

# 地震記事からの時空間的動向情報可視化についての取り組み

山田 隆志<sup>1</sup> 中野 純<sup>2</sup> 高間 康史<sup>2</sup>

1 都立科学技術大学 〒191-0065 東京都日野市旭が丘 6-6

2 首都大学東京 〒191-0065 東京都日野市旭が丘 6-6

E-mail: ytakama@cc.tmit.ac.jp

あらまし 地震記事に関するタグ付きコーパスから、地震に関する時空間的動向を可視化する地震情報可視化システムについて提案する。「日経平均株価」や「内閣の支持率」のように、動向情報の要約は近年注目を集めつつあるが、地震や台風などは時間的な動向情報だけでなく空間的な動向情報を含んでおり、両者を考慮した可視化が必要となる。提案システムでは、タグ付きコーパス及び Web から地震に関する情報を抽出して可視化する。空間情報を可視化する日本地図と、時系列その他を可視化する折れ線グラフを組み合わせることにより、ユーザのクエリに適した地震情報の可視化が可能であることを示す。

## 1. はじめに

近年ユーザが扱うことの出来る情報は、増加の一途を辿っている。これらの情報の中には、テキストによる表示よりもグラフなどの可視化表現を用いた方が直感的に理解しやすい動向情報も多く含まれる。動向情報とは、「日経平均株価」や「内閣の支持率」のように、数値が時間の経過とともに変動する情報のことを指す。動向情報の要約と可視化は近年注目を集めつつあり、活発に研究されているが[1][2]、対象として選択されているのは、時間情報と数値情報がほとんどである。本稿では、これら時系列情報に空間情報を加えた地震情報を対象とする。地震情報の中でも、指定した地域・期間に発生した一連の地震情報を時間的動向情報、震源地と各地の震度の関係、そして比較的狭い地域で集中的に発生する群発地震を空間的動向情報として定義する。動向情報の抽出を簡易化するために、文書として「動向情報の要約と可視化に関するワークショップ (MuST)」[1]により配布されているタグ付きコーパスを用いる。ユーザからのクエリに基づき地震記事に関する複数の文書から可視化に必要な動向情報を抽出し、日本地図と折れ線グラフを組み合わせることによりユーザの要求を満足させる地震情報可視化システムについて提案する。

本稿では、まず 2 節で提案システムの概要を、3 節、4 節で実際の地震情報の抽出と可視化について述べる。5 節で既存システムとの比較を行い、提案システムの有効性を考察する。

## 2. 提案システムの概要

提案システムの構成を図 1 に示す。システムは新聞記事からデータベースを作成する事前処理と、ユーザからの要求に基づき動向情報を提示するインタフェース部に大別できる。事前処理として、新聞記事データベースから (MuST コーパス) から地震に関する情報 (時間、マグニチュード、震度、地域など) をタグを利用して抽出し、地震情報データベースを構築する。群発地震についての判定もここで行う。

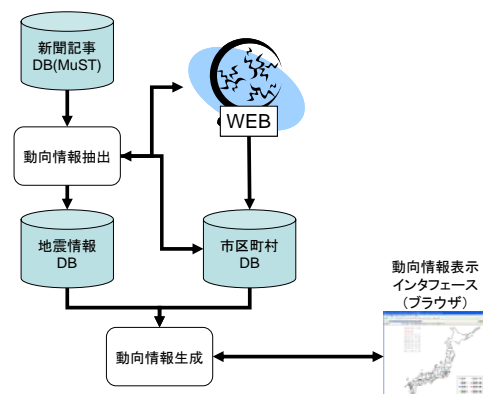


図 1 システム構成図

<input checked="" type="radio"/> マップ <input type="radio"/> グラフ	
マグニチュード	指定しない ▼
検索開始時刻	1998 年 01 月 01 日 00 時 00 分
検索終了時刻	1998 年 12 月 31 日 23 時 59 分
<input type="button" value="検索"/>	

