

Q&A サイトを用いた地域に対する関心の推移の可視化

藤田 秀之, 柴崎 真理子, 木實 新一, 有川 正俊

Visualization of changing interest in geographic areas based on a Q&A site

Hideyuki FUJITA, Mariko SHIBASAKI, Shin'ichi KONOMI and

Masatoshi ARIKAWA

Abstract: ネットワーク上には地理空間に関連する情報が大量に蓄積されているが、一般的には、それらの情報がどのようなユーザの要求・関心に基づいて生成されたかは明らかでない。しかしながら、Q&A サイトに投稿される“地域”に関する質問および回答は対象地域に対するユーザの要求・関心をより直接的に反映している。そこで本稿では、代表的な Q&A サイトである Yahoo!知恵袋を対象に、東日本大震災被災地域に関する質問・回答を例に、その内容を地図上に可視化し、ユーザの要求・関心の推移を明らかにする枠組みについて検討する。

Keywords: ソーシャルメディア(social media), 可視化(visualization), Q&A サイト(Question and Answer site), ヒートマップ(heat map)

1. 序論

近年、ネットワーク上には地理空間に関連する情報が大量に蓄積されている。しかしながら、これらの地理空間情報は、どのようなユーザの要求・関心に基づいて生成されたか明らかではなく、他の情報とともに更新されていく膨大な情報の中に埋もれ、その保有する情報量の価値が失われてしまう。

蓄積される地理空間関連情報の利活用が進み、有用な知識を共有・集約したり、または行動を支援するサービス(Ushahidi[1]; Andra Kay, 2010)や、地域に関する情報検索(以下、ローカルサーチとする)の効率性向上の試みが行われている。たとえば、Matt Jones. (2007)は、周辺にいるユーザが使用した検索語を参照することで検索の利便性が上が

ると考え、周辺ユーザの検索語を表示する検索インタフェースを考案した。

しかし地域に関する要求や関心は、社会情勢の変化や大きな事象の発生に対して広範囲のユーザにより生成され、急激に変動する場合もある。そのような要求や関心は一時的な空間的分布(Matt Jones, 2007)ではなく、その推移に着目すべきである。このような事象に対する推移に着目した関連研究には Twitter を用いた Sakaki 他(2010)がある。

本研究では、Q&A サイトに投稿される“地域”に関する質問および回答は対象地域に対するユーザの要求・関心をより直接的に反映していると考え、代表的な Q&A サイトである Yahoo!知恵袋[2]を対象に、東日本大震災被災地域に関する質問を収集し、その内容をユーザの要求・関心の推移を考慮して地図上に可視化する。

柴崎真理子 〒277-8568 千葉県柏の葉 5-1-5

東京大学 空間情報科学研究センター

Phone: 04-7136-4291

E-mail: ms@csis.u-tokyo.ac.jp

Yahoo!知恵袋とは、2008年時点で1261万人[3]の利用者がおり、質問回答形式による知識・情報共有サービスである。過去の質問と回答はネット上で蓄積・共有される。回答の質をあげるために、一番適切な回答をした回答者にポイントを与えることで、ユーザに適切な回答をするインセンティブを与えている。Yahoo!知恵袋に投稿される質問に、質問者の現在位置情報は付与されていない。

本研究では、以下の可視化ツールを提案する。Yahoo!知恵袋から、「地域」を対象とした「ある特定のキーワードを含む質問」のトレンドを抽出し、地図上に可視化することで、地理空間に関連するユーザの要求・関心およびその推移を表す。具体的には、2つの可視化ツールを提案する。1つ目はユーザが指定した検索語を含む質問の分布を、県レベルの地域別に可視化するツールである。2つ目は市町村レベルの地域別に、ヒートマップを用いて可視化するツールである。

タイムリーな推移の可視化により、地理空間に関連するユーザの要求・関心を、ただ時系列に更新され表示されていく膨大な情報の中から、互いに発見し共有しあい、ユーザ同士の有用なコミュニケーションにつなげることで、ローカルサーチの検索効率の向上を目指す。

2. 地域に関連するデータ取得

Yahoo!知恵袋サイトでは、位置情報や詳細地名による検索APIは提供されていない。しかし、条件指定による検索フォーム[4]とキーワードによる質問検索API[5]が提供されており、本収集システムではこれらを利用した。

2.1 県レベルのデータ取得

都道府県別の質問を取得するには、条件を指定する検索を用いた。この条件指定による検索は、キーワード、カテゴリ、質問・回答の投稿日時と

解決日時、エリア(都道府県別)などの条件を指定できる。

質問取得に用いた指定条件はキーワードと質問・回答の投稿日時・解決日時とエリアの3条件を用いた。収集方法は3段階からなる。(1)条件指定の検索フォームのリクエストURLのgetパラメータに、キーワード、期間、県のパラメータを設定し、html形式のデータを取得する。図1の赤文字はキーワード、青文字は期間、オレンジ色は県の条件指定を表している。

```
http://chiebukuro.search.yahoo.co.jp/search?p=
放 射 能
&vaop=a&search=all&id=&flg=3&dnum=0&noct=1
&dflg=4&dfrom_y=2011&dfrom_m=07&dfrom_d=01
&dto_y=2011&dto_m=07&dto_d=31&type=&geo=13
```

図1 リクエストURLの構造

(2)取得したhtml形式のデータを解析し、検索キーワードヒット数や質問IDリストを取得する。

(3)(1)の指定条件の期間を指定し、開始日をずらしてリクエストを繰り返すことで、各期間の集計データを取得する。

以上の方法により、地域に関する質問文のキーワードヒット件数と質問IDリストを取得した。

2.2 市町村レベルのデータ取得

市町村ごとの質問文の取得には、キーワードによる質問検索APIを用いた。質問検索APIは、質問本文に指定したキーワード(以下検索語とする)を含む質問の質問ID(各質問文固有のID)、質問本文、回答(ベストアンサー)、カテゴリ、投稿・回答日時などを取得する。収集システムでは、質問IDと質問本文、投稿・回答日時、投稿の種類(pcかmobile)を取得した。

これら上記の2つの質問文取得方法には長所と短所が存在する。

条件指定による検索フォームは、都道府県の指

定のみで市町村レベルの指定はできないが、2004年からさかのぼって質問を取得することができ、かつ条件指定に期間を設定できる。また、取得できる質問数に制限はない。

質問検索 API はクエリに市町村名を入れることで各市町村に関連する質問を取得できると考えられる。しかし、質問取得の条件に期間設定ができず、API を動かした時点から最大 1000 件のみの質問しか取得できない。そのため、1000 件を超える過去の質問は取得できないという制約がある。

3. 実装

3. 1 収集システム

3. 1. 1 県別の質問収集システムの実装

2. 1 節にて解説した Yahoo! 知恵袋の条件指定による検索フォームによる収集方法を利用する。

本収集システムでは、代表的な被災地域を含む岩手県、宮城県、福島県と関東近郊の茨城県、栃木県、千葉県、群馬県、東京都と大阪府、福岡県の 10 都道府県ごとの「原発」「原子力」「セシウム」をキーワードとした 1 ヶ月間ごとのヒット件数を 2010 年 3 月から 2011 年 7 月まで取得した。

条件指定による検索フォームのリクエスト URL を用いて取得した html ページから検索ヒット件数が何行目にあるか解析し、該当行を切り出すことで検索ヒット件数を取得する。ただし、検索ヒット件数が 1 件のみの場合、html ページの構造が変化し、検索ヒット件数の位置がずれるため、従来の方法では検索ヒット件数が 1 件の場合と 0 件の場合の区別ができない。そのため、検索ヒット件数が 1 件かどうか確認する check プログラムを動かし、検索ヒット件数が 1 件となる期間を調べる。

