

日常生活での遭遇履歴を用いた ユーザクラスタリング手法

玉井 祐輔^{†1} 高田 秀志^{†2}

携帯端末の自律的な近接無線通信により情報交換を行うアドホックコミュニケーションでは、ユーザ同士がつながりを認識できないため、他のユーザとの間でグループを形成することが難しい。本稿では、アドホックコミュニケーションにおける遭遇履歴を用いて、ユーザをグループ化することを可能にするクラスタリング手法を提案する。本手法では、遭遇履歴によるユーザ間のつながり「遭遇率」として表し、クラスタリングの指標とする。また、時間帯ごとにクラスタを作成することで、人々が様々なグループに属している状況を表現できるようにする。これにより、アドホックコミュニケーションにおいてユーザ同士のつながりを反映した情報交換を行うことが可能となる。

A User Clustering Method Based on Daily Encounter History

YUSUKE TAMAI^{†1} and HIDEYUKI TAKADA^{†2}

In the ad-hoc communication where the information is exchanged autonomously between mobile terminals with short range wireless network, it is difficult for users to create groups with other users because users can not recognize the connection between users. In this paper, a user clustering method based on users' encounter history in the ad-hoc communication is proposed. This method creates clusters with the measure called "encounter probability" that represents users relations based on users' encounter history. This method can also represent the situation where people belong to several groups by creating clusters based on time window. This clustering of users enables the information exchange considering the user connections in the ad-hoc communication.

^{†1} 立命館大学大学院理工学研究科

Graduate School of Science and Engineering, Ritsumeikan University

^{†2} 立命館大学情報理工学部

1. はじめに

現在我々が用いているコミュニケーション手段は、大きく「対面型コミュニケーション」、および「非対面型コミュニケーション」の2種類に分類することができる。上記の2種類のコミュニケーション手段は、情報伝達の形態などさまざまな点で違いが存在するが、コミュニケーションを行うにあたってグループが形成され、そのグループに基づいて活動が行われているという点においては変わりがない。そのため、他者とのコミュニケーションにおいて、グループは非常に重要な役割を持っていると考えられる。

近年では、上記2種類のコミュニケーション手段に加え、機器間近距離無線通信を用いて情報交換を行うアドホックコミュニケーションに対する注目が集まってきている。端末によって自律的に情報交換が行われるアドホックコミュニケーションでは、ユーザ同士がつながりを認識することが困難である。そのため、ユーザ自身が他のユーザとの間にグループを形成することが難しく、システムによってグループ形成を支援する必要がある。例えば、街中で情報交換をするような場合に、すべての人からの情報を均等に扱うのではなく、ユーザをグループ化し、グループ内の人からの情報を上位にランキングするなどを行うことで、より効果的な情報交換を行うことが可能となる。

そこで、本稿ではアドホックコミュニケーションでのユーザの遭遇履歴を用いてユーザをクラスタリングし、グループを生成する手法を提案する。本手法では、遭遇履歴を用いて一定の時間間隔ごとのユーザ同士の関係を表す「遭遇率」を算出し、クラスタリングに用いる。遭遇率を用いてクラスタを作成することで、人々が様々なグループに属している状況を表現する。これにより、ユーザが属するグループの判断を、コミュニケーションが行われたその時々で的確に行うことが可能になる。

また、手法の有効性を検証するために、大学内でのアドホックコミュニケーションを想定したシミュレーションを行い、シミュレーションによって得た結果をもとに生成したクラスタの特徴を分析した。その結果、行動パターンの類似したユーザを同一のクラスタに分類可能なこと、および、時間帯によるグループの変化に対応できることが確認された。

