

アドホックコミュニケーションにおける口コミ的情報伝播の要因推定

Estimation of a Factor of Word-of-mouth Communication-like Information Spread in Ad Hoc Communications

武井 瑞希 片上 大輔 新田 克己
Mizuki Takei Daisuke Katagami Katsumi Nitta

東京工業大学大学院総合理工学研究科
Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology.

We focus on the information spread in ad hoc communications, and propose an agent based simulation method using a word-of-mouth information transmission model. In this method, each agent represents a person who lives in a city. They move to various spots such as houses, offices or schools. When an agent meets another agent in a spot, information transfer occurs stochastically. We generate KeyGraph from results of simulation to extract concurrent relation between the agents. This graph shows the structure of social community. We estimate information source in social community using this simulation method and KeyGraph.

1. 研究背景・目的

近年, Web を介した情報交換(BBS,Blog,SNS など)が広く普及している. これらの情報交換は, ユーザが通信先を意識的に選択して行われる.

一方, ユーザの行動範囲に基づくローカルな情報の収集には, Bluetooth などの近距離無線通信を利用した直接遭遇する人との情報交換(アドホックコミュニケーション)が有効であるとされており, これを対象とした研究も盛んになってきている.

アドホックコミュニケーションでは, ユーザが相手を意識することなく自動的に情報の取得・配信が行われるため, 情報の行方をコントロールすることが困難である. しかし, 情報伝播の要因となる事象を発見することができれば, 悪意情報の蔓延防止や, 優良情報の効率的な取得・配信につながると考えられる.

情報発信源の推定手法として, データの共起性を利用した「組織の見えない黒幕」を探る手法[前野 2007]がある. この手法では, 組織におけるメールの送信記録などのコミュニケーション情報を基に人物間の関係性をネットワークとして可視化し, 弱いハブに注目することで組織における重要人物を発見する. しかし, この手法はコミュニケーション情報として送信記録という確定した情報を用いるため, 口コミ的情報伝播の特徴である, 「確率的な通信」「時間経過による情報の消失」という性質のコミュニケーション情報解析には適していない.

そこで本研究では, アドホックコミュニケーションの性質に適した情報伝播の要因事象の推定手法の提案を目的とし, エージェントベースシミュレーションを用いた口コミ的情報伝播現象のデータの生成と, シミュレーションにおける人の行動履歴データに基づく人同士の社会的つながりを表すネットワークの生成を行い, ネットワークを解析することで情報発信源を推定する.

2. 関連研究

2.1 アドホックコミュニケーション

近年, 近距離無線通信を利用して, 自分の直接遭遇する人との情報交換を対象とした研究が盛んになってきている. そのようなアドホック通信を利用した情報交換をアドホックコミュニケーションと呼ぶ.

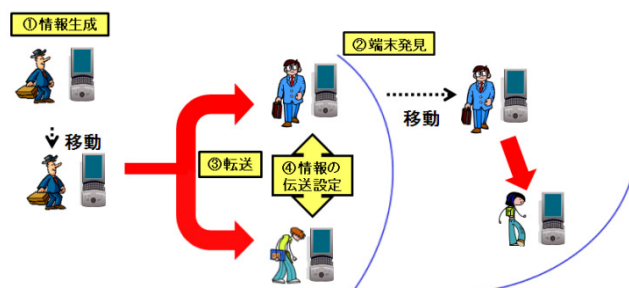


図1 アドホックコミュニケーション

ョンと呼ぶ.

アドホックコミュニケーションには, 地理的に近接した端末との通信をユーザが移動する毎に繰り返すことで, 地域を共有する相手から信頼性の高い情報を取り入れることができるというメリットがある.

例えば, 自分の行動範囲のお店等の店舗情報や施設に関する口コミ情報が近隣を歩いているだけで収集でき, 同じ居住地域の住人のお勧め等の意見を抽出し, さらに配信することが出来る[高橋 2003].

図1に, アドホックコミュニケーションの基本的な流れを示す.

