

様 式 C - 19、F - 19 - 1、Z - 19 (共通)

## 科学研究費助成事業

## 研究成果報告書



平成 29 年 5 月 9 日現在

機関番号：12102

研究種目：基盤研究(B) (一般)

研究期間：2014～2016

課題番号：26280037

研究課題名(和文)複合型並列計算環境を活用した大規模不均質データの実時間分析基盤

研究課題名(英文)Real-time Analytics Frameworks for Big Heterogeneous Data in Composite Parallel Computing Environments

研究代表者

北川 博之(KITAGAWA, HIROYUKI)

筑波大学・計算科学研究センター・教授

研究者番号：00204876

交付決定額(研究期間全体)：(直接経費) 13,200,000円

研究成果の概要(和文)：本研究では、複合型並列計算環境を活用した大規模不均質データの実時間分析基盤の構築に関する研究を行った。大規模不均質データ分析のためのデータモデル・データ処理記述、同モデル化を実現するシステムアーキテクチャ、GPUを利用し超並列データ分析処理、メニーコアプロセッサを用いた高性能データ分析、差分計算を用いた効率的データ分析処理、半構造データストリームに対する構造とキーワードを併用した検索、複数のテキストストリームの実時間照合、複数のリレーションストリームを対象としたキーワード検索、小粒度処理と大粒度処理を含むデータ分析の最適実行プラン生成等に関して研究成果を得た。

研究成果の概要(英文)：The goal is to develop real-time analytics frameworks for big heterogeneous data in composite parallel computing environments. The main research outcomes cover the following topics: data modeling and data description for big heterogeneous data analytics, system architecture implementing the data model, massively parallel GPU-based data analytics, high performance data analytics using many core processors, efficient data analytics using incremental computation, semistructured stream filtering using intrinsic structures and keywords, real-time matching of multiple text streams, keyword query processing of multiple relational streams, and generation of optimal data analytics plans including fine grain and coarse grain data processing.

研究分野：データ工学

キーワード：大規模不均質データ データ分析 複合型並列計算環境

## 1. 研究開始当初の背景

情報システムの扱うデータの量は指数関数的に増大し、テラバイトを超えるデータはもはや珍しくない。これらのデータの情報源は多種多様であり、種類や様式が極めて多岐に渡る不均質データである。一方、モバイルデバイスやセンサ等から実時間で大量のデータがストリームとして連続的に取得される。これらの実時間性をもつ不均質大規模データの効率的分析による新たな価値創造に大きな期待が寄せられている。一方、計算機資源の高度化も急激に進んでおり、従来の無共有型クラスタに加え、メニーコアや複数GPU等の演算加速機構や高速ノード間通信を有する複合型高性能並列計算機等が利用可能になりつつある。複雑な構造を有する大規模不均質データの実時間分析のためには、これら最新の計算機資源の活用が必須である。本研究は、ペタバイトを超える大規模不均質データに対する様々な実時間分析を、複合型高性能並列計算機を含む計算機環境（以下、複合型並列計算環境）を活用して支援する基盤技術の構築を目指した。

## 2. 研究の目的

並列計算機環境を活用したデータ分析処理に、MapReduce や SQL 風記述が可能な Hive 等が使われている。しかし、これらは並列処理プログラム記述の枠組みに過ぎず、複雑なデータ処理に対応可能なデータ管理・処理機構はもたない。また、バッチ処理が基本であり実時間分析に対応できない。さらに、無共有型クラスタ上でのデータ並列処理のみを実現しており、複合型並列計算機のもつ能力を活用したものではない。大規模データ分析システムとして、IBM の Jaql、カリフォルニア大の ASTERIX 等が開発された。これらは、JSON 等に基づく軽量な不均質データモデルを採用し、多様な分析処理を実現する。しかし、いずれもバッチ処理が基本であり実時間分析に対応できない。また、MapReduce 等と同様、無共有型クラスタ上でのデータ並列処理のみを実現している。一方、実時間ストリーム処理に関しては、これまで多くの処理エンジンが開発、実用化されてきたが、単純なタプル形式のデータのみを対象とするものや、対象データの扱いを利用者のプログラミングに委ねるものがほとんどであり、不均質ストリームデータの扱いを系統的にサポートするものは見当たらない。

本研究では、研究代表者らのこれまでの研究の知見に大規模並列計算を融合し、複合型並列計算環境を活用した大規模不均質データの実時間分析基盤の構築を目的とした。本研究では、特に、メニーコアを活用するためのタスク分割やタスク多重化、マルチGPUを用いた超並列検索や計算インテンシブ処理の加速等、従来の並列処理とは異なる手法を

