

Micro media を用いたパーソナライズ情報収集支援システムの提案

Personalized Information-Gathering System Using Micro Media

○澤田瑞穂¹, 戸田健², 尾崎信耶³Mizuho Sawada¹, *Takeshi Toda¹, Shinya Ozaki²

Abstract: Information on WEB keeps expanding, and to examine the enormous amount of information closely in a past, active search strategy, time and the labor are required. In the present study, the user doesn't do an active retrieval, and it proposes the search engine that offers desired information. The proposal system extracts the retrieval word by understanding user's situation by doing the text mining that pays attention to everything from information contributed by the user to not only the noun but also the verb, and achieves personalizing the retrieval and the retrieval result.

1. はじめに

WEB が普及した今日では、ソーシャルネットワークサービス(Social Network Service : SNS)や Blog, Twitter 等のマイクロメディアを用いた一般ユーザによる情報発信が容易なものとなり、今まではあまり重要視されていなかった一般ユーザの発信する情報が重要な情報源として利用されるようになった。その結果、WEB 上には様々な情報が溢れている。こうした状況を踏まえ、情報取得の面から WEB を見ると、WEB 検索エンジンを利用したとき一度に取得される情報が膨大なものになるが、そこからユーザが情報を取得するスキルはまだ不足している。これは明確なキーワードが見つからずに所望の情報が得られないためである。

そこで、本提案システムでは以上の観点を踏まえ、日常生活動作の一部となっている WEB 検索の時間をかけないで即座に有用な情報の取得を実現することでユーザの行動意欲の明確化を実現することを目的とする。また、既存のレコメンドサービス等で行う所望の情報元の登録作業を排除することで、パーソナル情報の推移にも十分対応できると考えられる。

提案システムでは、ユーザによってマイクロメディアに投稿された情報から、名詞だけでなく動詞に注目したテキストマイニングを行うことでユーザの状況を把握する事により検索語を抽出し、同時に検索結果の推測を行い、検索及び検索結果のパーソナライズを実現する。

2. 関連研究

現在、WEB 上において様々なレコメンドサービスの開発が進んでいる。各種商品購買サイトにおける商品の表示履歴や購入履歴から同様な購入履歴を持つ

ユーザの履歴に残った商品を薦める、「商品推奨サービス」^[1]をはじめ、Blog 内広告の内容を Blog の内容から自動的に選択する Yahoo Japan の「インタレストマッチ」^[2]が上げられる。また携帯端末上での情報レコメンドサービスとして、NTT DOCOMO による「i コンシェル」^[3]がある。しかし、これらのサービスでは、過去のユーザの所望の情報から提供情報を選択し提供するために、情報の新規性が失われている場合が考えられる。また、ユーザによる情報提供サイトの登録によって新規情報を得る事ができるが、ユーザ自身の所望の情報は常に変化しているものと考えられ、登録した情報配信が常に所望のものであるという事はない。

3. 提案システム

3.1 機能構成

提案システムの機能構成を図 1 に示す。なお、提案システムでは、ユーザの日常動作を重要視するので、普段から使用しているソーシャルメディアを従来どおり使用する。提案システムの機能構成は、図 1 より A. 外部ソーシャルメディアシステム、B. 情報集支援システムに大別される。ユーザは、あらかじめ端末上から自身が使用しているソーシャルメディアの URL を B. 情報収集支援システムに登録しておく。次に、システム構成要素の概要を説明する。

3.1.1 外部ソーシャルメディア

Twitter や Blog, SNS 等従来ユーザが使用していたものを従来通り用いる。

3.1.2 パーソナライズ情報収集支援システム

サーバ上のパーソナライズ情報収集支援システムは以下(a)~(d)の 4 つの部分から成る。

(a) 新着投稿情報取得部, (b) 形態素解析部, (c) パー

1 : 日大理工・学部・電気 2 : 日大理工・教員・電気 3 : 日大理工・院(前)・電気

ソナライズランキング部, (d) 情報検索部, (e) 表示インタフェース部

3.2 動作の流れ

前節で示したシステムの実施例を述べる. ユーザがあらかじめシステムに登録した外部サイトに記事を投稿するとサーバ上のが情報の取得を判断し, 外部ソーシャルメディアから新着記事の文字情報を取得, 形態素解析処理をする. 処理結果より得られた名詞+動詞に着目し, 重み付けを行い, 検索語の抽出を行う. 検索語と過去の関連語を元に外部 WEB 検索エンジンにて情報検索をし, 得られた結果を再びパーソナライズ情報収集支援システムで処理し, 上位数件を端末に送り端末上に表示する.

