

Title	投稿行動から見た電子ニュースfj
Author(s)	佐藤, 円; 篠田, 陽一
Citation	Research report (School of Information Science, Japan Advanced Institute of Science and Technology), IS-RR-94-0023: 1-24
Issue Date	1994-07
Type	Technical Report
Text version	publisher
URL	http://hdl.handle.net/10119/8367
Rights	
Description	リサーチレポート (北陸先端科学技術大学院大学情報科学研究科)

投稿行動から見た電子ニュース fj

佐藤円 篠田陽一

1994年7月

IS-RR-94-23

北陸先端科学技術大学院大学

情報科学研究科

〒923-12 石川県能美郡辰口町旭台15

{madoka, shinoda}@jaist.ac.jp

©Madoka Sato and Yoichi Shinoda, 1994

ISSN 0918-7553

要旨

本論文では、電子ニュース fj の現状を、投稿行動に関する統計的な調査から明らかにする。調査は、fj 全体の規模およびニュースグループ・サブジェクト・投稿者を軸にした規模と傾向について行なった。さらに、ニュースグループの統計量とニュースグループの特徴の関係を探り、特徴的なニュースグループを抽出した。

もくじ

1	はじめに	1
2	電子ニュース調査の概要	1
2.1	調査対象	1
2.2	データベースの作成	1
2.3	調査項目	2
2.4	用語説明	4
3	全体のバイト数・記事数・投稿者数と投稿傾向の調査	5
4	ニュースグループのバイト数・記事数・投稿者数の調査	6
4.1	ニュースグループのバイト数・記事数・投稿者数	6
4.2	バイト数・記事数の多いニュースグループ	6
4.3	バイト数・記事数の少ないニュースグループ	11
5	サブジェクトの特徴	12
6	投稿者の投稿傾向の調査	13
6.1	投稿回数とイニシャル投稿比	13
6.2	イニシャル投稿・フォロー投稿別投稿傾向	13
7	特徴的なニュースグループ	16
7.1	ニュースグループの統計量とニュースグループの特徴の関係	16
7.2	イニシャル投稿率	16
7.3	投稿者多様度	19
7.4	平均バイト数	21
7.5	特徴的なニュースグループ群	22
8	まとめ	22

1 はじめに

計算機ネットワークの発達は、電子メール、電子ニュースなどの新しいコミュニケーションメディアを我々にもたらした。このうち、電子ニュースは、多対多のマスコミュニケーションメディアであり、日本では、非商用ネットワークでは fj が、商用ネットワークでは PC-VAN や Nifty-Serve の電子掲示版が広く利用されている [1]。

fj とはインターネット上で、主として日本語での投稿が行なわれているニュースグループの総称であり、非商用である、実名を用いたネットワーク・コミュニケーションであるといった特徴を持つ。その規模は、1993 年の時点で、1 カ月の投稿量は、約 17000 件、約 36MB であり、毎日ほぼ新聞の全国紙に匹敵するデータ量のテキストが流通している。

現在の電子ニュースは、黎明期のものであり、今後サービスの内容や対象において様々な形態が出現すると考えられる。その中で、非商用ネットワークサービスは、重要な位置を占めていくと予想される。また、実名を用いたネットワーク・コミュニケーションは、今後のインターネットの拡充により、広く普及すると考える。このように、fj には、今後の電子ニュースを考えるうえで重要な要素が多い。しかし、その実態調査はほとんどなされていないと言っていよい。唯一の例外は、野島 [2] らによる、JUNET を対象とした記事数および投稿者数の調査および利用者へのアンケート調査である。この調査以来約 6 年間にわたり、体系的な調査研究は行なわれていないとともに、各投稿者の投稿傾向や同一話題への投稿量などの詳細な調査は行なわれたことがない。

このような背景より、我々は、fj を対象とした電子ニュースの現状調査を行なった [3]。今回の調査では、実際に流通している記事の各種統計量を調べることによって、投稿行動 (投稿量) から見た fj の現状を明らかにすることを主に行なった。

2 電子ニュース調査の概要

本節では、電子ニュースの調査の対象、および調査の過程で用いたデータベースの内容と作成方法について述べる。また、以降で用いる用語について説明する。

2.1 調査対象

今回の調査では、1993 年 10 月 1 日夜の時点で jaist-news.jaist.ac.jp¹ にスプールされていた fj の全データ (60 日分)² を対象とした。調査対象の記事数は 33534 記事であり、全データ量は 101,433Kbyte である³。

2.2 データベースの作成

以降の調査を容易にするために、それぞれの記事に対して、以下に示す 11 項目のデータからなるレコードを作成し、データベース化した。

¹北陸先端科学技術大学院大学において、ニュースをスプールしているマシン。

²ただし、fj.test は 3 日分、fj.jokes は 30 日分である。

³du コマンドにより測定。

1. ニュースグループ名
記事が格納されているディレクトリより切り出す。
2. 記事のファイル名
処理対象ファイルパス名より切り出す。
3. ニュースグループ名 (含クロスポスト)
ヘッダー部の**Newsgroups:**で始まる行より切り出す。クロスポストの場合は、';'で区切られた形で、ニュースグループ名が複数存在することになる。
4. 投稿者名
ヘッダー部の**From:**で始まる行より切り出す。
5. サブジェクト
ヘッダー部の**Subject:**で始まる行より切り出す。
6. reply
ヘッダー部の**In-Reply-To:**で始まる行が存在すれば、1 とする。存在しなければ、0 とする。
7. reference
ヘッダー部の**References:**で始まる行が存在すれば、1 とする。存在しなければ、0 とする。
8. 投稿日
ヘッダー部の**Date:**で始まる行より切り出す。
9. 本体の行数
ヘッダー部の**Lines:**で始まる行があれば、そこから切り出す。**Lines:**行がない場合は、プログラムによって求める。
10. ヘッダー部のバイト数
プログラムによって求める。
11. 本体のバイト数
ファイル全体のバイト数から、ヘッダー部のバイト数を引くことによって求める。

表1に、図1⁴に示した記事から抽出したレコードを示す。以下の調査は全て、このデータベースに対して、機械処理で行なったものである。

2.3 調査項目

今回調査した項目は、以下の通りである。

1. fj全体のバイト数・記事数・投稿数と投稿傾向の調査
2. 各ニュースグループのバイト数・記事数・投稿数・投稿傾向の調査
3. サブジェクトの特徴の調査
4. 投稿者の投稿傾向の調査

⁴この記事は、実際に投稿された記事ではなく、説明のために作成したものである。

Newsgroups: fj.ai
 Path: jaist-news:sato
 From: sato@is17e0s00.jaist.ac.jp (Satoshi Sato)
 Subject: WAL'94
 Message-ID: <SATO.93Dec9145713@is17e0s00.jaist.ac.jp>
 Sender: news@jaist.ac.jp (News System Administrator)
 Organisation: School of Information Science, JAIST, Hokuriku, Tatsunokuchi,
 Ishikawa, Japan.
 Distribution: fj
 Date: Thu, 9 Dec 1993 19:57:13 GMT

