

Title	KyotoSearch : An integrated system for geographic web search using web contents analysis(Abstract_要旨)
Author(s)	Lee, Ryong
Citation	Kyoto University (京都大学)
Issue Date	2003-09-24
URL	http://hdl.handle.net/2433/148500
Right	
Type	Thesis or Dissertation
Textversion	none

氏名	李 龍 ^{リョウ}
学位の種類	博士 (情報学)
学位記番号	情博第 87 号
学位授与の日付	平成 15 年 9 月 24 日
学位授与の要件	学位規則第 4 条第 1 項該当
研究科・専攻	情報学研究科社会情報学専攻
学位論文題目	KyotoSEARCH : An Integrated System for Geographic Web Search Using Web Contents Analysis (KyotoSEARCH : ウェブコンテンツ分析を用いた地域ウェブ検索のための総合システム)
論文調査委員	(主査) 教授 上林 彌彦 教授 河田 恵昭 教授 美濃 導彦

論 文 内 容 の 要 旨

本論文では、ウェブ情報を用いた地域情報検索システム KyotoSEARCH を提案し、そのための基盤研究および応用研究について述べているもので、6章から成っている。

第1章は序論で、ウェブ情報を用いた地域情報検索においてウェブページに対応する地域の確定が困難なために実用上の問題があることを取り上げ、従来の地理情報システムを用いたウェブ検索システムへの拡張が重要であることを示し、本研究の目的とアプローチを概説している。

第2章は、地域ウェブ情報検索において関連情報を連続的に検索することが困難である理由を説明し、ウェブ情報空間・実空間、キーワード情報空間を自由に情報ナビゲーションできるモデルを提案し、それに基づいて構築したシステム KyotoSEARCH を紹介している。このシステムの利用者インタフェースは、地図、地名と非地名の連想関連、ウェブページ検索結果リストから構成されている。利用者が関心を持っている地名または非地名に関連がある地図上の場所、関連キーワード（複数でよい）、検索結果リストを同時に表示する機能を備えている。また、このシステムを実際に利用した人々の評価による改良案について述べていて、このシステムが地域・言語に対する移植性が優れ、韓国のソウルを中心とした地域用にも適用していることを述べている。

第3章は、有用な情報を得るための分類とランキング手法について述べている。ウェブ情報を利用した地域検索結果には、重複したものや重要でないものがあり、それらの問題に対処するためのものである。提案された分類手法には、ウェブ情報における地名の性質と詳細度による区別を定義し、実験による評価も行っている。リンク構造を利用した PageRank と呼ばれるランキングアルゴリズムが使われているが、地域限定の場合に適用すると十分なリンク構造が得られない問題点がある。このため、地域とリンク構造の拡張による改良法を提案している。また、利用者が地域ウェブ情報のみを検索していくための地域指向モデルを提案している。さらに、これらの提案に対する実データを利用した実験による評価について説明している。

第4章は、地域ウェブ検索に適した、質問作成手段として地図インタフェースを用いる方法を提案し、実際にシステムを構築して評価している。従来は地図上での点や円（中心点と半径）として質問領域を定義する方法が用いられていたが、本論文では利用者が直感的利用できるようにするため、2次元領域質問処理手法を提案している。また、評価のために、ウェブから収集した（地名を含む）地域ウェブページを地図上に2次元最小対応矩形領域として対応させ、利用者の領域質問との地理演算による検索結果の再現率が低くなる場合を分析している。その結果、同じ地名が二つ以上の場所に使われる曖昧性、無駄な地理階層情報、ウェブページ内の重要ではない地名の検出が重要であることを示し、再現率をあげるためにウェブページの対応領域の最適化手法を提案している。実験によってこの方法により再現率があげられることを示している。

第5章は、地域ウェブ情報を用いた認知地理空間分析とモバイル端末での利用について述べている。ウェブ情報には、現実空間に対する人々の認識が反映されていると考えられ、地域ウェブ情報を分析することによるランドマーク計算法を提案

している。実際の計算では、ウェブページに出現する地名と近傍表現を計量し、ランドマーク対象が周囲の他のランドマーク対象によって影響を受けて認知度が落ちるという仮説で、実験によって評価をしている。結果として、周辺にランドマーク対象が少なくなるほど近傍表現の比率が高くなり、ランドマークとして認識されるようになることを示している。また、地域ウェブ情報をモバイル端末で使う場合には、有効なキャッシュモデルが必要であることを述べ、そのキャッシュ特性分析とPre-Fetchのモデルを提案している。モバイル端末は個人利用向けであるため、一回使われた情報はすぐには再利用されない特性を持っているので、利用者の移動情報、時間制約、個人趣向によって予め必要なデータを予測し、キャッシュに空きが生じた時にデータを取得するPre-Fetchが有効であることを述べている。

第6章は結論で、本論文のまとめと今後の展望が述べられている。

論文審査の結果の要旨

本論文は、ウェブ情報を用いた地域情報検索を目的としたシステム KyotoSEARCH の研究開発の結果を示したもので、地域ウェブ情報検索の質を向上させるための検索結果の分類とランキング、2次元領域質問処理、地域ウェブ情報の応用に関して研究した結果をまとめたものであり、得られた成果は以下のとおりである。

- 1) 従来のウェブ検索システムで地理情報検索を行う場合は、地域情報を区別して扱う検索機能を備えていなかったため効果的な検索ができないという問題があったが、提案された手法ではウェブ・地図・地域知識を統合したモデルによって、多様な地域情報をウェブから得られた地域知識を元に、意味的に分かりやすくまとめて検索することが可能になった。
- 2) 地域情報検索目的に合わせた分類法とランキング手法を提案し、重複している検索結果を省くことやより人気がある結果を選択的に得ることが可能になった。分類では地域ウェブ情報における地名の性質とページの形式を考慮すべきであることに、ランキング手法では地域情報検索のために従来のリンク構造基盤のアルゴリズムを改良したことに意義がある。
- 3) 地図をウェブ検索システムに対する質問作成インターフェースとして用い、2次元領域質問によるウェブ検索システムを実現した。また、ウェブページの内容に対応する地域情報計算に対する高精度な測定手法を提案し、2次元領域質問処理の効率を向上させている。
- 4) 地域ウェブ情報を利用して都市のイメージを求める新しい方法を提案し、実際の地域ウェブページの内容に出現する近傍表現を用いて、ランドマークの成立条件、特性化、干渉問題について論じている。また、モバイル端末上での地域ウェブ情報利用におけるキャッシュ手法として、従来のキャッシュ手法と異なる有効な方法を示している。

以上、本論文は地域情報に特化する形で、地域ウェブ検索に対する新しい検索モデルを提案し、実際のシステムを構築しその有用性を検証しており、学術的にも実用的にも極めて有意義であると言える。

よって、本論文は博士（情報学）の学位論文として価値あるものと認める。

また、平成15年8月5日に実施した論文内容とそれに関連した試問の結果合格と認めた。