

IRCにおけるグループの生起時間と規模の研究

松澤 智史 中山 雅哉

東京大学大学院工学系研究科 東京大学情報基盤センター

〒113-8658 東京都 文京区弥生 2-11-16

TEL 03-5841-3008: E-Mail t-matsu@kaynet.ccc.u-tokyo.ac.jp

概要

今日のインターネットとコンピュータの発展により、インターネットを介する情報交換、情報提供の場が格段に増えてきており、それらのニーズに答えるため、情報をやりとりするソフトウェアの研究も盛んになっている。またそれらのインフラとしてマルチキャストが誕生し発展してきた。本稿ではマルチキャストの研究で最近注目されているSGM(Small Group Multicast)における、具体的なインフラ・アプリケーション・ソフトウェアの設計に必要な情報交換グループの規模と生起時間を推定するために、現行のインターネット上でのグループ通信としての代表であるIRC(Internet Relay Chat)の上でのチャンネルにおけるユーザ数や、生起してから継続している時間を調べ、その結果約90%のグループは20人以下で構成されており、80%のグループは生起してから継続時間が13時間以下であることが、明らかとなった。

1 はじめに

コンピュータの発展とともに、個人による仕事・作業の生産性は目覚しく進歩してきた。それによってある同じ目標を持つ個人同士でのスムーズな情報交換や、情報共有などが注目をあびるようになってきた。これらを支援するソフトウェアとしてグループウェアと呼ばれるものが生まれ、近年注目を浴びている。

グループコミュニケーションには主に

- 発信元が少数で受信が多数の情報提供
- 一定数のメンバーによる相互情報交換

の2種が挙げられる。前者は例として掲示板などが該当し、後者はchatと呼ばれるソフトウェアなどがあたる。

また、これらのインフラとしてマルチキャスト技術が誕生し、多くの情報提供のソフトウェアを生み出している。

現在のマルチキャストでは前者の情報提供型のケースが多いが、最近注目をあびてきたSGM(Small Group

Multicast)[1][2]においては後者の一定数メンバーによる情報交換が主流となる。しかし、これまでグループの規模や生起時間について言及する研究は見当たらない。そこで本稿では、実際のインターネット上でのコミュニケーションが、どの程度の規模でどの程度の期間形成されるかを調査するためにこれらの調査を行い、IRC(Internet Relay Chat)[3]を対象し、この結果をもとにsmall groupのマルチキャストを実現開発する上での参考データの取得を試みることにする。

2 IRCを使ったグループ調査について

2.1 IRCとは

IRC(Internet Relay Chat)は1988年にフィンランドで開発され、1990年には日本でも利用が開始されたネットワークリアルタイム会議システムであり、複数のサーバが相互接続されて1つのIRCネットワークを形成している。IRCクライアントは、いずれかのサーバに接続することで、同時に世界中にいる多数の相手と会話を行なうことができ、いつでも好きな時に会話に参加したり会話から抜け出したりする事が可能となる。

そもそもIRCは海外でUNIX上で開発されたシステムであったが、現在ではさまざまな人々の努力によりUNIX, Windows, Macintoshなどのさまざまな環境から利用でき、日本語による会話を行なえるようなクライアント(ユーザが利用するプログラム)も数多く作られている。

IRCネットワークにはチャンネルというグループのコミュニケーション単位を持ち、同一のチャンネルにJOINしているユーザ同士で会話を行う。チャンネルはチャンネル名という文字列によって識別され、IRCネットワーク内で同一のチャンネル名を持つチャンネルは1つしか存在しない。

本稿ではこのIRCのチャンネルをグループとして見たことによってグループと生起継続時間の測定を行っている。

2.2 計測方法

IRC ネットワークには、異なるサーバ群で形成された複数の IRC サーバが存在している。本稿では代表的な IRC ネットワークのうち、OpenProjectNet, GimpNet と呼ばれる 2 つの IRC ネットワークを対象としてそれぞれの IRC ネットワークの各チャンネルに JOIN しているユーザ数を調べる。まず各 IRC ネットワークの中の 1 つのサーバを選んで接続し、10 分ごとに IRC の list コマンド¹を用いて、その IRC ネットワークに存在するチャンネルとユーザ数を取得するクライアントを作成した。

取得したデータは以下のようにチャンネル名、ユーザ数の形で 1 行ずつの形式とし、10 分毎に時刻をファイル名とした 1 つのファイルに出力される。

```
#gxsncmp,2  
#avalon,1  
#mgicchikn,5
```

この際取得するチャンネルはグローバルチャンネルのみとした。IRC の仕様上、チャンネルオペレータによるチャンネルの秘密化がされたチャンネルは list コマンドにより出力されないため、public に公開されたチャンネルのみの結果である。

これらの結果をもとに統計処理を行って結果を得る。ただしユーザ数が 1 であるチャンネルはグループとしての機能を果たしていないチャンネルであるため、統計データはユーザ数が 2 以上のものに限定している。