CALL FOR PAPERS
WAL (Workshop on Artificial intelligence toward Learning) '94

人工知能に関するワークショップを下記の要領で開催致します。奮ってご参加下さい。

(中略)

・また、合わせて、wolメーリングリスト(元気なAI研究者が数多く参加しています)への加入申し込みも承っております。

[問い合わせ先]
 〒923-12 石川県能美郡辰口町旭台15
 北陸先端科学技術大学院大学 情報科学研究科
 佐藤理史
 Phone: 0761-51-1221
 Fax: 0761-51-1149
 Email: sato@jaist.ac.jp

図 1: 記事の例

3

表 1: 記事例に対するレコード

1	fj/ai
2	1044
3	fj.ai
4	sato@is17e0s00.jaist.ac.jp
5	WAL'94
6	0
7	0
8	Thu, 9 Dec 1993 19:57:13 GMT
9	109
10	358
11	4317

2.4 用語説明

ここで、本論文で用いる用語を整理しておく。

- バイト数
記事のバイト数とは、記事からヘッダー部を除いた本体のデータ量をバイトで表したものである。ニュースグループのバイト数とは、そのニュースグループの全記事の本体のバイト数を合計したものである。fj のバイト数は、fj の全記事の本文部分のバイト数の合計を示す。
- 記事数
記事数とは、投稿された記事の数である。
- 投稿者数
投稿者数とは、記事を投稿した異なりの投稿者数である。
- イニシャル投稿
投稿者が、新しいサブジェクト名を設定して投稿した、新たな話題を提供する記事である。以下の条件を1つも満たさない記事はイニシャル投稿記事であるとみなす。
 - ヘッダー部に In-Reply-To:行が存在する
 - ヘッダー部に Reference:行が存在する
 - ヘッダー部の Subject:行が、Re: で始まる
- フォロー投稿
以前に投稿された記事へのリプライ記事である。イニシャル投稿記事ではない投稿は、フォロー投稿とみなす。
- サブジェクト
1つのイニシャル記事を基点に、そのイニシャル投稿へのリプライ記事およびリプライ記事へのリプライ記事の全てを含む記事グループを、サブジェクトと呼ぶ。1つのサブジェクトは、通常、イニシャル投稿者が設定したサブジェクト名を共有する。(サブジェクトの途中で、話題の変換に伴い、サブジェクト名を変更する場合がある。)
- イニシャル投稿率
記事数に対する、イニシャル投稿記事数の割合を示す値で、以下のように定義する。

$$\text{イニシャル投稿率} = \frac{\text{イニシャル記事数}}{\text{全記事数}}$$

- 平均バイト数
1記事あたりの平均バイト数である。
- 投稿者多様度
投稿者の多様度を示す値で、以下のように定義する。

$$\text{投稿者多様度} = \frac{\text{投稿者数}}{\text{全記事数}}$$

表 2: fj の月別バイト数・記事数・投稿者数

月	バイト数 (KB)	記事数	投稿者数
9 月	32671	15265	4121
10 月	38694	18269	4614
平均	35683	16767	4378

表 3: fj の日別バイト数・記事数・投稿者数

	バイト数	記事数	投稿者数
週日	1592K	797.05	200.75
土曜 (週日比)	691K (43.4%)	248.2 (31.1%)	64.4 (32.1%)
日曜・祝日 (週日比)	565K (35.5%)	181.17 (22.7%)	46.17 (23.0%)

表 4: 曜日別平均バイト数・記事数・投稿者数

曜日	バイト数 (KB)	記事数	投稿者数
月曜	1886	877.33	200.3
火曜	1638	848.25	212.5
水曜	1435	785.25	197.0
木曜	1521	780.75	200.5
金曜	1563	730.40	194.8
土曜	691	248.20	64.4
日曜	508	834.00	40.8

ニュースグループの全記事が、一人の投稿者によって投稿されている場合は、投稿者多様度は 0 に近い値をとる。逆に、全記事が、異なった投稿者により投稿されている場合は、投稿者多様度は 1 となる。

3 fj 全体のバイト数・記事数・投稿者数と投稿傾向の調査

fj 全体の規模を把握するため、バイト数・記事数・投稿者数について、1 ヶ月・1 日を単位に、調査を行なった。さらに、これらの値について、曜日別、週日・週末別の差異を調べた。それらの結果を表 2-5 に示す。

これらの調査結果から、以下のことが観察される。

- fj の 1 ヶ月の情報の規模は、約 36MB、約 17000 記事、約 4400 投稿者である (表 2)。
- fj の 1 日 (週日⁵) の情報の規模は、約 1.6MB、約 800 記事、約 200 投稿者である (表 3)。

⁵月曜・火曜・水曜・木曜・金曜。

表 5: 週日・土曜・日曜の投稿傾向の変化

曜日	1 記事あたりの 平均バイト数 (週日比)		1 投稿者あたりの 平均投稿記事数 (週日比)	
週日	1.99	(100%)	4.00	(100%)
土曜	2.78	(139.7%)	3.86	(96.3%)
日曜	3.12	(156.8%)	3.90	(97.5%)

- 週日においては、曜日によるバイト数・記事数・投稿者数の変化はみとめられない。週日と週末⁶を比較すると、バイト数・記事数・投稿者数とも、週末に減少がみられる(表4)。
- 1 投稿者あたりの投稿記事数は、週日・週末を通して殆んど差異がない。1 記事あたりのバイト数は、週末に約 1.5 倍に増加する(表5)。

4 ニュースグループのバイト数・記事数・投稿者数の調査

本節では、各ニュースグループのバイト数・記事数・投稿者数について行なった調査の結果を示す。さらに、バイト数・記事数の多いニュースグループ、少ないニュースグループについて、調査結果に基づいた考察を行なう。

4.1 ニュースグループのバイト数・記事数・投稿者数

fj の全ニュースグループについて、1 ヶ月あたりのバイト数、記事数、投稿者数を調査した。それらの結果を表6-8に示す。

これらの調査から以下のことが観察される。

- 1 ヶ月のバイト数が500KB以下のニュースグループ数が全ニュースグループ数の大部分(86.9%)を占め、20KB以下のニュースグループが、全ニュースグループ数の約1/4(23.1%)を占めている(表6)。
- 1 ヶ月の記事数が50以下のニュースグループの数が、全ニュースグループ数の約半分(51.9%)を占めている。さらに、1 ヶ月の記事数が10記事以下、つまり平均3日に1度の投稿しかないニュースグループが、全ニュースグループ数の約1/5(18.1%)を占めている(表7)。
- 1 ヶ月あたりの投稿者数が50人以下のニュースグループ数が全ニュースグループ数の71.9%を占めており、10人以下のニュースグループ数が全ニュースグループの27.5%を占めている(表8)。

4.2 バイト数・記事数の多いニュースグループ

ここでは、バイト数・記事数の多いニュースグループについてバイト数と記事数の相関および全バイト数・記事数に占める割合について調査した。調査結果を表9、表10に示す。

