

## 科学研究費助成事業 研究成果報告書

平成 26 年 5 月 25 日現在

機関番号：13901

研究種目：基盤研究(B)

研究期間：2010～2013

課題番号：22300034

研究課題名(和文)クラウド環境における時空間情報資源の動的統合利用に関する研究

研究課題名(英文)Dynamic Integration and Use of Spatio-temporal Information Resources in Cloud Environments

研究代表者

石川 佳治(Ishikawa, Yoshiharu)

名古屋大学・情報科学研究科・教授

研究者番号：80263440

交付決定額(研究期間全体)：(直接経費) 13,900,000円、(間接経費) 4,170,000円

研究成果の概要(和文)：クラウドコンピューティングでは、並列・分散環境において多数のコンピュータを連携してスケーラブルな情報処理が実現される。また、さまざまな情報の連携による高度なサービスが提供される。本研究では、このようなクラウド技術の発展を考慮に入れた大規模時空間データの管理・処理技術に着目した。時空間データの特性を踏まえたクラウド環境上でのデータベース問合せ処理技術や、クラウドシステムの利活用技術などを開発するのみならず、時空間・モバイルデータベースにおけるプライバシー保護などの応用技術についても開発を行った。

研究成果の概要(英文)：In cloud computing, many computers cooperate in a parallel and distributed environment and achieve scalable information processing. In addition, advanced services are provided based on integration of various information. In this research, we focused on management and processing technologies of large-scale spatio-temporal data considering the development of cloud computing. The research topics include spatio-temporal database query processing in cloud environment and utilization technologies of cloud systems considering the features of spatio-temporal data. In addition, we also developed application-level technologies such as privacy protection mechanisms in spatio-temporal and mobile databases.

研究分野：総合領域

科研費の分科・細目：情報学/メディア情報学・データベース

キーワード：クラウドコンピューティング 時空間データベース 情報統合 プライバシー保護 イベント処理 データストリーム 大規模データ 並列分散処理

### 1. 研究開始当初の背景

クラウドコンピューティング (cloud computing) は、低価格・高性能化が一段と進むコンピュータの状況と、ネットワークの一層の高速化を踏まえ、並列分散処理によりスケラブルなデータの処理を実現しようとするを一つの目的としている。また、サービスの側面からは、システムリソースをクラウド環境上に集約することで、さまざまなサービスの広範囲な提供も目指している。

クラウド技術は急速に進展しており、ビジネス分野を中心に大いに活用されているが、時空間データベースに関しては未だ取り組むべき課題が大きい。時間情報、空間情報を扱う時空間データベースでは、時間、空間、またそれらを組み合わせた時空間情報に関するセマンティクスが発生し、また、固有のデータ処理なども必要となる。加えて、時空間データの特性上、地図情報の活用などの問題や、時々刻々と取得されるデータストリームへの対応などの問題も発生する。このような点を考慮すると、クラウド時代における時空間データの管理や処理について、新たな技術開発が求められてきている。

本研究では、このような情報を踏まえて研究を開始した。

### 2. 研究の目的

本研究では、クラウド環境における大規模空間データの管理・処理技術に関する研究開発を行う。研究は、基盤技術、要素技術、応用技術に大まかに大別することができる。

- (1) 基盤技術の開発：クラウド環境における現在の研究開発状況を踏まえたシステム寄りの開発を行う。特に、Hadoop などの、クラウド環境上で主流の並列分散処理ソフトウェアなどを想定したアルゴリズムの開発がその一例となる。また、クラウド環境上では、大規模データの一括処理のみだけでなく、ストリーミング的に訪れるデータに対する処理・問合せも重要となっている。そのため、データストリームやイベントデータの処理技術やシステム技術の開発を行う。また、並列分散環境における時空間データの統合・流通に関する研究も実施する。
- (2) 要素技術の開発：大規模時空間データの処理や活用のために必要となる要素的な技術の開発を行う。一つには、新たな時空間データベース問合せの提案とその処理技術の開発が考えられる。近年、さまざまな応用領域において時空間データが活用されていることから、多様な問合せを考えることができるようになっており、そのような状況を踏まえた開発が必要となっている。また、データストリームや時系列データに対するパターンマッチなどの基礎的なアルゴリズムの開発なども対象として考えられる。
- (3) 応用技術の開発：時空間データを用いた

