

「地震」に関する新聞記事における 報道内容の経時変化調査

福 島 孝 博

Time lapse analysis of information in newspaper articles on earthquakes

Takahiro Fukushima

概 要

時系列変化を記録したデータにおいてその情報内容の変化を把握するため、大きな自然災害が発生した際に、どのような情報がいつ報道されるのかを詳しく調査する。自然災害としては、2004年10月に発生した新潟県中越地震を取り上げ、この地震の新聞での報道内容について、その内容を分類し時間経過に沿って整理した。

調査の結果は、地震情報、被害情報は発生当初に多く、援助情報は発生3日目から増加するなど、予想のとおり出現傾向であった。しかし、生活に関する情報は予想よりも多く現れ、しかも、災害発生からかなり早い段階から被災者にとっての生活情報が出現していたことが判明した。

1. は じ め に

本論文では、自然災害に関する報道、特に2004年10月23日に発生した新潟中越地震に関する新聞記事での情報の種類と時系列による情報の内容の変化について調査をし、その結果を報告する。

自然言語処理の分野では、あるとピックに関する複数の新聞記事をまとめる（要約する）システムに関する研究・開発が近年盛んに行われている（[1]）。新聞記事をまとめるにあたり、どのような情報がどのようなタイミングで出現するかを調査しておくことは、システムを開発

する上でも重要であり、同種類の情報をまとめるに当たり、参考となる。

また、新聞記事のように時系列を追って書かれている情報を、あるイベント（例、新潟中越地震）に関する情報に整理することは、新聞記事を利用したイベント時系列データマイニングにおいても注目されている（[2]）。時系列に沿って、複数のイベントがどのような関係にあるかを分かることは、時系列に整理された情報からのデータマイニングにとって重要なことである。

以下に、調査の対象となったデータに説明し、調査の方法、調査の結果、結果に関する考察を記述する。

2. 対象データ

新潟中越地震は、2004年10月23日に発生し、新聞報道は、翌日の24日から始まり、その後長期間にわたって報道が続いた。

調査の対象としたのは、2004年度後期（7月から12月）の毎日新聞記事データベース¹の新潟中越地震に関する記事だけを抽出した。抽出にあたっては、記事のタイトルに「新潟中越地震」とあるものだけを収集した。

結果は、10月24日から12月31日までの69日間の記事であり、発生時からの経過日数と記事数をグラフにすると以下のとおりである。

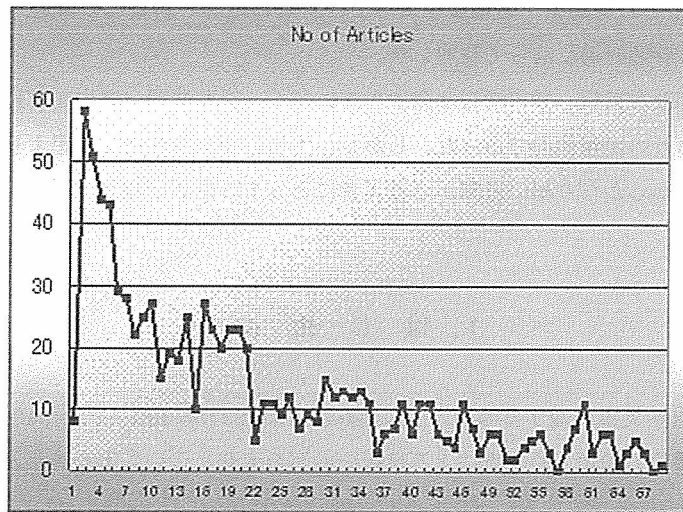


図1 新潟中越地震に関する新聞記事の経過日数と記事数

1 CD-毎日新聞2004データ集、毎日新聞社

グラフより分かるのは、地震発生の翌日（報道の初日）は、8記事と記事数がそれ程多くないが、翌日には記事数が58と最大となり、その後徐々に減っている。記事の総合計数は、881であった。

