

筑波大学大学院博士課程

システム情報工学研究科修士論文

トピックの位置変化を
地図上に表現する手法の開発

伊藤 聡

修士（工学）

（コンピュータサイエンス専攻）

指導教員 三末 和男

2014年3月

概要

位置を含むデータの分析においては、地図を用いた可視化手法の利用による分析支援が多く用いられる。位置付き時系列トピックデータの分析においても、地図を用いた可視化手法が有効であるが、従来の手法では一つの地図上に一つの時刻しか表すことができず、時間による頻度や位置の変化を地図上に表現することができなかった。

本研究では、Sun Burst Chartの外観をもとにした、円形と扇形を組み合わせた図形を地図上に描画する表現手法であるSun Burst Mapを開発した。Sun Burst Mapは、一つの地図上で頻度や位置の時間変化を把握可能な表現手法である。本手法は円周上に時間軸を割り当て、時刻ごとの頻度に対応した大きさの扇形を描画することによって、頻度の時間変化を表現している。この図形を位置に対応する地図上の地点へ配置することによって、一つの地図上で位置ごとの頻度の時間変化を表すことができ、様々な位置同士で頻度の時間変化を比較することで、位置の時間変化を把握することが可能となった。

本手法を利用した分析を行うためのツールを開発した。実データを用いたユースケースの提示により、本手法が位置付き時系列トピックデータの分析において有用であることを示した。

目次

第1章	はじめに	1
1.1	トピックデータ	1
1.2	位置付き時系列データ	1
1.3	位置付き時系列トピックデータ	2
1.4	可視化による分析と既存手法の問題点	3
1.5	本研究の目的	3
1.6	本研究の貢献	4
1.7	本論文の構成	4
第2章	関連研究	5
2.1	位置を持つデータの表現手法に関する研究	5
2.2	時系列データの表現手法に関する研究	5
2.2.1	直線上に時間軸を割り当てた表現手法	5
2.2.2	円周上に時間軸を割り当てた表現手法	6
2.3	位置付き時系列データの可視化に関する研究	6
2.3.1	Small Multiples を用いた表現手法	6
2.3.2	一つの地図上に時間変化を表した表現手法	6
2.4	位置付き時系列トピックデータの可視化に関する研究	7
第3章	対象データと分析要求	8
3.1	対象データ	8
3.2	対象データの定式化	9
3.3	分析の要求	9
第4章	表現手法: Sun Burst Map	11
4.1	表現手法の設計における要件	11
4.2	表現手法の概要	11
4.2.1	Sun Burst Chart について	12
4.3	マーカー表現の設計	13
4.3.1	時間と頻度の表現	13
4.3.2	扇形への色の割り当て	14
4.4	地図上への配置	14

第 5 章	分析ツールの開発	16
5.1	ツールの概観	16
5.2	ツールの機能	17
5.2.1	地図の操作	17
5.2.2	トピックの選択	17
5.2.3	パラメータの調節	17
5.2.4	頻度によるフィルタリング	19
5.2.5	時刻と頻度の情報を提示	19
5.2.6	マーカーのハイライト	19
5.3	実装	21
第 6 章	ユースケース	22
6.1	震災下のツイートデータ	22
6.2	観察	23
6.3	考察	26
第 7 章	議論	32
7.1	時間と頻度の表現方法について	32
7.2	Small Multiples との比較	32
7.3	今後の課題	33
7.3.1	複数のトピックの表現	33
7.3.2	マーカーの密集による視覚的混雑への対応	33
7.3.3	負の値の表現	33
第 8 章	まとめ	34
	謝辞	35
	参考文献	36

目次

1.1	Graduated Symbol Map に Small Multiples を適用した表現の例	3
4.1	Sun Burst Map による表現の例	12
4.2	マーカーについての説明	13
4.3	時系列への色の割り当て例	15
5.1	ツールの画面全体	16
5.2	ドロップダウンリストによるトピックの選択	18
5.3	マーカーの大きさを調節するパラメータの操作	18
5.4	頻度によるフィルタリング操作	19
5.5	ツールチップによる時刻と頻度の情報提示	20
5.6	マーカーのハイライトの例	20
6.1	トピック「津波」を表示したツール画面	24
6.2	図 6.1 からパラメータを調整した状態の地図ビュー	25
6.3	トピック「津波」に見られる扇形のパターン例: 黄色、黄緑色、緑色、水色 (マーカーの左下部分)、青色・紫色 (マーカーの左上部分) の 5 パターンが確認できる	26
6.4	北海道から東北にかけての地図ビューの表示	27
6.5	関東一帯の地図ビューの表示	28
6.6	西日本一帯の地図ビューの表示	29
6.7	図 6.3 の各パターンに対応する扇形の部分	30

表目次

1.1	トピックデータの例	1
1.2	位置付き時系列データの例	2
1.3	位置付き時系列トピックデータの例	2
3.1	本研究の対象データの例	8
6.1	震災下のツイートデータをもとに作成された位置付き時系列トピックデータの例	22

第1章 はじめに

1.1 トピックデータ

現実世界やインターネット上においては、人々の間で話題となる物や出来事が日常的に発生している。「地震が起こった」、「新型の携帯端末が発売された」などはその一例である。本論文では、このような話題となる物や出来事に繋がるような単語を、「トピック」として扱い、それがどのくらい話題になっているかを表す「頻度」と組み合わせて記録されたデータを、「トピックデータ」と呼ぶことにする。トピックデータの一例を、表 1.1 に示す。

表 1.1: トピックデータの例

トピック	頻度
地震	2525
iPhone	2014
発売	1878
一限	123
布団	140

トピックデータには、「人々の間でどんなことが、どの程度話題になっているのか」という情報が含まれている。こうしたデータを分析することによって、世の中で起こっている出来事や、トレンドの把握に役立てることができる。

1.2 位置付き時系列データ

位置付き時系列データとは、位置を表す緯度と経度の組と、それらに対応する時刻を合わせ持つデータである。例えば、利用者のいる位置と時刻を記録した GPS 記録データは、位置付き時系列データの一つである。位置付き時系列データの一例を、表 1.2 に示す。

位置付き時系列データには、地理的な位置の情報に加えて、時間による位置の変化に関する情報が含まれている。こうしたデータを分析することによって、地理空間上における分布や移動についての知見を得ることができる。

表 1.2: 位置付き時系列データの例

位置 (緯度, 経度)	時刻
(32.9, 132.7)	3/1 11:00
(33.2, 133.1)	3/1 13:00
(33.4, 133.2)	3/1 17:00
(33.5, 133.5)	3/1 19:00

1.3 位置付き時系列トピックデータ

本論文が対象とするデータは、上述した「トピックデータ」と「位置付き時系列データ」の双方の特徴を合わせ持つデータである。具体的には、「トピック」、「位置」、「時刻」、「頻度」の四つを属性として持つデータとなる。このようなデータを、本論文では「位置付き時系列トピックデータ」と呼ぶ。位置付き時系列トピックデータの一例を表 1.3 に示す。

表 1.3: 位置付き時系列トピックデータの例

トピック	位置 (緯度, 経度)	時刻	頻度
地震	(32.9, 132.7)	3/1 11:00	320
地震	(33.2, 133.1)	3/1 11:00	255
地震	(32.9, 132.7)	3/1 12:00	210
停電	(33.2, 133.1)	3/1 12:00	302
津波	(33.2, 133.1)	3/1 12:00	283
地震	(33.2, 133.1)	3/1 12:00	270
停電	(33.2, 133.1)	3/1 13:00	420
津波	(33.4, 133.2)	3/1 13:00	315
断水	(32.9, 132.7)	3/1 13:00	139

位置付き時系列トピックデータとして扱うことのできる情報には、さまざまなものが存在する。例として、Twitter¹などのマイクロブログ上のデータが挙げられる。一つのつぶやきに含まれる名詞を「トピック」、同じつぶやきに含まれる Geo タグの緯度経度を「位置」、つぶやかれた日時を「時刻」として扱い、単位時間ごとのトピックの出現回数を「頻度」とする。このような方法によって、マイクロブログ上のデータは位置付き時系列トピックデータの形に整理することができる。

位置付き時系列トピックデータを分析していくと、トピックの地理的な分布やその時間変化といった特徴を見ることができる。これによって、トピックデータでは分からなかった「現実世界における出来事やトレンドの地域性」に関する知見を得ることができる。

¹<https://twitter.com/>

1.6 本研究の貢献

本研究の貢献は、以下の二点である。一つ目の貢献は、トピックの頻度の値や位置の時間変化を一枚の地図上で表すことが可能な手法を開発したことである。この表現手法は一枚の地図上で収まる表現であるため、Small Multiples に比べて大きな画面領域を必要としない。また、円周上に時間軸を割り当てているため、離れた位置同士でも時間ごとの頻度を比較することが容易になっている。本手法によって、トピックの地理的な分布やその時間変化といった特徴をより簡単に把握し、知見を得ることが可能となった。

二つ目の貢献は、本手法を実装した位置付き時系列トピックデータの分析ツールを開発したことである。開発したツールには、本手法を実際の分析タスクに利用するために、ユーザー側が操作を行ういくつかの機能が実装されている。

1.7 本論文の構成

第2章では、本研究に関連する研究を四つの観点から紹介する。第3章では、本研究で対象とするデータとその分析要求について述べる。第4章では、本研究で開発した表現手法について述べ、第5章ではその手法を実装した分析ツールについて述べる。第6章では、本手法を用いた分析のユースケースを提示する。第7章では、表現手法についての議論を行い、第8章では本研究のまとめを述べる。

第2章 関連研究

本章では、本研究に関連する研究を四つの観点から紹介する。第一に位置を持つデータの表現手法に関する研究、第二に時系列データの表現手法に関する研究、第三に位置付き時系列データの可視化に関する研究、第四に位置付き時系列トピックデータの可視化に関する研究である。

2.1 位置を持つデータの表現手法に関する研究

位置を持つデータの可視化においては、地図を用いた表現手法が利用されることが多く、これまでに様々なものが提案されてきている [1]。

Graduated Symbol Map[2] は、大きさの異なる図形を用いて属性の値を表す地図である。図形の大きさが属性の値に比例し、地域の人口や駅の利用者数など、地理上の位置に関連付けられた量的な情報を表すのに役立つ。Graduated Symbol Map は多くの可視化手法に取り入れられており、本論文で提案する表現手法は、この表現の拡張である。同様に Graduated Symbol Map を拡張した手法として、Tominski らの手法 [8] がある。この手法では、図形同士をエッジで結ぶことによって、ネットワークデータを地図上に表現することを可能にしている。

Nguyen らの開発した Taggram[9] は、地図と TagClouds[10] を組み合わせた手法である。地図上に TagClouds を描画することによって、位置に関連するタグとその頻度を地図上に表現する。Ahern らの開発した World Explorer[11] も同様の表現を用いているが、写真のビューと組み合わせることによって、タグと位置に関連する写真を見ることができるようになっている。

さらに、地図上の道、川、公園といった構造物を全て文字列によって表現する Typographic Map[12] という表現手法がある。Afzal らは、この Typographic Map を自動生成する手法 [13] を開発している。

しかし、これらの手法は一つの地図上で一つの時刻の情報しか表すことができず、値の時間変化を表現することができないという問題がある。

2.2 時系列データの表現手法に関する研究

2.2.1 直線上に時間軸を割り当てた表現手法

時系列データを表現する最も基本的な手法として、折れ線グラフや面グラフがある。横軸に時間軸を、縦軸に値を割り当てて、値の時間変化を表現することができる。Lee らの開発し

た SparkClouds[14] は、TagClouds に折れ線グラフや面グラフを組み合わせて表示することにより、各タグに対応した値の時間変化を表すことができるようになっている。

