条件付きユーザ参加型センサネットワークにおける効率的なユーザ選択手法の提案

石 塚 宏 紀[†1,†4] 岩 井 将 行[†1]
戸 辺 義 人[†3,†4] 瀬 崎 薫[†2]

携帯電話の発展に伴って，街を往来する人々の携帯電話からセンサデータを収集するユーザ参加型センサネットワークの研究が行われている．ユーザ参加型センサネットワークの性能は，センサネットワーク管理者の複雑な要求に応えうる参加者の選択に左右される．既存の参加者選択手法は，管理者が指定したデータ収集領域内で活動する候補者の行動を分析することで最終的に参加者を選出する．しかしながら，候補者の携帯電話を利用した継続的な行動分析は，端末の電池を著しく浪費するため現実的ではない．そこで我々は，候補者が日常的に利用しているソーシャルメディアによって行動分析を行い，さらい投稿内容から日常的な観測習慣を推測することで参加者を効率的に選択可能なプラットフォーム *KEIAN* を提案する．本論文は，*KEIAN* のシステム設計とそのプロトタイプ実装について詳しく述べる．

An Efficient Recruitment Platform for Participatory Sensing using Social Media

Hiroki Ishizuka,[†1,†4] Masayuki Iwai,[†1]
Yoshito Tobe[†3,†4] and Kaoru Sezaki[†2]

Participatory sensing enables citizens with a mobile phone as sensors to observe phenomenons in an city. Performance of such approach has depended on quality of each participant's contribution. Therefore, recruiting of well-suited participants who can contribute directly to collect sensed data according to purposes of organizers is a significant challenge. Previous works elect participants from candidates which were analyzed geographic behaviors. However, behavior analysis in a phone is not feasible in terms of its battery wasting. In this paper, we propose *KEIAN* that has two contirbutions: First, participant's habit has been extracted from social media. Second, the platform elects participants taking advantage of sensing habits of their posts to social media.

1. はじめに

近年，携帯電話は，通話やメールのためのコミュニケーション機器としてだけでなく，内蔵した各種センサによって環境情報を収集するセンサ機器としても注目されている．また，携帯電話によるインターネット接続が容易になり，利用者は，Web サービスやクラウドと連携したアプリケーションを常時利用可能となった．このような携帯電話の変革に伴って，街を往来する個々人が携帯電話を用いて環境をセンシングし，それらを収集することでより拡張性の高いセンサネットワークを構築するユーザ参加型センサネットワーク[1)2)3)] の研究が盛んに行われている．

ユーザ参加型センサネットワークにおける参加者はセンサネットワークのデータ収集要件（対象データ，対象領域，対象時間帯）における自身の都合に応じてセンサネットワークに貢献する．そのため，ユーザ参加型センサネットワークの性能は，センサネットワーク管理者の複雑な要求に応えうる参加者の選択に大きく左右される．既存の参加者選択手法[4)5)] は，管理者が指定したデータ収集領域内で活動する候補者の行動を分析することで参加者を選出し，その後の貢献度によって参加継続の可否を検討している．本論では，センサネットワーク参加後の参加者評価ではなく，あくまで候補者の中から参加者を選択する手法に焦点を絞って議論する．候補者の行動分析から，データ収集要件に適した参加者を選択するアプローチは有効であるが，候補者の携帯電話を利用した継続的な行動分析は，端末の電池を著しく浪費するため非現実的である．また，ユーザ参加型センサネットワークにおいて，候補者がセンサネットワークへの参加を懸念するようなを行為は避けなければならない．

一方，自身の状況を短いメッセージで表現し，ユーザ同士がコミュニケーションを行うマイクロブログサービスが注目されている．その中でも我々は 2 つのサービスに注目している．まず，Twitter[⋆1] は自身の位置情報や撮影した写真をメッセージに付加する機能を有しており，

---

†1 東京大学 生産技術研究所
Institute of Industrial Science, The University of Tokyo
†2 東京大学 空間情報科学研究センター
Center for Spatial Information Science, The University of Tokyo
†3 東京電機大学 未来科学部
School of Science and Technology for Future Life, Tokyo Denki University
†4 独立行政法人 科学技術振興機構 戦略的創造推進事業
JST, CREST