福島〔3〕では、台湾地震に関する新聞記事における報道内容の分類を、記事のタイトル情報だけを使って行っている。本研究では、記事内容全体を対象としている。具体的には、毎日新聞記事データベースに登録の記事中の情報分類項目が \C 0\、\A D\、\A F\、\T 1\、\S 1\、\T 2\に現れる情報である。C 0は記事固有の番号であり、T 1は記事のタイトル（見出し）、T 2は記事の文章（1段落に相当）を示す。T 2には表や、写真も含まれる。また、\A F\では、日付と朝夕刊の区別が記載されている。

尚、A Dなどその他の項目は今回の調査内容に含めなかったが、参考情報としてデータに含めている。以下の新潟中越地震の毎日新聞における第一報（一部省略）を掲載する。

 \C 0\041024003

\A D\01

\A F\041024M01

\T 1\新潟中越地震：新潟で震度6強3回 10人死亡、640人以上けが、10人不明

\S 1\ '04. 10. 24 朝刊 1頁 写図有 (全1399文字)

\T 2\ ◇M 6. 8、上越新幹線脱線

\T 2\ 23日午後5時56分ごろ、新潟県を中心に強い地震があり、新潟県小千谷市で震度6強の揺れを観測した。気象庁によると、震源地は同県中越地方で、震源の深さは約20キロ、地震の規模を示すマグニチュード(M)は6・8と推定される。同県内で10人が死亡、10人が行方不明となっている。毎日新聞の24日午前0時半現在のまとめでは、けが人は640人を超えた。住宅倒壊や土砂崩れが各地で発生し、上越新幹線は浦佐―長岡間で脱線した。その後も震度6強2回を含む余震が相次いでいる。

(中略)

\T 2\ 気象庁によると、24日午前0時までに有感の余震は163回に達した。山本雅博・地震津波監視課長は「今後1週間は、大きな地震が起きる可能性がある」と警戒を呼びかけている。

\T 2\ 新潟県は23日、地震による被害の大きい小千谷市、長岡市など4市2町1村に災害救助法の適用を決めた。

 \T 2\

\T 2\ ■今回の地震と主な余震

\T 2\	発生時間	規模	震度
\T 2\本震	午後5時56分	M 6. 8	6強
\T 2\余震	午後6時3分	M 6. 2	5強
\T 2\	午後6時12分	M 5. 9	6強
\T 2\	午後6時34分	M 6. 3	6強
\T 2\	午後6時36分	M 5. 0	5弱
\T 2\	午後6時57分	M 5. 1	5強

「地震」に関する新聞記事における報道内容の経時変化調査

＼T 2＼ 午後 7 時 36 分 M 5. 2 5 弱

＼T 2＼ 午後 7 時 46 分 M 5. 9 6 弱

＼T 2＼ 写真説明 新潟県の浦佐－長岡間で脱線した上越新幹線「とき325号」＝23日午後 9 時 4 分、本社ヘリから佐々木順一写す

実施に詳細な調査の対象としたのは、881記事のうち、10月24日から10月30日までの7日間の記事とした。

また、調査にあたり、以下の点について考慮した。

- ・ T 1 (タイトル)、T 2 (記事の文章)、写真説明は別々に集計した。
- ・ T 2 は、T 2 (1段落に相当) 毎に、つまり、一段落毎に分析した。
- ・ 記事中の表は一まとめにして分析した。
- ・ 記事中の小見出し (◇などではじまる) は無視した。
- ・ 記事の中には、少数であるが、タイトル (T 1) と写真だけのものや、タイトル (T 1) はあるが著作権の関係で、記事の文章が掲載されていないものがある。また、タイトルには「新潟中越地震」の文句が入っているが、記事全体が全く地震に関する情報を含んでないものがある (例、サッカー J1 新潟の連勝ストップという内容の記事)。これらは調査の対象外とした。
- ・ 「社説」は対象外とした (該当するのは 3 記事、1025034、1026033、1028025)