Havre らは、時系列データの表現手法の一つである ThemeRiver[15] を開発した。この手法は、縦軸方向の幅によって値を表現する手法であり、積み上げ式面グラフの変形である。Wei らの開発した TIARA[16] は、トピックの時間変化を表す ThemeRiver の内部に TagClouds を埋め込むことによって、各トピックが持つキーワードの時間変化を同時に表現している。Cui らの開発した TextFlow[17] は、ThemeRiver にグラフィレイアウトを組み合わせることで、枝分かれや合流も視覚的に表現することを可能にしている。

これらの手法は、いずれも二次元平面上の直線に時間軸を割り当てたものであり、円周上に時間軸を割り当てる本研究の手法とは異なる。

2.2.2 円周上に時間軸を割り当てた表現手法

Shiroi らの ChronoView[18] は、円周上に時間軸を割り当て、二次元平面上の位置によって時刻の集合を表現している。また、Carlis らの開発した手法 [19] では、螺旋状の線に時間軸を割り当て、週、月、年といった連続する時間の周期性を表している。Dragicevic らの SpiraClock[20] は、螺旋状の表現にアナログ時計の表現を組み合わせた手法であり、直近の数時間程度の範囲における情報を提示する。

これらの手法は、周期性を比較することを目的として設計されており、時間変化を表現することを目的とした本研究とは異なる。

2.3 位置付き時系列データの可視化に関する研究

2.3.1 Small Multiples を用いた表現手法

Small Multiples[5] は、図を格子状に区切った領域に並べて表示する表現手法である。Maceachren らの開発した GeoVISTA Studio[21] は、Cholopleth Map、散布図、棒グラフなどの様々な図を、Small Multiples を利用して行列状に提示する。行と列がそれぞれ属性に対応しており、属性の特性によって用いる図を使い分けることができるため、効率的に多角的な分析を行うことができる。しかし、Small Multiples は行や列を多く表示しようとする、大きな画面領域が必要となるため、時系列などの多くの値を取りうる属性を俯瞰することが困難になる。

2.3.2 一つの地図上に時間変化を表した表現手法

Thakur らの開発した手法 [6] は、3D 空間上の XY 平面に地図を描画し、Z 軸に時間軸を割り当てることによって、位置ごとの値の時間変化を表現している。この手法では、時間軸の基準となる画面上の位置がずれてしまうために、離れた位置同士で時間ごとの頻度を比較することが難しいという問題点がある。

Scheepens らの開発した Density Map[22] は、移動の軌跡を地図上に表現する手法である。軌跡の密度を色の濃さによって表現し、異なる時間帯の軌跡に異なる色相を割り当てることで、時刻を区別している。Tominski らは、3D 空間上の Z 軸に移動の軌跡を積み重ねて表現することで、軌跡の持つ値の時間変化を表現可能にした手法 [23] を開発している。これらの手法では、位置に関連付けられた値を色によって表現しているが、本研究では図形の大きさによって表現する。

2.4 位置付き時系列トピックデータの可視化に関する研究

地図を含む複数の手法を組み合わせた可視化手法の中には、位置付き時系列トピックデータの可視化として捉えることのできるものが存在する。

Marcus らの開発した TwitInfo[24] は、Twitter のつぶやきデータからイベントの分析を行うために、折れ線グラフと地図を用いてつぶやきデータを可視化するツールである。特定のキーワードに関係するつぶやきを抽出し、折れ線グラフでつぶやき数の時間変化を、地図上へのピンの描画によって Geo タグの持つ位置を表現している。Cao らの開発した Whisper[25] は、ソーシャルメディアにおいて情報が拡散していく過程を可視化するツールである。トピックに関する情報を表すノードを円形の中心に、ユーザーを表すノードを周囲に配置し、リツイート関係をエッジとして両者を結んでいる。ユーザーを表すノードは、背景に描画された世界地図に対応した位置へ配置されるようになっている。Ferreira らの開発した BirdVis[26] は、鳥の生態系を調査するためのツールである。HeatMap を用いて地図上に鳥の生息数を表現し、Small Multiples で複数の地図を並べることによって、位置と生息数の時間変化を表現することができる。Dou らの開発した LeadLine[27] は、ThemeRiver、Graduated Symbol Map、TagClouds、Graph などの複数の表現手法によるビューを組み合わせることで、位置付き時系列トピックデータの分析を行うことが可能なツールである。特に頻度の高い時間帯を「イベント」と定義し、各イベントに着目して、さらに詳細な分析を行うことができる。Lundblad らの RoadVis[28] は、Graduated Symbol Map、折れ線グラフ、Parallel Coordinates[29] の三つのビューを用いて、各地の道路状況を視覚的に提示するツールである。Graduated Symbol Map と Parallel Coordinates で、ある時刻における全体的な道路状況を俯瞰し、折れ線グラフによって各道路状況の時間変化を見ることができ。これらに用いられた地図では、地図上の対応する位置にトピックの頻度を提示することができるが、提示されるのは時系列の一点における頻度や区間の合計頻度であり、頻度の時間変化を一つの地図上に表すことはできない。

Walker らの開発した手法 [30] は、半球状の地図と Parallel Coordinates を組み合わせることによって、位置に対応する複数の属性を一度に表現している。この手法では、Parallel Coordinates を構成する軸の一つに時間軸を割り当てることで時間変化を表現できるが、描画される線の数が膨大なものになり、視覚的に読み取ることが困難になるという問題がある。

第3章 対象データと分析要求

本章では、本研究が対象とするデータについて検討し、対象データの分析における要求を整理する。

3.1 対象データ

第1.3節で述べたように、本研究が対象とするデータは「トピック」、「位置」、「時刻」、「頻度」の四つを属性として持つ、位置付き時系列トピックデータである。表3.1に、本研究の対象データの一例を示す。

表 3.1: 本研究の対象データの例

トピック	位置 (緯度, 経度)	時刻	頻度
地震	(32.9, 132.7)	3/1 11:00	320
		3/1 12:00	210
	(33.2, 133.1)	3/1 11:00	255
		3/1 12:00	270
停電	(33.2, 133.1)	3/1 12:00	302
		3/1 13:00	420
津波	(33.2, 133.1)	3/1 12:00	283
	(33.4, 133.2)	3/1 13:00	315
断水	(32.9, 132.7)	3/1 13:00	139

本研究では、一つのトピックは複数の位置を持ち、それぞれの位置に時間とともに変化する頻度が割り当てられているものとする。トピックは文字列、位置は緯度と経度を表す実数の組、時刻は時系列上の区間、頻度は正の整数によってそれぞれ表される。

例えば、表3.1の「地震」というトピックは二つの位置を持っており、それぞれの頻度が時間とともに変化している。また、「停電」というトピックの(33.2, 133.1)という位置における3月1日11時の頻度など、表に記されていない部分もある。このような場合は、その頻度を0として扱うものとする。

3.2 対象データの定式化

先の節での説明に基づいて、本研究が対象とするデータの定式化を行う。トピックの集合を W で表す。 W の要素は、個々のトピックを表す文字列とする。

$$W = \{w_1, w_2, \dots, w_l\} \quad (3.1)$$

位置の集合を P で表す。 P の要素は、 -90 以上 90 以下の範囲にある緯度を表す実数と、 -180 より大きく 180 以下の範囲にある、経度を表す実数の組とする。

$$P = \{p_1, p_2, \dots, p_m\} \quad (3.2)$$

$$p_i \in [-90, 90] \times (-180, 180], (i = 1, 2, \dots, m) \quad (3.3)$$

時刻の集合を T で表す。 T の要素は、連続する区間とする。

$$T = \{t_1, t_2, \dots, t_n\} \quad (3.4)$$

頻度はトピック、位置、時刻の直積集合から正の整数への関数 f で表される。トピック w 、位置 p 、時刻 t における頻度を、 $f(w, p, t)$ で表す。

$$f : W \times P \times T \rightarrow \mathcal{N} \quad (3.5)$$

表 3.1 のデータを例にとって説明する。トピック「地震」を w_1 とし、位置 $(32.9, 132.7)$ を p_1 とする。また、3月1日の11時台、12時台、13時台をそれぞれ t_1 、 t_2 、 t_3 とする。このとき、 $f(w_1, p_1, t_1)$ は320、 $f(w_1, p_1, t_2)$ は210となる。 $f(w_1, p_1, t_3)$ は表に記されていない、つまりデータ中に含まれていないため、0となる。

3.3 分析の要求

位置付き時系列トピックデータの分析においては、「位置」、「時間」、「頻度」が大きな意味を持つ。まずは「時間」と「頻度」に着目して、頻度が時間とともにどのような変化をするのかを知りたい、という要求がある。

要求 1 頻度が時間とともにどのように変化していくのかを知りたい。

この要求に対しては、Excel のグラフ機能を利用するなどして、面グラフや折れ線グラフを見ることでも解決が可能である。しかし、これに「位置」の要素を加えた以下の様な要求は、従来の面グラフや折れ線グラフだけでは達成が困難となる。なぜなら、「位置」は取りうる値が多く、頻度の時間変化を多くの位置同士で比較することが難しいためである。

要求 2 様々な位置における頻度の時間変化を比較したい。

要求 3 位置と頻度の関係性についての特徴を知りたい。

要求 4 位置と頻度の関係性が、時間によってどう変化するかを知りたい。

さらに、「位置」は実世界の地理と深く関連しているため、面グラフや折れ線グラフをただ並べて示すだけでは、データの地理的な特徴を把握しきれない事がある。例えば、「10:00ではA地点よりB地点の方が頻度が高い」「12:00ではB地点よりC地点のほうが頻度が高い」というような事は把握できても、「10:00では頻度の高い地点が東の方に集中していたが、12:00には西の方へ移った」というような特徴を把握することは難しいと言える。

第4章 表現手法: Sun Burst Map

本章では、第3章で検討した要求をもとに、開発した表現手法である Sun Burst Map について説明する。本研究が対象とする位置付き時系列トピックデータは、実世界の地理と関連する「位置」の属性を持っている。そのため、可視化においては実世界の地理を表現することが可能な、地図を用いた手法が大いに有効であると考えられる。したがって、本研究では地図を用いた表現手法について考える。

4.1 表現手法の設計における要件

第3.3節で示した分析の要求を踏まえると、位置付き時系列トピックデータの可視化手法の要件としては以下の四つが考えられる。

要件1 頻度の時間変化を俯瞰できること。

要件2 多くの位置における頻度の時間変化を同時に比較できること。

要件3 位置と頻度の関係性についての特徴が読み取れること。

要件4 位置と頻度の関係性の時間変化が読み取れること。

4.2 表現手法の概要

本手法は、Graduated Symbol Map を拡張したものである。ひとつの Graduated Symbol Map では、ある一つの時刻における頻度は表せるが、頻度の時間変化を表現することはできなかった。そこで、Sun Burst Chart をもとにして、頻度の時間変化を表現することができるような円形の表現を開発し、地図上に配置することを考えた。本研究では、この地図上に配置する個々の図形を「マーカー」と呼称する。図4.1に、本手法による表現の例を示す。

本手法では、円周上に時間軸を割り当てており、円周から伸びる扇形の向きによって時間の区別ができるようになっている。また、頻度を扇形の面積によって表現している。開発したこの表現により、頻度の時間変化を俯瞰することができるようになった（要件1）。加えて、マーカーを地図上の対応する位置へ配置することにより、「頻度の高い位置は東北地方の沿岸部に多い」というような、位置と頻度の関係性についての特徴が読み取れるようになった（要件3）。

また、時間軸の基準となる向きが全てのマーカーで同じであるため、画面上のマーカーの配置において縦や横の位置が揃ってなくても、時間変化を比較することが可能である。したがって、多くの位置における頻度の時間変化を同時に比較することができる（要件2）。さらに、こうした比較作業を可能にしたことで、位置と頻度の関係性が時間によってどう変化するかを読み取ることも可能となっている（要件4）。

