

第 XIV 部

無線を用いた位置情報 プラットフォームの構築

第 14 部

無線を用いた位置情報プラットフォームの構築

第 1 章 はじめに

WiL WG は、無線 LAN 等を用いたワイヤレス位置情報プラットフォームの研究開発を進めることを目的とした WG である。2009 年度の WiL WG の活動としては、ポータルサイト Locky.jp における無線 LAN 基地局のデータベース収集、無線 LAN による位置推定応用の iPhone/iPod Touch アプリである「駅.Locky」、および Web サービスサイト <http://eki.locky.jp> の公開、Gaussian Mixture Model を用いた無線 LAN 位置推定に関する研究を行った。

第 2 章 無線 LAN 位置情報ポータルサイト Locky.jp の現状

Locky.jp では、2005 年のプロジェクト開始以来、これまでにボランティアの支援によって無線 LAN 位置推定のための無線 LAN 基地局の収集を進めてきた。2009 年 1 月時点で収集された無線 LAN 基地局数は 53 万であったのに対し 2010 年 1 月時点では 72 万を超えており、継続的にデータが収集できていることが確認できる。また、データ収集に参加しているボランティアユーザは 700 名を超えている。なお、2009 年度には、Locky.jp のデータを利用して位置推定が可能な iPhone/iPod Touch アプリケーション「iSpotter」が登場しており、これを用いて、データの利用・収集をしているユーザが多く含まれていることが予想される。

Locky.jp のサービスは 2005 年の稼働以来、サーバを更新してこなかったが、データ量の増加とハードウェアの老朽化が進んでおり、2009 年度にはサーバの停止や障害が数回発生した。2010 年度には更新の検討が必要である。このようなサービスを長期間継続的に行うためには、適切な更新計画が必要である。

第 3 章 無線 LAN 位置推定を用いた時刻表アプリ「駅.Locky」

「駅.Locky」は無線 LAN 位置推定を用いた iPhone/iPod touch 用時刻表アプリであり、<http://eki.locky.jp> はそれを支えるための時刻表収集・サポート Web サービスである。駅.Locky では、我々がアプリと Web サービスを提供し、アプリの利用に必要な時刻表情報と無線 LAN 情報はユーザに提供してもらう仕組みを実現している。本アプリでは、主に以下の機能が利用できる。(図 3.1)

- 無線 LAN 位置推定による最寄駅の検出
- 選択した駅の次の電車の発車までの時間のカウントダウン表示
- 時刻表一覧の表

ユーザは自分が必要な時刻表を特定のテキストフォーマットで記述し、駅.Locky の Web サイト [107] からアップロードを行う。アップロードされた時刻表情報は、全てのユーザがアプリからダウンロードして利用することが可能である。ユーザは任意の駅に対して時刻表情報を追加できるが、故意の悪戯等による情報の改ざんを防ぐため、他のユーザが追加した時刻表情報を編集及び削除することはできない。

駅.Locky では、正確な時刻表データベースを維持するために、有志のユーザ（サポータ）に対して編集及び削除の権限を与えている。時刻表の誤りについては、発見したユーザが Web サイト内の掲示板にて報告し、提供者本人またはサポータに自主的に修正対応を行ってもらう。

本アプリには無線 LAN 情報を学習する機能が実装されており、ユーザはアプリ上で選択している駅と、自分が今いる位置で取得している無線 LAN 情報 (BSSID、SSID) とを対応づけて学習させることにより、以後その無線 LAN 情報を取得した際に、対応づけられた駅を最寄駅として表示させることができるようになる。学習を行う際、ユーザは学習情報を公開するかしないかを選択することができ、公開

● 第 14 部 無線を用いた位置情報プラットフォームの構築



図 3.1. 駅 Locky のスクリーンショット

用として学習された無線 LAN 情報は、その後ユーザの任意のタイミングでサーバにアップロードすることで、他のユーザがダウンロードして最寄駅の検出に利用することが可能となる。

2009 年 10 月 25 日のサービス公開後約 2 ヶ月が経過した時点で、本アプリは 11 万件以上ダウンロードされ、12730 件の時刻表と、52273 件の無線 LAN 基地局情報がアップロードされている。国内全駅に対する時刻表のカバー率は 60%を超えており、首都圏では 80%近くに達している。また、時刻表アップロードユーザ数は 2472 名、無線 LAN のアップロードユーザ数は 4201 名である。

