

地域特有の話題発見につながるスマートフォン向け検索サービス: 発見探地図エリアダス

藤田 尚樹[†] 数原 良彦[†] 片岡 良治[†]

[†] 日本電信電話株式会社 NTT サイバーソリューション研究所 〒 239-0847 神奈川県横須賀市光の丘 1-1

E-mail: †{fujita.naoki,suhara.yoshihiko,kataoka.ryoji}@lab.ntt.co.jp

あらまし スマートフォンの普及に伴い、移動先で地域情報に簡単にアクセスできる位置情報サービスの需要が高まっている。我々はこれまでキーワードに加えて、ユーザが任意に指定した地理・時間クエリに関連する文書を取得する地理・時間検索の研究を進めてきた。本稿ではさらにユーザの注目している場所・時間に特徴的なキーワードの提示と、指定されたキーワードに対する文書の地理的分布を視覚的に表示する技術を新たに組み合わせることで、ユーザがスマートフォン上で場所・時間に関連した情報をより容易に、発見的に探すことができる時空間マップ型 Web 検索技術を開発した。また、本技術を適用した実験サービスについても述べる。

キーワード 情報検索, 地理情報検索, 携帯端末

Areadas: Search Service for Smartphone Users to Discover Area-Specific Topics

Naoki FUJITA[†], Yoshihiko SUHARA[†], and Ryoji KATAOKA[†]

[†] NTT Cyber Solutions Laboratories, NTT Corporation

1-1 Hikarinooka, Yokosuka-Shi, Kanagawa, 239-0847 Japan

E-mail: †{fujita.naoki,suhara.yoshihiko,kataoka.ryoji}@lab.ntt.co.jp

Abstract As the number of smartphone users grows rapidly, there is an increase in demand for location-based services that enables users to access area-specific information wherever they are. We have studied search methods that retrieve documents satisfying geographical and temporal queries as well as keywords specified by users. In this paper, we enhance the methods by adding two new components: one shows “area tags” which are keywords that represent topics specific to the area where a user focuses on, and the other visually displays geographical distributions of the documents relevant to the input keywords. As a result, the user can discover and access geographically and temporally specific information efficiently even when the user does not come up with a keyword. We also describe our trial service for smartphone users, called Areadas that adopts the method.

Key words information retrieval, geographic information retrieval, mobile device

1. はじめに

情報検索技術は人々の日常生活において必要不可欠なものとなっている。近年ではスマートフォンの普及に伴い、戸外での利用も増加しており、外出時にその場所に関する情報検索を行うという利用方法が自然に行われている。日常生活でウェブ上の情報を用いて効率的に行動を決定する習慣がより広まってくることが今後も予想されるため、時間や場所など状況に応じたユーザの検索要求に答える検索サービスに対する需要がますます高まっていくと考えられる。そのようなユーザの検索要求に

答えるサービスとして、たとえば Google Maps^(注1), Yahoo!ロコ^(注2), 30min^(注3), 食べログ^(注4), じゃらん^(注5) などがある。例えば Google Maps や Yahoo!ロコでは、現在地を中心とする地図を表示した状態で「ラーメン」というキーワードを入力することにより、現在地の近くのラーメン屋情報を検索することができ、30min ではブログ記事からユーザのロコミ情報を抽出し、

(注1) : <http://maps.google.com>

(注2) : <http://loco.yahoo.co.jp>

(注3) : <http://30min.jp>

(注4) : <http://r.tabelog.com>

(注5) : <http://www.jalan.net>

地図上にマッピングする。また、食べログやじゃらんでは利用予定日と地域を検索の条件に加えることで、目的の日時と地域で利用可能な店舗を見つけることが可能である。このように、ユーザは検索サービスに対して検索キーワードに加えて地理や時間に関する要求も入力することで目的の情報を取得することができる。これらは、店舗情報やレビューなどのコンテンツを検索対象文書と見なせば、地理や時間をクエリとして入力可能な情報検索システムと見なすことができる。

本稿では、このようにキーワードに加えて地理的要求と時間的要求を入力する情報検索システムを地理日時検索サービスと呼ぶことにする。また、ユーザが地理日時検索サービスに入力するキーワードで表現されるクエリを**検索キーワード**、地理的要求を表すクエリを**地理条件**、時間的要求を表すクエリを**日時条件**と呼ぶ。地理日時検索サービスでは、ユーザに検索キーワード、地理条件、日時条件を明確に指定して検索を行うことが求められる。上記の食べログやじゃらんなど、具体的に利用シーンが定められているサービスにおいては、“特定の日時”に“特定の地域”のレストランに興味を持つユーザは、自身の情報要求を明確にすることが容易にできる。しかしながら、一般的なウェブ文書やブログ文書などを対象にした地理日時検索では、利用シーンが限定されないため、明確な情報要求を持たないユーザはクエリ生成に戸惑い、具体的な地理条件、日時条件を定められないという問題が生じる。そのため、ウェブ文書やブログ文書に対する地理日時検索サービスが実現できたとしても、それを利用することがユーザにとって容易でなければ、有用なサービスとは言えない。

