Twitterにおける群衆の経験に基づく近接地域検索システム

若宮　翔子† 　李　　龍††,† 　角谷　和俊†

† 兵庫県立大学大学院 環境人間学研究科　〒 670–0092 兵庫県姫路市新在家本町 1–1–12
†† 独立行政法人 情報通信研究機構 NICT　〒 619–0289 京都府相楽郡精華町光台 3–5
E-mail: †ne11n002@stshse.u-hyogo.ac.jp, ††lee.ryong@gmail.com, †††sumiya@shse.u-hyogo.ac.jp

あらまし　都市空間における地域間の近接性は人々の生活における意思決定の際の重要な判断基準であり，地理的な距離だけでなく，移動にかかる時間や親近性などに影響される．実際に，今日の都市では交通ネットワークの複雑化や人々の活動領域の拡大などの影響により，従来の検索システムだけでは場所間の近接関係を把握することは困難である．この問題に対して都市空間で生活する人々の経験を借りて地域間の近接性を測定することができれば，なじみのない都市におけるユーザの行動支援などのために活用できると期待される．本研究では，ソーシャルネットワークに投稿されているライフログから取得可能な実空間における人々の経験を活用して，都市空間における人々の生活を中心とした地域間の近接性を測定し，近接地域検索に応用する手法を提案する．具体的には，Twitter のジオタグ付きツイートを用いて群衆の移動経験を地域間の移動距離，移動時間，移動量の観点で抽出し，ユーザ操作によってカスタマイズされた三つの尺度を統合して測定される地域間の近接性に基づき，ユーザが指定した地域の近接地域を検索して提示するシステムを提案する．

キーワード　地域近接性，Twitter，距離認知，群衆移動，位置ベース SNS，時空間データマイニング

Geo-social Neighborhood Area Search based on Urban District Proximity using Crowd Experience over Twitter

Shoko WAKAMIYA†, Ryong LEE††,†, and Kazutoshi SUMIYA†

† Graduate School of Human Science and Environment, University of Hyogo
1–1–12 Shinzaike-honcho, Himeji, Hyogo, 670–0092 Japan
†† National Institute of Information and Communications Technology
3–5 Hikaridai, Seika-cho, Sagara-gun, Kyoto, 619–0289 Japan
E-mail: †ne11n002@stshse.u-hyogo.ac.jp, ††lee.ryong@gmail.com, †††sumiya@shse.u-hyogo.ac.jp

1. はじめに

今日の都市における交通ネットワークの複雑化や人々の活動領域の拡大により，地域間を移動する際に，実際の距離とは異なる歪みを感じることがよくある．実際に，地理的には離れている地域であっても，交通アクセスが便利で移動時間がかからないためにより近くに感じることもあれば，地理的には近い地域であるにも関わらず，交通機関の乗り継ぎの悪さなどの原因で移動時間がかかるためにより遠くに感じることがある．また，なじみのある地域やある都市における中心的な役割を持つ地域をより近くに感じたり，あまり訪れることのない地域をより遠くに感じたりすることがある．このように，都市空間における近接性は，地理的な距離だけではなく，その場所間の移動にかかる時間や場所間の親近性といった複数の観点が関わることで歪んでおり，簡単に把握することが困難である．

一般的に，このような場所間の近接性を把握するためには，地図を見たり，地域検索システムを利用して調べたり，その都市に詳しい人々に話を聞いたりすることができる．しかしながら，単に地域間の地理的な位置関係を示している地図や，地域間の最短経路や最短移動時間を提示する従来の検索システムだけでは，人々が感じる都市空間の近接性を把握するには十分ではない．また，人々に距離感覚を尋ねたとしても，その友人がその都市全体を十分に知っているとは限らず，個々の限られた経験上での距離感覚には偏りが生じると考えられるため，この方法も都市空間の近接性を把握するには十分ではないといえる．

一方，近年普及しているソーシャルネットワーク（以降，SNS と記す）上には，実空間において人々が経験したことや感じたことなどを記述した膨大な量のライフログが投稿されている．

さらに，スマートフォンの普及や SNS の発達とともに，位置情報が付与されたライフログも大量に投稿されるようになっている．このようなライフログにより，都市空間におけるこれまでにない規模の人々の生活や経験を簡単に観察して集約することができるようになっている．さらに，集約した結果を応用することによを用いることで，都市での様々なイベント [5] や地域性 [6], [14] などを分析することが可能になっている．従って，このような群衆の経験を活用することで，都市空間の近接性を把握することが可能になると考えられる．

実際に，実空間における人々の経験は，その空間について理解したり，評価したりする際の重要な指標として様々な社会調査に役に立つと期待される．例えば，国土交通省の都市交通調査(注1)では，特定の都市空間における人々の日常的な移動パターンをはじめとする様々な活動パターンが調査されており，その分析結果は都市計画や都市開発などの場面で効果的に活用されている．しかし，このような社会調査を大規模に実施し，質問票を用いたアンケートなどを通して市民から収集したデータを分析し，都市空間における行動パターンを抽出するためには，多大な金銭的なコスト，時間，そして労力がかかる．例えば，都市交通調査の一つであるパーソントリップ OD (Origin-Destination) 量調査は，10 年に一度の頻度でしか実施されていない．最近の調査は平成 22 年の 10 月から 11 月にかけて実施されたが，その速報集計結果(注2)が発表されたのは調査実施から 2 年後の平成 24 年 12 月であった．さらに，詳細な調査集計結果については，平成 25 年 3 月の時点でも現在集計作業中となっている．このように，従来の社会調査で得られるデータでは，多様な人々の経験を大規模かつリアルタイムに把握することが難しく，地域間の近接性分析には不十分であると思われる．

本研究では，SNS から不特定多数の人々のライフログを取得して都市における人々の経験を抽出し，都市内の地域間の近接性を測定する．それを近接地域検索のために応用したシステムを提案し，ユーザ操作によりカスタマイズした近接性に基づく近接地域検索の実現を目指す．具体的には，Twitter(注3)から大量の位置情報付きライフログ（以降，ジオタグ付きツイートと記す）を取得し，群衆の移動経験を市区町村レベルの行政区域を単位とした地域間の移動距離，移動時間，移動量の観点で抽出し，三つの尺度を合わせた地域間の近接性を測定する．このとき，ユーザ操作により三つの尺度に優先度を付与し，ユーザの目的に合わせてカスタマイズした近接性に基づくインデクスマップを作成し，このインデクスマップ上でユーザが指定した地域の近接地域を検索する．