①PDA 等の情報端末をもったユーザが端末内に情報を生成し移動する. ②移動中に無線の通信範囲内に情報端末を持った他のユーザが認識されると, ③無線通信を通じて情報が転送される. ④情報を受け取ったユーザが情報の伝送設定を行うと, さらにそのユーザを発信源として情報が他社へ伝送されていくことになる.

2.2 コミュニケーション情報に基づく重要人物発見手法

前野らは, 電子メールによるコミュニケーションやテロ組織におけるコミュニティ内の送信記録に基づく「意思決定の見えない黒幕」の同定手法を提案している. 黒幕とは, 電子メールによるコミュニケーションにおいてメールを使わずに議論の誘導をした人物や, テロ組織において実行犯を支援した共謀者を指す. この研究では, コミュニティ内の送信記録を基にヒューマンネットワークを生成し, 「データ結晶化」の手法を用いることで, クラスタ間の潜在的なつながりを顕在化させるような事象をダミーノードとして出現させる. このダミーノードを基に, 送信記録からは観測できない, 組織のコミュニケーションと意思決定を誘導する見

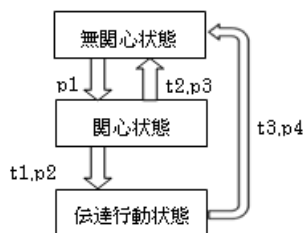


図2 ロコモの情報伝達モデル

えない黒幕の推定を行っている。

この黒幕推定手法を本研究で注目しているアドホックコミュニケーションにおける情報伝播に用いることができれば、情報伝播に強く影響を与えた人物の推定が可能となり、情報発信源の推定につながると考えられる。

しかし、この手法はコミュニケーション情報として交信記録という確定した情報を用いるため、ロコモの情報伝播の特徴である、「確率的な通信」「時間経過による情報の消失」という性質のコミュニケーション情報解析には適していない。

3. 研究方法

本研究では、エージェントベースシミュレーションを用いて実際に起こりうる人の行動とそれに伴うロコモの情報伝播現象のデータを生成する。さらにシミュレーションにおける人の行動履歴データから人同士の社会的つながりを表すネットワークを生成し、ネットワークを解析することで情報発信源の推定を行う。

3.1 シミュレーションによる情報伝播データの生成

情報伝播のデータを得るため、ロコモの情報伝達モデルを採用したエージェントベースシミュレーションを利用する。エージェントベースシミュレーションは、エージェントとその相互作用によるシミュレーションであり、社会の現象のプロセスを理解するための有力な手段とされている。

(1) ロコモの情報伝達モデル

従来のロコモ現象を表す代表的なモデルとして、SIRモデル[Kermack 91]が挙げられる。SIRモデルは病気の伝染モデルとして提案されているが、情報の伝播という点においても有効性が高いことが知られている。[林 07]では、ロコモの特徴である情報への関心保持状態をSIRモデルに加えたSHIRモデルを提案している。本研究で用いる情報伝達モデルは、このSHIRモデルを基にしている。モデルの概要を図2に示す。

本モデルにおけるユーザの情報保持状態は、無関心状態、関心状態、伝達行動状態の3つである。無関心状態は、ユーザが端末に情報を保持していない状態を表す。無関心状態のユーザが伝達行動状態のユーザと接近したとき、一定確率($p1$)で情報の伝送が行われ、無関心状態だったユーザは情報を保持することとなり関心状態へ遷移する。さらに一定時間($t1$)経過後に、関心状態のユーザには保持している情報を周囲に伝えたいという欲求が一定確率($p2$)で生じ、近接ユーザに情報を発信する伝達行動状態となる。また、情報を保持している関心状態および伝達行動状態のユーザは、一定時間($t2,t3$)経過後に一定確率($p3,p4$)で情報に対する興味を失い、無関心状態へ遷移する。

(2) エージェントベースシミュレーション

本研究では、シミュレーション言語としてSOARS[田沼 07]を用いる。SOARSは、モデル内の空間を「スポット」のネットワーク構造として表現するモデル化手法を用いる。「スポット」とはシミュレーション内でエージェントが移動できる空間のことを指し、この「スポット」を用いることで、様々な相互作用が容易に記述でき、本研究のように現実社会を想定した仮想社会の表現に適していると考えられる。

