

平成28年度 情報工学コース卒業研究報告要旨

石川 研究室	氏 名	河井 悠佑
卒業研究題目	大規模シミュレーションデータの対話的分析 のためのデータベース技術に関する研究	
<p>今日では、計算機の性能向上やストレージの容量拡大といった情報技術の発達により、科学の分野では大規模なシミュレーションが行われている。それに伴い、生成されるデータ量も非常に大きなものとなっている。こういった大規模なデータそのものを把握することは非常に困難であるため、ユーザが効果的に分析できる環境を提供する必要がある。大規模で複雑なデータであれば、一部のデータを検証しながら段階的に進めていくといった、探索的なアプローチの分析が求められる。したがって、ユーザの目的に合わせた多方面からの分析と、その分析要求に対して素早く応答可能な対話性を実現する必要がある。</p> <p>本研究では、特に時空間情報に基づいて行われた大規模な時空間シミュレーションに着目し、その効果的な管理及び分析技術の開発を目的とする。シミュレーションデータの分析では、主に時刻または空間に基づいたデータの参照が行われる。しかし、従来のリレーショナルDBMS (relational DataBase Management System) による時空間シミュレーションデータの管理は非効率である。</p> <p>そこで、本研究では近年登場した配列指向DBMS (array DBMS) による大規模な時空間シミュレーションデータの分析について提案する。配列指向DBMSは、多次元配列形式データの活用の特化したDBMSである。時空間情報を配列における次元と結び付けることによって、時空間情報に基づいた分析が効率的に行える。本研究で使用する配列指向DBMSはSciDBであり、その特徴はチャンク (chunk) 分割と並列分散処理である。チャンク分割により、配列上で近い位置にある要素をコンピュータのディスク上においても近い位置に保存するため、一部のデータを抽出、加工するような処理に優れる。更に、全体のフレームワークは並列分散処理に基づいているため、大規模なデータ全体に対する解析処理にも対応できる。</p> <p>こうした特徴を踏まえ、本稿ではSciDB上で時空間シミュレーションデータの管理及び分析を行った。扱った時空間シミュレーションモデルは首都直下型地震を想定した避難シミュレーションであり、地震発生から24時間分の人の位置や状態が記録されている。対象となる地域は周囲が川で囲まれているだけでなく、木造住宅が密集しているため、災害において火災などの危険性が高い地域である。このシミュレーションデータに対する分析例を挙げ、次元範囲に基づいた要素の抽出や集約的な計算が必要であることを確認した。そして、SciDBにおけるそれらの問合せについて実行時間を計測し、瞬時に応答可能であることを確認した。最後に実験結果を踏まえ、大規模な時空間シミュレーションデータの分析についての今後の展望を述べる。</p>		