図 2 インタフェース部の入力フォーム

インタフェース部では事前処理で作成した地震記事データベース、および市区町村の緯度・経度情報を格納したデータベースからユーザの要求を満たす動向情報表現を生成し、Web ブラウザを利用した動向情報表示インタフェースによりユーザに提示する。この時、図 2 に示す入力フォームだけでなく、可視化表現からインタラクティブに次の動向情報生成を行うことも可能である。

### 3.地震情報の抽出

一般的に、あるトピックに関する動向情報を抽出するには、そのトピックに関する文書を収集し、必要とされる数値情報・時間情報を抽出する。しかし、本研究では動向情報として、地震情報を対象としているので、これに加え地域情報も抽出の対象となる。地震情報の抽出によって得られるデータテーブルは以下のとおりである。1 行目は震源に関するデータ、2 行目は影響のあった地域に関するデータ、3 行目は地震と記事の対応に関するデータを指す。

(ID,マグニチュード,緯度,経度,群発地震の有無)
(ID,都道府県名,市区町村名,震度)
(ID,記事番号)

注:ID とは時間情報のことを指す

本稿では記事としてタグ付きコーパスを用いており、時間情報・マグニチュード・記事番号にはタグが付与されているので、これらの情報は XML パーサにより取得する。また、震源の緯度・経度に関する情報は、時間情報をもとに Web から補完する。地域・震度情報の取得に関しては 3.1 節で、群発地震

の判定については 3.2 節で述べる。

#### 3.1.地域・震度情報の取得

地震記事での地域・震度情報の表記には

(1) 震度→地域

(2) 地域→震度

の 2 種類のパターンがある。以下にその例を示す。

(1) 各地の主な震度は次の通り。▽震度 4 新潟県小千谷市，六日町，大和町，中里村

【毎日新聞 98. 2. 21】

(2) 1 回目の地震の揺れは首都圏にも広がり，横浜市中区で震度 3，千葉県館山市，東京都国分寺市，山梨県大月市などで震度 2 を観測した。

【毎日新聞 98. 4. 27】

本稿では、文書全体を形態素解析することによって、この 2 種類のパターンの判定を行う。文書より取得した市区町村名から都道府県名を一意に特定できない場合、震源地に最も近い都道府県との都道府県番号の差の絶対値が最小のものを、その市区町村の属する都道府県とする。また、市区町村が市区町村データベースに登録されていない場合、Web 上を検索し、緯度・経度の情報を取得することにより、これを強化する。

#### 3.2.群発地震の判定

上記の処理は、時間的な動向情報を取得するものである。一方、本研究では空間的な動向情報として、比較的狭い地域で集中的に発生する群発地震を対象としているので、各地震に対しこの判定を行う必要がある。本稿では、記事中に「群発地震」という表記がある場合、その地震を「群発地震候補」とし、前後 30 日、震源の緯度・経度が 0.1 度以内で異なる地震が発生している場合、それらを群発地震と判定する。

### 4.地震情報の可視化

前節の処理により構築した地震情報データベースからクエリに該当するデータテーブルを入手し、適

切な可視化表現を用いてユーザに提示するインタラクティブな地震情報可視化システムのプロトタイプを構築した。システムはJSPとして実装しており、Web上からの検索が可能である。空間的動向の可視化に関しては4.1節で、時間的動向の可視化に関しては4.2節で述べる。