我々は、検索キーワード、地理条件、日時条件のいずれかが欠けた情報要求を持つユーザに対して、欠けたクエリの入力を支援することができれば、ユーザが地理条件、日時条件を適切に指定することが可能となり、より簡便に地理日時検索サービスを利用することができると考える。それにより、地理日時検索を通常の検索サービスと同程度の情報要求で、より利便性の高いサービスとして実現することを目指す。

そこで、本研究では地理日時検索機能に地理日時検索支援機能を加えたアプリケーション開発を通じてユーザビリティの高い地理日時検索サービスについて検討を行った。

具体的には我々が既に開発した地理情報検索技術 [1] および日時指定検索技術 [2] を組み合わせることで基盤となる地理日時検索を実現し、ユーザが地理日時検索に必要なクエリの入力を簡便にする検索支援機能を追加することでスマートフォン向けアプリケーション「発見探地図エリアダス」を開発し、評価のためのフィールド試験を行っている。

以下、2章では開発方針、3章でシステムの主要機能、4章でフィールド試験について述べる。5章で「発見探地図エリアダス」についての評価を述べ、最後に6章でまとめる。

2. 開発方針

我々はウェブ上に存在する膨大な情報の中からユーザの入力した地理と日時に応じて情報を取捨選択できるスマートフォン向け地理日時検索サービスを実現するため、以下の要件を満た

すシステム開発を行う。

- 地図上の表示範囲を地理条件として指定可能である。
- 日時条件を指定可能である。
- 検索キーワード、地理条件、日時条件をクエリとする検索 (地理日時検索) が可能である。

地理条件を地図上の表示範囲で指定することを可能にするとは、Google Maps などで表示範囲を地理条件として検索を行うことが一般的になってきており、ユーザにとって直感的であると考えためである。

しかし、上記要件により地理日時検索を実現したとしてもユーザは検索キーワード、地理条件、日時条件を想起できなければ利用することができず、有用なサービスとは言えない。そのため、上記要件に加えて検索支援機能について検討を行い、本開発では検索キーワード、地理条件に対する検索支援機能を付与することとした。日時条件に関しては、サービスにおいて全期間を条件とする“日時指定なし”を許容することで代替する。**地理条件支援:** 指定した日時条件下で検索キーワードを“どこ”で検索すれば良いかで支援する。

検索キーワード支援: 指定した地理日時条件下でどのような検索キーワードで検索すれば良いか支援する。

この2つの支援機能は地理日時検索を利用するユーザにおける次のような場合を支援することを想定している。地理条件支援では3月に花見はどこへ行けば良いかわからないユーザに対して、“花見”と“3月”という検索条件で検索結果が得られる具体的な地域を提示することにより、適切な地理条件を選択できるように支援する。検索キーワード支援では、旅行中のユーザに対して、“現在地周辺”と“現在”という検索条件に加えて近隣の観光スポットや名物料理などを検索キーワードの候補としてユーザに提示することにより、その地域における話題発見を支援する。

この2つの支援機能を実現するため次の要件を満たすものとしてシステム設計を行う。

- (A) 日時条件下で表示している地図範囲内のキーワード分布図を表示
- (B) 地理日時条件下で表示された地図上に特徴的なキーワードを推薦

(A) は地理条件支援と対応している。これにより、ユーザが閲覧している地図範囲を与えられた地理条件の候補集合と見なし、その地図上における検索キーワードを含む検索対象文書の分布をヒートマップ図の要領で提示することで、ユーザが文書頻度が高い地域を地理条件として入力することが可能となる。

(B) は検索キーワード支援と対応している。これにより、ユーザが閲覧している地図範囲を地理条件と見なし、与えられた地理条件、日時条件に対して文書に特徴的に出現するキーワードを推薦することにより、ユーザが興味を持った検索キーワードとして選択することが可能となる。

本支援機能は情報を地図上に表示することで、ユーザが推薦キーワードや検索対象文書が地図上のどの範囲と対応付けられているのかを視覚的に理解することが可能となり、検索キーワードの選択及び地理条件の決定を地図を見ながら連続的に行