⋆1 http://www.twitter.com/

より幅広いコミュニケーションを気軽に楽しむことが可能である．次に，foursquare*2 は，携帯電話の GPS 機能を利用して，自身の現在位置を登録（チェックイン）し，友人や他人に現在位置を広く公開することで同じ空間を共有する人々とのコミュニケーションを図る．我々は，これらのサービスを利用して熱心に自身の状況や周辺環境を投稿しているユーザが，ユーザ参加型センサネットワークにおいて優良な参加者になりうると考えている．

そこで，我々は，候補者が日常的に利用している位置情報を含む投稿を利用して行動分析を行うことで，候補者の負担にならずにセンサネットワークのデータ収集要件への適性を判断する参加者選択プラットフォーム *KEIAN* を提案する．*KEIAN* における本研究への貢献を以下に述べる．

1) オープンなユーザ参加型センサネットワークのプラットフォーム設計

ソーシャルメディアを媒体として，ある条件下のセンサデータを要求する者と，その要求データを取得する可能性の高い候補者集合をマッチングするプラットフォームの設計．プラットフォームがセンサデータ要求者とそのセンサネットワーク参加者の匿名性を保証し，データのみを仲介することでセキュアな環境を提供．

2) 位置情報付き投稿からの行動分析

センサデータ収集要件に適した行動を行う候補者を選択するため，候補者がソーシャルメディアに対して日常的に投稿する位置情報を伴うメッセージを収集／解析し，隠れマルコフモデルを用いて候補者の行動モデルを構築．

3) 写真付き投稿からの観測習慣抽出

候補者がソーシャルメディアに対して写真を含む投稿を行う位置，時間，場所の属性などを分析し，カメラを用いたセンシングに対する観測習慣を抽出することで，センサデータ要求者の求める写真データの撮影可能性が高い候補者を選出．

我々はソーシャルメディアから約 1 年間に渡って収集した約 10 万件の位置情報付き投稿を利用して，任意の 5 人行動モデルを構築し，そのモデルについて簡易的な検証を行った．

本論は第 2 章にて関連研究を述べ，第 3 章にて *KEIAN* の設計について詳しく解説する．第 4 章にて実データを用いたモデル構築とその簡易検証について言及し，第 5 章にて全体に対するまとめを行う．

*2 https://foursquare.com/

2. 関 連 研 究

Sasanka 等[4] は，ユーザ参加型センサネットワークにおける参加者選択の重要性について言及し，参加者登用フレームワークを提案した．提案フレームワークは，3 つの工程で管理され，“絞り込み” 及び “査定” 過程にて，対象領域における時空間的なセンシングカバー率を最大化する候補者集合を選出し，最後に，選出された参加者の “再評価” を行って随時参加者を入れ替えることで，センサネットワーク全体の性能向上を図っている．Sasanka 等の手法は，行動情報の収集において候補者の携帯電話に導入された専用アプリケーションを利用しているため，候補者への負荷は大きいと考えられる．しかしながら，精度の高い行動情報を利用することで，マッチングの精度も向上すると考えられる．*KEIAN* は，参加者及びセンサデータ要求者とのコミュニケーション媒体にソーシャルメディア用いているため，オープンな環境として利用可能であるが，候補者の行動分析は，ソーシャルメディアに対する投稿のみに依存しているため，投稿内容の質によって精度が伴わない場合がある．しかしがなら，投稿が活発な候補者を優良な参加者として選択するため，自ずと行動分析の精度もある水準に保たれると推測される．

次に，センサを用いた行動分析に関する関連研究について考察する．GSM，GPS や WiFi を用いたユーザ位置トレースによる行動クラスタの特定に関する研究[6][7] は，クラスタの特定に GIS 情報を用いている．我々も GIS 情報としてユーザが投稿した forsquare のチェックイン箇所のカテゴリ情報を活用した行動分析を行っている．また，省電力でユーザの移動履歴を取得可能な SensLoc[8] は，ユーザの動静を消費電力が低い加速度センサで観測し，停止時に GPS 受信機の電源を落とし，移動時にはその移動速度に応じて GPS 情報の取得間隔を調整することで消費電力を低減している．しかしながら，ユーザが意図しない位置情報の取得を行っているため，このような専用アプリケーションの利用は，ユーザ参加型センサネットワークの候補者に懸念される可能性が高い．