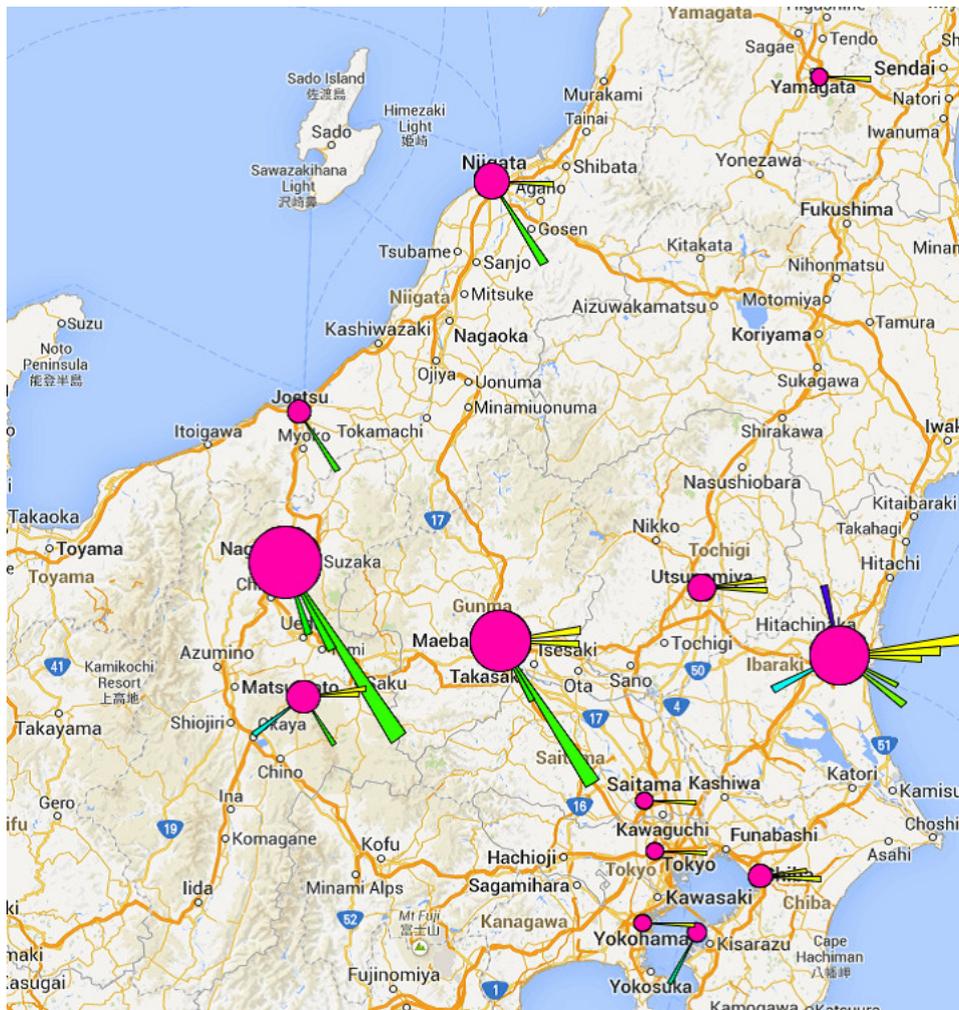


図 4.1: Sun Burst Map による表現の例

4.2.1 Sun Burst Chart について

Sun Burst Chart[7] は、円形表現を用いた階層構造表現手法の一つである。各ノードを扇形で表し、子ノードを親ノードの外側につなげて配置することによって、親子関係を表現している。扇形の大きさは、各ノードの重みに対応している。本研究が提案する表現手法は、Sun

Burst Chart の外観をもとにして設計されているが、円周上に時間軸を割り当てることによって、時間変化を表すことができる点が異なっている。

4.3 マーカー表現の設計

本手法では、マーカーによってトピックの頻度、およびその時間変化を表現している。一つのマーカーは、一つのトピックの一つの位置に対応しており、ここではトピック w の位置 p に対応するものとして説明する。以下、マーカーの設計について図 4.2 をもとに説明する。

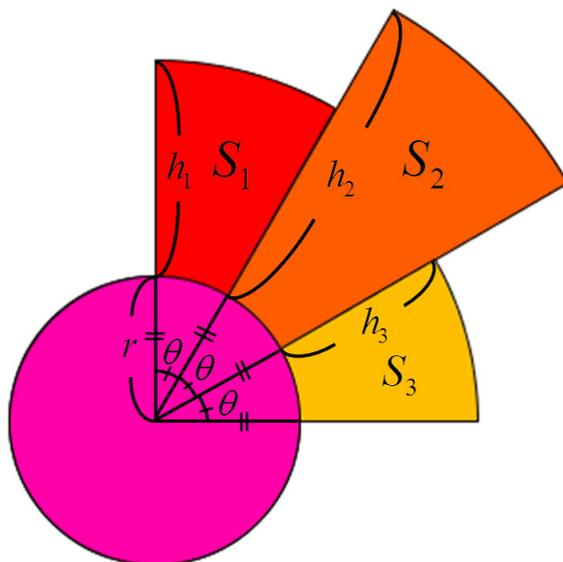


図 4.2: マーカーについての説明

4.3.1 時間と頻度の表現

マーカーは、中心に描画された一つの円と、その円周から伸びる複数の扇形によって構成されている。扇形は時刻の集合の要素数だけ描画され、中心角は全ての扇形で同じになるようにする。したがって、個々の扇形の中心角 θ は次のようになる。

$$\theta = \frac{2\pi}{|T|} \quad (4.1)$$

それぞれの扇形は一つの時刻に対応しており、真上を始点として時計回りの順番で割り当てられる。 i 番目の扇形の面積 S_i は、 i 番目の時刻における頻度に応じて決めるようにする。すなわち、 S_i は次の式で表される。 α は、扇形の大きさを調節するための係数である。

$$S_i = \alpha \cdot f(w, p, t_i), (\alpha \geq 0) \quad (4.2)$$

また、中心の円の面積が全時刻の頻度の合計に比例するようにする。よって、中心の円の半径 r は、次の式で表される。 β は α と同様に、中心の円の大きさを調節するための係数である。

$$r = \sqrt{\beta \cdot \sum_{j=1}^n f(w, p, t_j)}, (\beta \geq 0) \quad (4.3)$$

式 4.2、4.3 より、始点から i 番目の扇形の半径 h_i は、次のように表される。

$$h_i = -\sqrt{\beta \cdot \sum_{j=1}^n f(w, p, t_j)} + \sqrt{\beta \cdot \sum_{j=1}^n f(w, p, t_j) + \frac{2\alpha f(w, p, t_i)}{\theta}}, (\alpha, \beta \geq 0) \quad (4.4)$$

なお、頻度が 0 となった場合には半径が 0 となるため、扇形は描画されない。さらに、全ての時刻で頻度が 0 となった場合は、中心の円も描画されない。 α か β の値に 0 が設定された場合も、同様に描画されなくなる。

4.3.2 扇形への色の割り当て

それぞれの扇形には一つの色を割り当て、その色で内部を塗りつぶす。割り当てられる色の色相は扇形の順番によって決まり、彩度と明度の値は一定である。 i 番目の扇形に割り当てられる色相 C_i は、次のようになる。

$$C_i = \frac{3}{2}\pi \cdot \frac{i-1}{n-1}, (i = 1, 2, \dots, n) \quad (4.5)$$

色の割り当てに用いる色相は、色相環における 0 から $\frac{3}{2}\pi$ の範囲とする。これは色相の全範囲を割り当てた場合、1 番目の扇形と最後の扇形の色相が連続してしまうために、隣接していない時間が隣接しているかのように見えてしまうためである。一方で、割り当てる色相の範囲が狭いと、隣接する時間同士の区別が難しくなってしまう。そこで、本手法では適度に隣接する時間同士の区別が付きやすく、隣接していない時間が隣接しているように見えない割り当て方として、上記の範囲を採用した。色相の割り当て方を、図 4.3 に示す。

中心の円の色は特に定めないが、扇形に用いられる色相の範囲にない色を用いるのが良いと考えられる。これは、扇形と似通った色が用いられると、観察者に誤った印象を与える可能性があるためである。

4.4 地図上への配置

前述のように設計したマーカーを、位置に対応する地図上の地点へ配置する。異なる位置で頻度の時間変化の仕方が異なっていれば、対応するマーカーの扇形の数や向きも異なるも

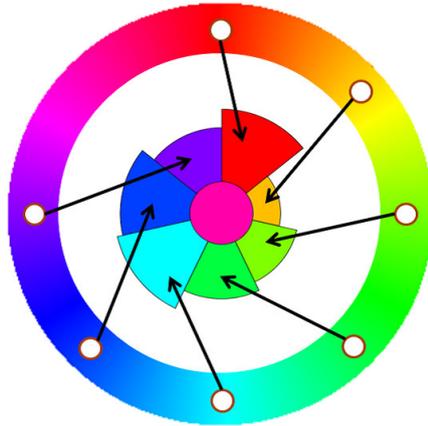


図 4.3: 時系列への色の割り当て例

のになる。一方で、頻度の時間変化の仕方が似通ってれば、対応するマーカーの扇形の数や向きも似通ったものになる。このように、地図上に配置されたマーカーを比較することにより、位置と頻度の関係性が時間によってどう変化するかを把握することができる。

第5章 分析ツールの開発

第4章で述べた Sun Burst Map を利用した、可視化によるトピック分析を行うためのツールを開発した。図 5.1 は、本ツールの画面全体を示した図である。本章では、分析を支援するために本ツールが備えた機能や、実装についての説明を行う。

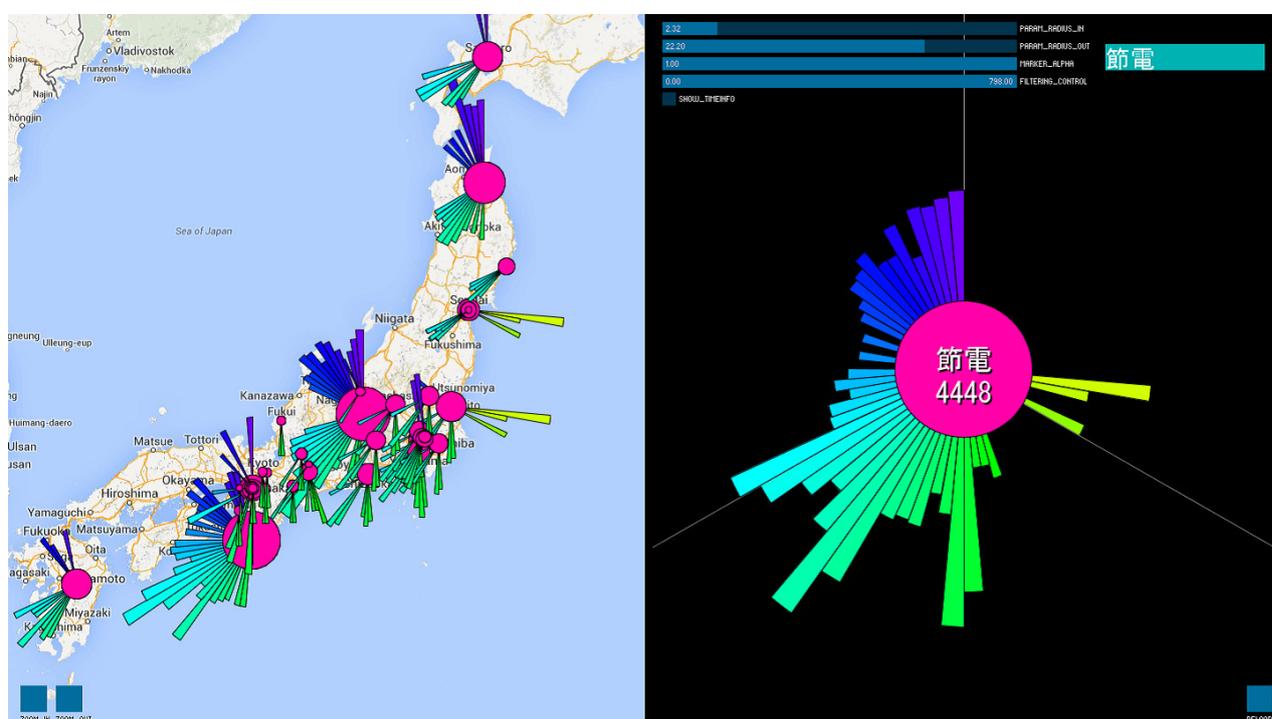


図 5.1: ツールの画面全体

5.1 ツールの概観

本ツールは、位置付き時系列トピックデータをもとに、Sun Burst Map を用いた可視化結果をユーザーに提示する。ユーザーは可視化結果を見ながら、インタラクションを通してデータの分析を行う。ツールの画面は中央で二分割されており、左側が地図ビュー、右側がトピックビューとなっている。地図ビューでは、一つのトピックについての可視化結果を Sun Burst

