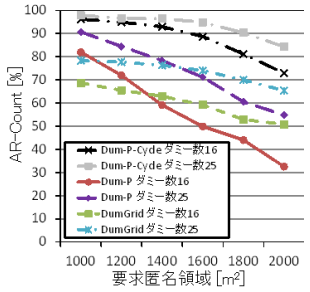

図 8 AR-Count (Dum-P-Cycle の周期 20000[s])

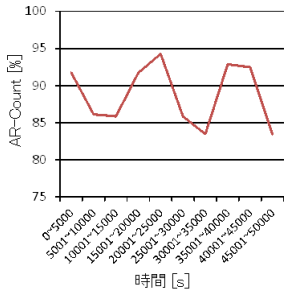

図 9 Dum-P-Cycle における AR-Count の時間経過 (ダミー数 16，要求匿名領域 1600^2[m^2]，周期 20000[s])

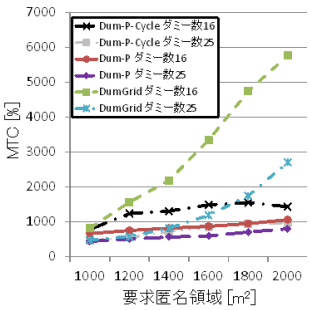

図 10 MTC (Dum-P-Cycle の周期 20000[s])

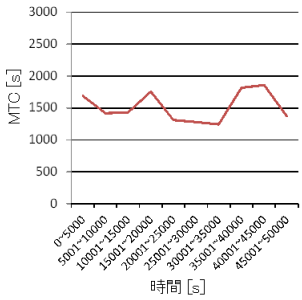

図 11 Dum-P-Cycle における MTC の時間経過 (ダミー数 16，要求匿名領域 1600^2[m^2]，周期 20000[s])

る．これは，Dum-P-Cycle は匿名領域を確保するための停止地点がダミーの移動経路のなかに複数あり，シミュレーションの時間を通して何度か匿名領域を大きくできたためである．さらに，その停止地点はダミーの移動経路を作成する際に最初に決定することで，ダミーがその地点を必ず経由しなければならなくなるため，ダミー同士が過度に狭い範囲に集まることが少なくなり，さらなる匿名領域の確保につながったものと考えられる．

Dum-P-Cycle において，ダミー数 25 の場合は，全ての要求匿名領域で，ダミー数が 16 の場合よりも，AR-Count の値は大きくなり，その値の差の平均は，5.6%となる．これは，ダミー数が多くなると，ダミーを広範囲に分散でき，匿名領域が大きくなるからである．

Dum-P-Cycle の手法で作成されたダミーによって作られる匿名領域が時間経過によってどのように増減しているかを確認するため，AR-Count の時間変化を調べた．その結果を図 9 に示す．ここで，要求匿名領域は 1600^2[m^2]，基準地点を設定する周期は 20000[s] とした．図 9 より，時間の経過によって AR-Count が増減を繰り返していることがわかる．具体的には，図 9 より，周期直後 (0[s]～5000[s]，20001[s]～25000[s]，40001[s]～45000[s]) に，高い AR-Count を達成できている．これは，Dum-P-Cycle では，複数の基準地点を周期的に設定することに起因する．これにより，Dum-P-Cycle では，ユーザの要求する匿名領域がユーザの総移動時間を通して確保できており，3.4.2 項で言及した Dum-P の問題点が解決できたことがわかる．

4.2.2 MTC

ユーザの追跡可能性の評価を行うため，さまざまな要求匿名領域において，MTC を調べた．その結果を図 10 に示す．ここで，Dum-P-Cycle の基準地点を設定する周期は 20000[s] とした．すべての手法において，要求匿名領域が小さい場合，ユーザやダミー間の距離が近くなり，それぞれの移動可能な範囲内に位置することが多く，MTC は小さくなる．一方，要求匿名領域が大きくなると，ユーザやダミー間の距離が離れるため，MTC も大きくなる．

ダミー数 16 および 25 共に，Dum-P-Cycle と DumGrid を比較すると，Dum-P-Cycle は全ての要求匿名領域において，DumGrid よりも MTC の値が小さく，その差の平均は 1768[s] である．一方，Dum-P-Cycle と Dum-P を比べると，Dum-P-Cycle は Dum-P よりも MTC の値は大きくなってしまう．これは，Dum-P-Cycle は基準地点を複数決定する (Dum-P よりも増やす) ことで，ダミーの移動経路のなかに含まれる交差を発生させるための共有地点が，Dum-P よりも少なくなってしまうためである．ただし，MTC の値の差の平均は，Dum-P-Cycle と Dum-P で 448[s] となり，DumGrid からの減少分の方が大きい．

Dum-P-Cycle において，ダミー数が 16 の場合と，25 の場合を比較すると，全ての要求匿名領域で，ダミー数が 25 の場合の方が MTC が小さくなり，その値の差の平均は 549[s] となる．これは，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内に入る可能性が高くなり，さらに，ユーザと停止地点を共有するダミーが増え，ユーザの交差回数が増えるからである．

