

スマートフォンの使用履歴に基づいた個人モデリング

OSHIKAWA, Eisuke / 押川, 英将

(発行年 / Year)

2013-03-24

(学位授与年月日 / Date of Granted)

2013-03-24

(学位名 / Degree Name)

修士(理学)

(学位授与機関 / Degree Grantor)

法政大学 (Hosei University)

2012 年度修士論文

スマートフォンの使用履歴に基づいた

個人モデリング

Individual Modeling

based on Records of Smartphone Uses

法政大学大学院 情報科学研究科 情報科学専攻

11T0012 押川英将 オシカワエイスケ

指導教員

法政大学大学院 情報科学研究科

西島 利尚

2013 年 2 月

概要

近年スマートフォンは広く普及しており、あらゆるところで利用されている。また、多くのスマートフォンアプリケーションが様々なサービスのために開発されている。他のコンテキストウェアアプリケーションと同様に、スマートフォンにおいても状況・環境により情報を推薦することはより良いユーザーインターフェースを設計するために重要なことである。ユーザーの習慣や場所などといったコンテキストを用いることによって、インターフェースの改良が可能になると考えられる。これを実現するためには、ユーザーの状況・環境とアプリケーション使用を関連付ける必要がある。そこで本研究は、スマートフォン使用履歴とそれを使用したときの時間、場所、移動状態を関連付けることによって、ユーザーのアプリケーション使用特徴の抽出を行うことを目的とした。本論文では取得したユーザーのアプリケーション使用時のログからの使用特徴の抽出、それをを用いた実験と評価について述べる。

Abstract

The smart phones are getting ubiquitous in these days. Meanwhile, a lot of smart phones based applications have been developed for various services. Like other context-aware applications, it becomes also an important issue to design better user interfaces suitable for special features and different demands. That is, a smart phone user interface may not be fixed, and can be automatically adjusted according to context information such as user's habit, preference, location, etc. It is thus necessary to provide a changeable interface layout associated with user's possible needs and current situations so that the user can use the smart phone more efficiently and conveniently. So, this research is to first record the use history of various applications used by a user, then extract relations between an application and its use time, location, movement state, action sequence, etc., and finally model the user characters using a smart phone. In this paper, application use log information acquisition and transformation are discussed, the application use features are analyzed, the experimental results are presented, and evaluations are given, respectively.

目次

| | |
|---------------------------------|----|
| 1. まえがき | 4 |
| 2. 関連研究 | 6 |
| 2.1. ヒューマンモデリング | 6 |
| 2.2. ユーザーのためのコンテキストを用いた利用支援 | 7 |
| 2.3. 携帯端末を用いたユーザーの行動判別 | 8 |
| 3. 研究の概要 | 9 |
| 4. ログの取得・変換 | 11 |
| 4.1. ログの取得 | 11 |
| 4.2. ログデータ形式の変換 | 13 |
| 5. アプリケーションの使用形態 | 14 |
| 5.1. アプリケーション使用特徴の可視化 | 14 |
| 5.2. アプリケーションのコンテキスト依存度による重み付け | 17 |
| 5.2.1. コンテキスト依存度による重み付けの概要 | 18 |
| 5.2.2. コンテキスト依存度による重み付けの計算 | 21 |
| 6. 使用特徴の計算と使用アプリケーションの推定 | 23 |
| 6.1. 特定のコンテキストにおけるアプリケーションの使用頻度 | 23 |
| 6.2. コンテキストの割合による使用アプリケーション推定 | 23 |
| 6.3. コンテキスト依存度による使用アプリケーション推定 | 24 |
| 7. 評価 | 26 |
| 7.1. 実験結果 ユーザー1 | 26 |
| 7.2. 実験結果 ユーザー2 | 27 |
| 8. まとめと今後の課題 | 31 |

1. まえがき

近年のスマートフォンの普及の拡大は著しい。スマートフォンとは携帯電話とパソコン・PDAの機能を組み合わせた多機能な携帯端末であり、iPhone や Android などに代表され、最近では技術の進歩によって端末の小型化やディスプレイの大型化が進んでいる。総務省発表の平成 24 年版情報通信白書によると、世界市場における携帯電話販売台数に占めるスマートフォンの割合は、2011 年（平成 23 年）は約 27% で、スマートフォンの利用割合は今後も拡大を続けて 2015 年（平成 27 年）には世界市場において 5 割を超える見通しとなっている[1]。また、スマートフォンはプリインストールされているアプリケーション以外にも、ユーザーが自由にアプリケーションをインストールして使用することができるのが特徴である。NTT DOCOMO では、「d マーケット」、 「d メニュー」といったサービス、KDDI では「au スマートパス」といったスマートフォンのアプリケーション定額サービスを通してコンテンツの拡大を図っている。有料アプリケーションのインストール数の平均は Apple Store で 5.4 個、Android Market (現 Google Play) では 1.8 個となっており、無料アプリケーションの合計インストール数の平均は Apple Store で 18.0 個、Android Market では 14.0 個となっている。こういった無料アプリケーションの開発を企業だけでなく個人でも自由に行うことができるのがスマートフォンの特徴でもあり、スマートフォンの普及の拡大とともにアプリケーションは今後も増えていくことが考えられる。しかし、スマートフォンは PC と比べてホーム画面に作成できるショートカットの数も少なく、ユーザーがアプリケーションを使用する際には、インストールされている多数のアプリケーションの中から選んで起動する必要がある。インストールされているアプリケーションが増えれば増えるほどに、端末の小型化やディスプレイの大型化による端末に搭載されているボタン数の減少も相まって、目的のアプリケーションをいち早く起動することは困難となっている。このようなスマートフォンの高機能化・多機能化に伴うユーザーに求められる操作の複雑化は、若者や IT に精通している人々にとっては操作過程が増加するとともに、お年寄りや IT が苦手な人々にとってはデジタルディバイドを広げる一方であり、これは無視できない問題であると考えられる。そしてこの問題を解決するためには、ユーザーインターフェースの改良やユーザーの特徴や趣味嗜好による使用アプリケーションの推薦を行うことにより、いち早く目的のアプリケーションの起動にたどり着くことが有効であると考えられる。

一方、ユビキタス・コンピューティング(Ubiquitous Computing)というコンセプトと、それを実現するための重要な要素としてコンテキストウェアネス(Context Awareness)という技術・概念

がある。ユビキタス・コンピューティングとは、現実世界のあらゆるものにコンピューターを組み込むことによって人の意識を介することなくコンピューター同士が協調動作する環境やコンセプトのことであり、コンテキストウェアネスとは、そのコンピューター同士の協調動作によって現実世界にあるモノの状況や状態を理解して捉える技術・概念のことである。ユビキタス・コンピューティングは1991年にアメリカのマーク・ワイザーによって提唱され [2][3]、その考え方から野村総合研究所が提唱したものがユビキタス・ネットワーク社会である。そして、ユビキタスネットワーク社会とコンテキストウェアネスを結びつけたものとしてコンテキストウェアコンピューティングがある。コンテキストウェアコンピューティングとはコンピューターが自動的に環境・状況の変化や環境・状況自体を認識してその結果を元に高精度なサービスを生み出すというものであり、人間においてもそれを活用することによってユーザーに対して環境や状況に応じての最適なコンピューターサービスの提供が可能となると考えられる[4]。

