

学位論文内容の要旨

学位申請者	榎 美紀 【理学専攻 平成25年度生】	要 旨
論文題目	リアルタイムデータアクセス処理機構の最適化	<p>近年、様々なデバイスから位置情報・気象・ソーシャル・センサーデータなど、膨大な量のデータが日々発信され続けており、地球上で生成されるデータは 2013 年から 2020 年の間で 4 兆 4,000 億ギガバイトから 44 兆ギガバイトへと 10 倍の規模に拡大すると予想されている。このような膨大なデータを利用して関連性を見つけたり、変化を検知したりと、様々な観点から分析する要求は年々高まっており、それらを実現するための処理基盤は大規模データをリアルタイムに、高速に扱えることが重要となる。</p> <p>そこで本論文では、リアルタイムなストリームデータを逐次的にストリーム処理する機構と、蓄積されたデータを処理する機構を兼ね備えたリアルタイム分析システムを提案した。大規模リアルタイムデータの一つであるソーシャルメディアデータを対象として、ソーシャルメディアデータがもつ特徴を利用したデータアクセスの最適化手法を、逐次的にストリーム処理する機構と、蓄積されたデータを処理する機構それぞれに提案した。</p> <p>リアルタイム分析を実施するための、逐次的にストリーム処理をする部分については、拡散モデルを用いて各メッセージの流行が収束するタイミングを早期に推定し、時間ウィンドウ幅をカスタマイズしメンテナンスする手法を提案して、Twitter の実データを用いてその効果を示した。重要度が低いと判断されたメッセージをフィルタリングして処理するデータ量をコントロールする手法を提案して有効性を明らかにした。</p> <p>提案システム内の蓄積されたデータを処理する部分については、データアクセスの処理性能を評価し、提案したデータアクセス最適化手法の効果を検証した。更新の影響が及ぶキャッシュのみを無効化する細粒度のキャッシュメンテナンス手法により性能が改善することを示した。またグラフデータアクセスの高速化のため、ビットマップのインデックスを導入し、行列を並び替えることにより、さらに圧縮率を高められる事を示した。</p> <p>本論文により大規模データをリアルタイムに、高速に扱う最適化手法が明らかにされ、大規模データを分析処理する有用性が広げられた。</p>
審査委員	(主査) 教授 小口 正人	
	教授 小林 一郎	
	教授 吉田 裕亮	
	教授 椎尾 一郎	
	教授 伊藤 貴之	