

(27-1)

## 平成 27 年度研究開発成果概要書

課 題 名 : ソーシャル・ビッグデータ利活用・基盤技術の研究開発

採 択 番 号 : 178B06

個別課題名 : 課題B 新たなソーシャル・ビッグデータ利活用・基盤技術の研究開発

副 題 : 交通ビッグデータに基づく運転者指向サービス基盤技術の研究開発

### (1) 研究開発の目的

ドライブレコーダ販売事業者である株式会社データ・テックの協力を受け、数千台規模で導入を行っている大手事業者等からドライブレコーダデータの提供を受け蓄積を開始する。小規模なデータを用いて、交通ビッグデータ処理基盤および運転者指向サービス基盤技術の予備実験及び基礎設計を行う。

### (2) 研究開発期間

平成26年度から平成27年度（2年間）

### (3) 実施機関

国立大学法人東京大学<代表研究者>（実施責任者 准教授 豊田正史）

### (4) 研究開発予算（契約額）

総額 120百万円（平成27年度 65百万円）

※百万円未満切り上げ

### (5) 研究開発課題と担当

課題 : 交通ビッグデータに基づく運転者指向サービス基盤技術の研究開発

1. 大規模ドライブレコーダデータの蓄積（国立大学法人東京大学）
2. 交通ビッグデータ処理基盤技術の構築（国立大学法人東京大学）
3. 運転者指向サービス基盤技術の研究開発（国立大学法人東京大学）

### (6) これまで得られた成果（特許出願や論文発表等）

		累計（件）	当該年度（件）
特許出願	国内出願	0	0
	外国出願	0	0
外部発表	研究論文	2	2
	その他研究発表	14	8
	プレスリリース・報道	0	0
	展示会	0	0
	標準化提案	0	0

### (7) 具体的な実施内容と成果

課題A「大規模ドライブレコーダデータの蓄積」:

ドライブレコーダの製造販売を行っている、株式会社データ・テックの協力の下、数千台規模でドライブレコーダの導入を行っている大手運送事業者からデータを収集、蓄積するシステムを構築、稼働させ、継続的にドライブレコーダデータを蓄積することができた。今年度は、2015年4月より継続的に、都内16営業所、1400人強のドライバーの運転記録を収集し、2016年3月現在も継続的にデータ蓄積を行っている。

また、大手バス事業者からもデータの提供を受けることで合意し、同システムでデータの収集と蓄積を開始した。23 営業所、2700 人強のドライバーの、1 か月分程度の運転記録を試験的に収集し、蓄積システムが問題なく稼働することを確認した。これにより、業種の異なる複数事業者のデータからなる、学術目的で最大規模のドライブレコーダデータアーカイブを構築することができ、今後も継続的に規模を拡大して運用していく準備が整った。

課題1 78A09、1 78B01 との連携に関し、フローティングカーデータ及びレダデータの大規模蓄積に関する検討を行い、多種データを統合的に利用した運転者試行サービス基盤技術の開発を推進する体制が整った。

