

東日本大震災時における個人属性に応じた情報ニーズの時系列評価に関する研究

京都大学大学院工学研究科 学生員 ○藤森崇浩
京都大学大学院工学研究科 正会員 小山真紀

京都大学大学院工学研究科 正会員 清野純史

1.研究の背景と目的

自然災害が起こった際に、被災者が必要とする情報は、それぞれの状況に応じて様々である。例えば、災害直後は災害の規模や被害状況、ライフラインの復旧度合などの情報ニーズがあるが、ある人にとってはそれらの情報のうち自分の居住地域を対象とした情報のみが必要となり、また別の人では家族がその時滞在している場所の情報も必要になるなどである。さらに被災から時間が経過するにつれて情報ニーズにも幅が出てくる。慢性疾患を抱えている人や高齢者にとっては病院や薬のある場所の情報、仕事を失った人にとっては失業対策の情報、子どものいる家庭にとっては学校の再開の情報がそれぞれに必要なと考えられる¹⁾。

しかしながら、現状の災害情報提供は被災者一般を対象としたものが多く、自分にとって必要な情報入手するにはそれらの情報を網羅した上で自ら取捨選択する必要がある。このような状況を踏まえると、個人属性や被災からのタイムフェーズを考慮し、被災者個人が必要とする情報に容易に到達できるようにするための支援システムが求められると考えられる。

本研究では、そのための基礎的研究として、震災時にどのような属性の人物がどのような情報ニーズを持っているかを明らかにすることを目的として、2011年東日本大震災後から2011年末までの朝日新聞記事のテキストマイニングを通じて情報ニーズの時系列評価を行った。

2.データと手法

2011年3月12日から12月31日までの朝日新聞の全文記事の中から、「地震」「震災」「津波」「被災」「避難」「原発」「原子力発電所」「計画停電」「節電」「帰宅困難」「帰宅難民」の語を一つでも含む記事を分析対象として抽出した。その上で、被災地の復興フェーズは初期ほど短期間で変化していくこと²⁾を考慮し、表1に示すようなフェーズにより新聞記事を区分し、各フェーズにおける記事内容の分析を行うこととした。

次に個人属性や情報ニーズ項目に関する「コーディングルール」を作成し、各記事に関連するコードを付与した。例えば「高齢」というコードは「高齢」「高齢者」「お年寄り」「老人」などの語を含む記事に付与される。個人属性に関するコードと情報や物資に関するコードの両方を用意し、それらの出現率の変化や相関を分析した。

新聞記事の解析はテキストマイニングツール KH Coder(2.b.29g)を用いた。形態素解析には MeCab を用いた。形態素解析のための辞書としては IPA 辞書と 2012 年 10 月 21 日時点での日本語版 Wikipedia の項目リストから作成したユーザー辞書を用いた。

表1 新聞記事のタイムフェーズ

フェーズ	期間	経過時間
フェーズ 1	3/12	～1 日後
フェーズ 2	3/13～3/14	～3 日後
フェーズ 3	3/15～3/18	～1 週間後
フェーズ 4	3/19～3/31	～3 月末
フェーズ 5	4/1～4/11	～1 ヶ月後
フェーズ 6	4/12～5/11	～2 ヶ月後
フェーズ 7	5/12～8/11	～5 ヶ月後
フェーズ 8	8/12～12/31	～2011 年末

3.コード出現率の時系列変化

情報や物資に関する各コードが付与された記事が、それぞれどのフェーズにおいて多かったかを表2に示す。矢印で示すフェーズが、コード出現率の高かった期間である。初期においては「地震被害」「帰宅・交通」「安否」「ライフライン」などのコードの出現率が高いという傾向がある。その後、「風呂」のような新たな問題が出現し始め、「学校の再開」「仮設住宅」といった日常生活への復旧に関する話題が多く取り上げられるようになっていく。このようにして、まずは被災地全体での情報ニーズの時系列変化を読み取ることができた。

キーワード 東日本大震災, 個人属性, 情報ニーズ, テキストマイニング

連絡先 〒615-8540 京都市西京区京都大学桂 C クラスター1 号棟 146 号室 TEL075-383-3253

表2 情報コードの出現率の時系列評価

フェーズ ニーズ	1	2	3	4	5	6	7	8
帰宅・交通	←→							
地震被害	←→	←→						
安否	←→	←→						
ライフライン	←→	←→						
食料	←→	←→	←→					
風呂			←→	←→				
学校の再開				←→	←→			
仮設住宅				←→	←→	←→		
義援金					←→	←→		
失業・就職					←→	←→		
復興					←→	←→	←→	←→