3. 1. 2 市町村別の質問収集システムの実装

2. 2 節で解説した質問検索 API による市町村レベルの質問収集システムを構築する。今回は一例

として、岩手県に関する質問文を収集し可視化を行う。質問本文に“地名語”を含む質問は、その“地域”に関する質問だと考えられる。したがって、質問検索 API の質問取得条件である検索語に“地名語”を設定し、質問文に“地名語”を含む質問を取得する。

以下のように実装した。

(1) 地名語リストの作成

質問検索 API の検索語に用いる地名語は、国立情報学研究所の北本朝展准教授と東京大学空間情報科学研究センターとの共同研究として開発を行っている「地名語辞書管理システム」で使用されている地名語辞書(国交省作成「街区レベル位置参照情報」[6])を用いた。この地名語辞書から対象とする県内の全市区町村名を抽出し、地名から「市」「区」「町」「村」を省いた地名語リストを作成する。全 76 語からなるリストが作成された。

(2) 質問検索 API による質問取得

質問検索 API を用いて、(1)で作成した地名語リストを検索語として質問を取得する。収集された岩手県に関する総質問数は 19932 件となる。取得された質問集合には、以下のように不要な質問も取得される。たとえば、地名「大迫」で取得された質問集合には、「大迫選手」という人名を含む質問もある。また、同じ地名で異なる地域を指している質問も取得されてしまう。たとえば、岩手県にある山形町は「山形」という検索語で取得されるため、「山形(県)」の質問も含まれてしまう。さらに、取得結果の重複も起こる。「二戸市軽米町」を含む質問文は、「二戸」と「軽米」の検索語で重複して取得している。

このような重複は質問 ID によって抽出され、削除できる。重複を削除した結果、岩手県の市町村名を含む質問数は 17301 件となった。

(3) ジオパースとジオコーディング

(2)で取得し重複を削除した質問本文に「地名語辞書管理システム」のジオパース(文章中から地名語抽出)とジオコーディング(地名語への緯度経度を付与する)機能を用いる。

このシステムのジオパース機能は、以下の特徴を備えている。文章の前後関係から地名語らしさを評価し、同時に前後関係から上位の地名語を推定できる。また、部分的な地名語を補間し、住所文字列を生成する。

そのため、(2)で示したような地名ではない検索語を含む質問にはジオコーディングができないために、ヒートマップ作成結果に影響を与えない。次に、同じ地名で異なる地域を指している質問も削除できる。ジオパース機能によって文章の前後から岩手県の「山形」町なのか「山形」県なのか判別され、ジオコーディングされるため、各市町村の検索ヒット件数の集計には影響を与えない。

Yahoo!知恵袋の質問検索APIと「地名語辞書管理システム」のジオパースAPIとジオコーディングAPIを用いて、市町村レベルの質問収集システムを作成した。

```
{
  Id: "1370031028",
  Content: "東京駅周辺にいる方で今から盛岡方面に行く方いませんか?!",
  CategoryPath: "地域|旅行、お出かけ|交通、地図|鉄道、列車、駅",
  PostedDate: "2011-08-29T00:41:58+09:00",
  SolvedDate: "",
  PostedDevice: "mobile",
  Locations: [
    - {
      name: "東京駅",
      lat: "35.681391",
      lng: "139.766103",
      count: 1
    },
    - {
      name: "盛岡駅",
      lat: "39.701547",
      lng: "141.136599",
      count: 1
    }
  ]
}
```

図2 市町村に関する質問収集システムの取得結果

この質問収集システムは、検索語に「市」「町」「村」「区」を省いた地名を使うことで、地名らし

き語を含む質問収集の網羅性を高めることを優先し、不要な質問を「地名辞書管理システムの」ジオパーサ機能を用いて、選択的に削除している。

3. 2 可視化ツールの実装

二つの可視化ツールを実装した。1つは、3. 1. 1節の県別の質問収集システムで集められたキーワードのヒット件数を時系列的な折れ線グラフで表現する。2つ目は、3. 2. 1節の市町村に関する質問収集システムで収集された地域に関する質問集合から、ユーザが入力したキーワードを含む質問の時系列的な件数推移を市町村ごとに地図上にヒートマップとして可視化する。

4. 実装結果

4. 1 県レベルの可視化

3. 1. 1節の県別の質問収集システムで収集した検索キーワードヒット件数を時系列な折れ線グラフで表現する。図3は「原発」と「原子力」の各キーワードのヒット件数合計の推移を表した。ここから、原発や原子力に関連する質問が福島に集中していることがわかり、5月以降から関心が離れていることが分かる。図4は「セシウム」のヒット件数である。放射性物質セシウムが社会的問題となった時期から福島県と東京都に関連する「セシウム」のヒット件数が増加していることが分かる。

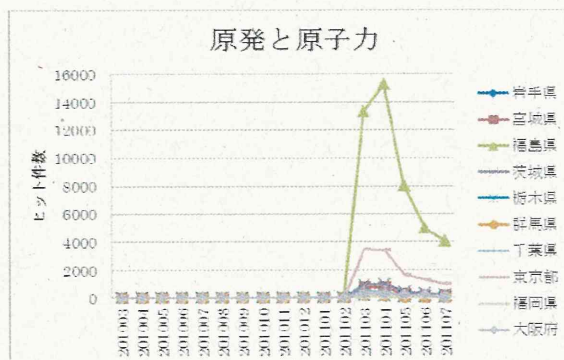


図3 「原発」と「原子力」の県別ヒット件数

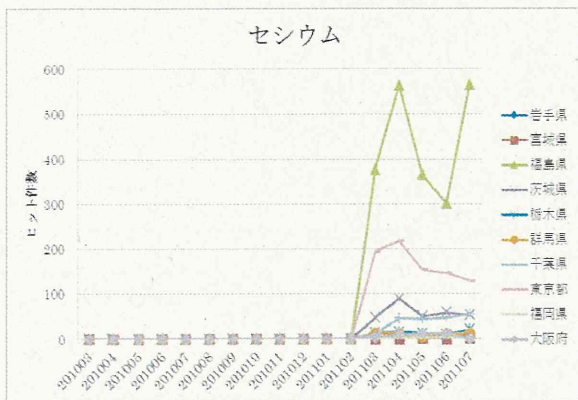


図4 「セシウム」の県別ヒット件数

4. 2 市町村レベルの可視化

3. 2節で実装した可視化ツールの市町村レベルの可視化のデモとして岩手県の質問集合から「安否」という検索キーワードを含む質問の時系列的な推移を、以下のヒートマップに可視化した。