Map を用いて提示する。トピックビューでは、一つのトピックの全位置の合計頻度と、その時間変化を Sun Burst Map のマーカーによって提示する。また、日付の境界となる部分の背景に直線を描画している。双方のビューには、一つの同じトピックについての情報が提示される。

5.2 ツールの機能

ユーザーによる分析作業を支援するために、各ビューから操作可能ないくつかの機能が実装されている。

5.2.1 地図の操作

地図ビュー上でマウスを右クリックすることにより、地図のスクロールを行うことができる。スクロールは、右クリックした地点がビューの中心となるように行われる。また、地図ビューの左下部分にはズームイン、ズームアウトのボタンが配置されている。各ボタンをクリックすることで、地図の拡大縮小を行うことができる。拡大縮小は地図に対してのみ行われ、マーカーの大きさは変化しない。

5.2.2 トピックの選択

トピックビューの右上部分には、現在提示されているトピックが表示されている。この部分はドロップダウンリストになっており、リストから選択することで、提示するトピックを切り替えることができる。リストの並び順は、トピックの合計頻度が高い順となっている。また、何らかの理由で正しくデータが読み込まれなかった場合（合計頻度が0になってしまう等）は、トピックビューの右下部分にあるリロードボタンを押すことで、データの読み込みを再実行することができる。

5.2.3 パラメータの調節

トピックビューの上部には、マーカーの描画に関わる各種パラメータを調節するためのスライダーが用意されている。一番上のスライダーは、マーカーの中心部分の円の大きさの倍率を調節する。二番目のスライダーは、マーカーの扇形部分の大きさの倍率を調節する。三番目のスライダーは、マーカーの透明度を調節する。ユーザーは、提示された結果の特徴やユーザーの見た情報に合わせて、これらのパラメータを調節することができ、適切なビューを見ることができる。

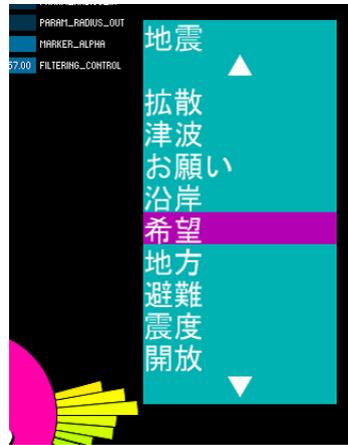


図 5.2: ドロップダウンリストによるトピックの選択

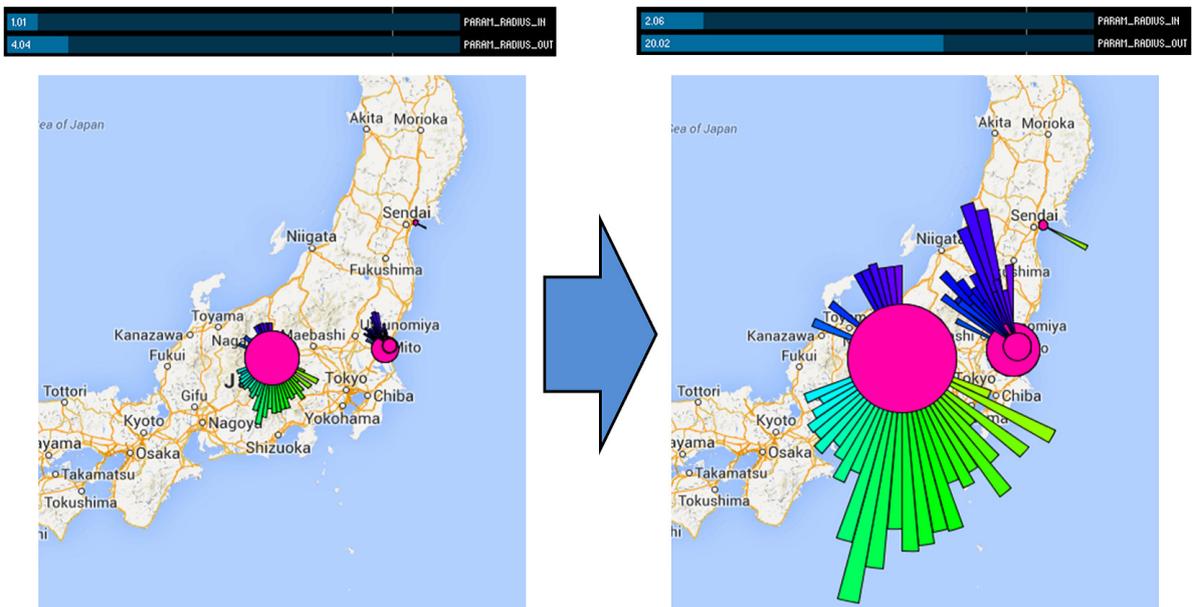


図 5.3: マーカーの大きさを調節するパラメータの操作

5.2.4 頻度によるフィルタリング

トピックビューの上部にある上から四番目のスライダーは、頻度によるフィルタリングを行うためのものである。表示範囲の上限、下限がともに操作できるようになっており、合計頻度が表示範囲外のマーカーは地図ビュー上に表示されなくなる。これにより、極端に大きなマーカーを非表示にして密集した小さなマーカーを見やすくしたり、逆に小さなマーカーを非表示にして、重なり合っただけで見えなくなっていたマーカーが見えるようにする、といったことが可能となる。

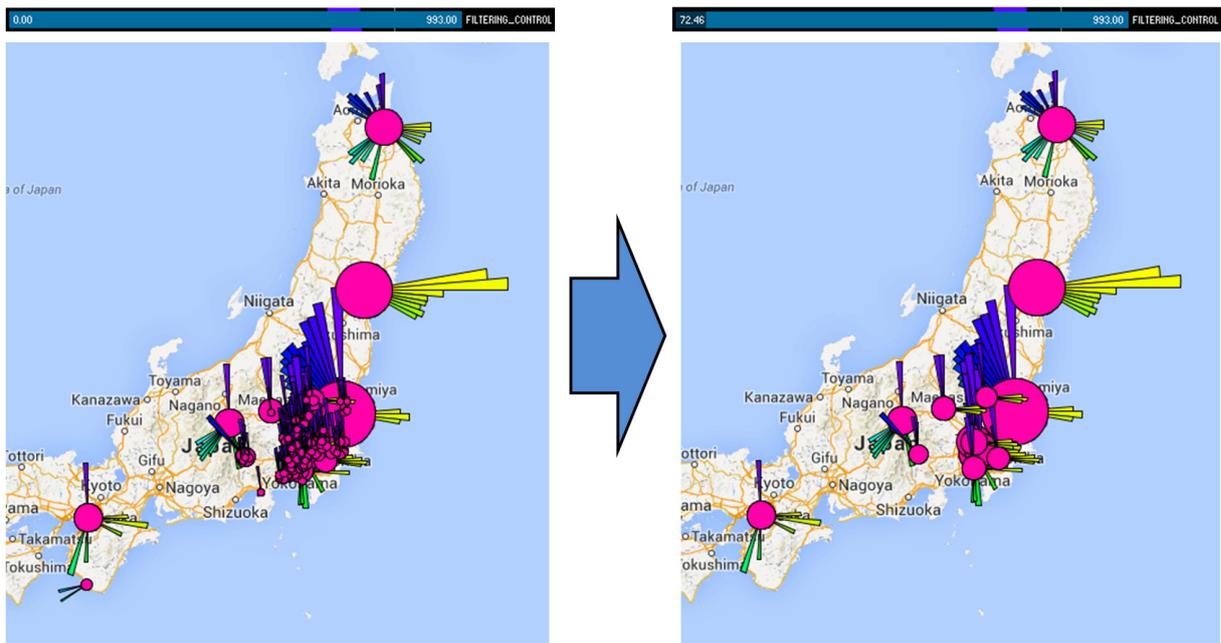


図 5.4: 頻度によるフィルタリング操作

5.2.5 時刻と頻度の情報を提示

トピックビューのスライダー群の下にあるスイッチをオンにすることで、時刻と頻度の情報を表示することができる。トピックビューに表示されたマーカーの中心からマウスポインタに向かって直線が引かれ、直線と重なっている扇形が選択状態になる。選択された扇形に対応する時刻と頻度が、マウスポインタの右上にツールチップで表示される。

5.2.6 マーカーのハイライト

上記の操作によって時刻と頻度の情報を提示している間は、地図ビュー上のマーカーのハイライトが同時に行われる。地図ビュー上のマーカーのうち、選択されている扇形と同じ時刻



図 5.5: ツールチップによる時刻と頻度の情報提示

に対応する扇形の輪郭が、太く表示される。同時に、ハイライトされた扇形を持つマーカーの中心円も、輪郭が太く表示される。

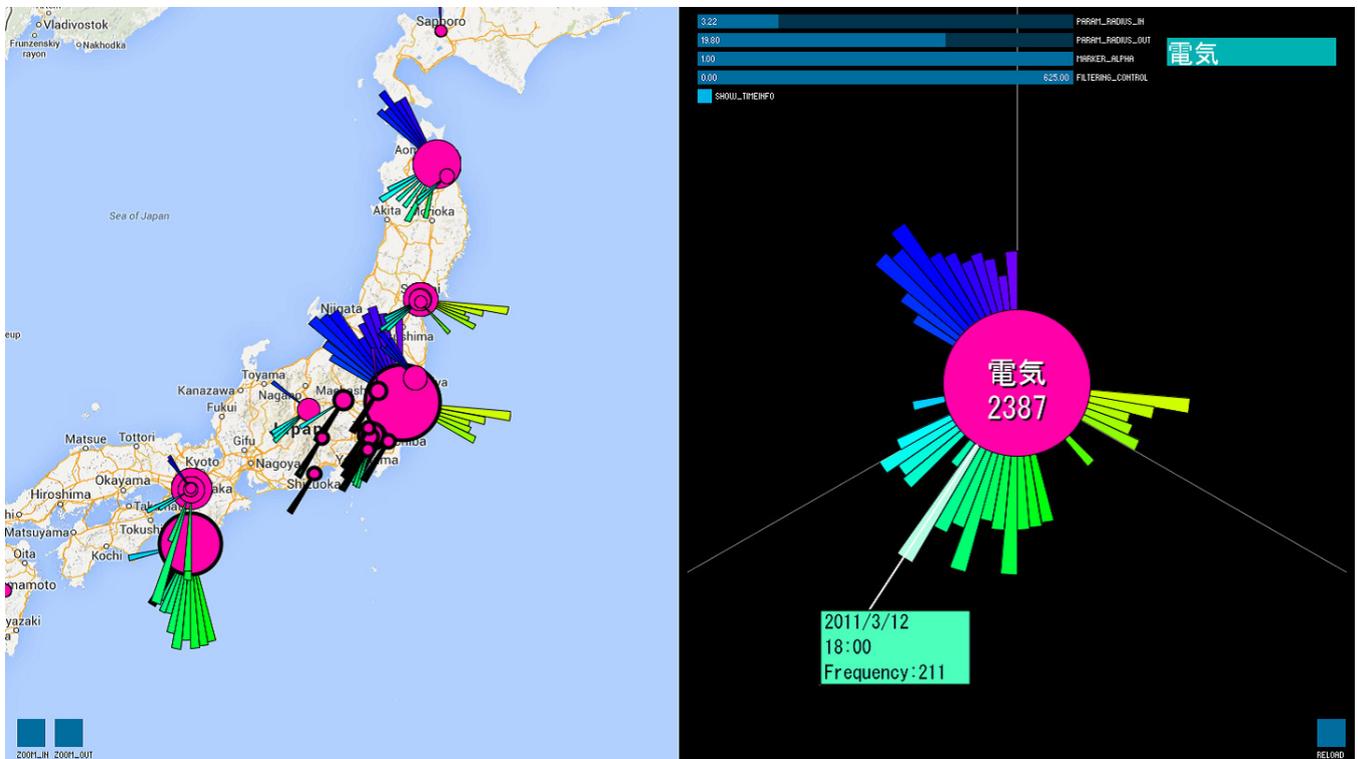


図 5.6: マーカーのハイライトの例

5.3 実装

本ツールは、Processing2.1¹ を用いて開発した。プログラムのサイズは 837 行である。データは端末上にインストールされた MySQL² に格納されており、データベースへのアクセスに拡張ライブラリの BezierSQLib³ を利用している。また、地図の描画には拡張ライブラリの GoogleMapper⁴ を利用している。

