

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 22 日現在

機関番号：34304

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500140

研究課題名(和文)ブロガーの熟知度分析およびセンチメント分析とその応用に関する研究

研究課題名(英文)Studies and Applications based on Analysis Bloggers' Knowledge Level and their Sentiments

研究代表者

中島 伸介 (NAKAJIMA, Shinsuke)

京都産業大学・コンピュータ理工学部・准教授

研究者番号：90399535

交付決定額(研究期間全体)：(直接経費) 3,600,000円、(間接経費) 1,080,000円

研究成果の概要(和文)：本研究では、精度および詳細度の高いブロガー分析手法の開発を行った。さらに、情報発信者および情報受信者としてのブロガー分析に基づく、信頼性の高い情報の判定技術、トレンド情報の抽出技術、情報発信者のプロファイリング技術とこれに基づく情報推薦技術の開発に取り組んだ。成果としては、研究協力を頂いている企業において、本研究の成果の一部を実サービスにて活用していただくなど、技術の実用化という意味でも成果を上げている。特に、流行トレンドの早期発見方式に関する研究では、難関国際会議での論文採択や推薦論文としての学会論文誌への採択を受けると共に、国内研究会発表にて3度受賞するなど、学会においても高く評価された。

研究成果の概要(英文)：In this study, we developed blogger analysis methods with high level of detail and accuracy. In addition, we also developed methods for discovering high credible information, extracting early-stage of trend information, profiling bloggers as sender of the information, and recommendation technique based on taking advantage of such methods. As the result, a part of our outputs was adopted in a real service of the collaborator company. Specifically, research papers of our studies about extracting early-stage of trend information were not only accepted as a competitive international conference paper and a journal paper and also awarded three prizes in domestic research workshops.

研究分野：総合領域

科研費の分科・細目：情報学、メディア情報学・データベース

キーワード：情報システム Webマイニング ブログマイニング 流行語発見

1. 研究開始当初の背景

Twitter 等を含めたブログコンテンツが広く普及したことにより、誰もが簡単に情報を発信および受信し、相互にインタラクションすることが可能となった。これにより、マスメディアから一般ユーザに流れるだけであった従来の情報の流れを大きく変化させてしまった。このように世論を反映したブログは、もはや単なるコミュニケーションツールではなく、マーケティングやトレンド分析および企業製品の評判情報調査を行うための貴重なデータとなっている。このように利用価値の高いブログ情報であるが、どのような特性のユーザが発信した情報なのかを明らかにすることにより、その情報の重要性や位置付けを判断することが可能となり、その利用価値はさらに高まる。例えば、株式に詳しい人のブログ記事は、そうでない人の記事よりもその価値は高いといえる。すなわち、対象トピックに関して詳しい人を予め見極めておくことで、そのトピックに関して価値の高い情報を見つけやすくなる。申請者はこれまでの研究にて、株式会社きざしカンパニーとの連携により、各ブロガーがどのような分野に詳しいのかということを表す“熟知度”の算出と共に、どのような感情(センチメント)で表現したのかを分析することに基づくブログランキングシステムの開発を行ってきた(図1参照)(稲垣ら、情報処理学会論文誌：データベース、Vol.3 No.3)。



図1：熟知度に基づくブログランキング

データ数としては、2010年3月10日現在で、約7,422,000名のブロガーにより投稿された、約173,715,000ブログエントリのデータを対象としている。すなわち、大量の実データに対しても対応可能な分析手法を開発済みである。しかし、分析精度の向上という面でまだまだ改良の余地があること、このブロガー分析技術を応用したアプリケーション開発という意味ではほとんど手が付けら

れていないことを踏まえて、さらに研究を進展させる必要があると考えた。

2. 研究の目的

上述の研究の背景を踏まえ、本研究課題の申請時当初の研究目的を以下に示す。

(1) 精度および詳細度の高いブロガー分析手法を開発する。熟知度を算出可能な領域数を拡充すると共に、熟知度およびセンチメント分析精度の更なる向上を目指す。

(2) 情報発信者としてのブロガーの熟知度やセンチメント分析結果に基づいた高信頼性情報の判定方式の開発を行う。情報発信者分析に基づく精度の高いランキングと、高信頼性情報の判定手法、および希少価値の高いニッチトレンドの分析技術の開発を行う。