本論文の構成は以下のとおりである．2 章では，本研究のアプローチとして SNS における群衆の経験を活用した近接地域検索システムについて概観し，関連研究を紹介する．3 章では Twitter からの群衆の移動経験に基づく近接地域検索手法を説明する．4 章では，実データを用いて行った実験とその結果について述べる．最後に，5 章で本研究のまとめと今後の課題と展開について述べる．

(注1)：国土交通省 都市交通調査：
http://www.mlit.go.jp/toshi/tosiko/index.html
(注2)：第 5 回近畿圏パーソントリップ調査速報集計結果：
http://www.kkr.mlit.go.jp/plan/pt/research_pt/h22/pt_h22.pdf
(注3)：Twitter: https://twitter.com/

2. 群衆の経験を活用した近接地域検索システム

本章では，SNS に投稿されている大量のライフログを活用して，大規模かつ動的に取得可能な都市における人々の経験を活用することの意義について述べる．また，群衆の経験を中心として近接地域を検索することができるシステムの構成について説明する．そして，関連研究を紹介して提案システムの特徴を述べる．

2.1 SNS における群衆の経験の活用

本研究では，SNS から取得した大量のライフログを集約し，都市における群衆の経験として再利用することを目的とする．実際に，SNS に発信される個々のメッセージは，一般的にコミュニケーションツールとして利用されているが，その後はただ単に過去のコンテンツとしてどんどん蓄積されていくだけである．しかしながら，多くの人々が日々発信しているライフログを集め，再利用のためにまとめることにより，実空間における人々の経験を大規模かつリアルタイムに把握することが可能になる．本稿では，このような群衆の経験を活用して，都市内の地域間の近接性を測定し，ユーザが指定した地域の近接地域を検索するためのシステムを提案する．

従来の地域検索システムを用いて近接地域を検索する場合，ある地点から何キロメートル以内，何分以内といった具体的な条件を指定する．しかしながら，特に，なじみのない都市においてこのような具体的な条件の指定は難しいと思われる．さらに，このようなシステムにおいて基準として用いられている距離や時間は，地域間の最短経路に基づく最短経路距離や最短経路時間であるが，実情とは異なる部分があると考えられる．例えば，最短経路の場合，多くの信号があるため，予想以上に移動時間がかかる場合があるかもしれない．また，日中は最短経路を利用する人が多いが，夜は治安が悪くなるため迂回して別の経路を利用する人が多くなるため，時間帯によって地域間の移動距離や移動時間が異なるという場合も考えられる．

これに対し，提案システムを用いることで，実空間における人々の様々な振る舞いも含めた実質的な移動経験に基づき，地域間の近接性を測定することで，近接地域を検索することができる．ここで，提案システムの利用シナリオを一つを紹介する．例えば，大阪駅周辺（大阪府大阪市北区）にある会社に勤めているユーザが，マイホームの購入を検討しているとする．このときの条件として，郊外で，都心へのアクセスがよく，実際にその間を行き来する人が多い場所を中心に調べたいと考えている．この条件を移動の観点で整理すると，会社がある地域を基点として，そこからの移動距離が離れており，移動時間は短く，かつ多くの人々がその間を移動している地域というように，移動空間距離，移動時間，移動量という三つの尺度で表現することができる．この移動経験から抽出することができる三つの尺度は，近接地域検索の際に一般的に用いられる尺度になりうる