2. コミュニケーション手段の現状

本節では、現状のコミュニケーション手段とその分類、および、特徴について述べる。現

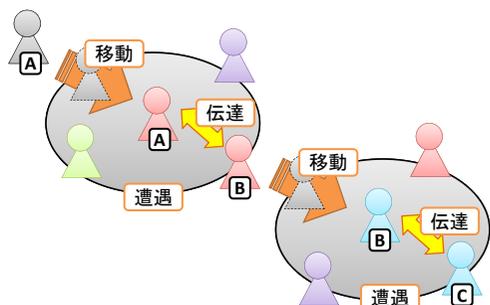


図1 対面型コミュニケーションの情報伝達の流れ

Fig. 1 The Way of Information Transmission in Face-to-face Communication

状のコミュニケーション手段は、大きく「対面型コミュニケーション」、および「非対面型コミュニケーション」の2種類に分類できる。

以下に、対面型、および、非対面型コミュニケーションの情報伝達形態とグループ形成のメカニズムについて述べた後、両者の特徴を比較する。

2.1 対面型コミュニケーション

対面型コミュニケーションは、対話に代表されるもっとも原始的なコミュニケーション手段であり、人との直接的な出会いによって成り立っている。

2.1.1 情報伝達形態

図1に対面型コミュニケーションによる情報伝達の流れを示す。

図1に示したように対面型コミュニケーションは、「移動」、「遭遇」、「伝達」という3つの過程を繰り返すことによって人から人へと情報が伝達されている。しかし、実際には対象者との間につながりを認識できた場合のみ情報の伝達が行われる場合が多く、友人や知人をはじめとする比較的親しい間柄の人同士の情報交換として用いられることが多いと考えられる。

2.1.2 グループ形成のメカニズム

市川らは1)において、日常生活におけるグループ形成活動を以下の3つの段階に分類し、それらの過程を繰り返し行うことによってグループ形成活動が行われているとしている。

(1) 対象の認識

会社などの同一のコミュニティに属する人をグループの対象として認識する。

(2) グループの形成

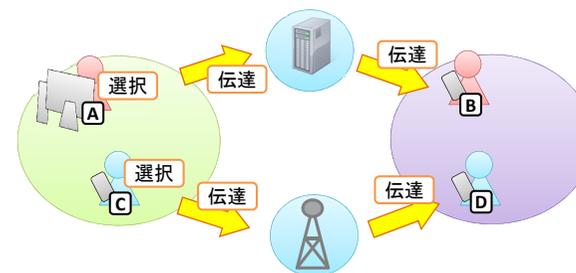


図2 非対面型コミュニケーションの情報伝達の流れ

Fig. 2 The Way of Information Transmission in Remote Communication

同一のコミュニティに属すると認識している人とのすれ違いや出会いの繰り返しのことによってグループを形成する。

(3) 情報交換

同一のグループに属する人との間で、会話が生まれ、情報交換が行われる。

ここで、市川らが1)で述べている日常生活は、本研究で述べる「対面型コミュニケーション」に非常に類似している。そのため、対面型コミュニケーションにおけるグループ形成も上記の過程によって成り立っていると考えられる。

また、対象が「同一コミュニティに属する人」に限られるため、友人や知人などの比較的親しい人との間でグループ形成が行われていると考えられる。このようにして形成されたグループは、ほとんどの場合、人によって暗黙的に認識されているのみである。

2.2 非対面型コミュニケーション

非対面型コミュニケーションは、電話やメール、チャットなどに代表されるコミュニケーション手段であり、遠隔地にいる人との間でもコミュニケーションを行うことを可能にする。以下、非対面型コミュニケーションについて詳しく検討を行う。

2.2.1 情報伝達形態

図2に非対面型コミュニケーションによる情報伝達の流れを示す。

図2に示したように、非対面型コミュニケーションでは、「選択」、「伝達」という2つの過程を経ることによってコミュニケーションが行われている。

非対面型コミュニケーションでは、遠隔地に存在する他者とコミュニケーションをとるため、コミュニケーションをとりたい相手を認識している必要がある。そのため、電話やメールなどの手段は、主に対面型コミュニケーションにおいて暗黙的に形成されたグループに属

する人同士でのコミュニケーション手段として用いられていたと考えられる。しかし、近年 mixi などの SNS (Social Networking Service) 普及により遠隔地に存在する人との間でユーザが明示的にグループを形成することが可能になり、従来に比べ非常に多くの人とコミュニケーションをとることが可能となった。

2.2.2 グループ形成のメカニズム

現在、SNS などのサービスが普及したことにより、我々は容易に「コミュニティ」と呼ばれるグループを形成することが可能となった。mixi などを利用して行われるグループの形成は、以下の 3 つの過程により行われていると考えられる。

- (1) グループの作成
趣味や嗜好、出身地などさまざまな指標をもとに、あるユーザがグループを作成する。
- (2) グループへの参加
作成されたグループに対して他のユーザが参加していくことで、グループの規模が拡大し、コミュニケーションをとることが可能になる。
- (3) 情報交換
システムによって提供される電子掲示板や、チャットシステムなどを使い、メンバーの中で情報交換が行われる。