そして、ユビキタス・ネットワーク社会においてスマートフォンは重要な端末であると言える。スマートフォンにはGPS機能や加速度センサー・傾きセンサー・地磁気センサーなど各種センサーが搭載されており、ユーザーの「どこで」「いつ」「何を」したかという情報を、人間の日常生活上の情報であるライフログとしてユーザーの意識を介することなく常に記録することが可能になると考えられる。また、コンテキストウェアネスの考え方と同様に、ユーザーのスマートフォンアプリケーション使用もそのユーザーの環境や状況(コンテキスト)に依存していると考えられる。たとえばユーザーによっては、朝外に出かける前には天気予報のアプリケーションを使用することが多いであろうし、駅の近くではスマートフォンを使う場合は電車の路線案内・乗り換え案内のアプリケーションを使用することが多いであろうと考えられる。よって、特定のコンテキストで使用されているアプリケーションを推定することで、ユーザーインターフェースの改良やユーザーの特徴や趣味・嗜好による使用アプリケーションの推薦が可能とすることにつながると考えられる。本研究ではそのための、ユーザーのアプリケーション使用履歴とコンテキストに基づいた特徴の抽出と、ユーザーの使用アプリケーションの推定を行う。

本論文の構成は次のとおりである。第2章では、関連研究について述べる。第3章では、本研究での研究の概要について述べる。第4章では、ログの取得と特徴抽出可能な形式への変換について述べる。第5章では、アプリケーションの使用形態の可視化と使用特徴について述べる。第6章では、使用アプリケーションの推定について述べる。第7章では、使用アプリケーション推定の実験評価について述べる。第8章では、まとめと今後の課題について述べる。

2. 関連研究

2.1. ヒューマンモデリング

人間の特徴や行動や振る舞いをモデル化することはヒューマンモデリングと呼ばれている。ヒューマンモデリングを行うメリットとして次の三つが挙げられる[5]。一つ目のメリットは、ヒューマンモデリングを行うことによって、現実世界においてすでに行われた人間の行動を理解し説明できるようになることである。このような目的で用いられるモデルは記述的モデルと呼ばれる。記述的モデルは何かしらの事故が起こったときに、その事故の原因の解析手法として利用される。これは、モデリングを行うことで、起きてしまった事故の記録から原因となった人間の行動を突き止めることにより、再発防止のための予防策を作るのに有効である。二つ目のメリットは、ヒューマンモデリングは物事が起こった後における人間の行動のモデリングを行うことができるだけでなく、これから行われる人間の行動を事前に推定することができることである。これは予見的モデルと呼ばれ、マン・マシン・インターフェースの分野において利用されている。マン・マシン・インターフェースの設計において評価の段階では、被験者に対して条件を変えながらの複数回の実験が必要であり大きな費用や労力が必要となる。そのような場合でも、ヒューマンモデルを利用した代替シミュレーションを行うことによって、少ない費用や労力で設計評価が可能となる。三つ目のメリットは、ヒューマンモデリングによって人間の行動が合理的であるかどうか、正しいかどうかなど、意思決定のための示唆を行うために利用できることである。このような目的で用いられるモデルは規範的モデルと呼ばれる。これは人工知能の研究分野で、正しい結果を出すシステムの開発を目的とする場合に利用されている。

このようなヒューマンモデリングに関する研究は複数存在するが、人間の特徴や行動や振る舞いのモデリングといってもそれらの分類は多岐にわたり、目的に応じて人間のどの特性のモデリングを行うか、そしてどのような手法でモデリングを行うかさまざまである。麻生は[6]、ユーザーの物理的状態（ユーザーの位置、顔の向き、視線方向、声の大きさなど）、記号的状態（表層的な発話内容、顔、表情など）、意味的状态（意図、感情、ユーザーの要求などのユーザーの思考に関する状態）、人格的状态（性格、知識量、嗜好、社会的地位など）を、ユーザーの状態を表す変数として挙げ、音声を使用した場合のユーザーの位置と状態の推定についての研究を行っている。井上らは航空管制システムにおける航空管制官の認知モデルを構築す

る研究を行っている[7]. 航空管制システムは、その規模の大きさにもかかわらず業務における航空管制官の果たす役割が大きい。そのため、エラーの発生メカニズムの分析を行うことや、それらに対して対応できるシステムを開発することが重要となっており、航空管制官のグループ内・チーム内の状況認識や判断・意思決定の認知プロセスをモデリングすることにより、その問題の解決を図っている。本村らは、子供の事故防止予防のための研究においてヒューマンモデリングを行っている[8]. 過去に発生した事故事例データから子供の行動や事故の関連性を抽出して、事故の発生状況の再現やシミュレーションを可能とすることを目的としている。このように、ヒューマンモデリングは危険察知による事故防止や行動推定による生活支援などを可能とするため、人間の生活上において非常に重要な技術であると言える。

2.2. ユーザーのためのコンテキストを用いた利用支援

アプリケーション使用特徴とコンテキストを関連付けた特徴モデルを作成することで、ユーザーの状況・環境に適したサービスを提供することができ、ユーザーのアプリケーション使用を支援することが可能になると考えられる。しかし必要なスマートフォンアプリケーションは、複数のユーザーの環境や状況が同じであったとしても、ユーザー個人によって異なるものである。そこで、ユーザー個人の置かれている環境や状況に依存しての情報推定・推薦が必要であると考えられる。このような、コンテキストからユーザーの行動や嗜好を推定・推薦する研究は複数存在する。

奥らは、ユーザーの状況に応じて適切な情報を提供する状況依存型情報推薦システムの提案をしている[9][10]. これはユーザーのコンテンツに対する嗜好を「時刻、天気、予算、同伴者」といったパラメータを持った形式として表現し、多様に変化する人の嗜好をモデル化する研究である。この研究では「時刻、天気、予算、同伴者」といった特徴をコンテキストとして扱っている。このモデルを用いることによって、これらのコンテキストに応じた飲食店に関する情報の推薦が可能となる。松本らは、携帯端末使用履歴からの位置情報・時間情報のコンテキストに基づく最適アプリケーションの推定方法の提案をしている[11][12]. この研究では、アプリケーションの使用頻度による推定法と、コンテキストを利用した推定法の有効性の検証を行っている。また松本らは、コンテキストから使用される可能性があるアプリケーションの優先順位を導き、それを用いたインターフェースの設計と評価を行っている[13][14].

2.3. 携帯端末を用いたユーザー行動判別

GPS やスマートフォンに搭載されている各種センサーを利用した移動状態判別の研究も行われている。上坂らは加速度センサーや地磁気センサーを用いて、歩行者の歩数、歩幅、進行方向を推定し累積的に現在位置を求める手法を提案している[15]。小林らは、携帯電話に搭載されている加速度センサー、マイク、GPS を複合的に用いてユーザーの移動状態を推定する方法を提案している[16]。青木らは GPS から位置情報が取得できなかった場合でも、取得できなかった区間において歩数計で行動状態を判別する研究を行っている[17]。

このように、スマートフォンでユーザーの移動状態の推定が可能になることで、ユーザーの移動状態もコンテキストとして扱い、ユーザーのアプリケーション使用モデルの構築を行う際に、より詳細な特徴を抽出が可能となる事が考えられる。たとえばユーザーによっては、ナビゲーションアプリを使用するのは電車やバスに乗っているときより、歩いているときや車の助手席に乗っているときに使用する頻度が高い場合があると考えられる。

そこで本研究では、従来研究で多く用いられているコンテキストである「位置情報」と「時間情報」に加えて「移動情報」もコンテキストとして扱い、ユーザーのスマートフォン端末の使用履歴をアプリケーション使用時における「位置情報」「時間情報」「移動情報」の3つのコンテキストと関連付けることにより、ユーザー個人のスマートフォンアプリケーション使用特徴の抽出を行う。