⁶土曜・日曜・祝日。

表 6: 1 カ月のバイト数別ニュースグループ数

バイト数 (KB)	ニュースグループ数 (全体比)	
1-20	37	(23.1%)
21-40	16	(10.0%)
41-60	11	(6.9%)
61-80	11	(6.9%)
81-100	7	(4.4%)
1-100	82	(51.3%)
101-200	26	(16.3%)
201-300	15	(9.4%)
301-400	11	(6.9%)
401-500	5	(3.1%)
1-500	139	(86.9%)
501-1000	14	(8.8%)
1001-1500	4	(2.5%)
1501-2000	0	
2001-2500	1	(0.6%)
2501-3000	1	(0.6%)
3001-3500	0	
3501-4000	1	(0.6%)
合計	160	(100%)

表 7: 1 カ月の記事数別ニュースグループ数

記事数	ニュースグループ数 (全体比)	
1-10	29	(18.1%)
11-20	21	(13.1%)
21-30	13	(8.1%)
31-40	14	(8.8%)
41-50	6	(3.8%)
1-50	83	(51.9%)
51-100	27	(16.9%)
101-150	14	(8.8%)
150-200	11	(6.9%)
1-200	135	(84.4%)
201-400	15	(9.4%)
401-600	4	(2.5%)
601-800	3	(1.9%)
801-1000	1	(0.6%)
1001-1200	1	(0.6%)
1201-1400	1	(0.6%)
合計	160	(100%)

表 8: 1 ヶ月の投稿者数別ニュースグループ数

投稿者数	ニュースグループ数 (全体比)	
1-10	44	(27.5%)
11-20	23	(14.4%)
21-30	21	(13.1%)
31-40	11	(6.9%)
41-50	16	(10.0%)
1-50	115	(71.9%)
51-60	7	(4.4%)
61-70	7	(4.4%)
71-80	4	(2.5%)
81-90	2	(1.3%)
91-100	3	(1.9%)
1-100	138	(86.3%)
101-200	13	(8.1%)
201-300	7	(4.4%)
301-400	1	(0.6%)
401-500	0	
501-600	0	
601-700	0	
701-800	0	
801-900	1	(0.6%)
合計	160	(100%)

表 9: 1 カ月のバイト数が 500KB 以上のニュースグループ

ニュースグループ名	バイト数 (順位)		記事数 (順位)	
fj.sources	3961	(1)	124	(43)
fj.binaries.msdos	2859	(2)	62	(69)
fj.mail-lists.x-window	2087	(3)	1273	(1)
fj.rec.tv	1412	(4)	935	(3)
fj.soc.men-women	1333	(5)	739	(4)
fj.rec.autos	1132	(6)	687	(5)
fj.rec.animation	1017	(7)	453	(9)
1MB 以上の合計	13801		4273	
1MB 以上の合計の対全体比	35.7%		23.4%	
fj.rec.games.video	974	(8)	610	(6)
fj.jokes	745	(9)	1121	(2)
fj.forsale	706	(10)	527	(7)
fj.soc.misc	687	(11)	297	(14)
fj.rec.rail	676	(12)	350	(13)
fj.life.in-japan	619	(13)	275	(16)
fj.rec.music	610	(14)	457	(8)
fj.sci.math	597	(15)	210	(25)
fj.mail	597	(16)	234	(21)
fj.rec.motorcycles	575	(17)	232	(22)
fj.sys.mac	556	(18)	421	(10)
fj.living	546	(19)	361	(11)
fj.rec.sports.soccer	520	(20)	356	(12)
fj.binaries.mac	515	(21)	11	(69)
500KB 以上の合計	22724		9735	
500KB 以上の合計の対全体比	58.7%		53.3%	

表 10: 1 ヶ月の記事数が 150 記事以上のニュースグループ

ニュースグループ名	記事数 (順位)		バイト数 (順位)	
fj.mail-lists.x-window	1273	(1)	2087	(3)
fj.jokes	1121	(2)	745	(9)
fj.rec.tv	935	(3)	1412	(4)
fj.soc.men-women	739	(4)	1333	(5)
fj.rec.autos	687	(5)	1132	(6)
fj.rec.games.video	610	(6)	974	(8)
fj.forsale	527	(7)	706	(10)
fj.rec.music	457	(8)	610	(14)
fj.rec.animation	453	(9)	1017	(7)
fj.sys.mac	421	(10)	556	(18)
fj.living	361	(11)	546	(19)
fj.rec.sports.soccer	356	(12)	520	(20)
fj.rec.rail	350	(13)	676	(12)
fj.soc.misc	297	(14)	687	(11)
fj.wanted	276	(15)	274	(41)
fj.life.in-japan	275	(16)	620	(13)
fj.rec.comics	257	(17)	389	(28)
fj.rec.food	255	(18)	328	(33)
fj.rec.sports.baseball	242	(19)	364	(30)
fj.rec.sports.ski	240	(20)	371	(29)
fj.mail	234	(21)	597	(16)
fj.rec.motorcycles	232	(22)	575	(17)
fj.rec.photo	220	(23)	448	(23)
fj.sys.pc98	216	(24)	301	(37)
fj.sci.math	210	(25)	597	(15)
fj.rec.sports.keiba	195	(26)	396	(27)
fj.rec.av	184	(27)	294	(38)
fj.sys.ibmpc	179	(28)	229	(45)
fj.rec.sports.prowrestling	179	(29)	421	(25)
fj.questions.misc	171	(30)	245	(43)
fj.rec.models	160	(31)	314	(34)
fj.soc.smoking	159	(32)	280	(40)
fj.comp.texhax	157	(33)	257	(42)
fj.sys.x68000	155	(34)	281	(39)
fj.questions.unix	154	(35)	221	(48)
fj.sci.lang	151	(36)	226	(46)
合計	13088		21029	
対全体比	71.6%		54.3%	

これらの調査結果から、以下のことが観察される。

- 1 ヶ月のバイト数における上位 21 グループ (バイト数が 500KB 以上、全ニュースグループ数の約 13% を占める) のバイト数・記事数を合計すると、fj の 1 ヶ月のバイト数の 58.7%、記事数の 53.3% を占める (表 9)。
- 1 ヶ月の記事数における上位 36 グループ (記事数が 150 以上、全ニュースグループ

表 11: 1 ヶ月のバイト数が 100KB 以下のニュースグループ

ニュース グループ数 (対全体比)	バイト数 合計 (対全体比)	記事数 合計 (対全体比)
82 (51.3%)	2840 (7.3%)	1709 (9.4%)

表 12: 1 ヶ月の記事数が 30 記事以下のニュースグループ

ニュース グループ数 (対全体比)	バイト数 合計 (対全体比)	記事数 合計 (対全体比)
65 (40.6%)	825 (4.5%)	2486 (6.4%)

ブ数の約 23% を占める) のバイト数・記事数を合計すると、fj の 1 ヶ月のバイト数の 54.3%、記事数の 71.6% を占める (表 10)。

- バイト数上位 21 グループからソースコード用の 3 つのニュースグループ⁷ を除いた 18 グループは、いずれも、記事数順の上位 25 位以内に属している。
- 記事数順の上位 36 グループ (1 ヶ月の記事数が 150 記事以上) では、どのニュースグループも、バイト数順の上位 50 位以内に属している。

