

氏名	中挟 晃介		
学位の種類	博士（工学）		
学位記番号	博甲第9407号		
学位授与年月日	令和2年3月25日		
学位授与の要件	学位規則第4条第1項該当		
審査研究科	システム情報工学研究科		
学位論文題目	時系列データに対するパターンマッチングと予測に関する研究		
主査	筑波大学 教授	理学博士	北川 博之
副査	筑波大学 教授	博士(工学)	天笠 俊之
副査	筑波大学 教授	博士(工学)	櫻井 鉄也
副査	筑波大学 教授	博士(工学)	山本 幹雄
副査	筑波大学 教授	博士(工学)	佐藤 哲司
(図書館情報メディア研究科)			

論文の要旨

近年、情報通信技術、センサ技術等の発展に伴い、センサデータ、ログデータ、SNS データといった時系列データが日々大量に生成・配信されている。時系列データは、データ間の時間的な順序が重要な意味を持つデータである。本論文では、各時系列を識別する属性、タイムスタンプ等の各レコードの時間的順序を表す属性、観測値や場所等のデータ固有の属性をもつレコード（行）群からなる時系列データを対象とし、その解析と予測に関する2つのテーマに関し研究を行っている。

時系列データの解析では、リレーショナルデータベース等に格納された大規模な時系列データに対するパターンマッチング（行パターンマッチング）を対象としている。具体的には、SQL に導入された行パターンマッチング機能におけるような、行の出現順序と行の属性値の両者に関わる条件が与えられるパターンマッチング処理の効率化について検討している。一般に、行パターンマッチングの処理コストは時系列データの長さに比例する。本論文では、行の属性値に関する条件を用いて事前にフィルタリングを行うことで、行パターンマッチングのコスト削減を図る手法を提案している。1つは、パターンに合致する可能性のある行を全く含まないシーケンス全体をフィルタリングで除去する Sequence Filtering である。もう1つは、パターンに合致する可能性のある行とその前後の行以外の行をフィルタリングで除去する Row Filtering である。本論文では、行パターンマッチング処理と提案するフィルタリング処理を PostgreSQL と Spark で実現し、人工データと実データを用いた評価実験により、これら2種類のフィルタリング手法によって行パターンマッチングの処理コストが削減できることを確認している。一方、コスト削減の効果はデータやパターンの性質に依存する。そこで、本論文では各手法による処理コストを見積もるためのコストモデルを提案し、事前に各手法の処理時間を一定程度見積もる

ことが可能であることを示している。

時系列データに対する予測に関しては、鉄道における各列車の各駅における運行遅延時間を記録した時系列データを対象とし、現時点までの状態から今後の各駅における遅延時間を予測する手法を示している。具体的には、予測対象列車の現時点より数駅前における遅延時間と乗車率の時系列データ、および予測対象列車の前後列車の遅延時間の時系列データから、予測対象列車がこれから停車／通過する予定の駅における遅延時間を予測するニューラルネットワークを用いた予測モデルを提案している。実在路線における実際の列車の遅延データを用いた実験を行い、モデルの構築単位や入出力の観点からの評価、他の予測モデルとの比較評価等により、提案モデルが精度が良い予測を可能とすることを示している。

審 査 の 要 旨

【批評】

近年、多様な時系列データが日々大量に生成・配信されており、その利活用は重要な課題である。本論文では、複数の属性を持ったレコードからなる時系列データの解析と予測に関する2つのテーマを扱っている。前者に関しては、リレーショナルデータベース等に格納された時系列データを対象に、行の出現順序に関する条件と行が持つ属性値に関する条件の両者が与えられるパターンマッチング処理の効率化手法を提案している。具体的には、リレーショナルデータベース等において一般的に実行可能で処理負荷があまり大きくない属性値に関する条件に基づく行選択処理を事前に実行し、行数を削減することで、出現順序に関わるパターンマッチング処理のコスト削減を行う2つの手法を提案している。また、PostgreSQLとSparkを用いた実験によりその有効性を確認している。また、処理コストを事前に見積もるためのコストモデルを併せて提案し、それにより適切な処理手法を選択することが可能であることを示している。時系列データの予測に関しては、ニューラルネットワークを用いた列車の運行遅延時間の予測を可能とするモデルを提案し、実際の列車運行データを用いることで、従来よりもより高精度の予測が可能であることを示している。行パターンマッチング処理の効率化に関して新たな手法を提案しその有効性を示している点や、列車遅延時間に関して従来よりも精度のよい予測手法を提案している点で、本論文は情報工学上の大きな貢献が認められる。

今後は、行パターンマッチングと他のリレーショナルデータベース演算が混在する問合せ処理の効率化や、天候や曜日等の他の要因を考慮した列車遅延時間の予測モデルへの発展等、残された課題に対する検討が進められることが期待される。

【最終試験の結果】

令和2年2月3日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。