#### 課題イ「交通ビッグデータ処理基盤技術の構築」:

本課題では、内閣府最先端研究開発支援プログラム (FIRST) において東京大学が開発した超高速データベースエンジンをベースとして、交通ビッグデータに適した処理基盤の構築を実現するべく、大規模ドライブデータを対象としたアーカイブ技術ならびに高速処理技術に関する検討を実施し、交通ビッグデータ処理基盤技術の基本設計・実装ならびに処理実証実験を実施した。アーカイブ技術に関しては、前年度から引き続き、技術を先行的に調査し、パブリッククラウドに於けるオブジェクトストレージ技術の調査ならびに新しい磁気ディスクドライブ技術である瓦書き磁気記録技術の調査を実施した。前者は、パブリッククラウドで利用が広がりつつあり、アーカイブに有力な技術と言えるが、現時点に於いてはアーカイブを全てパブリッククラウドに設置することは必ずしも現実的ではなく、また、瓦書き磁気記録技術も既に磁気ディスクで実用化はされているものの、旧来的なインターフェースを踏襲するにとどまっており、継続的に調査は行うものの、当面は一般的なブロックストレージ装置を用いてアーカイブを構築する方針を固めた。アーカイブ技術の検討を基に、処理基盤の基本設計としては、交通ビッグデータ、即ち大規模データが次々と到着する状況を想定し、直近のデータ解析を行うオンラインの空間と、過去のデータを蓄積するアーカイブ空間を物理的には分離して構成するものの、問い合わせの都度に人手で両方に対してこれを実行し、また、その結果を併合することは効率が低く、両空間を融合するためのシステムが欠かせない。このための主要なコンポーネントとして、ストレージ・索引管理機構ならびに高速問合せ処理機構を設計した。前者は、オンラインの空間に取り込まれたデータを基に、これをアーカイブ空間に移送し、また、両空間に於いて索引等のデータ構造管理を行う。他方、高速問合せ処理機構は、非順序型データベースエンジンを用いて、受付けた問合せを両空間にまたがって実行し、結果を返す。前年度に導入した交通ビッグデータ解析試験サーバ等の実験装置に関しては、本年度、処理サーバを5 ノードから7 ノードへ、アーカイブ容量を360TB から720TB へと拡充した。当該実験装置に於いては、OS やストレージ等のチューニングを実施し、性能試験を実施した。基本性能としては、各サーバで8千 IOPS 程度のストレージアクセススループット、ならびに、サーバ間では15万 OPS 程度のネットワークアクセススループットが確認されており、更に、交通ビッグデータを模擬した1TB から10TB 程度のデータセットを用意し、ドライバーの特性の運転操作を抽出する問合せを複数実行し、その性能を調査した。この結果、いずれも非順序型実行により10倍を超える高速性を確認し、交通ビッグデータに適した処理基盤の有効性を明らかにすることができた。

#### 課題ウ「運転者指向サービス基盤技術の研究開発」:

本課題では、運転者マネジメント支援や災害時情報提供サービス等の運転者指向サービスを実現するための交通ビッグデータ解析手法および可視化手法の確立を目指している。今年度は、運転者指向サービスのひとつの大きな目的である安全運転実現に向けて、ドライバーの運転経歴と運転操作特徴との関連性、並びに道路の特徴と事故履歴との関連性について解析を行った。

ドライバーの運転経歴と運転操作特徴との関連性に関しては、運転操作記録を数値地図にマッチングし、道路状況と運転操作を統合して解析する基礎的な実験を行った。操作に与える影響が大きいと考えられる道路幅の情報を利用し、道路幅ごとに運転操作分布を別々に特徴量化することで、運送業320人のドライバーの事故歴をF値0.82で分類することに成功した。これは、運転操作の分布のみの情報を利用した場合より大きく分類性能を向上させており、異種データ統合解析による有効性を確認することができた。

(27-1)

また、新しく得られたバスの運転操作記録についても提案の解析手法を適用する予備評価を行った。昨年度利用した運送業者とは業務内容も車両特性も大きく異なるため、運転操作分布の性質が大きく異なっているが、バスに合わせた適切な閾値で量子化することで同様の解析結果が得られ、異なる業種のドライバーについても、提案する分析手法が有効であることが確認できた。

ドライブレコーダデータの可視化手法に関しては、交差点等での事故の発生とそこでなされる運転操作との関連性に関する解析を行うため、運転操作記録を地図上に表示し、選択した領域内の操作に関する分布を表示・比較できる可視化システムを構築した。本システムを用いて、事故の発生した交差点としていない交差点との比較を行ったところ、運転操作の特徴量の分布について差異がみられることが確認できた。また、数値地図からは得られない見通しの良さ悪しや一時停止線等の有無といった道路状況に関する特徴が、事故の発生に関連することも確認できた。これらの知見を用い、機械学習により交差点のリスクを予測する予備実験を実施した。