

センサー間の隣接関係の自動推定の高精度化

Accurate Estimation for Sensor Network Topology

高橋 謙輔†

Takahashi Kensuke

菅原俊治†

Sugawara Toshiharu

1 はじめに

近年、情報通信端末の小型化が進みユビキタス社会への実現が着実に進んでいる。この社会では、大量のセンサーを配置し、環境中の情報を収集して実世界と有機的に結ばれた context-awareness なサービスを想定している。しかし、大量のセンサーを配置すると、個々のセンサーを手動で管理する労力は大きくなる。その問題を解決するには、センサー自身による自動設定が必要である。本研究では、この自動設定の中でも、実世界の位置関係を示し、各種サービスにとって必須の情報であるセンサー間の隣接関係に着目する。

隣接関係を推定するものとしては、無線強度を利用した [4] などがある。しかし、全てのセンサーが無線機能を備えているとは限らない。[3] では、無線機能などを前提とせず、センサーの反応データから隣接関係を推定するアルゴリズムを提案している。しかし、これらは環境中の人数が一定であること、行動パターンが同じなどを前提としており、応用を考えるとこの前提は必ずしも合致しない。これに対し [1] では、タイムスタンプ付きのセンサー反応データのみからの隣接関係の推定に成功しており、ユビキタス社会のアプリケーションの観点からは自然な状況での提案法と言える。しかし、実用性を考えるとその精度は十分ではなく、より効率的な推定法が必要である。

我々はこれまで (1) 環境中の人数などの前提条件を考えずに、(2) センサーから得られた時系列の反応データのみで、(3) 短期間の使用データで、より精度の高い隣接関係推定のアルゴリズムの提案をしてきた [5]。[5] では個々のセンサー間の反応時間間隔のみに注目しているのに対し、本研究ではセンサー全体の反応した順番に注目し、より高い精度で隣接関係を推定する新たなアルゴリズムを提案する。

2 問題の定義

2.1 センサーネットワークの表現方法

本研究では、センサー間のネットワーク構造を重み付き有向グラフ $G = (V, E)$ で表現する。ここで、各ノード $v_i \in V$ は実環境の各センサーに 1 対 1 対応している。また、 $e_{i,j} \in E$ は v_i から v_j に向かう経路を表し、その重みを $\tau_{i,j}$ とし、隣接尤度を表すものとする。本研究では、入力されたセンサーデータの解析によって $\tau_{i,j}$ を変化させ、隣接尤度を求めることで隣接関係を推定する。

2.2 入力データ

本研究では、センサーの前を人が通ると反応する人感センサーを用いた。よって入力データは、ある時間 t において反応したセンサーの識別子が得られ、これを $O_t = \{i_{t_1}, i_{t_2}, \dots, i_{t_n}\}$ 、 $i_{t_j} \in [1, N]$ と表す。ここで $N = |V|$ である。例を図 1 に示す。この例では時刻 $t = 1$ には 1, 3, 5 のセンサーが、時刻 $t = 2$ には 2, 3 のセンサーが反応したことを示している。得られたデータをタイムスパン T で区切り、 $O_t \sim O_{t+T-1}$ のデータを 1 回の解析で使用する。

†早稲田大学基幹理工学研究科情報理工学専攻

時刻 t における反応データ

$$O_t = \{i_{t_1}, i_{t_2}, \dots\}$$

(反応データ例)