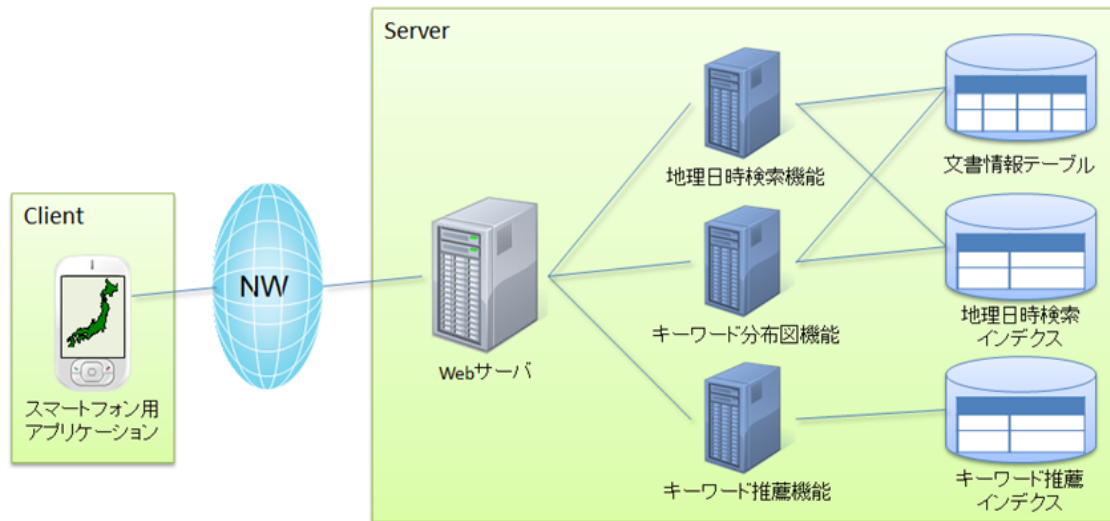


図 1 システム概要図

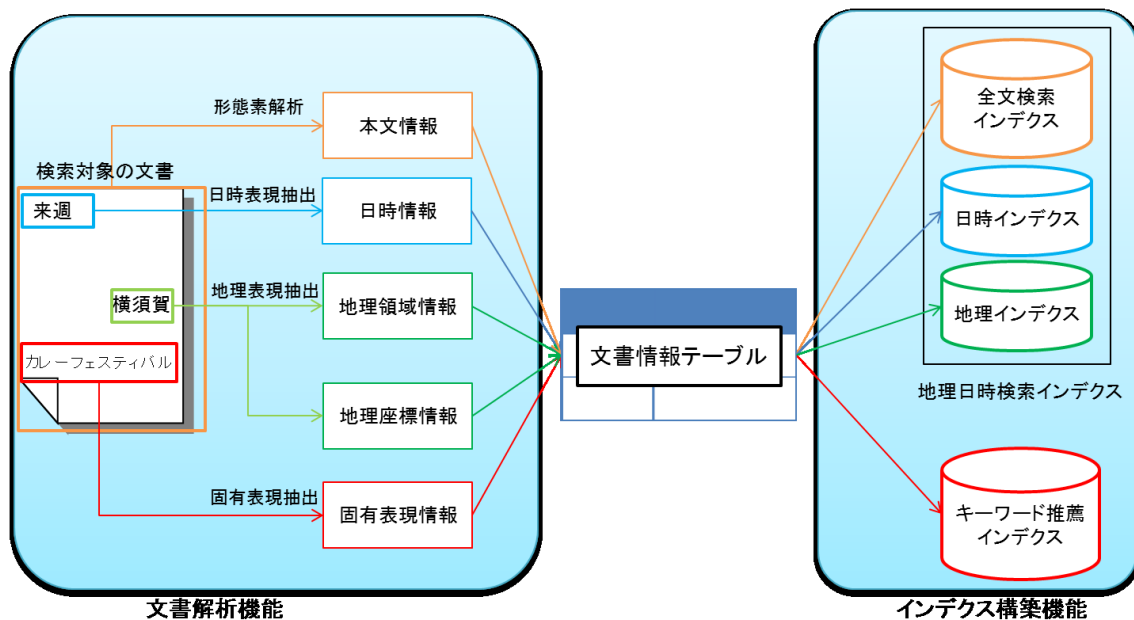


図 2 文書解析機能とインデックス構築機能の処理の流れ

えるようにしている。

システム構築にあたり、構成は利用者とのインタフェースとなるスマートフォン用アプリケーションと、アプリケーションとデータのやり取りを行うサーバとし、検索などの計算処理は可能な限りサーバ側で行うよう設計する。

3. 主要機能

本章では本システムの各機能について述べる。図 1 にシステム概要図を示す。本システムはクライアントとしてのスマートフォン用アプリケーションとサーバに分けられる。サーバは地理日時検索機能、キーワード分布図生成機能、キーワード推薦機能とそれぞれが問い合わせを行うインデックス及び検索対象文書の情報を保持する文書情報テーブルで構成される。本システ

ムでは検索対象文書から、インデックス及びテーブルを図 2 に示す流れで文書解析機能とインデックス構築機能によって構築する。

本章では以下、文書解析機能、インデックス構築機能、サーバ機能、クライアント機能について順に述べる。実際に開発したスマートフォン用アプリケーション「発見探地図エリアダス」については別途 4. 章で後述する。

3.1 文書解析機能

本システムは検索対象文書の本文を解析し、以下の情報を抽出し、それぞれを対象文書の情報として文書情報テーブルに保存する。文書情報テーブル内のデータ及びその抽出手法は以下の通りである。

文書 ID: 文書に対してユニークな ID を付与。

URL: 文書のデータをそのまま保存。

タイトル: 文書のデータをそのまま保存。

本文情報: 形態素解析を行い、全ての形態素を本文情報として抽出。形態素解析には日本語形態素解析器 JTAG [1] を用いる。

推薦キーワード情報: 文書から固有表現や日本語 Wikipedia^(注6) の見出し語を推薦キーワードの候補語として抽出。

地理座標情報: 本文から地名表現を特定し、その地名表現が対応する地理座標 (緯度・経度) 情報を抽出。地名表現及び座標特定には平野ら [2] の手法を用いている。

地理領域情報: 地理座標情報と同様に文書から地名表現を抽出し、その地名表現と対応する地理的領域を地理領域情報として抽出する。この際、地名表現に対応する地理領域情報は安田らの手法 [3] に従い、地名と緯度経度範囲情報の対応に基づいてあらかじめ定められたものとする。本システムでは地理領域情報を矩形で表している。