この結果、調査期間中の記事のうち13記事が対象外となり、248記事が対象データとなった。以下に朝夕刊別の記事数を示す。表中 (Time の列) のMは朝刊を、Eは夕刊を表しており、4桁の数字ははじめの2桁が月をその後の2桁が日付を表している。

Time	No of Articles	Time	No of Articles
1024M	8	1028M	27
1025M	35	1028E	13
1025E	20	1029M	20
1026M	29	1029E	6
1026E	18	1030M	22
1027M	34	1030E	6
1027E	10	Total	248

表 1 調査対象記事数

3. 調査・分析方法

調査対象となった記事での T 1、T 2、T 2 での写真説明文について、そこでの情報の内容

を分類を行った。情報内容の分類については、福島（[3]）を参考にし、そこでの分類を一部変更した。以下に本論文での情報の分類区分を示す。

番号	情報分類	詳細
1	地震情報	震度、震源地、規模、津波の危険性
2	余震情報	余震の程度、見通し
3	被害情報	被害が大きい場所 二次災害予防情報（ガス漏れ、通電火災など）
4	生活情報	水、食料、避難所状況、救援物資
5	交通情報	交通機関の復旧の情報
6	医療情報	医療を
7	生活再建情報	罹災証明、家の解体など
8	ボランティア情報	ボランティア活動情報
9	援助活動情報（国内）	地震が発生した国内の援助情報
10	援助活動情報（国際）	国際的援助活動情報
11	解説情報	一般的解説、過去の地震との比較など

表 2 記事中の情報の分類

情報分類 1, 2 は地震に関するものであり、3 は被害についての種々の情報である。3 から 8 は、大きく分けると援助、救助の情報に含まれるものであるが、被災者の視点から特に重要であるとされるため（[4]）、このように細分している。11 は地震、被害、援助の記事に見られる用語の解説や過去の地震との比較など、新潟中越地震に直接関わる情報ではないものが該当する。

4. 調査・分析結果

調査対象の記事について報道されている情報の分類をした結果は以下のとおりである。結果は、T1、T2、T2（写真説明）の3種類に分けて示す。

4. 1 タイトルにおける情報の分類

まず、記事のタイトル（T1）における情報の分類の結果を表3に示す。タイトルにおいては、援助情報が最も多く、次に被害情報が多くなっていることが分かる。

情報の中でも頻度が10回以上のものだけをまとめると、図2となる。

「地震」に関する新聞記事における報道内容の経時変化調査

	1	2	3	4	5	6	7	8	9	10	11
1024M	2	0	6	0	0	0	0	0	2	0	0
1025M	2	1	23	0	0	0	0	1	7	0	0
1025E	1	3	9	0	0	0	0	1	7	0	0
1026M	2	0	10	0	0	0	0	1	14	2	0
1026E	0	0	8	6	0	0	0	0	4	0	0
1027M	2	0	9	3	0	0	4	1	17	0	0
1027E	0	4	3	0	0	0	0	1	3	0	0
1028M	0	2	6	3	0	0	1	1	15	0	0
1028E	0	1	2	1	0	2	1	0	5	0	0
1029M	0	2	8	0	0	0	0	0	11	0	0
1029E	0	0	0	0	0	0	2	0	4	0	0
1030M	3	0	5	1	0	1	1	0	10	1	0
1030E	0	0	2	1	0	1	1	0	0	0	0
Total	12	13	91	15	0	4	10	5	102	3	1