このように、非対面型コミュニケーションでは対面型とは異なり、まず最初にグループが作成される。したがって、グループはユーザによって明示的に作成され、SNS などのサービスによって表現される。

2.3 各コミュニケーション手段の特徴比較

前節までに述べた内容をふまえて、対面型コミュニケーションと非対面型コミュニケーションの特徴の比較を行う。特徴の比較は「コミュニケーションのタイミング」、「コミュニケーションが起こる場所」、「コミュニケーションを行う対象」、および「形成されるグループの明示性」の 4 つの指標に基づいて行う。

表 1 に上記の指標に基づいて対面型および非対面型コミュニケーションの特徴を比較した結果を示す。この表からわかるように、非対面型コミュニケーションは対面型コミュニケーションに存在した制限を可能な限り排除した形態となっていると考えられる。また、対面型および非対面型のどちらのコミュニケーションにおいてもグループが形成され、そのグループを活用することでコミュニケーションが行われていることがわかる。

表 1 コミュニケーションの特徴比較
Table 1 Comparison of Communication Feature

	対面型	非対面型
コミュニケーションのタイミング	遭遇時のみ可能	いつでも可能
コミュニケーションが起こる場所	遭遇した場所	あらゆる場所
コミュニケーションを行う対象	友人や知人などの限られた人	サービスを利用可能な人
形成されるグループの明示性	暗黙的	明示的

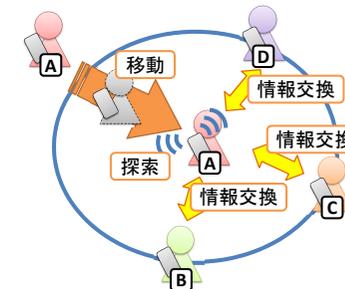


図 3 アドホックコミュニケーションの情報伝達の流れ
Fig. 3 The Way of Information Transmission in Ad-hoc Communication

3. アドホックコミュニケーション

アドホックコミュニケーションでは、機器間近距離無線通信を用いて情報交換が行われるため、無線 LAN などのインフラを必要としない。

3.1 情報伝達形態

アドホックコミュニケーションの情報伝達の流れを図 3 に示す。この図に示すように、アドホックコミュニケーションにおける情報伝達は「移動」、「探索」、および「情報交換」の 3 つの過程によって成り立っている。

ここで、「移動」はユーザによってもたらされるものであるが、その後の「探索」および「情報交換」は情報端末によって自律的に行われる場合が多い。また、「探索」はユーザが移動したタイミングで行われるものではなく、一定の間隔で周期的に行われている。したがって、図 3 に示した例の場合、A の端末が B, C, D の端末とそれぞれ情報交換を行い、その後すべてのユーザが一定期間以上その場にとどまっていた場合、再度探索がなされ、発見さ

表 2 アドホックコミュニケーションの特徴と類似するコミュニケーション
 Table 2 A Feature of Ad-hoc Communication and Similar Communication

	特徴	類似する型
コミュニケーションのタイミング	遭遇時のみ可能	対面型
コミュニケーションが起こる場所	遭遇した場所	対面型
コミュニケーションを行う対象	サービスを利用する遭遇した人	対面型 / 非対面型
形成されるグループの明示性		

れた各端末の間で再び情報交換が行われることになる。

以下で、アドホックコミュニケーションと第 2 節で述べた対面型および非対面型コミュニケーションとの比較を行い、アドホックコミュニケーションにみられる特徴の分析を行う。

3.2 アドホックコミュニケーションの特徴

2.3 節で特徴比較に用いた 4 つの指標に基づき、アドホックコミュニケーションの特徴を分析した結果を表 2 に示す。「形成されるグループの明示性」については、現在アドホックコミュニケーションの特徴を表すグループが形成されている例がないため、記載していない。

表 2 からアドホックコミュニケーションは多くの面で対面型コミュニケーションと類似していることがわかる。しかし、コミュニケーションを行う対象の面では対面型および非対面型コミュニケーションの両者の間に存在する中間的なコミュニケーション手段であると考えられる。そのため、対面型のグループ形成メカニズムをもとに、非対面型のメカニズムを併せたグループ形成を行うことで、アドホックコミュニケーションの特徴を表すグループを形成することが可能であると考えられる。