日時情報: 本文中に含まれる「明日」「来週」「今月」といった日時表現を特定し、各日時表現が示す期間を日単位 (開始日・終了日) で抽出する。ここで日時情報は、文書の更新日時そのものではなく、日時表現が表す期間である。日時表現の特定と、期間への変換は廣嶋らの手法 [4] を用いる。

3.2 インデクス構築機能

地理日時検索機能を実現するため、3.1 で得られた文書情報テーブルの情報をを用いてあらかじめ全文検索インデクス、地理インデクス、日時インデクスの3つのインデクス構築を行う。また、キーワード推薦機能を実現するためのキーワード推薦インデクスは、推薦キーワード情報と日時情報、地理座標情報、地理領域情報を用いて構築する。

インデクス構築機能の概要は図2の右側の通りである。

全文検索インデクス: 文書情報テーブル内の本文情報と文書 ID を元に、全文検索インデクスを構築する。本インデクスは一般的な全文検索で用いられる転置インデクスを用いることができる。

地理インデクス: 文書解析機能によって得られた地理領域情報と文書 ID を元に、与えられた地理領域条件に対して、重なりのある地理領域情報を含む文書 ID を取得するための地理インデクスを構築する [3]。

日時インデクス: 文書情報テーブル内の日時情報と文書 ID を元に、与えられた日時条件が表す期間に対して、重なりのある日時情報を含む文書 ID を取得するための日時インデクスを構築する [5]。

キーワード推薦インデクス: 文書情報テーブル内の文書 ID、固有表現情報、地理座標情報、日時情報を元に、与えられた日時条件が表す期間に対して重なりのある日時情報を持ち、与えられた地理領域に含まれる地理座標を持つ文書に含まれる推薦キーワードとその頻度情報を取得するためのキーワード推薦インデクスを構築する。インデクスは緯度・経度・日時の3次元をそれぞれ一定の単位で分割した3次元グリッドで表現し、各グリッドに対して推薦キーワードの情報を対応付けた。3次元グ

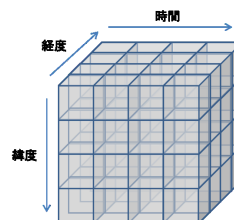


図3 キーワード推薦インデクスの3次元グリッドのイメージ

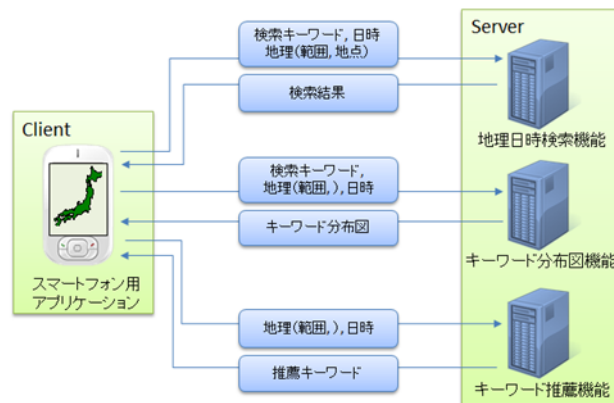


図4 アプリケーションと各機能間のデータフロー

リッドのイメージを図3に示す。キーワード推薦インデクスは緯度、経度、時間という3次元をあらかじめ定められた一定の値で離散化した各ブロックに対して、キーワードとそれに対応する頻度情報を保持している。これを用いることで、地理条件と日時条件が与えられた際に、与えられた条件を満たすグリッド集合に含まれるキーワードと頻度情報を元に推薦キーワードとそのスコアの計算を行うことができる。

サーバ機能は地理日時検索機能、キーワード分布図生成機能、キーワード推薦機能で構成される。サーバとクライアント間における各機能のデータフローを図4に示す。

3.3 地理日時検索機能

地理日時検索機能は、検索キーワード、地理 (地点、範囲)、日時の情報を含むクエリを入力として受け取り、ランキングされた検索結果集合とそれぞれの概要文を返す。

検索は一般的な検索キーワードへの適合性に基づくキーワードスコア、安田らの手法 [3] により入力された地理 (座標、領域) と文書が含む地理領域情報との地理的適合性を表す地理スコア、廣嶋らの手法 [4] により入力された日時に対する時間的適合性を表す日時スコアを算出し、それぞれのスコアの積に基づいて最終スコアを算出し、その降順に文書のランキングを行う。日時クエリが入力されない場合には、地理日時検索スコアの計算から日時スコアを外して地理日時検索スコアを計算する。

概要文生成では、戸田らの手法 [6] を改変し、地理の情報を可能な限り含むようにした。

3.4 キーワード分布図生成機能

ユーザが検索キーワードをどこで検索すると良いかわからない状況において、キーワードの地理的な分布情報を提示し、分布の高い地域をユーザが地理条件として選択できるように支援

(注6) : <http://ja.wikipedia.org>

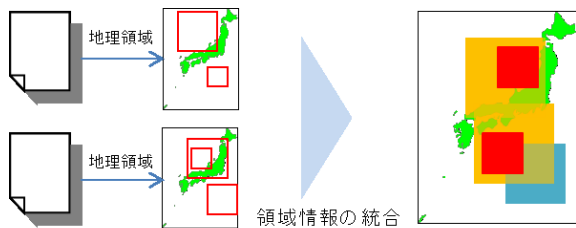


図 5 キーワード分布図生成機能の概念図

することでサービスの有用性を高められると考える。

そのためキーワード分布図生成機能は、検索キーワードと地理領域、日時の情報を入力として受け取り、地図上のどの領域について書かれた文書中に検索キーワードが多く出現するかという情報をヒートマップの要領で提示する。そのためキーワード分布図生成機能は、システムの出力画面を有限な数の矩形メッシュに分割し、各メッシュにおける検索キーワードの頻度情報に基づく分布情報を計算し、出力する。キーワード分布図生成機能のイメージを図 5 に示す。