「安否」という語を含む質問数の推移は、2011年3月11日に発生した東日本大震災時に、ユーザらの知人の安否情報に対する要求を表していると考えられる。

図5のヒートマップでは、3月には岩手県全域に「安否」を含む質問が分布しているが、時間が経つにつれ、津波による不明者の多い沿岸地域のみ分布が残っている。

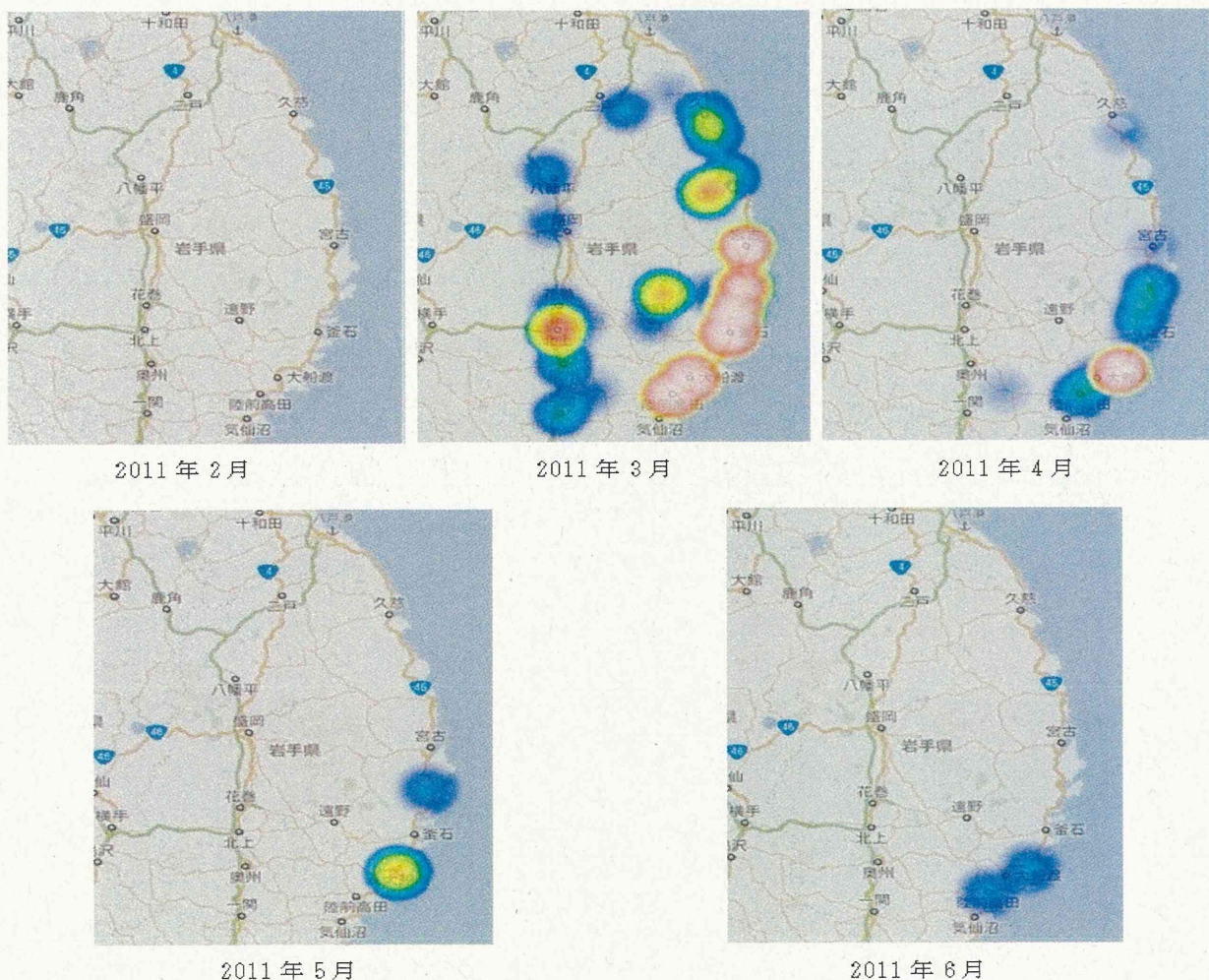


図5 「安否」に対する岩手県市区町村レベルのヒートマップ

4. 3 結果の概要と考察

本研究で扱った岩手県の例に見られるように、東日本大震災のような社会的関心の高い、大きな事象が発生したとき生成される地域に対する要求や関心は、事象の変化に伴い即座に変動する。これらの推移を時系列なヒートマップ、折れ線グラフで地図上に可視化し、他のユーザに緊急性の高い回答が求められている要求や関心を示すことで、優先的に回答を求め、より質の高い回答ができるだろうユーザを質問者と結びつけることが可能である。

このように、ヒートマップおよび時系列の可視化を統合してマクロのパタンから、要求や関心の度合いが高い情報に焦点をあてることができれば、ローカルサーチ等においてもユーザの要求や関心に合致する情報を提供することが容易になると考える。

人と地理空間情報のインタラクションの枠組みにおいて、ユーザの要求や関心の推移を考慮すれば、災害時などの大きな事象が発生した場合に、発信される膨大な情報の中から、コミュニケーションや情報伝達に使いやすいユーザフレンドリーな地理空間情報システムが構築することが可能と考える。

5. 今後の課題

本システムで収集された質問がどれほど正確に地名語辞書管理システム(謝辞参照)によってジオコーディングされているのか、もしくは質問内容に基づいた可視化がされたか定量的な評価が必要である。また、デモで用いたキーワードはいずれも東日本大震災に密接に関連した語であるため、ユーザの要求や関心の推移を比較的表すヒートマップを地図上に可視化できたと考える。「セシウム」など

とは違い、より日常的に用いられる語を用いたユーザの要求や関心の可視化の精度についても検討する。

謝辞

本研究では、国立情報学研究所の北本朝展准教授と東京大学空間情報科学研究センターとの共同研究として開発を行っているジオコーダを利用させていただきました。

参考文献

- [1] <http://www.usahidi.com/> 参照日時 2011/08/28
- [2] <http://chiebukuro.yahoo.co.jp/> 参照日時 2011/08/28
- [3] http://www.netratings.co.jp/news_release/2008/04/Newsrelease20080422.html 参照日時 2011/08/28
- [4] <http://chiebukuro.search.yahoo.co.jp/advanced?fi=common-navi> 参照日時 2011/08/28
- [5] <http://developer.yahoo.co.jp/webapi/chiebukuro/chiebukuro/v1/questionsearch.html> 参照日時 2011/08/28
- [6] <http://nlftp.mlit.go.jp/isj/> 参照日時 2011/08/28

Andra, Key, 2010. Ushahide and Crowdmap: micro-Streaming as time-binding media. 3PM Journal of Digital Research & Publishing, 116-125

Matt, Jones, George, Buchanan, Richard, Harper, Pierre-Louis, Xech, 2007. Question Not Answers: A Novel Mobile Search Technique. CHI 2007 Proceeding / Mobile Interaction.

Takeshi, Sakaki, Makoto, Okazaki, Yutaka, Matsuo, 2010. Earthquake Shake Twitter Users: Real-time Event Detection by Social Sensors. WWW2010

PhotoField : 地図ベースのスライドショー制作ツールと

ユーザスタディ

藤田秀之、有川正俊

2011

東京大学空間情報科学研究センター

CSIS DAYS 2011 「全国共同利用研究発表大会」

PhotoField: 地図ベースのスライドショー制作ツールとユーザスタディ

藤田 秀之, 有川 正俊

東京大学 空間情報科学研究センター

連絡先: <fujita@csis.u-tokyo.ac.jp> Web: <http://www.s-it.org/photofield/>

(1) **背景と目的:** Google Maps のマイマップ機能を代表として、ユーザが背景地図上に複数の対象をマッピングし、手軽に公開できるサービスが一般化した。こうしたサービスは、場所のリスト(例えば飲食店リスト、観光スポットリスト等)を整備・検索する辞書的なコンテンツに適しており、閲覧者は地図上の対象を検索・選択して閲覧する。一方で、地図をベースとして、映像のように「再生」して視聴する、ストーリー形式のコンテンツがいくつか登場している。コンテンツの再生時には、作り手の定めたストーリーに従い、地図上の複数の対象が順にフォーカスされる。Google Earth のナレーション付きツアーはその代表であり、ユーザにより多くの印象的なコンテンツが制作され始めている。ストーリー形式の場合、辞書的な形式と比較し、作り手の個性や創造性がより発揮されたコンテンツの生産が期待できる。

