

携帯端末におけるデータマイニング法

佐藤 寛和[†] 國井 利泰[†]

法政大学情報科学部デジタルメディア学科[†]

1, はじめに

近年, 日本において携帯電話は急速に普及し, その機能も単なる移動通話端末という領域を超え, Web アクセスが可能になるなどモバイル端末としての要素が増えてきている. これにより携帯端末から, Web 上の情報へのアクセスが可能となってきている. しかし, Web 上には膨大な情報が存在し, 演算能力や通信帯域の小さい携帯端末ではその情報量を十分に利用することは困難である. そこで, Web 上の膨大な量のデータから, いかにして有効な情報を取り出していくかが重要となる.

携帯端末とはその名の通り, 人々が常に携帯して, 特定の場所と時間において利用するものである. 携帯端末は一般的な PC 端末などと違い, 位置情報に大きく依存している. そこで, 位置情報をキーとして Web 上から情報を抽出し, 携帯端末の持つ位置情報と対応を取ることで, 携帯端末において効果的に情報を取得できる.

2, 携帯端末向け情報抽出

2, 1, 位置情報と Web データのマッチング方法

携帯端末の位置と Web 上の情報との対応付けは以下のステップで行う.

Step1: Web Page 上から, 住所・地名など位置に関するキーワードを探しそれを抽出する.

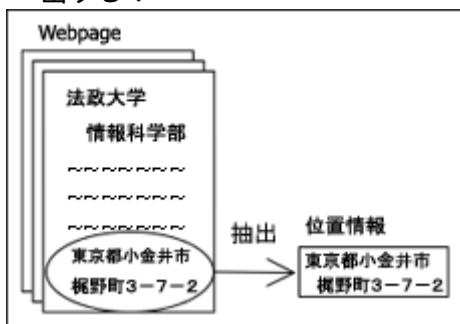


Fig1

Step2: Step1 で抽出した位置情報に対応するデータ部分を抽出して対応を取る. データ部を「法政大学情報科学部」とすると, 「法政大学情報科学部 = 東京都小金井市梶野町 3-7-2」という様に対応付けられる.

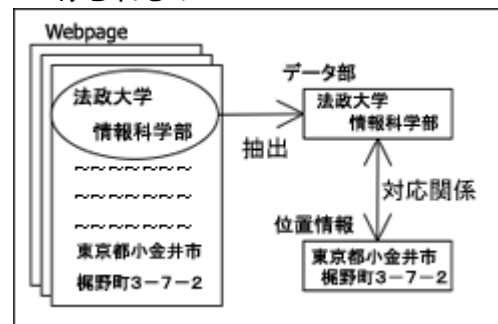


Fig2

Step3: 位置情報を GPS[1]情報に変換する. これにより, 「法政大学情報科学部 = 経度:139.30.23.75, 緯度:35.42.6.50」となる.

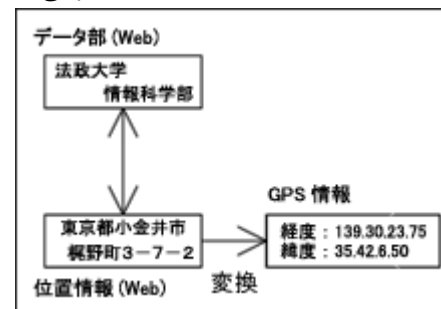


Fig3

Step4: 携帯端末の GPS 機能を利用し端末の位置情報を取得する. これにより, 「携帯端末 = 経度:139.30.23.75, 緯度:35.42.6.50」となる.

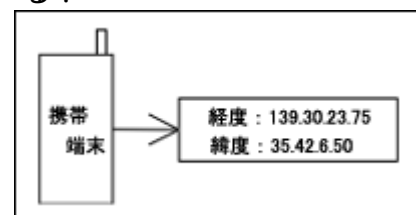


Fig4

On a Data Mining Method for Mobile Terminals
Hirokazu Sato[†], Tosiyasu L.Kunii[†]

[†]Department of Digital Media Science, Graduate School of Computer and Information Sciences, Hosei University

Step5: Web 情報と携帯端末を両者の位置情報によって対応付けることにより, Web 情報と携帯端末との対応を取ることが出来る.