#### 4.1.空間的動向の可視化

システムに対し地震検索条件とグラフ種として「マップ」を入力すると、図3のような空間的動向の可視化表現が得られる。

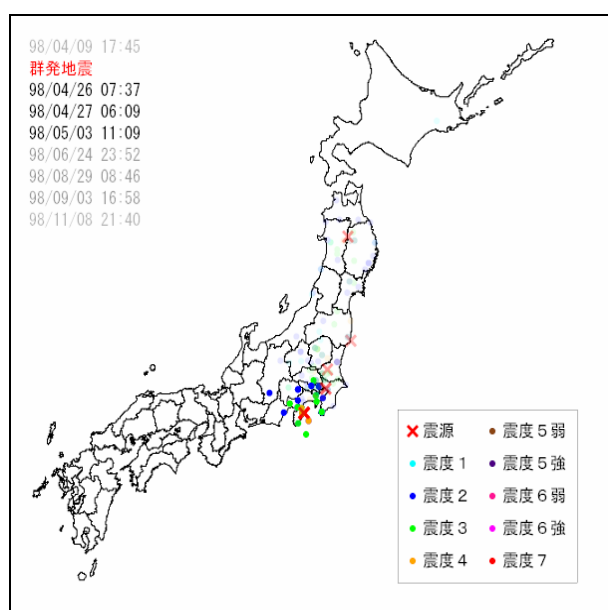


図3 空間的動向の可視化

図3では複数の該当する地震の情報として震源と被害地域を日本地図上へ描画している。ユーザが日本地図上の左上に描画されている時間、もしくは日本地図上の震源をクリックすると、システムはその地震に関して詳細な情報を持った記事を提示する(図4)。

図3において、選択された地震に群発地震が含まれていれば、該当する地震の時間表記の前に群発地震を表すキーワードを表示し、ユーザにこれを通知する。

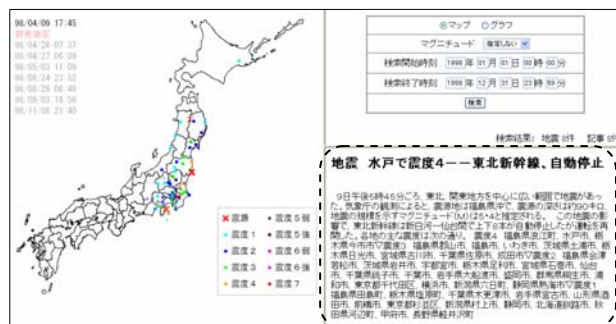


図4 地図上からの地震記事の参照例

#### 4.2.時間的動向の可視化

システムに対し地震検索条件とグラフ種として「グラフ」を入力すると、図5のような時間的動向の可視化表現が得られる。

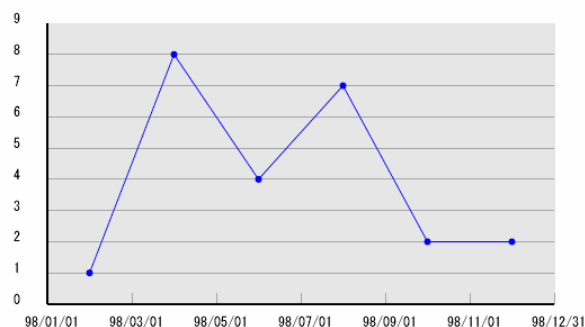


図5 時間的動向の可視化

図5では横軸に時間、縦軸に地震回数をそれぞれ対応付けることで時間的な動向情報を折れ線グラフ上に描画している。また、図3においてユーザが群発地震を通知するキーワードをクリックすると、システムは一連の地震に関する情報を、横軸に時間、縦軸にマグニチュードを対応付けた可視化表現を用いて提供する(図6)。

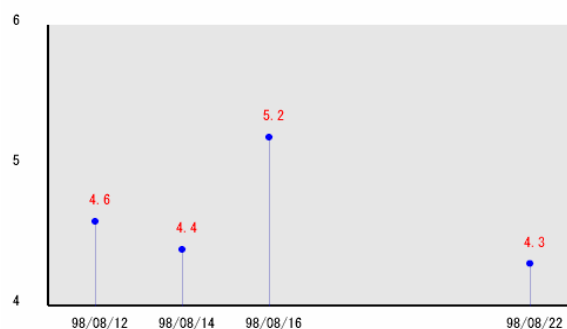


図6 群発地震の可視化

## 5.既存システムとの比較

提案システムの有効性を考察するために、提案システムと気象庁の震度データベース検索システム[3]で同一の地震を出力した。それぞれのシステムから得られた結果を図 7,8 に示す。

