

参加型センシングのためのデータ精度を考慮した仮想センサの提案

Proposal of an Accuracy-aware Virtual Sensor for Participatory Sensing

山本 淳† Atsushi Yamamoto 川崎 仁史† Hitoshi Kawasaki 倉沢 央† Hisashi Kurasawa 佐藤 浩史† Hiroshi Sato 中村 元紀† Motonori Nakamura 松村 一† Hajime Matsumura

1 はじめに

各ユーザがセンシングしたデータを収集して利用する参加型センシングが注目されている [1][2]。参加型センシングでは、センサを固定的に設置するのではなく、移動するユーザにセンサを携帯させることが特徴である。これにより、広大な空間の環境情報を低コストに収集できるという利点がある。しかしその一方で、特定の空間に存在するセンサの実体がユーザの移動とともに入れ替わり、その結果として収集されるデータの品質もゆらぐという問題がある [3]。参加型センシングを利用したアプリケーションの普及には、センサの実体が入れ替わることを隠蔽するとともに、要求された品質でデータを安定的に提供する仕組みが求められる。

センサの実体が入れ替わることを隠蔽する方法については、特定の空間でセンシングされたデータを仮想センサにマッピングする方法が提案されている [4]。しかし、この方法では、仮想センサのデータ品質については考慮されていない。一方、データの品質を考慮した仮想センサについても提案されている [5]。この方法では、データの品質としてカバー率を取り上げ、その制御に焦点を当てている。しかし、参加型センシングでは、個々のデータが信頼できるとは限らないため、カバーされているだけでは十分とはいえない。データの精度についても考慮する必要がある。

そこで本稿では、データの精度を考慮した仮想センサを提案する。具体的には、統計的な母平均の信頼区間の幅をデータ精度の尺度として用い、その信頼区間の幅を要求された一定の幅以下に制御するための手法を示す。また、シミュレーション実験により、提案手法の有効性を示す。

2 データ精度を考慮した仮想センサの提案

本章では、仮想センサのデータ精度を制御するための基本的なアイデアについて述べる。また、アイデアを実現する上での課題への対処について述べる。

2.1 基本的なアイデア

参加型センシングで収集される個々のデータは多分にばらつきを含んでおり、必ずしも信頼できない。アプリケーションが知りたいのは、ばらつきを含んだデータそ

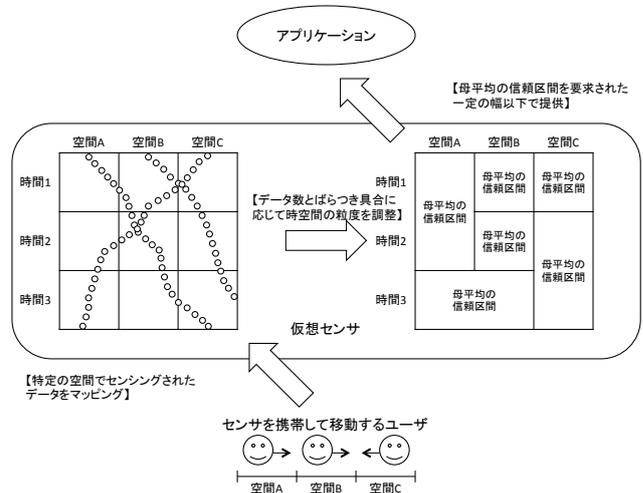


図 1 データ精度を考慮した仮想センサ

のものではなく、データからばらつきを取り除いた母平均についてである。そこで、複数のデータから一定の信頼度のもとで母平均を含む区間を統計的に推定した結果を仮想センサの出力とし、その信頼区間の幅をデータ精度の尺度として用いることとする。

母平均の信頼区間の幅は、信頼度を一定とした場合、ばらつき具合が大きくてもデータ数が多くなれば狭くなるという性質がある。したがって、仮想センサのデータ精度を制御するためには、ばらつき具合に応じて収集するデータ数を制御すればよいことになる。しかし、参加型センシングでは、いつでもどこでデータがどれくらい収集されるかは、各ユーザの移動に依存し制御できない。そこで、実際に収集されたデータ数とばらつき具合に応じて、時空間の粒度（領域の大きさ）を事後に調整することで、各領域における母平均の信頼区間の幅が要求された一定の幅以下になるように制御することとする。

図 1 に、以上の基本的なアイデアに基づいた、データ精度を考慮した仮想センサの概観を示す。このデータ精度を考慮した仮想センサにより、特定の空間に存在するセンサの実体が入れ替わることをアプリケーションから隠蔽するとともに、アプリケーションに対して要求された品質でデータを安定的に提供することを実現する。

