

平成24年度北海道大学情報基盤センター共同研究成果報告書

1. 研究領域番号 A3 大規模データ科学
2. 研究課題名 大規模複雑類似性データの解析法に関する総合的研究
3. 研究期間 平成24年4月23日 ~ 平成25年3月31日

4. 研究代表者

氏名	所属機関・部局名	職名	備考
宿久 洋	同志社大学 文化情報学部	教授	

5. 研究分担者

氏名	所属機関・部局名	職名	備考
水田 正弘	北海道大学 情報基盤センター	教授	
寺田 吉吉	大阪大学 基礎工学研究科	D2	
谷岡 健資	同志社大学 文化情報学研究科	D1	
山下 陽司	同志社大学 文化情報学研究科	M2	
光廣 正基	同志社大学 文化情報学研究科	M1	

6. 共同研究の成果

近年の計算機の発達により、大規模複雑データに関するデータ解析法の需要が高まってきている。特に、社会現象や心理現象を扱う様々な学問分野あるいは応用分野で、2者関係を示すデータ（類似性データ）は多種多様に存在し、それぞれの分野の性質に対応した解析法が必要である。本研究では、大規模かつ複雑な類似性データに注目する。計算の複雑化・計算量の増加・解釈困難性等の課題を解決し、従来の解析法を拡張した新たな解析法の提案を目的として総合的な研究に取り組む。そして、1. 大規模かつ複雑な類似性データの理論的特徴付け、2. 新規手法の提案、3. 分析結果の評価法の提案、を中心に研究に取り組んだ。

特筆すべき成果としては、研究計画の2. についての研究成果である。総合的に本研究を行った結果、多次元尺度構成法、クラスター分析、社会ネットワーク分析等の従来の解析法が、大規模複雑な類似性データに対しても適用できるような新しい手法へと拡張することができ、新しい解析法として提案することができた。さらに、シンボリックデータや多相多元データのような類似性データに対しても、新たな手法を提案することができた。現在、大規模複雑な類似性データに対する解析法が少ない中、様々な解析法をこのようなデータに適用できるように拡張できたことは、本研究課題の大きな成果である。また、提案手法を人口データによるシミュレーションだけでなく、大規模複雑な実データへの適用も行い、既存手法との比較も行っている。手法適用の結果、実データに対する解析法の有用性を確認できたことも良かった点であると考える。

本年度は、「大規模複雑類似性データ」をテーマにして研究討論会を1回行い、集中的に最新の研究成果の相互レビューを実施した。また、本研究課題に関連した論文も公表したり、国内外での学会発表等も行ったりすることで、応用分野への普及の促進を目指した。主な研究成果は以下のとおりである。

(研究成果のつづき)

- ・ Tanioka, K. and Yadohisa, H. (2012): Effect of data standardization on the result of k-means clustering, Challenges at the interface of Data Analysis, ComputerScience, and Optimization (W, Gaul et al. eds), Springer, Heidelberg, p59-p67.
- ・ Saito, Y. and Yadohisa, H. (2012): Visualization of asymmetric clustering result with digraph and dendrogram, Challenges at the interface of Data Analysis, ComputerScience, and Optimization (W, Gaul et al. eds), Springer, Heidelberg, p151-p159.
- ・ Tanioka, K. and Yadohisa, H. (2012): Three-way asymmetric hierarchical clustering based on regularized similarity model, COMPSTAT 2012: Proceedings in Computational Statistics, Limassol, Cyprus. p789-p800
- ・ Kitano, M. and Yadohisa, H. (2012): An overlapping clustering method for signed graphs, COMPSTAT 2012: Proceedings in Computational Statistics, Limassol, Cyprus, p391-p402, (Best Paper Award: ERS IASC Young Researchers Award).
- ・ Tanioka, K. and Yadohisa, H. (2012): Subspace hierarchical clustering for three-way three-mode data using quadratic regularization, Procedia Computer Science, 12, Washington, USA, p248-p253
- ・ Tanioka, K. and Yadohisa, H. (2012): Subspace hierarchical clustering for three-way three-mode data using quadratic regularization, to appear in CAS2012, Washington D.C., USA.
- ・ Tanioka, K. and Yadohisa, H. (2012): Three-way subspace hierarchical clustering based on entropy regularization method, to appear in GfKI2012, Hildesheim, German.
- ・ Tanioka, K. and Yadohisa, H. (2012): Three-mode subspace clustering for considering effects under noise variables and occasions, The 4th Japanese-German Symposium on Classification, Kyoto, Japan.
- ・ Terada, Y. and Yadohisa, H. (2012): Asymmetric multidimensional scaling with generalized hyperellipse model, The 4th Japanese-German Symposium on Classification, Kyoto, Japan.
- ・ Yamashita, Y. and Yadohisa, H.(2012) : Similarity Measures Between Series of Candlesticks, The 2d Institute of Mathematical Statistics Asia Pacific Rim Meeting, Tsukuba, Japan.
- ・ Yamashita, Y. and Yadohisa, H.(2012) : Similarity Measure and Clustering Algorithm for Candlestick Valued Data, Joint meeting of Japanese and Italian Classification Societies, Capli, Italy.
- ・ Terada Y. and Yadohisa H. (2012): Multidimensional scaling with non-concentric hypersphere and hyperbox models for percentile-valued dissimilarity data, Third Workshop in Symbolic Data Analysis, Madrid, Spain.
- ・ 谷岡健資, 宿久洋 (2012): 正則化に基づく 3 相 3 元部分空間階層的クラスタリング法とその解析結果の視覚化法, 2012 年度統計関連学会連合大会講演報告集, p230 (於 北海道大学).
- ・ 谷岡健資, 宿久洋 (2012): 単相 3 元非対称 (非) 類似度データに対するクラスタリング法について, 日本行動計量学会 第 40 回大会発表論文抄録集, p469-470 (於 新潟県立大学).
- ・ 北野道春, 宿久洋 (2012): 符号付有向グラフを用いた非対称データの視覚化, 日本行動計量学会 第 40 回大会発表論文抄録集, p281-282 (於 新潟県立大学).
- ・ 光廣正基, 谷岡健資, 宿久洋 (2012): 視覚化法を用いた野球配球の特徴把握 ~多元データを用いたアプローチ~, 日本計算機統計学会 第 26 回シンポジウム講演論文集, p31-32 (於 東京大学).
- ・ 谷岡健資, 宿久洋 (2013): 3 相 3 元データに対する変量および条件の重みを考慮した超距離の推定法について, 第 22 回大規模データ科学に関する研究会, (於 北海道大学).
- ・ 谷岡健資, 宿久洋 (2013): 3 相 3 元データに対する変量および条件の重みを考慮した階層的クラスタリング法について, 日本分類学会第 31 回大会, (於 中央大学).