

匿名型 P2P 通信における 流通情報の信頼度評価に関する検討

奥居 哲₁₎₂₎ 鈴木 裕利₁₎₂₎

1)中部大学情報科学研究所 2)中部大学工学部

E-mail: {okui, yuris}@cs.chubu.ac.jp

あらまし 匿名型Peer-to-Peer通信において不確実情報が蔓延する問題を分析するために、ピアが相互の信頼性を評価する際にとる行動を収集する実験システムを構築した。予備実験の結果、正しい情報に到達するピアほど、より多くのピアと接触を持つ傾向が観察された。一方、メッセージの受信数と正答率との間に明白な関係は認められなかった。

Avoiding unreliable messages in anonymous peer-to-peer communication

Satoshi OKUI₁₎₂₎ Yuri SUZUKI₁₎₂₎

1) Research Institute for Information Science, Chubu University

2) College of Engineering, Chubu University

E-mail: {okui, yuris}@cs.chubu.ac.jp

Abstract: Anonymous peer-to-peer networks tend to be useless owing to thousands of unreliable messages spreading over them. To approach this problem, we introduce a laboratory system, and examine how peers evaluate each other. Our preliminary experiments have so far indicated that the peers getting correct answers make more contacts, rather than receiving more messages.

1. はじめに

本プロジェクトは、各種モバイル機器を利用した Peer-to-Peer(以下 P2P)型のネットワークを構成し、それを介した地域住民間における、災害情報の収集と提供を可能にするシステムの構築を目的としている。

これまでに、匿名型 P2P メッセージ交換システムの基本部分の構築・実装を終え、現在、予備実験を実施中である。

本稿では、この基本システムの概要と、これを用いた予備実験について報告する。以下、2章では、本研究が対象とする「匿名型 P2P メッセージ交換」について説明し、3章では、構築したメッセージ交換システムの概要と特徴につい

て解説する。さらに、4章で予備実験の結果に関するこれまでの分析状況について報告し、最後に、5章で今後の計画について述べる。

2. 匿名型 P2P メッセージ交換

2.1 P2P システム[1][2][3]

P2P とは、「ピア(peer)と呼ばれる対等なネットワークエージェント間において、直接的に情報がやりとりされるネットワークサービスの形態」を指す言葉である。ここで「対等に」とは、クライアント・サーバ(Client-Server: 以下 CS)方式における情報の提供者(サーバ)と要求者(クライアント)のように明確な役割分担が存在し

ないという特徴を意味する。また「直接的に」とは、CS方式におけるサーバのような、クライアント間の通信を仲介するものを持たないという特徴を意味する。

2.2 匿名性

前項で述べたP2Pの形態は、ネットワークの障害や負荷集中によるサービス停止に陥りにくいという利点を持つが、その一方で、一元管理可能なCS方式に比べて、情報管理が困難になると考えられる。特に、情報の発信源を特定することが難しく、匿名性が高くなるといえる。この特徴は、プライバシー保護の観点からは利点ともいえるが、犯罪の温床となる危険性も高い。また、本研究が想定する災害時の情報伝達においては、過失や悪意による流言蜚語の発生が容易に予想される。そこで、信頼できるピアと信頼できないピアとを、どのように識別するかが重要な課題となる。この識別のために、指標として何が有効かを明らかにし、利用者が適切に流通情報に対して判断をするための手段を提供することが、本研究の中心課題となる。この課題に関して、いくつかの研究が行なわれているが、まだ実験段階であるといえる[4][5]。

2.3 純粋型と混成型

P2Pは、さらに、純粋型と混成型に分類される(図1参照)。純粋型は機能的にも役割的にも差別がない純粋に対等なピアからなるタイプである。これに対して混成型は、一部の機能や役割を分担するタイプである。例えば、ピアのアドレスを分配するためのピアや、情報の索引機能を持つピアを有する場合等である。これらは、それぞれの形態の代表的アプリケーションの呼称をとって、前者が「Gnutella型」、後者が「Napstar型」とも呼ばれる。