組み合わせたスケーラブルな処理を目指した。実時間データ分析には、ストリームとして取得される実時間データのモニタリング等の比較的小粒度の処理と、既に蓄積された大規模データとストリームデータ等を連携させた大粒度処理があると考えられる。本研究では、不均質データを対象とした、1) ストリーム処理をベースとした超多数の小粒度処理の並列同時実行技術、2) 並列差分計算をベースとした超大粒度処理の高速化、3) 小粒度処理と大粒度処理を融合した多様な実時間大規模分析の実現を目的とした。

## 3. 研究の方法

A. 大規模不均質データ分析フレームワークに関する研究項目

【研究項目 A-1: データモデル・データ処理記述】

JSON 等の軽量なモデルを用いてスキーマのない不均質データをモデル化すると共に、多様なデータ分析処理フローを記述可能なデータ処理記述系を開発する。その際、ストリーム処理を基本とする小粒度処理と差分処理に基づく大粒度処理が混在する実時間分析処理を統一的に記述するための手法を明らかにする。

【研究項目 A-2: システムアーキテクチャ】  
複合型並列計算環境において、上記のデータモデルやデータ処理記述を実装するシステムアーキテクチャ、大規模不均質データ処理機構のあり方を明らかにする。

B. 大粒度・小粒度分析処理の高速化に関する研究項目

【研究項目 B-1: 大粒度分析の並列処理】  
複合型並列計算環境を利用した不均質データの並列分析アルゴリズム、データ処理記述の解析に基づく自動並列化の手法、メニーコアを活用するためのタスク分割手法、マルチGPUを用いた計算インテンシブ処理の高速化等について明らかにする。

【研究項目 B-2: 並列差分処理】  
大粒度実時間分析を効率化する上で重要となる差分計算に基づく大規模分析処理の実現方法の手法を明らかにする。

【研究項目 B-3: 分散ストリーム処理に基づく小粒度処理】

複合型並列計算環境を利用した不均質データストリームの分散処理方式、データ処理記述解析に基づく自動分散ストリーム処理手法を明らかにする。

C. 大粒度・小粒度分析処理の融合に関する研究項目

【研究項目 C: 大粒度・小粒度分析処理の融合】

実時間制約や分析粒度要求等を加味したデータ処理記述解析により、大粒度・小粒度分析を融合したデータ分析処理フローの実現方法を明らかにする。

#### 4. 研究成果

##### A. 大規模不均質データ分析フレームワークに関する研究項目

###### 【研究項目 A-1: データモデル・データ処理記述】

不均質データの代表として JSON を取り上げ、ストア型 JSON データとストリーム型 JSON データを対象に、ストリーム処理を基本とする小粒度処理と差分処理に基づく大粒度処理が混在する多様なデータ分析処理フローを記述可能なデータ処理記述方式とその演算処理体系を定式化した。さらに、データ分析処理フロー記述を、ユーザ定義処理やタスク等の概念を導入して拡張した処理体系を定式化した。[論文 5, 学会発表 5, 11, 16, 24]

長いテキストを要素内容とする XML データを取り上げ、ファセット検索に基づく不均質データフレームワークの開発を行った。[論文 10, 17, 学会発表 28]

【研究項目 A-2: システムアーキテクチャ】上記でモデル化した処理体系を実現するシステムアーキテクチャとして JsFlow を提案し、そのプロトタイプシステム実装を行った。JsFlow では、JSON データに対する小粒度処理と大粒度処理が混在する処理記述を解析し、実際に記述された分析処理を実行可能である。プロトタイプシステム開発では、汎用的なデータ処理基盤である Spark や Flink を用い、複数のタスクを対象とした評価実験を行った。[論文 6, 学会発表 5, 11, 16, 24]

##### B. 大粒度・小粒度分析処理の高速化に関する研究項目

###### 【研究項目 B-1: 大粒度分析の並列処理】

GPU を利用した大規模不均質データ並列分析処理方式の開発を行った。具体的には、Canopy クラスタリング、集合間類似結合、SIFT 特徴量を用いた類似画像検索、ラベル伝搬を利用したグラフクラスタリング等の問題において GPU を用いた高速化の有効性を示した。[論文 3, 6, 13, 14, 16, 学会発表 7, 12, 15, 19, 21, 22, 26]

