

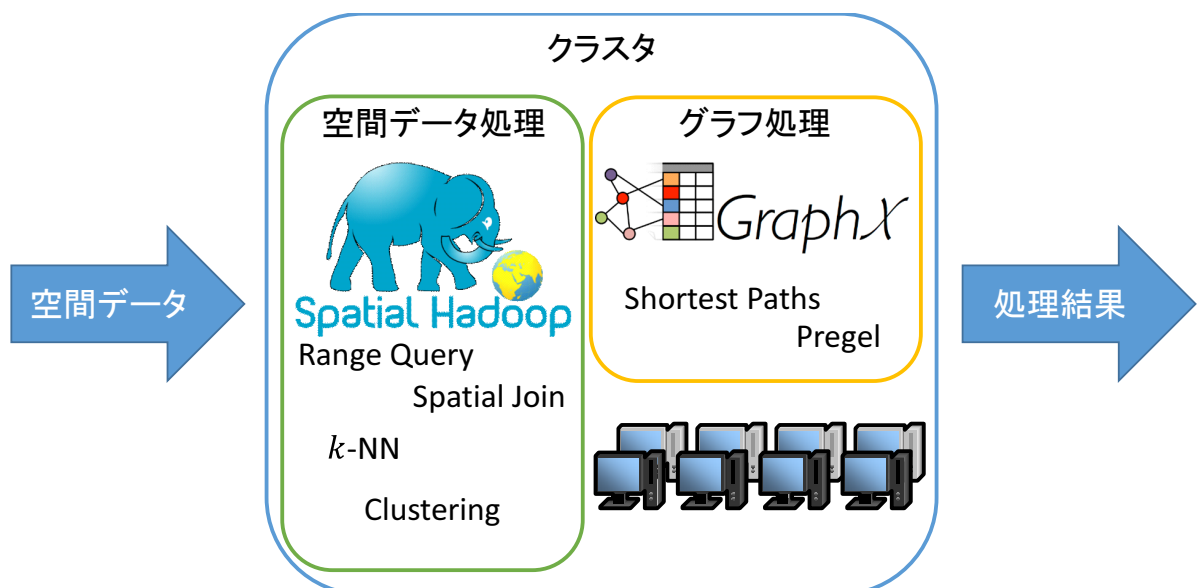
平成26年度 情報工学コース卒業研究報告要旨

石川 研究室	氏 名	瀧 本 祥 章
卒業研究題目	並列分散処理技術を用いた 最人気経路の探索に関する研究	

スマートフォンなどのモバイルデバイスの普及によって、生成される空間データはここ最近で爆発的に増加したため、これらを効率的に処理することが求められている。このようなビッグデータの分析には、HadoopをはじめとするMapReduceフレームワークに基づく並列分散処理技術が非常に有用である。しかし、Hadoopは空間上の性質を考慮していないため、その処理の効率性が制限されるという問題がある。また、空間データの処理では道路ネットワークなどに関して、グラフ処理が必要になることがある。反復処理が前提となるグラフ処理に関しては、一回ごとの処理でHDFSに書き出す必要があるHadoopは効率的とは言えない。

そこで、本研究では、空間データの処理をSpatialHadoopで行い、グラフ処理をGraphXで行うことによる、最新の並列分散データ処理技術に基づく処理の効率化を提案する。SpatialHadoopとは、Hadoopの包括的な拡張である。従来のHadoopにPointやRectangleなどの空間データを扱うための型や、グリッドインデックスやR木などの空間インデックスを実装することにより、空間データを容易かつ効率的に扱えるようにしている。また、GraphXはApache Sparkの構成要素の1つで、グラフ処理を並列分散環境で行うためのフレームワークである。グラフ処理中の反復処理ごとにHDFSに書き出す必要がなく、効率的にグラフ分析を行うことができる。

本稿では、その適用事例としてChenらによる最人気経路の探索手法の効率化を図る。Chenらによる最人気経路の探索手法は道路ネットワークに関する情報を利用せず、移動軌跡データから直接的に移動情報を抽出する。その過程でクラスタリングによる交差点の発見やグラフ処理を行う必要があるため、SpatialHadoopやGraphXを適用し、効率化を図る。



提案手法イメージ図