多様なアプリケーションを考え、そこで必要となる応用技術の開発を進める。一つの課題となるのは時空間アプリケーションで発生するプライバシーの問題である。実際の応用ではプライバシーの問題は深刻であるので、現実の要求に即した技術開発を進める。その他、基盤技術、要素技術の開発で行った技術開発を応用する上で必要となる手法等の開発なども行う。

### 3. 研究の方法

- (1) Hadoop を用いた大規模空間データの問合せ処理方式の開発：クラウド環境におけるデータ処理のフレームワークとして主流になっている Hadoop 上で、空間データに対する問合せおよびデータ処理を効率的に処理するためのアルゴリズムを開発する。また、問合せ処理を容易に記述可能とするためのインタフェースとしての問合せ言語の開発も行う。
- (2) 空間データベース技術に基づく新たな問合せ処理技術の開発：空間データベースにおいては、提供するサービスに応じてさまざまな問合せが開発されているが、本テーマではその延長として、さらに有用な問合せとその効率的な処理手法の開発を行う。
- (3) 曖昧な位置情報に基づく空間データベース問合せ処理：GPS のノイズなど、さまざまな環境で正確な位置情報が把握できない状況がある。曖昧な位置情報であっても、その情報を効果的に用いて問合せ処理を行うための技術開発を行う。特に、位置情報が確率分布として表現される場合の問合せ処理手法を開発する。
- (4) 位置情報サービスにおけるプライバシー保護技術の開発：空間データベースの応用技術の一つとして、モバイルユーザの位置に応じて情報を提供する位置情報サービスがある。位置情報を用いることでユーザにきめ細かなサービスができる反面、ユーザの位置というプライバシー情報が漏れることになる。そのため、プライバシー保護技術の開発を行う。
- (5) 時空間データベースにおけるデータストリーム処理およびイベント処理技術の開発：時空間データベースにおける応用を考慮する上では、センサ等から時々刻々と入力されるデータストリームの処理が重要となっている。また、より高次の概念であるイベント検出とその利用についても研究開発が盛んに進められている。このような流れを踏まえた問合せ処理技術の開発を行う。
- (6) 時空間データを対象とした P2P データ統合に関する研究：P2P (Peer-to-Peer) 技術は、柔軟なネットワークを動的に構築するための技術である。P2P 環境を対象として、時空間データを流通させたり統