本研究の最終目標は、こうした地図ベースのストーリー形式のコンテンツ制作支援である。例えば、分かりやすいコンテンツ、あるいは印象的なコンテンツに共通する特徴や、そのようなコンテンツの制作に有用なソフトウェアの機能を探ることを目標としている。本稿では、地図ベースのスライドショー制作ツールのユーザスタディを行い、上記に関する知見を得ることを目的とする。

(2) **対象ソフトウェア:** 先行研究で提案・公開した、マッピングした複数の写真によりスライドショーを作成するソフトウェア(PhotoField)を利用した。スライドショーの並びやナレーションを作り手が定めるため、地図をベースとしつつストーリー形式のコンテンツを作成できる。起承転結のようなストーリー性を与えることも可能である。コンテンツの閲覧は、再生、一時停止、1フレーム進める・戻る等の操作で行う。写真の切り替え時には、写真と写真の間を移動するような3次元アニメーションが自動生成される。

(3) **ユーザスタディと結果:** 上記のソフトウェアによりコンテンツを制作してもらった実験を行い、制作された80のコンテンツ、ソフトウェアの操作ログ、アンケート結果を分析した。特徴的な結果は次のとおり。

- ・過去に撮影した写真、新たに撮影した写真、ウェブで探した写真等が組み合わせて利用される。
- ・スライドショーの写真の並びを検討する際、地図が利用されている。

特に、スライドショーの写真の並びに関して、

- ・撮影時刻順の並びが多い一方、それ以外の並びも数割を占める。
- ・空間的連続性に従う並び(空間順)が頻出する。空間順には複数のパターンがある。
- ・非撮影時刻順の多くの部分は空間順である。

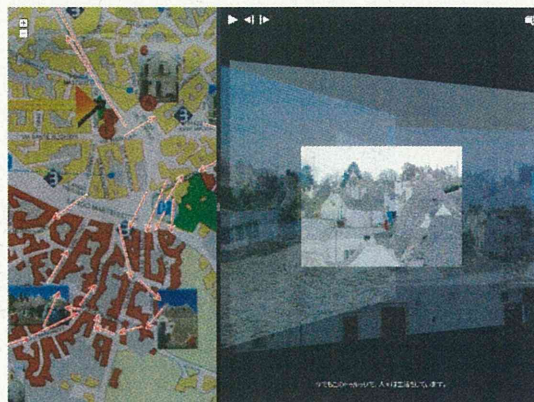


図1: 地図ベースのスライドショー制作ツール
(ソフトウェア名: PhotoField)

地図上の矢印記号が各写真に対応する。矢印はどこからどこを撮影した写真かを表す。

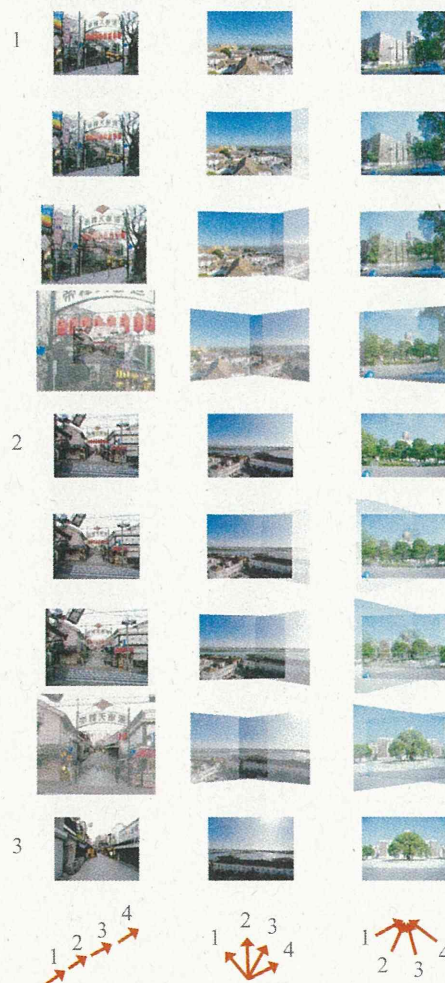


図2: 典型的な空間順の並びと生成されるアニメーション

タブレット PC を用いたデータ収集端末の開発

笹尾知世、木實新一、有川正俊、藤田秀之

2011

東京大学空間情報科学研究センター

CSIS DAYS 2011 「全国共同利用研究発表大会」

タブレット PC を用いたデータ収集端末の開発

笹尾知世¹, 木實新一², 有川正俊², 藤田秀之²

¹ 東京大学大学院 新領域創成科学研究科, ² 東京大学 空間情報科学研究センター

連絡先: <sasaotomoyo@csis.u-tokyo.ac.jp>

- (1) **動機:** 近年, 携帯型デバイスの高性能化と普及に伴い, GPS トラッキングやセンサネットワークなど空間データを自動的に収集する技術の一般利用が注目されている。しかし, これらの技術を用いて収集できるのは位置や時刻あるいは数値データに基づく単純な情報が主体である。場所や人に関する定性的なデータ収集は効率化が難しく, 現在でもアンケートやインタビューといった古典的な手法が多く用いられている。デジタル技術を用いて, 定性的なデータを収集するシステムとしては, PDA や携帯電話を用いた経験サンプリングシステムが知られているが, アンケートやインタビューにおけるインタラクティブな情報取得過程を支援するには至っていない。
- (2) **アプローチ:** 古典的なデータ収集法の中でもアンケートの枠組みに着目し, タブレット PC を用いたアンケートシステムを開発し, 形成的評価(formative evaluation)を行うことによって, インタラクティブなアンケートシステムにより得られるデータの質とアンケートを回答する際の負担等について調査する。
- (3) **意義:** タブレット PC を用いたアンケートシステムを開発し, 実験を行った結果, 情報の質と心理的経験の質に影響を与える要素を明らかにし, またデバイスに適したデータ収集インタフェースのデザインのための概念的な枠組みを示すことができた。また, タブレット PC を用いた定性的な空間データの収集の可能性を示すことができた。
- (4) **特徴:**
- タブレット PC を用いたアンケートシステムのプロトタイプを, Flex/ActionScript と PHP を用いて開発した。なお, 使用したタブレット PC は Acer ICONIA TAB W500 (Windows7)であり, 10.1 インチのマルチタッチ対応ディスプレイ, 照度センサ, 加速度センサ, Bluetooth, IEEE802.11 を搭載している。
 - 位置入力, 自由入力(思ったこと・絵), 選択入力(単数・複数), 並び替えを行う 6 種類のインタフェースを実装した(図 1)。ユーザは主に指を使って操作を行い, 入力したデータは PHP/Apache により実装したアンケートサーバに自動的にアップロードされる。
 - アンケート回答中の発話とヒアリングから得られたデータを分析した結果, タブレット PC を用いたアンケートで取得したデータの質や, アンケート回答者の心理的経験の質に影響を及ぼす可能性がある 5 つの要素(コンテンツ固有の操作性, GUI(Graphical User Interface)のユーザビリティ, メディア固有の操作性, 記憶・思考の補完性, 社会的インタラクション)が浮上した。
- (5) **今後の展開:** データ収集システムのプロトタイプを拡張し, ユーザが入力する質問文や選択肢によって自動生成するアンケートシステムの実装・利用実験を行う予定である。