さらに，Dum-P-Cycle において，時間経過によって追跡可能性がどのように変化しているかを確認するため，MTC の時間変化を調べた．その結果を図 11 に示す．ここで，要求匿名領域は 1600^2[m^2]，基準地点を設定する周期は 20000[s] とした．図 11 より，時間の経過によって MTC が増減を繰り返していることがわかる．Dum-P-Cycle では，共有地点は基準地点を設定する周期以外の時間に決定されるため，周期の近くの時間には交差が発生しにくくなる．ここではダミーの基準地点を設定する周期を 20000[s] としていたため，周期直後 (0[s]～5000[s]，20001[s]～25000[s]，40001[s]～45000[s]) に，MTC は大きくなる．一方，それ以外の時間は MTC は小さくなる．これにより，Dum-P-Cycle では，MTC をユーザの総移動時間を通して低減できており，3.4.2 項で言及した，時間が経過しないと交差が発生しにくいという Dum-P の問題を解決できたことがわかる．

4.2.3 周期による影響

基準地点を設定する周期の変化による Dum-P-Cycle の AR-Count および MTC の変化を調べた．その結果を図 12 および図 13 に示す．ここで，要求匿名領域は 1600^2[m^2] と

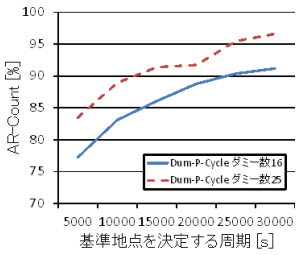

図 12 基準地点の設定周期と AR-Count (要求匿名領域 $1600^2[m^2]$)

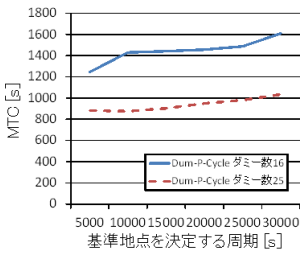

図 13 基準地点の設定周期と MTC (要求匿名領域 $1600^2[m^2]$)

した．ダミー数 16 および 25 共に，基準地点を設定する周期が短くなると，ダミーの移動経路に含まれる基準地点が増え，匿名領域確保のための移動をする機会が増える．しかし，図 12 より，周期が長くなるにしたがって，AR-Count は大きくなることがわかる．これは，Dum-P-Cycle は，ユーザが要求する匿名領域の大きさと等しいグリッド領域を作成し，その中でユーザおよびダミーの存在数が少ない領域 (セル) 内に基準地点を設定することに起因する．つまり，周期 5000[s] のように基準地点が過度に増えると，総移動時間を通してダミーがユーザの要求する匿名領域の中に収まってしまう傾向があると考えられる．その結果として，基準地点を多く設定したにも関わらず，要求匿名領域を満たすことが難しくなる．

また，ダミー数 16 および 25 共に，基準地点を設定する周期が短くなると，ダミーの移動経路に含まれる基準地点が増え，それに伴い，決定できる共有地点が減る．しかし，図 13 より，周期が長くなるにつれて，MTC は大きくなることがわかる．これは，上述のように，周期が短くなると，ユーザと全てのダミーで作られる匿名領域が小さくなることに起因する．周期が短くなると，ユーザとダミーおよびダミーとダミーの距離が近くなり，共有地点で交差を発生させる以上に，意図しない交差が発生するため，MTC が短くなったものと考えられる．

これらの結果より，周期が小さい場合は，匿名領域は確保しにくいが，追跡可能性は低下しやすく，周期が大きい場合は，匿名領域は確保しやすいが，追跡可能性は低下しにくいことがわかる．そのため，ユーザの匿名領域の確保に対する要求が大きい場合は周期を大きく設定し，追跡可能性の低下に対する要求が大きい場合は，周期を小さく設定することで，ユーザの位置プライバシに対する要求に細かく対応することができる．

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，ダミーの広範囲な分散を考慮したユーザ位置曖昧化手法 (Dum-P-Cycle) を提案した．Dum-P-Cycle では，ユーザの位置を移動時間を通して曖昧化できるように，あらかじめ予測されたユーザの行動を参考にして，停止しながら移動するダミーの行動スケジュールを生成する．この際，周期的にダミーをユーザや他のダミーの存在が少ない場所に移動させることで匿名領域を確保する．さらに周期以外の時間帯でユーザや他のダミーの停止地点を共有させることにより，交差を発生させ，追跡可能性を低下させる．

評価実験の結果，提案手法である Dum-P-Cycle は，先行研究 [9] における提案手法 (DumGrid) に比べ，匿名領域の確保と追跡可能性の低下に有効であることを確認した．また，先行研究 [4] における提案手法 (Dum-P) よりも，ダミーを広範囲に分散させ，総移動時間を通してユーザの位置を曖昧化できていることを確認した．

今後は，実際に人の目でユーザとダミーの動きを見た際に，どの程度ユーザを曖昧化できているか，視認性の評価を行う予定である．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 (CORE 9) 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] M. Duckham and L. Kulik: Simulation of Obfuscation and Negotiation for Location Privacy, *In Proc. CONSIT*, pp. 31–48, 2005.

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

[4] 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.

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

[6] 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.

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

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

[9] A. Suzuki, M. Iwata, Takahiro. H, X. Xie, and S. Nishio: A User Location Anonymization Method for Location based Services in a Real Environment, *In Proc. GIS*, pp. 398–401, 2010.

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