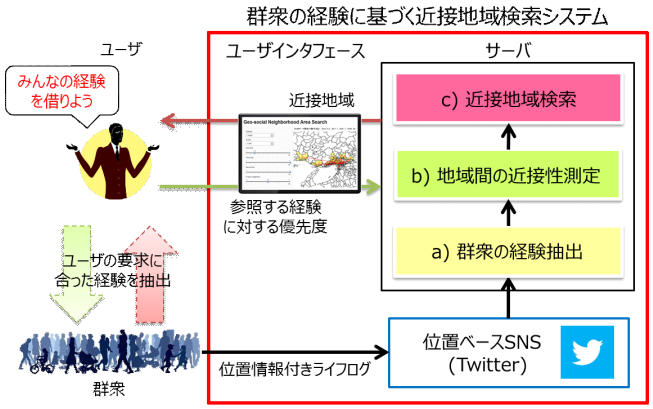


図 1　群衆の移動経験を活用した近接地域検索システムの構成

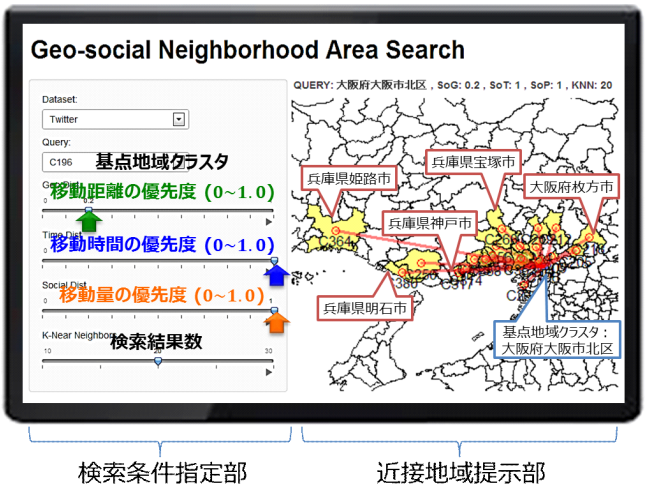


図 2　提案する近接地域検索システムのユーザインタフェース

と考えられる．ただし，ユーザの要求に合わせた近接性表現のために，三つの尺度に対する重みをユーザ操作により直感的に調整できるようにすべきである．この例では，移動空間距離の優先度は低く，移動時間や移動量の優先度は高くなるように指定することで，ユーザの要求に合う近接性を表現することができる．具体的に，図 2 のユーザインタフェースの検索条件指定部において，基点地域クラスタ ($C196$) を選択し，移動空間距離の優先度を 0.2，移動時間と移動量の優先度をそれぞれ最大の 1.0，検索結果数を 20 件と指定することで，今回のユーザの要求に基づく近接性を表現する．これにより，図 1 のサーバでは，ユーザ操作に基づき生成されるクエリに基づき，群衆の経験を利用して近接地域クラスタを検索する．そして，図 2 のユーザインタフェースの近接地域提示部に示すように，検索結果を地図上で表現する．この検索結果を見ると，大阪府内の大阪市以外の市や兵庫県宝塚市，兵庫県神戸市，兵庫県明石市や兵庫県姫路市などのように新幹線や新快速が停まる大きな駅があったり，電車一本で行ける路線があったりとアクセスに便利で，実際に多くの人々が通勤などの目的で移動しているような地域クラスタを発見することができる．このように，SNS を通して取得可能な群衆の経験を活用し，ユーザの目的に合わせた近接基準を用いた近接地域検索システムを用いることにより，従来の最短経路や最短移動時間に基づく検索システムでは抽出することが困難な近接地域を検索して提示することが可能になる．

2.2　近接地域検索システムの構成

本研究では，都市空間における人々の移動経験を用いて，近接地域検索を行うシステムを開発する．このシステムは，図 1 に示すように，ユーザが検索時に利用するユーザインターフェースと，ユーザからの要求を受けて実際に検索処理を行うサーバから構成されている．ユーザインタフェースは，図 2 のように，検索条件指定部と近接地域提示部から構成されている．サーバは，a) 群衆の経験抽出部，b) 地域間の近接性測定部，c) 近接地域検索部から構成されている．

a) 群衆の移動経験抽出部

Twitter に代表される SNS に投稿されている群衆のライフログを取得し，都市における群衆の移動経験を抽出する（図 1 a)）．今回は，Twitter に投稿されているツイートの中でも位置情報付きのツイートを取得して，日常生活に不可欠であり実空間とのつながりが強い群衆の移動経験をモニタリングする．具体的には，群衆の移動経験を移動距離，移動時間，移動量という三つの観点で抽出する．

b) 地域間の近接性測定部

ここでは，a) 群衆の経験抽出部で抽出した群衆の移動経験に基づき，地域間の近接性を求める（図 1 b)）．このとき，必ずしもすべての経験を利用する必要はなく，むしろユーザの意図や目的に合う経験を選択的に用いて近接基準を表現することが重要である．そこで，本稿で提案するシステムでは，群衆の移動経験から抽出する移動空間距離，移動時間，移動量の三つの尺度に対する優先度を，図 2 の検索条件指定部に示すように，ユーザがそれぞれ対応するスライドバーでインタラクティブに指定しながら近接性を測定することができるユーザインタフェースを提案する．それぞれの優先度は，0 から 1.0 の範囲に設定することができ，三つの優先度付き尺度を組み合わせることによって，ユーザの意図や目的に合わせた近接性を測定する．

c) 近接地域検索部

b) 地域間の近接性測定部で求めた近接性に基づき，ユーザが指定した地域の近接地域を検索する（図 1 c)）．図 2 の検索条件指定部では，ユーザが近接地域検索の基点となる地域クラスタを指定するために，地域クラスタの ID リストから一つの地域クラスタを選択できるようにしている．4 章の実験で用いたプロトタイプシステムでは，234 個の地域クラスタの ID リストの中から一つの地域クラスタを選択できるようにしている．また，近接地域をユーザが決めた数だけ検索して表示するために，検索結果数を指定することができるようにしている．また，検索された近接地域の提示には，分かりやすさのために，図 2 の近接地域提示部のような地図を用いる．

2.3　関 連 研 究

スマートフォンの普及と SNS の発達に伴い，多くの人々が位置情報付きのライフログや写真のようなソーシャルメディアコンテンツを発信している．近年，このようなジオタグ付きコンテンツから都市空間における群衆の経験を抽出して活用する研

究が盛んに行われている．Zheng らは，GPS 機能を搭載した携帯端末を通して実空間での人々の移動軌跡とメッセージから予測される行動を分析し，実空間に存在する POI を人々の活動によって関連付けることにより，ある行動に適した場所の推薦を行うための手法を提案した [12]．Leung らは， GPS 端末のデータから取得した移動軌跡を用いてユーザ，行動，場所の関連性を分析することで，ユーザや行動に基づき協調的に位置推薦を行う手法を提案した [7]．Choudhury ら [1] や Kurashima ら [4] は，写真共有サイト Flicker に投稿されている時空間タグ付き写真からランドマークを中心とした群衆の移動履歴を分析し，旅行ルートとしてのランドマーク間の関係性を分析した．岡ら [13] は，ジオタグ付きツイートを用いて，その発信密度が多い地域を地図変形により強調することで，イベントやホットスポットを時系列的に可視化するインタフェースを提案した．Mislove ら [8] は，ツイートのメッセージから人々の感情を抽出し，空間的および時間的な感情の変化を地図変形によって表現するシステムを開発した．我々の過去の研究 [6], [14] では，位置ベースソーシャルネットワークに投稿されているジオタグ付きライフログから人々の行動パターンを抽出し，そのパターンに基づき地域をグルーピングすることで地域の特徴付けを行った．さらに，同様のデータセットを用いて都市空間における群衆移動をモニタリングし，地域間の認知距離とその都市空間における各地域の影響力を測定することにより，人々の生活を中心とした認知地図を作成するための手法も提案している [10], [11]．これらの研究に対し，本稿では，Twitter から取得した群衆のライフログから都市空間における群衆の移動経験から移動距離，移動時間，移動量という三つの尺度を抽出し，ユーザ操作により三つの尺度をカスタマイズして近接基準を求めて活用した近接地域検索システムを提案しているという点で特徴的である．

