

## モバイルユーザ向け情報選別配信技術

福島 俊一

NEC インターネットシステム研究所  
〒630-0101 奈良県生駒市高山町 8916-47

本論文では、モバイルサーチエンジン WithAir とモバイル情報配信エンジン TPOCAST の例を中心に、モバイルユーザ向けの情報選別配信技術を紹介する。携帯電話などのモバイル端末は、画面サイズが小さく、キー操作も不便であることから、ユーザが欲しい情報を簡単な操作で得られる情報選別技術が重要である。WithAir は、インターネット上の大量の i モードページを選択的に自動収集し、地域別に自動分類する技術、および、ユーザ不特定の状況のもとで、ユーザの検索語や目的を先読みして絞り込む技術に特長がある。TPOCAST は、ユーザが特定された状況のもとで、T(時間)・P(位置)・O(嗜好)を考慮してユーザと情報(コンテンツ)をマッチングする技術、および、ユーザの行動から T・P を考慮したユーザ嗜好(O)を学習する技術に特長がある。

### Information Retrieval and Filtering Technologies for Mobile Users

Toshikazu Fukushima

Internet Systems Research Laboratories, NEC Corporation

t-fukushima@cj.jp.nec.com

This paper introduces information retrieval and filtering technologies for mobile users and describes a mobile search engine WithAir and a mobile information filtering engine TPOCAST as examples. Information selection technologies are very important for the mobile users because such mobile appliances as cellular phones have small size of the screen and inconvenience of the key operation. WithAir has the advantages that it can automatically collect i-mode pages selectively from the Internet and classify them based on their regional information. It has another novel function of page navigation based on anticipating the search keyword and purpose of unspecified users. TPOCAST has the advantages that it has TPO based matching functions between a specified user and contents, where T means time, P means position, and O means preference. It can learn user's preference from his actions.

#### 1 はじめに

携帯電話、車載端末、PDA などのモバイル端末が普及し、i モードに代表されるような、モバイル端末からの情報アクセスが、日常的に利用されるようになってきた。モバイル端末は、PC と比較して、画面サイズが小さいことやキー操作が不便なことが、情報アクセスにおける障害となる。すなわち、モバイルユーザにとって、数個の検索語を入力することさえ煩わしく、小さい画面で検索結果リストの内容を 1 つ 1 つ吟味する作業もまた煩わしい。したがって、モバイルユーザ向けの情報提供サービス[1][2][3]のためには、ユーザが欲しい情報を簡単な操作で得られるようにする情報選別技

術がいつそう重要なものになる。

よりの確に情報を選別するためには、ユーザの情報要求・関心を、システム側がいかにして把握するかがポイントになる。また、ユーザを特定できる状況か、特定できない状況かに応じて、とり得るアプローチが異なってくる。

まず、ユーザを特定できない状況でのアプローチについて概観する。

一般のサーチエンジンが、このタイプの情報提供サービスに該当する。サーチエンジンにおいてユーザは通常、2 種類の方法で、情報を絞り込んでいくことができる。ジャンル分けされたディレクトリをたどる方法と、検索語・検索式を入力して条件を限定する方法である。Web ページのサーチエ

ンジンのみならず、i モードページなどを対象としたモバイル端末用の検索エンジン(例えば、OH!NEW?や Yahoo!モバイルなど)でも、同様の方法が用いられている。しかし、規模が大きくなってくると、小さい画面でディレクトリの全体構造を把握することが困難になり、無駄な試行錯誤的探索が増えてくる。また、検索語入力が煩わしいことは言わずもがなである。

このような問題に対して、ユーザの典型的な検索語や目的を先読みすることで、モバイル端末におけるキー操作の不便さを軽減するアプローチが考えられる。これを実現したシステムとして WithAir[4]がある。

次に、ユーザを特定できる状況でのアプローチについて概観する。

この場合、ユーザの属性、関心、位置情報などを取得することが可能となり、その都度、ユーザに検索語や検索式を入力させなくとも、ユーザが欲しがるであろう情報を自動的に・継続的に配信するサービス(いわゆるプッシュ型サービス)が提供できる。どのような条件を利用して情報を選別するかの違いによって、次のようなアプローチが考えられている。

- (a) ルールベース型：ユーザが自分の関心事あるいは欲しい情報の内容や配信タイミングなどの条件を事前に記述しておくタイプ。また、情報を提供する側が、個々の情報や情報集合に対して配信対象とするユーザの属性や位置・移動の条件を記述しておくタイプ。
- (b) 位置ベース型：配信対象の情報を位置と対応付けて用意・格納しておき、ユーザの位置をトリガとして、ユーザにその周辺位置に関わる情報を配信するタイプ。
- (c) 嗜好学習型：ユーザの行為(購買履歴など)からユーザの嗜好(関心事、欲しい情報の特徴など)をシステムが自動的に推定(学習)し、ユーザの嗜好にマッチする情報を選別配信するタイプ。