メニーコアプロセッサを用いたデータ分析高速化についても研究を推進した。具体的には、構造的類似度に基づくグラフクラスタリング、集合間類似結合の高速化である。[論文 1, 学会発表 1, 4, 6]

大規模グラフに対する ObjectRank 計算の高速化手法を開発した。[学会発表 3, 8]

###### 【研究項目 B-2: 並列差分処理】

差分計算に基づく大規模分析処理として、連続的な外れ値検出手法を開発した。一つは、動的な環境における角度ベースの外れ値検出手法の開発である。もう一つは、動的なデータ集合に対する対話的外れ値分析フレームワーク DIO の開発である。[論文 15, 学会発表 10, 23]

移動体等の並列差分処理環境におけるトップk検索において重要となる、逆近傍検索の手法を開発した。[論文 2, 学会発表 2]

###### 【研究項目 B-3: 分散ストリーム処理に基づく

##### 小粒度処理】

不均質データストリームの例である XML ストリームを対象に、XPath で指定された構造に関する条件と、キーワードで指定された内容に関する条件を併用した、XML ストリーム分析の手法等を開発した。[論文 11, 18, 学会発表 27, 29]

複数のテキストストリームを高速に照合するシステムの開発を行った。具体例として、Web 上のオンラインニュースと Twitter ストリームを照合し、各ニュースに対する関連ツイートを獲得するシステムの開発を行った。また、関連するトピックとして、オンラインニュースの地域局所的な注目度の分析の手法を開発した。[論文 4, 9, 12, 学会発表 9, 17, 20, 25]

複数のリレーションストリームを対象としたキーワード検索において、キーワード数やリレーション数が増加しても性能劣化が生じにくい新たな手法の開発を行った。[論文 7, 学会発表 13, 18]

##### C. 大粒度・小粒度分析処理の融合に関する研究項目

###### 【研究項目 C: 大粒度・小粒度分析処理の融合】

上記 JSON データを対象としたデータ分析フレームワークにおいて実時間制約や分析粒度要求等を加味したデータ処理記述の解析により、小粒度処理と大粒度処理を適切に切り分け、並列差分処理や中間データの維持管理をも考慮した実行プランを生成するための基本方式を開発した。上記 JsFlow において、計算資源制約や分析粒度要求等を加味したデータ処理記述の解析により、小粒度処理と大粒度処理を適切に切り分け、並列差分処理や中間データの維持管理をも考慮した実行プランを生成するための基本方式の開発を行い、実験評価によりその有効性を確認した。[論文 5, 学会発表 5, 11, 16, 24]

#### 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 18 件)

1. Tomokatsu Takahashi, Hiroaki Shiokawa, Hiroyuki Kitagawa, "SCAN-XP: Parallel Structural Graph Clustering Algorithm on Intel Xeon Phi Coprocessors", Proc. 2nd ACM SIGMOD Workshop on Network Data Analytics (NDA 2017), 査読有, 2017. (to appear)

2. Yuyang Dong, Hanxiong Chen, Jeffrey Xu Yu, Kazutaka Furuse, Hiroyuki Kitagawa, "Grid-Index Algorithm for Reverse Rank Queries", Proc. 20th International Conference on Extending Database Technology (EDBT 2017), 査読有, 2017, pp.306-317.

<https://openproceedings.org/2017/conf/e>

dbt/paper-11.pdf

3. Mateus S. H. Cruz, Yusuke Kozawa, Toshiyuki Amagasa, Hiroyuki Kitagawa, "Accelerating Set Similarity Joins Using GPUs", Transactions on Large-Scale Data and Knowledge-Centered Systems XXVIII: Special Issue on Database and Expert Systems Applications (TLDKS), 査読有, vol.9940,2016,pp.1-22.

DOI. 10.1007/978-3-662-53455-7\_1

4. Saki Nagaki, Yuto Yamaguchi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Local Attention Analysis and Prediction of Online News Articles in Twitter", International Workshop On Mobile Ubiquitous Systems, Infrastructures, Communications, And Applications (MUSICAL 2016), 査読有, 2016, pp.131-136.

DOI.10.1145/3004010.3004042

5. Hirotoshi Cho, Hiroaki Shiokawa and Hiroyuki Kitagawa, "JsFlow: Integration of Massive Streams and Batches via JSON-based Dataflow Algebra", Proc. 19th International Conference on Network-Based Information Systems (NBIS 2016), 査読有,2016, pp.188-195.