3.3 グループ形成のメカニズム

前節で述べた特徴に基づき、アドホックコミュニケーションの特徴を活かせるグループ形成メカニズムについての考察を行う。

表 2 からわかるように、アドホックコミュニケーションにおいてコミュニケーションをとる対象は実際に遭遇した人である。そのため、グループの作成も遭遇した人との間で行われることが望ましい。したがって、アドホックコミュニケーションでのグループ形成は、「同一のサービスを利用する人とのすれ違いや出会いの繰り返し」によって行うことが適切であると考えられる。しかし、アドホックコミュニケーションは端末間による自律的なコミュニケーションであるため、ユーザ自身が対象者との出会いやすれ違いを認識することは困難である。そのため、システムによってグループ形成を支援する必要があると考えられる。

そこで、本研究ではアドホックコミュニケーションの情報伝達のさいに行われる「探索」

に着目し、端末が探索し発見した他の端末の履歴をユーザ間の遭遇履歴として扱うことで、アドホックコミュニケーションにおけるユーザ間のつながりの強さを数値化する。このようにして数値化したつながりの強さに基づきシステムによってユーザをクラスタリングし、グループを作成することで、アドホックコミュニケーションの特徴を表すグループを作成することが可能であると考えられる。

4. 遭遇履歴に基づくグループ形成

本研究では、アドホックコミュニケーションにおいて端末が探索し、発見した他端末の履歴を、ユーザ間の遭遇履歴とみなし、遭遇履歴に基づいてユーザのグループを形成することを考える。本手法では、まずユーザ間のつながりの強さを数値化する。さらに、数値化されたつながりの強さをを用いて「遭遇率」を算出し、算出された遭遇率に基づきユーザのクラスタリングを行う。

以降、4.1 節で本研究で想定する情報交換のモデルについて述べた後に、4.2 節にて遭遇率の算出について述べ、4.3 節にてユーザクラスタリング手法について述べる。

4.1 情報交換モデル

アドホックコミュニケーションはサーバ等のインフラを必要としない。そのため、ユーザはアドホックコミュニケーションを用いてあらゆる場所で他のユーザと情報交換を行うことが可能である。しかし、端末に保存されている遭遇履歴をもとにユーザをクラスタリングする場合、全ての端末に保存されている遭遇履歴が必要になる。そのため、サーバ等をまったく利用することなく、全ての端末の遭遇履歴を利用したクラスタリングを行う仕組みを実現することは不可能であると考えられる。

そこで本研究では、あらゆる場所で情報交換を行うことが可能であるというアドホックコミュニケーションの特徴を失わないように、できる限りサーバとの通信を必要としないモデルを考える。そのため、サーバに対する情報の集約、クラスタリング処理、および、グループ情報の端末への送信は、ユーザが端末を使用している可能性が低い深夜から明け方などに定期的に一定の周期で行うことを想定する。このようにして端末に送信されたグループ情報は、ユーザ同士が端末を持って遭遇し、情報交換を行うさいに用いられる。

4.2 遭遇率の算出

本研究では、探索によって発見した端末の識別子とその時刻を端末に保存し、ユーザ間の遭遇回数を遭遇履歴として扱う。さらに、この遭遇履歴から一定の時間間隔によって与えられた期間におけるユーザ同士の遭遇頻度を表す指標である「遭遇率」を定義し、以下の式に

よって算出する．

$$(\text{ユーザ } i \text{ とユーザ } j \text{ の遭遇率}) = (C_{ij}|t_s, t_e) * \frac{1}{(t_e - t_s) * \frac{60}{d_s}}$$

ここで，

- d_s : アドホックコミュニケーションにおける端末探索間隔 (単位: 分)
- t_s : 期間開始時刻 (単位: 時)
- t_e : 期間終了時刻 (単位: 時)
- $(C_{ij}|t_s, t_e)$: t_s から t_e までの期間におけるユーザ i とユーザ j の遭遇回数