3. 研究の概要

本研究は、ユーザーの使用アプリケーションを推定すること目的とする。そこで、アプリケーション使用履歴とコンテキストの関係性からの特徴抽出に焦点を当てた。ユーザーのスマートフォンアプリケーション使用時のコンテキストを取得し、コンテキストからの使用アプリケーション推定を行う。本研究ではアプリケーション使用時の「位置情報」「時間情報」「移動情報」の3つをユーザーのコンテキストとして扱う。今回は特徴抽出に焦点を当てるため、使用アプリケーションの情報および移動状態の情報の取得は、ユーザー自身で判断して与えることで実験を行う。また、位置情報と時間情報はスマートフォンで取得を行うが、位置情報に関しては、GPSが利用できない場所などで位置情報の取得が不可能な場合には、ユーザー自身で地図上にタグを打つことにより位置情報の取得を行う。そして、これらの取得した情報を元にアプリケーション使用特徴の抽出を行う。また、使用されるアプリケーションごとにもコンテキストとの関連性が存在すると考えられ、その関連性を考慮するためにアプリケーションごとにコンテキストへの依存度を算出する。それぞれのコンテキストにおいてのアプリケーションの使用頻度、アプリケーションに対するコンテキストの割合、およびアプリケーションのコンテキストへの依存度をユーザー個人のアプリケーション使用特徴として抽出する。その3つからアプリケーション使用時の「位置情報」「時間情報」「移動情報」のそれぞれにおいて使用されている可能性の点数付けを行う。そして、その点数を元に特定のコンテキストでのアプリケーション使用の順位付けを行うことで、そのコンテキストでのユーザーの使用するアプリケーションを推定する。今回は、従来研究で行われている「位置情報」「時間情報」からの特徴抽出と「位置情報」「時間情報」「移動情報」のからの特徴抽出の2パターン、それぞれでアプリケーションのコンテキストへの依存度を考慮した場合としない場合においてアプリケーション使用推定と検証を行った。図1に使用アプリケーション推定のイメージ図を示す。

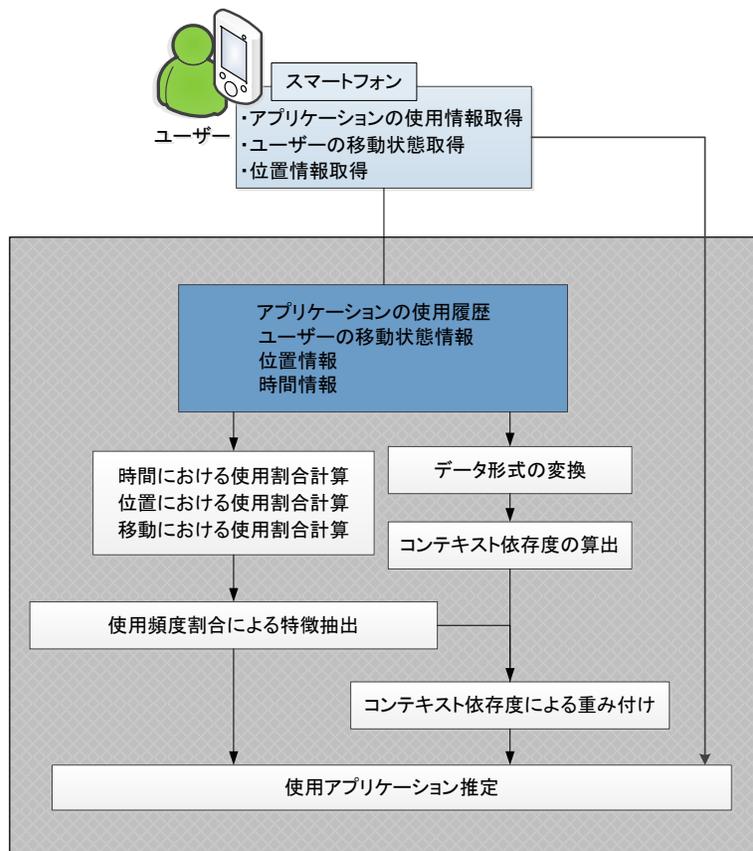


図1. 使用アプリケーション推定のイメージ図

4. ログの取得・変換

この章では本研究での実験の際のログの取得と変換について述べる。ログは、「使用アプリケーション名」「位置情報」「時間情報」「移動情報」の4つについて取得を行い、時間情報についてはアプリケーションのコンテキストに対する依存度の計算を可能とするデータ形式への変換を行う。

4.1. ログの取得

ユーザーがアプリケーションを起動するときに、「使用するアプリケーション名」「位置」「移動状態」をユーザー側から送信することでログの取得を行う。使用アプリケーション名および移動状態の確認はユーザー自身で行ってデータを与えることで取得を行う。本研究の実験においてログの送信は NHN Japan 株式会社が提供しているインスタントメッセージングサービス LINE を用いて行う。

また、実験に利用するアプリケーション使用時のコンテキスト「位置情報」「時間情報」「移動情報」の3つと、使用しているアプリケーションについて、それぞれ次のように取得と保存を行う。

①位置情報

スマートフォンに搭載されている GPS を用いて位置情報の取得を行う。GPS で取得する位置情報は住所情報および、数値情報として保存する。屋内などで GPS により位置情報が取得できない場合は、ユーザーが直接地図上から現在地にタグを打つことで位置情報の取得を行う。アプリケーション使用特徴のコンテキストとして位置情報を扱う場合には、緯度・経度といった数値での情報に加え、その緯度・経度の示す場所が何処であるかという情報も重要であると考えられる（例:ユーザーのアプリケーション使用位置が〇〇大学である場合、〇〇大学内であれば緯度経度が違っていても同じ位置のコンテキストとして考える）。そこで本研究では、アプリケーション使用時のコンテキストとしての位置情報には住所情報（都道府県名・市町村名・町名・丁目・番地）を利用し、使用アプリケーションのコンテキストへの依存度を計算する場合の位置情報には GPS で取得した緯度・経度の数値情報を利用する。

②時間情報

ユーザーからログが送信されてきた時間（時・分）をアプリケーションの使用起動時として保存し、アプリケーションの使用開始時をアプリケーション使用時のコンテキストとして使用する。また、1日を1時間ごとに区切って24個の時間帯に分類し、1日周期で同様の時間情報のコンテキストとして扱う。本研究では、アプリケーション使用時のコンテキストとしての時間情報にはこの1日を24個の時間帯に区切ったものを利用し、使用アプリケーションのコンテキストへの依存度を計算する場合の時間情報には周期性を持たせたものに変換したものを利用する。

③移動情報

移動情報はユーザーがアプリケーションを使用しているときの移動方法の種類を利用し、ユーザー自身が移動方法を判断し直接与えることで取得する。また、移動状態は静止、歩行、電車、車の4つに分類して保存する。なお、電車内・車内でユーザーが静止している場合もしくは歩行している場合であっても、静止状態、歩行状態として移動方法の判断を行うのではなく、それぞれの電車・車の移動状態であると判断する。

④アプリケーション情報

アプリケーションの名前を取得する。ただし、ログの送信時に使用するアプリケーションは使用アプリケーションとはせず、ログの送信は行わない。

4.2. ログデータ形式の変換

時間情報については、アプリケーションのコンテキストへの依存度を算出する場合に、その計算を可能とするデータ形式への変換を行う。アプリケーションのコンテキストへの依存度を算出する場合に利用する時間情報は、時・分であり、一日を一周とした周期的な情報として扱われ、また依存度の算出時には時間の範囲を考慮して計算を行う。そのため、24時間の始点と終点を一致していなければならない。しかし、取得した時間情報は単に数値としての情報であり、そのまま時間情報として扱う場合は、たとえば0:01と23:59が2分差の情報としてではなく、ほぼ24時間の差がある情報とみなしてしまう。時間情報に周期性を持たせるには0:00と24:00を始点と終点として同位のものとして扱う必要がある。そのために、取得した時間を2次元の連続値として扱い、始点と終点を一致させ周期性を持たせるための変換を行う。24時間を一周とした円としてみなし、円上にプロットされたアプリケーション使用の時間のx軸座標、y軸座標を計算することで変換を行い、その座標を2次元の時間情報とする。