3. Twitter からの群衆の移動経験に基づく近接地域検索

本章では，図 4 a) のように Twitter から取得した大量の位置情報付きライフログを用いて，市区町村を単位とした地域クラスタ間の群衆の移動経験を移動距離，移動時間，移動量の観点で抽出し，図 4 b) のように移動行列を作成し (3.1 節)，図 4 c), d) のようにユーザ操作によりカスタマイズした地域クラスタ間の近接性を測定してインデクスマップを作成し (3.2 節)，図 4 e) のようにそのインデクスマップ上で近接地域を検索して提示する (3.3 節) までの具体的な手法を説明する．また，提案手法を実装したシステムについて紹介する (3.4 節)．

3.1 Twitter を用いた群衆移動モニタリング

本稿では，都市における群衆の移動経験を大規模かつリアルタイムに抽出するために， ソーシャルネットワークを代表する Twitter に投稿・共有されている位置情報付きのメッセージ（以降，ジオタグ付きツイートとよぶ）を収集する．そのために，Twitter が提供している Streaming API(注4) と我々が過去の研究で開発した地理的ツイート収集システム [5] を用いて，対象地域内で発信されているジオタグ付きツイートを収集する．各ツイートには，ユーザ ID，発信日時，位置情報，テキストメッセージが付与されているため，これらのメタデータを用いることで人々の経験を把握することが可能である．今回は，ユーザ ID，発信日時，位置情報を用いて人々の移動経験を抽出する．具体的には，収集したジオタグ付きツイートをユーザ ID ごとにまとめて発信日時でソートした上で，あるユーザが連続して発信したツイートの位置情報が同一ではない場合に移動があったものとして判定する．ただし，今回は地域クラスタ間の関係性を求めるために，地域クラスタ間の移動をモニタリングする必要がある．そこで，あるユーザが連続して発信したツイートの位置情報がそれぞれ異なる地域クラスタ内にある移動のみを扱う．このようにして抽出されたすべての移動セグメントについて，各移動セグメントの頂点に関わる二つの地域クラスタ間の移動距離，移動時間，移動量を算出する．そして，この結果を集約して地域クラスタ間の群衆の移動データとして移動距離，移動時間，移動量を求める．また，本研究の目的として地域クラスタ間の近接性を測定するために，3 種類の移動データに基づき地域クラスタ間の移動空間距離，移動時間距離，移動量距離の三つの尺度を求め，これらを行列成分とした移動空間距離行列 M_S，移動時間行列 M_T，移動量行列 M_F の三つの移動行列を作成する．以下に，三つの尺度の定義と算出方法を示す．

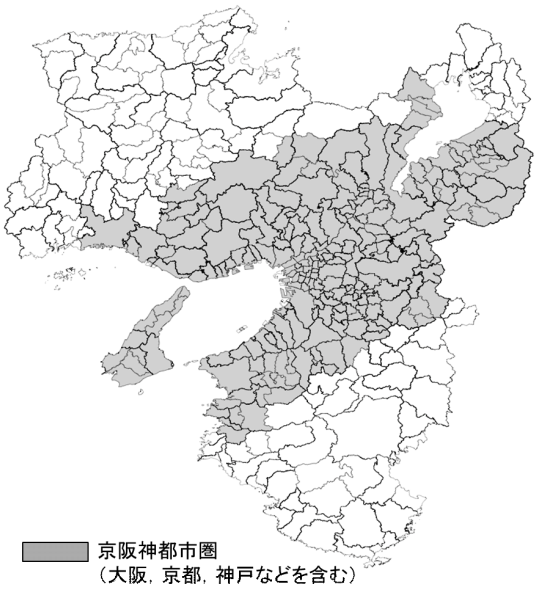


図 3　京阪神都市圏における行政区域に基づく地域

- 移動空間距離 $M_S[c_s, c_e]$

地域クラスタ間の群衆の移動距離に基づき算出される距離とする．ただし，2 地域クラスタ c_s と c_e 間の移動セグメントは複数存在するため，その平均的な移動距離を求める必要がある．今回は，移動に関わるツイートの密度に基づき重心を計算し，その重心間の距離 $|c_s.center - c_e.center|$ を移動空間距離とする．なお，今回用いているジオタグ付きツイートに含まれる位

(注4)：Twitter Streaming API: https://dev.twitter.com/docs/streaming-apis

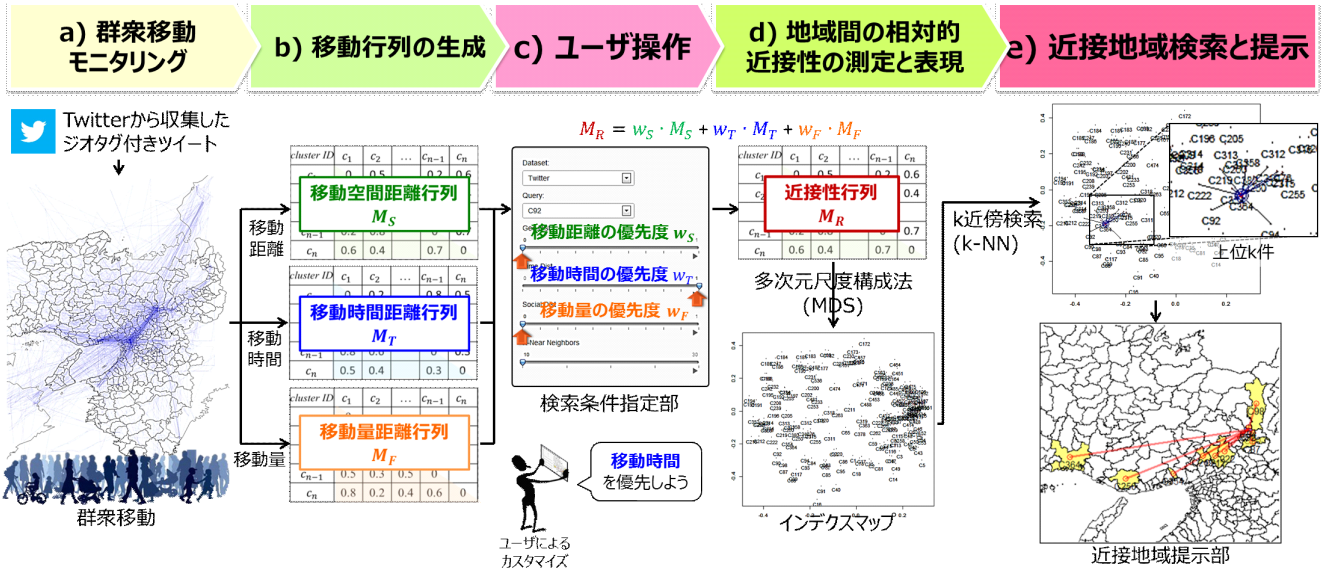


図 4　群衆移動に基づく近接地域検索プロセス

置情報は，正確な地理座標である場合もあれば，テキストでおおまかな地名が記述されている場合もある．後者の場合，正確な地理座標が不明であるため，地名を該当する地域に割り当て，その地域の中心の地理座標を与えた．