合したりすることを可能にするための技術開発を行う。

#### 4. 研究成果

- (1) Hadoop を用いた大規模空間データの問合せ処理方式の開発：特に全  $k$  最近傍問合せについて、Hadoop 上で問合せ処理を効率的に実施可能なアルゴリズムを開発した。この問合せは、たとえば「データベース中のそれぞれの駅について、もより駅を 5 件ずつ列挙せよ」といったものである。また、Hadoop 上の空間データ問合せ言語として、Pig をベースとした言語についての検討を行った。
- (2) 空間データベース技術に基づく新たな問合せ処理技術の開発：空間スカイライン問合せの考え方を参考にして、方向に基づく周辺問合せを形式化し、効率的なアルゴリズムを開発した。問合せ点の近くにある点を、各方向まんべんなく検索する手法である。移動環境にも適応できるアルゴリズムとなっている。また、空間索引を用いてアイテムの最適な組合せを求める組合せ問合せについての開発も行った。
- (3) 曖昧な位置情報に基づく空間データベース問合せ処理：特に、位置情報がガウス分布で表現される場合について、研究開発を実施した。移動ロボットの位置表現などでガウス分布による位置表現がしばしば現れる。本研究では、確率的最近傍問合せや確率的範囲問合せを効率的に処理するための索引構造や問合せ処理手法の開発を行った。
- (4) 位置情報サービスにおけるプライバシー保護技術の開発：特に着目したのは、モバイルユーザの位置情報だけでなく、属性（例：性別、年齢）も用いる位置情報サービスである。位置情報だけでなく属性情報も合わせて匿名化可能な手法の提案を行った。良く知られている  $k$  匿名化の概念を拡張している。
- (5) 時空間データベースにおけるデータストリーム処理およびイベント処理技術の開発：センサなどから取得された曖昧なデータストリームを想定して、確率的データストリームに対するパターンマッチ手法を提案し、その実現技術を開発した。また、高次の複合イベントに対する処理技術、特にオントロジの活用による背景知識の導入を狙った。これら以外に、時系列データ（データストリームの一種）に対する柔軟なパターンマッチを効率的に実行する手法を開発した。
- (6) 時空間データを対象とした P2P データ統合に関する研究：特に、P2P 環境における情報流通を対象とし、情報流通のトレーサビリティに関する研究開発を行った。P2P 上の各ピア上で情報交換に関する履歴情報を保持しておき、トレーサ要

求が発生した際には、ピアどうしの協力により追跡問合せを実行することによりそれを実現する。

#### 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 13 件)

Tingting Dong, Yoshiharu Ishikawa, Chuan Xiao. Probabilistic Range Querying over Gaussian Objects. *IEICE Transactions on Information and Systems*, Vol. E97-D, No. 4, pp. 694-704, Apr. 2014 (査読有)

DOI: 10.1587/transinf.E97.D.694

石川 佳治. 大規模データアナリティクスに対する研究動向と展望. 電子情報通信学会論文誌, Vol. J97-D, No. 4, pp. 718-728, 2014 年 4 月. (招待論文, 査読無)

Machiko Toyoda, Yasushi Sakurai, Yoshiharu Ishikawa. Pattern Discovery in Data Streams under the Time Warping Distance. *The VLDB Journal*, Vol. 22, No. 3, pp. 295-318, Jun. 2013. (査読有)

DOI: 10.1007/s00778-012-0289-3

加藤 翔, 石川 佳治. 確率的なイベントストリームに対するイベントパターン問合せ. 日本データベース学会論文誌, Vol. 12, No. 1, pp. 55-60, 2013 年 6 月. (査読有)

Chuan Xiao, Jianbin Qin, Wei Wang, Yoshiharu Ishikawa, Koji Tsuda, Kunihiko Sadakane. Efficient Error-tolerant Query Autocompletion. *Proceedings of the VLDB Endowment (PVLDB)*, Vol. 6, No. 6, pp. 373-384, Apr. 2013. (査読有)

Xi Guo, Chuan Xiao, Yoshiharu Ishikawa. Combination Skyline Queries. *Transactions on Large-Scale Data- and Knowledge-Centered Systems VI*, pp. 1-30, Sept. 2012. (査読有)

横山 拓也, 石川 佳治, 鈴木 優. Hadoop 環境における空間分割による並列全  $k$  近傍問合せ処理. 日本データベース学会論文誌, Vol. 11, No. 1, pp. 25-30, 2012 年 6 月. (査読有)

Yanwei Xu, Yoshiharu Ishikawa, Jihong Guan. Efficient Continual Top- $k$  Keyword Search in Relational Databases. *Journal of Information Processing*, Vol. 20, No. 1, pp. 114-127, IPSJ, Jan. 2012. (査読有)

Xi Guo, Baihua Zheng, Yoshiharu Ishikawa, Yunjun Gao. Direction-Based Surrounding Queries for Mobile Recommendations. *The VLDB Journal*, 20(5), pp. 743-766, Oct. 2011. (査読