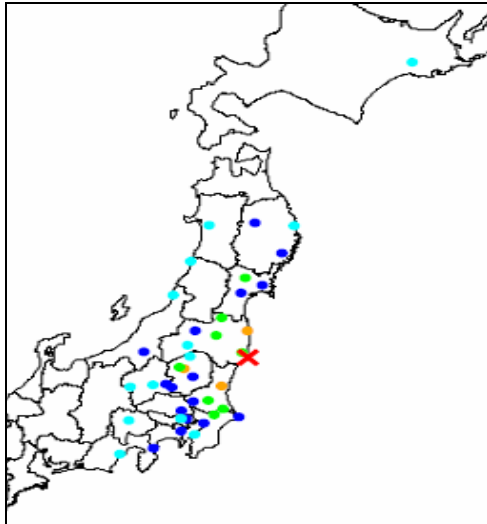


図 7 提案システム

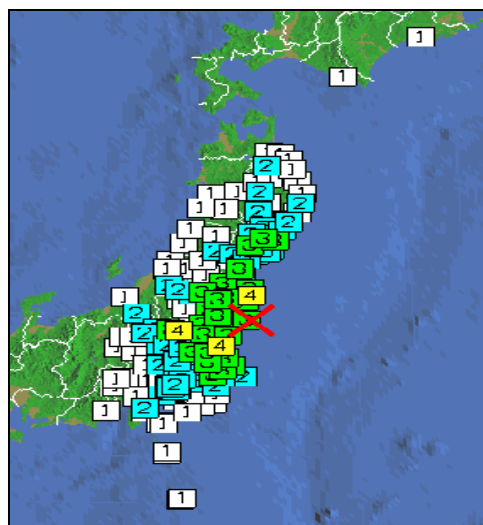


図 8 気象庁のシステム

それぞれの出力結果を見比べてみると、図 8 の方が図 7 よりも地域に関する情報が多いことがわかる。しかし、図 7 は震度と地域の情報を新聞記事から抽出しているため、地震に関する主要な情報は含んでおり、地震の影響の広がりといったマクロな動向を把握するには十分である。また、地震についての動

向を知るためには単一の地震だけでは不足しており、提案システムのように複数の地震について同じ空間へ描画することが求められる。さらに、地震に関する時間的動向を折れ線グラフによって表現することで、ユーザはこの変化を直感的に理解することが可能となる。提案システムでは、得られた可視化表現をもとに、システムへのユーザインタラクションを通じて群発地震の動向などさらなる可視化表現を得ることができ、よりユーザの興味に応えることが可能と考える。

## 6.おわりに

本稿では、ユーザからのクエリに基づき MuST コーパスから地震に関する複数の文書を収集し、その中から適切な動向情報を抽出し、複数の可視化表現を用いてユーザに提示する地震情報可視化システムについて提案した。提案システムにより、ユーザは地震に関する時間的動向・空間的動向をそれぞれに適した可視化表現を用いて知ることが可能である。今後は、地震情報抽出時に影響のあった市区町村の情報を Web から補完することで、システムに対する情報量の増加を目指す。また、現在のシステムでは検索時に指定できる空間情報は全国と都道府県だけであるが、地域情報を階層化して定義することにより地方毎の動向情報生成なども検討する。

## 参考文献

- [1] 加藤恒昭, 松下光範, 平尾努: 動向情報の要約と可視化に関するワークショップの提案, 情報処理学会研究報告, Vol.2004, No.108, pp.89-94, 2004
- [2] 難波英嗣, 国政美伸, 福島志穂, 相沢輝昭, 奥村学: 文書横断文間関係を考慮した動向情報の抽出と可視化, 情報処理学会研究報告, Vol.2005, No.73, pp.67-74, 2005
- [3] 気象庁 震度データベース検索  
<http://www.jma.go.jp/jp/quake/>