2.4 共有型と交換型

P2Pの形態は、情報の共有と交換の違いによっても分類される。

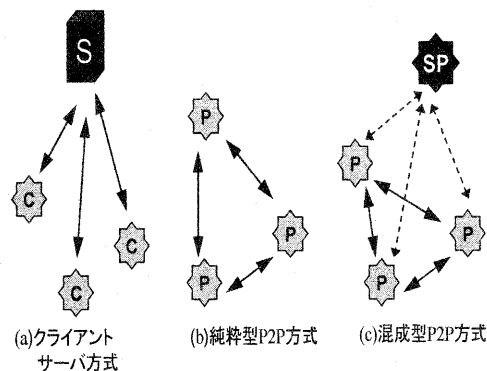


図1. ネットワークシステムの構造の種類

共有方式は、ピア間で、情報の同期をとり、あたかも共通の情報を共有しているようにイメージされるタイプである。情報の一元化が要求されるグループウェア等はこの方式を採用している[6]。しかし、トラフィックが大きくなるため、広域の通信に適切な方式とはいえない[7]。

次に、交換方式は、必要に応じて非同期に情報を交換するタイプである。ピア毎に所持する情報は異なるが、インターネットでの不特定多数による広域の通信が可能となるタイプである。Gnutella等のファイル交換システムは、この方式を採用している。

2.5 匿名型P2Pメッセージ交換

以上のように、P2Pシステムは、いくつかの方式に分類されるが、本研究においては、災害時の広域・不特定多数の情報交換に適合する、匿名の交換型システムを対象とする。そして、情報収集の容易性等から、混成型のシステムを利用する。また、次章で述べる本研究の実験システムにおいては、主として交換されるデータは会話等のテキスト情報(以下:メッセージ)とする。

3. 実験システム

本章では、本プロジェクトの実験環境として利用するために実装したメッセージ交換システムについて述べる。

3.1 システムの概要

本システムは、不特定多数の匿名の利用者がネットワーク上でメッセージの交換を行う際の行動を収集・記録する実験環境である。50名前後の利用者が一度に80文字以内程度の短いメッセージを交換するような状況での使用を主に念頭において設計されている。その基本的な構成要素はピア(peer)とメッセージ(message)である。これについては3.2.で述べる。

本システムは、その主機能であるメッセージの送信に加えて、会話の相手を探索する機能、会話の相手を選択する機能、相手の信頼度を評価する機能、評価値を交換する機能を備えている。これらは、それぞれコマンドとして利用者に提供される。これらの機能については3.3.で述べる。

コマンドの使用状況(以下：行動記録)は、実験・分析のため、収集・保存される。行動記録の内容については3.4.で述べる。

3.2. システムの基本構成

3.2.1 ピア

本システムの中心的な構成要素である。以下の2種類のピアが存在する。

(1)汎用ピア(general peer)

システムの利用者によって操作され、メッセージの交換を担うエージェントである。通常、単にピアと呼ぶ。

ピアはアドレスとピア識別子を持つ。前者のアドレスとは、ネットワーク上で通信相手を特定するための情報である。後者のピア識別子とは、ピアの同一性を保持するために設けられた各ピア固有の情報である。

ピアはアジェンダ(agenda)とよぶハッシュ表を持つ。アジェンダは、周辺のピアに関するさまざまな情報を恒久的に記憶・管理するのに用いられる。本稿の予備実験では、特に信頼度の評価値を記憶・管理するのに用いられる。アジェンダは永続化されており、ピアを再起動した

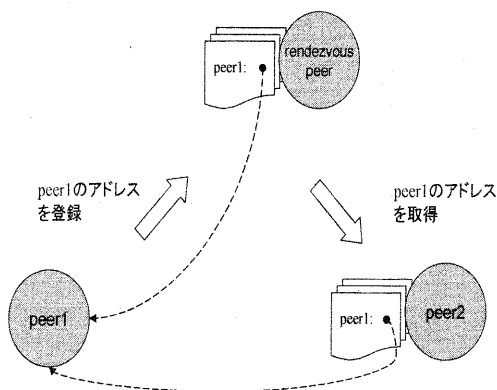


図2. ランデブー・ピアによるアドレスの配送

ときにも前回利用したアジェンダが用いられる。

(2)ランデブー・ピア(rendezvous peer)