有)  
豊田 真智子, 櫻井 保志, 石川 佳治 . 部分シーケンスマッチングのためのストリームアルゴリズム . 電子情報通信学会論文誌 D, Vol. J94-D, No. 7, pp. 1058-1070, 2011年7月 . (査読有)  
Fengrong Li, Yoshiharu Ishikawa. Using Materialized Views to Enhance a Traceable P2P Record Exchange Framework. *Journal of Advances in Information Technology*, Vol. 2, No. 1, pp. 27-39, Jan. 2011. (査読有)  
飯島 裕一, 石川 佳治 . あいまいな位置情報に基づく最近傍問合せの処理手法 . 電子情報通信学会論文誌, Vol. J93-D, No. 6, pp. 781-794, 2010年6月 . (査読有)  
Fengrong Li, Yoshiharu Ishikawa. Query Processing in a Traceable P2P Record Exchange Framework. *IEICE Transactions on Information and Systems*, Vol. E93-D, No. 6, pp. 1433-1446, Jun. 2010. (査読有)

[学会発表](計43件)

高橋 正和, 築井 美咲, 稲葉 鉄平, 石川 佳治 . オントロジーを利用したイベント処理システムの提案 . 情報処理学会第76回全国大会, 2014年3月 (査読無)  
姜 仁河, 杉山 武至, 石川 佳治 . 人気経路の推薦のための大規模移動軌跡データ処理 . 情報処理学会第76回全国大会, 2014年3月 . (査読無)  
杉浦 健人, 早矢仕 新, 石川 佳治 . 確率的ストリームにおけるグループを用いたパターン問合せ . 情報処理学会第76回全国大会, 2014年3月 . (査読無)  
趙 セイ, 董 テイテイ, 石川 佳治 . 参加型センシングにおけるプライバシー保護手法 . 情報処理学会第76回全国大会, 2014年3月 . (査読無)  
早矢仕 新, 杉浦 健人, 董 ていてい, 石川 佳治 . パーティクル表現を用いた曖昧位置情報に対する空間問合せ処理 . 第6回データ工学と情報マネジメントに関するフォーラム (DEIM 2014), 2014年3月 . (査読無)  
稲葉 鉄平, 高橋 正和, 築井 美咲, 石川 佳治 . オントロジーに基づくLBSN上でのイベント検出 . 第12回情報科学技術フォーラム (FIT 2013), 2013年9月 . (査読無)  
早矢仕 新, 杉浦 健人, 董 ていてい, 石川 佳治 . 曖昧な移動軌跡に対する範囲問合せ . 第12回情報科学技術フォーラム (FIT 2013), 2013年9月 . (査読無)  
Tingting Dong, Chuan Xiao, Xi Guo, Yoshiharu Ishikawa. Processing Probabilistic Range Queries over Gaussian-based Uncertain Data. *The*