- 移動時間距離 $M_T[c_s, c_e]$

地域クラスタ間の群衆の移動時間に基づき算出される距離とする．移動空間距離の場合と同様に，2 地域クラスタ c_s と c_e 間の移動セグメントは複数存在するため，その平均的な移動時間を求める必要がある．今回は，c_s から c_e への平均移動時間を算出し，移動時間距離とする．なお，移動時間にばらつきがあるため，その影響を抑えるために，平均移動時間に対数 log をかけた値を用いる．さらに，移動時間が長すぎるものや短すぎるものはノイズである可能性が高い．今回は日常的な生活に伴う移動を扱うということで，経験的に最長移動時間を 24 時間，最短移動時間を 15 分として，それ以外のものはノイズとした．

- 移動量距離 $M_F[c_s, c_e]$

地域クラスタ間の群衆の移動セグメントの物量に基づき算出される距離とする．具体的には，地域クラスタ c_s から c_e への移動セグメントの総数を用いて移動量距離を算出する．具体的には，移動量が多いほど 2 地域クラスタ間の親近性は強くなるという仮説 [9] に基づき，移動セグメントの総数の逆数を求める．

なお，4 章の実験で用いたプロトタイプシステムでは，より高速なシステム応答のために，あらかじめ三つの移動行列を生成しておき，その後はその行列を再利用している．この移動行列の生成は，Twitter から収集した膨大なジオタグ付きツイートを集約して行うため，空間データの管理や検索を容易に行うことができるデータベースの一つである MySQL を用いて行っている．

3.2　地域クラスタ間の近接性測定

次に，地域クラスタ間の近接性を測定するために，3.1 節で生成した三つの移動行列を用いて近接性行列を作成する．ここで，三つの移動行列成分の単位はそれぞれ異なるため，以下の式を用いて正規化する．

$$nr(M) = \frac{M[c_s, c_e]}{max(M)}$$

そして，正規化した三つの移動行列を以下の式により統合し，近接性行列 M_R を作成する．この式において，w_S, w_T, w_F は三つの移動行列に対してそれぞれ付与する重み付け変数 $(0 \leqq w \leqq 1.0)$ である．提案システムでは，ユーザ操作によりこの重み付け変数を指定することができるようになっており，ユーザカスタマイズの地域クラスタ間の近接性行列を作成することができる．

$$M_R = w_S \cdot nr(M_S) + w_T \cdot nr(M_T) + w_F \cdot nr(M_F)$$

近接性行列成分 $M_R[c_s, c_e]$ は，二つの地域クラスタ間の絶対的な距離に基づく近接性である．都市空間における地域の構造を把握するためには，複数の地域間の近接性を分析する必要がある．そこで，生成した近接性行列に基づき相対的な近接性を算出する．このとき，都市空間における絶対的な近接性に近似させるために，絶対的な近接性が高い地域同士は相対的に近くに配置し，反対に絶対的な近接性が低い地域同士は相対的に遠くに配置する必要がある．そこで，多次元尺度構成法 (Multi-Dimensional Scaling; MDS) [3] を適用し，地域クラスタ間の相対的な近接性を求める．このアルゴリズムは，対象間の絶対的な距離に基づき，多次元空間上の各対象の座標値を求めて，点間の距離が観測された元の距離と最も一致する点を決定する．そして，対象間の相対関係に基づくインデクスマップを作成する．

実際に，提案方式に MDS を適用し，2 次元のインデクスマップを作成する方法について説明する．MDS では 2 対象間の距離について距離の公理の対称性を前提としているため，行列成分の値が対称ではない近接性行列をそのまま処理することができない．そこで今回は，地域クラスタ間の近接性を平均近接性に置き換えることにより，対称近接性行列 M'_R に変換し，MDS

を適用する．これにより，MDS は各地域クラスタの 2 次元空間上の座標値を算出し，相対的な近接性に基づき地域クラスタを配置し，図 5 の真ん中にあるようなインデクスマップを作成する．

3.3 インデクスマップを用いた近接地域検索

最後に，ユーザの目的に基づきカスタマイズした近接性に基づき，任意の地域クラスタ（以降，基点地域クラスタとする）の近接地域クラスタを検索する．そのために，3.2 節で作成した地域クラスタ間の近接関係の空間的な歪みが反映されているインデクスマップに k-NN (k-Nearest Neighbor) アルゴリズム [2] を適用し，基点地域クラスタと群衆の移動経験を通して近接している上位 k 件の近接地域クラスタを抽出する．図 5 の真ん中がインデクスマップであり，インデクスマップ上で基底地域クラスタの上位 20 個の近接地域クラスタを線でつないで示している．しかし，この表現では地域クラスタ間の地理的な位置関係が不明であるため，図 5 の右側に示すように，地図を用いて該当する地域クラスタに色付けして強調することで，ユーザが理解しやすいように表現する．今回のユーザに提示する地図の生成には，国土交通省が提供している国土数値情報の一つである行政区域データを利用している．

4. 実　　験

本章では，提案手法に基づきプロトタイプシステムを実装し，実際に Twitter から取得したデータセットを用いて近接地域を検索した結果について説明する．このプロトタイプシステムのインタフェースは，JavaScript をベースに作成しており，ウェブベースでのアクセスを可能にしている．そのため，スマートフォンのような携帯端末などからのアクセスも可能であり，いつでもどこでも簡単に利用することができるようになっている．

