

審査の結果の要旨

氏 名 小島 要

細胞内システムを理解することは、がん細胞メカニズムの理解、薬剤応答の推定、新規薬剤の開発等に有用であり、そのために遺伝子発現量から遺伝子間の制御関係を表現する遺伝子ネットワークの推定を行うことが望まれてきた。これは、近年のマイクロアレイ技術の発展から大規模な遺伝子発現データの計測が可能となり、異なる条件下での遺伝子発現量の変化を比較することが可能になったことによる。現在では遺伝子ネットワークの推定が細胞内システム解明のための基礎を与える方法として認められているが、既存研究では大規模で複雑な系に対処することが実現できていない。本論文は大規模で複雑なデータに対する遺伝子ネットワークを推定する方法論について多面的に研究を行うもので、これまで厳密解が得られなかった規模まで計算を可能とする推定、そしてシステムの理解に有用なネットワークレイアウトについて、モデリング法とそのための効率的なアルゴリズムを提案し、実際のデータに適用して知見を得るという大きな研究成果をあげたものである。

まず、ノックダウンマイクロアレイの遺伝子発現データ等の非時系列データから遺伝子ネットワークを求める新しい探索アルゴリズムが提案されている。このアルゴリズムは、遺伝子ネットワーク推定をベイジアンネットワークの構造学習により行う。既存研究では、条件付き独立性の検定による構造推定法と情報量規準をスコアとした構造探索を行う方法が提案されているが、有向閉路を避けるために検定結果と異なるエッジの方向付けが行われる等の問題があった。また、情報量規準をスコアとしたベイジアンネットワークの最適構造推定問題はNP困難であり、最適構造の探索は動的計画法による効率計算手法と超並列による計算を行った場合でも数十ノード程度が限界であった。本論文では、条件付き独立性の検定やその他の統計的手法により推定された無向グラフが上位構造として与えられたもとで、上位構造の部分グラフとなるような最適構造を探索する方法を提案している。特に、上位構造が与えられたもとの最適ベイジアンネットワーク構造の学習において、上位構造を密に連結した部分グラフに分割し、各部分グラフでの最適ベイジアンネットワーク構造を結合する方法を与え、非有向閉路性を保つための必要十分条件を数理的に解析している。これにより、初めて平均次数4程度で数百ノードからなる最適ネットワークの探索を可能とし、定常状態での遺伝子ネットワークを大規模系で推定する方法を与えた。この方法をヒト臍帯静脈内皮細胞でのノックダウン遺

伝子発現データに適用し、96遺伝子からなる遺伝子ネットワークの推定を行い、生物学的知識との比較を行っている。

次に、時系列からなる遺伝子発現データからの遺伝子ネットワーク推定が、薬剤刺激やヒートショック反応時の制御関係の解析に有効であることから、この推定問題に取り組んでいる。既存手法で線形性やデータ計測時点の等間隔性が仮定されている問題を克服するため、非線形関係まで抽出できるようなノンパラメトリックモデルをスパース学習により推定する方法を提案している。同時に、不等間隔に計測されたデータからも正確な遺伝子ネットワークの推定を行うために、ベクトル自己回帰モデルを状態空間表現することで、不等間隔データを扱うことが可能なモデル、さらにスパース学習法を用いたパラメータの推定法も提案している。提案手法を正常ヒト小気道上皮細胞からの時系列遺伝子発現データとEGF 受容体への刺激と抗がん剤イレッサとして知られるゲフィティニブの投与を行ったヒト小気道上皮細胞からの時系列遺伝子発現データに適用した結果、2つの状態の細胞において推定されたネットワークを比較することでEGF 受容体を介さずにイレッサから影響を受ける遺伝子の候補が同定された。

また、遺伝子制御ネットワークがパスウェイに関連したモジュールから構成されることに着目し、モジュール内では比較的密に、モジュール間では疎に制御関係が存在することを仮定し、変分ベイズ法の枠組から時系列遺伝子発現データから遺伝子の制御関係と遺伝子モジュールを同時に推定する手法を開発した。提案手法をヒトの子宮頸癌細胞由来のHeLa 細胞から測定された94の癌関連及び細胞周期関連遺伝子の時系列遺伝子発現データに適用することで、遺伝子ネットワークと2つの遺伝子モジュールが同時に同定された。

そして、遺伝子ネットワークからの情報抽出を行うために、遺伝子ネットワークや生物パスウェイを対象とした可視化技術として、数種のネットワークレイアウト手法について、たんぱく質の細胞内の複雑な局在情報を考慮できる方式を提案している。よい局所最適解を得るために、既存の探索手法において時間計算量を増加させることなくノードの位置交換を考慮したアルゴリズムを開発した。さらに、特にスパースなグラフ構造において有効なアルゴリズムを開発して時間計算量を削減し、実時間アプリケーションに耐え得るグリッドレイアウトを実現した。また、細胞局在情報を表すコンパートメントの拡大・縮小と移動を探索ステップにおいて考慮することで、より良いレイアウトを得ることが可能となった。提案レイアウトアルゴリズムをヒト小気道上皮細胞のデータから推定された遺伝子ネットワークに適用し、各遺伝子に対応したタンパク質の細胞内局在情報を用いた遺伝子ネットワークのレイアウトが可能となり、遺伝子ネットワークの特徴を捉えることができている。

なお、本論文の一部は共同研究によって得られたものであるが、申請者が主体的に研究して得られた成果であることを確認している。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。