駅 Locky の公開後、相模鉄道相鉄本線や東京メトロ丸の内線で路線全体のダイヤ改正が行われたが、これについても掲示板にて報告が行われ、サポータにより全て修正されており、ユーザとの協力によりサービスが維持できている。

第 4 章 Gaussian Mixture Model を用いた無線 LAN 電波情報のモデル化

無線 LAN による位置推定を行うためには、アクセスポイントの位置や受信電波強度の分布など、各システムに応じた事前収集データが必要となる。特に、屋内においては高い精度を得るために、そのデータ量は大規模なものになってしまう。大規模なデータは、端末にも負担がかかり実用的ではないため、軽量化する必要がある。

我々は、Gaussian Mixture Model(以下 GMM と

呼ぶ)を用いて事前収集データをモデル化し軽量化する手法を提案する。GMM とは、複数の Gauss 分布を線形結合させて重ね合わせたものである。十分な数の Gauss 分布を用い、また線形結合する重みの係数と共に各分布の平均と共分散を調整すれば、ほぼ任意の連続な密度関数を任意の精度で近似できる。2次元の Gauss 分布を用いることにより、無線 LAN の電波環境を近似することができる。

$$\mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{2\pi}|\boldsymbol{\Sigma}|} \exp\{-(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\} \quad (1)$$

式 (1) は Gauss 分布を表している。この Gauss 分布 K 個を重ね合わせた GMM は以下の式 (2) で表される。

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (2)$$

$$\sum_{k=1}^K \pi_k = 1 \quad (3)$$

ここで、個々の Gauss 分布は混合要素といい、平均 $\boldsymbol{\mu}_k$ と分散共分散行列 $\boldsymbol{\Sigma}_k$ を持つ。また、 π_k は混合係数と呼ばれ、各混合要素の重みを表す。式 (3) のように、全混合係数の総和は 1 となる。GMM の形状は、個々の平均、分散共分散行列、混合係数により決まる。

4.1 無線 LAN 情報の変換

モデル化を行うために、観測した無線 LAN 情報を GMM にあてはめる。しかし無線 LAN 情報は、観測した位置の受信電波強度を記録した点の集合になっているため、点の密度と電波強度の間に関連性はな

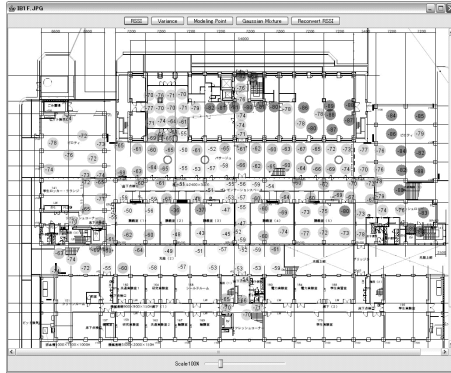


図 4.1. 変換前の無線 LAN 情報

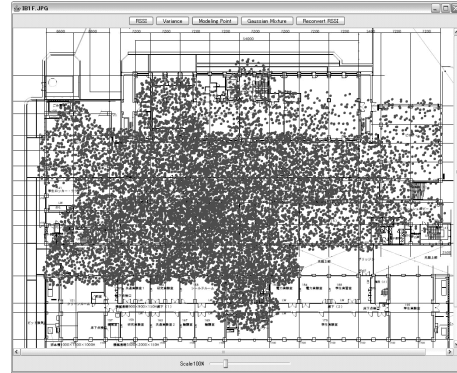


図 4.2. 変換後の観測点情報

い。GMM は密度関数であるため、このままでは正しく電波環境をモデル化することはできない。そこで、無線 LAN 情報の点の集合を各観測位置の電波強度に基づく密度を持った点の集合に変換する必要がある。また、電波の伝播特性は基準点（アクセスポイントの位置）で不連続なのに対し、GMM は連続な関数である。ゆえに、GMM を用いることで完全に電波環境を表すことはできない。しかし、これは極めてアクセスポイントに近い位置での影響なので、位置推定を行う上ではほとんど障害にはならないと考えられる。