最後に位置情報付き Tweet を用いた行動調査に関する研究[9] では，各 Tweet における位置情報によるクラスタリングを行い，その位置における属性情報を Tweet の内容から推測する手法を提案している．しかしながら，行動モデルの構築には至っていない．また，Tweet に付加された GPS による位置情報は都市部において誤差が大きいため，クラスタリングにもノイズが多く存在する．*KEIAN* は，foursquare の情報を用いているため，位置情報は正しく，さらに行動の履歴に基づくモデルも構築しているため，行動分析においては，優れていると言える．

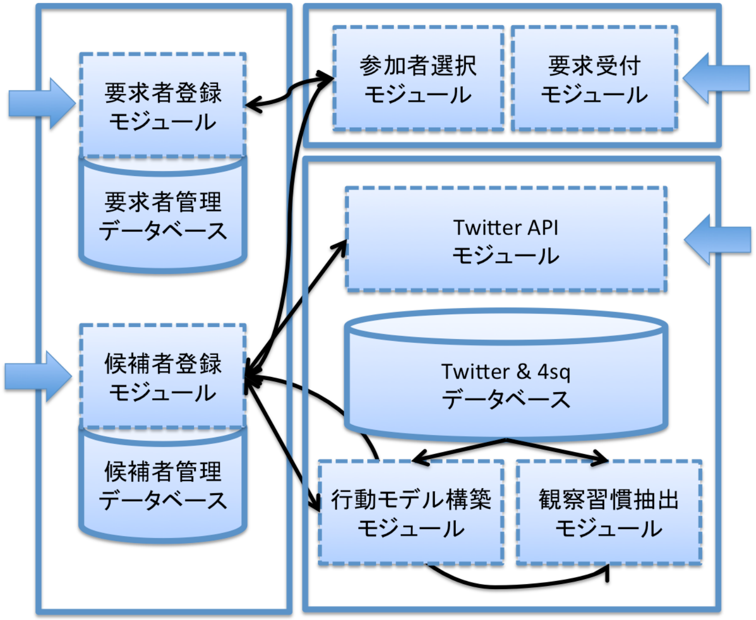


図 1 KEIAN のシステム構成
Fig. 1 A System Architecture of *KEIAN*

3. KEIAN の設計

本章では，ユーザ参加型センサネットワークにおける参加者選択プラットフォームである *KEIAN* の設計の詳細について述べる．まず，オープンなプラっとフォームとして設計された *KEIAN* の概要とそのシステム構成について解説し，その後，主要機能である隠れマルコフモデルを用いた位置情報付き投稿からの行動モデル構築手法について解説し，最後に，写真付き投稿からの観測習慣抽出手法について解説する．

3.1 システム構成

KEIAN のシステム構成図を**図 1** に示す．*KEIAN* は，Web サービスとして機能し，外部との通信はすべて HTTP にて行われる．また，候補者と要求者の情報は，内部で保持し，互いに匿名性を保持したままユーザ参加型センサネットワークの構築が可能な仕組みをとっている．システム構成は，大きくユーザ管理システム，参加者選択システム，ソーシャルメディア情報管理システムの 3 つのシステムから構成されている．それぞれの機能について以下で詳しく解説する．

3.1.1 ユーザ管理システム

KEIAN におけるユーザの定義は，ユーザ参加型センサネットワークの参加者となりうる**候補者**と *KEIAN* に対して条件付きセンサネットワークのための人材，及びセンサデータ自体を問い合わせる**要求者**の 2 種類である．本プラットフォームは，オープンなシステムを目指しているため，ユーザの登録は，ソーシャルメディアの ID を利用して行う．

候補者の情報は，登録した Twitter ID 及び，行動分析結果による行動モデルのパラメータ，観測習慣のパラメータが管理され，センサネットワークの構築要求があった場合，参加者選択システムによって候補者情報の走査が行われる．

要求者の情報も，Twitter ID が管理され，過去のセンサデータ要求履歴が管理されている．主に参加者選出状況 URL や要求結果 URL の送信に用いられる．

3.1.2 参加者選択システム

本システムは，参加者選択モジュールと要求受付モジュールから構成されている．*KEIAN* における要求 Q は，センシング対象領域 $(lat_1, lon_1, lat_2, lon_2)$，センシング対象時間帯 $(T_s, T_e, DayofWeek)$，センシング対象データ $TypeofData$，要求データ数もしくは期間 $\{Cout|Time\}$ で構成されている．要求受付モジュールは，Q を受け取り管理するとともに，参加者選択モジュールにも Q の情報を送信する．参加者選択モジュールは，Q の内容に基づき，候補者データベースより候補者選択を行う．