表3 タイトル（T1）における情報の分類

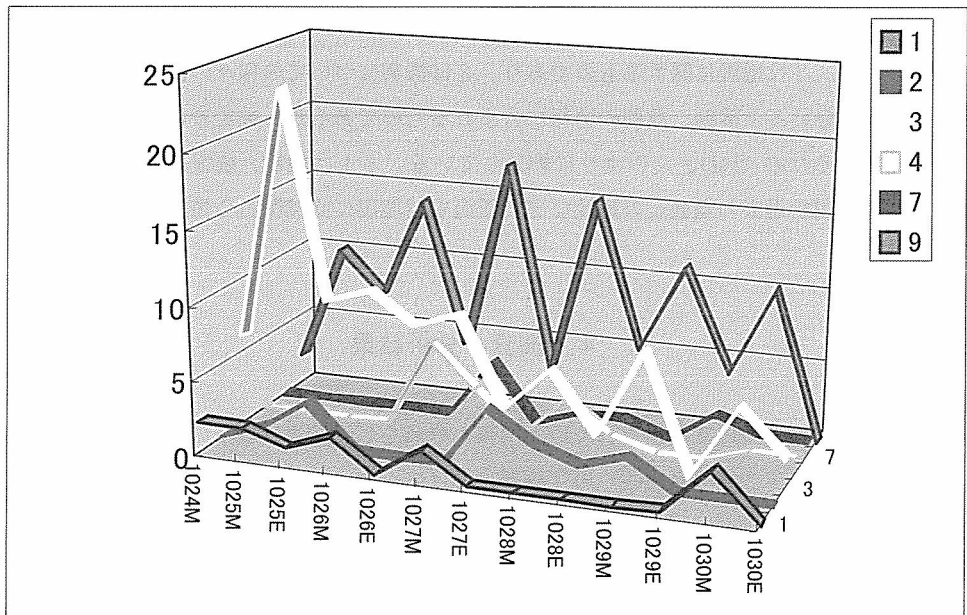


図2 タイトル（T1）における高頻度の情報の出現変化

図2からわかるように、被害情報（3）は、地震発生当初に多く、その後段々と減っている。一方、国内の援助情報（9）は、程度の差はあるが、1週間をとおしてコンスタントに出現している。

4. 2 記事本文（T 2）における情報の分類

次に、記事の本文にあたるT 2における情報を分類した。T 2は、本文の1段落に相当し、記事により多くのT 2を含むものあれが、1個だけの記事もあった。

	1	2	3	4	5	6	7	8	9	10	11
1024M	5	1	41	0	0	0	0	0	5	0	5
1025M	4	6	70	25	2	11	0	5	13	0	13
1025E	4	4	33	23	3	8	0	6	15	0	1
1026M	4	1	38	42	7	0	0	25	38	2	4
1026E	0	3	23	42	2	2	0	6	15	0	0
1027M	7	1	27	45	8	23	11	8	25	0	6
1027E	0	13	13	3	2	0	0	1	9	0	0
1028M	0	10	22	22	8	1	8	5	39	0	12
1028E	0	1	1	10	0	8	1	0	18	0	3
1029M	2	4	35	17	6	6	5	0	22	1	3
1029E	0	0	0	0	0	0	7	0	12	0	1
1030M	9	3	24	21	6	3	7	1	13	2	5
1030E	0	0	4	2	0	3	2	0	0	0	1
Average	2.69	3.62	25.5	19.4	3.38	5	3.15	4.38	17.2	0.38	4.15
Total	35	47	331	252	44	65	41	57	224	5	54