以上のことを考慮し、一定の間隔で電波強度の値が記録された点を、伝播特性に合った密度になるように変換する手法を提案する。まず、それぞれの観測点に記録されている電波強度の値により、その周辺に点を増加させる。このためには、それぞれの観測点ごとの支配領域を設定し、その領域の中で点を増加させる必要がある。このことから、それぞれの観測点を持つ領域同士が重ならず、かつ空白ができないように領域の設定をする必要があると考えた。そこで、ボロノイ分割を利用して点の変換を行う。まず、ボロノイ分割により各観測点を持つ領域を設定し、その後で受信電波強度の値に応じて点を増加させた。ただし、各観測点を持つ領域の面積も考慮する必要がある。観測した電波強度の値が高く、領域の面積が広いほど、多くの点を増加させた。

ある BSSID について以下のように観測されたとする。

$$O = \{o_1, \dots, o_N\} \quad (4)$$

$$o_n = (x_n, y_n, \alpha_n) \quad (5)$$

ここで、 N は観測点の数、 x_n 、 y_n は観測を行っ

た座標、 α_n は受信電波強度である。これらの観測点にボロノイ分割 [128] を適用して、それぞれの観測点を持つ領域 V_n を決定する。ただし、基準となる観測点からの距離が、あらかじめ設定した値より大きくなった場合、領域内には含めないようにした。

$$V_n = \{v_1, \dots, v_{M_n}\} \quad (6)$$

$$(n = 1, \dots, N) \quad (7)$$

ここで、 $v_1 \sim v_{M_n}$ は式 (8) のように、その領域に含まれる座標であり、 M_n がその領域 V_n の大きさとなる。

$$v_m = (x_m, y_m) \quad (8)$$

$$(m = 1, \dots, M_n) \quad (9)$$

この領域の中から受信電波強度 α_n と領域の大きさ M_n により、ランダムに v_m を選択する。選択する数 S_n は以下の式 (10) に基づく。

$$S_n = (\alpha_n + C) \frac{M_n}{\text{Min}_{i=0}^N M_i} \quad (10)$$

式 (10) における C は電波強度を正数に換算する定数である。また Min は、全観測点の中で最も小さいボロノイ領域の大きさを返す関数とする。以下にボロノイ分割を利用して観測点を増加させたものを示す。図 4.1 が事前収集データで、図 4.2 が増加させた結果である。

4.2 EM アルゴリズム

変換された観測点情報を GMM にあてはめる際に、最尤解を求めるための強力な手法として、EM アルゴリズムと呼ばれるものがある。変換された観測点情報を $\{x_1, \dots, x_N\}$ とすると、このデータ集合は、第 n 行を x_n^T とする $N \times D$ 行列 X となり、式 (2)

より、対数尤度関数は以下のように表される。

$$\begin{aligned} \ln p(\mathbf{X} | \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ = \sum_{n=1}^N \ln \left\{ \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right\} \end{aligned} \quad (11)$$

EM アルゴリズムでは、GMM の各パラメータ（平均、分散共分散行列、混合係数）について、この対数尤度関数を最大にすることが目的である。EM アルゴリズムは、E ステップと M ステップと呼ばれる 2 つの更新手続きが繰り返し行われる。以下に無線 LAN データのモデル化で利用する、EM アルゴリズムの繰り返し過程をまとめる。

1. 初期化
2. E ステップ
3. M ステップ
4. 式 (11) の対数尤度関数を計算し、収束基準を満たしていなければ、2 に戻る。

4.2.1 初期化

最初に、個々の Gauss 分布の平均、分散共分散行列、混合係数を初期化する。収束基準を満たすまでの繰り返し回数を減少させるために、初期化には K-means アルゴリズムを用いる。K-means アルゴリズムとは、多次元空間上の点集合について、各データが属するグループまたはクラスターを同定するための手法である。K-means アルゴリズムにより見つけたクラスターのサンプル平均とサンプル共分散を、各 Gauss 分布の平均と分散共分散行列として用いる。また、混合係数には各クラスターに属する点の割合を用いる。K-means アルゴリズムを用いて初期化を行うことにより、ランダムに初期化を行った場合と比べて EM アルゴリズムの繰り返し回数を減少させることができる。