これらのうち(a)は、いかにきめ細かにルールを記述しておくかが、ユーザから見た情報の適合性を決定する。このルールのメンテナンスを人手で行うことは大きな労力を要する。また、(b)のような位置情報のみによる選別には限界がある。ユーザの周辺位置に関わる情報を単純に何でも配信すると、ユーザの関心にマッチしない不適合情報が多数送られることになってしまう。(c)はショッピングサイト

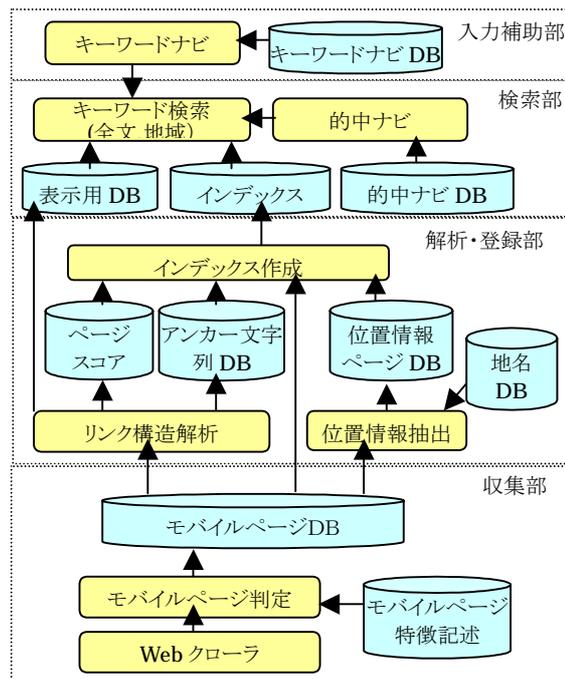


図1 WithAirのシステム構成

における商品レコメンデーションなどで実績のある手法であり、(a)で問題となったメンテナンス負荷も軽減できる。しかし、モバイルユーザへの情報配信を考えた場合、ユーザの位置情報・行動情報も活用することで、より高精度な情報選別を実現し得るはずである。

このような問題に対して、(a)(b)(c)をさらに発展させ、T(時間)・P(位置)・O(嗜好)に基づいた情報選別とユーザ嗜好の学習を実現したシステムが TPOCAST[6][7]である。

## 2 モバイル検索エンジン WithAir

ユーザ不特定の状況下でモバイルユーザ向けに情報選別機能を提供する代表的なシステムとして、河合・赤峯・喜田・松田・福島の開発した WithAir[4]を紹介する。

WithAir によれば、ユーザは、インターネット上の大量のモバイルページ(i モードページ)を対象にして、携帯電話から簡単な操作でキーワード検索が行える。

### 2.1 モバイルページの選択的収集・分類

WithAir の情報選別技術の第1の特長は、インターネット上の大量の i モードページを選択的



図2 WithAir での先読み検索の画面例

に自動収集し、地域別に自動分類することである。一般の Web ページのなかから i モードページを自動判別することで、大量で鮮度の高い情報の自動収集を実現している。

i モードページの自動判別には、ページタイプ判別技術[5]を応用した。ページサイズ、使用されている文字コード、タグの種類、ページ本文や URL 中のキーワードなどに着目して、i モードページらしさのスコアを計算する。

WithAir では、収集した i モードページに対して、さらにリンク構造解析と位置情報抽出を行うことで、ページスコアと地域別分類を求める(図1における収集部と解析・登録部)。リンク構造解析では、ページ間のリンク関係から各ページの人気の高さやリンク元キーワード(アンカー文字列)を求める。位置情報抽出では、ページ中に記述された住所、市外局番、郵便番号などの表現を抽出し、その出現頻度を考慮した上で、そのページがどこの地域に関わる情報であるかを判別する。

## 2.2 検索語・目的の先読み検索

WithAir の情報選別技術の第2の特長は、ユーザの検索語や目的を先読みすることで、キーワード検索時のユーザ操作数を削減できることである。ユーザを特定せずに、過去の検索語の使用頻度や、検索語の共起語の傾向(どのような語が同時に検索に用いられたか)を蓄積しておき、検索語や目的の先読みを利用する。

以下、図2の画面例を参照しながら、先読み検索の機能概要を説明する。

図2における「人気キーワード」のメニューが検索語先読み機能を意味する。この例では、「ま行」を選択することで、読みが「ま行」で始まる人気キーワード(よく使用される検索語)のリストが得られ