本研究における都市型モデルでは、都市における各施設(家、会社、学校など)をスポットで表し、住民をエージェントで表す。各エージェントに社会的役割(会社員、主婦、学生など)を設定し、エージェントは与えられた日常行動のルールに従ってスポット間の移動を行うものとする。以下に、シミュレーションで用いる基本的な設定を示す。

- エージェントの役割
 - Man(男性就業者): 仕事をもつ男性
 - Woman1(女性就業者): 仕事をもつ女性
 - Woman2(専業主婦): 仕事を持たずに家事につく女性
 - Child(学生): 学校に通う男女
- 移動先の設定
 - Home(家): 全エージェントの移動先
 - Office(会社): ManとWoman1の移動先
 - Shop(店): Woman2の移動先
 - School(学校): Childの移動先
 - Station(駅): 遠方へ移動するエージェントの移動先

さらに、本シミュレーションにロコモの情報伝達モデルを導入することで、都市における情報伝播現象を発生させる。各エージェントは情報保持状態に関する状態変数を設定され、条件に応じて無関心状態、関心状態、伝達行動状態の3つの状態への遷移が生じる。エージェントの初期状態は無関心状態とし、無関心状態のエージェントは伝達行動状態のエージェントと同一スポットに存在するとき、一定確率で情報を受けとり関心状態へと遷移する。シミュレーション開始時に伝達行動状態のエージェントを設定することで情報伝播が生じ、シミュレーション終了時に複数の伝達行動状態エージェントが発生することになる。

また、シミュレーションの出力として各エージェントの移動履歴データを取得する。移動履歴データは、各スポットに存在するエージェントを15分刻みで記録したものとなっている。このデータを解析することで、同一スポットに存在する頻度(共起頻度)の高いエージェント群を抽出することが出来る。共起頻度の高いエージェント間では情報伝播が生じる可能性が高いといえるため、共起頻度の高いエージェント群は情報伝播の過程に強く影響を及ぼしていると考えられる。

3.2 シミュレーション結果からのKeyGraph抽出

本研究では、KeyGraph[大澤 99]を用いてシミュレーションにおけるエージェントの行動履歴データからのネットワーク抽出を行う。KeyGraphは、バスケット単位で区切られたデータにおける各要素の共起情報を基に要素同士の関連性を可視化する技術である。同一バスケットに同時に出現する頻度の高い要素同士は関連が強いと解釈される。エージェントの移動履歴データにおいて、要素を各エージェント、バスケットを各スポットと定義することで、同一スポットに存在する頻度の高い(共起頻度の高い)エージェント同士がリンクでつながれるKeyGraphを生成することができる。

共起頻度の計算式を以下に示す。エージェント e_i, e_j の共起性を表す尺度 $C(e_i, e_j)$ を、エージェント e_i, e_j が同一スポットに存在した頻度 $F(e_i \cap e_j)$ で定義している。