まず、ユーザが入力した検索キーワードと地理日時条件に適合する文書 ID 集合を全文検索インデクス、地理インデクス、日時インデクスを用いて取得する。

次に取得した文書 ID 全てに対して文書情報テーブルから全ての地理領域情報を取得し、各地理領域情報と重なりがあるメッシュ全てについて、その文書内の検索キーワードの頻度から分布スコアを求め、頻度を加算する。

この際、出力画面に対して閾値以上の大きさの地理領域情報は画面全体に対して頻度を加算することになるため除外する。計算された分布情報を元にヒートマップの画像ファイルを生成し、クライアントに送信する。

3.5 キーワード推薦機能

ユーザが地理領域と日時が決まっているが何で検索して良いかわからない状況において、その地理日時条件下で特徴的なキーワードを推薦し、ユーザが興味を持ったキーワードを検索キーワードとして指定できるように支援することでサービスの有用性を高められると考える。

そのためキーワード推薦機能は地理領域と日時の情報を入力として受け取り、その条件下で推薦するキーワードを返す。推薦するキーワードは指定された地理日時条件下で特徴的なキーワードとなるよう、全期間・地理的領域で均等に出現した場合の期待値に対して、指定された地理日時条件での実値の比率が高いキーワードを出力する。データマイニング分野でよく用いられる指標としてリフト値 [7] があるが、全グリッド (緯度・経度・日時で分割した単位) 数に対して各キーワードの総出現頻度は小さいため、各キーワードのスコアが同じような値となってしまう、この用途には適さない。そのため、本システムではキーワードの特徴を表す指標としてポアソン分布を仮定した累積確率を用い、その値を特徴スコアとした。

キーワード w の特徴スコア $S(w, r)$ は、全検索対象文書数を n 、 w の全文書中での出現頻度を s 、指定された地理日時条件下での全てのキーワードの出現頻度を k 、 w の頻度を r としたと

き次の式で計算できる：

$$S(w, r) = \sum_{x=0}^r \frac{\exp(-\frac{sk}{n}) (\frac{sk}{n})^x}{x!}$$

推薦するキーワードは指定地理日時条件に含まれるグリッド全てのキーワードの中から特徴スコアの値が大きいキーワードを出力する。

3.6 スマートフォン用アプリケーション

クライアントアプリケーションはユーザが指定した検索キーワード、地理領域、地理座標、日時の情報をクエリとして地理日時検索機能に問い合わせを行い、サーバから返ってきた情報を表示する機能を持つ。また、キーワード分布図機能、キーワード推薦機能に対しても、それぞれ必要な情報をクエリとして問い合わせ、応答結果をキーワード分布図、推薦キーワードとして表示する。データの送受信には通常の HTTP プロトコルによる通信を用いる。

4. 発見探地図エリアダス

本システムをユーザがいつでも、どこでも利用できるようにクライアントアプリケーションとして Android アプリケーションの開発を行った。そして、今後の地理日時検索の精度及びユーザビリティの向上を目的として、ログ収集とユーザからのフィードバックを得るために、2011 年 12 月 15 日から 2012 年 6 月末までの計画でフィールド実験を開始した。開発したアプリケーションは**発見探地図エリアダス**という名称で Android マーケット^(注7)で無償提供している。

図 6 に発見探地図エリアダスの画面と利用例を示す。図 6(a) は基本画面である。ユーザは表示される画面を操作して地理的領域を指定することができる。地図インタフェースには Google Maps API を用いて表示しており、ドラッグによる地図上の移動、ピンチイン・ピンチアウトの操作で地図の縮小・拡大が可能である。

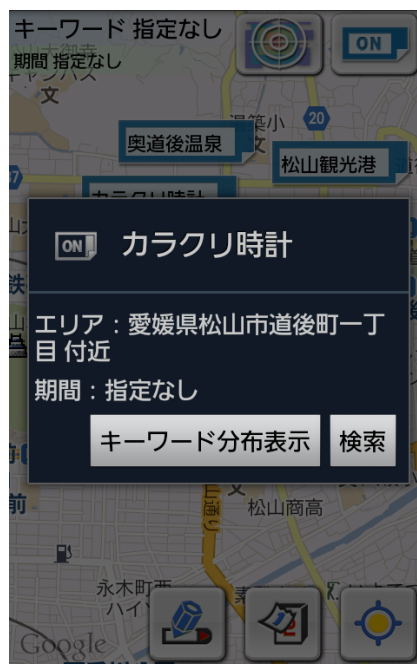
図 6(a) では、愛媛県松山市の松山城近くの地理範囲に対して提示された推薦キーワードの例である。このように発見探地図エリアダスは、ユーザが現在閲覧している地図の表示範囲内および指定日時において特徴的なキーワードの提示を行う。ユーザはその中から自分の興味を持ったキーワードをタップすることにより、選択することができる。キーワードを選択すると図 6(b) の画面が表示される。この例では、提示されたキーワードの中から「カラクリ時計」を選択した場合の画面例を示している。このように、ユーザ自身が検索キーワードを入力することなくキーワード分布図表示および地理日時検索を行うことができる。