5. アプリケーションの使用形態

アプリケーションはユーザーの行動の目的により利用され、例えばカメラのように非日常なものに遭遇した場合に利用するようなものもあれば、メールのようにいつでもどこでも利用されるものも存在する。よってアプリケーションには、使用時に特定のコンテキストに依存しているものと依存していないものがあると考えられる。そこで、アプリケーション使用特徴の可視化をそれぞれのコンテキストについて行い、コンテキストと使用アプリケーションの関連性について示す。また、アプリケーションのコンテキストへの依存度を考慮することにより使用特徴が現れることが考えられる。そこで、使用アプリケーション推定を行うための依存度の計算を行う。

5.1. アプリケーションの使用特徴の可視化

アプリケーションの使用特徴の検証を行うために、例として3つのアプリケーションについてのそれぞれのコンテキストにおけるアプリケーション使用時の可視化を行った。アプリケーションAはログ数が101個、アプリケーションBは14個、アプリケーションCは13個であった。図4は縦軸に緯度、横軸に経度を取り、アプリケーション使用位置をプロットした図である。アプリケーションAは様々な場所で使用されており、位置に依存せずに使用されるアプリケーションであることが分かる。逆に、アプリケーションB、Cは2箇所の特定の位置で使用されており、位置に依存して使用されるアプリケーションであることが分かる。

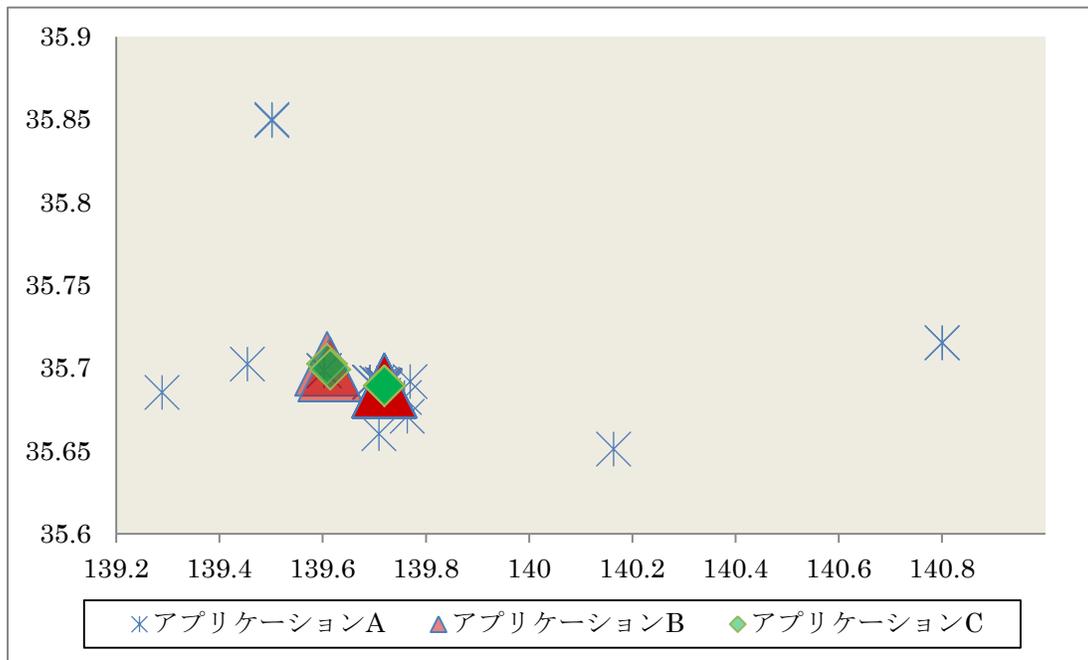


図4. アプリケーション使用位置

図5に、それぞれのアプリケーション使用時を時間軸にプロットした図を示す、アプリケーションAは起床時間中であればどの時間帯でも使用されている。アプリケーションBはまばらではあるが一日全体に分布が広がっており、どの時間帯にでも使用されている。アプリケーションCは昼から夕方にかけて使用されている場合が多い。このことから、アプリケーションA、Bに比べてアプリケーションCは時間に依存して使用されるアプリケーションであると言える。

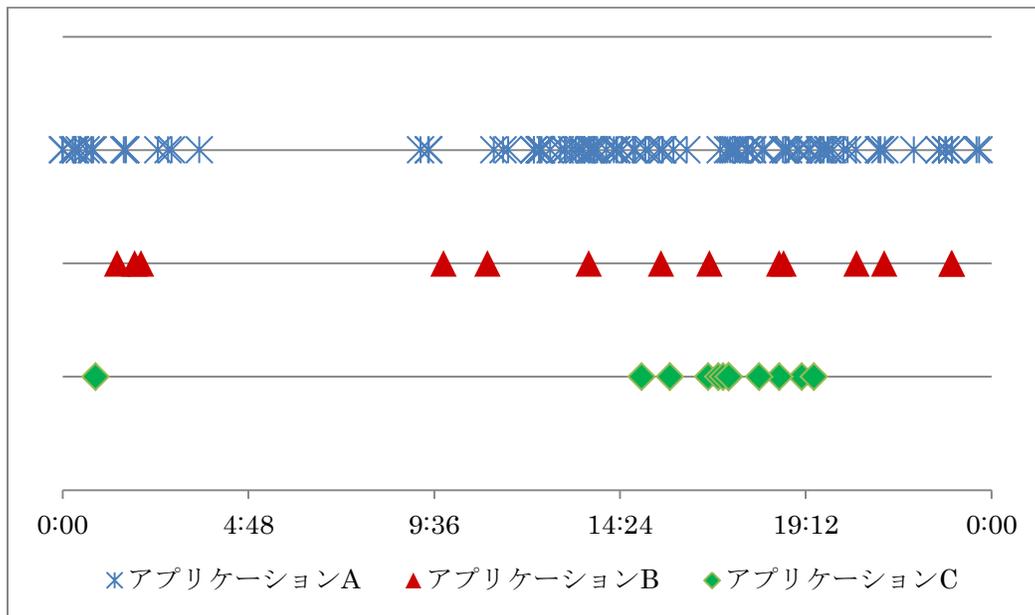


図 5. アプリケーション使用時間

図 6 にそれぞれの移動状態においてのアプリケーションの使用比率を示す。アプリケーション A はどの移動状態においても使用されているが、アプリケーション B、アプリケーション C は使用される移動状態に偏りが見られる。これらの図から、アプリケーション A はどのような状況でも使用される可能性が高く、コンテキストへの依存度が低いアプリケーションであると言える。アプリケーション B、C はコンテキストへの依存度が高いアプリケーションであると言える。