3.3 パーソナライズランキング部

本研究において最も重要なパーソナライズ部について述べる. パーソナライズランキングは, パーソナライズ機能としての働きをする. このランキングでは独自にいくつかの動詞をあらかじめ選出しておき, 重みを与えておく. ユーザのソーシャルメディア上での発言から抽出された動詞についてあらかじめ選出されている動詞とマッチしたときに, マッチした抽出された動詞に伴って抽出されている名詞に選出されている動詞に与えられている重みを与えることで, その重みに応じて名詞のランキングを作るというものである. ユーザが〇〇コーヒーが好きだ. と発言すると, システムが発言から動詞+名詞を抽出する. あらかじめ独自に選出してある 5 の重みを持った「好き」という動詞と, 発言の「好きだ」という動詞がマッチしたことで, 抽出された名詞「〇〇コーヒー」は 5 の重みを与えられ, 名詞情報がランキングとして残る. このランキングを元に検索語の決定や検索結果のフィルタリングを行う. また, このシステムはユーザのリアルタイムの発言で変化するランキングである. このため, 常に変化するパーソナル情報に対応することができると考えられる.

4. 評価方法

提案システムを実際に試作し, 実環境における評価を通して本用途における提案システムの有効性について検証を行う.

5. おわりに

本稿では, マイクロメディアを用いたユーザのリアルタイムな投稿から独自に検索語のパーソナライズを行いユーザに意識させない形で WEB 検索をする

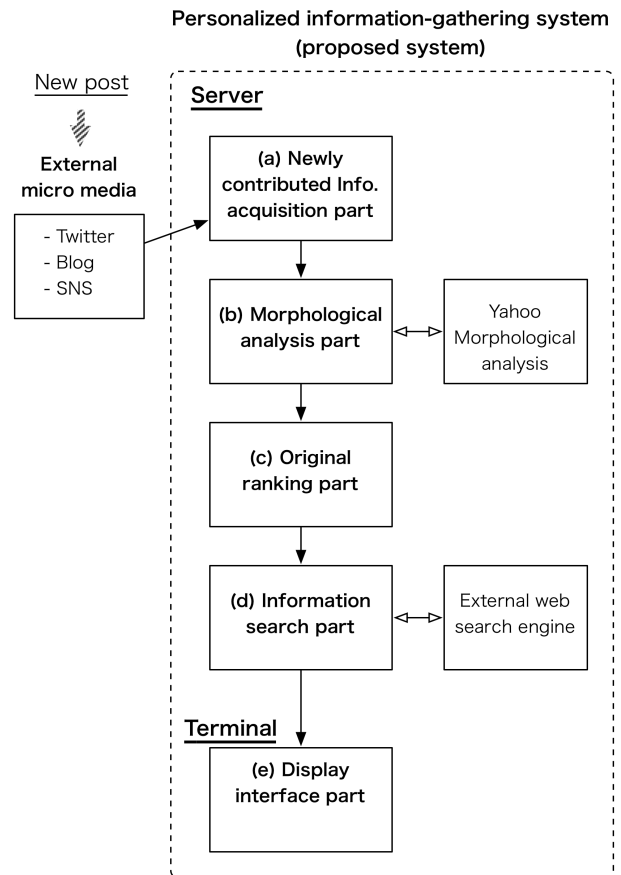


Figure.1 Original Ranking

ことで, 時間と労力がかかるようになってきた従来の WEB 検索の方法を解決する方法を提案した. 今後は, 実際にシステムを試作し, 実験を通してシステムの最適化を図る. また, 今後はシステムを発展させて発行情報に付与される GPS 位置情報からユーザの行動を分析することや, 発言時間に合わせた情報を提供することも検討し, より高度なパーソナライズ情報収集支援システムを提案する予定である.

6. 参考文献

- [1] Amazon.com:おすすめ商品;
<http://www.amazon.co.jp/gp/help/customer/display.html?nodeId=779360> (2011)
- [2] 株式会社エヌ・ティ・ティ・ドコモ:i コンシェル;
<http://www.nttdocomo.co.jp/service/customize/iconcier/> (2011)
- [3] ヤフー株式会社:インタレストマッチ;
<http://listing.yahoo.co.jp/service/int/index.html?o=IM0024> (2011)