(3) 情報受信者としてのブロガーの熟知度やセンチメント分析結果に基づいた情報推薦方式の開発に取り組む。熟知度やセンチメントの分析により、そのブロガーの関心や興味等の嗜好情報を抽出することが可能となる。このブロガー分析に基づいた、ブロガーを対象とした新規性の高い情報推薦方式の開発を行う。

3. 研究の方法

(1) 精度および詳細度の高いブロガー分析手法を開発

ブロガーが書いた過去のブログ記事を解析することに基づく熟知度分析手法の検証実験結果を踏まえ、精度および詳細度の高いブロガー分析手法の設計を行うと共に、分析システムを開発し、実験および考察を行う。その上で、分析対象ブロガーを“専門家”や“熱狂的マニア”といったブロガーのタイプを分類する手法についても実現を目指す。

(2) 情報発信者としてのブロガー分析に基づく情報システムの開発

情報発信者としてのブロガーの熟知度やセンチメント分析結果に基づいた情報システムとして、ブロガーおよびブログ記事のランキング、信頼性の高い情報の判定方式の開発に取り組む。さらに、このブログランキングシステムの実証実験と、熟知度とセンチメント分析結果に基づく信頼性の高い情報の判定方式の実装および実験に基づく検証を行う。その後、アルゴリズムの改良を行った後に、実用システムを想定した実装に取り組む。加えて、特定のコミュニティのみで盛り上がっているよう希少価値の高いニッチなトレンド分析技術の開発を行う。

(3) 情報受信者としてのブロガー分析に基づく情報システムの開発

各ブロガーの特性分析を行うシステム(blogram.jp)を活用し、これを用いた情報推薦方式の検討を行う。具体的には、特性分析システムにて得られたブロガーのプロファイルデータの適切な表現方法について検

討する．コンテンツベースフィルタリングおよび協調フィルタリングに基づく情報推薦実験システムを構築する．さらに実験に基づいてブロガーのプロファイルデータの妥当性や問題点について検討する．

4．研究成果

(1) 精度および詳細度の高いブロガー分析手法を開発

ブロガー熟知度を算出可能な領域数を拡充すると共に，熟知度算出時に使う関連語辞書の改良，およびセンチメント分析精度の向上を実現した．これにより，これまでよりも精度および詳細度の高いブロガー分析を実現した．なお，領域数としては約 11,000 の小領域と，これらをまとめた 122 の大領域における熟知度分析が可能となった．なお，ブログデータは常時収集および解析を行っており，今後も引き続き解析および手法の改良を継続する予定である．本技術の一部は，研究協力者である株式会社きざしカンパニーが運営する Web サイト blogram.jp にて活用されており，まだ限定的ではあるが，本技術の実利用を実現している．

(2) 情報発信者としてのブロガー分析に基づく情報システムの開発

流行語の早期発見方式および流行語を早期に発言するブロガー発見方式の開発において，新規性が高く画期的な手法を提案した．執筆した論文が難関国際会議の 1 つである ACM Hypertext2012 に採択されると共に，学術論文誌に採択された．また，国内の研究会発表において合計 3 度受賞するなど，学会においても高く評価された．今後も引き続き実用化に向けた取り組みを推進していく予定である．

(3) 情報受信者としてのブロガー分析に基づく情報システムの開発

情報受信者のコンテキストや振る舞いに応じた推薦手法についても検討すると共に，応用システムの 1 つとして，投稿型レシピ検索システムを取り上げ，新たなレシピ推薦手法の開発にも取り組み，評価実験を通じて提案手法の有効性を確認した．このように，ブログに限らず，ユーザの投稿型 Web コンテンツにおいては本研究課題にて開発された技術を応用することが可能である．最近では，ブログに限らず，Twitter や SNS，商品レビューデータなど，一般ユーザが情報を発信可能な環境が爆発的に拡大しており，我々が行った情報発信者および情報受信者分析も基づいたサービスの可能性は非常に高いと考えており，本研究課題の意義は非常に高いと考えている．

5．主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 2 件)

1. 中島伸介，張 建偉，稲垣陽一，中本レン，(WebDB11 推薦論文) 大規模なブログ記事時系列分析に基づく流行語候補の早期発見手法，情報処理学会論文誌データベース (TOD56)，6(1),1-15 (2013-01-23)，2013 年 1 月。(査読有)
<http://ci.nii.ac.jp/naid/110009514611>
2. Shinsuke NAKAJIMA，Adam JATOWT，Yoichi INAGAKI，Reyn NAKAMOTO，Jianwei ZHANG，Katsumi TANAKA: 'Finding Good Predictors in Blogosphere Based on Temporal Analysis of Posting Patterns', DBSJ Journal, Vol.10, No.1, pp.13-18, June 2011. (査読有)
http://dbsj.org/journal/dbsj_journal/dbsj_journal_vol_10_no_1_13_18/

