

モバイルサーチエンジン WithAir の試作と評価

4W-6

河合英紀 赤峯享 喜田弘司 松田勝志 福島俊一

NEC インターネットシステム研究所

1. はじめに

近年、携帯電話をはじめとするモバイル端末で閲覧可能な Web ページの数が急増しており、サーチエンジンへの要求が高まっている。

鮮度を保って大量のページを網羅的に検索するためには、人手に頼らずモバイルページのみを自動収集する必要がある。また、モバイル端末は入力が困難で画面も小さいため、ボタンの押下回数など操作コストを抑えた入力補助機能や効率のよいナビゲーション機能が必要である。さらに、モバイル端末での利用形態を考慮すると、出張や旅行など外出先の位置に関係した情報を検索する機能も重要である。

本稿では、(1)モバイルページのみを自動収集し、人気度でのランキング、及び地域情報への分類によって大量のページを検索可能にすること、(2)目的のページに到達するまでの操作コストを最小にすることを設計方針としてモバイルサーチエンジン WithAir を試作し、その基本性能と利便性を評価した結果を報告する。

2. WithAir の概要

モバイルサーチエンジン WithAir の画面例を図 1 に示す。本サーチエンジンには、次の 4 つの特長がある。

(1) モバイル専用クローラと高精度な全文検索機能

モバイル専用クローラは、「ページタイプ判別技術」[1]を用いて、ディレクトリ名や特定タグの有無、ページサイズなどを総合的に分析し、選択的にモバイルページを収集する。単純に全文検索するだけでは、結果が膨大になるため、リンク構造解析[2]で算出したページスコアで高精度なランキングを行い、リンク先の内容を端的に表現可能なページ要約文[3]を抽出して表示している(図 1 の画面(C)下部、画面(E)のリンク文字列)。

(2) 地域情報検索機能

地域情報検索機能は、モバイル専用クローラが収集したページに記述されている地名、市外局番、郵便番号などの位置表現を抽出し、その出現頻度に応じて

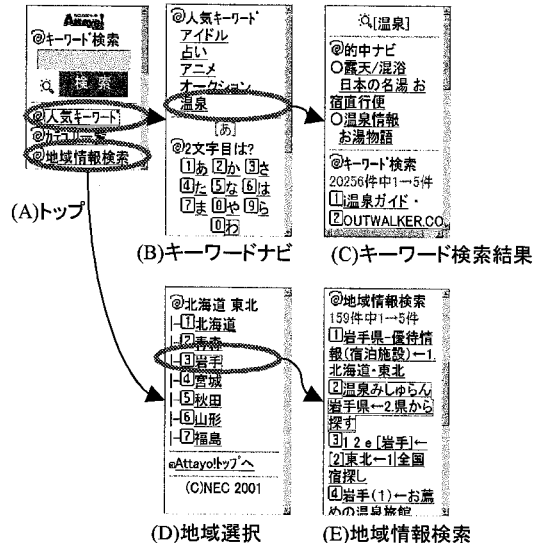


図 1 WithAir の画面例

の地域の情報を自動判別して地域別にインデックスを作成する[4]。検索時は、利用者が階層構造をたどって地名を選択すると(図 1 の画面(D))、システムがその地域のタウン情報や観光情報を優先的に検索して表示する(図 1 の画面(E))。

(3) キーワードナビによる入力補助機能

キーワードナビでは、「あ行」「か行」など 1 文字ずつキーワードの読みの行を入力することによって、その読みで始まる人気キーワードで検索できる。例えば、「温泉」というキーワードで検索する場合、従来のテンキー入力では、「入力モードに切り替え、1111(お)、000(ん)、3333(せ)、000(ん)、文字確定(おんせん)、変換、変換確定(温泉)、文字列確定、入力モードを解除、検索」と、テンキーと決定キーの押下(ストローク)が 22 回必要であったのに対し、キーワードナビでは「キーワードナビに切り替え、1(あ行)」の入力だけで図 1 の画面(B)のように人気キーワードの候補として「温泉」が現れるため 3 ストロークで検索できる。

(4) 的中ナビによる目的先読みナビゲーション機能

的中ナビは、入力された検索キーワードから利用者の典型的な検索目的を先読みして、人手で厳選した優

良サイトを提示する[5]。図 1 の画面(C)上部では、キーワード「温泉」の典型的な検索目的として「露天/混浴」などのこだわり検索ができるサイトと、宿泊施設などの一般的な「温泉情報」が得られるサイトを表示している。