このことから、ソースコード用ニュースグループを除外すると、fj においては、「バイト数の多いニュースグループは記事数も多い」ということが成り立つのがわかる。また、その逆の「記事数の多いニュースグループは、バイト数も多い」ということも成り立つことがわかる。

4.3 バイト数・記事数の少ないニュースグループ

1 ヶ月のバイト数が 100KB 以下のニュースグループのグループ数・バイト数合計・記事数合計を、表 11 に示す。また、1 ヶ月の記事数が 30 記事以下 (1 日の平均記事数が 1 記事以下) のニュースグループについても同様に、表 12 に示す。

これらの表から、以下のことが観察される。

- fj の約半数のニュースグループ (82 ニュースグループ、全ニュースグループ数の 51.3%) は、1 ヶ月のバイト数が 100KB 以下である。これらのニュースグループのバイト数合計は、全ニュースグループのバイト数合計の 1/10 以下 (7.3%) である。
- fj では、約 2/5 のニュースグループ (65 ニュースグループ、全ニュースグループ数の 40.6%) において、1 ヶ月の記事数が 30 記事以下 (1 日の平均記事数が 1 記事以下) である。これらのニュースグループの記事数合計は、全ニュースグループの合計記事数の 1/25 以下 (4.2%) である。

⁷fj.sources, fj.binaries.msdos, fj.binaries.mac。これらのグループは、バイト数は多いが、この上位 21 グループ中の他のニュースグループと比較すると、記事数が明らかに少ないこのことにより、ソースコード用ニュースグループの 1 記事あたりのバイト数は、他のニュースグループの記事に比較して、非常に多いことがわかる。

表 13: 記事数別のサブジェクト数

記事数	サブジェクト数 (対全体比)	
1	5825	(54.2%)
2	1798	(16.7%)
3	868	(8.1%)
4	524	(4.9%)
5	392	(3.6%)
6	276	(2.6%)
7	183	(1.7%)
8	148	(1.4%)
9	122	(1.1%)
10	87	(0.8%)
1-10	10223	(95.4%)
11-20	367	(3.4%)
21-30	81	(0.7%)
31-40	36	(0.3%)
41-50	13	(0.1%)
51-60	6	
61-70	5	
71-80	8	
81-90	5	
91-100	4	
101-200	3	
201-300	2	
合計	10753	(100%)

表 14: サブジェクトの存続期間と記事数 (30 記事以上のサブジェクト)

	期間	記事数	1日あたりの平均記事数
平均	20.8 日	55.1 記事	2.7 記事
最高値	61 日	286 記事	8.4 記事

5 サブジェクトの特徴

fjにおけるサブジェクトの調査として、各サブジェクトの記事数と存続期間を調べた。その結果を、表 13、表 14に示す。

これらの表から、以下のことがわかる。

- サブジェクトの記事数には、ばらつきがある。
- 全サブジェクトの半数以上 (54.2%) が、1 記事から成る (表 13)。
- サブジェクトは、平均で約 21 日間、最高で 61 日⁸と比較的長い期間にわたって存続している (表 14)。

⁸全調査期間である。

- 同一サブジェクトの記事数は、平均で約 55 記事、最高で 286 記事である。期間中の記事数合計が約 34000 記事であることから、サブジェクトへの記事集中は、最高でも記事数合計の 1%未満である。

サブジェクトは、おおよそ以下の 3つの種類に分類できると考えられる。

- 「アナウンス」型
会告の通知、お知らせなど。通常、1つの記事で完結する。
- 「質問-応答」型
各種質問、情報提供要求と、それに対する応答。記事数は 1 以上である。
- 「議論・おしゃべり」
ある問題や論点について、2 以上の参加者が投稿を繰り返し、議論やおしゃべりを行なうもの。記事数はかなり多くなる。

サブジェクトの記事数にばらつきがあるという事実は、fj ではこれらの種類のサブジェクトが混在していることを示唆する。また、1つのサブジェクトの記事集中が fj の全記事数に占める割合が、かなり低いという事実は、1つのサブジェクトへの多数記事の集中投稿という現象があまり起きていないということを示している。

6 投稿者の投稿傾向の調査

6.1 投稿回数とイニシャル投稿比

調査期間 1ヶ月間の全投稿者について、投稿回数とイニシャル投稿率を調べた。その結果を表 15に示す。

この表より、以下のことが観察される。

- fj においては、1ヶ月に 1回投稿の投稿者が、全投稿者の約 1/2(49.3%) を占める。
- 平均イニシャル投稿率は、1ヵ月あたりの投稿回数 1回の投稿者グループにおいて、最も高い値を示す (51.4%)。
- ごく少数の例外 (投稿回数 101-110 回の 1名、投稿回数 150-160 回の 2名) を除外すると、投稿回数が増加するに従い、平均イニシャル投稿率が減少している。
- 1ヶ月の投稿回数が 11 回以上 (平均 3 日に 1度以上投稿) の投稿者は、全体の 1/10(7.4%) にすぎない。

6.2 イニシャル投稿・フォロー投稿別投稿傾向

投稿回数月 30 回以上の各投稿者について、イニシャル投稿・フォロー投稿のそれぞれにおける、記事数、ニュースグループ数、バイト数、1記事あたりの平均バイト数を調査した。その結果を、表 16 に示す。また各投稿者において、記事数・バイト数・1記事あたりのバイト数は、イニシャル投稿とフォロー投稿のどちらの投稿で高い値を示すかを調査した。その結果を表 17に示す。この表より、以下のことが観察される。

- 投稿回数月 30 回以上の全投稿者の総計では、記事数、バイト数とも、イニシャル投稿よりフォロー投稿の方が多い。

表 15: 1ヶ月の投稿回数と投稿傾向

投稿回数	投稿者 数 (対全体比)	記事 数 (対全体比)	平均イニシ ヤル投稿率
1	2276 (49.3%)	2276 (12.5%)	51.4%
2	858 (18.6%)	1716 (9.4%)	39.6%
3	377 (8.2%)	1131 (6.2%)	31.7%
4	256 (5.5%)	1024 (5.6%)	29.5%
5	138 (3.0%)	690 (3.8%)	27.4%
6	114 (2.5%)	684 (3.7%)	23.7%
7	100 (2.2%)	700 (3.8%)	24.0%
8	69 (1.5%)	552 (3.0%)	24.5%
9	48 (1.0%)	432 (2.4%)	15.7%
10	35 (0.8%)	350 (1.9%)	17.1%
1-10	4271 (92.6%)	9555 (52.3%)	42.7%
11-20	205 (4.4%)	2944 (16.1%)	19.4%
21-30	63 (1.4%)	1567 (8.6%)	18.4%
31-40	21 (0.5%)	731 (4.0%)	15.0%
41-50	18 (0.4%)	814 (4.5%)	12.7%
51-60	16 (0.3%)	882 (4.8%)	10.1%
61-70	3 (0.1%)	208 (1.1%)	3.4%
71-80	8 (0.2%)	594 (3.3%)	4.8%
81-90	3 (0.1%)	253 (1.4%)	4.8%
91-100	2	188 (1.0%)	4.8%
101-110	1	110 (0.6%)	13.6%
111-120	1	112 (0.6%)	4.5%
151-160	2	311 (1.7%)	26.8%
合計	4614 (100%)	18269 (100%)	