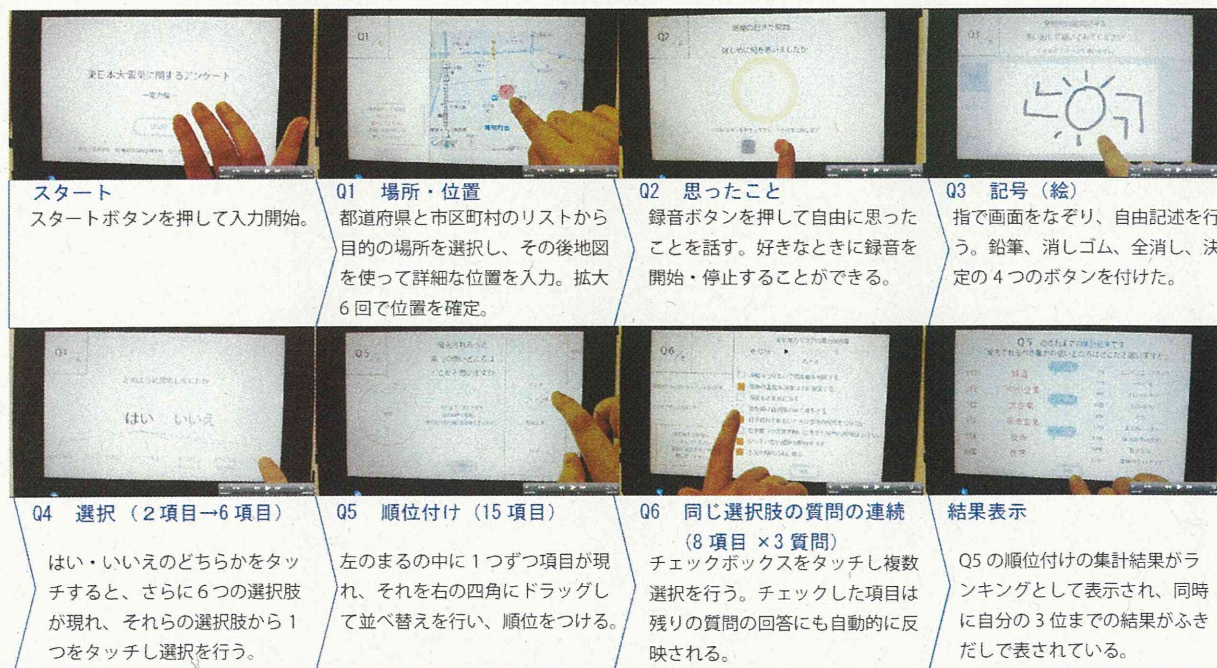


図 1: 開発したプロトタイプ

利用者自身がコンテンツを作成するパーソナル LBS の設計

と開発

鍛冶秀紀、有川正俊

2011

東京大学空間情報科学研究センター

CSIS DAYS 2011 「全国共同利用研究発表大会」

利用者自身がコンテンツを作成するパーソナル LBS の設計と開発

鍛冶 秀紀, 有川 正俊

東京大学 空間情報科学研究センター

連絡先: <kaji@csis.u-tokyo.ac.jp> Web: <http://ptalk.csis.u-tokyo.ac.jp/>

- (1) **動機:** スマートフォンの登場により位置情報サービス(LBS: Location Based Service)の利用が拡大し, 様々な LBS が展開されるようになった. 利用者自身が情報を提供できる LBS も数多くあるが, 多くはプロバイダが用意した情報に対して補足的な情報を与えるものであり, 利用者自身がコンテンツ全体を作成できるサービスではない. 利用者自身がコンテンツをデザインし作成できる環境の実現が LBS コンテンツの豊かな発展につながると考えている.
- (2) **アプローチ:** スマートフォンアプリケーションは, LBS コンテンツの利用閲覧を容易にしたが, 複雑なコンテンツの編集作業には適さない. 本研究では, LBS コンテンツを作成する場合, フェーズごとにスマートフォンとウェブブラウザのアプリケーションを使い分ける枠組みを導入する. 具体的には, Apple 社の iPhone アプリケーションと Adobe 社の Flash を用いた Web アプリケーションを併用する, 独自開発した LBS プラットフォーム pTalk により, 現地での記録と, 記録した情報の整理を効率的に行うことができる環境を提供する.
- (3) **意義:** 商用の LBS では取り扱われないような個人的なメモやローカルな情報を, 利用者自身が LBS コンテンツに加工し, 情報提供できるパーソナル LBS を実現した. これにより, 専門知識が乏しい利

用者が簡単に LBS コンテンツを作成でき, ウェブの発展と同様に, 量の点から LBS コンテンツを充実させ, 質の点から堅実な展開へとつながられる.

- (4) **特徴:**
 - iPhone アプリケーションと Flash Web アプリケーションによる効率的なコンテンツの閲覧, 作成
 - スマートフォン用のアプリケーションでそのまま音声録音することで, より臨場感のあるコンテンツが作成可能
 - 音声と折れ線を組み合わせた音声ナビゲーション
 - スマートフォンでのネットワーク接続の必要性を最低限に押さえるためのデータ取得と, データキャッシュ
- (5) **結果:** 東京大学大学院新領域創成科学研究科社会文化環境学専攻および青山学院大学総合文化政策学部総合文化政策学科の授業において pTalk を利用した LBS コンテンツ作成演習を行い, コンテンツの作成, 利用に関する実証実験を行い, 提案した枠組みの有効性を示した.
- (6) **その他:**
 - pTalk は無料で公開する予定である.
 - iPhone アプリケーションは現在 AppStore への登録申請準備中.



図 1: pTalk 画面 (iPhone と Flash Web) の例

築地場外市場の音声ガイドを Apple 社 iPhone アプリケーションおよび, Adobe 社 Flash Web アプリケーションで表示している. 画面の制約の大きい iPhone ではアイコンマーカーだけの表示だが, Web インタフェースでは写真付きのマーカーや作成時の移動ログが表示される.

Q&A サイトにおけるユーザの要求・関心の時空間的な推移
の可視化

柴崎真理子、藤田秀之、木實新一、有川正俊

2011

第4回知識共有コミュニティワークショップ

情報社会学会

Q&A サイトにおけるユーザの要求・関心の時空間的な推移の可視化

Visualization of changing interest and need in geographic areas based on a Q&A site

柴崎 真理子 (しばさき まりこ・Shibasaki Mariko)¹・藤田 秀之 (ふじた ひでゆき・Fujita Hideyuki)³

木實 新一 (このみ しんいち・Konmi Shin' ichi)³・有川 正俊 (ありかわ まさとし・Arikawa Masatoshi)⁴

¹ 東京大学新領域創成科学研究科社会文化観光学専攻 修士課程 ・ ² 東京大学空間情報科学研究センター 助教

³ 東京大学空間情報科学研究センター 准教授 ・ ⁴ 東京大学空間情報科学研究センター 教授

[Abstract]

We developed a framework to analyze how information on geographical space on the interest is generated from what kinds of user request and interest in “Yahoo! Q&A site”. With this system, we can easily grasp spatio-temporal transition of user’s request and interests. By analyzing association with news, association of social movements and transition of user’s request and interests could be more clearly understood..