図 6(c) は図 6(b) において検索ボタンを選択した際の地理日時検索結果の例を示している。この例では、選択されたカラクリ時計という検索キーワードと、閲覧している地理範囲、そして提示されたキーワードに紐づけられた地理座標情報を地理条件として地理日時検索を行う。検索結果は与えられた条件に

(注7) : <https://market.android.com/>



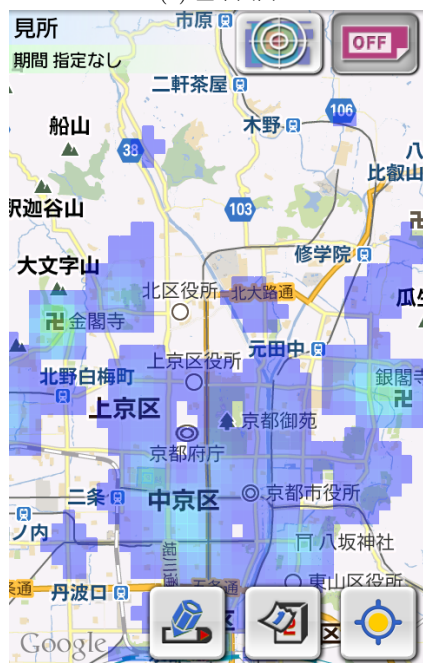
(a) 基本画面



(b) 推薦キーワード選択時の画面



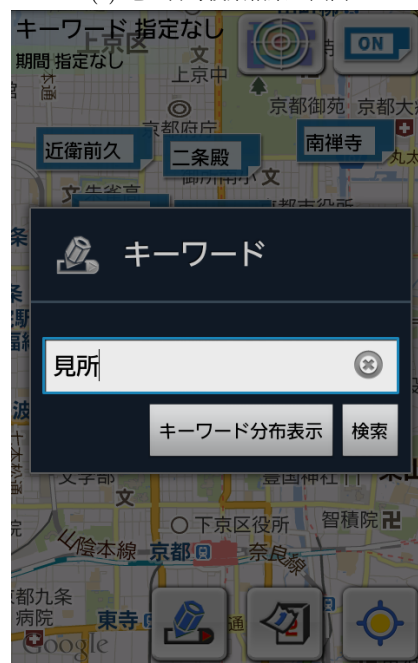
(c) 地理日時検索結果の画面



(d) キーワード分布図表示例



(e) 日時条件指定の画面



(f) キーワード入力画面

図 6 発見探地図エリアダス画面例

従って検索ランキングを行い、検索キーワードと地理条件を考慮したスニペットを提示している。

図 6(d) はキーワード分布図の表示画面例である。この例では「見所」という検索キーワードに対する京都市周辺のキーワード分布図を示している。色が濃くなっているところに、検索キーワードに該当する検索結果となる文書がより多く存在することを表している。ここでユーザが地図をタップすることにより「見所」を検索キーワード、タップした位置を地理条件とした地理日時検索を行うことができるため、タップした位置に応じて異なる検索結果を閲覧することができる。

図 6(e) は、日時条件指定の画面である。基本画面から右下に

3つ並んでいるアイコンのうち、真ん中にある日めくりカレンダーを模したアイコンを選択すると日時指定の画面が表示される。ここに提示されているように、日時条件は年単位、月単位、日単位で設定が可能である。

図 6(f) は、キーワード入力画面である。基本画面から右下に3つ並んでいるアイコンのうち、左にある鉛筆を模したアイコンを選択するとキーワード入力画面に遷移する。地図上に表示された推薦キーワードを選択する以外にも、ユーザが好みの検索キーワードを入力することが可能である。ユーザは自分で入力した検索キーワードに対して、キーワード分布図の表示、地理日時検索の実行を選択することができる。

表 1 検索支援機能の利用比率

検索方法	比率
(a) 推薦キーワードを選択した後に検索	0.731
(b) キーワード分布図を表示した後に検索	0.244
(c) 自分で入力したキーワードで検索	0.126

5. 評価

本章では我々が開発した発見探地図エリアダスについて、フィールド実験を通じて得られたユーザログを元に検索支援機能の評価を行うと共に被験者を用いた評価実験を通じてシステムの有効性の検証を行った。以下、それぞれについて述べる。

5.1 検索支援機能の評価

画面に表示される推薦キーワードは、ユーザが選択した地理日時条件下において、どのような検索キーワードを用いて検索すればよいかわからないユーザに対して、当該条件下で特徴的に出現するキーワードを提示することで地理日時検索のクエリ入力支援を行っている。そのため、推薦キーワードを選択することでユーザの検索行動が生じていれば提示した推薦キーワードは有用であったといえる。

また、キーワード分布図はユーザが入力した検索キーワードについてどのような地理条件を入力すればより多くの文書を得られるかをユーザに視覚的に提示することによって、地理条件の入力の支援を行う。そのため、推薦キーワードと同様に、キーワード分布図によってユーザの検索行動が生じていればキーワード分布図が地理条件の入力に有用であったといえる。