表 4 記事本文（T 2）における情報の分類

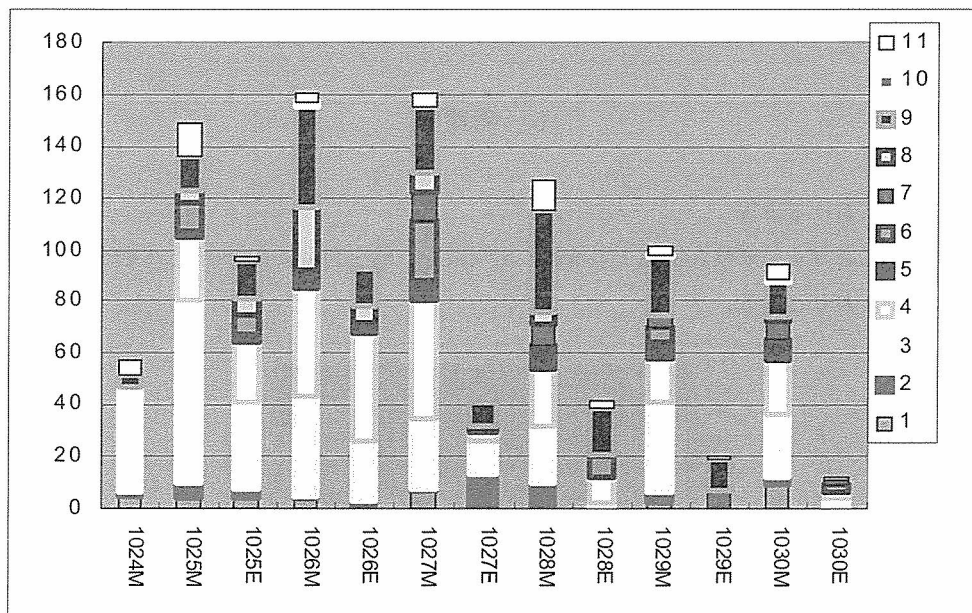


図 3 記事本文における情報の分類（累積グラフ）

「地震」に関する新聞記事における報道内容の経時変化調査

タイトルにおいてと同様に、被害（3）、援助（9）情報が多いが、生活情報（4）も多く、頻度では二番目に多くなっている。

4. 3 写真説明における情報の分類

第三の情報として、T2における写真説明文について分類を行った。これにより、全てではないが、紙面に現れる写真の大部分がどのようなものであったのかが分かる。表5では、朝刊夕刊をまとめ日付毎に集計して頻度を計算している。

	1	2	3	4	5	6	7	8	9	10	11
1024	0	0	4	0	0	0	0	0	0	0	0
1025	2	0	20	5	0	1	0	0	3	0	0
1026	0	0	2	9	1	0	0	2	1	0	0
1027	0	2	5	4	0	0	0	0	2	0	0
1028	0	0	2	5	0	0	0	0	9	0	0
1029	0	0	4	2	0	0	1	0	0	0	0
1030	1	0	3	1	0	0	2	0	1	0	0
Total	3	2	40	26	1	1	3	2	16	0	0

表5 写真説明文における情報の分類

ここでも、被害と援助に関するものが多いが、被害の写真は、25日に圧倒的に多く、生活に関する写真は、25日以降コンスタントに現れ、援助の写真は、28日にピークがある。

5. 考察

タイトルにおける情報

タイトルにおける頻度数10回以上の情報の出現数を、朝夕刊毎に100%として割合を算出すると図4のとおりとなる。

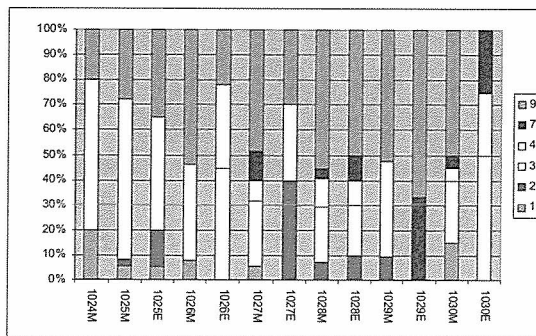


図4 タイトルにおける高頻度情報の構成（朝夕刊別）

タイトルにおいては、割合として、地震発生当初は被害情報が多く、時間が経つにつれて援助情報が多くなっている。1999年発生の台湾大地震の新聞記事の分析〔3〕と比較すると、大体の傾向は同じであるが、生活情報がより多く見られる。また、被害に関する情報は、台湾大地震の報道では、発生から3日目以降はかなり少なくなるが、新潟中越地震の場合は、23日以降も余震がかなり発生し〔5〕、余震による被害が続いたため、被害情報の出現回数が急激に減らずにいることが分かる。

