

日本—米国 国際共同研究「ビッグデータと災害」 平成 27 年度 年次報告書	
<b>研究課題名（和文）</b>	災害時応用のための効率的かつスケーラブルなビッグデータの収集・解析・処理
<b>研究課題名（英文）</b>	BDD: Efficient and Scalable Collection, Analytics and Processing of Big Data for Disaster Applications
<b>日本側研究代表者 氏名</b>	原 隆浩
<b>所属・役職</b>	大阪大学 大学院情報科学研究科・教授
<b>研究期間</b>	平成 27 年 4 月 1 日～平成 30 年 3 月 31 日

## 1. 日本側の研究実施体制

氏名	所属機関・部局・役職	役割
原 隆浩	大阪大学・大学院情報科学研究科・教授	問題定義、方法論立案、機構の設計・拡張
西尾 章治郎	大阪大学・総長	機構の実装、実証実験

## 2. 日本側研究チームの研究目標及び計画概要

初年度は、本研究課題の基礎的な技術開発を目的として、(1)高次元データの複数センサストリームの圧縮、(2)高次元データ検索のための索引機構、(3)複数の情報源から発生する高次元データのモニタリングの 3 つの課題について研究を推進する。日本側が課題(3)について、米国側が課題(1)および(2)について、中心的に取り組む。

この年度は、月に 1 回の頻度でオンラインミーティングにより、進捗状況および研究の方向性の確認を行う。また、半年に 1 回程度の頻度で、一方の研究チームが他方を訪問し、研究開発内容の確認・意見交換および今後の方向性について詳細に議論する。両国の PI は、提案課題の要素技術について相補的かつ高い専門性を有しており、実システムの開発経験も豊富である。そのため、両者が連携すると大きな相乗効果が期待できる。

### 3. 日本側研究チームの実施概要

本研究課題の初年度である平成 27 年度は、複数の情報源から発生する大規模な高次元データに対して、効率的なモニタリングを行う機構に付いて研究開発を実施した。具体的には、Top-k 検索（指定条件から算出されるスコアが上位 k 個のデータを取得）、k 最近傍検索（指定位置に近い上位 k 個のデータを取得）、MaxRS 検索（指定条件から算出されるスコアの合計が最も高くなるデータ集合を包含する領域を発見）などの高度な検索とそのいくつかの派生形を対象として、単発検索および連続検索を効率的に実行する技術を考案した。さらに、外れ値などの検出に有効と考えられる検索技術として、検索結果の多様性を考慮した Top-k 検索技術についても研究開発を実施した。考案技術の有効性を、多様な人工データと実データの両方を用いて、多角的に検証した。これらの研究成果は、分散・並列データベース分野で著名な国際論文誌（2 編）やデータベース分野の難関国際会議を含めて、6 編の論文誌および 3 編の国際会議論文として公表されている。このように、初年度は研究計画に対して順調に進捗し、成果公表に関しては予想以上の業績を達成することができた。

平成 27 年度はさらに、平成 28 年度から開始するワークパッケージ 5「マイクロブログの実社会センシング応用のための機構」についても、準備的な研究を開始した。

米国側の PI である Sanjay Madria 教授とは、米国で 2 回、日本で 1 回、直接会って研究打合せを実施すると共に、オンラインミーティングやメールでも頻繁に意見交換を行い、両者の担当ワークパッケージにおける考案技術について議論を行い、さらに今後の具体的な連携について検討を行った。

さらに、平成 28 年 3 月 27 日～31 日に、Sanjay Madria 教授および他数名の研究者と共にクローズドなセミナー（NII 湘南会議）「Big Data: Challenges and Opportunities for Disaster Recovery」を主催し、災害対応のためのビッグデータ応用について、情報システム、地理学、人流解析の専門家らと研究・技術討論を行った。