3.1.3 ソーシャルメディア情報管理システム

本システムは，Twitter API を用いて候補者登録時に取得した Twitter ID から候補者の投稿を収集し，データベースへの保存する．投稿が保存されたデータベースに対して，行動モデル構築モジュールは，常に位置情報付き投稿を監視し，行動モデルを学習する．また，観察習慣抽出モジュールは，写真付き投稿を監視し，観察習慣に反映させる．2 つもモジュールは，一定の間隔で構築したモデルのパラメータを候補者管理データベースへ登録／更新する．

3.2 位置情報付き投稿からの行動モデル構築手法

位置情報付き投稿からの行動モデル構築は，*KEIAN* における主要な機能である．我々は，ユーザの 1 日の行動は，訪れた場所の属性遷移によってモデル化可能であると考える．例えば，ある学生は，朝，コンビニ A(*SHOP* 属性) で飲み物を購入後，駅 B(*TRAEL* 属性) から電車に乗って登校する．大学の最寄り駅である駅 C(*TRAVEL* 属性) に到着し，大学

D($UNIVERSITY$ 属性) にて授業を受ける．お昼休みに，定食屋 E($FOOD$ 属性) で食事をとり放課後，アルバイト先 ($WORK$ 属性) に移動して働く．終業後，駅 B から電車に乗り，駅 A に到着後帰宅するといった属性遷移で 1 日の行動をモデル化できる．また，我々の日常生活は，立ち寄るコンビニや定食屋は変化すれど，ある属性遷移を習慣的に繰り返すこととなる．

3.2.1 隠れマルコフモデルを用いた行動モデル構築

我々は，隠れマルコフモデルを用いて場所の属性遷移を表現することで，候補者の行動モデルを構築する．本手法において，行動モデルの構築には，各候補者の foursquare のチェックイン情報を用いる．foursquare のチェックイン情報には，訪れた場所の名称 $Name$，位置情報 Loc，カテゴリ情報 Cat が含まれている．Cat を属性情報として利用し，$Name$ を観測データとして捉えることで隠れマルコフモデル M を構築する．M は次の 5 つの組 $M = (S, Y, A, B, \pi)$ で定義される．

S： 内部状態の集合，$S = \{Cat_i\}$ 本手法において，内部状態はチェックイン情報のカテゴリ情報を用いる．

Y： 出力シンボルの集合，$Y = \{Name_i\}$ 本手法において，離散で有限個の出力としてチェックインした場所の名称を用いる．

A： 状態遷移確率の集合，$A = \{a_{ij}\}$; a_{ij} は，状態 s_i から状態 sum_j への遷移確率を表し，$\sum_j a_{ij} = 1$ を満たす．

B： 出力確率の集合，$B = \{b_i(k)\}$; b_{ij} は，状態 s_i において出力シンボル k を出力する確率であり，$\sum_k b_i(k) = 1$ を満たす．

π： 初期状態確率の集合，$\pi = \{\pi_i\}$ は，初期状態が s_i である確率であり，$\sum_i \pi_i = 1$ を満たす．

M において，各属性の遷移には時間的な変化も考慮しなくてはならない．foursquare によるチャックインの間隔や頻度は当然日によって異なるため，事前に調査した不特定多数のチェックイン情報から数時間ごとの平均，及び分散値を算出したところ，3 時間ごと，計 6 個の M_i を生成することがバランスよく投稿が存在することが明らかとなった．

3.2.2 モデルを用いた行動推定

候補者がある時間 T において，複数の出力シンボルが存在する領域 A に存在する確率を求めることで，データ収集要件に適した候補者を選出することが可能となる．つまり，T の時間における行動モデル M_i を用いて，$T - M_i$ の観測開始時刻 T_{M_i} に至るステップ後に A 内に存在する出力シンボルの集合 Y' の和が解となる．

