

演算加速装置を用いた遺伝子ネットワーク推定アプリケーションの最適化

岡山理科大学 情報理工学部学部 木戸 善之

目的： 遺伝子制御ネットワーク推定アプリケーションINGORは、B-spline回帰に基づくベイズネット推定に加え、欠損補完、相関のスコア計算も含むため、遺伝子数やサンプル数の増大に伴い計算時間が増加する。本研究は、実用規模（例：1万遺伝子級）での計算を現実的にすることを狙い、ベクトルプロセッサおよびGPGPUでの高速化可能性と最適化方針を明確化する。

内容：まずプロファイリングによりINGORの実行時間を支配する関数群（今回の対象はB-spline計算とスコア計算関数）を特定し、ベクトル化率と負荷の内訳を分析した。次に、線形代数演算をVE-BLAS/VE-LAPACKへ置換し、データ配置やメモリアライメントを調整してベクトル実行効率を改善した。自動ベクトル化が難しい複雑ループにはループアンローリング等を適用し、スカラー命令効率の向上も含めて総合的に最適化した。

結果：大規模条件（10,540遺伝子×500サンプル）では、CPU版の約761分に対し、ベクトル版は自動最適化で約317分、手動アンローリング併用で約203分まで短縮し、実用規模で顕著な高速化を実証した。一方、小規模ではオーバーヘッドの影響でCPUが有利な場合もあり、規模依存性が示された。また可視化では全エッジ表示が過密になるため、実務では高スコアエッジの抽出表示が必要である点も整理された。GPGPUでの最適化については現在検証を行っており、国際会議投稿に向けて準備中である。

利用した計算機 SQUID 汎用CPUノード群

ノード時間 1920.51 時間
使用メモリ 24 GB
並列化 (並列化については将来課題)

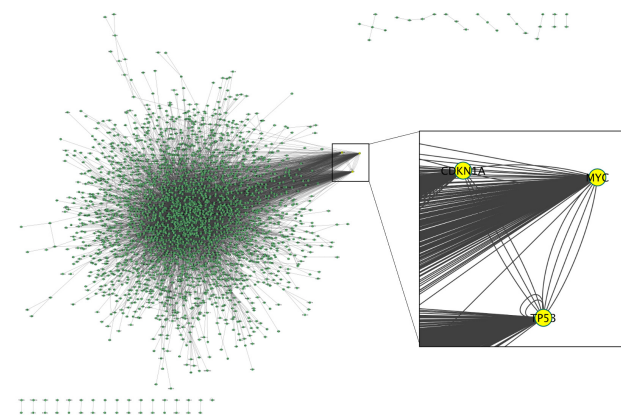


図 遺伝子ネットワーク推定結果の可視化