

様 式 Z - 7

平成 2 6 年度科学研究費助成事業 実績報告書（研究実績報告書）

1. 機関番号

1	4	6	0	3
---	---	---	---	---

 2. 研究機関名 奈良先端科学技術大学院大学
3. 研究種目名 基盤研究(B) 4. 研究期間 平成 2 4 年度 ~ 平成 2 7 年度
5. 課題番号

2	4	3	0	0	0	5	7
---	---	---	---	---	---	---	---
6. 研究課題名 高次元・構造化データに適したリンク解析的類似度尺度の研究

7. 研究代表者

研究 者 番 号	研究 代 表 者 名	所 属 部 局 名	職 名
9 0 3 1 1 5 8 9	シンボ マサシ 新保 仁	情報科学研究科	准教授

8. 研究分担者

研究 者 番 号	研究 分 担 者 名	所属研究機関名・部局名	職 名
3 0 4 6 7 6 9 1	ハラ カズオ 原 一夫	国立遺伝学研究所・生命情報研究センター	特任研究員

9. 研究実績の概要

本年度の主な研究内容・成果は以下のとおりである：

- 重み付き全域森に基づくグラフ密度尺度 (sum-over-forests density index) を開発した。本手法は全域木 (spanning tree) に基づく密度尺度計算法の拡張であり、この手法に、物理学の枠組みに倣って温度パラメータを導入することで、森の重み付けを用途に応じて制御可能である。また、全域森定理 (spanning forest theorem) を応用し、closed form によって計算可能という好ましい特徴がある。
- 対訳抽出タスクにおいて、ハブ (hub) が出現し、精度を悪化させていることを明らかにした。また、同タスクにおいてグラフスパース化がもたらす精度向上は、ハブが抑制されることによってもたらされることがわかった。また、前年度開発したセンタリングが、同タスクに対して有効であり、最新のグラフ伝搬に基づく対訳抽出よりも高精度であることを実証した。
- 高次元空間のみならず、多数の事例を持つデータセットにおいてもハブが出現することを明らかにした。このようなハブは、やはり k 近傍分類 (k-nearest neighbor classification) タスクの精度を悪化させるが、センタリングでは削減することができないことがわかった。このため、昨年開発した重み付きセンタリング法を発展させ、その抑制に有効な新しい手法を開発した。

10. キーワード

- (1) データマイニング (2) リンク解析 (3) 類似度 (4) 高次元データ
- (5) (6) (7) (8)

（注）・印刷に当たっては、A 4 判（縦長）・両面印刷すること。

（ 1 / 5 ）

11. 現在までの達成度

(区分) (3) やや遅れている。

(理由)

本課題の研究成果を含む、リンク解析法に関する書籍を出版予定であるが、海外共著者との調整・出版社との交渉で遅れが生じたため、出版年が 2015 年に延期された。しかしながら、契約自体は成立しており、現在は執筆についてもほぼ完成し、微細な加筆修正・索引・文献一覧の作成・整理と校閲を残すのみである。

それ以外の研究については、ほぼ計画どおりである。

12. 今後の研究の推進方策 等

(今後の推進方策)

書籍原稿に加筆修正・校閲を行い脱稿する予定である。

(次年度使用額が生じた理由と使用計画)

(理由)

本課題の研究成果を含む、リンク（ネットワーク）解析法に関する書籍を出版を企画し・執筆中であるが、海外共著者（本課題連携研究者）との調整および出版社との交渉で遅れが生じたため、出版年が 2015 年に延期された。これにともない、予定していた校閲費用に剰余が生じた。

また、最近のディープラーニングとそれにともなうベクトル表現学習技術の急速な発展にともない、本課題が対象にするベクトル空間上の類似度尺度についても、その影響調査の必要が生じた。

(使用計画)

書籍原稿の校閲と、共同研究者との打ち合わせ、および調査に用いる。

13.研究発表(平成26年度の研究成果)

(雑誌論文) 計(1)件 うち査読付論文 計(1)件

著 者 名		論 文 標 題			
Mathieu Senelle, Silvia Garcia-Diez, Amin Mantrach, Masashi Shimbo, Marco Saerens, Francois Fous		The sum-over-forests density index: Identifying dense regions in a graph			
雑 誌 名		査読の有無	巻	発行年	最初と最後の頁
IEEE Transactions on Pattern Analysis and Machine Intelligence		有	36	2014	1268-1274
掲載論文のDOI(デジタルオブジェクト識別子)					
doi:10.1109/TPAMI.2013.227					

(学会発表) 計(6)件 うち招待講演 計(0)件

発 表 者 名		発 表 標 題	
重藤 優太郎, 新保 仁, 松本 裕治		ベクトルのスパース化を用いたk近傍法におけるハブの軽減	
学 会 等 名		発 表 年 月 日	発 表 場 所
情報処理学会研究報告 第216回自然言語処理研究会 第101回音声言語情報処理研究会 合同研究会		2014年05月22日	東工大大岡山キャンパス (東京都目黒区)

発 表 者 名		発 表 標 題	
濱口 拓男, 新保 仁, 松本 裕治		条件付きロジスティック分布を用いた重み付き多タスク学習	
学 会 等 名		発 表 年 月 日	発 表 場 所
情報処理学会研究報告 第216回自然言語処理研究会 第101回音声言語情報処理研究会 合同研究会		2014年05月22日	東工大大岡山キャンパス (東京都目黒区)

発 表 者 名		発 表 標 題	
Kazuo Hara, Ikumi Suzuki, Kousaku Okubo, and Isamu Muto		Annotating cohesive statements of anatomical knowledge toward semi-automated information extraction	
学 会 等 名		発 表 年 月 日	発 表 場 所
7th International Conference on Knowledge Discovery and Information Retrieval		2014年10月21日～2014年10月24日	Rome, Italy

発 表 者 名		発 表 標 題	
濱口 拓男, 新保 仁, 松本 裕治		文脈自由文法に基づいた復元的ニューラルネット	
学 会 等 名		発 表 年 月 日	発 表 場 所
第17回情報論の学習理論ワークショップ		2014年11月16日～2014年11月19日	名古屋大学 (愛知県名古屋市)

発 表 者 名		発 表 標 題	
Kazuo Hara, Ikumi Suzuki, Masashi Shimbo, Kei Kobayashi, Kenji Fukumizu, Miloš Radovanović		Localized centering: Reducing hubness in large-sample data	
学 会 等 名		発 表 年 月 日	発 表 場 所
29th AAAI Conference on Artificial Intelligence		2015年01月25日～2015年01月30日	Austin, Texas, USA

発 表 者 名		発 表 標 題	
小林 雄太, 新保 仁, 松本 裕治		自由エネルギー距離によるハブの抑制	
学 会 等 名		発 表 年 月 日	発 表 場 所
言語処理学会第21回年次大会		2015年03月17日～2015年03月20日	京都大学 (京都府京都市)

〔図書〕計(0)件

著 者 名		出 版 社	
書 名		発 行 年	総ページ数

14. 研究成果による産業財産権の出願・取得状況

(出願) 計(0)件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	出願年月日	国内・外国の別

(取得) 計(0)件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	取得年月日	国内・外国の別
				出願年月日	

15. 備考

--