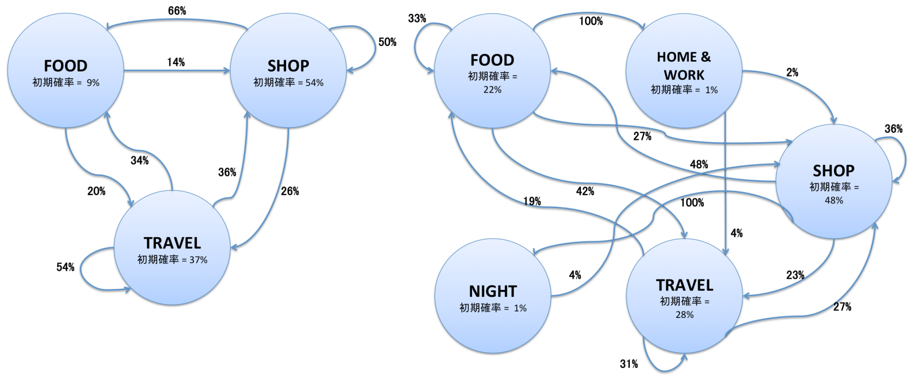


図 2 簡易実験における n=10(左), n=2(右) における行動モデル

Fig. 2 State transitions of a user ¡n=10(left) or n=2(right)¿

3.3 写真付き投稿からの観測習慣抽出手法

ユーザ参加型センサネットワークを利用した収集対象データは現実的にカメラで撮影した画像データが大半を占めると考えられる．そこで，我々は，候補者の写真付き投稿と行動分析結果を用いて，時間帯と位置属性に対する写真投稿頻度を調査した．調査方法は，写真付き投稿に位置情報が付加されている場合，投稿時間に，その場所に存在した場合の位置属性情報を行動モデルから推測し，その候補者が投稿された時間に写真を撮影する位置属性を記憶する．この情報を蓄積することで，候補者がある時間において，どの場所属性で写真を撮影する傾向があるかを判断できる．この写真撮影属性率を用いて参加者選択をすることでより，収集対象データを画像としたセンサネットワークに対して，優良な参加者を選択することが可能となる．

4. 簡 易 実 験

我々は，2010 年の 7 月から現在まで約 1 年間，Twitter 社が提供する API を用いて東京を中心とした 30km 圏内の位置情報付き Tweet 情報を収集してきた．その結果， 現在，約 12 万件の位置情報付き Tweet を保持している．その情報内において，foursquare を日常的に利用している (1 週間に 10 回のチェックイン) ユーザは，9715 人存在する．それら

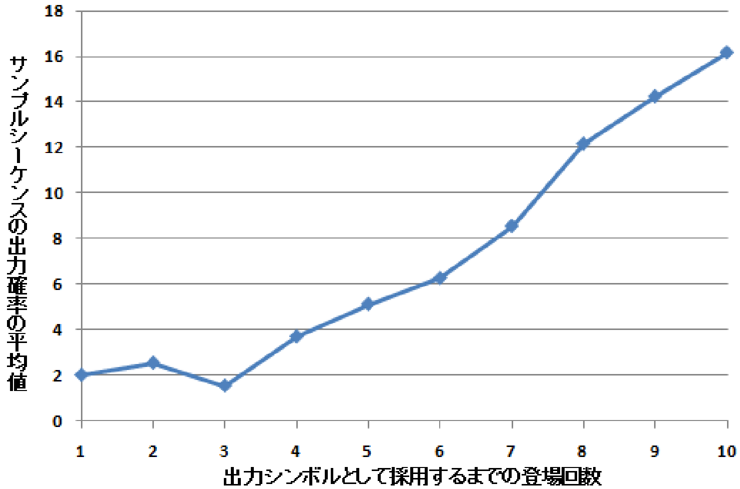
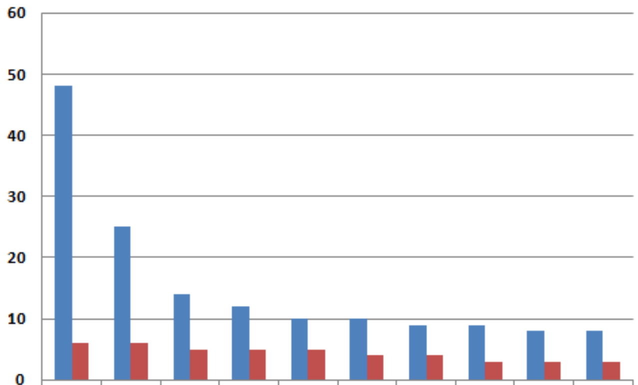


図 3 各出力シンボルが n 回登場する場合の HMM におけるシーケンス出力確率
Fig. 3 A Probability of HMM when each symbols appear at least n times.