2.2 実現上の課題への対処

2.2.1 単純な粒度調整手法の問題

時空間の粒度を調整する単純な手法として、全ての領域における母平均の信頼区間の幅が要求された一定の幅

† 日本電信電話株式会社 NTT 未来ねつと研究所
NTT Network Innovation Laboratories, NTT Corporation

以下になるように、時空間を一様な大きさの領域に機械的に分割する手法が考えられる。しかし、この単純な粒度調整手法には以下の二つの問題がある。

- (1) 各領域の大きさが必要以上に大きくなる。
参加型センシングでは、時間や空間によって収集されるデータ数やばらつき具合は異なる。母平均の信頼区間の幅はデータ数とばらつき具合に依存するため、時空間を一様な大きさの領域に分割する手法では、一部の領域を除き大部分の領域にとって領域の大きさが必要以上に大きくなる可能性がある。母平均の信頼区間は領域ごとに算出するため、領域の大きさが必要以上に大きくなることは、データの粒度が必要以上に粗くなることを意味し、好ましくない。言い換えると、要求されたデータ精度を満足する範囲で、各領域の大きさは小さいほうが好ましい。
- (2) 各領域内の母平均の分散が大きくなる。
時間的・空間的に近ければ、母平均の差がないとは限らない。このため、時間的・空間的な近さだけで機械的に領域化する手法では、母平均の差が極端にあるものを一つの領域にしてしまい、領域内の母平均の分散が大きくなる可能性がある。領域ごとの母平均は領域内の母平均の平均となるため、領域内の母平均の分散が大きくなることは、データの精度と引き換えに失われるものが大きくなることを意味し、好ましくない。言い換えると、同じ大きさの領域であっても、各領域内の母平均の分散は小さいほうが好ましい。

2.2.2 提案の粒度調整手法

図 2 に、上記の二つの問題に対処した時空間の粒度調整手法のフローチャートを示す。提案の粒度調整手法における対処のポイントは以下の二つである。

- (1) 時空間をデータ数とばらつき具合の実態に適合させた非一様な大きさの領域に分割する。
参加型センシングでは、時間や空間によって収集されるデータ数やばらつき具合は異なるので、時空間を分割する領域の大きさも、一様な大きさとするのではなく、データ数とばらつき具合の実態に適合させた非一様な大きさとする。具体的には、時空間を単位領域に分割した後、母平均の信頼区間の幅が要求された一定の幅以下になるまで、隣接領域を統合していく、ということを全ての領域に渡って適用する。これにより、要求されたデータ精度を満足する範囲で、各領域の大きさを小さく保つことを実現する。
- (2) 母平均の差の有無を検査した上で母平均の差がないものを優先的に領域化する。
時間的・空間的に近ければ母平均の差がないとは限らないので、時間的・空間的な近さだけで機械的に

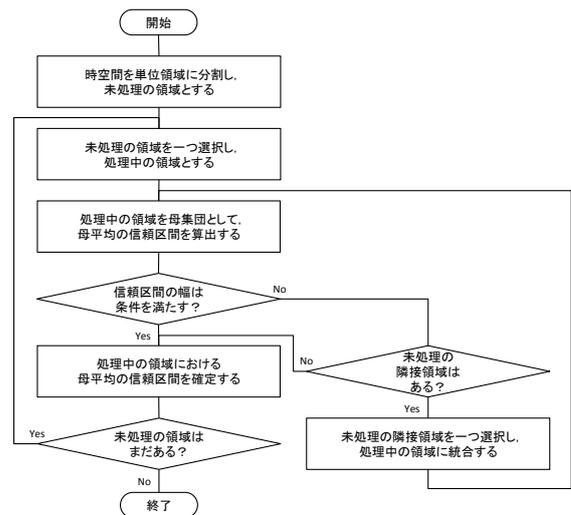


図 2 時空間の粒度調整手法

領域化するのではなく、母平均の差の有無を検査した上で母平均の差がないものを優先的に領域化する。具体的には、隣接領域を統合する際に、ランダムに隣接領域を選択するのではなく、母平均の差についての統計的な検定を行い、 p 値が大きいものを優先的に選択する。これにより、同じ大きさの領域であっても、各領域内の母平均の分散を小さく保つことを実現する。

3 シミュレーション実験

本章では、以下の二つの評価を目的として行ったシミュレーション実験について述べる。

- 仮想センサのデータ精度を制御するための基本的なアイデアの評価
時空間の粒度調整により、母平均の信頼区間の幅を一定の要求幅以下に制御できるか、また信頼区間に母平均を含む割合を一定の信頼度に保てるか。
- アイデアを実現する上での課題への対処の評価
提案の粒度調整手法により、単純な粒度調整手法に比べて、各領域の大きさを小さく保てるか、また各領域内の母平均の分散を小さく保てるか。