である．例えば，端末探索間隔が 15 分，開始時刻が 14 時，終了時刻が 16 時，14 時から 16 時のユーザ A と B の遭遇回数の合計が 3 回であった場合の遭遇率は 0.375 となり，この値を 14 時から 16 時におけるユーザ A と B の関係の強さとする．これは，ユーザ A とユーザ B が 14 時から 16 時の期間の 37.5% の期間遭遇していたことを表している．したがって，この遭遇率を用いることにより特定の時間帯におけるユーザ間のつながりの強さを算出することが可能となる．

このように時間間隔によって定められた一定の期間における遭遇率を用いてクラスタを作成することで，人が様々なグループに属している状況を表現することが可能になる．

4.3 遭遇率に基づくユーザクラスタリング

前節で述べた手法で算出した遭遇率をユーザ間の関連の強さとして扱い，ユーザのクラスタリングを行う．現在，クラスタリングを行う手法は数多く研究されており²⁾³⁾，それらは大きく「階層的クラスタリング手法」および「分割最適化クラスタリング手法」の 2 種類に分類できる．上記の手法はどちらかが常にすぐれているわけではなく，クラスタリングを行うデータの特徴などによって使い分けの必要があるとされている．

特に，分割最適化クラスタリング手法で多く用いられている k-means 法⁴⁾ では，クラスタが超球状の形状であり，クラスタ中の要素数がどれも等しいということを暗黙的に仮定している．そのため，上記の仮定に反する構造の抽出は困難であるとされており，Rajeev らによって実際に仮定に反する構造を持つデータセットに対して k-means 法を適用した場合の結果が示されている⁵⁾．

本研究ではユーザのクラスタリングを目的とするため，生成されるクラスタ中の要素数が等しくなるという仮定を満たしているとは考えづらい．そのため，今回ユーザクラスタリングを行う手法として階層的クラスタリング手法を用いる．

階層的クラスタリング手法によるクラスタリングは，以下のアルゴリズムによって行わ

れる．

- (1) データ集合に存在する全てのデータを要素 1 つのクラスタとする．
- (2) 全てのクラスタの中から，距離がもっとも近いクラスタ対を探し，1 つのクラスタに併合する．
- (3) (2) をデータ全体が 1 つのクラスタに併合されるまで繰り返す．

階層的クラスタリング手法によってユーザをクラスタリングすることで，最終的にユーザは 1 つのクラスタに併合される．そこで，本手法では任意に与えられる閾値を用いて，1 つに併合されたクラスタから適切なクラスタを抽出する．

5. シミュレーション実験

本節では，手法の有用性を検証するために行ったエージェントベースシミュレーションについて述べる．

5.1 シミュレーションモデル

本シミュレーションは，学部 3 回生から修士 2 回生までの学生を対象とした大学内でのアドホックコミュニケーションを想定し，半期の授業が行われる期間である 16 週間を実施期間とした．

学部 3 回生や 4 回生などは，修士の学生と比べて時間帯によるグループの変化が比較的頻繁に起こっていると考えられる．一方で修士 2 回生はほぼ研究室内にいるため，時間帯によるグループの変化がめったに起こることはない．したがって，上記の環境でシミュレーションを行うことで，時間帯によるグループの変化に違いのあるユーザが同時に存在する環境において，適切なグループ形成が行えているかを検証することが可能であると考えられる．

5.2 使用言語

今回シミュレーションモデルを構築するにあたり，シミュレーション言語として SOARS⁶⁾ を用いた．SOARS では，現実世界における場所を「スポット」として表現し，エージェントに対して行動ルールを定めてスポット間を移動させることで，現実世界での人の場所移動を再現した社会的なシミュレーションを容易に構築することができる．

本シミュレーションでは，学内に存在する教室などの場所，および，学生が移動に用いると考えられる駅やバス停などをスポットとし，学生をエージェントとして表現した．

5.3 事前調査

SOARS を用いて今回検証に用いる学内でのアドホックコミュニケーションを想定したモデルを構築するにあたり，エージェントのルール，および，エージェントが訪れるスポット

表 3 場所の合計と研究室への来訪頻度
Table 3 The Number of Places and
Visiting Probability of Lab

	訪れる場所の合計数	研究室を訪れる頻度
学部 3 回生	10ヶ所	32%
学部 4 回生	3ヶ所	75%
修士 1 回生	4ヶ所	80%
修士 2 回生	1ヶ所	85%
合計	18ヶ所	