4.2.2 E ステップ (expectation step)

E ステップとは、その時点での各混合要素の平均、分散共分散行列、混合係数を用いて負担率の計算を行うことである。負担率とは、混合要素 k が収集データ $\mathbf{x}_n$ を負担する度合いを示したものであり、以下の式 (12) で表される。

$$\gamma(z_{nk}) = \frac{\pi_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{i=1}^K \pi_i \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)} \quad (12)$$

4.2.3 M ステップ (maximization step)

M ステップでは、式 (12) で計算した負担率を用いて個々の混合要素の平均、分散共分散行列、混合係数を更新する。更新式を以下に示す。

$$N_k = \sum_{n=1}^N \gamma(z_{nk}) \quad (13)$$

$$\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) \mathbf{x}_n \quad (14)$$

$$\boldsymbol{\Sigma}_k = \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) (\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^T \quad (15)$$

$$\pi_k = \frac{N_k}{N} \quad (16)$$

この計算を行うとき、最初に式 (14) で新しい平均値を計算し、その値を式 (15) の分散共分散行列の更新で用いる。

4.2.4 対数尤度関数の計算

E ステップと M ステップのパラメータ更新手続きは、いつも式 (11) の対数尤度関数を増加させる。対数尤度関数または各パラメータの変化量が、あるしきい値より小さくなったときに EM アルゴリズムが収束したと判定する。この収束条件を満たしていなければ、再び E ステップと M ステップを行う。本研究では、対数尤度関数の変化量が 10% 以下になったときを収束条件とした。

実際に無線 LAN の事前収集データにボロノイ分割と EM アルゴリズムを適用した様子を図 4.3 に示す。ここでは、混合分布数は 8 としている。GMM の混合数をうまく調整することで、図のように通路や部屋に沿って電波強度の値を高精度で表すことができる。このように、GMM を用いることにより、無線 LAN の電波環境を一つの基地局に対して一つの式で近似することができる。

4.3 パーティクルフィルタと GMM を用いた位置推定

無線 LAN のデータを Scene Analysis 方式を用いて位置推定を行う主な方式としてベイズ推定を利用する手法がある。しかし、ベイズ推定を用いる場合、全ての候補位置での尤度を計算し、最も尤度の高い位置を推定結果とするため計算量が非常に多くなる。位置推定サービスは、ユーザに対して現在位置をリアルタイムに提示するものであるため、計算時間がかかりすぎることは望ましくない。また、2 次元の

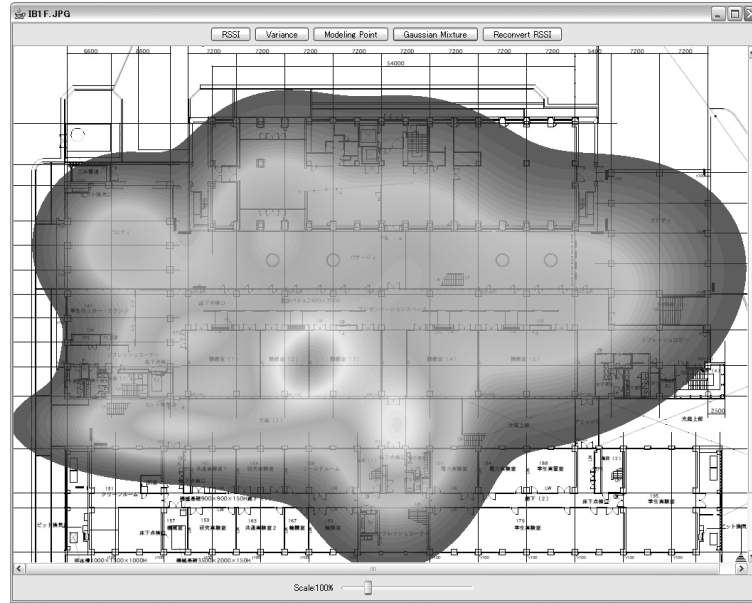


図 4.3. モデル化された無線 LAN データ (混合分布数: 8)