3.1 実験方法

シミュレーション用のデータを以下の方法で生成した。

- アメダスの東京都内 8 地点（気温が観測されてない地点と島を除く）のある 1 日の 1 時間ごとの気温の観測データを母平均とする 24 時間 × 8 地点の時空間配列を用意した。
- 50 人のユーザごとに、1 時間あたりのデータ数とばらつき具合（母標準偏差）をそれぞれ [5, 20] と [0.5,

1.0] の範囲からランダムに決定し、上記時空間配列の 8 地点を時間軸に沿って 24 ステップ分だけランダムウォークさせた。

- 上記各ステップでは、上記時空間配列の対応する母平均と上記ユーザごとに決定した母標準偏差の正規分布に従う乱数を、上記ユーザごとに決定したデータ数分だけ生成した。

なお、今回の実験では、カバー率が 100% の場合のみをシミュレーション用のデータとして用いた。

上記の方法で生成した同一のデータに対して以下の異なる手法をそれぞれ適用した。

- 時空間の粒度調整なし (non-adjusted)
単位領域 (1 時間 × 1 地点) ごとに母平均の信頼区間を算出するだけの手法である。
- 時空間の粒度調整あり (naive)
2.2.1 項で説明した単純な粒度調整手法である。
- 時空間の粒度調整あり (proposed1)
2.2.2 項で説明した提案の粒度調整手法 (対処 1 のみ) である。
- 時空間の粒度調整あり (proposed1+2)
2.2.2 項で説明した提案の粒度調整手法 (対処 1 と対処 2 の両方) である。

なお、今回の実験では、要求幅は 0.2、信頼度は 0.95 として手法を適用した。

以上のデータの生成と手法の適用を 10 回繰り返し、以下の評価項目を手法ごとに集計した。

- 母平均の信頼区間の幅
- 信頼区間に母平均を含む割合
- 各領域の大きさ
- 各領域内の母平均の分散

3.2 結果と考察

図 3 と図 4 に、母平均の信頼区間の幅と信頼区間に母平均を含む割合の結果を示す。図 3 より、時空間の粒度調整あり (naive, proposed1, proposed1+2) は、時空間の粒度調整なし (non-adjusted) に比べて、母平均の信頼区間の幅を一定の要求幅 0.2 以下に制御できていることがわかる。また、図 4 と図 3 の比較から、時空間の粒度調整あり (naive, proposed1, proposed1+2) は、時空間の粒度調整なし (non-adjusted) に比べて、母平均の信頼区間の幅を一定の要求幅 0.2 以下に制御しつつも、信頼区間に母平均を含む割合を一定の信頼度 0.95 に保てていることがわかる。

図 5 と図 6 に、各領域の大きさと各領域内の母平均の分散の結果を示す。図 5 より、提案の粒度調整手法 (proposed1, proposed1+2) は、単純な粒度調整手法 (naive)

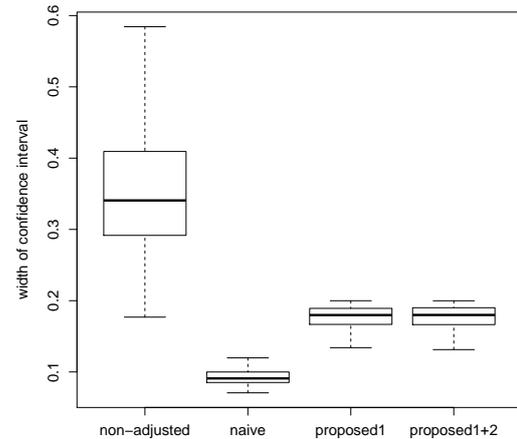


図 3 母平均の信頼区間の幅 (要求幅は 0.2)

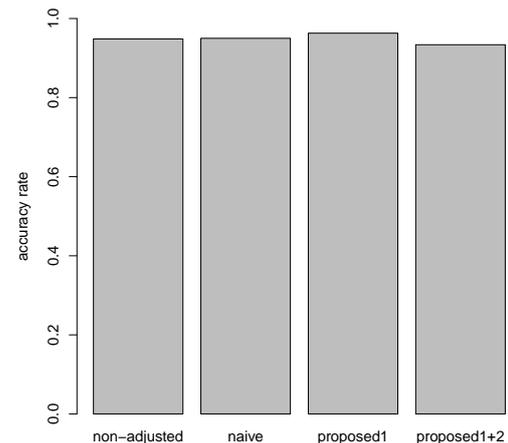


図 4 信頼区間に母平均を含む割合 (信頼度は 0.95)