[キーワード]

ソーシャルメディア、Q&A サイト、Web ニュース、可視化、時系列、ヒートマップ、地域情報

1. はじめに

近年、ソーシャルメディアを用いたネットユーザの地域情報の発信が盛んになり、ネットワーク上には地域に関連する情報が大量に蓄積されている。こうした情報の中でも他のユーザ、あるいは他の人々の関心・要求の対象に関する情報は重要である。多くの人は自らの行動や意見を定める上で他の人々が今何に関心を持っているか、それがどのように変化しつつあるのかを重要視している ([1]池田謙一、2010)。たとえば、風評やうわさの伝搬を考える上でも、人々の関心の所在、その時間的、空間的な変化を把握することは重要である。特に地域情報は個別の場所や地域に特有の関心・要求を反映していることから有用性が高いと言える。

地理空間関連情報の共有・集約を用いた行動支援サービスとしては、[2]Ushahidi(2011)などが挙げられる。また、ネットユーザが発信した現在位置に関連する情報の推移に関する関連研究では Twitter を用いた[3]Sakaki 他(2011)の取り組みがある。周辺ユーザとの興味・関心の共有によってユーザの検索効率がどのように変化するかを研究した[4]Jones(2007)がある。

本研究では、Q&A サイトに投稿される“地域”に関する質問及び回答は対象地域に対するユーザの要求・関心をより直接的に反映していると考えた。そこで、代表的な Q&A サイトである Yahoo!知恵袋を対象に、東日本大震災被災地域である岩手県に関する質問と Yahoo!Japan サイトの WEB ニュースを収集し、Yahoo!知恵袋において表出したユーザの地域に対する要求・関心がどのようにユーザ間で共有されているのか分析すると同時に、ユーザの要望・興味を時系列的な推移を可視化するシステムを実装し、WEB ニュースとの推移の関連性を考察した。

2. データ収集方法

2.1 地域に関連するデータ収集

Yahoo!ディベロッパーネットワークにて提供されている質問検索 API を用いて、岩手県に関連する質問と回答を収集した。この質問検索 API では、質問本文に指定した検索語を含む質問の質問 ID と質問本文、投稿・回答日時、投稿ツールの種類とその質問に対するベストアンサーを取得した。

2011 年 1 月から 2011 年 10 月 8 日の期間における岩手県に関する質問を収集するために、質問検索 API の質問

¹ ms@csis.u-tokyo.ac.jp

取得条件である検索語に岩手県の“地名語”を設定し、質問文に岩手県の“地名語”を含む質問と回答を取得する。この地名語リストには、国立情報学研究所の北本朝展准教授と東京大学空間情報科学研究センターとの共同研究として開発を行なっている「地名語辞書管理システム」で使用されている地名語辞書（国交省作成「街区レベル位置参照情報」（国土交通省、2011））を用いた。

地名語には「市」「町」「村」を省くことで「盛岡」などといった「市」を省いた地名を含む質問も収集可能としたが、人名など地名ではないものを含む質問文も取得される。これらの“地名語”を含まない質問集合は、「地名語辞書管理システム」のジオパース（文章中から地名語抽出）とジオコーディング（地名語への緯度経度を付与する）機能によって選別される。「地名語辞書管理システム」は文章の前後関係から地名語らしさを評価し、地名語とそれ以外の語を判別する。「地名語辞書管理システム」によってジオパースすることでそれ以外の語を含む質問集合を除外できる。質問検索 API で収集された岩手県に関する総質問数 12,366 件をジオパース機能により 7,148 件に絞り込むことができた。

2. 2 WEB ニュースの収集

Yahoo!ディベロッパーネットワークにて提供されているトピックス見出しアーカイブ API を用いて、2011 年 1 月から 2011 年 10 月 1 日の期間内で「国内」と「地域」カテゴリに配信された 5,387 件の WEB ニュースの見出し、トピック名、掲載開始時間、掲載終了時間、Yahoo! Japan のトップページに掲載されていた時間（TopicPickupTime）、掲載期間内に出たアクセス数から割り出した PV 指標（PvTotal）を取得した。

3. Yahoo!知恵袋におけるユーザコミュニケーションの分析

2.1 において取得した岩手県に関連する質問 7,148 件を質問本部の内容別に以下の質問タイプに分類した。

表-1 質問本文の分類

	件数
現地情報	5,176
意見情報	749
その他	1,223

ここで、岩手県に関する質問タイプ別の時系列的な推移をグラフで表現する。図-1 を用いて全体的な岩手県の質問数が 2011 年 1 月から 10 月がどのように推移したかを考察すると、1 月と 2 月では岩手県に関する質問数は 100 件ほどだったか 3 月の東日本大震災発生以降、3 月の質問件数は 2,683 件と増え、それ以降 400 件から 600 件の値をとり、8 月には再び 1,249 件ほどに倍増し、10 月になって 146 件となる。3 月に急増した現地情報は主に、岩手県の被災者の安否情報や避難状況、交通状況に対する情報提供の要望である。よって質問件数の推移は岩手県で起こった社会的な関心を引きつけるイベントに関わっていると考えられる。

次に、これらの現地情報型、意見情報型の質問が回答（ベストアンサー）を得られるまでにどれほどの時間がかかっているのか検証する。現地情報型の質問は“現時点”における情報を要望していることから、質問者はこの質問に対する回答を自らが必要とする“期間中に”得ることを目的にしていると考えられる。図-2 は質問タイプ別の回答を得るまでにかかった月ごとの平均的な期間（日数）の推移を表している。震災に関連する情報提供の要請など、社会的関心の高い質問では回答は早く行われるため、現地情報の回答を得るまでの平均時間は短いと考えたが、図-2 の 2011 年 3 月と他の月を比較する限り、その傾向は認められなかった。

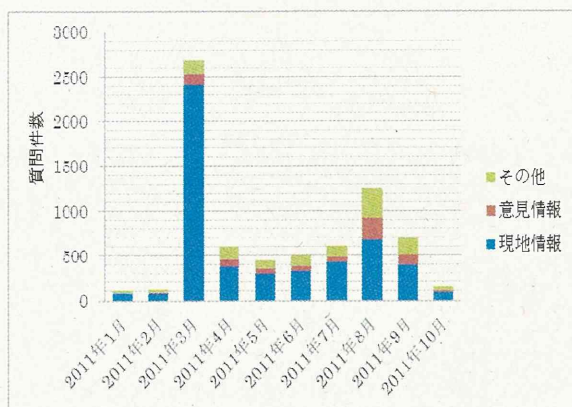


図-1 質問内容別の質問件数の推移

図-2 質問内容別の回答を得るまでにかかった平均時間の推移

4. 可視化システムの実装

2.1 にて取得した質問集合の時間的な推移をヒートマップとして可視化するシステムを実装した。本システムでは、月と検索語を入力し、検索語を質問本文に含む質問を質問本文に含まれる地名語に対応する緯度経度に基づいて可視化する。

可視化された時系列的な推移が WEB ニュースから抽出した岩手県関連のトピックの推移と対応しているか考察した。WEB ニュースの“注目度”に関する定義に基づいて地域別の WEB ニュースのトピックの抽出を行い、その推移を可視化する。WEB ニュースの注目度として、Yahoo!ディベロッパーサイトから提供されている TopicPickupTime (Yahoo! Japan のトップページに掲載されていた時間)、PvTotal (掲載期間内に出たアクセス数から割り出した PV 指標) を用い、以下のような定義式を作成した。

$$\text{TopicPickupTime} * \text{PvTotal} = \text{WEB ニュース記事の注目度}$$

この注目度は WEB ニュースの title もしくは topicname に含まれる地名語ごとに算出し、対数化する。例えば、topicname が「岩手・宮城の震災被害」ならばジオパースによって抽出された地名語「岩手」「宮城」それぞれにこの WEB ニュースの注目度が付与される。このように WEB ニュースの対象地域ごとの注目度を算出し、月間のニュース全体の注目度に対する各地域の注目度の割合 (%) と title、topicname のキーワードを出した。本研究では岩手県の時系列的な推移に着目しているため、岩手県に関連する WEB ニュースの注目度の推移のみを可視化した。