汎用ピアのアドレスを配信するサーバ・ピアである。それ自身はメッセージ交換を行わない。ローカル・ネットワーク上に1個以上存在し、そのアドレスは周辺のピアにあらかじめ周知されている。3.3.1で詳細を述べる。

3.2.2 メッセージ

ピアによって交換される情報であり、ヘッダとボディからなる。ボディは、伝達内容である任意のテキスト情報からなる。ヘッダは、現行システムでは以下の情報を含んでいる。

- ・ 配信元のピア識別子
- ・ 配信先のピア識別子
- ・ メッセージ識別子

ここで、メッセージ識別子とは、メッセージを唯一に識別するための情報である。

3.3. コマンド

本システムが備えるコマンドのうち、4章の予備実験で使用する主要な5種類のコマンドについて述べる。

3.3.1 周辺のピアの探索(FIND コマンド)

周辺のピアのアドレスを探索するために、ランデブー・ピアへ問い合わせを行うコマンドである。

ピアは起動時に既知のランデブー・ピアの1個に対して自分のピア識別子とアドレスを登録

する。そこで、必要に応じてFINDコマンドを実行してランデブー・ピアに問い合わせを行うことにより、周辺のピアのアドレスを取得できる。例として図2に、ランデブー・ピアを通じてpeer1のアドレスがpeer2に配信される仕組みを示す。

なお、探索によって取得したアドレスはキャッシュされる。したがって、ランデブー・ピアへの問い合わせは頻繁に行う必要はない。ランデブー・ピアとの通信に障害が生じて、各ピア間の通信は、キャッシュに残されたアドレスの利用により、継続が可能である。

3.3.2 宛先の設定

(DESTINATION コマンド)

メッセージ送信コマンドの実行に先立って宛先をセットするコマンドである。

3.3.3 メッセージの送信(SEND コマンド)

メッセージを作成し送信するコマンドである。

3.3.4 信頼度の評価と更新

(UPDATE コマンド)

他のピアに対する主観的な信頼度を保存し、更新する機能である。信頼度とは、利用者自身が他のピアをどの程度信頼しているかを、百分率で主観的に評価した値である。保存した信頼度は随時更新できる。更新とは、新しい評価値の入力が生ずると、保存されている旧信頼度と入力された新信頼度の加重平均をとり、その値を新しい信頼度として保存することである。ここで使用する加重平均の重み、および、信頼度の初期値は実験毎に設定できる。

3.3.5 信頼度の問い合わせ(ASK コマンド)

他のピアが周辺のピアをどのように評価しているか(言わば「近所の評判」)を知るための機能である。

初めてメッセージを交換するピアに対する信頼度の評価は非常に難しいと考えられる。このような場合に、対象とするピアを、「他のピアがどの程度信頼しているか」の情報を得るために用いるのが本コマンドである。

本コマンドは、問合せ先ピアと信頼度を知りたい対象ピアの、両者の識別子を指定して使用する。問い合わせの結果は画面に表示される。このとき、問い合わせ元のピアのアジェンダに保持されている信頼度は自動的に更新されない。表示結果に信頼がおけると利用者が判断した場合のみ、UPDATEコマンドをあらためて明示的に実行する。

3.3.6 信頼度の表示(LIST コマンド)

アジェンダに保存されている現在の信頼度情報を一覧表示するコマンドである。

3.4 行動記録

現行システムで収集可能な行動記録には以下の内容が含まれる。

- (1)交換したメッセージの内容
- (2)送信コマンドを実行した時刻と送信先のピア識別子
- (3)探索コマンドを実行した時刻と更新先のランデブー・ピアのアドレス
- (4)信頼度更新コマンドを実行した時刻、更新対象のピア識別子、および入力した評価値
- (5)信頼度問い合わせコマンドを実行した時刻と問い合わせ先のピア識別子、および問い合わせ対象のピア識別子
- (6)メッセージを受信した時刻と送信元のピア識別子
- (7)信頼度表示コマンドを実行した時刻