に比べて、各領域の大きさを小さく保てていることがわかる。これは、図 5 と図 3 の比較からもわかるように、提案の粒度調整手法 (proposed1, proposed1+2) は、単純な粒度調整手法 (naive) に比べて、母平均の信頼区間の幅を一定の要求幅以下に制御する上で、各領域の大きさが必要以上に大きくなってないことを意味する。また、図 6 より、提案の粒度調整手法 (proposed1, proposed1+2) は、単純な粒度調整手法 (naive) に比べて、各領域内の母平均の分散を小さく保てていることがわかる。図 6 と図 5 の比較から、このうち、単純な粒度調整手法 (naive) と提案の粒度調整手法 (proposed1) の差分は各領域の大きさが小さくなったことに起因するものであり、提案の粒度調整手法 (proposed1) と提案の粒度調整手法 (proposed1+2) の差分が 2.2.2 項の対処 2 を施したことに起因するもの

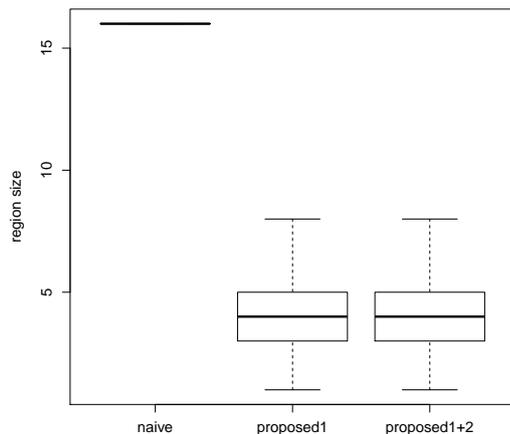


図5 各領域の大きさ

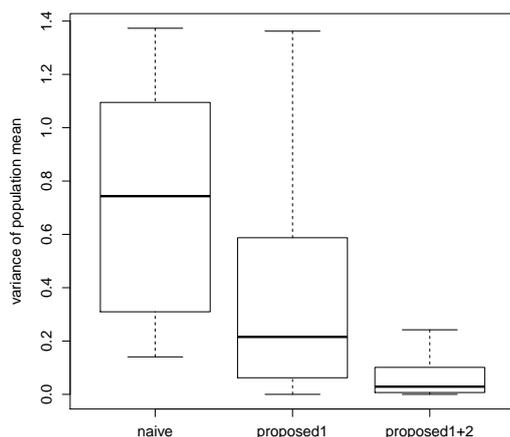


図6 各領域内の母平均の分散

であることがわかる。したがって、提案の粒度調整手法 (proposed1+2) は、単純な粒度調整手法 (naive) に比べて、たとえ各領域の大きさが同じであっても各領域内の母平均の分散を小さく保てると推測する。

以上より、時空間の粒度調整により、母平均の信頼区間の幅を一定の要求幅以下に制御でき、また信頼区間に母平均を含む割合を一定の信頼度に保つことができるといえる。また、提案の粒度調整手法により、単純な粒度調整手法に比べて、各領域の大きさを小さく保つことができ、また各領域内の母平均の分散を小さく保つことができる。つまり、仮想センサのデータ精度を制御するための基本的なアイデアと、アイデアを実現する上での課題への対処の有効性が確認できたといえる。

4 おわりに

本稿では、データの精度を考慮した仮想センサとして、統計的な母平均の信頼区間の幅をデータ精度の尺度として用い、その信頼区間の幅を要求された一定の幅以下に制御するための手法を提案した。データの精度を考慮した仮想センサにより、センサの実体が入れ替わることを隠蔽するとともに、要求された品質でデータを安定的に提供できるようになる。これにより、参加型センシングを利用したアプリケーションの普及が期待できる。

提案手法は、データ数とばらつき具合に応じて時空間の粒度 (領域の大きさ) を調整することでデータ精度を制御することを基本的なアイデアとし、要求されたデータ精度を満足する範囲で各領域の大きさを小さく、また各領域内の母平均の分散を小さく保てるという特徴をもつ。シミュレーション実験により、提案手法の有効性を確認した。今後の課題として、シミュレーション実験ではなく、実データ・実アプリケーションを用いた提案手法の評価がある。

参考文献

- [1] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M.B. Srivastava, "Participatory sensing," Proc. First Workshop on World-Sensor-Web (WSW'2006), Oct. 2006.
- [2] A.T. Campbell, S.B. Eisenman, N.D. Lane, E. Miluzzo, R.A. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn, "The rise of people-centric sensing," IEEE Internet Computing, vol.12, no.4, pp.12–21, July-Aug. 2008.
- [3] 小林亜令, 木實新一, "携帯電話を用いたセンシングの可能性と課題," 情報処理, vol.51, no.9, pp.1157–1163, Sep. 2010.
- [4] A. Kansal and F. Zhao, "Location and mobility in a sensor network of mobile phones," Proc. 17th International workshop on Network and Operating Systems Support for Digital Audio & Video (NOSSDAV 2007), June 2007.
- [5] H. Weinschrott, F. Dürr, and K. Rothermel, "Streamshaper: coordination algorithms for participatory mobile urban sensing," Proc. 7th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS 2010), pp.195–204, Nov. 2010.