審 査 の 結 果 の 要 旨

氏　　名　　小林　諭

本論文は、「ネットワークログの因果解析による障害の原因究明支援技術に関する研究」と題し、日本文で記されており、全体で10章から構成される。大規模ネットワークの運用支援および効率化を目的に、運用システムログからイベント間の因果関係を自動抽出してトラブルシューティングに活用する手法を提案したものである。システムログはシステムの障害の原因究明を行う上で最も重要なデータであるが、一方で大規模なシステムではデータの量も莫大なものとなり活用が困難となる。このシステムログ中のイベント間の関連性を因果関係の形で抽出することができれば、オペレータがトラブルシューティングのためシステムの振る舞いを把握することが容易になり障害対応が効率化されると考えられる。システムログ中の因果関係は既存の技術では計算処理の負担から限られた期間のデータで行われてきた。これに対し本論文ではより広範囲のデータを対象とする探索的な因果解析を行い、得られる情報を平時と対比することでより重要性の高い情報に着目することを可能としている。さらに実運用されている大規模ネットワークシステムのログデータに適用し、手法の有用性と得られる因果情報のトラブルシューティング上の価値についての検証を行った。

第1章「序論」では、本論文で扱う大規模ネットワークシステムの運用の現状について解説し、その問題点と改善のための課題を整理している。そしてこの課題を解決するための提案手法の概要を示し、その実現により得られる社会への貢献を議論している。

第2章「運用とログ」では、本論文が前提としているシステムログのネットワークシステム運用における立場と価値について述べている。システムログのうち特に広く用いられているsyslogの動作について解説し、またその機能に当たるログの記録と収集、そしてその活用手段の現状を整理している。

第3章「関連研究と既存技術」では、システムログの活用のための既存技術について記述している。まず商用化されているログ解析製品やサービスについて整理し、これらが主に相関などの基本的な技術から構成されていることを示している。次に研究段階の技術についてログの前処理と得られたログに基づく解析の双方について説明している。特に本論文の提案手法と関わりの深い因果解析を行なっている論文を取り上げ、提案手法との手法選択の違いが広範囲のデータの因果解析を想定しているかどうかという点に由来していることを示している。

第4章「理論的背景」では、本論文の提案手法の軸となっている因果解析について、その基礎となっている因果推論の考え方、および因果解析を効率的に行うためのPCアルゴリズムという手法について解説している。さらにこのPCアルゴリズムを利用するため

の条件付き独立検定手法や、PCアルゴリズムと類似する因果DAG推定手法などの周辺技術について紹介している。

第5章「提案手法」では、PCアルゴリズムに基づく因果解析手法を提案手法として解説している。この提案手法は主に、ログを時系列として扱うためのLog template生成手法、時系列を解析に適した性質のデータに加工する前処理、PCアルゴリズムでスパースなデータを扱うための検定手法、得られる因果のトラブルシューティング上の重要性を推定する後処理、の4つの要素技術から構成される。これらの要素技術は本論文の目的である広範囲のデータを対象とする因果解析を効率的に行う上で重要なものである。

第6章「データセット」では、本論文で検証および評価で用いた教育機関向け大規模ネットワークであるSINETのシステムログについて、このデータセットが持っている特徴を示し以降の各実験を理解する上での助けとしている。

第7章「要素技術の検証」では、提案手法の各要素技術の手法選択の妥当性を示すため、それぞれの手法の性能や既存手法に対する優位性を示している。またパラメータの検証や発生するFalse positiveの調査などを行なっている。

第8章「評価」では、第6章で示したデータセットにおいて提案手法で検出された因果関係について調査を行なっている。得られる因果関係がシステムの振る舞いとして妥当なものであることを、オペレータの知識に基づく因果の分類、および複数のケーススタディを通して解説している。さらに、運用トラブルチケットとの対比によって得られる因果の多くが記録されている大規模な障害において重要な情報を指摘していることを示した。一方でケーススタディにはトラブルチケットには見られない事象が含まれており、提案手法がオペレータや利用者の気づかない障害を見つけることで障害を未然に防ぐ助けとなることをも示唆している。

第9章「議論」では本論文で行なった各実験から得られる知見を整理するとともに、そこから導かれる今後の課題および改善案について議論している。またリアルタイム解析での利用や因果を用いた原因イベントの特定など提案手法の応用手段についても言及している。

第10章「結論」では論文全体での技術的、あるいは学術的貢献について再度まとめ、この研究の応用が期待される技術とその価値について議論している。

以上を要するに、本論文は、システムログの探索的な因果解析という要求に対する従来技術では実現困難であった課題について、新しいシステム構成を提案し、それに由来する性能改善および効率化を実証評価している。この実現により、因果に基づくログ活用は柔軟性を大きく向上し、システム運用の効率化および安定化に大きな貢献が期待される。 大規模インターネットシステムに関する運用技術の研究開発に関して実践的で先駆的な貢献と認められ、情報理工学における創造的実践の観点で価値が認められる。

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。