る。この例では「無料」から「モーニング娘。」までの5件が表示される。このなかにユーザが検索したいと考えていた検索語が含まれていれば、それを選択することで検索を実行できる。含まれていなければ、さらに2文字目の行を指定することで人気キーワードリストの絞り込みが行える。このように検索語の先頭数文字を選択するだけで検索が実行でき、携帯電話における検索語入力の手間を軽減できる。

さらに、検索結果は、「的中ナビ」という目的別リストと、通常タイプのキーワード検索結果リストという2種類に分けて表示される。「的中ナビ」が目的先読み検索機能を意味する。入力された検索語に応じて典型的な目的と各々の該当サイトが列挙される。図2の例では、「モーニング娘。」という検索語に対して「公式サイト」「着メロ」「CD」などが典型的な目的として表示されている。このような目的先読み機能によって、ユーザは、検索結果をさらに絞り込むために検索語を追加入力しなくてもいいという手間を軽減することができる。

## 2.3 モバイルサーチエンジンの実用化事例

ここで紹介した WithAir は「BIGLOBE サーチ Attayo ケイタイ」という名称で、既に実用サービスとして公開されている。携帯電話から attayo.jp という URL を入力することで利用できる。

前述の情報選別技術の評価に関しては、次のような結果が得られている[4]。まず、i モードページ判別精度は、適合率99%、再現率88%である。これによって i モードページの自動収集が可能となり、130万ページという規模の i モードページ収集が実現できた。従来のモバイルサーチエンジンでは人手で i モードページを選別登録していたため、規模が大きいものでも、登録件数は数万サイト程

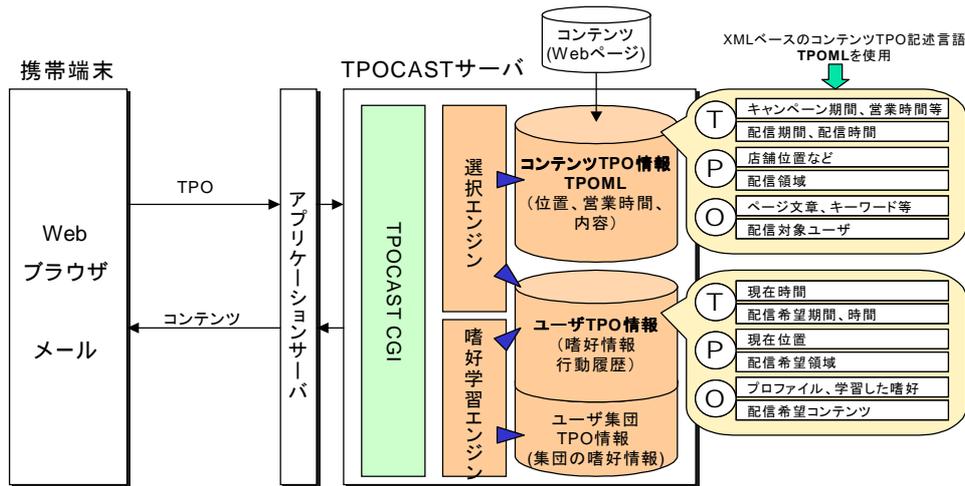


図3 TPOCAST を利用したシステム構成

度であった。また、先読み検索を利用することで、キーワード検索時のキー押下回数は平均 4.4 である。先読み検索なしの従来型と比較したところ、約 1/5 に軽減できていることが確認できた。

### 3 モバイル情報配信エンジン TPOCAST

ユーザを特定できる状況のもとでモバイルユーザ向けに情報選別機能を提供する代表的なシステムとして、二瓶・伊東・茶園の開発した TPOCAST[6][7]を紹介する。

TPOCAST によれば、ユーザは、レストラン情報、観光情報、店舗情報、広告などのいわゆる POI (Point of Interest) 情報を、そのとき、その場で、ユーザの好みにマッチした形で自動的に受け取ることができる。

#### 3.1 TPO に応じて情報を選別

TPOCAST の情報選別技術の第 1 の特長は、T(時間)・P(位置)・O(嗜好)を考慮してユーザと情報(コンテンツ)をマッチングすることである。TPOCAST(図 3 参照)では、このような TPO ベースのマッチングを実現するために、まず、各コンテンツ(POI)に対して、コンテンツ TPO 情報を登録しておく。コンテンツ TPO 情報は、POI に関する営業時間(T)、店舗位置(P)、内容説明文・キーワード(O)などの TPO 属性、および、配信対象ユーザを限定する TPO 条件を、XML ベースの言語 TPOML (Contents TPO Markup Language) で記述したものである。一方、ユーザに対しては、