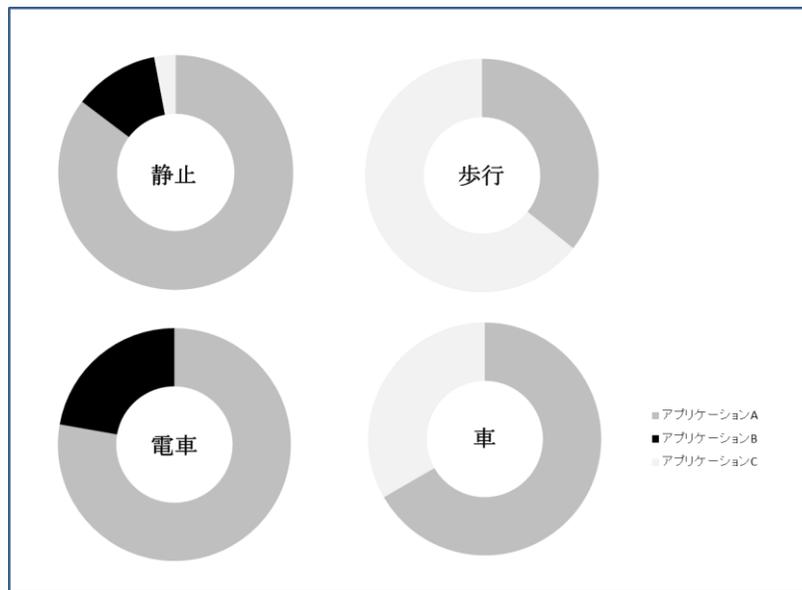


図 6. 移動状態におけるアプリケーション使用割合

これらの結果からそれぞれのアプリケーションでコンテキストに対する依存度としての使用特徴が現れることが確認できる。

5.2. アプリケーションのコンテキスト依存度による重み付け

5.1. の結果から、アプリケーションには使用時に特定のコンテキストに依存するものと依存しないものが存在することが分かる。また、アプリケーションがコンテキストの種類自体に依存していることも考えられる。コンテキストに依存しているアプリケーションはユーザーが使用するとき、その環境や状況からのアプリケーション使用を重要視している可能性が高く、使用アプリケーションを推定する際にはその依存度を考慮して推定することが有効だと言える。たとえば、全体での使用頻度が高く、かつコンテキストに依存しないアプリケーションはいつでもどこでも利用されているため、常にアプリケーション使用推定の順位で上位にランク付けされる可能性が高い。そのため、コンテキストに依存しておりそのコンテキストで使用される可能性が高いにもかかわらず、使用推定の順位で下位にランク付けされてしまうアプリケーションが存在する場合がある。そこで、ここでは推定するアプリケーションに対して、使用推定をする際にコンテキストの依存度によって重み付けを行う。特定のアプリケーションの使用点

が、全体の分布に比べて一つもしくは複数のエリアや時間帯に密集して存在する場合そのアプリケーションはそのコンテキストの種類に依存して利用されていると言える。そこで依存度を割り出すために、それぞれのコンテキストにおいてユーザーの使用した各アプリケーションの使用点の密集度の計算を行う。本研究ではアプリケーションの密集度をそれぞれのコンテキストに対する依存度とする。依存度 D は、特定のアプリケーション内での（各データ点において範囲半径 r 以内にデータ点 2 つの組み合わせが存在する数の合計） / （組み合わせの数の最大）を計算することで求める。

5.2.1. コンテキスト依存度による重み付けの概要

ここでは依存度 D の導き方の概要について位置情報の例を用いて説明する。ユーザーが特定のアプリケーションを使用した時の位置情報を①、②、③、④の 4 つの地点であるとする。

まずは位置情報の分布が図 7 の場合について説明する。図 7 は特定のアプリケーション使用時において地点①、②、③、④のそれぞれがお互いにすべて半径 r 以内に存在する場合である。各地点同士の組み合わせが 12 通り、各地点で半径 r 以内に他の点が存在する数が合計 12 通りであるため、この場合の依存度は $D=1$ となる。

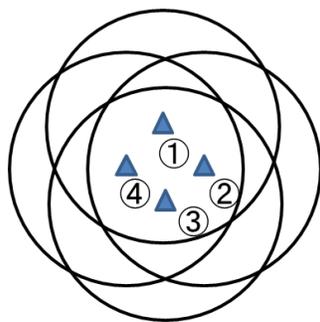


図 7 特定のアプリケーション使用時における位置情報の分布の例(1)

次に位置情報の分布が図 8 の場合について説明する。図 8 は特定のアプリケーション使用時において、地点①では半径 r 以内に他の地点が存在せず、地点②、③、④ではそれぞれがお互いにすべて半径 r 以内に存在する場合の位置情報の分布である。各地点同士の組み合わせが 12

通り，地点①では半径 r 以内に他の点が存在する数が 0 ，地点②，③，④で半径 r 以内に他の点が存在する数が 6 通りであるため合計 6 ，よってこの場合の依存度は $D = 6/12$ となる。

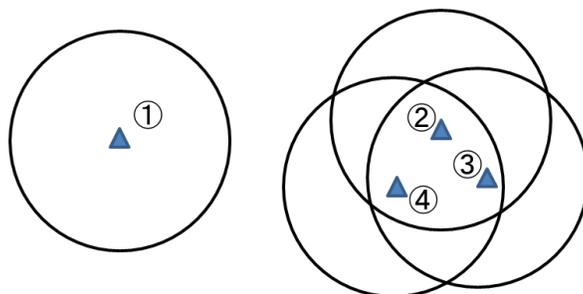


図 8 特定のアプリケーション使用時における位置情報の分布の例(2)

次に位置情報の分布が図 9 の場合について説明する．図 9 は特定のアプリケーション使用時において，地点①，②では半径 r 以内に他の地点が存在せず，地点③，④ではそれぞれがお互いにすべて半径 r 以内に存在する場合の位置情報の分布である．各地点同士の組み合わせが 12 通り，地点①，②では半径 r 以内に他の点が存在する数が 0 ，地点③，④で半径 r 以内に他の点が存在する数が 2 通りであるため合計 2 ，よってこの場合の依存度は $D = 2/12$ となる。

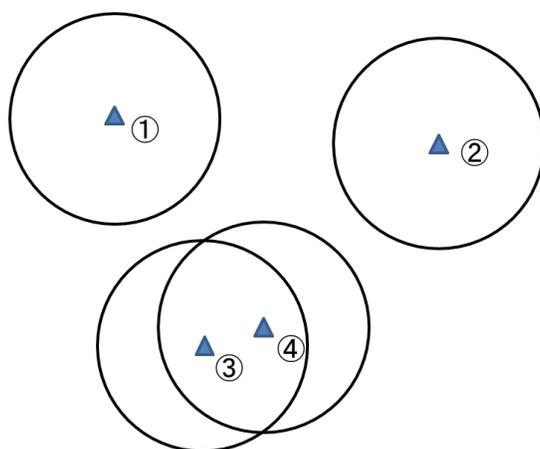


図 9 特定のアプリケーション使用時における位置情報の分布の例(3)

次に位置情報の分布が図 10 の場合について説明する．図 10 は特定のアプリケーション使用時において，地点①，②がそれぞれお互いに半径 r 以内に存在し，③，④がそれぞれお互いに半径 r 以内に存在する場合の位置情報の分布である．各地点同士の組み合わせが 12 通り，地点①，②，③，④すべてで半径 r 以内に他の点が存在する数が 4 通りであるため合計 4，よってこの場合の依存度は $D = 4/12$ となる．

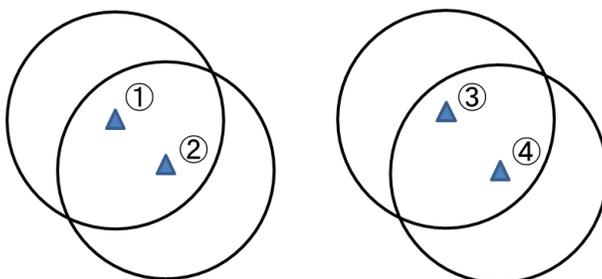


図 10 特定のアプリケーション使用時における位置情報の分布の例(4)