4. 予備実験

本章では、前述した本研究の基本システムを使用した予備実験について述べる。

4.1 実験内容

実験は、匿名(ハンドルネーム)を用いる被験者が基本システムを用いて会話(メッセージ交換)を行なうものである。その際の被験者の行動記録を収集する。被験者は中部大学工学部3年生18名であり、実験環境として中部大学工学部デジタルラボラトリのマシンを使用し、

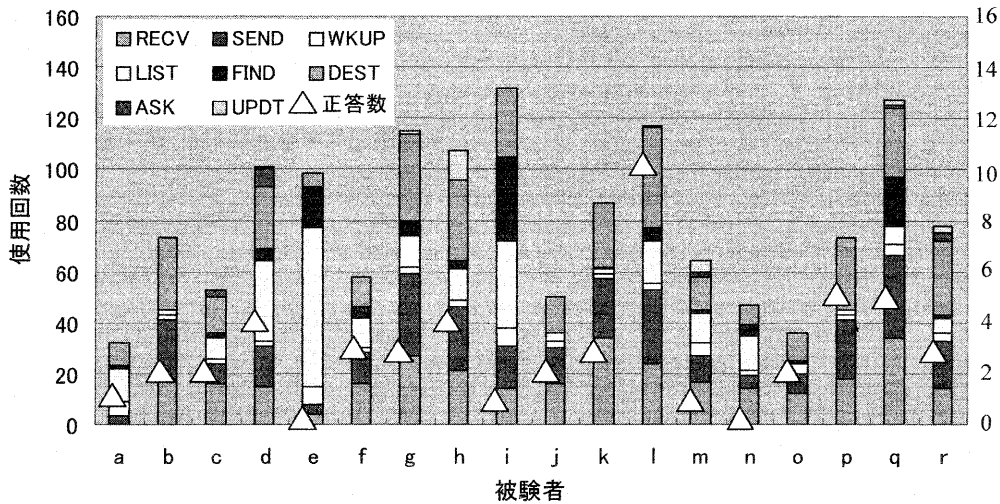


図4. 実験 I における被験者別使用コマンドの内訳と正答数

以下の実験課題を実施した。

・ピア特定課題

被験者に無作為に割り当てたハンドルネームを用いて会話を実施する。その会話内容と信頼度情報から、「各ハンドルネームを使用している被験者が誰なのか」を特定する課題である。会話の内容や信頼度の評価方法に制限は設けない。つまり、真実と異なる発言や評価が許される。実験時間は30分である。