	1	2	3	4	5	6	7	8	9	10
出力シンボル数	48	25	14	12	10	10	9	9	8	8
内部状態数	6	6	5	5	5	4	4	3	3	3

図 4 n の増加に伴う出力シンボル数と内部状態数
Fig. 4 The Number of symbols and states according to the increase of n.

のユーザ内からランダムに 10 人選出し，隠れマルコフモデルを用いた行動モデルの検証を行った．モデルの構築に 2010 年 7 月から 2010 年 12 月までの 5ヶ月間の情報を利用し，モデルの検証に 2011 年 5 月 15 日までの行動からランダムに 100 日分の行動シーケンスを用意し，モデルにおける行動シーケンスの出力確率を検証した．モデル生成において，出力シンボルの個数がそのまま精度に大きく関わる．さらにシンボル数は内部状態の数にも影響を及ぼす．そこで，簡易実験として，5ヶ月間で n 回登場した出力シンボルを採用してモデルを構築した際のシーケンス出力確率を調査した．**図 2** に n=2 と n=10 のときのあるユーザの行動モデルを示す．また，10 人の行動モデルにおける n とシーケンス出力確率の調査結果を**図 3** に示す．また，n の増加に伴う出力シンボル数と内部状態数の変化を**図 4** に示す．調査結果のグラフより，n の増加にともなってサンプルのシーケンス出力確率が増加していることがわかる．また，その要因として，出力シンボルや内部状態数の現状が見て取れる．このことから今後は，モデルを細分化し，状態数を絞ってより制度の高いモデルを構築する必要がある．

5. む す び

本論では，ユーザ参加型センサネットワークにおいて，センサネットワーク管理者の要求に応えうる参加者の選択が必要であることを述べ，効率的な参加者選択プロトタイプ *KEIAN* を提案した．*KEIAN* は，候補者の負担を軽減するために，候補者が日常的に行なっているソーシャルメディアへの投稿情報から候補者の分析を行い，さらに写真付き投稿による観察習慣の分析を行うことで，センサネットワーク管理者が求めるデータを効率的に取得する可能性の高い参加者を選択する．簡易実験により，出力シンボルと状態数の最適化を行い，より精度の高いモデルを構築する必要があることが明らかとなった．

謝 辞

位置情報付き Twitter 情報の収集において，多大なるご協力をいただいた酒巻智宏氏に心から感謝すると共に厚く御礼申し上げます．

参 考 文 献

1) Campbell, A., Eisenman, S., Lane, N., Miluzzo, E., Peterson, R.: People-centric urban sensing. In: Proceedings of WiCOM, IEEE (2006) 18-32
2) Burke, J., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., Srivastava, M.: Participatory sensing. In: Proceedings of WSW, ACM (2006) 1-5
3) 石塚宏紀，木實新一，戸辺義人，瀬崎薫，" 携帯端末で撮影した写真の利用を前提とする動的ストリート画像フロー生成機構の設計"，電子情報通信学会 論文誌 Vol J94-D No.1，pp.178-190
4) S. Reddy, D. Estrin, and M. Srivastava.: Recruitment Framework for Participatory Sensing Data Collections. In Proceedings of Pervasive, Springer (2010) 138―155
5) S. Reddy, D. Estrin, M.H. Mani, and M. Srivastava.: Examining Micro-Payments for Participatory Sensing Data Collections. In: Proceedings of Ubicomp, ACM (2010)
6) Ashbrook, D., Starner, T.: Using GPS to learn signicant locations and predict movement across users. Personal and Ubiquitous Computing (2003) 275-286
7) Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L.: Discovering personal gazetteers: an interactive clustering approach. In: Proceedings of GIS, ACM(2004) 266-273
8) D. H. Kim, Y. Kim, D. Estrin, M. Srivastava.: SensLoc: sensing everyday places and paths using less energy. In: Proceedings of Sensys, ACM (2010) 43-56
9) 酒巻智宏，岩井将行，瀬崎薫，“マイクロブログのジオタグを用いたユーザの行動パターンの推定に関する研究”，電子情報通信学会 信学技報 Vol. 110 No.400 (2010) 37-42