4. 個人属性ごとの情報ニーズの時系列変化

続いて、個人属性ごとの情報ニーズを評価するため、各フェーズにおいて個人属性コードとの共起語(コードの付された記事の中で用いられている語)を分析した。共起の指標としては Jaccard 係数を用いた。例えばコード X と単語 Y の間の Jaccard 係数を求める際、どちらか一方が出現した回数を $|X \cap Y|$ 、両方が出現した回数を $|X \cap Y|$ とすると、Jaccard 係数は $|X \cap Y| / |X \cup Y|$ で与えられる。よって、この値が大きいほど単語 Y はコード X の記事の中で良く用いられているということであり、コード X に固有な単語となる。一例として「会社員」という個人属性コードにおける分析結果を図1に示す。

会社員	フェーズ							
順位	1	2	3	4	5	6	7	8
1	帰る	帰る	客	妻	秘書	経歴	男性	男性
2	ツイッター	過ごす	男性	長女	建設会社	本人	主婦	女性
3	構内	一夜	来る	知人	理事	新顔	女性	会社
4	橋内	駅	列	男性	主婦	新顔	長男	主婦
5	足止め	バス	通勤	宅	男性	公	回答	現
6	ホテル	疲れる	待つ	自宅	衆院議員	回数	会社	会社役員
7	出張	電車	車	実家	投票	元	投票	元
8	ロビー	JR	朝	住む	現	年齢	有権者	新
9	帰れる	妻	開店	身	新	回答	仕事	建設会社
10	試験	自宅	開店	家	農業	秘書	数字	公
11	公衆電話	足止め	用意	家族	会社	投票	妻	経歴
12	タクシー	シート	ス	車	政党	現	長女	市議
13	電車	女性	並ぶ	娘	元	カッコ	息子	無職
14	並ぶ	帰れる	体育館	長男	鈴木	理事	新顔	長男
15	後期	歩く	帰る	両親	支持	候補	自宅	自宅
16	小中学校	帰宅	妻	子ども	選挙区	会社	経歴	中村
17	配る	不安	バー	逃げる	町長	立候補	職場	長女
18	運休	軽傷	困る	南相馬市	理事長	現職	声	数字
19	一夜	毛布	外す	心配	議長	原則	開票	定数
20	乗り場	途中	家	不安	不動産会社	無所属	現	帰宅
記事数	5	20	36	107	76	182	272	278

図1 個人属性コードとの共起語の例

フェーズ1, 2では「帰る」「足止め」「タクシー」「電車」「運休」などといった語が並び、帰宅時の交通手段に苦労している様子が見て取れる。フェーズ3では「通勤」「朝」「待つ」などといった語に変わり、帰宅から通勤へと記事内容が変わっていることが分かる。それ以降の共起語は選挙関連などの震災とは関連の無い語に移行していった。これより、「会社員」はフェーズ1から3までの間に帰宅・通勤時の交通手段に関する情報ニーズがあったということを読み取ることができる。同様の分析を他の個人属性においても行うことにより、各個人属性に応じた情報ニーズの時系列評価を行った。

5. 結論と今後の展望

本研究では、新聞記事にコードを付与することで被災地全体での大まかな情報ニーズの時系列変化を把握することや、個人属性コードの共起語の分析を行うことで個人属性に応じた情報ニーズの時系列評価をすることができた。ただし、本研究で用いた新聞記事データは、津波や原発事故を含むという意味で特殊な東日本大震災に関するものであったため、将来の災害時にも応用できるよう過去の新聞記事データに対しても同様の分析を行い、災害規模にも応じた形での情報ニーズのモデル化を検討することとする。また、新聞紙上からは障害者や妊婦といった個別の事情は抽出しにくいということも明らかになったため、情報ニーズのモデル化に関してはアンケート調査や被災地での聞き取り調査の結果も併用していくこととする。今後は、本研究によって得られた情報ニーズのモデルを利用し、あらかじめ自分の個人属性を入力しておく、被災時にその個人属性と被災からのタイムフェーズに応じて必要になってくると考えられる情報あるいはその情報へのアクセス方法を提供するスマートフォン向けのアプリの開発を目指している。スマートフォンというプラットフォームを利用することで位置情報など利用者その人自身の情報を加味することができるようになる事が期待される。

参考文献

- 1) サーベイリサーチセンター：「宮城県沿岸部被災地アンケート」, 2011.
- 2) 小山真紀, 翠川三郎：市町村における地震時の意思決定支援に向けた災害応急対応モデル化の試み, 自然災害科学 25(1), pp51-70, 2006.