$$C(e_i, e_j) \equiv F(e_i \cap e_j).$$

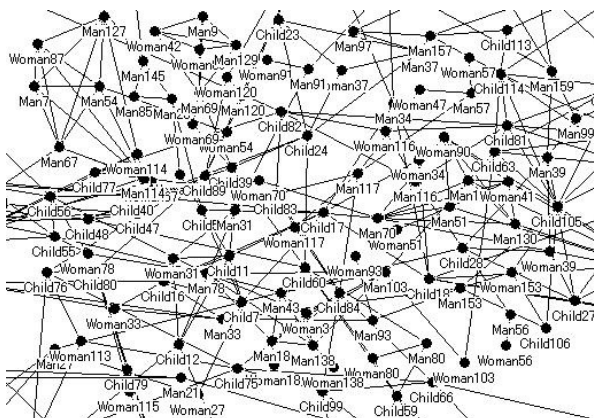


図3 0:00~24:00におけるエージェントの共起情報

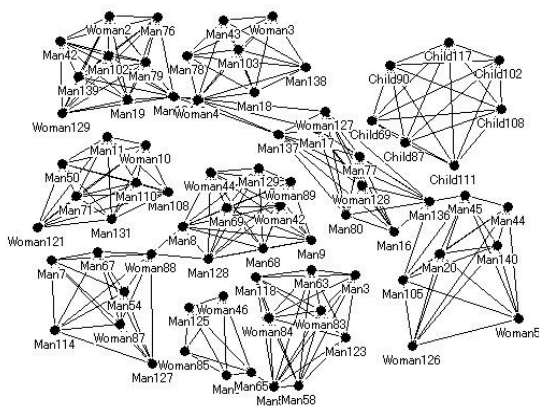


図4 9:00~17:00におけるエージェントの共起情報

各エージェントをノードで表し、全エージェントについてCを導出し、上位n個のCに対応するエージェントのノード間にエッジを張ることでKeyGraphを生成する。(nは任意)

ここで生成するKeyGraphは、都市における住民の社会的つながりを表すヒューマンネットワークといえる。

3.3 KeyGraphの解析による情報発信源の推定

(1) ネットワーク分析に基づく情報発信源の推定

エージェントの移動履歴データから得られるKeyGraphを基に、情報受信エージェントがもつ共起関係を辿り、情報の発信源となったエージェントの推定を行う。

従来手法[前野 07]では、コミュニケーション履歴における時間の概念を考慮する必要がないため、全コミュニケーション履歴から1種類のKeyGraphを生成し解析を行っている。

しかし、本研究におけるエージェントの移動履歴データに基づくヒューマンネットワーク構造は常に一定ではなく、時間の経過により変化する。典型的な例として、家庭をもつ男性就業者の場合、日中は会社のコミュニティと強い関係をもつが、夜や休日には家庭や地域のコミュニティと関係をもつ可能性が高いことなどが挙げられる。

図3は、従来手法に倣い、1日のエージェントの移動履歴データの全てを基に生成したKeyGraphである。本研究におけるエージェントの移動履歴は、従来研究におけるコミュニケーション履歴に対応すると考えられる。このKeyGraphでは、全時間帯の共起関係を一括してリンクとして表示しているため、時間経過により変化しているコミュニティ構造を観測することが困難である。

一方、図4は1日のエージェントの移動履歴データのうち、昼間(9:00~17:00)の移動履歴データのみを基に生成したKeyGraphである。

このKeyGraphでは、複数のクラスタが形成されていることが明確に観測でき、これらのクラスタは会社や学校など日中に形成されるコミュニティの存在を示唆している。伝達行動状態エージェントが特定のコミュニティに所属している場合、情報の伝播はコミュニティ内で生じる可能性が高いため、情報発信源の推定を行うためには各時間帯におけるコミュニティ構造を観測することが重要であると考えられる。

したがって、本研究では、時間経過によるネットワーク構造の変化に注目し、時間帯毎の共起関係を分析することで、情報発信源候補の推定を行う。

3.4 情報発信源推定アルゴリズム

生成したKeyGraphを解析し、情報発信源の推定を行う。

以下にアルゴリズムを示す。

1. シミュレーション終了時刻 T に伝達行動状態のエージェント群を A_T とする。
2. A_T が情報を受信したと予測される期間 $t_i = (T - x) \sim (T - y)$ ($i = 0, 1, \dots$) ($0 < x < y$) に対応するKeyGraphにおいて、 A とリンクするエージェント群を A_i とする。
3. A_i が情報を受信したと予測される期間 $t_{i+1} = (t_i - x) \sim (t_i - y)$ に対応するKeyGraphにおいて、 A_i とリンクするエージェント群を新たな A_i として更新する。
4. t がシミュレーション開始時刻(0:00)を含むまで3を繰り返す。
5. t がシミュレーション開始時刻(0:00)を含むとき、 t に対応するKeyGraphにおける A_t をシミュレーション開始時刻の情報発信源候補として出力する。