表 16: 投稿回数月 30 回以上の投稿者のイニシャル・フォロー投稿別投稿傾向

投稿者	I	イニシャル投稿				フォロー投稿			
		N	G	B	P	N	G	B	P
ogata	0.36	57	13	219	4.34	103	17	136	1.82
koichi	0.18	27	15	45	2.16	124	31	157	1.77
hifumi	0.04	5	4	6	1.77	107	13	194	2.32
void	0.14	15	8	51	3.91	95	24	98	1.53
nishida	0.04	4	1	4	1.47	91	4	128	1.91
tsutaki	0.05	5	4	2	0.99	88	18	72	1.32
nobuo	0.00	0	0	0	0.00	87	19	179	2.56
kono	0.04	3	3	4	1.98	81	28	128	2.08
h_siobar	0.11	9	1	5	1.02	73	5	61	1.34
toh-hei	0.01	1	1	0	0.74	78	13	165	2.61
n201201	0.14	11	3	18	2.10	66	19	262	4.46
ysato	0.07	5	2	3	1.11	71	6	130	2.34
hisadome	0.00	0	0	0	0.00	74	13	92	1.74
miyano	0.10	7	3	9	1.76	66	6	81	1.73
ten	0.01	1	1	1	1.74	72	10	103	1.93
e8301	0.06	4	2	2	0.98	67	4	57	1.35
g93067tn	0.00	0	0	0	0.00	71	1	157	2.72
ikedo	0.00	0	0	0	0.00	70	5	179	3.06
tasawa	0.07	5	2	12	2.83	64	2	78	1.72
shin	0.03	2	1	2	1.39	67	3	107	2.10
e9015	0.00	0	0	0	0.00	60	1	35	1.08
igeta	0.02	1	1	0	0.69	58	7	92	2.08
ryuk	0.08	5	3	9	2.22	54	10	101	2.36
suzuki	0.02	1	1	2	2.43	57	10	144	3.02
rt	0.04	2	1	5	3.24	55	6	121	2.71
bandou	0.04	2	1	3	2.23	55	4	91	2.15
ippi	0.00	0	0	0	0.00	56	6	66	1.68
hironobu	0.13	7	4	25	4.01	48	12	57	1.68
s916541	0.02	1	1	1	2.00	53	9	102	2.42
tss	0.00	0	0	0	0.00	54	13	54	1.49
t-oshiba	1.00	53	3	2628	50.08	0	0	0	0.00
alceste	0.04	2	2	2	1.64	51	21	58	1.64
shoji	0.08	4	2	4	1.50	48	5	70	1.96
funa	0.12	6	3	4	1.18	46	13	44	1.45
wind	0.02	1	1	4	4.06	51	12	72	1.92
inaba	0.04	2	1	5	3.19	49	3	95	2.44
wataru	0.34	17	7	236	14.36	33	6	76	2.82
f51586	0.12	6	5	15	3.06	44	7	155	4.03
wakita	0.00	0	0	0	0.00	50	12	75	1.99
masda	0.04	2	2	1	1.09	47	8	31	1.17
hirano	0.17	8	7	16	2.52	40	11	70	2.25
atlan	0.02	1	1	1	1.34	46	14	53	1.66
morioaka	0.09	4	3	26	6.98	42	6	212	5.56
nao	0.07	3	3	2	1.21	43	12	53	1.73
tosaki	0.00	0	0	0	0.00	45	5	79	2.25
hira	0.02	1	1	1	1.64	44	4	131	3.48
iwasa	0.18	8	7	8	1.50	37	11	47	1.77
shioya	0.72	31	1	47	2.00	12	2	25	2.59
takehiro	0.38	16	6	19	1.69	26	5	54	2.59
yamamura	0.07	3	1	6	2.55	39	6	78	2.50
g0f084	0.00	0	0	0	0.00	42	6	107	3.04
jiro	0.07	3	3	4	1.70	39	11	35	1.39
takagi	0.00	0	0	0	0.00	41	5	104	3.05
umeyama	0.00	0	0	0	0.00	41	3	67	2.12
harigaya	0.03	1	1	1	1.88	39	9	43	1.61
numomura	0.15	6	3	9	2.03	34	5	87	3.06
ryu	0.00	0	0	0	0.00	40	2	76	2.39
fuminori	0.00	0	0	0	0.00	39	7	26	1.17
ohta	0.03	1	1	2	2.42	37	13	48	1.79
kasama	0.00	0	0	0	0.00	37	11	55	2.00
kaleb	0.00	0	0	0	0.00	37	1	45	1.72
issei	0.67	24	3	38	2.07	12	5	21	2.21
tamura	1.00	35	5	35	1.51	0	0	0	0.00
sheeran	0.03	1	1	0	0.89	34	2	88	3.09
kim	0.03	1	1	1	1.80	34	15	45	1.81
wd	0.00	0	0	0	0.00	34	4	69	2.52
komeda	0.06	2	2	30	15.42	31	10	27	1.38
sasayama	0.27	9	5	7	1.25	24	9	34	1.92
masatake	0.06	2	2	3	1.87	30	7	33	1.60
uno	0.03	1	1	1	1.12	31	6	25	1.32
nakashim	0.03	1	1	14	14.22	30	7	25	1.33
aoki	0.00	0	0	0	0.00	31	7	34	1.59
mchiba	0.71	22	1	493	22.91	9	4	21	2.88
yamayasu	0.00	0	0	0	0.00	31	4	25	1.30
yuki	0.06	2	2	2	1.58	29	11	39	1.83
tacchan	0.00	0	0	0	0.00	30	7	37	1.75
wasisaka	0.03	1	1	1	1.12	29	7	32	1.61
akiba	0.10	3	2	5	2.05	27	3	34	1.76
kurata	0.17	5	4	5	1.60	25	14	27	1.59
lala	0.00	0	0	0	0.00	30	14	33	1.60
enami	0.23	7	4	106	15.71	23	8	20	1.37
合計		475		4210		3908		6267	

- I イニシャル率
- N 記事数
- G ニュースグループ数
- B バイト数 (KB)
- P 1記事当りのバイト数 (KB)

表 17: 月 30 回以上の投稿者の投稿傾向

記事数	バイト数	1記事あたりの 平均バイト数	投稿者数
i > f	i > f	i > f	3
i > f	i > f	i < f	2
i > f	i < f	i > f	0
i > f	i < f	i > f	0
i < f	i > f	i > f	4
i < f	i > f	i < f	0
i < f	i < f	i > f	19
i < f	i < f	i < f	53
合計			81