最後に位置情報の分布が図 11 の場合について説明する．図 11 は特定のアプリケーション使用時において，地点①，②，③，④すべてで半径 r 以内に点が存在しない場合の位置情報の分布である．各地点同士の組み合わせが存在しないので，この場合の依存度は $D = 0$ となる．

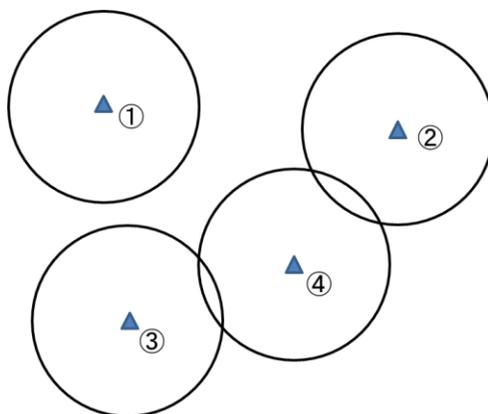


図 11 特定のアプリケーション使用時における位置情報の分布の例(5)

利用頻度割合のみの計算だと、地点③において例(3)と例(4)の場合が値となってしまう。しかし、例(3)の分布に比べて例(4)の分布は偏っているので、アプリケーション使用特徴を考える上で二つの場合を同じ値としてしまうのではなく、二つの分布を区別した特徴を抽出することが有効であるといえる。そこで使用アプリケーションのコンテキストに対する依存度を考慮することにより、この違いを特徴として抽出する。

5.2.2. コンテキスト依存度による重み付けの計算

特定のアプリケーションでの時間情報の場合の密集度 D_1 は、そのアプリケーション内での（各データ点において範囲半径 r 以内にデータ点 2 つの組み合わせが存在する数の合計） / （組み合わせの数の最大）を計算することで求める。まず、特定のアプリケーションすべての中で a 番目の特定のコンテキスト (x_a, y_a) に対しての、 (x_a, y_a) を中心とした範囲半径 r 以内に存在する特定のアプリケーションすべての中で n 番目のデータ点 (x_n, y_n) との組み合わせ

の数を計算する、そのデータ点の集合を X_a としたとき、

$$X_a = \{(a, n) | r > \sqrt{(x_a - x_n)^2 + (y_a - y_n)^2}\}$$

である。また、自身を除いた組み合わせを求めるので、特定のコンテキスト (x_a, y_a) においての他データ点 (x_n, y_n) との組み合わせの数は $|X_a| - 1$ 個となる。その組み合わせの計算を特定のアプリケーション内でのすべてのコンテキストで行い、それらを合算する。また、組み合わせの数の最大は $N(N-1)$ 個となる。すなわち密集度 D_1 は、

$$D_1 = \frac{\sum_{a=1}^N (|X_a| - 1)}{N(N-1)} \quad (N \geq 2)$$

で表される．ここで， x は時間情報の 2 次元への変換後 x 軸座標， y は時間情報の 2 次元への変換後の y 軸座標， n は 1 から N までの自然数， N はそれぞれのアプリケーションの使用数である．使用されたのが 1 回のアプリケーション ($N=1$) の場合は $D_1 = 1$ とする．位置情報の場合の密集度 D_2 も時間情報と同様に求める．位置情報の場合の (x, y) は (経度，緯度) である．

また，移動状態の場合はコンテキストとして利用する移動状態の種類が決まっているため，移動情報の場合の密集度 D_3' は， N を特定のアプリケーションの使用数としたとき，それぞれの移動状態で $S = |N - (\text{その移動状態の数})|$ を計算し，すべての移動状態分の S を合算することで求める．すなわち密集度 D_3' は，

$$D_3' = \frac{\sum_{i=1}^b S_i}{N(N-1)} \quad (N \geq 2)$$

で表される．ここで b は移動状態の種類数であり，今回は $b=4$ である．使用されたのが 1 回のアプリケーション ($N=1$) の場合は $D_3'=1$ とする．また，移動状態は種類数が限られているために，種類数より取得したログ数が多い場合は依存度の最低値が 0 にならないので， D_1 ， D_2 と同様に D_3' を 0 から 1 の間の値とするために，正規化を行い D_3 を求める．本研究の使用アプリケーションの推定では，以上で求められる密集度 D_1, D_2, D_3 をそれぞれ時間情報，位置情報，移動情報におけるアプリケーションのコンテキストに対する依存度として使用する．

6. 使用特徴の計算と使用アプリケーションの推定

この章では使用アプリケーションの推定方法について述べる。ユーザーのアプリケーション使用特徴を抽出するために、まず「位置情報」「時間情報」「行動情報」の3つのコンテキストにおいてのアプリケーション使用頻度を計算する。今回はアプリケーションの使用開始時をアプリケーション使用時のコンテキストとして使用するため、特定の時間情報、位置情報、または移動情報ごとのアプリケーション使用のログ数から使用頻度を求める。それをもとに特定のコンテキストでのアプリケーション使用の順位付けを行い、ユーザーの使用するアプリケーションの推定を行う。今回はアプリケーションの使用頻度に対して、コンテキストでの割合による重み付けと、コンテキスト依存度による重み付けをすることで、2パターンの使用推定を行う。

6.1. 特定のコンテキストにおけるアプリケーションの使用頻度

特定のコンテキストにおける特定のアプリケーションの使用頻度を求める。時間情報の場合の使用頻度 F_1 は、特定のコンテキストの時間 T においての特定のアプリケーションの使用割合である。すなわち、

$$F_1 = \frac{W_T}{N_T}$$

ここで、 W_T は特定の時間コンテキスト T においての特定のアプリケーション W の使用数、 N_T は特定の時間コンテキスト T においてのアプリケーション使用総数である。ただし、特定のコンテキストが存在しないときは $F_1=0$ とする。位置情報の場合の使用頻度 F_2 、移動情報の場合の使用頻度 F_3 も同様に求める。位置情報においての T は住所の情報、移動情報においての T は歩行、静止、電車、車といったそれぞれの移動状態である。

6.2. コンテキストの割合による使用アプリケーション推定

特定のコンテキスト内での特定のアプリケーションの使用頻度に対して、「一つのコンテキスト全体での特定のアプリケーション使用数」に対する「特定のコンテキストでの特定のアプリ

リケーション使用数」の割合で重み付けをすることによって使用アプリケーションの推定を行う。時間情報の場合のアプリケーションに対するコンテキストの割合 R_1 は、特定のアプリケーションにおける「時間情報のコンテキスト全体の数」に対する「特定の時間情報のコンテキストの数」の割合である。すなわち、

$$R_1 = \frac{T_W}{N_W}$$

ここで、 T_W は特定のアプリケーション W において時間情報の特定のコンテキスト T が存在する数、 N_W は特定のアプリケーション W の使用総数である。位置情報の場合の割合 R_2 、移動情報の場合の割合 R_3 も同様に求める。位置情報においての T は住所の情報、移動情報においての T は歩行、静止、電車、車といったそれぞれの移動状態である。

以上で求められる F_1, F_2, F_3 および R_1, R_2, R_3 を利用して使用アプリケーションの順位を推定するための点数 P_R を求める。点数 P_R は、それぞれのコンテキストにおいて使用頻度 F_1, F_2, F_3 に対し割合 R_1, R_2, R_3 による重み付けをして加算することで求める。すなわち、

$$P_R = \sum_{p=1}^3 F_p R_p$$

となる。そしてすべてのアプリケーションにおいて、それぞれの使用推定時のコンテキストで点数 P_R を計算する。点数 P_R が高い方から上位としてアプリケーション推定の順位を求める。