4. 情報発信源推定アルゴリズムの実行例

時間経過によるネットワーク構造の変化に注目し、時間帯毎の共起関係を分析することによる情報発信源候補推定アルゴリズム適用の実行例を示す。

4.1 手順

(1) シミュレーションによる情報伝播データの生成

都市モデルにおいて初期情報保持者となる伝達行動状態エージェントを1人設定し、1日の間に生じるロコミの情報伝播のデータを生成する。

シミュレーション設定を以下に示す。

- 都市モデル概要
 - 都市における各施設(家、会社、学校など)をスポットで表し、住民エージェント(会社員、主婦、学生など)が日常行動のルールに従ってスポット間の移動を行う過程で情報伝播現象が生じる。
- 都市の構成
 - エージェント数: 440
 - スポット数: 286
- シミュレーション期間
 - 1日(1ステップ=15min)
- ロコミの情報伝達モデル設定
 - $p_1=0.1, p_2=0.5, p_3=0.5, p_4=1.0$
 - $t_1=8h, t_2=8h, t_3=8h$

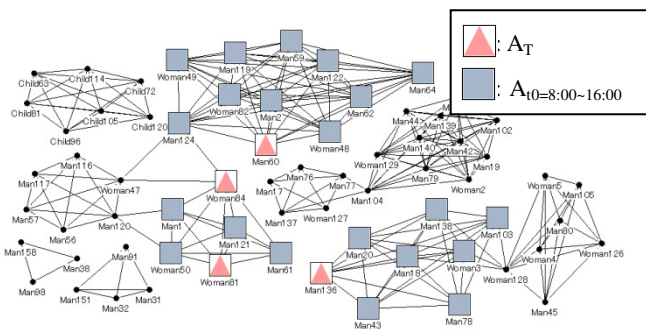


図 5 $t_0=8:00\sim 16:00$ におけるエージェントの共起情報

- シミュレーション開始時の伝達行動状態エージェント - Man1(男性就業者)

上記設定でシミュレーションを実行した結果、終了時(24:00)における伝達行動状態エージェントは 5 名であった。

(2) シミュレーション結果からの KeyGraph 抽出

シミュレーションから得られたエージェントの行動履歴における共起情報を基に KeyGraph を生成する。

関心状態から伝達行動状態への遷移期間、伝達行動状態から無関心状態への遷移期間はともに 8 時間であるため、ある時刻における伝達行動状態のエージェントが情報を受信した時間は該当時間の 8~16 時間前の期間であると予測できる。したがって、8 時間単位で分割した行動履歴データから KeyGraph を生成することで、情報の伝播過程を追うことができると考えられる。

(3) KeyGraph に基づく情報発信源推定アルゴリズムの実行

生成した KeyGraph を用いて、3.4 節の情報発信源推定アルゴリズムの手順に従い情報発信源候補の推定を行う。

4.2 結果

図 5 はシミュレーション終了時に伝達行動状態であった 5 人のエージェント A_T が情報を受信したと予測される時間帯 $t_0=8:00\sim 16:00$ の KeyGraph の一部である。この KeyGraph において A_T とリンクしているエージェント群を情報発信源候補 $A_{t_0=8:00-16:00}$ と推定している。

図 6 は $A_{t_0=8:00-16:00}$ が情報を受信したと予測される時間帯 $t_1=0:00\sim 8:00$ の KeyGraph の一部である。この KeyGraph において $A_{t_0=8:00-16:00}$ とリンクしているエージェント群を新たな情報発信源候補 $A_{t_1=0:00-8:00}$ として更新する。

図 6 の KeyGraph はシミュレーション開始時の行動履歴データを含んでいるため、この時点での情報発信源候補 $A_{t_1=0:00-8:00}$ がシミュレーション開始時の伝達行動状態エージェントの候補として導出される。

$A_{t_1=0:00-8:00}$ として 33 人のエージェントが導出された。 $A_{t_1=0:00-8:00}$ の中には Man1 が含まれており、推定した情報発信源候補の中に実際の情報発信源が含まれていた。