DOI. 10.1109/NBIS.2016.1

6. Yuta Kusamura, Yusuke Kozawa, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "GPU Acceleration of Content-based Image Retrieval based on SIFT Descriptors", Proc. the 5th International Workshop on Advances in Data Engineering and Mobile Computing in conjunction with NBIS2016(DEMOc2016), 査読有,2016, pp.342-347.

DOI. 10.1109/NBIS.2016.55

7. Savong Bou, Toshiyuki Amagasa, Hiroyuki Kitagawa, "An Improved Method of Keyword Search over Relational Data Streams by Aggressive Candidate Network Consolidation", Proc. 27th International Conference on Database and Expert Systems Applications (DEXA 2016), 査読有,2016, pp. 336-351.

DOI. 10.1007/978-3-319-44403-1\_21

8. Kento Kawai, Hiroyuki Kitagawa, "Collaborative Filtering with Implicit Feedbacks by Discounting Positive Feedbacks", Proc. 2nd IEEE International Conference on Multimedia Big Data (IEEE BigMM 2016), 査読有,2016, pp.41-48.

DOI. 10.1109/BigMM.2016.30

9. 大西 誠, 北川 博之, "オンラインニュースに関連するツイートのリアルタイムな収集",日本データベース学会和文論文誌, 査読有, Vol.14, No.4, Article No.15, 2016.

10. Takahiro Komamizu, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Facet-value

Extraction Scheme from Textual Contents in XML Data", International Journal of Web Information Systems (IJWIS), 査読有,2015, Vol. 11, Iss. 3, pp.270-290.

DOI.http://dx.doi.org/10.1108/IJWIS-04-2015-0012

11. Savong Bou, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Path-based Keyword Search over XML Streams", International Journal of Web Information Systems (IJWIS), 査読有,2015, Vol. 11, Iss. 3, pp.347-369.

DOI.http://dx.doi.org/10.1108/IJWIS-04-2015-0013

12. Sei Onishi, Yuto Yamaguchi, and Hiroyuki Kitagawa, "Real-Time Relevance Matching of News and Tweets", Proc. 23th International Conference on Cooperative Information Systems (CoopIS 2015), 査読有,2015, pp. 109-126.

DOI. 10.1007/978-3-319-26148-5\_7

13. Mateus S. H. Cruz, Yusuke Kozawa, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "GPU Acceleration of Set Similarity Joins", Proc. 26th International Conference on Database and Expert Systems Applications (DEXA 2015), 査読有,2015, pp. 384-398.

DOI. 10.1007/978-3-319-22849-5\_26

14. Yusuke Kozawa, Fumitaka Hayashi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Parallel Canopy Clustering on GPUs", Proc. 26th International Conference on Database and Expert Systems Applications (DEXA 2015), 査読有,2015, pp. 334-348.

DOI. 10.1007/978-3-319-22849-5\_23

15. Hao Ye, Hiroyuki Kitagawa, and Jun Xiao, "Continuous Angle-based Outlier Detection on High-dimensional Data Streams", Proc. the 19th International Database Engineering & Applications Symposium (IDEAS 2015), 査読有,2015, pp. 162-167.

DOI. 10.1145/2790755.2790775

16. 林 史尊, 小澤 佑介, 天笠 俊之, 北川 博之, "GPU を用いた Canopy クラスタリングの高速化", 日本データベース学会和文論文誌, 査読有, Vol. 13-J, No. 2, 2015,pp. 13-18.

17. Takahiro Komamizu, Toshiyuki Amagasa and Hiroyuki Kitagawa, "Extracting Facets from Textual Contents for Faceted Search over XML Data," Proc. 16th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2014), 査読有,2014, pp. 420-429.

DOI. 10.1145/2684200.2684294

18. Savong Bou, Toshiyuki Amagasa and Hiroyuki Kitagawa, "Filtering XML Streams by XPath and Keywords," Proc. 16th International Conference on Information Integration and Web-based Applications &

Services (iiWAS 2014), 査読有, 2014, pp. 410-419.

DOI. 10.1145/2684200.2684309

〔学会発表〕(計 29 件)

1. Tomokatsu Takahashi, Hiroaki Shiokawa, Hiroyuki Kitagawa, "SCAN-XP: Parallel Structural Graph Clustering Algorithm on Intel Xeon Phi Coprocessors", 2nd ACM SIGMOD Workshop on Network Data Analytics (NDA 2017), 2017 年 5 月 19 日, シカゴ(アメリカ合衆国).