*13th International Symposium on Spatial and Temporal Databases (SSTD 2013)*, pp. 410-428, Munich, Germany, Aug. 2013. (査読有)  
石川 佳治, 杉山 武至, 鈴木 優 . Hadoop上の空間データ問合せ処理フレームワーク . 情報処理学会第75回全国大会, 2013年3月 . (査読無)  
高橋 正和, 稲葉 鉄平, 石川 佳治 . 時空間イベントを扱う複合イベント処理システム . 情報処理学会第75回全国大会, 2013年3月 . (査読無)  
加藤 翔, 石川 佳治 . 確率的なイベントストリームに対するイベントパターン問合せ . 第5回データ工学と情報マネジメントに関するフォーラム (DEIM 2013), 2013年3月 . (査読無)  
加藤 翔, 石川 佳治 . 確率的イベント系列からの複合的イベント検出モデルについて . 情報処理学会データベースシステム研究会, 2012年12月 . (査読無)  
Masanori Mano, Xi Guo, Tingting Dong, Yoshiharu Ishikawa. Privacy Preservation for Location-Based Services Based on Attribute Visibility. *International Workshop on Information Management in Mobile Applications (IMMoA 2012)*, pp. 33-41, Istanbul, Turkey, Aug. 2012. (査読有)  
加藤 翔, 石川 佳治 . 確率的データストリームに対する問合せ処理手法 . 電子情報通信学会データ工学研究会, 2012年8月 . (査読無)  
豊田 真智子, 櫻井 保志, 石川 佳治 . データストリームにおける効率的なパターン検出 . 情報処理学会データベースシステム研究会, 2012年8月 . (査読無)  
Takuya Yokoyama, Yoshiharu Ishikawa, Yu Suzuki. Processing All  $k$ -Nearest Neighbor Queries in Hadoop. *The 13th International Conference on Web-Age Information Management (WAIM 2012)*, pp. 346-351, Harbin, China, Aug. 2012. (査読有)  
Yanwei Xu, Jihong Guan, Yoshiharu Ishikawa. Scalable Top-k Keyword Search in Relational Databases. *The 17th International Conference on Database Systems for Advanced Applications (DASFAA 2012)*, Part II, pp. 65-80, Busan, Korea, Apr. 2012. (査読有)  
杉山 武至, 横山 拓也, 鈴木 優, 石川 佳治 . Hadoopを用いた大規模空間シミュレーションのためのフレームワークの提案 . 情報処理学会第74回全国大会, 2012年3月 . (査読無)  
加藤 翔, 董 ていてい, 早矢仕 新, 石川 佳治 . 確率的な位置情報に基づくイベント問合せ . 情報処理学会第74回全