¹<http://processing.org/>

²<http://www.mysql.com/>

³<http://bezier.de/processing/libs/sql/>

⁴<http://googlemapper.pt.vu/>

第6章 ユースケース

本章では、Sun Burst Map の有用性を示すために、ユースケースの紹介を行う。今回行った分析は、2011年3月11日に発生した東日本大震災におけるトピックの分析である。分析には、Twitter のツイートを整理し、位置付き時系列トピックデータとしてまとめられたデータを用いた。実際の分析作業を通して、データからどのような知見が得られるのかを調査する。

6.1 震災下のツイートデータ

表 6.1: 震災下のツイートデータをもとに作成された位置付き時系列トピックデータの例

トピック	位置 (緯度, 経度)	時刻	頻度
地震	(32.9, 132.7)	3/11 15:00-15:59	320
地震	(33.2, 133.1)	3/11 15:00-15:59	255
地震	(32.9, 132.7)	3/11 16:00-16:59	210
停電	(33.2, 133.1)	3/11 16:00-16:59	302
津波	(33.2, 133.1)	3/11 16:00-16:59	283
地震	(33.2, 133.1)	3/11 16:00-16:59	270
停電	(33.2, 133.1)	3/11 17:00-17:59	420
津波	(33.4, 133.2)	3/11 16:00-16:59	315
断水	(32.9, 132.7)	3/11 16:00-16:59	139

利用するデータは、Twitter のツイートをもとにして作成された位置付き時系列トピックデータである。Takahashi らは、Twitter は実世界の出来事を捕捉するセンサとしての役割を果たすことができると述べており [31]、本データからも実世界で起こった出来事に関する知見が得られることが期待される。

対象となるのは、2011年3月11日0時から2011年3月14日0時までのツイートである。対象のツイートを形態素解析し、名詞に分類された単語を「トピック」として扱う。また、Geocoding¹を利用して、地名に分類された単語から得られた緯度、経度を「位置」として扱う。そして、ツイートされた日時を「時刻」とし、「トピック」として扱われる単語の登場回数を位置ごとに1時間単位で集計する。集計の結果、「トピック」として扱われる単語が、単位時間あたりに登場する回数を「頻度」として扱う。

¹<http://www.geocoding.jp/>

本データの一例を表 6.1 に示す。本データには、レコードが 25590 件、異なる位置が 750 件、異なるトピックが 3146 件含まれている。また、単位時間ごとの頻度において、0 でない最小値は 10、最大値は 4199 である。

6.2 観察

ツールを起動し、トピック一覧から表示するトピックを選ぶ。図 6.1 は、トピック「津波」を表示したツール画面である。円の大きさのパラメータは 1.00、扇形の大きさのパラメータは 5.62 に設定した。右側のトピックビューを見ると、3 月 11 日の後半に扇形が連続している区間（黄色の部分）があり、この区間の先頭から数えて 2 番目と 3 番目の扇形が飛び抜けて大きいことが分かる。時間情報を表示して確認すると、この扇形に対応する時刻は 2011 年 3 月 11 日の 16 時台と 17 時台であり、その頻度はそれぞれ 2350、3651 であった。また、同区間の 4 番目と最後の扇形も他と比べて大きく、それらに対応する時刻は 2011 年 3 月 11 日の 18 時台と 23 時台である。同様に頻度を確認すると、18 時台は 964、23 時台は 1113 であった。

左側の地図ビューを見ると、マーカーが北海道から九州までの太平洋側を中心に広く分布している。全体的に見て、県庁所在地に位置するマーカーの方が、他の地域に位置するマーカーよりも大きい傾向にあることが分かる。また、多くのマーカーでは、少ない数の扇形が 1、2 方向に集中して表れている様子が見て取れる。しかし、日本のほぼ中央となる位置に最も大きなマーカーがあり、このマーカーだけはトピックビュー側とほぼ同じような形で、広い時間帯に扇形が表れている。

図 6.2 は、図 6.1 の状態から大きさのパラメータを調節することで、マーカーを見やすくしたものである。円の大きさのパラメータは 1.61、扇形の大きさのパラメータは 23.02 に設定した。前述の最も大きなマーカーは日本の中心に位置し、他と比べて明らかに大きいために、「日本」という地名から得られた位置と頻度を表していると判断した。「日本」よりも詳細な粒度での位置について観察するために、このマーカーをフィルタリング機能で非表示にしている。この図を見ると、分布しているマーカーから出ている扇形には、大きく分けて 5 つのパターンがあることが分かる。5 つのパターンとは、黄色の扇形、黄緑色の扇形、青緑色の扇形、水色の扇形、青色・紫色の扇形である。それぞれのパターンの例を図 6.3 に示す。

図 6.4 は北海道から東北にかけて、図 6.5 は関東一帯、図 6.6 は西日本一帯の地図ビューの表示をそれぞれ示したものである。これらの図を見ると、図 6.3 に示された 5 つのパターンは、それぞれが地理的な分布において異なる特徴を示していることが見て取れる。黄色の扇形は、北海道から本州、四国、九州の太平洋側や離島まで全国的に広く分布しており、最も集中している地域は関東である。黄緑色の扇形は関東には少なく、北海道、岩手県、宮城県や東海以西の太平洋側に多く分布しており、最も集中している地域は宮城県付近である。青緑色の扇形は、黄色や黄緑に比べて数が大きく減少し、北海道、岩手県、宮城県、福島県、大阪府、高知県に見られる。水色の扇形は最も数が少なく、北海道と岩手県に見られる。青色・紫色の扇形は 1ヶ所に集中しており、茨城県に見られる。