・行動記録の内容

- (1)メッセージを送信した対象ピアのハンドルネームとメッセージの内容
- (2)メッセージを受信した対象ピアのハンドルネームとメッセージの内容

(3)他のピアに対する信頼度評価の履歴

(4)信頼度問合せの履歴

(5)使用コマンドの履歴

4.2 実験結果

実験の結果,1471 件の行動を記録し、現在分析中である。そこで本稿では、実験結果の正答率に着目した考察について、途中経過を簡単に報告する。

図4に実験 I における被験者別の行動（使用コマンド）の内訳と正答数を示す。これらの行動履歴と正答数の相関に着目し検討した結果、以下の傾向が観察された。

- (1)送受信数(SEND コマンドと RECV イベントの合計数)が多いほど正答数が増える傾向があ

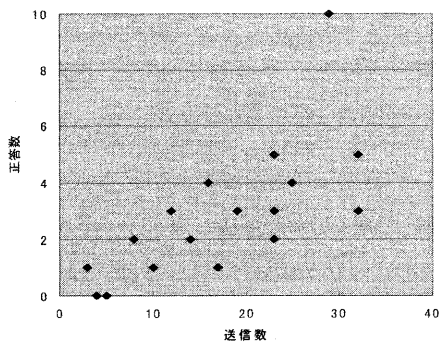


図5. 被験者別の送信数と正答数

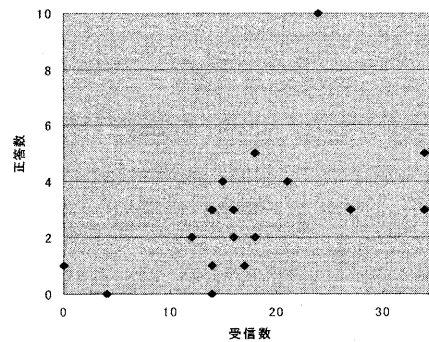


図6. 被験者別の受信数と正答数

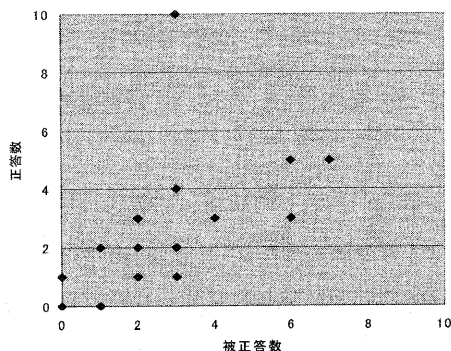


図7. 被験者別の被正答数と正答数

り、さらに、その傾向は送信数との関係においてより強くなる。図5に正答数と送信数、図6に正答数と受信数の関係を示す。

(2)正答数(他ピアを特定した数)の多い被験者は、被正答数(他ピアから特定された数)も多くなる傾向がある。図7に正答数と被正答数の関係を示す。

5. 今後の計画

前章で述べた予備実験の考察をふまえ、今後は、以下のように研究を進める予定である。また、実験結果については、人工社会における情報の流通という視点から、シミュレーションを利用して行なわれている研究の成果との比較も検討している[8][9]。

5.1 信頼度の指標としてのメッセージ

交換数

予備実験の結果から、正答率の高いピアはより多くのメッセージを交換していることがわかる。このことから、メッセージ交換の総数を、ピアの信頼度の指標として利用する可能性が考えられる。そこで、基本システムにピアのメッセージ交換総数を計量し公開する機能を取り入れ、メッセージ交換総数と信頼度評価との相関を調べる実験を実施する。

5.2 信頼度の指標としての連結数

予備実験の結果から、正答率の高いピアはより多くのピアとコンタクトを持っていることが観

察される。そこで、基本システムにピアが接触を持った相手ピアの数(連結数)を計量し公開する機能を取り入れ、連結数と信頼度評価との相関を調べる実験を実施する。また他のピアの仲介によって間接的に接触をもった相手の数(n 次連結数)に対しても同様の実験を実施する。

謝辞：実験に協力頂いた中部大学工学部デジタルラボラトリの関係者と被験者各位に感謝する。

参考文献

- [1] P2Pコンピューティング, 伊藤直樹, ソフト・リサーチ・センター, 東京, 2001.
- [2] P2Pインターネットの世紀, 河内正夫(監修), 小柳恵一(編著), 電気通信協会, 東京, 2002.
- [3] 山崎重一郎, "P2Pネットワークシステム," 人工知能学会誌, vol.16, no.6, pp.834-840, Nov.2001.
- [4] 臼井幸弘, 高橋寛幸, 吉開範章, "ネットコミュニティにおける信頼評価法としての評判システムの効果に関する実証的考察," 電子情報通信学会研究報告, vol.102, no.505, pp.19-24, Dec.2002.
- [5] 荒牧久美子, 麓真祈子, 箕浦美智子, 佐藤浩史, 河原正治, "超流通型P2Pアプリケーションにおける認証について," 情報処理学会研究報告, vol.2002, no.117, pp.29-36, Dec.2002.
- [6] Edwards. J : Peer-to-Peer Programing on Groove, Addison-Wesley, 2002.
- [7] 鳥居大祐, 板倉陽一郎, 田中裕一郎, 横澤誠, 篠原健, "Peer-to-Peerアーキテクチャのメッセージ共有への応用," 情報処理学会研究報告, vol.2003, no.1, pp.30-35, Jan.2003.
- [8] コンピュータの中の人工社会(マルチエージェントシミュレーションモデルと複雑系), 山影進, 服部正太(編), 共立出版, 東京, 2002.
- [9] 長谷川敦士, 植田一博, "人工社会における発生と消滅-インターネット上のPeer-to-Peerコミュニケーションにおける情報伝播の影響-, " 認知科学, vol.8, no.4, pp.417-430, Dec.2001.