

課題名 (タイトル) :

大規模遺伝子ネットワーク推定、およびシステム解析のためのプログラムの研究開発

利用者氏名 : ○宮野 悟
島村 徹平

所属 : 社会知創成事業 次世代計算科学研究開発プログラム
次世代生命体統合シミュレーション研究推進グループ
データ解析融合研究開発チーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

本研究課題は、現在開発中の次世代スーパーコンピュータを利活用するためのプログラム「次世代計算科学研究開発プログラム」において、データ解析融合チーム研究開発課題の一つとして掲げている大規模遺伝子ネットワーク推定ソフトウェアを開発することを目的とする。特に、L1 正則化法に基づく遺伝子ネットワーク推定ソフトウェア SiGN-L1 の研究開発を行う。これにより、全遺伝子を対象とした網羅的なネットワークを「地図」として用い、薬の被影響遺伝子の同定、作用点の推定、副作用の予測・回避、創薬ターゲット・毒性関与パスウェイ探索などへの応用を目指す。

2. 具体的な利用内容、計算方法

これまで開発を行ってきた Weighted Lasso、Recursive Elastic Net、Relevance-Weighted Recursive Elastic Net と呼ばれる L1 正則化アルゴリズムを用いて、遺伝子発現データからグラフィカルガウシアンモデル、またはベクトル自己回帰モデルのパラメータを推定することにより、全遺伝子 (約 2 万) を対象とした大規模遺伝子ネットワークを推定する。

3. 結果

本年度は、これまで開発を行ってきた R 言語によるプログラム、およびバッチ処理による高並列プログラムが RICC システム上で使用可能であることを確認した。また、各アルゴリズムのさらなる高速化、完全 MPI 化に向けて、一部のプログラムの FORTRAN 言語によるコードの変更を行い、そのプログラムが RICC システム上でコンパイルが可能であることを確認した。

4. まとめ

RICC 上で L1 正則化法に基づく遺伝子ネットワークソフトウェア SiGN-L1 (R 言語によるプログラム) が利用可能であることを確認した。また、さらなる高速化・完全 MPI 化に向けて、プログラムコードの一部を FORTRAN 言語でコーディングし、RICC 上でコンパイルが可能であることを確認した。

5. 今後の計画・展望

来年度はプログラムの高速化、および完全 MPI 化に向けて、全プログラムの FORTRAN 言語によるコードの変更を進める。また、これまで本チームで開発を行ってきた SiGN-BN、SiGN-SSM、SiGN-L1 を一つのソフトウェア SiGN として、統合的に開発を進める。

平成 22 年度 RICC 利用研究成果リスト

【論文、学会報告・雑誌などの論文発表】

【国際会議などの予稿集、proceeding】

【国際会議、学会などでの口頭発表】

1. Tamada, Y., Shimamura, T., Yamaguchi, R., Imoto, S., Nagasaki, M., Miyano, S., SiGN: Large-Scale Gene Network Estimation Environment for High Performance Computing, The 2010 Annual Conference of the Japanese Society for Bioinformatics, Dec 14th, 2010, Centennial Hall Kyushu University School of Medicine, Fukuoka, Japan.
2. Shimamura, T., Imoto, S., Niida, A., Nagasaki, M., Yamaguchi, R., Miyano, S., Modulator-Dependent System Changes Unravel Cancer Heterogeneity, The 2010 Annual Conference of the Japanese Society for Bioinformatics, Dec 14th, 2010, Centennial Hall Kyushu University School of Medicine, Fukuoka, Japan.
3. 島村徹平、井元清哉、山口 類、宮野 悟、がんの多様な鍵分子を探索するネットワーク推定法、日本計算機統計学会 第 24 回シンポジウム、2010 年 11 月 11 日、箕面市立メイプルホール、大阪
4. 島村徹平、薬剤耐性機構に関わる鍵分子を明らかにするネットワーク推定法、大分統計談話会第 42 回大会、2010 年 10 月 22 日、富士通大分システムラボラトリ、大分
5. 島村徹平、井元清哉、山口 類、宮野 悟、係数変動構造方程式モデルに基づくネットワークプロファイリングによるがんシステムの多様性解析、2010 年度統計関連学会合同連合大会、2010 年 9 月 6 日、早稲田大学、東京
6. 島村徹平、パスウェイアクティビティに基づくがんの多様性解析、第 11 回データ解析融合ワークショップ、2010 年 8 月 27 日、東京大学医科学研究所、東京

【その他】

1. Shimamura, T., Imoto, S., Niida, A., Nagasaki, M., Yamaguchi, R., Miyano, S., Estimation of patient-specific gene networks related with a cancer hallmark, Molecular Medicine Tri-Conference 2011, Feb 24th, 2011, Moscone North Convention Center, San Francisco, USA.
2. Shimamura, T., Imoto, S., Yamaguchi, R., Nagasaki, M., Miyano, S., Gene network comparison based on multiple time-course microarray data, 2010 Joint Wellcome Trust / CSHL Conference Systems Biology: Networks, Aug 13th, 2010, Wellcome Trust, Hinxton, UK.