5. 可視化の結果

まず、岩手県に関連する WEB ニュースの注目度の推移は表-2 のようになった。3月から震災、復興に焦点が当たっていることが分かる。

表-2 岩手県に関連する WEB ニュースの注目度の推移

見出し・トピック名にある地名	1月	2月	3月	4月	5月	6月	7月	8月	9月
岩手県	0.0	0.0	1.7	4.3	3.7	5.4	5.8	5.3	2.7
岩手郡	0.0	0.0	0.0	0.0	0.0	0.0	7.1	3.0	11.3
大槌町	0.0	0.0	3.2	0.0	0.0	2.5	2.8	5.3	0.0
陸前高田市	0.0	0.0	0.0	0.0	0.0	2.5	3.1	0.0	3.1
釜石市	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	3.4
宮古市	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6
キーワード			被災地, 復興	震災, 孤児	支援物資	震災被害	震災被害	町長, 選挙	震災, 高台
			ライフライン	遺児	救援物資	復興	復興, 津波	小中	防災センター
			震災被害		販売鯨肉	政治家	仮設住宅		

次に、ヒートマップとして可視化するシステムの実装結果を述べる。検索語には「安否」を用いた。「安否」情報は震災報道によって喚起されたユーザーの代表的な要望の一つである。Q&A サイトから、この要望がどのように推移しているのかが可視化した。図3より、Q&A サイトにおける質問が社会的なイベントによって生成され、推移

することが可視化されている。

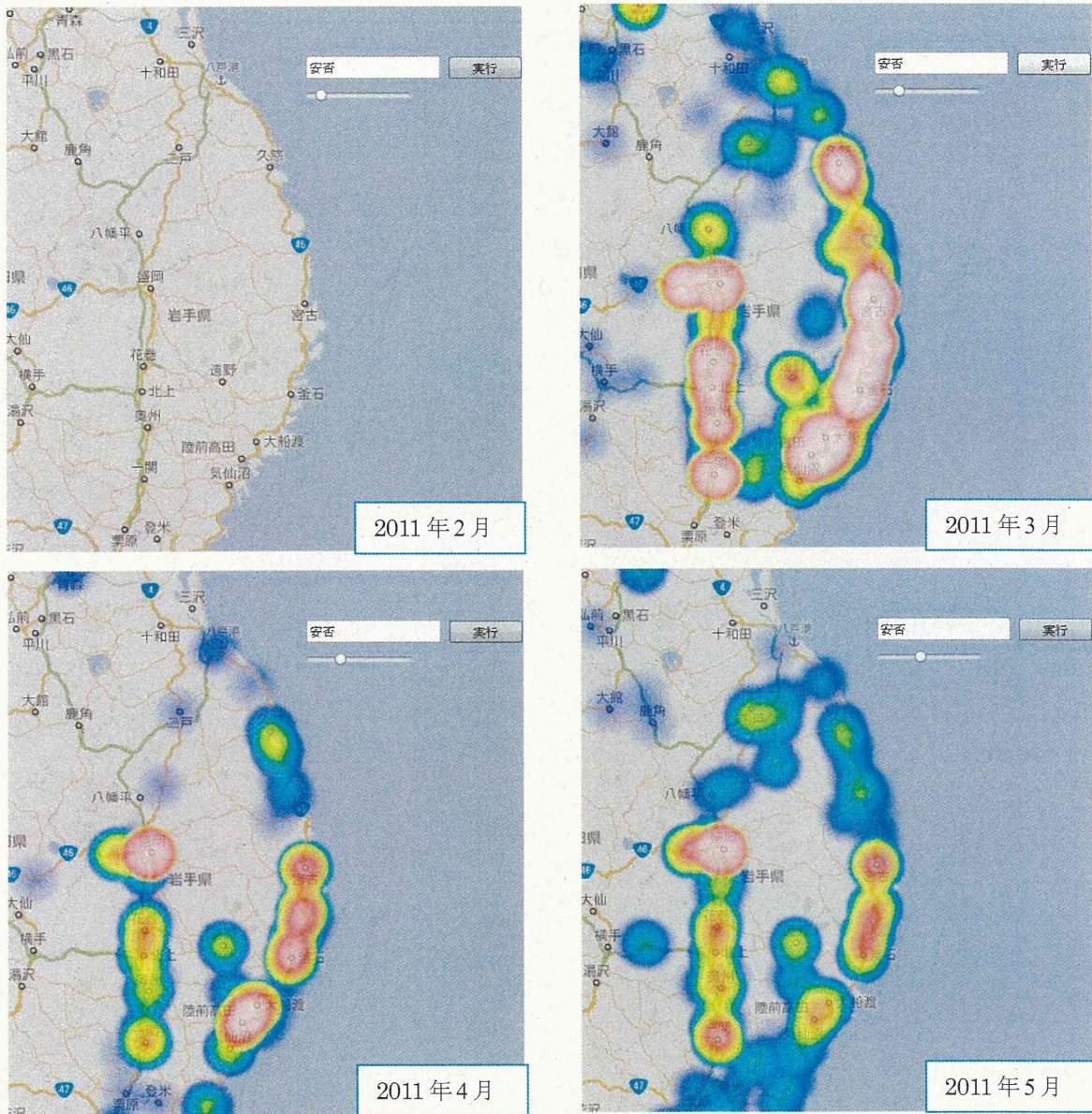


図-3 岩手県における「安否」を含む質問件数の推移

6. 今後の展望

ユーザの時空間的位置に依存した要求や関心が、Q&A サイトにおいて可視化できたことにより、今後 WEB 上における地域情報の共有サービスについて展開をしていきたい。

[参考文献]

- [1] 池田謙一 『クチコミとネットワークの社会心理—消費と普及のサービスイノベーション研究』 東京大学出版、2010 年
- [2] Ushahidi, Ushahidi Home Page. <http://www.usahidi.com/>. 参照日時:2011/10/28
- [3] Sakaki, T., Okazaki, M., Matsuo, Y., 2010. Earthquake Shake Twitter Users: Real-time Event Detection by Social Sensors. Proc. WWW2010.
- [4] O'Jones, M., Buchanan, G., Harper, R., Xech, P.-L., 2007. Question Not Answers: A Novel Mobile Search Technique. Proc. CHI 2007.