$$O_1 = \{1, 3, 5, \dots\}$$

$$O_2 = \{2, 3, \dots\}$$

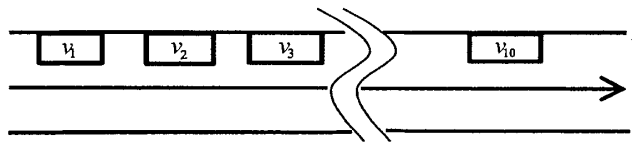
図 1: 時刻 t におけるセンサーの反応データ

図 2: 想定される経路

2.3 隣接関係推定の概要

室内にセンサーが取り付けられる例を、図 2 に示す。矢印の通り人が歩いて、時刻 t_1 に $O_{t_1} = \{v_1\}$ のデータが入力されたとする。次の時刻 t_2 には $O_{t_2} = \{v_2\}$ が、更に t_3 には $O_{t_3} = \{v_3\}$ が得られると予想できる。このように、ある時刻にセンサーが反応した場合、そのセンサーと隣接関係にあるセンサーは時系列的に反応するという性質を利用し、局所的な隣接関係を推定でき、これらを合わせて全体のセンサーの隣接関係を推定する。例えば、図 2 の場合はデータ $O = \{O_{t_1}, O_{t_2}, O_{t_3}\}$ より $V = \{v_1, v_2, v_3\}$ 、 $E = \{e_{1,2}, e_{2,3}\}$ という局所的なグラフ構造 (サブグラフ) を作成することができる。このように局所的に作成されたサブグラフのエッジに重みを与え、センサー全体で構成されているグラフ構造 (メイングラフ) のエッジにその重みを付加することで全体の隣接関係を推定できる。しかし、以下のような課題が生じる。

1. 環境中で複数人が同時に行動した場合、隣接関係のないセンサー同士が時系列的に反応する場合がある。
2. センサーの誤反応や反応漏れなどにより、隣接関係にあるセンサー同士が時系列的に反応しない場合がある。

本研究ではこれらの問題を解決するために、隣接関係の推定において 2 つのフェーズを用いる。1 つ目はデータセット作成フェーズであり、第 3 節で述べる入力データの特徴を用いてセンサーからの入力データの信頼性を推定する。2 つ目は隣接関係推定フェーズであり、[1] と同様、フェロモンモデル [2] を用いてデータセット作成フェーズで作成されたデータをもとに隣接関係を推定する。[1] では隣接関係のみに注目してフェロモンを付加しているのに対して、本研究では非隣接関係にも注目してフェロモンモデルを用いた。

3 反応頻度と精度

本研究では、反応データの特徴として「反応頻度の高いタイムスパンにおけるデータの精度は低い」という仮説を立てた。環境中で行動している人数が多いほど、センサーの反応データから隣接関係を得られる精度が低くなることを想定し

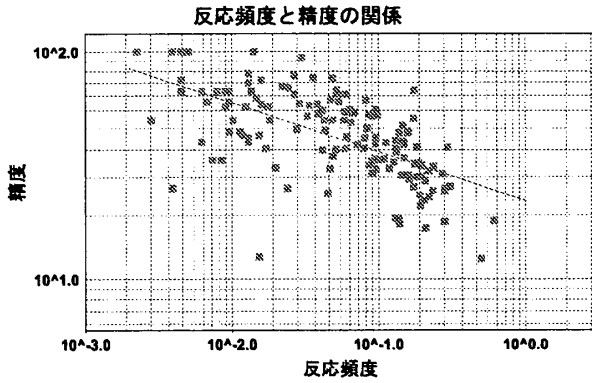


図3: 反応頻度と精度の関係

ている。しかし、本研究では環境中の人数は仮定しない。そこで、タイムスパン T における反応頻度が高ければ、その T の間に環境中で行動していた人数は多いと推定し、その時の反応データの精度は低くなると仮定する。以下で、この仮説の妥当性について検証する。

タイムスパン T の反応データ $O_t \sim O_{t+T-1}$ の反応頻度 $freq(t)$ を、

$$freq(t) = 100 \times \frac{\sum_{i=t}^{t+T-1} |O_i|}{T \times N} \quad (1)$$

と定義する。なお、 N は環境中のセンサー数である。

次に、反応データの精度を定義する。本研究では第2.3節で述べたように、入力データからサブグラフを作成して隣接関係を推定する。例えば、 $O_{t_1} = \{v_1, v_2\}, O_{t_2} = \{v_3\}$ の場合にはサブグラフのエッジとして $E = \{e_{1,3}, e_{2,3}\}$ が作成される。このように、反応データ $O_t \sim O_{t+T-1}$ の中で、ある時刻にセンサーが反応した後に反応したセンサーとのエッジを作成し、その作成されたエッジの中で実際に隣接関係にあるものの割合を精度とする。ここで、サブグラフによって作成されたエッジの集合を E_{sub} 、実際に隣接関係にあるエッジの集合を $E_{correct}$ とすると、精度 $acc(t)$ は、

$$acc(t) = 100 \times \frac{|E_{sub} \cap E_{correct}|}{|E_{sub}|} \quad (2)$$

とする。実際のセンサーデータとセンサー配置のレイアウト図をもとに求めた結果を図3に示す。図3は横軸に反応頻度の対数を、縦軸に精度の対数をとったものであり、この時の相関係数は -0.78 であった。これらのことより、反応頻度と精度はべき乗則で近似できる。