4.1 データセット

今回は，滋賀県，京都府，大阪府，兵庫県，奈良県，和歌山県を含む京阪神都市圏を対象都市として，対象都市内の市区町村レベルの行政区域を単位とした 234 の地域クラスタを用いた．次に，Twitter から対象都市内で発信されたジオタグ付きツイートを取得するために，Twitter の Streaming API と我々の過去の研究で開発した地理的ツイート収集システム [5] を用いて，日本の近畿地方を含む領域（経度 = [134.122433, 136.337186]，緯度 = [33.210804, 36.385050]）を収集対象領域に設定し，2012 年 9 月 16 日から同年 10 月 7 日の約 3 週間に発信されたジオタグ付きツイートを取得した．その結果，253,793 人のユーザが発信した 3,417,737 件のツイートを収集することができ，これらから計 154,366 もの 2 地域クラスタ間の移動セグメントを抽出することができた．

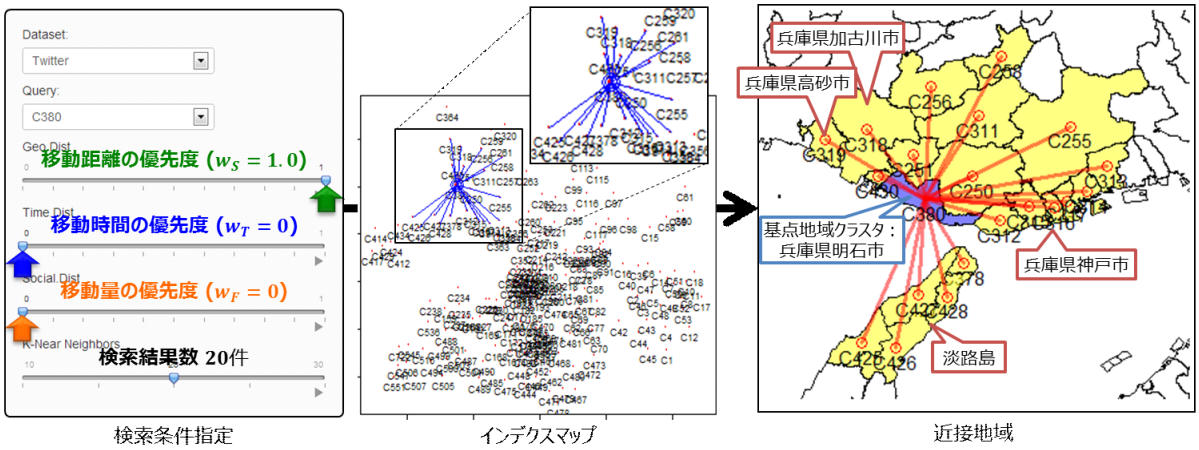
4.2 三つの尺度をそれぞれ優先したときの近接地域検索

4.1 節で紹介したデータセットを用いて，3 章で説明した提案手法に基づき，地域クラスタ間の移動距離，移動時間，移動量を抽出して，移動空間距離，移動時間距離，移動量距離の三つの尺度をそれぞれ算出して移動空間距離行列，移動時間距離行列，移動量距離行列を生成した．そして，ユーザ操作により重み付けされた三つの移動行列を統合して近接性行列を生成し，MDS によって求めた近接性に基づきインデクスマップを生成した．そして，このインデクスマップにおいて，ユーザが指定した基点地域クラスタの近接地域クラスタを k-NN アルゴリズムを用いて検索し，その結果を地図を用いて提示した．

まず，図 5 は，兵庫県明石市 (C380) を基点地域クラスタとして，三つの移動尺度をそれぞれ優先したときに検索される結果について比較した．図 5 の左側はユーザが検索条件を指定するためのユーザインタフェース，図 5 の真ん中はユーザの指定条件によりカスタマイズされた近接性に基づき生成されたインデクスマップ，図 5 の右側はインデクスマップ上で検索された近接地域クラスタの上位 20 件を地図上で表現した結果である．図 5 (a) の移動空間距離を優先した場合には，地理的に近接している地域クラスタを検索することができた．例えば，淡路島の市区町村，神戸市，加古川市などである．図 5 (b) の移動時間を優先した場合には，兵庫県姫路市，兵庫県神戸市内のいくつかの区，大阪府内のいくつかの市区，京都府京都市中京区や京都府京田辺市など，新快速や新幹線が停車するために短い時間で移動することができるような地域クラスタを検索することができた．一方，図 5 (a) の移動空間距離を優先した場合に検索された淡路島内の地域クラスタは検索されなかった．実際に，淡路島へ行くためのルートは限られており，移動手段は車かバスしかない．そのため，実験結果として検索された地域に行くよりも移動時間がかかるといえ，それを反映した結果が得られた．また，図 5 (c) の移動量を優先した場合には，兵庫県神戸市，大阪府大阪市北区，京都府京田辺市など，新快速や新幹線が停車する大きな駅がある地域クラスタや，基点地域クラスタの兵庫県明石市よりも東方向（大阪方面）にある地域クラスタを多く検索することができた．実際に，この基点地域クラスタからの移動は東へ向かう移動が多く，それを反映した結果を得ることができたといえる．

