

氏 名	Savong Bou		
学 位 の 種 類	博 士 ( 工 学 )		
学 位 記 番 号	博 甲 第 8076 号		
学位授与年月日	平成 29 年 3 月 24 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	A Study on Keyword Search over Structured and Semi-structured Data Streams		
	(構造データおよび半構造データストリームに対するキーワード検索に関する研究)		
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(理学)	加藤 和彦
副 査	筑波大学 教授	学術博士	和田 耕一
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)		
副 査	筑波大学 准教授	博士(工学)	天笠 俊之

## 論 文 の 要 旨

情報通信技術の急速な発展、インターネットや無線ネットワークの普及、情報通信機器の小型化、低価格化に伴い、さまざまな応用において、リアルタイムかつ継続的に送信される「データストリーム」が注目を集めている。特にビッグデータや Internet of Things (IoT) などにおいて主要な情報源として認識されており、活発な研究開発が行われている。本論文では、多様なデータストリームの中でも、特に XML (Extensible Markup Language) 形式のストリーム (XML ストリーム) 及びリレーショナルデータ形式のストリーム (関係ストリーム) を対象に、キーワードを用いた新たな検索手法を提案している。

XML ストリームに対するキーワード検索とは、ストリームとして連続的に送信される XML に対して、検索キーワードを全て含む最小の部分木を連続的に検出する処理である。本論文では、利用者が XML ストリームの構造に関する一定程度の知識を有する状況を想定し、指定された部分構造の範囲内にキーワードが出現した場合にだけ検索結果が得られるような枠組みを提案している。構造の指定には既存の XPath 式を用い、これとキーワードを組合せることで、検索条件の指定を可能にする。静的な XML データに対するこのような検索に関しては既存の研究があるが、XML ストリームを対象としたものはこれまでにない。本論文では、XML ストリームに対する XPath 処理手法である YFilter と XML ストリームに対するキーワード検索手法である CKStream を組合せた新たな処理アルゴリズムを提案するとともに、実験による評価を行い提案手法が有効であることを検証している。

一方、関係ストリームに対するキーワード検索では、外部キーで接続された複数の関係表に対応す

るタプルが連続的にストリームとして送信される状況を想定し、検索キーワードを全て含む最大サイズ  $T_{\max}$  以内のタプル集合（タプルネットワーク）を連続的に検出する。既存手法では、検索キーワードの個数あるいはタプルネットワークの最大サイズ（ $T_{\max}$ ）が大きくなると性能が急速に劣化するという問題がある。これは、関係スキーマ、検索キーワード、 $T_{\max}$  から、検索結果が取りうる全てのタプルネットワークのトポロジーとキーワードの出現パターン（候補ネットワーク (Candidate Network (CN))）を列挙し処理プランを生成するが、既存手法では、キーワード数や  $T_{\max}$  が大きくなると CN が組合せ爆発を起こし処理プランが巨大になってしまうためである。この問題に対し、本論文では、CN の組合せ爆発を回避することが可能な新たな処理プラン MX-Structure を提案するとともに、MX-Structure を利用した問合せ処理アルゴリズムを提案している。実験による評価により、検索キーワードや  $T_{\max}$  を増やしても、従来手法に比べて性能劣化を抑えられることを示している。

## 審 査 の 要 旨

### 【批評】

今日、ストリームデータの利活用技術の高度化はデータ処理における重要な課題となっている。本論文では、XML ストリーム及び関係ストリームに対するキーワード検索手法に対する新たな手法を提案している。XML ストリームに対する提案手法は、XML ストリームの構造概要に関する一定の知識を有する場合は有用な手法であり、その効率的な処理方式を提案し、有効性を実験によって示している。また関係ストリームに対しては、従来手法での性能劣化が顕著な検索キーワード数や検索結果のネットワークサイズが大きい状況に対して対応可能な新たな処理方式を提案し、実験的にその有効性を確認している。このように、本論文は、ストリームに対するキーワード検索の高度化に貢献しており、情報工学上の新規性及び有用性が認められる。今後、実際の応用への適用評価やより複雑な検索条件に対応した拡張に関する研究等が進められることを期待する。

### 【最終試験の結果】

平成 29 年 2 月 6 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。