表 4 自宅の場所の割合
Table 4 Location of Home

地域	割合
神戸	10%
大阪	25%
京都	10%
南草津	55%

を作成する必要がある。そこで、同研究室の学生に対して日常の行動を問う事前調査を行い、収集したデータをもとにルールを決定した。

5.3.1 調査内容

協力者に普段の生活での 1 週間の行動を想起してもらい「学内で訪れる場所」、「自宅を出発する時間や帰宅時間」、「教室や研究室を訪れる時間や出る時間」、および「自宅の地域」の調査を行った。

5.3.2 調査結果

上記の内容の調査を同研究室の学生に対して行った結果、23 名の学生からの回答を得ることができた。

回答により得たデータのうち回生によって大きく異なる結果になったものとして、訪れる場所の合計数と研究室を訪れる頻度を表 3 に、回生に依存しないものとして、自宅の地域の割合を表 4 に示す。ここで、研究室を訪れる頻度とは各回生の学生が平日 5 日間の中で研究室を訪れた日数を合計し、確率として求めたものである。また、自宅の地域は「神戸」、「大阪」、「京都」、「南草津」の 4 種類に分類した。

5.3.3 基本設定

本シミュレーションでは各エージェントに「B3」、「B4」、「M1」、および「M2」の 4 種類の役割を与え、立命館大学情報理工学部情報システム学科の 1 学年の学生数および、同大学大学院理工学研究科の 1 学年の学生数を参考に役割「B3」「B4」のエージェント各 150 名、「M1」「M2」のエージェント各 40 名を用意した。

また「教室」、「駅/バス停」、および「研究室」の 3 種類のスポットを用意し、種類「教室」のスポット数は、前節のアンケート結果の 18ヶ所から研究室を除いた 14ヶ所とした。また、種類「駅/バス停」は学生の自宅の地域 4ヶ所にバス停を加えた 5ヶ所とし、種類「研究室」に関しては、エージェント数 380 を均等に分割できるように、これを割り切れる数

として 10ヶ所とした。

5.3.4 エージェントの設定

各エージェントを作成するさいに、エージェントの行動決定に用いる以下のパラメータを設定した。

- 授業リスト：エージェントが受講する授業のリスト。事前調査をもとに設定。
- 研究室来訪確率：エージェントが研究室を訪れる確率。表 3 をもとに設定。
- 研究室来訪時刻：研究室を訪れる時刻。事前調査をもとに設定。
- 研究室出発時刻：研究室を出て帰宅する時刻。事前調査をもとに設定。
- 自宅の地域：普段生活している自宅の地域。表 4 をもとに設定。
- 研究室：エージェントが所属する研究室。各役割のエージェントを均等に分配。

また、エージェントの移動にさいし、人の行動の不確定さを表現するために、上記のパラメータに加え以下に示す 3 種類のパラメータのうち 1 つを「タイプ」として設定した。各パラメータが設定される割合は正規分布に従い「TypeA」を 80%、「TypeB」を 15%、「TypeC」を 5%とした。

- TypeA：目的時刻の 10 分前に目的地につくように移動を行うエージェント。
- TypeB：目的時刻に目的地につくように移動を行うエージェント。
- TypeC：目的時刻の 10 分後に目的地につくように移動を行うエージェント。

さらに、上記のタイプに加え、エージェントに毎朝目的時刻からどの程度遅れて行動するかを決定を行わせることで、行動の不確定さを表現する。この決定も正規分布に則るものとし、80%の確率で目的時刻どおりに、15%の確率で目的時刻から 10 分遅れて、5%の確率で目的時刻から 20 分遅れて行動するようにした。

このようにして、機械的に同じ時刻に行動を開始するのではなく、その日その日によって少しずつ異なった時刻にエージェントが行動を開始するようにした。

上記のパラメータを用いてエージェントは毎朝その日の行動パターンを決定し、スポット間を移動する。

5.3.5 環境設定と遭遇履歴の取得

本シミュレーションでは、アドホックコミュニケーションにおける端末探索間隔を 10 分として設定し、10 分毎の端末探索結果を遭遇履歴として用いる。

ここで、同一スポット上に存在するエージェントは探索により必ず発見可能であるとし、10 分毎に各スポットに存在するエージェントを観測した結果を遭遇履歴とした。

5.4 遭遇履歴に基づくクラスタリング

シミュレーションより得た遭遇履歴に対して、提案手法によるクラスタの抽出を行った。また、今回はクラスタを抽出するさいの閾値として、全ユーザの遭遇率が、以下に示す式(1)で求められる遭遇率平均 $E(D)$ を中心とした正規分布に基づくと仮定し、式(2)によって算出される閾値 L を用いた。