心停止発生地点に着目した AED の最適配置

片岡裕介，浅見泰司，浅利靖，郡山一明

2011

日本循環器病予防学会誌

解説

心停止発生地点に着目したAEDの最適配置

片岡裕介¹⁾ 浅見泰司²⁾ 浅利 靖³⁾
郡山一明⁴⁾

I. はじめに

突然の心停止状態への救急措置に有効とされるAED（自動体外式除細動器）の設置に際しては、より多くの装置を用意する努力もさることながら、その適切な配置地点について慎重な検討が必要とされる。AEDの効果が注目されつつある現在の社会的状況を鑑みても、その設置地点に関する議論が今後求められていくであろう。AEDの設置場所の選定や評価にあたっては、様々な方法の可能性が考えられるが、現実を単純化した条件のもとでの一つの結果であるにせよ、理論的に得られた設置地点が示されることの意義は大きいと考える。また、保健医療分野におけるGISの利活用が進展するなか、現象の地理的・空間的側面に着目した問題解決のアプローチの重要性が高まることも期待される。

AED設置に関する研究として、地域の人口密度と心停止が発生した場所との関連についての調査¹⁾や、集客施設で必要とされるAEDの設置数に関する数理モデルの提案²⁾などがあげられる。当然ながらAEDの配置地点に関する議論においては様々な観点からの検討が必要とされるが、この解説では、地域の視点からみた過去の心停止発生地点の分布状況と、生存退院率にもとづく救命の時間的要因を考慮した、AEDの最適配置を検討する方法についての研究事例³⁾を紹介する。

II. 方 法

地域全体としてAEDの適切な配置地点を求めるために、過去の心停止発生地点と生存退院率にもとづいた配置方法について述べる。ここでは、弘前地区消防事務組合が2002年1月から2004年12月までの3年間に扱った内因性心肺停止で、原因疾患が心疾患と推定された411例の心停止発生地点を用いて、解説する方法を適用した場合のAEDの最適配置地点を求める。

まず、過去に心停止が発生した地点の分布状況により、地域内のAEDの需要量の密度分布が決定されるものとする。とはいえ、心停止が発生した地点のみに需要があるとするのは、同一地点で再発する可能性を考えても適切ではない。そこで発生地点周辺も考慮し、過去に数多く発生した地域を潜在的に需要が高い地域であると見なすこととする。従って各心停止発生地点周辺のAEDの需要量は、心停止発生地点で最大であり、周囲に拡がるにつれて減衰する確率密度分布で表現することができる。ただし、ごく短期間のうちでは、心停止の発生が偶然的であることを免れないため、需要点数については需要を代表するのに十分な数である必要がある。

需要量の地域全体の確率密度分布についてはカーネル法を用いて推定する⁴⁾。カーネル法とは、点分布を形成する各点の中心に、カーネルと呼ばれる密度分布の山を置き、それら全てのカーネルの合計をもって全体の確率密度関数を求める密度推定法である。図1の下側に5個の点分布があったとする。図1では、この点分布をカーネル法によって密度分布に変換した場合の、平面上で密度

*1 大分大学福祉科学研究センター
(〒870-1192 大分市大字旦野原700番地)

*2 東京大学空間情報科学研究センター

*3 弘前大学医学部救急・災害医学講座

*4 救急救命九州研修所

受付日 2010年8月6日・受理日 2011年4月19日

の高低を表す等値線（下側）、およびそれを立体化したもの（上側）を表現している。例えば中心部分では、近距離にある2点の影響により、密度分布がその周辺と比べて高くなっていることがわかる。

心停止発生地点は、自宅などの日常生活の多くを過ごす場所や、外出先など様々であると考えられるが、発生地点の空間的傾向に着目することで、地域全体のAEDの潜在的な需要量を表す密度分布がカーネル法で推定される。具体的には、近辺で心停止が多く発生している場所は、AEDの需要が潜在的に高いと判断されることになる。なお、カーネル関数のパラメータとなるバンド幅選択に際しては、方程式解によるプラグイン法⁹⁾を採用した。プラグイン法は、確率密度関数の推定値がデータから推定されるという点に特徴があり、一般に最良なバンド幅選択法が発見されてはいない中で、多方面で有効な選択法であるとされる⁹⁾。

次に、各地点の需要量に対して、装置による効果の程度を表す「供給効果」を定義する。何よりも装置が有効に使用されるためには、心停止発生時点から限られた時間内に処置が施される必要がある。その際、各点から一定距離内にある領域に含まれる需要量を最大化する配置問題として知られる、最大被覆問題（Maximal Covering Location Problem）があげられる⁷⁾。またAEDの場合では、心停止発生から処置に至るまでの時間に、救命確率が大きく依存するとされるため⁸⁾、装置から近い場所であるほど救命の効果は高いとできる。こ

の装置からの距離と救命確率との関係を定式化するにあたり、処置までの時間に依存する「生存退院率」を用いた。「生存退院率」は処置による蘇生の成功率を意味し、倒れてから除細動の開始までに要する時間が1分経過するごとに約10%低下するとされることから⁸⁾、分速80mで装置に向かって直進するとしたうえで、救命確率を図2のように定めた。その際に、往復で10分を要する地点、すなわち装置から400m離れた地点で生存退院率がほぼ0になるとした。

本稿では、心停止が発生した現場に、一次救命を行う者とは別に装置を取りに行く者が存在することを前提としている。この装置を取りに行く救助者の置かれた状況を考えて、設置場所を知っている装置については、装置に向かって取りに行くことが可能である。ところが、当然ながら救助者は装置の全設置点を把握しているわけではない。その場合、近くに設置されている装置が多いほど、設置の効果は上昇するものとも考えられる。よっ

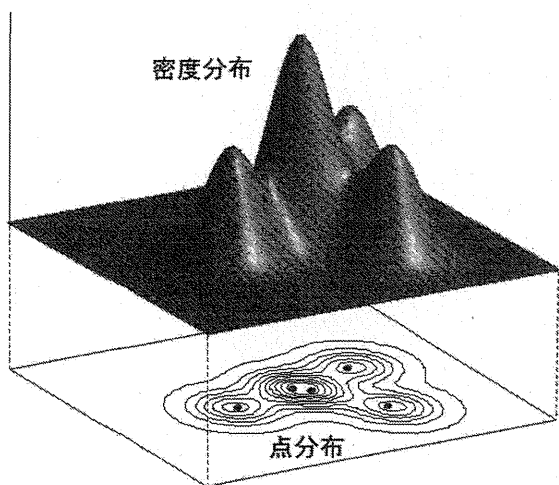


図1 点分布から密度分布への変換

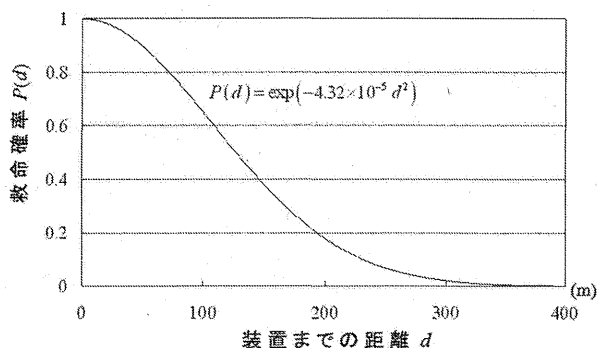


図2 AED装置からの距離と救命確率の関係

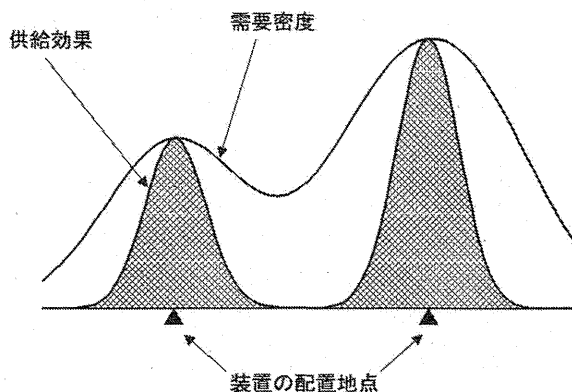


図3 需要密度に対する供給効果