6.3. コンテキスト依存度による使用アプリケーション推定

$F_1, F_2, F_3, R_1, R_2, R_3, D_1, D_2, D_3$ を利用して、コンテキスト依存度を考慮した上での使用アプリケーションの順位を推定するための点数 を求める。点数 P_D はそれぞれのコンテキストにおいて、使用頻度割合に対しコンテキスト依存度による重み付けを行い、加算して求める。すなわち、

$$P_D = \sum_{p=1}^3 F_p R_p D_p$$

となる。そしてすべてのアプリケーションにおいて、それぞれの使用推定時のコンテキストで点数 P_D を計算する。点数 P_D が高い方から上位としてアプリケーション推定の順位を求める。

7. 評価

ユーザー2人対してそれぞれログの取得を行い、そのログを学習データとテストデータに分けて実験と評価を行った。なお今回は「位置情報」「時間情報」からの使用アプリケーション推定、「位置情報」「時間情報」「移動情報」のからの使用アプリケーション推定の2パターン、それぞれでアプリケーションのコンテキストへの依存度を考慮した場合としない場合において検証を行った。また、ユーザーのアプリケーション使用はコンテキストに依存するが、特定のコンテキストで単一のアプリケーションしか使用されないという可能性は低い。そこで本研究では、学習データから各コンテキストにおける各アプリケーションの使用可能性の点数表を作成し、それをテストデータにおけるコンテキストと照らし合わせて、特定のアプリケーション使用推定の順位のランク付けを行い、実際にユーザーが使用したアプリケーションが何位にランク付けされたかを判断することで評価を行った。順位の最高位は1であり最下位はユーザー個人が使用しているアプリケーションの総種類数である。

使用アプリケーションのコンテキストへの依存度を求める場合の半径 r であるが、今回の実験では位置情報の場合は $r = 0.0001$ として計算を行った。緯度・経度のメートル法における距離は各々の地点の緯度によって異なるが、東京付近では緯度での 0.0001 度はおよそ 11m 、経度での 0.0001 度はおよそ 9m である。また、時間情報では前後1時間を範囲内として使用アプリケーションのコンテキストへの依存度を求めた。よって $r = \sqrt{((1 - \cos(\pi/12))^2 + \sin(\pi/12)^2)}$ として計算を行った。ちなみに $r \approx 0.26105$ である。

7.1. 実験結果 ユーザー1

ユーザー1に対し6週間分のログの取得を行い、308個のアプリケーション使用データを得た。収集したログからランダムに約2/3の210個を抽出して学習データとし、残りの98個をテストデータとして使用アプリケーション推定結果の評価を行った。テストデータ中の使用アプリケーションは14種類、移動状態は静止、歩行、電車、車の4つであった。位置情報・時間情報をコンテキストとした場合の使用アプリケーション推定の順位の平均を表1に示す。

表 1. 位置情報・時間情報での使用アプリケーション推定順位の平均（ユーザー 1）

| | 依存度なし | 依存度あり |
|-----------|-------|-------|
| 専用ブラウザ | 7.00 | 6.25 |
| ソーシャルアプリA | 6.00 | 5.00 |
| エージェントアプリ | 9.00 | 9.00 |
| メッセージャー | 3.00 | 3.25 |
| ソーシャルアプリB | 11.00 | 10.00 |
| アラーム | 13.00 | 12.00 |
| スケジューラー | 12.50 | 6.00 |
| ブラウザ | 4.07 | 4.93 |
| メール | 1.59 | 1.76 |
| ニュース | 13.00 | 10.00 |
| ゲーム | 8.85 | 3.86 |
| 地図 | 7.67 | 7.00 |
| 電話 | 4.00 | 6.00 |
| 時刻表・路線案内 | 8.00 | 5.00 |
| 全アプリケーション | 4.28 | 3.85 |

コンテキスト依存度を考慮して推定することによって、全アプリケーションでの順位平均は依存度無しでの推定より良い結果を示すことができた。また、依存度なしの場合で下位にランク付けされたスケジューラーやゲームといったアプリケーションが、依存度を考慮することで順位を上げていることが分かる。これは、スケジューラーは自宅と職場付近で使用されることが多く、依存度無しでの推定の場合にはどの場所でも使用されていてかつ高い頻度で使用されている他のアプリケーションの推定順位が上位にランク付けされることにより下位になってしまったが、依存度を考慮することによって位置情報におけるスケジューラーの使用可能性の点数が増えたからであると考えられる。次に、表 2 に位置情報・時間情報・移動情報をコンテキストとした場合の使用アプリケーション推定の順位の平均を示す。

表 2. 位置情報・時間情報・移動情報での使用アプリケーション推定順位の平均（ユーザー 1）

| | 依存度無し | 依存度込み |
|-----------|-------|-------|
| 専用ブラウザ | 8.00 | 7.00 |
| ソーシャルアプリA | 6.00 | 5.00 |
| エージェントアプリ | 9.00 | 9.00 |
| メッセージャー | 2.54 | 2.62 |
| ソーシャルアプリB | 11.00 | 10.00 |
| アラーム | 13.00 | 12.00 |
| スケジューラー | 7.00 | 6.50 |
| ブラウザ | 3.14 | 3.21 |
| メール | 1.07 | 1.20 |
| ニュース | 10.00 | 11.00 |
| ゲーム | 5.85 | 5.71 |
| 地図 | 7.00 | 8.30 |
| 電話 | 4.00 | 4.00 |
| 時刻表・路線案内 | 9.50 | 9.50 |
| 全アプリケーション | 3.58 | 3.49 |

コンテキストへの依存度での重み付けを行った場合、全アプリケーションでの順位でわずかに向上が見られたが、それぞれのアプリケーションでみると大きな順位の変動は見られなかった。これは、関連情報を増やしすぎたためにアプリケーション推定がばらついたことが理由の一つであると考えられる。また、表 1 と表 2 を比較すると、移動情報もコンテキストとして扱い順位予測を行うことにより、コンテキスト依存度を考慮した場合・しない場合のどちらでも、位置・時間のみコンテキストでの推定より順位を向上させることができた。

7.2. 実験結果 ユーザー 2

ユーザー 2 に対し 6 週間分のログの取得を行い、184 個のアプリケーション使用データを得た。収集したログからランダムに約 2/3 の 123 個を抽出して学習データとし、残りの 61 個をテストデータとして使用アプリケーション推定結果の評価を行った。テストデータ中の使用アプリケーションは 8 種類で、移動状態は静止、歩行、電車 3 つであった。位置情報・時間情報をコンテキストとした場合の使用アプリケーション推定の順位の平均を表 3 に示す。

表 3. 位置情報・時間情報での使用アプリケーション推定順位の平均（ユーザー 2）

| | 依存度なし | 依存度込み |
|-----------|-------|-------|
| カメラ | 1.25 | 1.00 |
| ゲームA | 1.00 | 2.00 |
| ゲームB | 1.00 | 1.00 |
| 電話 | 3.50 | 3.50 |
| ブラウザ | 4.00 | 3.67 |
| メール | 1.06 | 1.17 |
| メモ帳 | 2.00 | 4.00 |
| 路線 | 2.50 | 2.50 |
| 全アプリケーション | 1.64 | 1.79 |