- イニシャル投稿数がフォロー投稿数よりも多い投稿者は、81人中5人である。
- 1記事あたりの平均バイト数は、フォロー投稿の方が多。

最後の事実は、フォロー投稿の場合は、同サブジェクトの以前の記事を文中に参照して投稿するケースが多いことが影響していると考えられる。

7 特徴的なニュースグループ

7.1 ニュースグループの統計量とニュースグループの特徴の関係

fjの各ニュースグループは、それぞれ異なった目的とそれに付随する特徴を持っている。fj.forsaleの目的は、売買の仲介をすることであり、そのニュースグループに流れる記事のほとんどは、売買の掲示記事に限定されるといった特徴を持っている。

以下では、このようなニュースグループ毎の特徴とニュースグループ毎の統計量との間に何か関係があるかどうかを調べてみた。具体的には、各ニュースグループに対して、イニシャル投稿率、投稿者多様度、平均バイト数、の3つの統計量を求め、これらの値とニュースグループの特徴の関係を調査した。なお、調査は、全ニュースグループのうち、1ヶ月の記事数が31記事以上のニュースグループ(全ニュースグループ数の約60%、全記事数の約96%)を対象にした。調査対象としたデータは、1993年10月のものであるが、必要に応じて、11月および12月のデータについても調査した。

イニシャル投稿率と投稿者多様度、平均バイト数とイニシャル投稿率、投稿者多様度と平均バイト数をプロットしたものを、それぞれ図2、図3、図4に示す。

7.2 イニシャル投稿率

イニシャル記事とは、投稿者が自らサブジェクト名を設定して投稿した記事である。イニシャル投稿率は、ニュースグループ中の全記事数に対するイニシャル記事数の比率である。図2より、多くのニュースグループにおいて、イニシャル投稿率は0.06 ~ 0.5付近に集中していることがわかる。イニシャル投稿率が、この範囲外の値をとるニュー

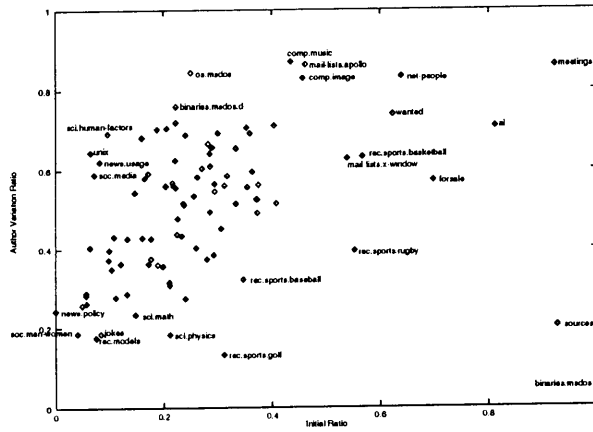


図 2: イニシャル投稿率と投稿者多様度

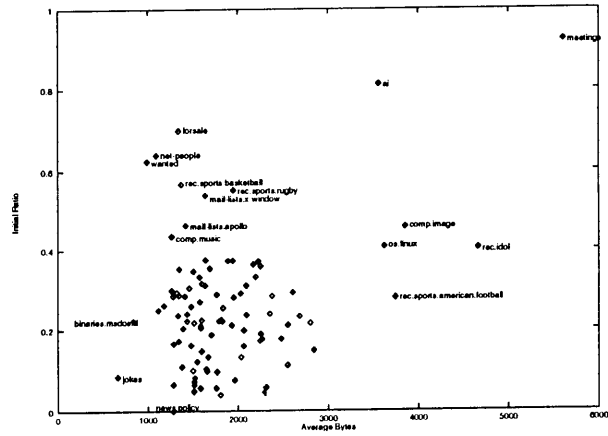


図 3: 平均バイト数とイニシャル投稿率

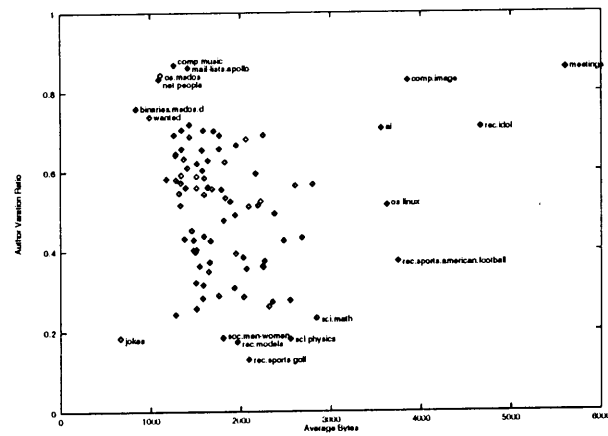


図 4: 投稿者多様度と平均バイト数

表 18: 特徴的なイニシャル投稿率を示すニュースグループ

ニュースグループ名	イニシャル投稿率			
	10月	11月	12月	平均
fj.news.policy	0.00	0.07	0.05	0.04
fj.soc.culture	0.04	0.04	0.03	0.04
fj.soc.men-women	0.04	0.05	0.03	0.12
fj.rec.tv	0.05	0.10	0.15	0.10
fj.living	0.06	0.10	0.08	0.08
fj.news.group.soc	0.06	0.07	0.01	0.14
fj.soc.misc	0.06	0.05	0.01	0.04
fj.soc.smoking	0.06	0.00	0.05	0.04
fj.unix	0.06	0.61	0.12	0.26
fj.mail-lists.x-windows	0.54	0.51	0.46	0.50
fj.rec.sports.rugby	0.55	0.41	0.35	0.44
fj.rec.sports.basketball	0.57	0.43	0.25	0.42
fj.wanted	0.62	0.56	0.62	0.6
fj.net-people	0.64	0.02	0.62	0.42
fj.forsale	0.70	0.72	0.72	0.71
fj.ai	0.81	0.64	0.35	0.60
fj.meeting	0.92	0.94	0.98	0.95
fj.sources	0.93	0.76	0.95	0.88
fj.binaries.msdos	1.00	1.00	1.00	1.00

スグループは、19グループある。これらのニュースグループに対しては、11月・12月のイニシャル投稿率を、さらに調査した。その結果を表18に示す。

この表から、fj.rec.tv、fj.living、fj.unix、fj.mail-lists.x-windows、fj.rec.sports.rugby、fj.rec.sports.basketballの6つのグループのイニシャル投稿率は、10月においては特徴的な値を示していたが、3ヶ月の平均値では、多くのニュースグループと同じイニシャル投稿率(0.06～0.5)をとったことがわかる。これらのニュースグループが、一時的に低いイニシャル投稿率をとった理由としては、他の月に比べて、長い議論を引き起こす話題があったこと(fj.unix、fj.living)などが挙げられる。逆に、高いイニシャル投稿率をとった理由としては、ニュースグループが新設であったこと(fj.rec.sports.rugby、fj.rec.sports.basketball)、長い議論を引き起こす話題があまりなかったこと(fj.mail-lists.x-windows)などが考えられる。このことから、あるニュースグループで、通常とは異なる事態(議論を呼ぶ話題の提供、新しいニュースグループの誕生、ニュースグループの分割)などが起きると、そのニュースグループのイニシャル投稿率は、他のニュースグループと比べて特異な点をとる、ということが出来るだろう。