2.3 計測期間

GimpNet の測定期間は 2003 年 3 月 10 日 18 時 10 分～20 日 7 時 50 分、OpenProjectNet の計測期間は 2003 年 3 月 10 日 18 時 30 分～12 日 2 時 40 分である。

3 測定結果

3.1 チャンネルごとのユーザ数推移

前章で示した方法により取得した各 IRC ネットワークにおける測定開始時からのチャンネル (グループ) への参加ユーザ数を図 1 と図 2 に示す。

図 1 は GimpNet のユーザ数推移、図 2 は OpenProjectNet のユーザ数推移である。x 軸には、Channel 各自に付与したチャンネル ID を示し、y 軸は 10 分単位の計

¹ list コマンドにより、現在その IRC ネットワークに存在するチャンネルの一覧と、そのチャンネルに JOIN しているユーザ数を引き出すことができる。

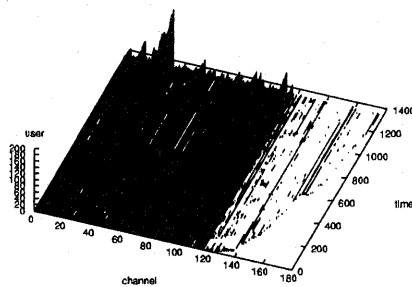


図 1: GimpNet のチャンネルごとのユーザ数推移

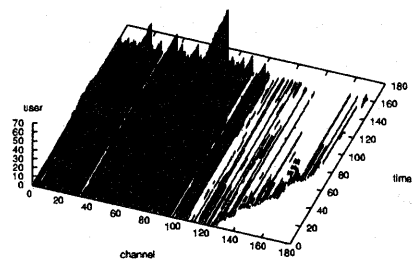


図 2: OpenProjectNet のチャンネルごとのユーザ数推移

測開始時からの経過時間、z 軸はそれに対応するユーザ数を表している。

このグラフによって常時多数ユーザを抱えるチャンネルと、短時間で少数のユーザで構成される一時的なグループチャンネルの傾向を視覚的にとらえることができる。

3.2 ユーザ数とチャンネル数の変移

前節の図 1 と図 2 に示したように、同じチャンネルでも時刻に応じて参加する人数が変動するため、計測期間中の各時刻における各チャンネルの参加人数を別々のグループとしてとらえた時のユーザ数を示すと図 3 と図 4 のようになる。

図 3 は GimpNet におけるユーザ数とそのユーザ数で構成されたチャンネルの度数を表しており、x 軸がユーザ数で y 軸はユーザ数ごとのチャンネル数を示している。1 チャンネルの最大数は 3 月 12 日の 2 時に記録した #gnome の 189 人であり、チャンネルあたりの平均ユーザ数は 10.3 人、標準偏差は 17.11 人となった。総度数は 164213 個であった。

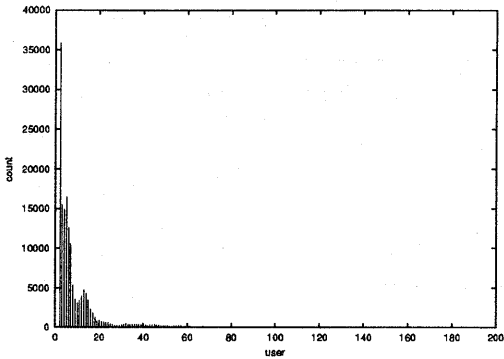


図 3: GimpNet のユーザ数とその度数

表 1 は図 3 におけるそのユーザ数以上の累積度数が 50%, 60%, 70%, 80%, 90% を超えたものを示したものである。

この結果から 20 人以下のチャンネルの累計が全体の 90% 以上となることがわかる。

表 1: GimpNet の累計とその割合

user	count の累計	全体の割合
5	84534	51%
7	108167	66%
9	117346	71%
13	132858	81%
20	148192	90%

図 4 および表 2 は、OpenProjectNet での同様の集計結果である。

1 チャンネルの最大数は 3 月 12 日の 2 時に記録した #slashdot の 68 人であり、チャンネルあたりの平均ユーザ数は 6.12 人、標準偏差は 7.29 人となった。チャンネル総数は 21863 個であった。

表 2: OpenProjectNet の累計とその割合

user	count の累計	全体の割合
4	12798	59%
5	14673	67%
6	15950	73%
8	17904	82%
13	19721	90%