3. 性能評価

モバイルサーチエンジンの基本性能として、モバイル専用クローラと地域情報検索機能の評価結果を示す。

(1) モバイル専用クローラ

今回は、モバイルページとしてiモードで閲覧可能なページを収集するようチューニングを行った。モバイル専用クローラの基本性能を表1に示す。適合率はiモードページと判断したページ数(1251件)中、実際にiモード端末で表示可能であったページ数(1246件)の割合である。再現率は、人手でピックアップしたiモードページ(1326件)に対して実際にiモードと判定できたページ(1169件)の割合である。

表 1 モバイル専用クローラの基本性能

収集ページ数	適合率	再現率
130万ページ	99%	88%

(2) 地域情報検索機能

位置情報抽出・検索モジュールの基本性能を表2に示す。モバイル専用クローラで収集したiモードページ130万ページのうち、37万ページが位置に関する情報を含んでいた。位置適合率は、ある特定地域の位置情報を含むと判別したページ(15件)のうち、判別が正しかったページ(14件)の割合である。また、位置再現率は、人手でピックアップした位置情報を含むページの集合(1957件)のうち、位置情報を含むと正しく判別できたページ(1480件)の割合である。さらに、地域情報適合率は、地域情報検索結果の上位10件に含まれているタウン情報や観光情報に関するページの割合である。

表 2 地域情報検索の基本性能

位置情報ページ数	位置適合率	位置再現率	地域情報適合率
37万ページ	93%	76%	91%

4. 利便性評価

検索結果のランキング方法や入力補助機能、ナビゲーション機能も含めた総合的な利便性の指標として、トップページから目的のページにたどり着くまでに必要なテンキーと決定キーの最小ストローク数を測定した。表3に「旅行先のホテル」「新幹線の時刻表」「アーティストの着メロ」など、モバイルでの典型的な検索課題に対して、

WithAir と、一般的な3種のモバイルサーチエンジン(OH!NEW?[6]、Yahoo!モバイル[7]、Infoseek[8])におけるキーワード検索とカテゴリ検索の最小ストローク数を比較した結果を示す。表3より、従来サービスでは、キーワード入力がない分、最小ストローク数はカテゴリ検索の方が少なかった。一方、WithAir では、キーワード検索でもカテゴリ検索と同程度の最小ストローク数であった。カテゴリ検索では、カテゴリ体系を熟知していないと、必ずしも最小ストロークで所望のページにたどり着けないことを考慮すると、WithAir によるキーワード検索が最も操作コストが低いといえる。

表 3 モバイル検索の最小ストローク数比較

	キーワード検索	カテゴリ検索
WithAir	4.4	5.0
OH!NEW?	26.2	5.6
Yahoo!	21.6	5.0
Infoseek	21.4	6.2

5. おわりに

モバイルサーチエンジンWithAirを試作し、基本性能と利便性の評価を行った。その結果、iモードページ自動判別で適合率99%、地域情報分類で適合率93%の精度を実現した。また、従来サービスと比較して最も操作コストが低いことを確認した。今後は、実サービス[9]上で、利用状況を測定しながら、評価・改良を進めていく。

参考文献

- [1]松田ほか、文書タイプ分類による問題解決向きWWW検索システムの開発と評価、情処研報 FI-53-2、1999年
- [2]高野ほか、サイテーションエンジン リンク解析を用いたWWWランキングシステム、情処研報 DBS-120-2、2000年
- [3]赤峯ほか、モバイル指向WWWサーチエンジンWithAirの開発(1) - システム構成 -, 情処 62 全大 6W-8、2001年
- [4]喜田ほか、モバイル位置指向サーチエンジンの開発、情処 61 全大 1U-2、2000年
- [5]河合ほか、モバイル指向WWWサーチエンジンWithAirの開発(2) - ナビゲーション機能 -, 情処 62 全大 6W-9、2001年
- [6]OH!NEW?, <http://ohnew.co.jp/i/>
- [7]Yahoo!モバイル, <http://mobile.yahoo.co.jp/>
- [8]Infoseek, <http://iseek.infoseek.co.jp/>
- [9]BIGLOBE サーチ Attayo!ケイタイ, <http://attayo.jp/i/>