3ヶ月の調査期間を通してイニシャル投稿率が恒常的に特異な値をとったニュースグループを、表19に示す。

ニュースグループのイニシャル投稿率が低い値をとるということは、ニュースグループ中に、以前の記事に対するリプライ記事が多いということである。1つの話題について多くのリプライ記事が続く「議論・おしゃべり」型のニュースグループでは、イニシャ

表 19: イニシャル投稿率において特徴的なニュースグループ

イニシャル投稿率	ニュースグループ名	イニシャル投稿率		
		10月	11月	12月
0 ~ 0.06	fj.news.policy	0.00	0.07	0.05
	fj.soc.culture	0.04	0.04	0.03
	fj.soc.men-women	0.04	0.05	0.03
	fj.news.group.soc	0.06	0.07	0.01
	fj.soc.misc	0.06	0.05	0.01
	fj.soc.smoking	0.06	0.00	0.05
0.06 ~ 0.5	多くのニュースグループが、この間にある			
0.5 ~ 0.9	fj.wanted	0.62	0.56	0.62
	fj.net-people	0.64	0.62	0.62
	fj.forsale	0.70	0.72	0.72
	fj.meeting	0.92	0.94	0.98
	fj.sources	0.93	0.76	0.95
	fj.binaries.msdos	1.00	1.00	1.00

ル投稿率は低い値をとる。表 19に示すイニシャル投稿率が低い6つのニュースグループは、いずれも「議論・おしゃべり」型である。

一方、ニュースグループのイニシャル投稿率が高い値をとるということは、新しく話題を設定する記事が多く、以前の記事に対するリプライ記事が少ないことを示す。つまり、このニュースグループに投稿される記事の多くは、同じニュースグループ中の他の記事との繋がりがなく、新しい情報を表すものと考えられる。このような特徴は、「アナウンス」型ニュースグループ、「ソースコード用」ニュースグループに見られる。イニシャル投稿率が高いニュースグループのうち、fj.sources と fj.binaries.msdos は、ソースコード用ニュースグループであり、fj.meetings、fj.net-people、fj.wanted、fj.forsale は、「アナウンス」型ニュースグループである。

7.3 投稿者多様度

投稿者多様度は、1つのニュースグループ中の投稿者数と記事数の比率であり、投稿者の多様度を示す。図 4より、多くのニュースグループにおいて、投稿者多様度は0.2 ~ 0.8 付近に集中していることがわかる。投稿者多様度が、この範囲外の値をとるニュースグループは12グループある。これらのニュースグループに対しては、11月と12月の投稿者多様度をさらに調査した。その結果を、表 20に示す。

この表から、fj.binaries.msdos.d、fj.rec.models、fj.sci.physics、fj.soc.men-women、fj.comp.image、fj.os.msos、fj.comp.music、の7つのグループの投稿者多様度は、10月においては、特徴的な値を示していたが、3ヶ月の平均値では、多くのニュースグループと同じ投稿者多様度(0.2 ~ 0.8)をとったことがわかる。これらのニュースグループの投稿者多様度が、一時的に低い値をとった理由としては、多くの人が興味を抱き、質問や議論をするような話題提供が少なかったことなどが考えられる。逆に、一時的に高い値をとった理由としては、多くの利用者が興味を抱くような話題提供が考えられる。このことから、

表 20: 投稿者多様度が特異な値を示したニュースグループ

ニュースグループ名	投稿者多様度			
	10月	11月	12月	平均
fj.binaries.msdos.d	0.05	0.45	0.25	0.25
fj.rec.sports.golf	0.13	0.19	0.16	0.16
fj.rec.models	0.18	0.23	0.14	0.18
fj.sci.physics	0.18	0.35	0.43	0.32
fj.soc.men-women	0.19	0.27	0.26	0.24
fj.jokes	0.19	—	0.13	0.16
fj.comp.image	0.83	0.78	0.69	0.80
fj.net-people	0.83	0.90	0.93	0.89
fj.os.msdos	0.84	0.72	0.75	0.77
fj.mail-lists.apollo	0.86	—	0.89	0.88
fj.meetings	0.86	0.81	0.93	0.87
fj.comp.music	0.87	0.60	0.67	0.80

表 21: 投稿者多様度において特徴的なニュースグループ

投稿者多様度	ニュースグループ名	投稿者多様度		
		10月	11月	12月
0 ~ 0.2	fj.rec.sports.golf	0.13	0.19	0.16
	fj.jokes	0.19	—	0.13
0.2 ~ 0.8	多くのニュースグループが、この間の値をとる			
0.8 ~ 1.0	fj.net-people	0.83	0.90	0.93
	fj.mail-lists.apollo	0.86	—	0.89
	fj.meetings	0.86	0.81	0.93

あるニュースグループで、通常とは異なる事態(活発な議論を呼ぶ話題の提供)などが起きると、そのニュースグループの投稿者多様度は、他のニュースグループと比べて特異な点をとる、ということが出来るだろう。3ヶ月の調査期間を通して投稿者多様度が恒常的に特異な値をとったニュースグループを、表 21 に示す。

ニュースグループの投稿者多様度が恒常的に低い値をとるということは、投稿が固定的なメンバーに限られ、同一投稿者が何度も投稿を行なっていることを意味している。投稿者多様度は、1 投稿者の投稿回数が多いほど 0 に近づく。投稿者が限定される専門性・趣味性の高いニュースグループでは、投稿者多様度は低い値をとる。投稿者多様度の低い 2 つのニュースグループ、fj.rec.sports.golf、fj.jokes は、専門性・趣味性の高いニュースグループであるため、投稿者が固定的になっていると考えられる。

一方、ニュースグループの投稿者多様度が恒常的に高い値をとるということは、多様な投稿者が投稿しているということである。全ての記事が異なる投稿者によって投稿された場合、投稿者多様度は 1 になる。同一投稿者が、「アナウンス」型の記事を短期間に何度も投稿することは、あまりないと考えられる。したがって、アナウンス型の記事を中心とするニュースグループでは、投稿者多様度は、高い値をとる。表 21 より、投稿者多様度が高いニュースグループは、fj.net-people、fj.mail-lists.apollo、fj.meetings、で

表 22: 平均バイト数が特異な値を示したニュースグループ

ニュースグループ名	イニシャル投稿率			
	10月	11月	12月	平均
fj.jokes	664	—	737	701
fj.binaries.msdos.d	841	1608	1689	1379
fj.wanted	991	1159	1012	1054
fj.net-people	1094	943	1029	1022
fj.ai	3562	2666	2188	2805
fj.os.linux	3624	2147	2467	2746
fj.rec.sports.am.ftball	3746	3143	2605	3164
fj.comp.image	3855	2069	1354	2426
fj.rec.idol	4663	2674	2644	3327
fj.meetings	5604	4298	4244	4709
fj.sources	31940	25373	1188165	415159
fj.binaries.msdos	46119	45938	1817826	636628