- 国大会, 2012年3月。(査読無)
- 横山 拓也, 石川 佳治, 鈴木 優. Hadoop環境における空間分割による並列全k近傍問合せ処理. 第4回データ工学と情報マネジメントに関するフォーラム (DEIM 2012), 2012年3月。(査読無)
- 21 眞野 将徳, 石川 佳治. 位置情報サービスにおける属性の可観測性を考慮したプライバシー保護手法. 第4回データ工学と情報マネジメントに関するフォーラム (DEIM 2012), 2012年3月。(査読無)
- 22 Yoshiharu Ishikawa. Adaptive Spatial Query Processing Based on Uncertain Location Information. *The 7th International Workshop on Databases in Networked Information Systems (DNIS 2011)*, Aizu, Fukushima, Dec. 2011. (招待講演, 査読無)
- 23 Yoshiharu Ishikawa. Adaptive Spatial Query Processing for Supporting Mobile Users' Decisions. *The 3rd International Workshop on Mobile Information Retrieval for Future (MIRF2011)*, Tokyo, Nov. 2011. (招待講演, 査読無)
- 24 Fengrong Li, Yoshiharu Ishikawa. Simulation Based Analysis for a Traceable P2P Record Exchange Framework. *The 4th International Conference on Data Management in Grid and P2P Systems (Globe 2011)*, pp. 49-60, Tournouse, France, Sept. 2011. (査読有)
- 25 加藤 翔, 石川 佳治. RDB技術に基づくストリームデータ問合せ処理. 第10回情報科学技術フォーラム (FIT 2011), 2011年9月。(査読無)
- 26 Xi Guo, Yoshiharu Ishikawa. Multi-Objective Optimal Combination Queries. *The 22nd International Conference on Database and Expert Systems (DEXA 2011)*, Part I, pp. 47-61, Tournouse, France, Aug.-Sept. 2011. (査読有)
- 27 Kazuki Kodama, Tingting Dong, Yoshiharu Ishikawa. An Index Structure for Spatial Range Querying on Gaussian Distributions. *The Fifth International VLDB Workshop on Management of Uncertain Data (MUD 2011)*, pp. 1-7, Seattle, WA, Aug. 2011. (査読有)
- 28 眞野 将徳, 石川 佳治. 位置情報サービスにおける属性を考慮したプライバシー保護について. 電子情報通信学会データ工学研究会, 2011年8月。(査読無)
- 29 横山 拓也, 石川 佳治. 空間分割による並列k近傍問合せの提案. 情報処理学会データベースシステム研究会, 2011年8月。(査読無)
- 30 堀田 孝司, 石川 佳治, 眞野将徳. RDBを用いた移動履歴からの移動パターン問合せ処理手法. 第3回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2011), 2011年3月。(査読無)
- 31 兒玉 一樹, 石川 佳治. ガウス分布に対する空間索引のGiSTを用いた実現. 第3回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2011), 2011年3月。(査読無)
- 32 眞野 将徳, 石川 佳治. 位置の匿名性に関する研究について. 第8回情報学ワークショップ (WiNF 2010), pp. 147-152, 2010年12月。(査読無)
- 33 横山 拓也, 石川 佳治. 大規模空間情報処理に対応するための分散処理フレームワークPigの拡張. 第8回情報学ワークショップ (WiNF 2010), pp. 221-224, 2010年12月。(査読無)
- 34 Masanori Mano, Yoshiharu Ishikawa. Anonymizing User Location and Profile Information for Privacy-aware Mobile Services. *The 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks (LBSN 2010)*, pp. 68-75, San Jose, CA, Nov. 2010. (査読有)
- 35 石川 佳治. 時空間・センサデータベースに関する研究. Webとデータベースに関するフォーラム (WebDB 2010), 2010年11月。(ポスター発表, 査読無)
- 36 眞野 将徳, 石川 佳治. 位置情報サービスのためのプライバシーを考慮した匿名化処理について. 第9回情報科学技術フォーラム (FIT 2010), 2010年9月。(査読無)
- 37 横山 拓也, 石川 佳治. 大規模空間情報処理のための分散コンピューティング基盤に関する検討, 第9回情報科学技術フォーラム (FIT 2010), 2010年9月。(査読無)
- 38 Xi Guo, Yoshiharu Ishikawa, Yunjun Gao. Direction-Based Spatial Skylines. *The 2nd International Workshop with Mentors on Databases, Web and Information Management for Young Researchers (iDB Workshop 2010)*, pp. 11-22, Tokyo, Japan, Aug. 2010. (査読無)
- 39 Yanwei Xu, Yoshiharu Ishikawa, Jihong Guan. Efficient Continuous Top-k Keyword Search in Relational Databases. *The 11th International Conference on Web-Age Information Management (WAIM 2010)*, pp. 755-767, Jiuzhaigou, China, Jul. 2010. (査読有)
- DOI: 10.1007/978-3-642-14246-8\_71
- 40 Fengrong Li, Yoshiharu Ishikawa.

Query Processing with Materialized Views in a Traceable P2P Record Exchange Framework. *WAIM 2010 International Workshop (XMLDM 2010)*, pp. 246-257, Jiuzhaigou, China, Jul. 2010. (査読有)

- 41 Xi Guo, Yoshiharu Ishikawa, Yunjun Gao. Direction-Based Spatial Skylines. *Proceedings of Ninth International ACM Workshop on Data Engineering for Wireless and Mobile Access (MobiDE 2010)*, pp. 73-80, Indianapolis, IN, Jun. 2010. (査読有)
- 42 堀田 孝司, 石川 佳治. 確率的移動軌跡に対する移動パターン問合せ処理手法. 電子情報通信学会データ工学研究会, 2010年6月.(査読無)
- 43 Yoshiharu Ishikawa. Traceable P2P Record Exchange. *The 5th Korea-Japan Database Workshop (KJDB 2010)*, Jeju Island, Korea, May 2010. (招待講演, 査読無)

## 6. 研究組織

### (1) 研究代表者

石川 佳治 (ISHIKAWA, Yoshiharu)  
名古屋大学・大学院情報科学研究科・教授  
研究者番号: 80263440