4 提案アルゴリズム

4.1 データセット作成フェイズ

データセット作成フェイズの目的は、複数人が同時に行動することによって隣接関係にないセンサー同士が時系列的に反応した場合や、センサーの反応漏れなどの誤った反応データの重みを小さくし、入力データの信頼性を推定することである。

最初に入力データ $O_t \sim O_{t+T-1}$ を用いてサブグラフを作成する。ここで、サブグラフは重み付き片方向有向グラフとする。[5]では各センサー間の反応間隔時間に注目し、反応時間間隔が短ければ隣接関係尤度が高いと判断した。それに対し本研究では、 O_i から時系列的に入力データを解析し、反応したセンサーとその次に反応したセンサーを

エッジでつなぐ。そして、あるセンサーが反応した後に α 秒間新たなセンサーが反応しなければ、そのサブグラフは完成し、その次に反応したセンサーから新しいサブグラフを作成する。例えば $O_{t_1} = \{v_1\}, O_{t_2} = \{v_2\}, O_{t_3} = \{v_3\}, O_{t_3+\alpha+1} = \{v_{10}\}, O_{t_3+\alpha+2} = \{v_{11}\}$ の入力データがあったとすると、 $G_1 = \{\{v_1, v_2, v_3\}, \{e_{1,2}, e_{2,3}\}\}, G_2 = \{\{v_{10}, v_{11}\}, \{e_{10,11}\}\}$ といった2個のサブグラフが作成される。このように、 $O_t \sim O_{t+T-1}$ のデータを用いて複数のサブグラフを作成する。本研究での手法では、センサー間の反応間隔時間が短くても、その間に反応しているセンサーがあれば隣接関係に無いと推定した。

次に、このサブグラフのエッジに重み付けをする。第3節で述べたように反応頻度と精度はべき乗則と推定できる。そこで、タイムスパン T の反応頻度 $freq(t)$ を用いて、

$$w(t) = a \times freq(t)^b \quad (3)$$

と定義する。この $w(t)$ をサブグラフのエッジの重みとする。なお、本研究では図3の近似式より $a = 1.84, b = -0.3$ とした。

データセット作成フェイズでは、局所的なセンサー間の隣接関係を推定し、反応データの特徴を用いて、入力データの精度の重み付けによって、隣接関係推定の精度を向上させる。

4.2 隣接関係推定フェイズ

隣接関係推定フェイズでは、データセット作成フェイズで作成したサブグラフをもとにメイングラフの各エッジの重みを変化させ、実際の隣接関係を推定する。

このフェイズでは、サブグラフ毎にエージェントをサブグラフ作成時に最初に追加されたノードから、最後に追加されたノードまでエッジを通して移動させる。この時スタート地点に割り当てるエージェント数は環境中のセンサー数とする。エージェントの移動先は、サブグラフのエッジの重みとメイングラフのエッジの重みの多寡によって確率的に決定する方法と、サブグラフのエッジの重みの多寡のみで確率的に決定する方法の2種類ある。前者は過去の解析の結果に依存して移動先を決定するが、それだけでは局所解に陥る可能性があり、それを防ぐために後者の手法を用いる。