また 13 人以下のチャンネルの累計が全体の 90% 以上を示すことがわかる。

これらの結果から、ほとんどのチャンネルは非常に少ないユーザ数で形成されたグループであり、その規模は 20 人以上で約 90% 以上をカバーすることができる。

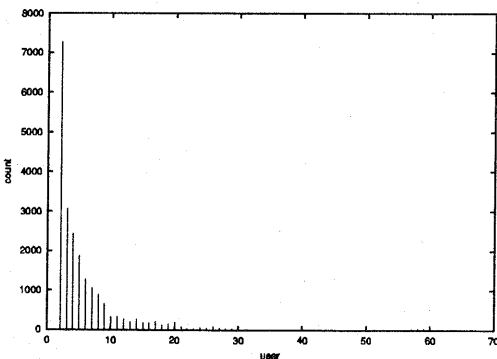


図 4: OpenProjectNet のユーザ数とその度数

3.3 グループの継続時間

次に、グループがどの程度の時間の間構成されているかを調査する。ここでのグループは、あるチャンネルが作成されてからユーザがいなくなるまでに形成されるコミュニティを 1 つとして考えることにした。すなわち同じチャンネルでも一旦ユーザがいなくなり、また後ほど形成された場合は別グループとみなすことにしている。

図 5 は、GimpNet においてグループが生成してから継続時間とその頻度を表したグラフである。また図 6 は

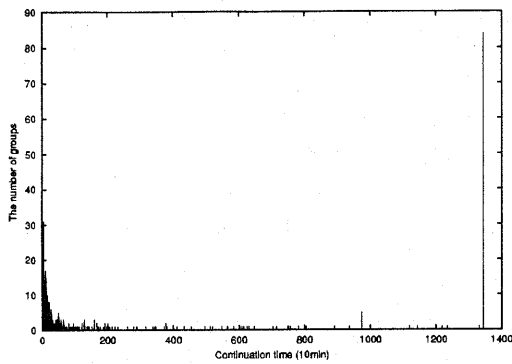


図 5: GimpNet のグループの継続時間

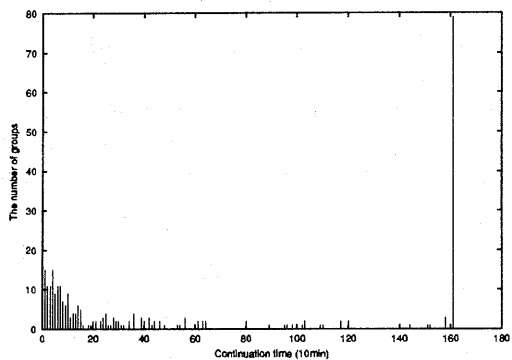


図 6: OpenProjectNet のグループの継続時間

OpenProjectNet のグループ継続時間とその頻度を表したグラフである。x 軸は継続時間を 10 分単位で表した数値であり、y 軸はそれに対応するグループ数である。

各々の図において右端に現れている頻度の高いデータは計測開始から終了までずっと継続しているグループを示している。このようなグループ(チャンネル)は、GimpNet では全 613 グループ中の 84 グループがこれに該当し、OpenProjectNet では 286 グループ中 79 グループが該当していた。GimpNet の残り 529 グループを対象に、全体の 50-90%を占めるグループ数に対応する継続時間は表 3 の通りである。

表 3: GimpNet の継続時間以内に該当するグループの累計とその全体の割合

継続時間	グループの累計	全体の割合
150	271	51%
220	319	60%
450	371	70%
830	424	80%
2060	477	90%

830 分以下の継続時間のグループ数は、全体の 8 割を占める 424 グループ存在していた。

OpenProjectNet の残り 207 グループを対象に同様のことを行くと、表 4 に示す結果を得られる。

表 4: OpenProjectNet の継続時間以内に該当するグループの累計とその全体の割合

継続時間	グループの累計	全体の割合
100	105	51%
150	127	61%
270	145	70%
420	168	81%
800	187	90%

継続時間が 800 分以下のグループは 187 グループで全体の 9 割を占める。

これらの結果により、IRC でのグループは、ある程度の規模の数だけ常時(継続時間不明)が存在するが、それ以外のグループは約 13 時間(800 分)以下に約 80%が集中することが明らかとなった。

4 まとめ

本稿では 2 つの IRC ネットワークのユーザ数とグループ数、グループの継続時間などからグループコミュニケーションの規模や継続時間を調査し、約 20 人以下のグループが 90%以上であり、約 13 時間以下のグループが 80%以上となることが明らかとなった。

IRC という特定のアプリケーションでの調査が、インターネットにおけるグループコミュニケーション全体の形態を示すものではないが、一つの目安となる参考データとなると考えている。

これによりこのような短い期間誕生するチャンネルには、従来のような情報提供型のマルチキャストではなく、small group による新しい体系のマルチキャストの設計指針とすることが可能となると考えている。

参考文献

- [1] R. Boivie, N. Feldman: Small Group Multicast, Internet Draft, work in progress, draft-boivie-sgm-01.txt, July 2000
- [2] R. Boivie, N. Feldman, Ch. Metz: Small Group Multicast: A New Solution for Multicasting on the Internet, Internet Computing, Vol. 4, No. 3, May/June 2000
- [3] J. Oikarinen, S. D. Reed: Internet Relay Chat Protocol, RFC1459, May 1993