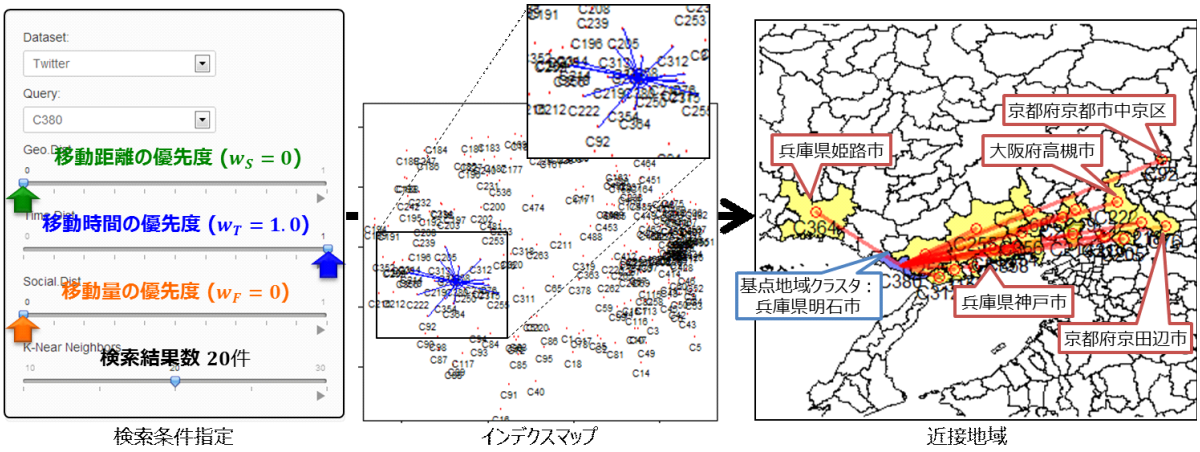
4.3 近接地域検索システムの使用例

次に，提案システムのユーザ使用例を示す．ユーザが，兵庫県明石市 (C380) を基点地域クラスタとして，移動時間と移動量を優先して上位 20 件の近接地域を検索しているとする．このユーザは，まず，地理的に狭い領域内で条件に合う近接地域クラスタを見つけようと思い，三つの移動尺度をすべて優先して検索を行った．その結果，図 6 (a) に示すように，兵庫県明石市と地理的に近接している兵庫県加古川市や兵庫県神戸市内のいくつかの区を中心とした近接地域クラスタを検索することができた．また，移動時間や移動量を優先しているため，地理的には離れている兵庫県明石市や大阪府内のいくつかの市区についても検索することができた．次に，図 6 (a) の結果を見たユーザが，移動距離に関する制約を緩めたときに，どのような地域クラスタが検索されるかを知りたいと思い，移動空間距離の優先度だけを少し下げて検索を行った．その結果を図 6 (b) に示す．図 6 (a) の結果と比較すると，兵庫県明石市と地理的に近接している兵庫県神戸市内のいくつかの区を中心とした近接地域クラスタと，大阪府内のいくつかの市区を中心とした近接地域クラスタという二つの地理的なまとまりが得られた．また，関西国際空港も検索されるなど，兵庫県明石市よりも東側



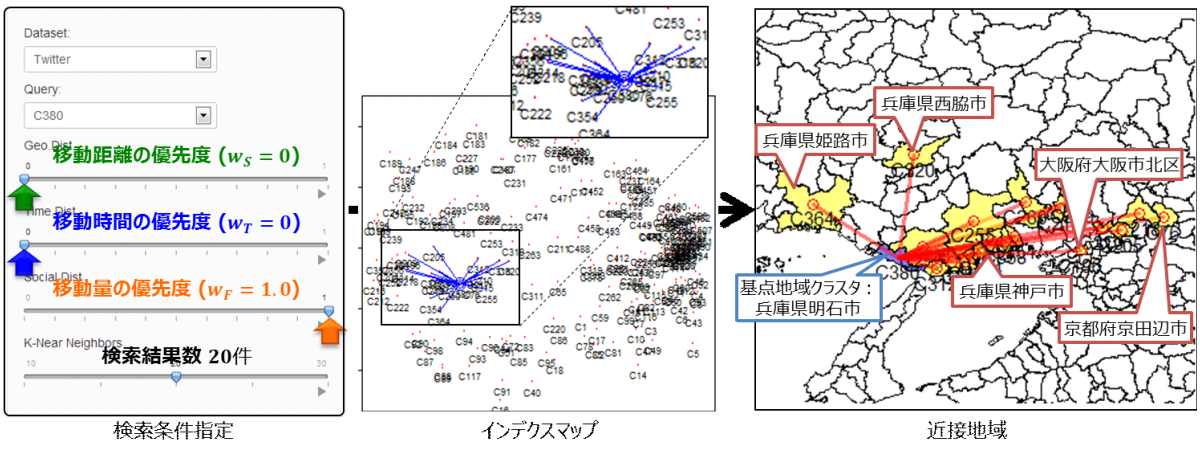
検索条件指定　インデクスマップ　近接地域

(a) 移動空間距離を優先したとき ($w_S = 0, w_T = 1.0, w_F = 1.0$)



検索条件指定　インデクスマップ　近接地域

(b) 移動時間を優先したとき ($w_S = 1.0, w_T = 0, w_F = 1.0$)



検索条件指定　インデクスマップ　近接地域

(c) 移動量を優先したとき ($w_S = 1.0, w_T = 1.0, w_F = 0$)

図 5　兵庫県明石市 (C380) の近接地域検索結果（(a) 移動空間距離を優先した場合，(b) 移動時間を優先した場合，(c) 移動量を優先した場合）

に広がる地理的領域内の地域クラスタが検索された．最後に，移動距離に関する制約をなくしたときの近接地域を調べるために，移動空間距離の優先度を 0 にして検索を行った．その結果，図 6 (c) に示すように，図 6 (a) や図 6 (b) の結果と比べて広域領域の近接地域クラスタが検索されたことが分かる．具体的には，西は兵庫県姫路市，東は京都府京田辺市を検索することができた．このように，ユーザは提示される結果を見ながらインタラクティブに三つの尺度の優先度を変えることができ，カスタマイズした近接基準による近接地域を検索することができる．

5. おわりに

本稿では，SNS に投稿されている群衆のライフログから群衆の移動経験を抽出し，ユーザの要求に合わせて地域間の近接性を測定し，求めた近接性を近接地域検索に応用したシステムの