次に、各サブグラフのエッジを移動したエージェントの数に応じてメイングラフに重みを付加する。また、この時に最大値と最小値を設定して重みを制限することで探索が発散することを防ぐ。更に、一定の割合で過去の結果を割り引き新しい情報が比較的強く反映する。ここで、サブグラフのエッジ $e_{i,j}$ の重みを $\eta_{i,j}(t)$ 、 $e_{i,j}$ を通過したエージェント数を $A_{i,j}(t)$ 、メイングラフのエッジ $e_{i,j}$ の重みを $\tau_{i,j}(t)$ 、割引率を ρ とすると以下の式になる。なお、本研究では $\tau_{max} = 500, \tau_{min} = 0, \rho = 0.3$ を用いた。

$$\begin{aligned} \tau'_{i,j}(t) &= \rho(A_{i,j}(t) \times \eta_{i,j}(t)) + (1 - \rho)\tau_{i,j}(t) \quad (4) \\ \tau_{i,j}(t+1) &= \begin{cases} \tau_{max} & \text{if } \tau'_{i,j}(t) > \tau_{max} \\ \tau'_{i,j}(t) & \text{otherwise} \end{cases} \quad (5) \end{aligned}$$

次に隣接尤度が低いと思われるエッジの重みを減少させる。第4.1節で作成したサブグラフのノードは、時系列的に反応したセンサーの集合なので、比較的隣接尤度が高いものの集合であるのみなせる。逆に、各サブグラフのノードとそれ以外のノードにおける隣接関係の尤度は低いのみなせる。そこで、サブグラフのエッジの集合を E_{sub} 、メイングラフのエッジの集合を E_{main} とすると非隣接関係と推定できるエッジの集合 E_{not} は $E_{not} \stackrel{\text{def}}{=} E_{main} - E_{sub}$ となる。この

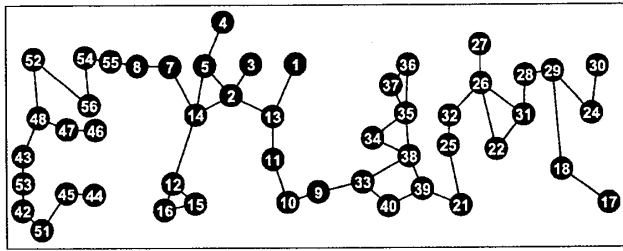


図 4: センサーネットワークの正解 Topology Map

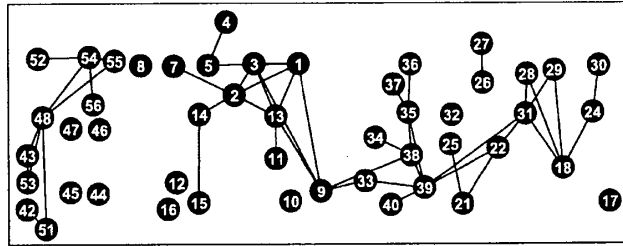


図 5: 提案手法による解析終了時の Topology Map

時 $e_{i,j} \in E_{not}$ の重み $\tau_{i,j}(t)$ に対し,

$$\tau'_{i,j}(t) = (1 - \rho)\tau_{i,j}(t) - \rho\eta_{i,j}(t) \quad (6)$$

$$\tau_{i,j}(t+1) = \begin{cases} \tau_{min} & \text{if } \tau'_{i,j}(t) < \tau_{min} \\ \tau'_{i,j}(t) & \text{otherwise} \end{cases} \quad (7)$$

と $\tau_{i,j}(t)$ を変化させる。

4.3 隣接関係の推定

これまでの解析から、メイングラフの各エッジの重み $\tau_{i,j}(t)$ が決定された。このエッジの重みが隣接関係尤度を表し、 $\tau_{i,j}(t) + \tau_{j,i}(t)$ の値が閾値 $th(t)$ より大きければそのエッジは隣接関係にあるとする。この $th(t)$ は解析ごとにメイングラフの各エッジの重みのヒストグラムを作成し、階級の大きいほうから解析した時に、立ち上がっている箇所の重みをその値とする。

4.4 誤差率の求め方

実験によって得られたセンサー間の隣接関係を、隣接行列 $A = \{a_{i,j}\}$ (隣接関係ならば $a_{i,j} = 1$, 非隣接関係ならば $a_{i,j} = 0$) で表す。レイアウト図より得られる正解の隣接関係を行列 $A' = \{a'_{i,j}\}$ とすると、実験によって得られた隣接関係の誤差率 err を以下の式で求める。なお、 m は行列の要素の個数である。

$$err = 100 \times \left(\frac{1}{m^2 - m} \right) \sum_{i,j,i \neq j} (a_{i,j} - a'_{i,j})^2 \quad (8)$$

