

L_013

スモールワールドネットワークにおける最適な検索手法の考案に向けた IRCネットワークの調査

Survey of IRC network for design of the best retrieval technique in small-world network

長谷川 卓*
Suguru Hasegawa

松澤 智史†
Tomofumi Matsuzawa

武田 正之†
Masayuki Takeda

1. まえがき

インターネットの普及により、ネットワークを介したコミュニケーションが活発に行われるようになった。また、近年の計算機の発達と通信速度の大幅な向上によりサーバを介さずにコミュニケーションを行う P2P ネットワークに注目が集まるようになってきた。情報を一元で管理するサーバが存在しない P2P ネットワークにおいて、ノード間のコミュニケーションを円滑にするために必要な要素として、目的のピアを短時間で検索する方法の考案があげられる。本稿では small-world 構造を持つネットワークに着目し、構造を生かしたネットワークの検索手法を考案するためのシミュレーションを行い、その結果から検索方法について考察する。シミュレーションに用いたネットワークは IRC [2] ネットワークを用いた。IRC ネットワークは文献 [4] によると、small-world 構造を持つとされる。本稿は第2章では small-world 構造について、第3章では実験と考察、第4章でまとめを述べる。

2. small-world とは

ネットワークの規模が巨大になると、小規模なネットワークには見られない特徴が出現する。その一つとして small-world 構造がある。この構造はノード間のパス長が短いにもかかわらず、ノードが高度にクラスタリングされている構造となっている。現実中存在する small-world 構造を持つネットワークの例として WWW、俳優の競演関係、単語の共起関係などがあげられる。[1]

文献 [4] によると、IRC ネットワークをユーザー一人ひとりをノード、同じチャンネルに参加しているユーザー同士をリンクで結ぶとしてネットワークを定義すると、ノード間の平均パス長は 2.93 でクラスタリング係数が 0.82 と極端な small-world 構造を持つことが判明している。本稿ではこの平均パス長が短いことを利用し、自然に形成されたネットワーク構造を生かして短時間でネットワークの大域に情報が伝播出来るのではと考えた。そこでこの平均パス長を導出したネットワークを利用して情報伝播のシミュレーションを行った。

3. IRC ネットワークを利用した情報伝播率の調査

3.1 ネットワークの構成の取得

ある時間の IRC ネットワークの構成を取得するために IRC ネットワークにクライアントを接続して、サーバに names コマンドを送りその結果から構成を得る、と

いう手法を用いた。names コマンドの結果として、サーバにコマンドを送った時点で存在するパブリックなチャンネルとそこに参加するユーザーの一覧を取得することが出来る。なお、全てのやりとりはテキストベースで行われる。ログ取得を行った IRC ネットワークは日本で運用されている Friend Chat † とし、2004 年 8 月から 12 月までのログを利用した。

3.2 情報伝播のシミュレーション

names のログを元にシミュレータを用いてネットワークを構築し、情報の伝播率を調べた。シミュレータは ns § を利用した。

情報を伝播させる方法として本稿では単純な擬似的ブロードキャストのフラッディングを用いた。以下では伝播させる情報をクエリと呼ぶ。クエリを自分とリンクを張っているノードに渡し、さらに隣にクエリを渡すということを繰り返してフラッディングを実現する。さらに既に受け取ったクエリは隣に渡さない、いくつかのノードを経由したらそこでクエリを渡すのを打ち切る、というルールの下にフラッディングをシミュレータ上で実装した。

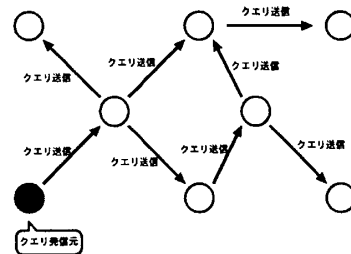


図 1: フラッディングの模式図

3.3 シミュレーションの結果

フラッディングのシミュレーションを行った結果以下に示す。シミュレーションの方法として、ログをランダムに抽出し、そのログ一つ一つでネットワークを構築し、フラッディングを試みた。フラッディングを開始するノードは試行毎にランダムに抽選し、一つのネットワーク構成で数十回のフラッディングを行い、クエリが全体の何パーセントまで行き渡るかを調査した。なお、クエリを渡すのを打ち切る転送回数は 3 から 6 までとした。最低の回数を 3 としたのは、この IRC ネットワークの平均パス長が約 3 であると判明しているからである。

*東京理科大学大学院理工学研究科情報科学専攻
†東京理科大学理工学部情報科学科

†<http://www.friend.td.nu/>

§<http://www.isi.edu/nsnam/ns/>

表 1: 月別最大到達ノード数の割合の平均

月	最大到達ノード数の割合の平均	サンプル数
8月	48%	185
9月	44%	219
10月	51%	108
11月	55%	64
12月	54%	129
合計	49%	705