ここで再びトピックビューを見ると、それぞれの扇形のパターンに対応する時刻が分かる。

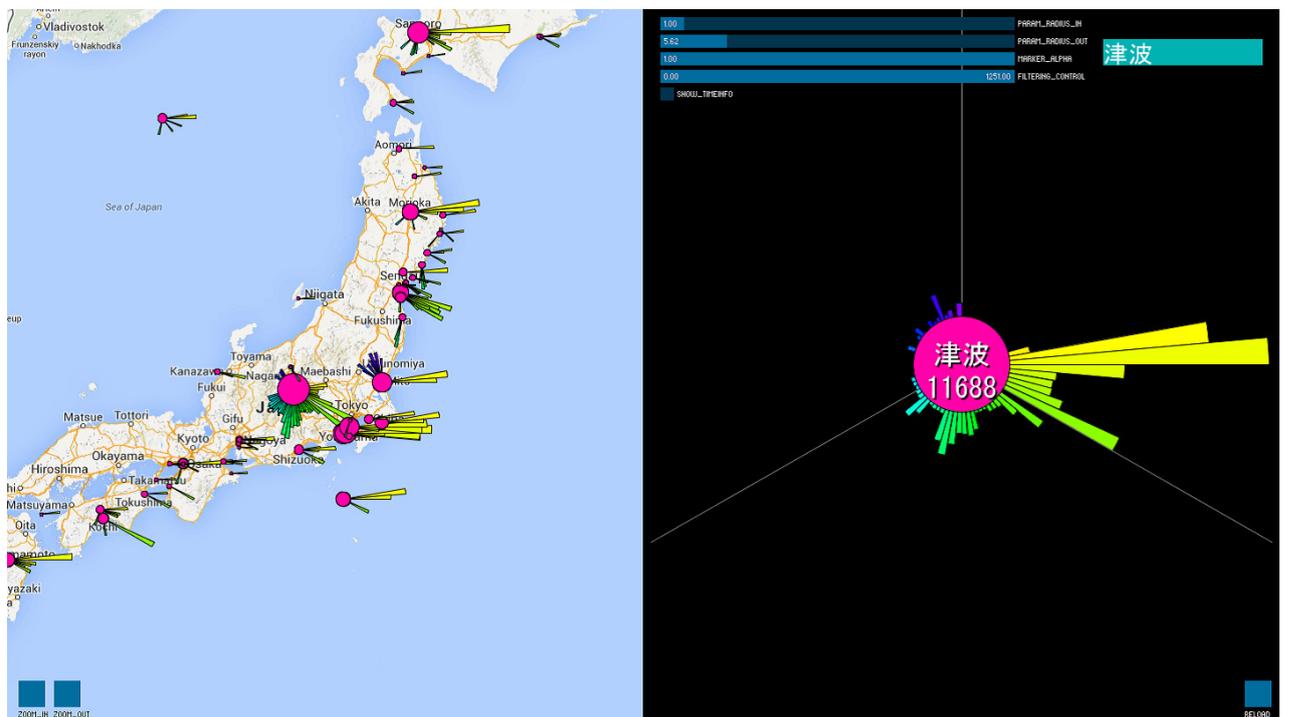


図 6.1: トピック「津波」を表示したツール画面

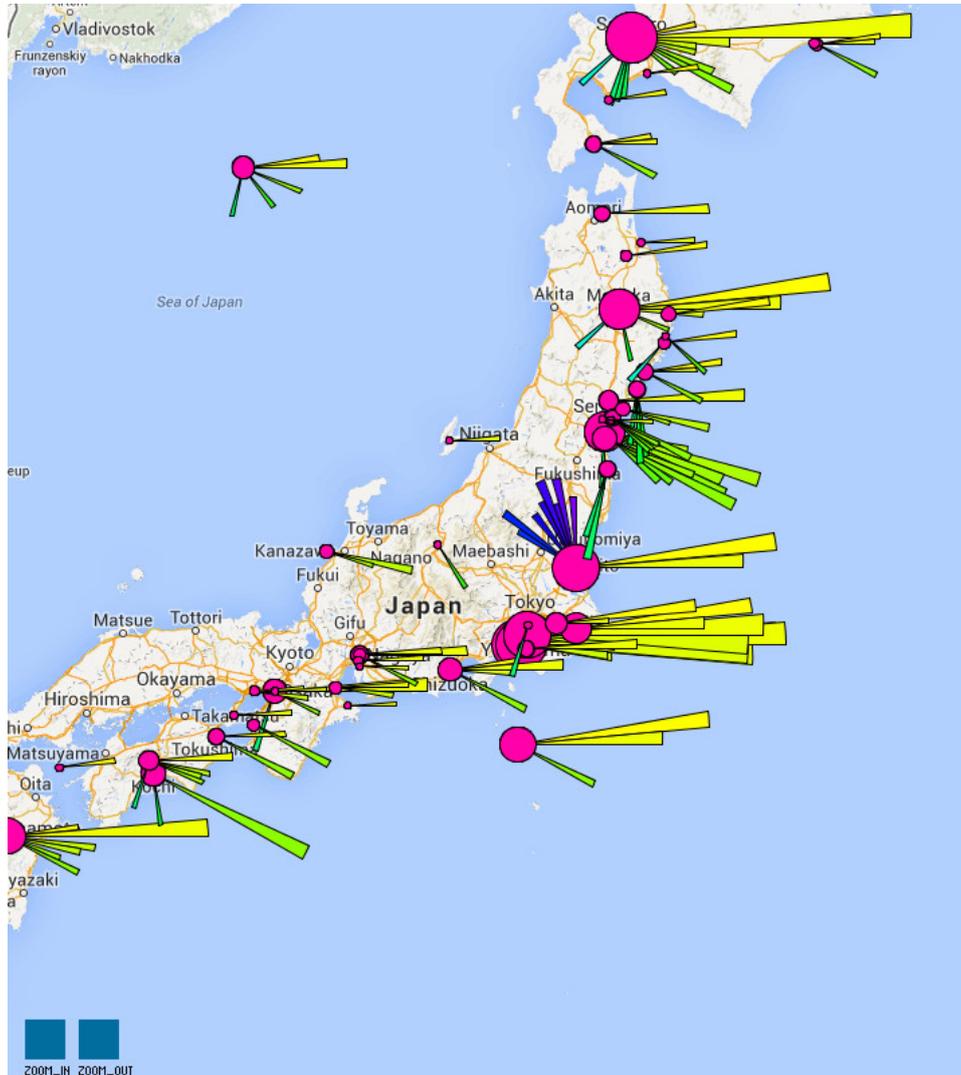


図 6.2: 図 6.1 からパラメータを調整した状態の地図ビュー

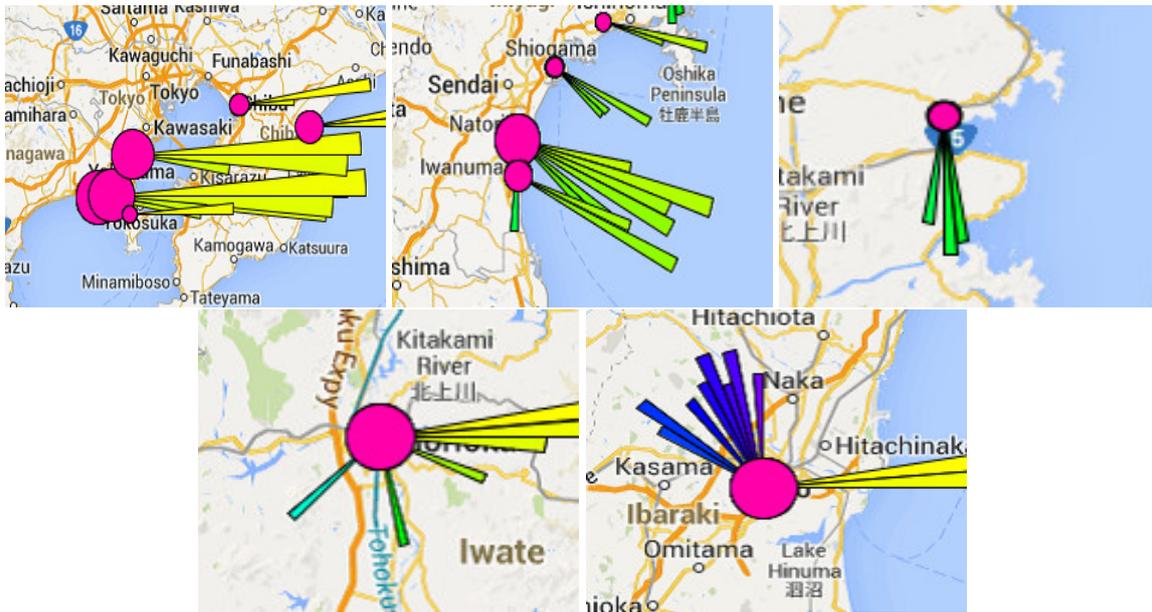


図 6.3: トピック「津波」に見られる扇形のパターン例: 黄色、黄緑色、緑色、水色（マーカーの左下部分）、青色・紫色（マーカーの左上部分）の5パターンが確認できる

図 6.7 は、各パターンが対応する時刻の範囲を示したものである。黄色の扇形が対応する時刻は、2011 年 3 月 11 日の 15 時台から 19 時台にかけてである。黄緑色の扇形が対応する時刻は、2011 年 3 月 11 日の 20 時台から 12 日の 2 時台にかけてである。青緑色の扇形が対応する時刻は、2011 年 3 月 12 日の 9 時台から 15 時台にかけてである。水色の扇形が対応する時刻は、2011 年 3 月 12 日の 20 時台から 21 時台にかけてである。青色・紫色の扇形が対応する時刻は、2011 年 3 月 13 日の 9 時台から 23 時台にかけてである。

6.3 考察

全体の合計頻度で見ると、3 月 11 日の 16 時台から 17 時台にかけて、頻度が飛び抜けて大きいことが分かった。この時間帯は、2011 年 3 月 11 日 14 時 46 分の東北地方太平洋沖地震発生から一、二時間ほど経過した時間帯である。また、地図ビューを見ると、北海道から本州、四国、九州にかけての都道府県庁所在地の他、太平洋沿いの沿岸地域や、伊豆諸島などの離島でもこの時間帯の頻度が高いことが分かった。次に合計頻度が高いのは、3 月 11 日の 23 時台である。16 時台から 17 時台では最も頻度の高かった関東地方や、岩手県や青森県の沿岸地域では、一転してこの時間帯の頻度が低くなり、一方で北海道や西日本の沿岸地域では、この時間帯の頻度が高くなっていることが分かる。ただし、宮城県では 16 時台から 17 時台よりも、この時間帯の方が頻度が高くなっている。