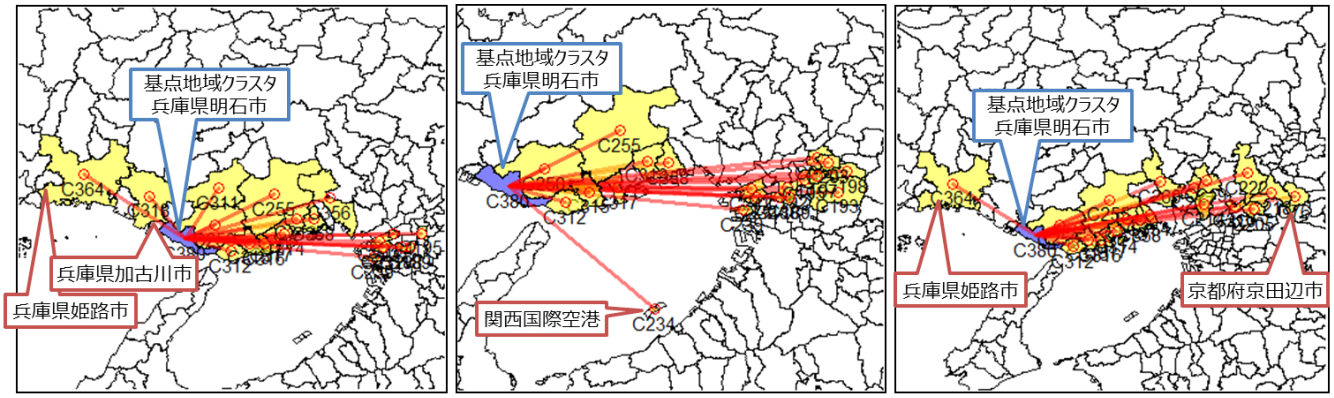


図 6　兵庫県明石市 (C380) の近接地域検索結果（移動空間距離の優先度を変更した場合）

提案を行った．特に，Twitter のジオタグ付きツイートを用いて，市区町村を単位とした地域間の群衆の移動経験を移動距離，移動時間，移動量の観点で抽出して，近接性の測定を行った．実験では，ユーザの要求に応じてカスタマイズした近接性に基づき近接地域を検索した結果について示し，考察を行った．

今後の課題として，今回は群衆の移動経験による地域間の近接性を測定し，定量的に近接地域を検索するシステムを提案したが，人々が日々の生活を通して抱いている地域間の近接感を SNS の群衆のライフログから抽出し，定性的に近接地域を検索するための手法について検討することを計画している．ユーザ属性や移動タイプを考慮して近接性の分析を行う予定である．例えば，テキストメッセージを分析することにより，ユーザの性別や年代といった属性，通勤・通学や観光といった移動タイプを分類し，それぞれの近接性に基づく近接地域検索が可能になると考えられる．また，直感的な近接性表現のためのユーザインタフェースの設計・開発について研究する予定である．本稿では，近接地域検索結果として，判定された地域を地図上で強調して表示したが，これでは，地理的な距離を強調した表示になってしまうため，地理的な距離が遠いにも関わらず，移動時間がかからなかったり移動量が多い場合に，それを直感的に示すことが困難である．そのため，近接性に基づき検索された近接地域を三つの尺度を考慮して表現するための手法について検討する必要があると考えている．また，被験者を用いた提案システムの評価についても，今後実施する予定である．

謝　　辞

本研究の一部は，平成 24 年度特別研究員奨励費 (24.9154) によるものである．ここに記して謹んで感謝の意を表する．

文　　献

[1] Choudhury, M. D., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R., and Yu, C. Automatic Construction of Travel Itineraries Using Social Breadcrumbs. In Proc. of ACM Conference on Hypertext and Hypermedia (HT), pp. 35-44 (2010).

[2] Cover, T., Hart, P. Nearest neighbor pattern classification. IEEE Transactions on Information Theory, Vol. 13, Issue: 1, pp. 21-27 (1967).

[3] Kruskal, J., Wish, M. Multidimensional scaling. In: Sage University Papers on Quantitative Applications in the Social Sciences, pp. 07-011 (1978).

[4] Kurashima, T., Iwata, T., Irie, G. and Fujimura, K. Travel Route Recommendation Using Geotags in Photo Sharing Sites. In Proc. of ACM Conference on Information and Knowledge Management (CIKM), pp.579-588 (2010).

[5] Lee, R., Wakamiya, S., Sumiya, K. Discovery of unusual regional social activities using geo-tagged microblogs. World Wide Web Special Issue on Mobile Services on the Web 14(4), pp. 321-349 (2011).

[6] Lee, R., Wakamiya, S., and Sumiya, K. Urban Area Characterization based on Crowd Behavioral Lifelogs over Twitter. Personal and Ubiquitous Computing, pp. 1-16 (2012).

[7] Leung, K. W.-T., Lee, D. L., and Lee, W. -C. CLR: A Collaborative Location Recommendation Framework based on Co-Clustering. In Proc. of ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR), pp. 305-314 (2011).

[8] Mislove, A., Lehmann, S., Ahn, Y.Y., Onnela, J.P., Rosenquist, J.N. Understanding the Demographics of Twitter Users. In Proc. of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM' 11), pp. 133-140 (2011).

[9] Pocock, D. The cognition of intra-urban distance. A summary of Scottish Geographical Magazine 94-1, pp. 31-35 (1978).

[10] Wakamiya, S., Lee, R., and Sumiya, K. Crowd-sourced Cartography: Measuring Socio-cognitive Distance for Urban Areas based on Crowd's Movement. In Proc. of the 4th International Workshop on Location-Based Social Networks (LBSN 2012), pp. 935-942 (2012).

[11] Wakamiya, S., Lee, R., and Sumiya, K. Looking into Socio-cognitive Relations between Urban Areas based on Crowd Movements Monitoring with Twitter. DBSJ Journal, Vol. 11, No. 2, pp.19-24 (2012).

[12] Zheng, V. W., Zheng, Y., Xie, X. and Yang, Q. Collaborative Location and Activity Recommendations with GPS History Data. In Proc. International Conference on World Wide Web (WWW), pp.1029-1038 (2010).

[13] 岡瑞起，橋本康弘，笹原和俊，池上高志. カートグラムを用いた空間のアフォーダンスの可視化インタラクション 2011 (2011).

[14] 李龍，若宮翔子，角谷和俊. Tweet 分析による群衆行動を用いた地域特徴抽出. 情報処理学会論文誌 : データベース (TOD54), Vol.5, No.2, pp. 36-52, 2012 年 6 月.