〔学会発表〕(計 25 件)

1. Seiya Tomonaga，Shinsuke Nakajima，Yoichi Inagaki，Reyn Nakamoto，Tsukasa Ogura，and Jianwei Zhang, Measurement of Bloggers' Buzzword Prediction Ability Based on Analyzing Frequency of Early Mentions of Past Buzzwords, International MultiConference of Engineers and Computer Scientists 2014, ICICW2014 pp462-467, Hong Kong, China, March 12th, 2014.
2. Mayumi Ueda，Syungo Asanuma，Yusuke Miyawaki，and Shinsuke Nakajima, Recipe Recommendation Method by Considering the User's Preference and Ingredient Quantity of Target Recipe, International MultiConference of Engineers and Computer Scientists 2014, ICICW2014 pp519-523, Hong Kong, China, March 12th, 2014.
3. 関 匠吾，奥 健太，中島伸介，ユーザモデル・アイテムモデルの統合に基づく Context-Aware 推薦システムの実験的評価，第 6 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2014) B8-5，淡路市，兵庫県，2014 年 3 月 5 日．
4. 朝永聖也，中島伸介，稲垣陽一，中本レン，張 建偉，流行先読みブロガー発見のための流行語分析手法，第 6 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2014) C3-4，淡路市，兵庫県，2014 年 3 月 3 日．
(DEIM2014 優秀インタラクティブ賞)
5. 並木賢太郎，朝永聖也，中島伸介，稲垣陽一，中本レン，田中昇太郎，張 建偉，ブログ投稿数の時系列変化を考慮した対象キーワードの影響度分析手法の提