本研究では、発見探地図エリアダスのフィールド実験を開始した 2011 年 12 月 15 日から 2012 年 1 月 10 日までの検索ログを解析し、推薦キーワードとキーワード分布図が地理情報検索における検索支援機能として有用であるかの検証を行った。具体的には、全ユーザの全ての検索行為において、(a) 推薦キーワードを選択した後に検索した回数および (b) キーワード分布図を表示した後に検索を行った回数、そして (c) ユーザが自分で検索キーワードを入力し、キーワード分布図を用いることなく検索を行った回数の分析を行った。

結果を表 1 に示す。ここで、表 1 の (a), (b), (c) の総和が 1 にならないことに注意する必要がある。これは推薦キーワードを選択した後にキーワード分布図を表示し、検索を行った場合に、その回数が (a) と (b) 両方にカウントされるためである。

表 1 より、推薦キーワードを用いての検索は全体の 7 割以上となっており、推薦キーワードがユーザに対して検索を想起させる役割として有用だと考えられる。また、全ての検索行為の 4 分の 1 程度はキーワード分布図を表示してから検索を行っており、キーワード分布図がユーザに対してある程度有効に機能していると考えられる。ただし、現在収集しているログでは、キーワード分布図のどの色の箇所でも検索を行ったかわからないため、ユーザがキーワード分布図によって本当に地理条件の入力支援を行っているか検証することができない。また、検索結果として表示された文書数をログとして取得できずキーワード分布図によって選択された地理条件が適切で

表 2 評価実験の結果

	平均満足度	発見に要する時間 (秒)
発見探地図エリアダス	3.4	167.9
Google Maps	2.8	194.9

あったかも検証できていない。

今後も収集ログを拡張した上で継続的にユーザ履歴を収集し、ユーザの検索結果に対するクリック履歴やランディングページなどの情報も含めて解析していくことで、キーワード分布図の有効性を評価すると共に、2つの検索支援機能が地理日時検索の満足度を高める効果があるかということを併せて明らかにしていく必要がある。

5.2 被験者を用いた評価実験

我々は、ユーザの意図した情報の取得を容易にすることを目指し、地理日時検索機能に加えて検索支援機能を持つサービスとして発見探地図エリアダスの開発を行った。そのため、本システムが本当にユーザの情報検索に対して有用であるかを検証するため、具体的な利用シーンを想定して被験者を用いた評価実験を行った。

本実験ではシナリオを下記のように設定した。

「ある地域を訪れた際に、ついでに寄りたいと思うスポットを 3 つ見つける」

このシナリオにおいて、各被験者が与えられたシステムを用いて検索行為を行った際の発見したスポットに対する満足度の評価と発見に要する時間の計測を行った。地域によって評価軸は大きく変わるため、あらかじめ観光などで訪れる可能性の高い国内 4 地域を訪問先の候補とした。各被験者に対して、候補の中から 2 地域を割り当て、発見探地図エリアダスと Google Maps の 2 つのシステムを用いて実験を依頼した。この際、被験者が出身地や居住経験のある地域などは外すように地域の割り当てを行った。各被験者の 1 地域の探索行動について、3 つのスポット発見までに要する時間を計測し、発見したスポットに対する満足度を被験者自身が 5 段階で評価した。本評価実験では、4 人の被験者を用意し、地域と検索手段の組み合わせは偏りが生じないように行った。評価実験の結果を表 2 に示す。表 2 の平均満足度は、1 地域における 3 つのスポットに対する満足度の平均値について全地域、全被験者の平均を取ったものである。発見に要する時間も同様に、全地域、全被験者の平均を取ったものである。

結果より、発見探地図エリアダスによって発見したスポットに対する満足度の平均は 3.4 となり、Google Maps で発見したスポットに対する満足度の平均である 2.8 に比べて高い値となった。これより発見探地図エリアダスによって、より満足度の高い地域特有のスポット検索が可能になることが示唆された。また、発見までに要した平均時間が発見探地図エリアダスの方が短いということから、地域特有の情報により簡便にアクセス可能であると伺える。しかし、今回の評価は評価者間の満足度の分散が大きいため、満足度の平均値に統計的有意差が確認できなかった。また、訪問対象のスポットに対する満足度で評価

```
<TOPIC ID="GeoTime-0026">
<DESCRIPTION LANG="JA">いつ、どこで、スペースシャトル
コロンビアの事故は起こった？</DESCRIPTION>
<NARRATIVE LANG="JA">ユーザーはどのような状態でスペース
シャトルコロンビア号が爆発したのかを知りたがっている。
</NARRATIVE>
<QUERYDATE YYYYMMDD="20051231"></QUERYDATE>
</TOPIC>
```

図 7 NTCIR-9 GeoTime 日本語トピックの例

を行ったが、ユーザが今まで知らなかった地域特有の話題発見に対する評価になっていない。今後は、評価対象地域と評価者を増やすとともに、地域特有の話題発見につながるかという観点を含めて本システムの性能を検証する予定である。

6. 関連研究

情報検索において地理情報検索に関する研究は古くから取り組まれており [8] [9], 日時情報検索も地理情報検索とは独立して研究が進められてきた [10] [11] [12]. しかしながら、地理と日時を同時に考慮した情報検索についての取り組みは比較的新しく、2010 年には NTCIR-8 において GeoTime と呼ばれる地理日時検索の評価タスクが初めて開催された [13]. 2011 年の NTCIR-9 でも引き続き GeoTime タスクが開催されている^(注8).