記事本文おける情報

記事本文（T2）においては、11種類の全ての情報が出現した。以下にそれらについて、情報の種類ごとに考察する。

まず、地震に関する情報（地震自体は1、余震は2）だけを抽出すると以下の図のとおりとなる。

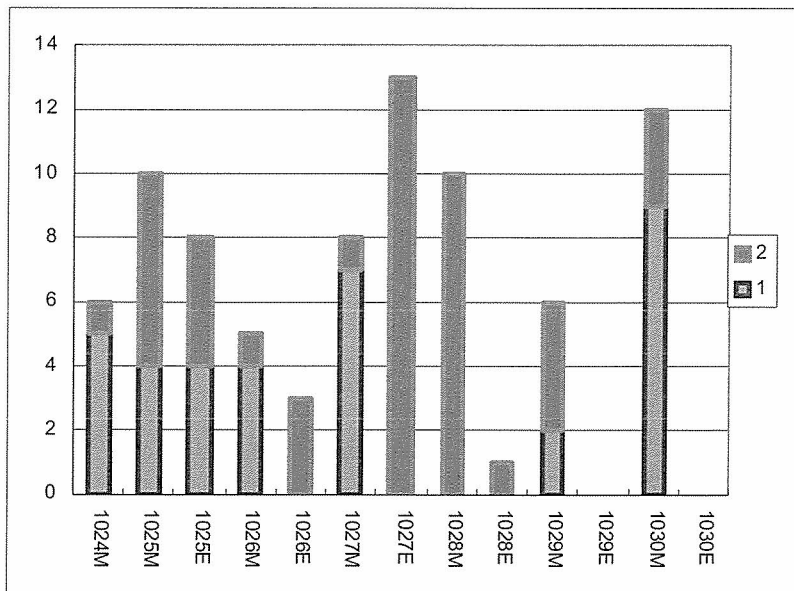


図5 本文中の地震情報の頻度

当然のことながら、地震自体に関する情報は、地震発生当初に報道されているが、発生から1週間経った10月30日に多く見られた。これは、発生した地震の原因などより詳しい情報が判明して報道されたものである。これに対して、余震は断続的に発生したため、連日の報道となっていることが分かる。

次に、被害（3）に関する情報の頻度の時系列による推移をグラフにすると以下のとおりと

「地震」に関する新聞記事における報道内容の経時変化調査

なり、発生翌日が最も多く、その後は漸減している。ここにおいても被害情報が単純に減っていかないのは、余震による被害が発生し続けたためである。

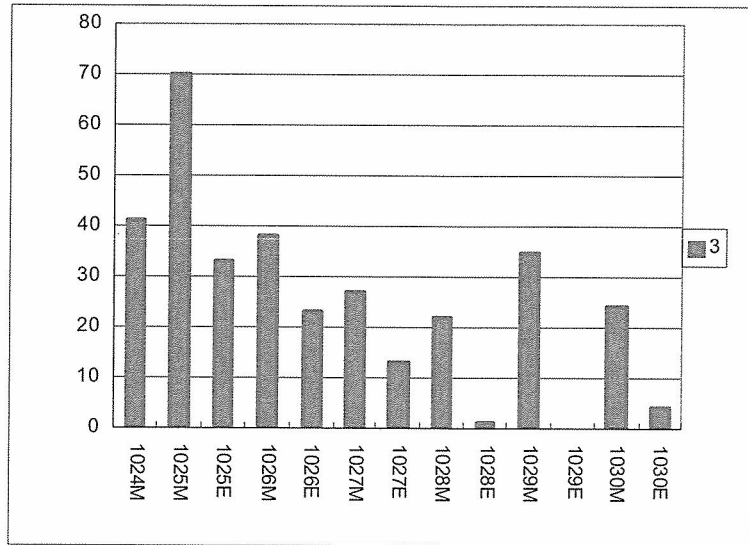


図6 被害情報の頻度（時系列推移）

被災者の生活に関する情報は、避難場所における生活に関するものだけでなく、交通、医療、生活再建、ボランティアについて分類した。(分類区分の4から8)