$$E(D) = \frac{1}{l_D} \sum_{x_1 \in D} \sum_{x_2 \in D} P_E(x_1, x_2) \tag{1}$$

$$L = E(P_E) + \sigma(P_E) \tag{2}$$

ここで、

- $P_E(x_1, x_2)$: ユーザ x_1 とユーザ x_2 の遭遇率
- l_D : データセット D のリンク数

である。

正規分布にもとづいたデータセットに対して式(2)の閾値を用いることで、併合されたクラスタの遭遇率平均が、データセット全体のおよそ 68%以上であった場合に、クラスタとして抽出されることになる。

5.5 クラスタリング結果と考察

シミュレーションにより得た履歴をもとに作成したクラスタに対して、ユーザがどのように分類されているかの調査を行った。

調査は、生成されたクラスタを階層的な情報を矩形領域の集合として 2 次元画面に展開する手法であるツリーマップ⁷⁾を用いて表現することで行った。ツリーマップを用いることで、構造が複雑化した場合においても限られた領域内に表示することが可能である。

5.5.1 生成されたクラスタに見られる特徴と妥当性

図 4 に生成されたクラスタをツリーマップを用いて表現し、各エージェントを表す一つの矩形を所属研究室によって色分けした場合の結果を示す。図の太線は、抽出されたクラスタの境界である。

図 4 から、同一の研究室に属するユーザがクラスタとして分類されていることがわかる。したがって、本手法を用いてクラスタを生成することによってユーザを適切なグループに分類することが可能であると考えられる。

5.5.2 時間帯によるグループ変化の表現

時間帯によるグループの変化を表現できているかを調査するために、期間 10 時~12 時において作成されたクラスタ、および、期間 16 時~18 時において作成されたクラスタに対

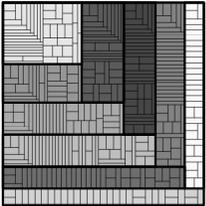


図 4 クラスタのツリーマップ表現
Fig. 4 Tree-Map Representation of Clusters



図 5 期間 10 時~12 時のクラスタのツリーマップ表示
Fig. 5 Tree-Map Representation of Clusters
from 10AM to noon

図 6 期間 16 時~18 時のクラスタのツリーマップ表示
Fig. 6 Tree-Map Representation of Clusters
from 4PM to 6PM

し、各エージェントを「所属研究室」、および、「役割」をもとに色分けし、ツリーマップとして表現した。図 5 に期間 10 時~12 時の場合の結果を、図 6 に期間 16 時~18 時の場合の結果を示す。

図 5 および図 6 より、時間帯によって異なる構成のクラスタが生成されていることが確認できる。また、図 5 において所属研究室により色分けした場合ではばらつきが見られるクラスタが存在するが、学年によって色分けした結果を見ると、同様の役割を持つエージェントが同一のクラスタに属していることがわかる。したがって、時間帯によって適切な指標に基づいたクラスタの生成が行えていると考えられる。

5.6 考察

今回、シミュレーションによる実験を行うにあたり、本手法を用いることで時間帯によって変化するグループを表現することが可能であるか、またそのようにして表現したグループの結果が妥当なものであるかの検証を行うことに重点を置いた。そのため、あえて多様な種類のエージェントが存在する環境ではなく、学生という 1 つの種類に限定したモデルを構築し、その中で検証を行った。その結果、ある程度妥当なクラスタの生成を行うことが可能なこと、および、時間帯によるグループ形成が可能なことが確認できた。

今後は、さらに多様な種類のエージェントが存在する環境での実験を行い、多様な種類の場所へ人が訪れるような状況においても適切にクラスタを抽出することが可能であるか、また、時間帯によるグループの変化を表現可能であるかなどの検証を行っていく必要があると考えられる。

6. 関連研究

本手法の関連研究として、武井らの「アドホックコミュニケーションにおける口コミの情報伝播の要因推定」⁸⁾ や、茂木らの「移動情報を活用したモバイル・コミュニティ形成機構」⁹⁾ があげられる。