GMM は、確率密度から座標を解析的に導くことができない。そこで本研究では、任意の確率密度関数を扱うことができ、計算量の調節も容易な Particle Filter[42] を位置推定アルゴリズムとして用いている。Particle Filter は複数の離散的な Particle を用いる。Particle は、その位置情報と重みから構成されており、移動と、観測データに基づく重みの更新を行った後、その重みによるリサンプリング (Particle の再構築) により、重みの小さいものは破棄され、尤もらしいものだけが残っていく。提案手法では、この手続きを繰り返し行い、重みの大きな Particle が集まっている位置を推定結果とする。Particle の数を操作することにより、計算時間の調整が可能である。以下でアルゴリズムの詳細について述べる。

時刻 t における無線 LAN の観測データ O_t は、BSSID β と受信電波強度 α の組で表される。

$$O_t = \{o_1, \dots, o_M\} \quad (17)$$

$$o_m = (\beta_m, \alpha_m) \quad (18)$$

ここで、 M は観測した基地局の数を表す。また、時刻 t における N 個の Particle は、存在する画像 (フロア) 名と座標、Particle の重みで表される。

$$A_t = \{a_t^1, \dots, a_t^N\} \quad (19)$$

$$a_t^n = (Image_t^n, \mathbf{p}_t^n, Weight_t^n) \quad (20)$$

ここで、 N は Particle の数を表す。また、 $Image_t^n$

はフロア (画像) 名、 $\mathbf{p}_t^n$ は座標、 $Weight_t^n$ は Particle の重みを表す。Particle Filter を用いた位置推定アルゴリズムの手順は以下ようになる。

1. Particle の初期化 (初回のみ): 事前位置にランダムに割り当て
2. Particle の移動: ランダムに移動
3. 重みの更新 (無線 LAN の観測): GMM により、電波強度を求め、存在確率を計算。
4. Particle のリサンプリング
5. 推定位置の提示

この手順を繰り返し実行する。ユーザの初期位置に関する情報が存在する場合には、Particle の位置により重みに差をつけることも可能である。

図 4.4 は、本手法を用いて推定した結果である。左図は、パーティクルを赤点で、実際の位置を青点でそのまま表示したもので、右図は、パーティクルが想定される精度を赤い円で表したものである。位置推定の結果を赤い円で表すことにより、ユーザにとってわかりやすい提示が可能になることを期待している。実験結果では、従来のベイズ推定や仮想無線 LAN 基地局を用いた手法と比較しても十分に精度の高い結果が得られている。

●第 14 部 無線を用いた位置情報プラットフォームの構築

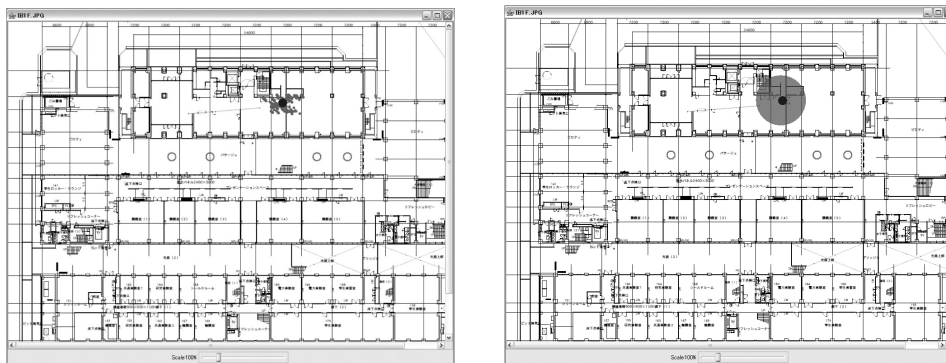


図 4.4. Particle Filter での推定

第 5 章 まとめ

本 WG では、位置依存情報の利活用のための基盤プラットフォーム構築を目指し、様々な無線 LAN 位置情報システムの構築と応用を進めている。

2009 年は実証実験の成果に基づき、実社会にソフトウェアを提供し、実際に 12 万回以上ダウンロードされるアプリケーションを提供することができた。

また、屋内の無線 LAN 位置推定に関しても、GMM を用いた手法を提案し、他手法と比較してコンパクトなデータにも関わらず高精度であることが確認できた。

今後はこれまでの研究をさらに発展させ

- 屋内のための位置依存情報データベースの構築手法の確立
- 位置情報を用いたウェブの事前データキャッシュ手法
- 無線 LAN 屋内位置推定手法の実端末への搭載と応用
- 大規模環境に適応可能な位置依存情報プラットフォームの構築

に関する研究を行う予定である。