これらの事から、震災発生直後において人々は、まず太平洋側の津波を中心とした全国の

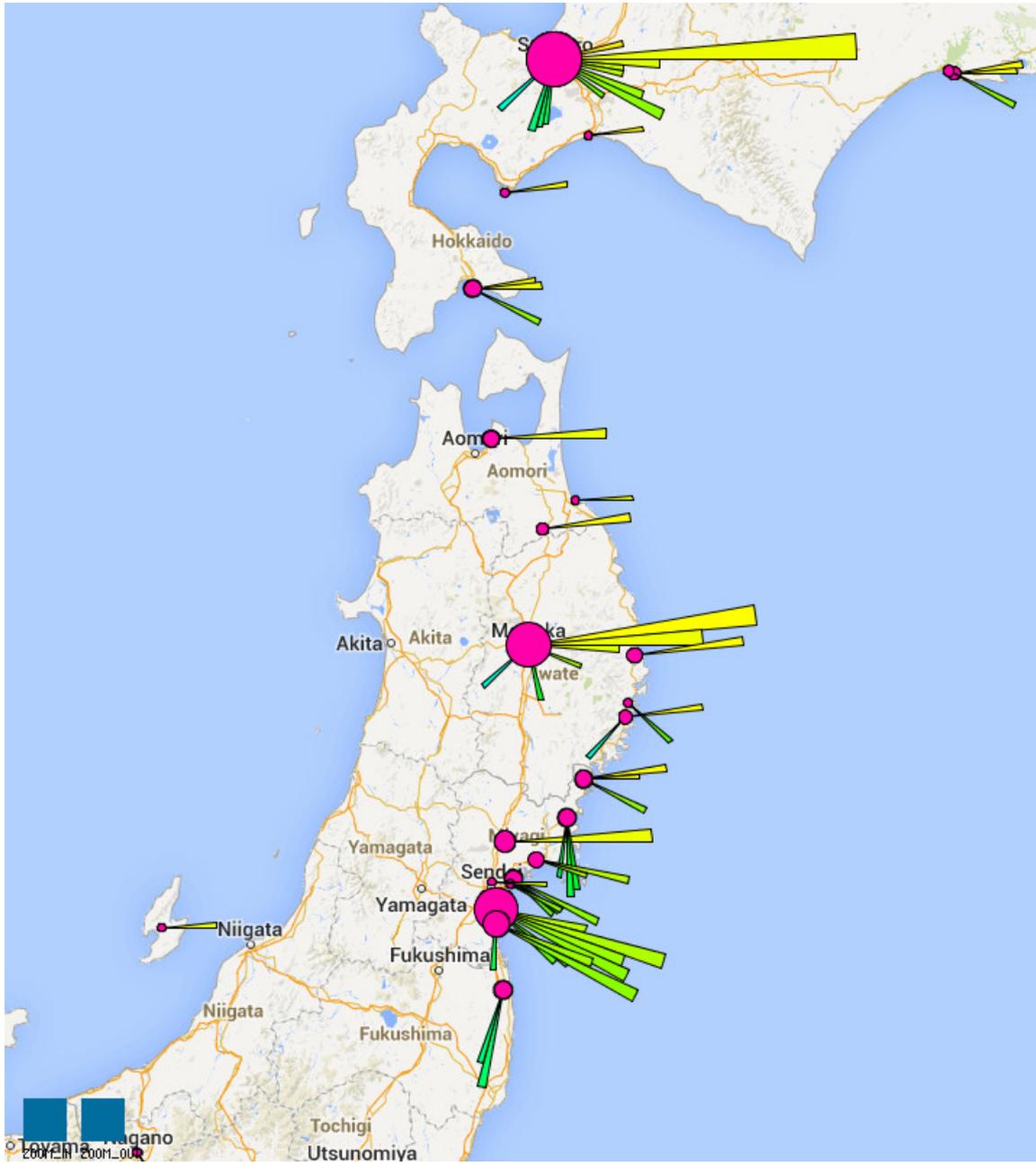


図 6.4: 北海道から東北にかけての地図ビューの表示

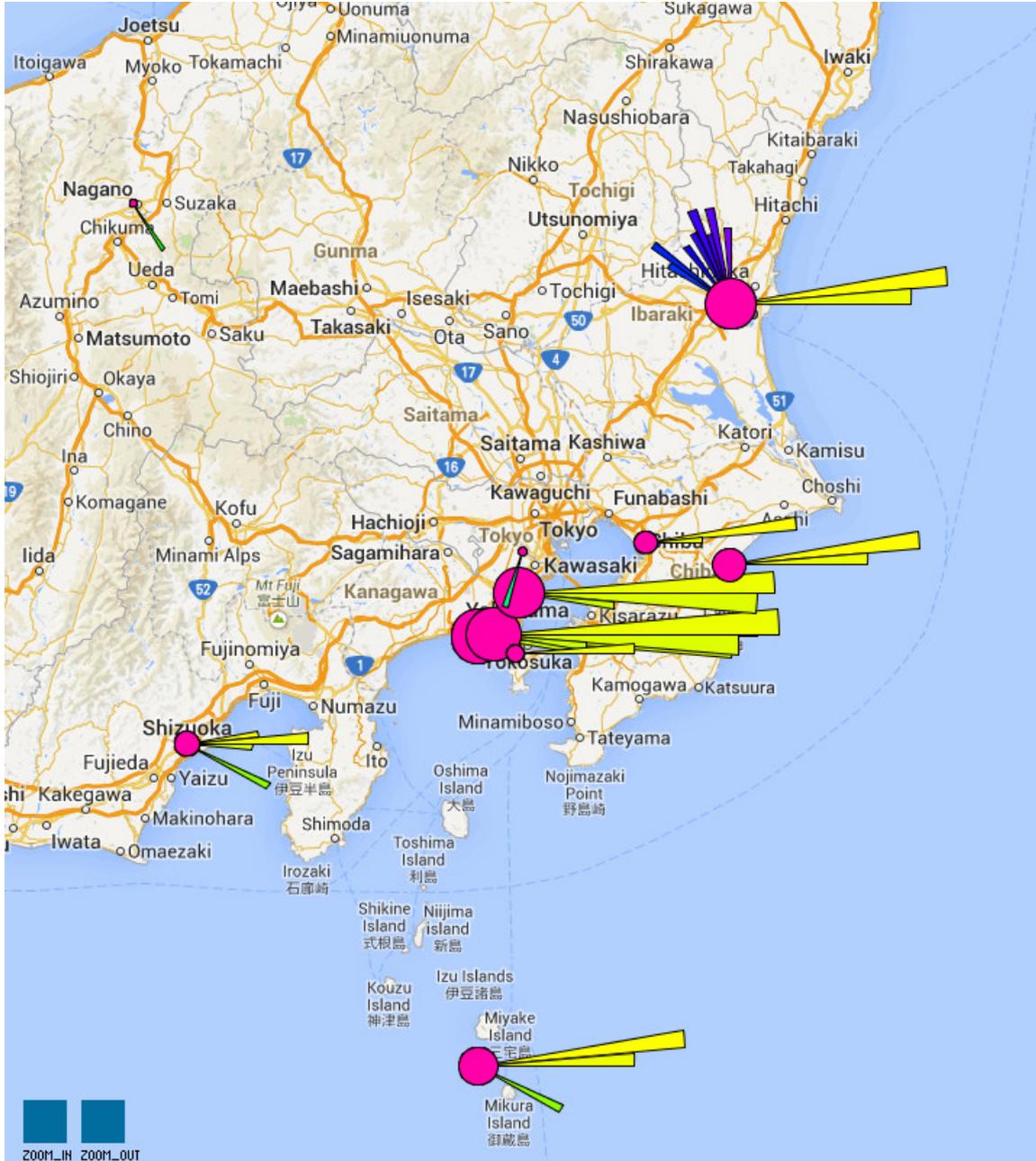


図 6.5: 関東一帯の地図ビューの表示

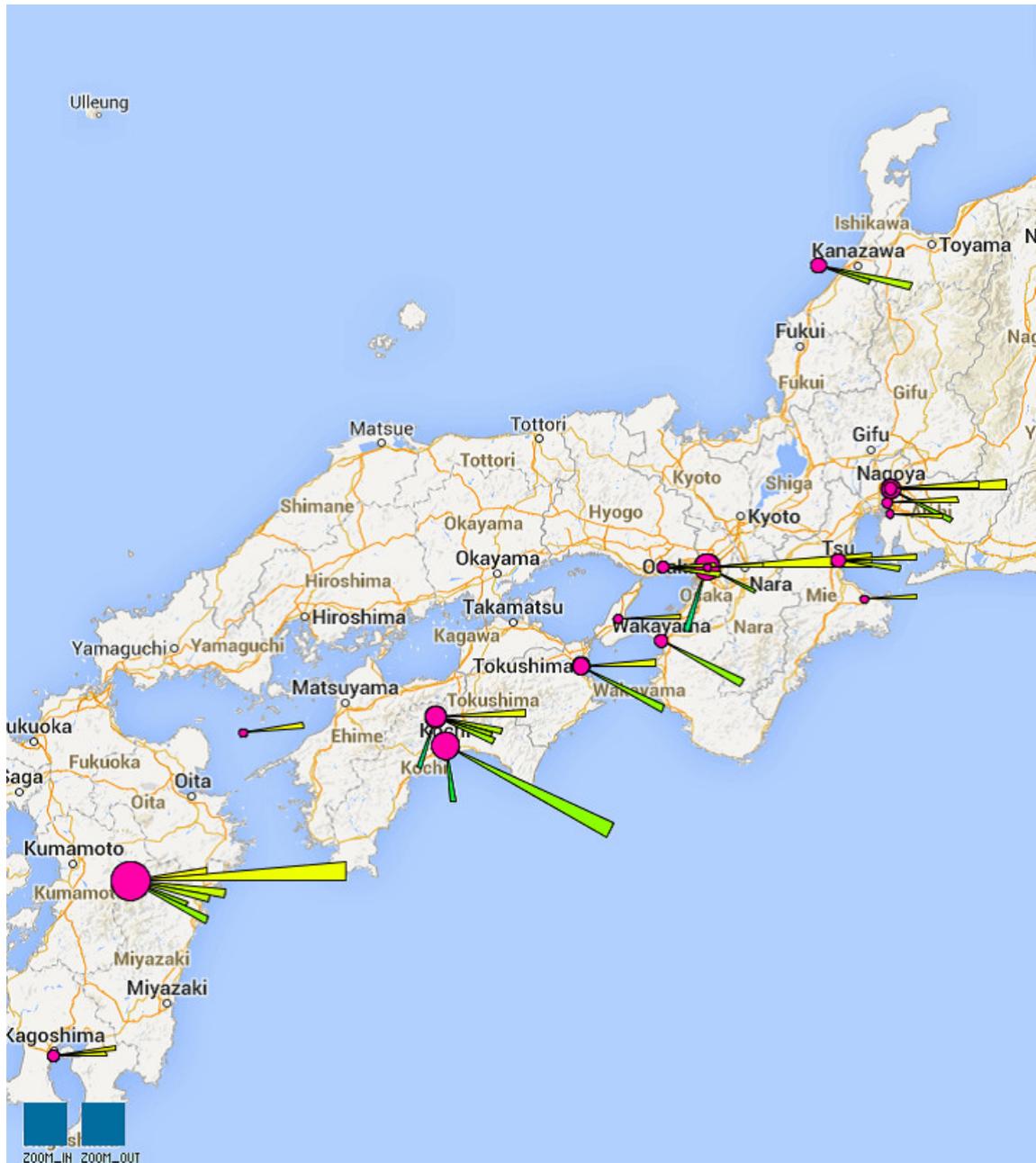


図 6.6: 西日本一帯の地図ビューの表示

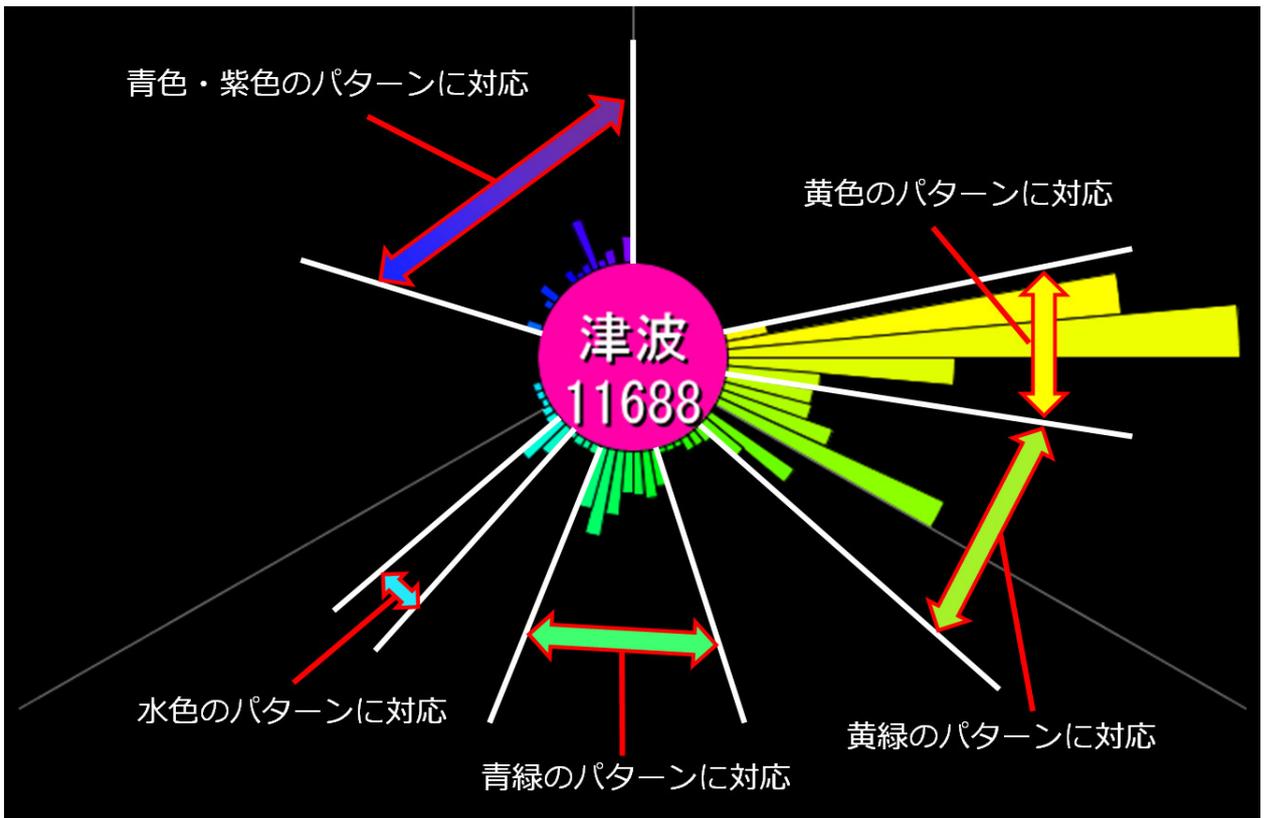


図 6.7: 図 6.3 の各パターンに対応する扇形の部分

津波に興味を持ち、時間が経つに連れて宮城県の津波へと対象がフォーカスしていったと考えられる。震源から離れた西日本や北海道の頻度が高くなったことについては、津波の襲来が東北や関東よりも遅かったために、興味のピークも遅れて訪れたと考えられる。このように、人々の興味が広い地域から狭い地域へと集まっていく様子が分かった。

さらに、3月13日の9時台から23時台にかけては、頻度の高い位置の分布に他とは異なる特徴が見られる。この時間帯の頻度が高い地域は茨城県一箇所のみであり、他の地域はほとんど頻度が0となっている。

これは、これまで話題となっていなかった茨城県の津波に対して、地震発生から二日経って人々が興味を持ったという事であると考えられる。二日間のタイムラグが発生した理由としては、人々の興味が宮城県の津波に集中していたために、同じ時間に津波が襲来したはずの茨城県については見逃されていたという可能性が考えられる。このように、トピックによっては人々の興味の「死角」となる地域が存在することが分かった。

第7章 議論

第6章で紹介したユースケースを踏まえ、本手法の有用性についての議論を行う。

7.1 時間と頻度の表現方法について

本研究では、円周上に時間軸を割り当て、円の外側に頻度に応じた大きさの扇形を描画する「マーカー」によって、頻度とその時間変化を表現した。この「マーカー」による表現は、属性の値に応じて図形の大きさを変化させるという点で、Graduated Symbol Map との相性がよく、適用が容易である。

また、この表現では時間の基準となる扇形の向きが全てのマーカーで統一されている。それに加えて時刻に色相を割り当てたことにより、地図上においてマーカーの縦位置や横位置が揃っていなくても、頻度の比較を行うことが可能となった。この点も、画面上の縦位置や横位置を揃えることが難しい地図上での可視化において、有効に作用していると考えられる。

ただし、この表現は従来の Graduated Symbol Map よりも、個々の位置における表現の領域が広がっているため、大きな頻度を表す場合にマーカー同士が重なる可能性を高めてしまっている。扇形の大きさを調節するパラメータを設け、ユーザー側が操作できるようにするといった工夫によって、この問題の解決を試みている。