4.3 考察

本章において行った解析により、人口 440 人のエージェントが存在する都市において情報発信源の候補者群 $A_{t_1=0:00-8:00}$ を 33 人に絞り込むことが出来た。 $A_{t_1=0:00-8:00}$ の中には実際の発信源である Man1 が含まれていたが、33 人の A_t の中から Man1 を発信源として特定するまでには至っていない。

しかし、図 6 の KeyGraph において $A_{t_1=0:00-8:00}$ に相当するノードが存在しているクラスタに注目すると、 $A_{t_1=0:00-8:00}$ の多くは 2

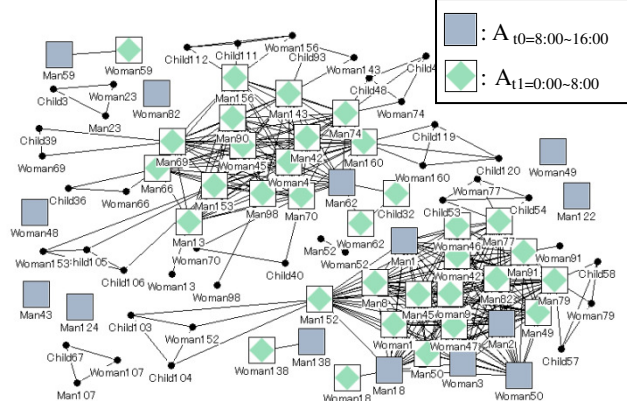


図 6 $t_1=0:00\sim 8:00$ におけるエージェントの共起情報

つのクラスタに含まれていることが観測できる。実際の発信源である Man1 はこのうち 1 つのクラスタに含まれており、このクラスタは同時刻に駅を介して出勤する就業者のコミュニティを示していた。このことから、情報発信源が所属していたコミュニティを推定するという点において本手法が有効である可能性が示唆されていると考えられる。

5. まとめ

本論文では、アドホックコミュニケーションにおけるロコミの情報伝播の要因推定手法提案のために、エージェントベースシミュレーションを用いた実際に起こりうる人の行動とそれに伴うロコミの情報伝播現象のデータの生成と、基礎的な情報発信源推定手法の提案を行った。

提案手法を用いて予備的に行った情報発信源の推定の結果から、情報発信源が所属していたコミュニティを推定するという点において本手法が有効である可能性が示唆されると考えられる。

今後は、実社会における統計データなどの実データをシミュレーションに組み込むことで、より現実に近い情報伝播現象を再現し、実用性のある情報発信源推定手法の提案をしていく予定である。

参考文献

- [Kermack 91] W.O.Kermack, A.G.McKendrick: Contributions to the mathematical theory of epidemics-I, Bulletin of Mathematical Biology, Springer, 1991.
- [前野 07] 前野 義晴, 大澤 幸生: コミュニケーションから探る組織の見えない黒幕, 人工知能学会論文誌 Vol.22, No.4, pp.389-396, 2007.
- [高橋 03] 高橋健一, 雨宮聡史, 鍾国強, 岩尾忠重, 雨宮真人: コミュニティとユーザプロファイルに基づいた情報は威信システム, 第二回情報科学技術フォーラム(FIT2003), 一般講演論文集第 4 分冊, pp.49-51, 2003.
- [林 07] 林 剛史, 木嶋 恭一: マスコミとの相乗効果に注目したクチコミマーケティングのエージェントベースモデリング, 東京工業大学修士論文, 2007.
- [田沼 07] 田沼 英樹, 出口 弘: エージェントベース社会シミュレーション言語 SOARS の開発, 電子情報通信学会誌 Vol.J90-D, No.9, pp.2415-2422, 電子情報通信学会, 2007.
- [大澤 99] 大澤 幸生, ネルス・ベンソン, 谷内田 正彦: 語の共起グラフの分割・統合によるキーワード抽出, 電子情報通信学会論文誌 Vol.J82-D-I, No.2, pp.391-400, 2007.