依存度を考慮して使用アプリケーション推定を行った場合の方が依存度を考慮しない場合より全アプリケーションで順位の平均が低い結果となった。これは全体での利用頻度が高かった「メール」が順位を下げたことが主な原因だと考えられる。次に、表 4 に位置情報・時間情報・移動情報をコンテキストとした場合の使用アプリケーション推定の順位の平均を示す。

表 4. 位置情報・時間情報・移動情報での使用アプリケーション推定順位の平均（ユーザー 2）

| | 依存度なし | 依存度込み |
|-----------|-------|-------|
| カメラ | 2.25 | 1.75 |
| ゲームA | 1.00 | 1.00 |
| ゲームB | 1.00 | 1.00 |
| 電話 | 3.00 | 2.25 |
| ブラウザ | 2.00 | 2.00 |
| メール | 1.00 | 1.00 |
| メモ帳 | 3.00 | 6.00 |
| 路線 | 3.50 | 3.00 |
| 全アプリケーション | 1.57 | 1.52 |

メモ帳の使用時のみ依存度を考慮した場合に大きく順位を下げてしまったが、他のアプリケーションにおいては依存度を考慮することにより、同順位もしくは順位の向上がみられた。また表 3 と表 4 を比較するとユーザー 1 と同様にユーザー 2 でも、移動情報もコンテキストとして

扱い順位予測を行うことにより、コンテキスト依存度を考慮した場合・しない場合のどちらでも、位置・時間のみコンテキストでの推定より順位を向上させることができた。

ユーザー1の場合は移動状態を含めることによって、さらに依存度を考慮することによって、アプリケーション推定順位平均の向上が見られた。しかしユーザー2の場合は、位置情報・時間情報のみでの使用アプリケーション推定では依存度を考慮することによって順位平均が下がってしまったが、位置情報・時間情報・移動情報での使用アプリケーション推定順位の平均は、依存度を考慮することによって向上することができた。これはただ単に情報量を増やせば精度が上がるわけではないということである。今後順位をさらに向上させるにはユーザー個人に合わせた最善のコンテキストの組み合わせ、利用するコンテキストの変換などが必要になると考えられる。

8. まとめと今後の課題

本研究では、位置情報・時間情報・移動情報とアプリケーション使用履歴を関連付けることにより、ユーザー個人の使用アプリケーション推定を行い、推定されたアプリケーション使用可能性順位を元に評価を行った。また、位置情報・時間情報・移動情報とアプリケーション使用履歴の関係だけでなく、アプリケーションのコンテキストへの依存度を考慮することで精度の向上を図った。ユーザー1, 2ともに、依存度を考慮してさらに位置情報を含んでの推定を行うことで、全アプリケーションでの順位の平均が上昇した。これによって、より使いやすいユーザーインターフェースの設計が可能になると考える。また、依存度を考慮して推定を行った場合には、元々使用頻度が高いアプリケーションは順位が下がる傾向があったが、それらはショートカットを作成するなどしてホーム画面に常駐させておき、コンテキストに依存するアプリケーションを推定し表示するなどの応用が考えられる。

今後の課題について以下に述べる。

・位置情報の分類

コンテキストとしての位置情報には住所、依存度の計算には緯度・経度を利用したが、住所と数値情報だけでなく位置ごとに属性分け(例えば「駅」「飲食店」など)をすることで、位置情報を意味がある情報として捉えることができると思う。また、本研究では位置情報をアプリケーション利用時に取得して点として扱ったが、常時取得してユーザーの移動の流れとしてみることで、新たな特徴を抽出できることが考えられる。

・ログの取得期間

本研究ではユーザー2人6週間のログの取得を行ったため、時間情報は1日を1周期として扱ったが、長期的なログの取得を行うことにより、曜日や月初め・月末といった情報も時間情報として扱うことが考えられる。

文献

- [1] 総務省, 平成 24 年版 情報通信白書, 「スマートフォン・エコノミー」～スマートフォン等の普及がもたらす ICT 産業構造・利用者行動の変化～, 2012.
- [2] M. Weiser, “The Computer for the Twenty-First Century”, Scientific America, pp94-104, September, 1991.
- [3] M. Weiser and J. S. Brown, “The Coming Age of Calm Technology”, P. J. Denning and R. M. Metcalfe, Eds., Beyond Calculation: The Next Fifty Years of Computing, 1996.
- [4] 飯田一朗, 森田俊彦, ”ヒューマンセントリックコンピューティングの全体像”, FUJITSU. 62, 5, p. 489-494, 2011.
- [5] 古田一雄 “ヒューマンモデリングの現状と課題”, 人工知能学会誌, Vol.13, No.3 356-363, 1998.
- [6] 麻生英樹, “パーティクルフィルタを用いたベイズ推論のユーザインターフェースへの応用”, 電子情報通信学会技術研究報告 信学技報, NC2005-51, 2005.
- [7] 井上諭, 青山久枝, 菅野太郎, 古田一雄, “航空路管制業務におけるチームのヒューマンモデリングに関する研究”, 電子情報通信学会技術研究報告 信学技報, SSS2007-23, 2007.
- [8] 本村陽一, 西田佳史, 山中龍宏, 北村光司, 金子彩, 柴田康徳, 溝口博, “子供の事故防止のための事故サーベイランシシステムへの取り組み～確率的知識の抽出と活用～”, 電子情報通信学会技術研究報告 信学技報, KBSE2005-23, 2006.
- [9] 奥健太, 中島伸介, 宮崎純他, ”状況依存型ユーザ嗜好モデリングに基づく Context-Aware 情報推薦システム”, 情報処理学会論文誌: データベース, Vol.48 No.SIG11(TOD34), 2007.
- [10] 奥健太, 中島伸介, 宮崎純他, 植村俊亮, 加藤博一, ”情報推薦におけるユーザの価値判断基準モデルに基づくコンテキスト依存型ランキング方式”, 情報処理学会論文誌: データベース, Vol.2, No.1 57-80, 2009.
- [11] 松本光弘, 清原良三, 沼尾正行, 栗原聡, “携帯電話におけるユーザ操作パターンを用いたアプリケーション推薦方式の提案”, マルチメディア, 分散, 協調とモバイル(DICOMO2009) シンポジウム In DICOMO2009, pp. 1149-1155, 2009.
- [12] 松本光弘, 清原良三, 沼尾正行, 栗原聡, “位置情報を含む携帯端末利用履歴からのコンテキストに基づく最適アプリケーション推定法の提案”, 情報処理学会研究報告, Vol.2010-MPS-80 No3, 2010.

- [13] 松本光弘, 清原良三, 沼尾正行, 栗原聡, ”携帯電話における時空間的利用履歴を基にしたアプリケーション推薦システム”, 情報処理学会 研究報告, 2008-MPS-72, 2008
- [14] 松本光弘, 清原良三, 福井秀徳, 沼尾正行, 栗原聡, ”ユーザのコンテキストから携帯電話の操作を予測するシステムの構築と評価”, The 22nd Annual Conference of the Japanese Society for Artificial Intelligence, 2008
- [15] 上坂大輔, 村松茂樹, 岩本健嗣, 横山浩之, ”手に保持されたセンサを用いた歩行者向けデッドレコニング手法の提案”, 情報処理学会論文誌, Vol.52 No.2 558-570, 2011.
- [16] 小林亜令, 岩本健嗣, 西山智, ”釈迦: 携帯電話も用いたユーザ移動状態推定・共有方式”, 情報処理学会 研究報告, 2008-MBL-45, 2008.
- [17] 青木政勝, 瀬古修一, 西野正彬, 山田智広, 武藤伸洋, 阿部匡伸, ”GPS 未計測区間における移動手段判定手法の検討”, 情報処理学会 研究報告, 2008-UBI-20, 2008.