2. Yuyang Dong, Hanxiong Chen, Jeffrey Xu Yu, Kazutaka Furuse, Hiroyuki Kitagawa, "Grid-Index Algorithm for Reverse Rank Queries", 20th International Conference on Extending Database Technology (EDBT 2017), 2017 年 3 月 24 日, ベネチア(イタリア).

3. 佐藤 朋紀, 塩川 浩昭, 北川 博之, "大規模グラフに対する逐次的なノードの枝刈りを用いた ObjectRank の高速化", 情報処理学会第 79 回全国大会 (IPSJ 全国大会 2017), 2017 年 3 月 16 日, 名古屋大学(愛知県名古屋市).

4. 菅野 健太, 天笠 俊之, 北川 博之, "メニーコアプロセッサを用いた集合間類似結合の高速化", 情報処理学会第 79 回全国大会 (IPSJ 全国大会 2017), 2017 年 3 月 16 日, 名古屋大学(愛知県名古屋市).

5. 長 裕敏, 塩川 浩昭, 北川 博之, "ストリーム処理とバッチ処理の統合と実行最適化", 第 9 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 2017), 2017 年 3 月 6 日, 高山グリーンホテル(岐阜県高山市).

6. 高橋 知克, 塩川 浩昭, 北川 博之, "メニーコアプロセッサを用いた構造的類似度に基づくグラフクラスタリングの高速化", 第 9 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 2017), 2017 年 3 月 6 日, 高山グリーンホテル(岐阜県高山市).

7. 草村 優太, 天笠 俊之, 北川 博之, "SIFT 特徴量を用いた画像検索の GPU による高速化", 第 9 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 2017), 2017 年 3 月 7 日, 高山グリーンホテル(岐阜県高山市).

8. 佐藤 朋紀, 塩川 浩昭, 山口 祐人, 北川 博之, "大規模グラフに対する ObjectRank の高速化", 第 9 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 2017), 2017 年 3 月 8 日, 高山グリーンホテル(岐阜県高山市).

9. Saki Nagaki, Yuto Yamaguchi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Local Attention Analysis and Prediction of Online News Articles in Twitter", International Workshop On Mobile Ubiquitous Systems, Infrastructures, Communications, And Applications (MUSICAL

2016), 2016 年 11 月 28 日, 広島国際会議場(広島県広島市).

10. 坂詰知完, 北川 博之, 天笠 俊之, "データの動的な変化に対応可能な対話的外れ値分析", 第 9 回 Web とデータベースに関するフォーラム (WebDB Forum 2016), 2016 年 9 月 14 日, 慶應義塾大学(神奈川県横浜市).

11. Hirotooshi Cho, Hiroaki Shiokawa and Hiroyuki Kitagawa, "JsFlow: Integration of Massive Streams and Batches via JSON-based Dataflow Algebra", 19th International Conference on Network-Based Information Systems (NBIS 2016), 2016 年 9 月 8 日, オストラバ(チェコ).

12. Yuta Kusamura, Yusuke Kozawa, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "GPU Acceleration of Content-based Image Retrieval based on SIFT Descriptors", 5th International Workshop on Advances in Data Engineering and Mobile Computing in conjunction with NBIS2016 (DEMoC 2016), 2016 年 9 月 8 日, オストラバ(チェコ).

13. Savong Bou, Toshiyuki Amagasa, Hiroyuki Kitagawa, "An Improved Method of Keyword Search over Relational Data Streams by Aggressive Candidate Network Consolidation", 27th International Conference on Database and Expert Systems Applications (DEXA 2016), 2016 年 9 月 6 日, オストラバ(チェコ).

14. Kento Kawai, Hiroyuki Kitagawa, "Collaborative Filtering with Implicit Feedbacks by Discounting Positive Feedbacks", 2nd IEEE International Conference on Multimedia Big Data (IEEE BigMM 2016), 2016 年 4 月 20 日, 台北(台湾).

15. 草村 優太, 小澤 佑介, 天笠 俊之, 北川 博之, "GPU を用いた類似画像検索の高速化", 情報処理学会第 78 回全国大会 (IPSJ 全国大会 2016), 2016 年 3 月 11 日, 慶應義塾大学(神奈川県横浜市).

16. 長 裕敏, 塩川 浩昭, 北川 博之, "高水準言語で記述可能なストリーム処理とバッチ処理の統合フレームワーク", 第 8 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 2016), 2016 年 2 月 29 日, ヒルトン福岡シーホーク(福岡県福岡市).