図 7 に NTCIR-9 GeoTime 日本語トピックの例を示す。GeoTime タスクではユーザの検索意図を表すトピックは基本的に「いつ」「どこで」という明確な情報要求を含んでおり、対応する正解が存在する。GeoTime タスクではユーザが検索意図がある程度明確にした状態における地理日時検索システムを想定しているため、本研究で想定しているユーザが明確な検索意図を持たない状態においてユーザが部分的に検索意図を入力するという利用シーンを考慮していない。

また、GeoTime タスクで扱うトピックは基本的にウェブ検索における navigational クエリ [14] のように対応する正解が存在するような検索意図である。一方、発見探地図エリアダスが想定する利用シーンでは、ユーザが「この地域に特有の話題は何か」というウェブ検索における informational クエリ [14] に相当する情報要求を持って検索行動を行う。このように、ユーザの行動支援を実現するという点において本研究の目指す地理日時検索システムの方向性と異なる。

日時情報検索の既存研究としては、河合ら [12] は検索キーワードに対する検索結果をウェブ検索エンジンから取得し、そこから日時表現を抽出することにより、過去・未来のイベント情報を俯瞰する検索システムを提案している。河合らの手法では、検索キーワードに適合する文書をウェブ検索エンジンを利用して取得した後に時間表現の抽出を行っているが、我々のシステムは事前に日時インデックスを構築している点で異なる。我々のシステムでは事前にインデックスを構築することにより、大規模な検索対象の文書群に対して高速な日時条件を指定した

キーワード分布図の表示、検索結果の取得が可能となる。

7. おわりに

本稿では、地理日時検索サービスにおいてユーザが明確な検索意図を持たない状況においてユーザの検索行動を支援する地理日時検索支援機能を提案し、これを実現したスマートフォン向けサービス発見探地図エリアダスのシステム開発とフィールド実験について述べた。評価実験では、訪問先のスポット検索のシナリオに基づく被験者実験を通じて、Google Maps に比べて発見探地図エリアダスがより短い探索行動でより高い満足度を得られることを示した。また、サービスの検索ログを解析し、推薦キーワードがユーザの検索行動を支援していることを確認した。今後は日時条件に関する検索支援機能の実現を目指すと共に、検索ログ及びユーザからのフィードバックの解析を行い、キーワード分布図の効果測定をはじめとする本システムの評価と地理日時検索サービスのさらなる精度向上を目指す。

文 献

- [1] T. Fuchi and S. Takagi. Japanese morphological analyzer using word co-occurrence - jtag. In *Proc. COLING-ACL*, pp. 409–413, 1998.
- [2] 平野徹, 松尾義博, 菊井玄一郎. 地理的距離を用いた地名の曖昧性解消. 第 70 回情報処理学会全国大会, 2008.
- [3] 安田宜仁, 戸田浩之. 検索位置のごく周辺を対象とした地理情報検索. 人工知能学会論文誌, Vol. 23, No. 5, pp. 364–373, 2008.
- [4] 廣嶋伸章, 別所克人, 小池義昌, 片岡良治. 記述された日時の有効範囲を考慮した日時指定検索. In *Proc. WebDB Forum2010*, 2010.
- [5] 廣嶋伸章, 戸田浩之, 松浦由美子, 片岡良治. ブログ記事からの特定の時間帯を表す時間表現の抽出. 第 16 回言語処理学会年次大会, 2009.
- [6] 戸田浩之, 安田宜仁, 松浦由美子, 片岡良治. 地理情報検索における概要文生成法. 人工知能学会論文誌, Vol. 24, No. 6, pp. 494–506, 2009.
- [7] P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison-Wesley, 2005.
- [8] R. R. Larson. Geographic information retrieval and spatial browsing. *Information Retrieval*, pp. 81–124, 1996.
- [9] C. B. Jones and R. S. Purves. Geographical information retrieval. *Int. J. Geogr. Inf. Sci.*, Vol. 22, pp. 219–228, 2008.
- [10] I. Mani, J. Pustejovsky, and B. Sundheim. Introduction to the special issue on temporal information processing. Vol. 3, pp. 1–10, 2004.
- [11] R. Baeza-Yates. Searching the future. In *Proc. ACM SIGIR Workshop MF/IR*, 2005.
- [12] 河合英紀, Adam Jatowt, 田中克己, 國枝和雄, 山田敬嗣. Chronoseeker: Web からの過去・未来情報のオンデマンド検索エンジン. In *Proc. WebDB Forum2009*, 2009.
- [13] F. Gey, R. Larson, N. Kando, J. Machado, and T. Sakai. NTCIR-GeoTime overview: Evaluating geographic and temporal search. In *Proc. NTCIR-8 Workshop Meeting*, pp. 147–153, 2010.
- [14] Andrei Broder. A taxonomy of web search. *SIGIR Forum*, Vol. 36, pp. 3–10, 2002.

(注8) : <http://metadata.berkeley.edu/NTCIR-GeoTime/description.php>