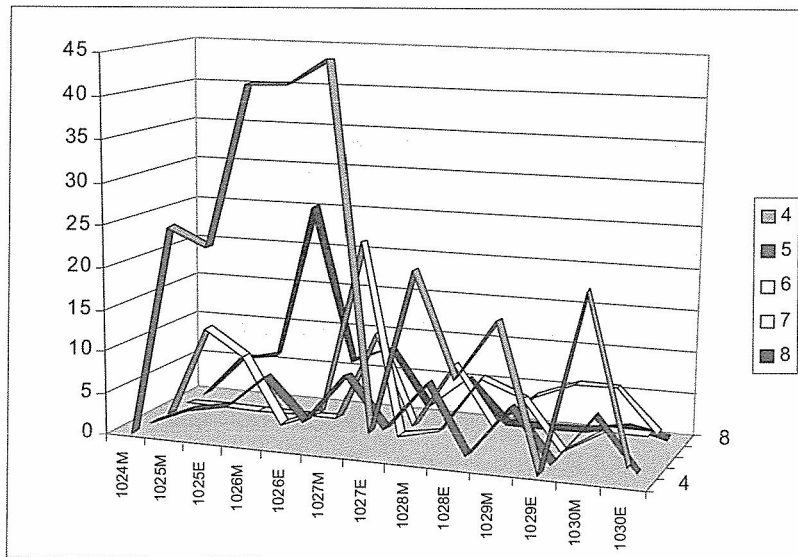


図7 生活情報の頻度

生活情報（４）は、その多くが25, 26, 27日に出現しているが、交通情報（５）については、頻度は少ないがコンスタントに表れている。医療情報（６）は発生４日目の27日に、生活再建情報（７）は27日ピークがあるがその後も続いて現れている。また、ボランティア情報（８）は、26日に集中して出現している。これは、図８に示す援助情報（９、10）がコンスタントに現れているのと比較すると、ボランティアの体制が整っており、地震発生から短時間の間に情報が報道されているためと考えられる。

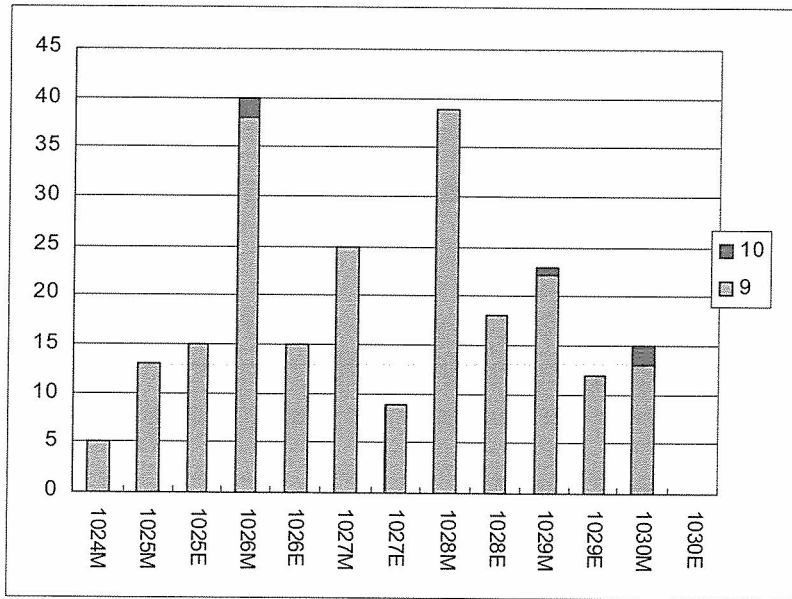


図 8 援助情報の頻度

記事本文の出現する情報について、個別に見たが、各情報の頻度の最大値が表れる日付を色で示すと以下ようになる。このグラフでは、頻度の最大値が表れる日付のセルを黒で、二番目に大きい値が現れる日付を灰色で示している。