武井らは 8) で、アドホックコミュニケーションにおける情報の発信源を特定するために、アドホックコミュニケーションでの行動履歴データからネットワークの抽出を行って利用している。行動履歴からネットワークを抽出することは本手法と類似していると考えられる。しかし、武井らの手法では発信源の特定を目的としているため、実際に抽出されたネットワークの特徴よりも、ユーザが細かい単位のコミュニティに属することを重視している。そのため、パラメータを変更してネットワークの抽出を行うなどといったことは行われていない。

また、茂木らは 9) において、現実世界を移動するユーザを想定し、そのユーザの移動情報を活用することで「モバイル・コミュニティ」と定義するコミュニティを提供することを目的としている。そのため本手法と想定するユーザや環境は同様のものであり、従来までのコミュニティに対して新たに現実世界の情報を取り入れようとしている点においても類似していると考えられる。

しかし、茂木らの手法ではコミュニティをユーザの意図によって生成されるものとしているため、コミュニティの生成はユーザによって能動的に行われるものである。一方、本手法においてグループは、端末によって自律的に行われる探索の履歴から生成される。そのため、本手法とは生成されるコミュニティの特徴が異なると考えられる。

7. おわりに

本論文では、アドホックコミュニケーションにおける遭遇履歴に基づいてユーザのクラスタリングを行い、グループを作成する手法を提案した。

本手法では端末の探索履歴をユーザ間の遭遇履歴として扱い、遭遇履歴をもとに算出した遭遇率に基づいてユーザクラスタリングを行う。また、人が様々なグループに属している状

況を考慮するために、一定の時間間隔ごとの遭遇率を算出する。そうすることで、ユーザの 24 時間の行動から時間帯によってさまざまなクラスタを作成することが可能となる。

また、本手法の有効性を検証するために学内でのアドホックコミュニケーションを想定したエージェントベースシミュレーションを行い、その結果に対して本手法を適用し、グループの作成を行った。その結果、ユーザを適切なグループに分類することが可能なこと、および、時間帯によってグループを生成することで、人の日常生活で起こっているグループの変化を表現可能であることが確認できた。

今後は、より多様な種類の人が存在する環境にて本手法を適用し、その有効性および妥当性の検証を行っていく必要があると考えられる。

参考文献

- 1) 市川裕介, 針生剛男. 動的に変化する階層化コミュニティ支援サービスの研究. 情報処理学会第 55 回全国大会講演論文集, Vol.55, No.4, pp. 139-140, 1997.
- 2) D.PELLEGRINI. X-means : Extending kmeans with efficient estimation of the number of clusters. *ICML2000*, Vol.17, pp. 727-734, 2000.
- 3) 吉田尚史, 岡子泰三, 清水康, 北川高嗣. ドキュメントデータ群を対象とした文脈依存動的クラスタリングおよび意味的データマイニング方式. 情報処理学会論文誌. データベース, Vol.41, No.1, pp. 127-139, 20000215.
- 4) J.A. HARTIGAN. A k-means clustering algorithm : algorithm as 136. *Appl. Stat.*, Vol.28, pp. 126-130, 1979.
- 5) Sudipto Guha, Rajeev Rastogi, and Kyuseok Shim. Cure: an efficient clustering algorithm for large databases. *Information Systems*, Vol.26, No.1, 2001.
- 6) 田沼英樹, 出口弘, 清水哲男. Soars : 新しいエージェントベースシミュレーション言語の開発. 情報処理学会論文誌. プログラミング, Vol.46, No.6, p.63, 2005.
- 7) Brian Johnson and Ben Shneiderman. Tree-maps: a space-filling approach to the visualization of hierarchical information structures. In *VIS '91: Proceedings of the 2nd conference on Visualization '91*, pp. 284-291, Los Alamitos, CA, USA, 1991. IEEE Computer Society Press.
- 8) 武井瑞希, 井上大輔, 新田克己. アドホックコミュニケーションにおける口コミの情報伝播の要因推定. 人工知能学会全国大会 (第 22 回) 論文集, No. 3C2-02, 2008.
- 9) 茂木信二, 堀内浩規, 小田稔周. 移動情報を活用したモバイル・コミュニティ形成機構. 情報処理学会研究報告. ITS, [高度交通システム], Vol. 2000, No.42, pp. 13-18, 2000.