- 案, 第6回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2014) P1-3, 淡路市, 兵庫県, 2014年3月3日.
6. 寺田有佑, 朝永聖也, 小倉 僚, 中島伸介, 稲垣陽一, 中本レン, 張 建偉, 対象トピックとの関連度の高いプロガーコミュニティ判定方式の実験的評価, 第6回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2014) P1-4, 淡路市, 兵庫県, 2014年3月3日.
 7. 朝永聖也, 中島伸介, 稲垣陽一, 中本レン, 小倉 僚, 張 建偉, 流行語に対する早期言及頻度分析に基づくプロガー先読み度判定手法の提案, 情報処理学会 第158回データベースシステム研究発表会, Vol.2013-DBS-158 No.1, 京都大学, 京都府, 2013年11月26日.
(情報処理学会DBS研究会 学生奨励賞)
 8. 関 匠吾, 奥 健太, 中島伸介, ユーザ・アイテム統合モデルに基づくContext-Aware 推薦システムの開発と評価, 第5回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2013) B10-3, 福島, 2013年3月5日.
 9. 宮脇佑介, 河合由起子, 中島伸介, 閲覧ウェブページ内の興味範囲推定に基づくユーザ嗜好情報の抽出, 第5回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2013) B10-1, 福島, 2013年3月5日.
 10. 小倉 僚, 朝永聖也, 中島伸介, 張 建偉, 稲垣陽一, 中本レン, 対象トピックへの関心度の高いプロガーコミュニティ判定方式, 第5回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2013) P1-2, 福島, 2013年3月4日.
 11. 浅沼駿佑, 中川明莉沙, 宮脇佑介, 上田真由美, 中島伸介, 食材の嗜好と使用分量を考慮したレシピ推薦システム, 第5回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2013) P3-4, 福島, 2013年3月3日.
 12. 朝永聖也, 中島伸介, 張建偉, 稲垣陽一, 中本レン, 流行語の事前言及頻度分析に基づくプロガー先読み度判定手法の提案, 第5回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2013) C1-2, 福島, 2013年3月3日.
(DEIM2013 優秀インタラクティブ賞)
 13. Shinsuke Nakajima, Jianwei Zhang, Yoichi Inagaki and Reyn Nakamoto. Early Detection of Buzzwords Based on Large-scale Time-Series Analysis of Blog Entries, 23rd ACM Conference on Hypertext and Social Media (ACM Hypertext 2012), pp.275-284, Milwaukee, WI, USA, June 27th, 2012.
 14. Jianwei Zhang, Yukiko Kawai, Tadahiko Kumamoto, Shinsuke Nakajima, and Yuhki Shiraishi, "Diverse Sentiment Comparison of News Websites over Time", Proc. 6th International KES Conference on Agents and Multi-agent Systems - Technologies and Applications (KES AMSTA 2012), LNAI 7327, pp. 434-443, Dubrovnik, Croatia, June 26th, 2012.
 15. 朝永聖也, 中島伸介, Adam JATOWT, 稲垣陽一, Reyn NAKAMOTO, 張 建偉, 田中克己. ブログ記事の時系列分析に基づくプロガー先読み度分析手法の提案. 第3回ソーシャルコンピューティングシンポジウム (SoC2012), SoC2012 講演論文集 pp.79-84, 東京, 2012年6月23日.
 16. 濱岡祐美, 上田真由美, 中島伸介. コスメアイテム口コミサイト構築のための種類別の評価観点の抽出手法の提案. 電子情報通信学会 第22回Webインテリジェンスとインタラクション研究会, WI2-2012-15, pp.45-46, 東京, 2012年3月17日.
 17. 朝永聖也, 中島伸介, Adam JATOWT, 張 建偉, 稲垣陽一, Reyn NAKAMOTO, 田中克己. 投稿パターンの時系列分析による先読みプロガー発見手法. 電子情報通信学会 第22回Webインテリジェンスとインタラクション研究会, WI2-2012-12, pp.39-40, 東京, 2012年3月17日.
 18. 関 匠吾, 奥 健太, 張 建偉, 中島伸介, ユーザモデル・アイテムモデルの統合に基づくContext-Aware 推薦システム, 第4回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2012) A4-1, 神戸, 2012年3月4日.
 19. 中川明莉沙, 上田真由美, 高畑麻理, 中島伸介, 好き嫌いラベル付き食材分量を考慮したレシписコア算出方式, 第4回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2012) A1-2, 神戸, 2012年3月3日.
 20. 宮脇佑介, 張 建偉, 河合由起子, 中島伸介, Web 閲覧コンテキストおよびマウスイベントを考慮したユーザプロファイル抽出方式の提案, 第4回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2012) A1-1, 神戸, 2012年3月3日.
 21. 上田真由美, 高畑麻理, 中島伸介. レシピ閲覧・摂食履歴を用いた嗜好の抽出. Web とデータベースに関するフォーラム (WebDB Forum 2011), 情報処理学会シンポジウムシリーズ ,3G-1-2, 東京, 2011年11月5日.
 22. 中島伸介, 稲垣陽一, 山岡千夏, 張 建偉, ナカモト レン, 桑原 雄. 大規模

なブログ記事時系列分析に基づく流行語候補の早期発見手法の検討．Web とデータベースに関するフォーラム (WebDB Forum 2011), 情報処理学会シンポジウムシリーズ, 1G-1-1, 東京, 2011 年 11 月 4 日．

23. Mayumi UEDA, Mari TAKAHATA, and Shinsuke NAKAJIMA, User's Food Preference Extraction for Personalized Cooking Recipe Recommendation, 2nd Workshop on Semantic Personalized Information Management: Retrieval and Recommendation (SPIM 2011), pp98-105, Bonn, Germany, October 24th, 2011.
24. 宮脇佑介, 中島伸介, 張 建偉, 河合由起子, ユーザの Web インタラクション履歴の分析に基づくソーシャルリレーション発見手法の提案, 平成 23 年度 情報処理学会関西支部 支部大会, 大阪, C-02, 2011 年 9 月 22 日．
25. 関 匠吾, 張 建偉, 中島伸介, ユーザコンテキストを利用したユーザ嗜好及びアイテム利用パターン分析に基づく情報推薦システム, 平成 23 年度 情報処理学会関西支部 支部大会, C-05, 大阪, 2011 年 9 月 22 日．

6. 研究組織

(1) 研究代表者

中島伸介 (NAKAJIMA, Shinsuke)

京都産業大学・コンピュータ理工学部・准教授

研究者番号：90399535

(2) 研究分担者

張 建偉 (ZHANG, Jianwei)

筑波技術大学・産業技術学部・助教

研究者番号：20635924

(平成 24 年度より研究分担者)