表1は月ごとのクエリの最大到達率の平均と、全体の平均を示している。最小の到達率についてはどの月もほぼ0%となった。すなわち、全体として平均的に全ノード中の40%から60%までクエリが行き渡るか、0%と全く行き渡らないという2極的な結果になった。これはIRCネットワークに複数のチャンネルがかぶさることによって出来る大規模なクラスタが存在し、その周りに参加ユーザ数が少数のチャンネルが多数存在することが伺える。また、転送の打ち切りの回数を3から6まで動かして試行を行ったが、転送の回数を4以上では、クエリが行き渡る最大のノード数の割合は一致した。

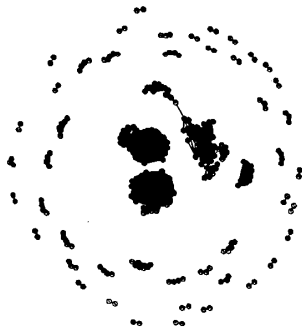


図 2: 分断されたネットワーク例

また、到達率が極端に低い場合も見受けられた。この場合のネットワークの構成を視覚化したものを図2に示す。通常は接続されると思われる中心のコミュニティが分断しているのが見てわかる。これはネットワークの参加人数が少ない場合において発生するようであるが、詳細な原因については不明である。

4. 考察

本稿でのシミュレーションはフラッディングをもとにクエリの伝播率を調べた。シミュレーションの結果よりクエリが行き渡る最大の割合はおよそ40%から60%という結果が出た。このクエリが行き渡った最大のノード数はネットワーク内に一つだけ存在する大規模コミュニティのノード数と等しい。つまり、ノードがこのコミュニティに接続していないと検索漏れの対象となってしまう。逆に言えば、大規模コミュニティのクエリ伝播に関しては極めて安定しており、クエリの転送回数も4という短い回数で十分であることが判明した。これはIRCネットワークの平均パス長が短いことから転送回数も少

なくなったということがいえる。

また、到達率が低いネットワーク構成の視覚化により、大規模コミュニティが分断されてしまう状態があることが確認された。この分断はフラッディングによってクエリを伝播させる方法において致命的な構造となる。実際のネットワークで実装の際にこの分断を克服する機構が必要となる。

4.1 改善案

クエリの到達率が60%以上にならない原因は、参加人数が少ないコミュニティが多数存在しているからと考えられる。そこで、この小規模コミュニティの参加ユーザと大規模コミュニティとを接続させる補助的な機構を考えなければならない。現在考えている改善案として、全ユーザの数%を一つのチャンネルに参加させ、このチャンネルを橋渡しとしてフラッディングを行う方法がある。この補助的なチャンネルの構造によって、小規模コミュニティのユーザも検索の対象とすることが出来る。これは同時にネットワークの分断の予防にもつながると考えられる。どれだけの数をその補助チャンネルに参加させればよいかについては現在も調査中であり、またこの手法だけではなくその他の構造も考慮に入れて改善案を考えていく予定である。

5. まとめ

本稿では small-world 構造を持つネットワークにおける最適な検索手法の考案のために、実際に small-world 構造をなす IRC ネットワークの構成を利用してフラッディングのシミュレートや構成の視覚化を行った。その結果として、IRC ネットワークには大規模なコミュニティの存在と、それに接続されない多数のコミュニティが存在し、単純なフラッディングのみでは大域まで行き渡らせるのは難しく、補助的な構造を付け加える必要があることが判明した。

今後の予定として、調査としてはIRCネットワークを利用したが、この構造を完全なサーバレスなP2Pへ応用したコミュニケーションツールの構築などを考えている。本稿で取り上げたIRCネットワークのようなコミュニティを単位としたネットワークを構築し、その上で今回シミュレーションを行ったフラッディングによる情報伝播の実装や、その上での改善点について実際に考察していきたいと考えている。

参考文献

- [1] M.E.J. Newman: The structure and function of complex network
- [2] J.OkkarinenS, D.Read:Internet Relay Chat Protocol, RFC1459, May 1993
- [3] 長谷川 卓, 松澤 智史, 武田 正之:複雑なネットワークにおける最適な検索手法の考案に向けたIRCネットワークの調査, 東京理科大学平成17年度卒業論文
- [4] 井出和典, 松澤 智史, 武田 正之:IRCにおけるユーザ関係の分析とその応用例, 情報処理学会 電子情報通信学会 FIT2005, Sep, 2005