「地震」に関する新聞記事における報道内容の経時変化調査

	1	2	3	4	5	6	7	8	9	10	11
1024M											
1025M											
1025E											
1026M											
1026E											
1027M											
1027E											
1028M											
1028E											
1029M											
1029E											
1030M											
1030E											

表6 頻度の最大値出現の時系列変化

色付けにより、よりはっきりと読み取れることをまとめる。

- 地震情報（1，2）は、発生当初よりも後になって多く現れている。
- 被害情報（3）は、報道当日とその翌日の朝刊が多くなっている。
- 生活情報（分類の4から8）は、26日朝刊から28日朝刊に出現頻度のピークが集中している。
- 援助情報（9，10）は、26日の朝刊以降に出現頻度のピークがあるが、二番目に多い日付との差があり、ばらついて出現している。

6. まとめ

本論文では、毎日新聞記事データベースにある新潟中越地震に関する新聞記事のうち、地震の報道当日から一週間の記事について、報道内容の分類を行い、その結果についての考察を行った。

地震情報、被害情報、援助情報は、予想のとおり出現傾向であったが、生活に関する情報が多く現れ、しかも、災害発生からかなり早い段階から医療やボランティアの情報が出現していた。

地震に関する情報は、新聞に現れるものだけではないが、マスメディアの代表の1つである新聞記事においてどのように報道がなされているかを調査した。このような分類の結果は、出来れば、地震防災のためのデータベース([6],[7])と関連付けされることが望ましいと考える。

時系列データから情報抽出を考えると、今回の調査で判明したことを考慮して、時系列に対

応したテンプレート（抽出すべき重要な情報の項目）をモデル化することが必要である。また、モデルかをしたとして、このモデルが他の地震の報道や、他の種類の自然災害の報道に関しても有効であるのかを検討していく必要がある。

本研究は、追手門学院大学「特色ある個人研究費」（2005年度）を用いてなされたものである。

参考文献

- [1] Document Understanding Conferences, <http://duc.nist.gov/>
- [2] 広瀬千夏, 岩沼宏治, 鍋島英知, 「背景記事集合の類似度に基づく新聞記事のクラスタリング」電子情報通信学会, 信学技報 IEICE Technical Report NLC2005-110, 2006-02.
- [3] 福島孝博, 「地震」に関する新聞記事からの情報の調査と分類、追手門学院大学『英語文化学会論集』第15号 pp21-26, 2006年3月20日発行.
- [4] メモリアル・コンファレンス・イン神戸 (編), 土岐憲三, 河田恵昭, 林春男 (監修), 『12歳からの被災者学』, 日本放送出版協会, 2005.
- [5] 多々納裕一, 梶谷義雄, 土屋哲, 「新潟中越地震の社会経済的影響」京都大学防災研究所年報第48A 2005年4月.
- [6] 川方裕則, 「災害に対して安全に暮らすための知識を管理する防災データベース」電子情報通信学会誌, Vol.86 No.5 2005.
- [7] Hironori Kawakata, Yshiaki Kawata, Haruo Hayashi, Takeyoshi Tanaka, Kenneth C. Topping, Katsuya Yamori, Paul Yshitomi, Go Urakawa, and Tomohiro Kugai, Building an Integrated Database Management System of Information on Disaster Hazard, Risk, and Recovery Process, Annuals of Disaster Prevention Research Institute, Kyoto University, No.47C, 2004.