7.2 Small Multiples との比較

Small Multiples では、複数の Graduated Symbol Map を並べて表示することで、頻度や位置の時間変化を表現する。多くの時刻を一度に表そうとすると大きな画面領域が必要となるために、Small Multiples で時系列全体の俯瞰を行うことは困難である。本手法では、一つの地図によって時系列全体を表現することが可能であるため、多くの時刻を表すのに大きな画面領域を必要としない。したがって、時系列や位置を広範囲で俯瞰するような場合において、本手法は Small Multiples よりも有効であると言える。

一方で、それぞれの時刻について厳密に頻度や位置を調べたい場合には、一つの地図で一つの時刻を提示する Small Multiples の方が有効である事もある。本手法は、データの時系列や位置を広範囲で俯瞰できるようにしており、頻度や位置の分布が特徴的な部分を発見することができる。個々の時刻や位置についてのより詳細な分析に繋げていくための発見をする手段として、本手法を利用することができると考えられる。

7.3 今後の課題

7.3.1 複数のトピックの表現

本手法では一つの地図上で時系列全体を表現し、広範囲の位置を同時に表すことが可能となっている。しかしながら、トピックについては現状、一度に一つのトピックしか表すことができていない。解決策としては、表現のもととなった Sun Burst Chart や Nightingale の手法 [32] のように、扇形を階層状にして複数のトピックについての頻度を同時に表現する方法があると考えられる。他にも、Small Multiples を利用することで、異なる数個のトピックについての可視化結果を提示することは可能である。

7.3.2 マーカーの密集による視覚的混雑への対応

データの特徴や規模によっては、地図上の一部分にマーカーが密集し、視覚的な分析が困難になることが起こると考えられる。本手法では、ユーザーによる手動でのパラメータ操作や、フィルタリングという手段によってこの問題の解決を図っている。しかし、この手段でもマーカーの密集による視覚的混雑を完全に解消することは難しい。より効果的に視覚的混雑を解消するためには、描画されるマーカーの近くに別のマーカーが存在する場合に、双方のマーカーを一つに統合して合計の頻度を提示する、などの描画手法を導入していく必要がある。

7.3.3 負の値の表現

本手法では、扇形によって表現する頻度の値は0以上であるものとしている。これが0以下の値であっても表現できるようになれば、特定の時刻における頻度などを基準に指定して、差分情報の提示をすることなどが可能になり、表現の幅を広げることができる。考えられる手法としては、マーカーの中心円の内側に向けて扇形を描画する、正の値と負の値で異なる色を割り当てる、などが挙げられる。

第8章 まとめ

本研究では、位置付き時系列トピックデータの分析を支援するために、一つの地図上でトピックの頻度や位置の時間変化を把握可能な表現手法である **Sun Burst Map** を開発した。本手法は、円周上に時間軸を割り当て、対応する時刻における頻度を面積に割り当てた扇形を描画することによって、頻度の時間変化を表現している。また、これらのマーカーを位置に対応する地図上の地点に配置することで、頻度の時間変化を位置ごとに比較することを可能にしている。

Sun Burst Map を利用して位置付き時系列トピックデータの分析を行うための、分析ツールを開発した。分析作業を支援するために、トピックの選択、マーカーの各種パラメータの操作、頻度によるフィルタリングなどの機能が実装されている。震災下における Twitter のツイートをもとに作成された位置付き時系列トピックデータを用いて、本手法のユースケースを提示した。ユースケースの提示により、本手法が位置付き時系列トピックデータの分析支援において、有用であることを示した。

Twitter のつぶやきデータ以外にも、位置付き時系列トピックデータとして表すことが可能なデータは様々なものが存在する。そのようなデータの分析に本手法を利用することにより、従来よりも効率的な分析を行っていくことが可能であろうと考えられる。

謝辞

本研究を行うにあたり、指導教員である三末和男准教授には、丁寧なご指導と助言を頂きました。また、田中二郎教授、志築文太郎准教授、高橋伸准教授には、ゼミを通して多くの貴重なご意見を頂きました。心より感謝を申し上げます。

インタラクティブプログラミング研究室の皆様には、ゼミでの発表や研究室の日常生活の中で、様々なご意見を頂きました。特に NAIS チームの皆様には、ゼミでの発表以外にも研究の相談を聞いて頂くなど、大変お世話になりました。大学院博士前期課程の二年間研究を続けることができたのは、皆様のおかげです。深く感謝しております。

本研究のユースケースで使用した震災下のツイートデータは、株式会社富士通研究所から提供して頂きました。また、研究を進めるにあたって、富士通研究所の皆様と議論をして頂き、貴重なご意見を頂くことができました。本当にありがとうございます。

最後に、学生生活を送るにあたって多大な援助を頂いた家族や、学生生活を実りあるものとしてくれた友人をはじめ、大学でお世話になった全ての皆様に、心より感謝を申し上げます。本当にありがとうございました。

参考文献

- [1] Martin Nöllenburg. Geographic Visualization. In *Human-Centered Visualization Environments*, pp. 257–294. Springer Berlin Heidelberg, 2007.
- [2] James John Flannery. *The graduated circle: A description, analysis, and evaluation of a quantitative map symbol*. University of Wisconsin–Madison, 1956.
- [3] Michael Friendly and Daniel J Denis. Milestones in the history of thematic cartography, statistical graphics, and data visualization, 2011. Retrieved 2001, 5.11.
- [4] Heat map. http://en.wikipedia.org/wiki/Heat_map. (Accessed on December 23, 2013).
- [5] Edward R Tufte and PR Graves-Morris. *The visual display of quantitative information*, Vol. 2. Graphics press Cheshire, CT, 1983.
- [6] Sidharth Thakur and Andrew J. Hanson. A 3D visualization of Multiple Time Series on Maps. In *Proceedings of 14th International Conference Information Visualization (IV2010)*, pp. 336–343. IEEE, 2010.
- [7] John Stasko and Eugene Zhang. Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. In *Proceedings of IEEE Symposium on Information Visualization (InfoVis2000)*, pp. 57–65. IEEE, 2000.
- [8] Christian Tominski, Jonathan E Donges, and Thomas nocke. Information Visualization in Climate Research. In *Proceedings of 15th International Conference Information Visualisation (IV2011)*, pp. 298–305. IEEE, 2011.
- [9] Dinh-Quyen Nguyen and Heidrun Schumann. Taggram: Exploring Geo-data on Maps through a Tag Cloud-Based Visualization. In *Proceedings of 14th International Conference Information Visualisation (IV2010)*, pp. 322–328. IEEE, 2010.
- [10] Stanley Milgram. Psychological maps of Paris. *Environmental psychology: People and their physical settings*, pp. 104–124, 1976.
- [11] Shane Ahern, Mor Naaman, Rahul Nair, and Jeannie Hui-I Yang. World Explorer: Visualizing Aggregate Data from Unstructured Text in Geo-Referenced Collections. In *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, pp. 1–10. ACM, 2007.

- [12] Typographic Maps. <http://store.axismaps.co.uk/>. (Accessed on January 9, 2014).
- [13] Shehzad Afzal, Ross Maciejewski, Yun Jang, Niklas Elmqvist, and David S Ebert. Spatial Text Visualization Using Automatic Typographic Maps. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, pp. 2556–2564, 2012.
- [14] Bongshin Lee, Nathalie Henry Riche, Amy K Karlson, and Sheelagh Carpendale. Spark-Clouds: Visualizing Trends in Tag Clouds. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 6, pp. 1182–1189, 2010.
- [15] Susan Havre, Elizabeth Hetzler, Paul Whitney, and Lucy Nowell. ThemeRiver: Visualizing Thematic Changes in Large Document Collections. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 8, No. 1, pp. 9–20, 2002.
- [16] Furu Wei, Shixia Liu, Yangqiu Song, Shimei Pan, Michelle X Zhou, Weihong Qian, Lei Shi, Li Tan, and Qiang Zhang. TIARA: A Visual Exploratory Text Analytic System. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 153–162. ACM, 2010.
- [17] Weiwei Cui, Shixia Liu, Li Tan, Conglei Shi, Yangqiu Song, Zekai Gao, Huamin Qu, and Xin Tong. TextFlow: Towards Better Understanding of Evolving Topics in Text. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 12, pp. 2412–2421, 2011.
- [18] Satoko Shiroi, Kazuo Misue, and Jiro Tanaka. ChronoView: Visualization Technique for Many Temporal Data. In *Proceedings of 16th International Conference on Information Visualisation (IV2012)*, pp. 112–117. IEEE, 2012.
- [19] John V Carlis and Joseph A Konstan. Interactive visualization of serial periodic data. In *Proceedings of the 11th annual ACM symposium on User interface software and technology*, pp. 29–38. ACM, 1998.
- [20] Pierre Dragicevic and Stéphane Huot. SpiraClock: a continuous and non-intrusive display for upcoming events. In *Proceedings of CHI'02 extended abstracts on Human factors in computing systems*, pp. 604–605. ACM, 2002.
- [21] Alan MacEachren, D Xiping, Frank Hardisty, Diansheng Guo, and Gene Lengerich. Exploring High-D Spaces with Multiform Matrices and Small Multiples. In *Proceedings of IEEE Symposium on Information Visualization (InfoVis2003)*, pp. 31–38. IEEE, 2003.
- [22] Roeland Scheepens, Niels Willems, Huub van de Wetering, and Jarke J van Wijk. Interactive Visualization of Multivariate Trajectory Data with Density Maps. In *Proceedings of IEEE Pacific Visualization Symposium (PacificVis2011)*, pp. 147–154. IEEE, 2011.

- [23] Christian Tominski, Heidrun Schumann, Gennady Andrienko, and Natalia Andrienko. Stacking-Based Visualization of Trajectory Attribute Data. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, pp. 2565–2574, 2012.
- [24] Adam Marcus, Michael S. Bernstein, Osama Badar, David R. Karger, Samuel Madden, and Robert C. Miller. TwitInfo: Aggregating and Visualizing Microblogs for Event Exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 227–236. ACM, 2011.
- [25] Nan Cao, Yu-Ru Lin, Xiaohua Sun, David Lazer, Shixia Liu, and Huamin Qu. Whisper: Tracing the Spatiotemporal Process of Information Diffusion in Real Time. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, pp. 2649–2658, 2012.
- [26] Nivan Ferreira, Lauro Lins, Daniel Fink, Steve Kelling, Christopher Wood, Juliana Freire, and Claudio Silva. BirdVis: Visualizing and Understanding Bird Populations. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 12, pp. 2374–2383, 2011.
- [27] Wenwen Dou, Xiaoyu Wang, Drew Skau, William Ribarsky, and Michelle X. Zhou. LeadLine: Interactive Visual Analysis of Text Data through Event Identification and Exploration. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology (VAST2012)*, pp. 93–102. IEEE, 2012.
- [28] Patrik Lundblad, Jonas Thoursie, and Mikael Jern. Swedish Road Weather Visualization. In *Proceedings of 14th International Conference Information Visualisation (IV2010)*, pp. 313–321. IEEE, 2010.
- [29] Alfred Inselberg and Bernard Dimsdale. Parallel coordinates. In *Human-Machine Interactive Systems*, pp. 199–233. Springer, 1991.
- [30] James Walker, Zhao Geng, MW Jones, and Robert S Laramee. Visualization of Large, Time-Dependent, Abstract Data with Integrated Spherical and Parallel Coordinates. *Eurographics Conference on Visualization (EuroVis2012)–Short Papers*, pp. 43–47, 2012.
- [31] Tetsuro Takahashi, Shuya Abe, and Nobuyuki Igata. Can Twitter Be an Alternative of Real-World Sensors? In *Human-Computer Interaction. Towards Mobile and Intelligent Interaction Environments*, pp. 240–249. Springer, 2011.
- [32] Florence Nightingale. Diagram of the causes of mortality in the army in the east. *Reproduced at <http://www.florence-nightingale-avenging-angel.co.uk/Coxcomb.htm>*, 1858.