17. 大西 誠, 山口 祐人, 北川 博之, "オンラインニュースとツイートのリアルタイムマッチング手法", 第 8 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 2016), 2016 年 3 月 1 日, ヒルトン福岡シーホーク(福岡県福岡市).

18. Savong Bou, Toshiyuki Amagasa and Hiroyuki Kitagawa, "Efficient Keyword Search over Relational Data Streams", 第 8 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 2016), 2016 年 3 月 1 日,

ヒルトン福岡シーホーク（福岡県福岡市）。

19. 小澤 佑介, 天笠 俊之, 北川 博之, "GPUを用いたラベル伝搬法によるグラフクラスタリングの高速化", 第8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 2016), 2016年3月2日, ヒルトン福岡シーホーク（福岡県福岡市）。

20. Sei Onishi, Yuto Yamaguchi, and Hiroyuki Kitagawa, "Real-Time Relevance Matching of News and Tweets ", 23th International Conference on Cooperative Information Systems (CoopIS 2015), 2015年10月30日, ロードス(ギリシャ)。

21. Mateus S. H. Cruz, Yusuke Kozawa, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "GPU Acceleration of Set Similarity Joins", 26th International Conference on Database and Expert Systems Applications (DEXA 2015), 2015年9月2日, バレンシア(スペイン)。

22. Yusuke Kozawa, Fumitaka Hayashi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Parallel Canopy Clustering on GPUs", 26th International Conference on Database and Expert Systems Applications (DEXA 2015), 2015年9月1日, バレンシア(スペイン)。

23. Hao Ye, Hiroyuki Kitagawa, and Jun Xiao, "Continuous Angle-based Outlier Detection on High-dimensional Data Streams", 19th International Database Engineering & Applications Symposium (IDEAS 2015), 2015年7月15日, 慶應義塾大学(神奈川県横浜市)。

24. 長 裕敏, 王 岩, 北川 博之, 天笠 俊之, "半構造データに対するストリーム処理とバッチ処理の統合フレームワーク", 情報処理学会第77回全国大会 (IPSI 全国大会 2015), 2015年3月19日, 京都大学(京都市)。

25. 大西 誠, 北川 博之, "オンラインニュースに関連するツイートのリアルタイムな収集", 第7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 2015), 2015年3月4日, 磐梯熱海ホテル華の湯(福島県郡山市)。

26. Silqueira Hickson Cruz Mateus, 小澤 佑介, 天笠 俊之, 北川 博之, "Efficient Set Similarity Join on Graphics Processors", 第7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 2015), 2015年3月2日, 磐梯熱海ホテル華の湯(福島県郡山市)。

27. Savong Bou, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "An Improved Method of Keyword Search over Relational Data Streams", 第7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 2015), 2015年3月2日, 磐梯熱海ホテル華の湯(福島県郡山市)。

28. Takahiro Komamizu, Toshiyuki Amagasa and Hiroyuki Kitagawa, "Extracting Facets from Textual Contents for Faceted Search over XML Data," 16th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2014), 2014年12月4日, ハノイ(ベトナム)。

29. Savong Bou, Toshiyuki Amagasa and Hiroyuki Kitagawa, "Filtering XML Streams by XPath and Keywords," 16th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2014), 2014年12月4日, ハノイ(ベトナム)。

〔その他〕  
 ホームページ等  
<http://www.kde.cs.tsukuba.ac.jp/>

## 6. 研究組織

### (1) 研究代表者

北川 博之 (KITAGAWA HIROYUKI)  
 筑波大学・計算科学研究センター・教授  
 研究者番号: 00204876

### (2) 研究分担者

天笠 俊之 (AMAGASA TOSHIYUKI)  
 筑波大学・計算科学研究センター・教授  
 研究者番号: 70314531

塩川 浩昭 (SHIOKAWA HIROAKI)  
 筑波大学・計算科学研究センター・助教  
 研究者番号: 90775248

早瀬 康裕 (HAYASE YASUHIRO)  
 筑波大学・システム情報系・助教  
 研究者番号: 40423090

渡辺 知恵美 (WATANABE CHIEMI)  
 筑波大学・システム情報系・助教  
 研究者番号: 20362832

### (3) 連携研究者

高橋大介 (DAISUKE TAKAHASHI)  
 筑波大学・計算科学研究センター・教授  
 研究者番号: 00292714

渡邊陽介 (YOUSUKE WATANABE)  
 名古屋大学・未来社会創造機構・准教授  
 研究者番号: 80532944