ユーザの現在時間(T)、現在位置(P)、嗜好(O)や、ユーザの希望する配信条件などをユーザ TPO 情報として記録する。TPOCAST の選択エンジンは、ユーザ TPO 情報とコンテンツ TPO 情報を T・P・O の 3 軸で比較、スコアリングすることで、各ユーザの TPO にマッチしたコンテンツを選別する。

#### 3.2 ユーザの行動から嗜好を学習

TPOCAST の情報選別技術の第 2 の特長は、ユーザの行動から T・P を考慮したユーザ嗜好(O)を学習することである。TPOCAST では、ユーザの行動に対して、その時間(T)と、位置(P)と、行動に対応するコンテンツの特徴(キーワードやカテゴリ)と、行動の種類に応じたウェイトという 4 つ組を、ユーザ嗜好情報としてユーザ TPO 情報のなかに蓄積する。行動の種類としては、現状、「閲覧」「訪問」「購入」の 3 種類を区別している。「閲覧」はモバイル端末上でのコンテンツ表示、「訪問」はコンテンツに関連する場所の訪問、「購入」はコンテンツに関連する場所での購入を意味する。

なお、使い始めたばかりの時期や初めて訪れた土地などで、ユーザ嗜好の学習が進んでいない状態では、ユーザ嗜好情報としてユーザ集団 TPO 情報も併用する。ユーザ集団 TPO 情報は、あるユーザ層(例えば 20 代の女性など)の代表的な嗜好情報である。学習の進行に応じて、ユーザ集団 TPO 情報のウェイトを下げ、本来のユーザ嗜好情報のウェイトを高めていく。



図4 TPOCASTによる情報配信の画面例

### 3.3 TPOCASTによる情報配信サービス事例

TPOCASTの適用事例として、松江観光GIS実証実験[7]について紹介する。この実証実験は、国土交通省、日本観光協会が主体となり、GISを利用したIT利用の可能性を実証することを目的に、2000年12月末から2001年3月末まで松江市で実施されたものである。松江市を訪れた観光客に個人情報(氏名、性別、年齢など)を登録してもらった上でPDAを貸し出し、TPOCASTは、観光施設や飲食店など合計345件の観光情報を、ユーザ(観光客)のTPOを考慮してPDAに配信するサービスを提供した。プッシュ型の自動配信サービス(図4)に加えて、ユーザが望むタイミングで配信をリクエストするプル型サービスも提供した。プッシュ型のみならずプル型の場合にも、TPOを考慮した情報選別を行う。また、観光情報の閲覧、クーポンの使用、電子チケットの購入といった行動からユーザ嗜好を学習することも実施した。

実証実験では、TPOCASTによる観光情報配信(プッシュ、プル)のほかにも、歩行者ナビ、バス停案内、電子チケット、クーポン、スタンプラリーほか、様々なサービスが提供されたが、アンケート結果では、TPOCASTによる2つのサービスが満足度の1位・2位を占めるという評価が得られた。また、ログ分析では年齢層によって閲覧するコンテンツの傾向が異なり、ユーザ嗜好を考慮した情報配信の重要性が確認された。

## 4 おわりに

モバイルユーザ向け情報選別配信技術に関し

て、ユーザ不特定タイプのWithAirとユーザ特定タイプのTPOCASTを中心に紹介した。モバイル利用環境におけるユーザの情報要求・関心を、様々な視点からシステムに取り込んで活用することがポイントになる。ユーザの行動情報や実世界情報のセンシング技術の発展・普及に連動して、今後さらに、情報選別配信技術の研究発展と日常への普及が期待される。

## 参考文献

- [1] 島、位置情報流通のプラットフォーム、情報処理学会誌、Vol.42、No.4、特集：位置情報を利用したモバイルコンピューティング、2001年。
- [2] 市村・二瓶・坂田・茶園・倉島、位置情報サービス—位置情報を用いた通知サービスの発展へ向けて—、情報処理学会誌、Vol.42、No.12、特集：モバイルインターネット、2001年。
- [3] 時津・高橋、インターネットITSプロジェクト(実験編)、情報処理学会誌、Vol.43、No.4、特集：インターネットと自動車、2002年。
- [4] 河合・赤峯・喜田・松田・福島、モバイルサーチエンジンWithAirの試作と評価、情報処理学会研究報告、FI-64-11、2001年。
- [5] 松田・福島、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、情報処理学会研究報告、FI-53-2、1999年。
- [6] 伊東・二瓶・茶園、モバイル情報配信プラットフォームTPOCast、ケータイ・カーナビの利用性と人間工学シンポジウム、2001年。
- [7] 二瓶・茶園・伊東、モバイル情報配信プラットフォームTPOCAST—松江観光GIS実証実験結果報告—、情報処理学会第63回全国大会、2R-3・4、2001年。