ある。fj.net-people と fj.meetings は、「アナウンス」型のニュースグループであると考えられる。fj.mail-lists.apollo は、メイリングリストに投稿された記事を転載したニュースグループである。ニュース上に記事が転載されるメイリングリストでは、参加者が能動的であり、メンバーが記事を読む率も高く、活発に利用されていると予想される。そのため、投稿者多様度が高い値をとると考えられる。

7.4 平均バイト数

平均バイト数とは、1記事あたりのバイト数である。図4より、多くのニュースグループにおいて、平均バイト数は、1100 ~ 2800 付近に集中していることがわかる。平均バイト数がこの範囲の値をとらないニュースグループは、12グループある。これらのニュースグループに対しては、11月および12月の平均バイト数をさらに調査した。その結果を表22に示す。

この表から、fj.binaries.msdos.d、fj.os.linux、fj.comp.image、の3つのニュースグループの平均バイト数は、10月においては特徴的な値を示していたが、3ヶ月の平均値では、多くのニュースグループと同じ1100 ~ 2800の値をとることが分かった。これらのニュースグループが、10月の時点で特徴的な値をとった理由として、話題の性質により一時的に短い記事が増えたこと (fj.binaries.msdos.d)、一時的に会告記事(会議プログラムを含む)数が増えたこと (fj.comp.image)、FAQ記事が投稿されたこと (fj.os.linux)、が挙げられる。また、fj.rec.sports.american.football については、10月の平均バイト数は3746で、他のニュースグループに比べて特徴的な値をとっていたが、11月は3143、12月は2605と、徐々に他のニュースグループ群の平均バイト数に近い値となっている。これは、ニュースグループが1993年9月に新設されたため、10月には一時的に、自己紹介やニュースグループへの希望が書かれた長い記事が多く投稿されたことや、アメリカンフットボールのシーズン開始時期が10月であるため、シーズン中のスケジュールなどが多数投稿されたことが原因であると考えられ

表 23: 平均バイト数において特徴的なニュースグループ

平均バイト数	ニュースグループ名	平均バイト数		
		10月	11月	12月
0 ~ 1100	fj.jokes	664	—	737
	fj.wanted	991	1159	1012
	fj.net-people	1094	943	1029
1100 ~ 2800	多くのニュースグループが、この間の値をとる			
2800 ~	fj.meetings	5604	4298	4224
	fj.sources	31940	25375	1188165
	fj.binaries.msdos	46119	45938	1817826

る。このことから、平均バイト数においても、あるニュースグループで、通常とは異なる事態(グループの新設、分割、FAQの投稿)などが起きると、そのニュースグループの平均バイト数は、他のニュースグループと比べて特異な点をとる、ということが出来るだろう。また、いくつかのニュースグループにおいては、特殊な投稿(特定の会議の会告記事、スポーツシーズンの開始にあたって投稿されるスケジュールなど)のために、ある月に限って平均バイト数が増えることがある。

3ヶ月の調査期間を通して、平均バイト数が恒常的に特異な値をとったニュースグループを、表 23 に示す。

一方、平均バイト数が高い値をとるということは、記事の長さが長いということである。このような性質を示すのは、ソースコード用ニュースグループである。平均バイト数の高い3つのニュースグループのうち、fj.sources、fj.binaries.msdos は、ソースコード用ニュースグループである。残りの fj.meetings は、プログラムやアプリケーションフォームを含む会告記事が流れるというが、記事の長さを長くしていると考えられる。

7.5 特徴的なニュースグループ群

これまでの調査から明らかになった特徴的なニュースグループのまとめを、表 24 に示す。

8 まとめ

本稿では、投稿行動(投稿量)から見た fj の現状を明らかにするために行なった調査について述べた。その結果を以下にまとめる。

- 1ヶ月の情報の規模は、約 36MB、17000 記事、4400 投稿者である。
- 1日の情報の規模は、週日において、約 1.6MB、800 記事、200 投稿者である。これは、全国紙朝刊 1 紙分に相当する。
- バイト数・記事数・投稿者とも、週日に比べ週末には減少するが、1 投稿者あたりの投稿バイト数は、週末に約 1.5 倍に増加する。
- バイト数・記事数が多いグループと少ないグループが混在している。全バイト数の約 1/2 は、バイト数・記事数ともに上位の 30 グループで占められている。また、

表 24: fj における特徴的なニュースグループ

ニュースグループ名	記事の特徴	統計上の特徴
fj.news.policy fj.soc.culture fj.soc.men-women fj.news.group.soc fj.soc.misc fj.soc.smoking	議論やおしゃべりの記事が中心である	イニシャル率が低い
fj.rec.sports.golf fj.jokes	専門性・趣味性の高い記事が中心である	投稿者多様度が低い
fj.wanted fj.net-peole fj.forsale	短い情報を提示する記事が中心である	イニシャル率が高い
fj.meetings	会告記事が中心である	イニシャル率が高い 投稿者多様度が高い 平均バイト数が高い
fj.sources fj.binaries.msos	ソースコードが中心である	イニシャル率が低い 平均バイト数が高い
fj.mail-lists.apollo	メイルリストの記事を転載するニュースグループである	投稿者多様度が高い

全バイト数の約 1/5 は、ソースコード用ニュースグループで占められている。

5. サブジェクトの記事数にはばらつきがある。このことは、「アナウンス」型・「質問 - 応答」型・「議論・おしゃべり」型のサブジェクトが混在していることを示唆する。
6. 全サブジェクトの約半数は、1 記事で構成されている。
7. 1 つのサブジェクトへの記事集中が、fj 全体の記事数に対して占める割合はかなり低い。(調査期間中最も記事数の多かったサブジェクトの記事数は、調査期間中の全記事数の 0.8% である)
8. 1 ヶ月に 1 回の投稿者は、全投稿者数の約 1/2 を占める。頻繁投稿者 (平均 3 日に 1 度以上投稿する投稿者) の全投稿者数に対する比率は、1 割未満である。投稿回数が増えるに従い、1 投稿者の全投稿中に占めるイニシャル投稿の割合は減少する。
9. ニュースグループの統計量から、そのニュースグループの特徴を推測することが出来る。また、ニュースグループに、通常とは異なる事態が起こると、そのニュースグループのイニシャル率・投稿者多様度・平均バイト数の値が変化する。
10. 特徴的なニュースグループとしては、「議論・おしゃべり」型ニュースグループ、「専門性・趣味性」ニュースグループ、掲示情報用ニュースグループ、会告記事用ニュースグループ、ソースコード用ニュースグループ、メイルリスト転載ニュースグループ、がある。

参考文献

- [1] 川上, 川浦, 池田, 古川, 電子ネットワーキングの社会心理, 誠信書房, 1993.
- [2] 野島久雄, 二上俊嗣, JUNET における利用者環境の変化, 情報処理学会第 35 回全国大会, Vol.3, No.10, pp589-590, 1987.
- [3] 佐藤円, 電子ニュースにおけるダイジェスト機構の提案と実現, 修士論文, 北陸先端科学技術大学院大学情報科学研究科, 1994.