5 実験

今回の解析実験の使用データとして、大阪大学産業科学研究所の沼尾/栗原研究室に取り付けられたセンサーネットワーク (図 4) における 2006/7/7~2006/7/17 の 10 日間分のセンサー情報を頂いて使用した。その結果図 5 の Topology Map を得ることができた。また、誤差率の時系列変化を図 6 に示す。図 6 より解析終了時の誤差率は提案手法が 4.33%、[1] の手法が 34.0%、[5] の手法が 6.25% となり、[5] と比べ 30% の減、特に [1] と比べると 1/8 程度に抑えられ、既存研究より精度が格段に向上したことがわかる。

誤差率の時系列変化

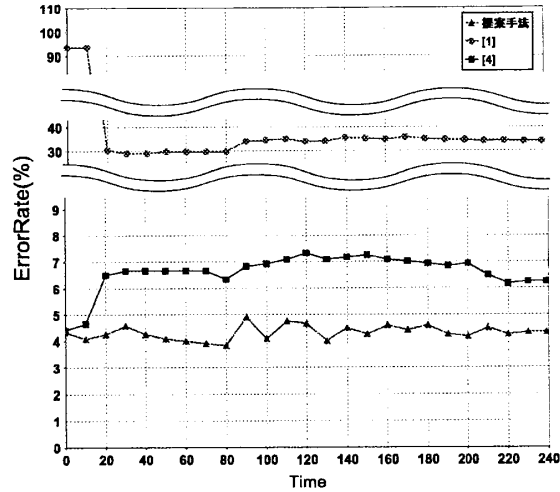


図 6: 誤差率の時系列変化

6 結論

本研究ではセンサーネットワークにおけるセンサー間の隣接関係の推定法として、各センサー反応データを時系列的に利用する手法を提案した。

本提案の特徴として、(1) 入力データの信頼性を推定し、信頼性に応じて重み付けを行い、(2) 隣接関係推定方法としてフェロモンモデルを用いたことが挙げられる。(1) の特徴により、環境中での複数人の同時行動におけるセンサーの反応やセンサー自身の誤反応による入力データを結果に反映されにくくできた。(2) の特徴により、解析時に過去の情報をもとにした解析を行うことができ、また、一定の割合でランダムな探索を行い、局所解に陥ることを防げた。更に、一定の割合で結果を割り引くことで、古い情報からの反映を弱くしている。その結果、提案アルゴリズムでは既存研究と比べ、高い精度でセンサー間の隣接関係を推定できることが確かめられた。

今後の課題としては、他の環境で得られたデータに対しても提案手法が有効かを検証したい。

謝辞

本研究は総務省戦略的情報通信研究開発制度 (SCOPE) の助成を受けている。

参考文献

- [1] 玉置洋, 福井健一, 沼尾正行ほか, “Ant Colony Optimization によるセンサー隣接関係抽出法の提案,” 知識ベースシステム研究会 Vol.77, pp.21 - 28 (2007-3)
- [2] Marco Dorigo and Gianni Di Caro, “The Ant Colony Optimization Meta-heuristic,” *New ideas in optimization*, McGraw-Hil Ltd., UK, pp. 11 - 32 (1999)
- [3] Dimitri Marinakis and Gregoru Dudek, “Topological Mapping through Distributed, Passive Sensors,” *IJCAI-07* (2007-1)
- [4] A. Savvides, CC Han and MB Strivastave, “Dynamic fine-grained localization in ad-hoc networks of sensors,” *Proc. ACM/IEEE MOBICOM 2001*, pp. 166-179 (2001-4)
- [5] 高橋謙輔, 菅原俊治, “時系列センサーデータを使用したセンサー間の隣接関係の推定法の提案,” 信学技報「人工知能と知識処理」 Vol.107 No.523, pp.85 - 90 (2008-3)