「携帯端末 = GPS 情報 (Mobile) = GPS 情報 (Web 情報) = 法政大学情報科学部」となる.

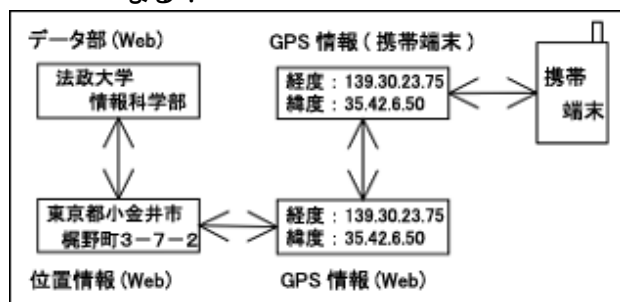


Fig5

このような 5 ステップを踏むことにより, 位置情報を接点として, 携帯端末の位置に関連した情報を Web 上から取得できる.

2, 2, HTML からの位置情報抽出

2, 1, の Step1 において Web 上から位置情報を抽出するが, HTML は一般的に平易な自然言語で記述されている. そこで, HTML 文書中から自然言語処理により, 住所など位置情報を取得する必要がある. 多くの場合, 地名は複合名詞になっているため, 複数の連続する名詞を抽出し, 地名の辞書を用いてマッチングすることで地名を抽出するという方法[2]がある. このように Web Page 中の文章の構文解析をすることで位置情報を抽出することが可能となる.

2, 3, GPS

GPS とは「Global Positioning System」の略で, 地球上空を周回する 24 個の衛星の位置を基準として, 地球上での 3 次元位置を検出する衛星測位システムである. GPS の民間利用に対しては, SA(Selective Availability)によって強制的に精度を落とされていたが, 2002 年 5 月に解除され大幅に精度が向上した. GPS はカーナビゲーションなどに利用されているが, 最近では GPS による測位機能が搭載された携帯端末も販売されている.

2, 4, 住所の GPS 変換

2, 1, の Step3 で, 住所を GPS 情報に変換している. このような, 空間情報を座標に変換する技術をアドレスマッチング技術[3]という. 日

本は欧米の様に住所体系に規則性がなく複雑なため, 住所と経度緯度の対応表を作成してアドレスマッチングを行う方法が一般的である.

3, 応用例

この手法の応用例としては, Web 情報を利用した周辺情報サービスがある. 現在, 携帯電話サービスの一つとして, 周辺の地域情報などを提供する情報サービスは存在している. しかしこれらの情報源は各サービス会社それぞれが保持するデータに限られている. そこで本手法を利用すると, Web 上の情報から携帯端末の位置に関係した情報を抽出することができるため, Web を情報源としたサービスが可能になる.

4, まとめ

本論文では, GPS 搭載携帯端末における効果的な Web 情報の抽出方法として, 携帯端末の位置情報をキーとして Web 上から情報を抽出する手法について提案した. これにより, 位置依存という携帯端末の特性を生かした情報抽出が可能となる.

今後の課題として,

- Web Page において, どの部分をデータ部として認識するのか
 - Web Page 上において, データ部分と位置情報部分の対応をどう取るのか
- がある. また, 情報源を Web としているため, Web 情報の意味論的階層の分析と抽出を進めている. Web 情報の分類指標として正確性のランキングも考慮中である.

6, 参考

- [1] 安田 明生, 「GPS の現状と展望」, 電子情報通信学会誌 Vol.82 No.12, pp.1207-1215 (2002)
- [2] 相良 毅, 有川 正俊, 坂内 正夫, ジオリアランス情報を用いた空間情報抽出システム, 情報処理学会論文誌, 41/SIG6(TOD7), pp69-80. (2000)
- [3] Karen C.Hanna, R.Brian Culpepper, GIS in Site Design, John Wiley & Sons, Inc. (1998)