

部門・分野名： 計算情報学研究部門・計算知能分野

1. メンバ

教授 北川博之

講師 天笠俊之

講師 川島英之

2. 概要

計算科学において、大規模データの管理や活用は極めて重要な課題となっている。計算情報学研究部門計算知能分野は、データ工学関連分野の研究開発を担当している。具体的には、異種データベースや多様な情報源を統合的に扱うための情報統合基盤技術、データ中に埋もれた知識や規則を発見するためのデータマイニング・知識発見技術、インターネット環境において様々なデータを統合的に扱うための XML 関連技術等の基盤技術の研究を行った。また、地球生物環境研究部門と連携して、気象データベースの運用や知識発見等に関する応用的な研究も推進した。

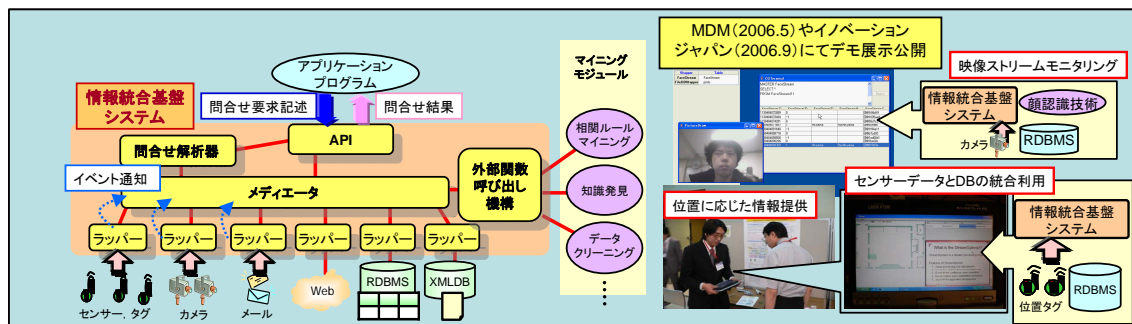
3. 研究成果

【1】情報統合基盤技術

(主な研究費：科研費基盤研究 A, 科研費特定領域研究, JST CREST, 科研費若手研究 B)

(1) 情報統合基盤システム (北川, 天笠, 川島)

異種の分散したデータベースや情報源を統合的に扱うための基盤技術・システム・応用の研究開発を行った。特に、従来型のデータベースや Web 等のみではなく、センサー、位置情報源等の連続的に情報を提供するストリーム情報源をも統合対象とすることができる基盤システム StreamSpinner を研究開発した。StreamSpinner は、リレーショナルモデルをベースとした情報統合処理機能をするが、ビデオや音声等の連続メディアの統合処理にも適用可能である。また、あらかじめ提供される基本演算子に加えて、応用目的向けのプログラムを外部関数として問合せから呼び出す機能を実現している。これによって、時系列データの類似検索や、カメラ映像ストリームに対して解析機能を組み合わせた情報統合等が可能である。



さらに、複数ノード上で StreamSpinner を協調動作させることにより、分散環境におけるストリーム処理を実現することができる。また、ノード障害が発生した場合でも持続的に統合処理を実現するための機能の研究も推進した。

(2) センシングデータベース基盤 (川島, 北川)

実世界監視データ等の時系列データを DBMS で管理する需要が増加しているが、このような要求に対応するための DBMS には、高速データ挿入機能、データ解析機能、連続的問合せ機能等が要求される。これらの機能を実現するセンシングデータベース基盤として KRAFT を開発した。UPS 付きマシンのメモリを永続的記憶装置とみなすことで実現する高速データ挿入機能、時系列データに対する類似検索関数、FFT 関数、集約関数等の解析機能が特徴である。

(3) 確率的ストリーム処理技術 (川島, 北川)

実世界で生起する事象は不確実であり、確率的に表現され得る。そのような確率的データストリームを処理するために、Lineage を用いる手法を提案した。同手法により、複合事象の確率を安全かつ効率的に計算される。処理過程において確率値はブール式により保持される。最終段階に残されたデータについてのみ Lineage を用いた確率計算が実行される。また、確率推論の一技法であるベイジアンネットワークをリレーショナルデータストリーム処理システムと統合する研究も行った。確率的データを生ずるベイジアンネットワークをリレーショナルデータベースの枠組みの中で扱うため、ベイジアンネットワークをオブジェクトとして実現し、それを操作するための演算子を定義している。

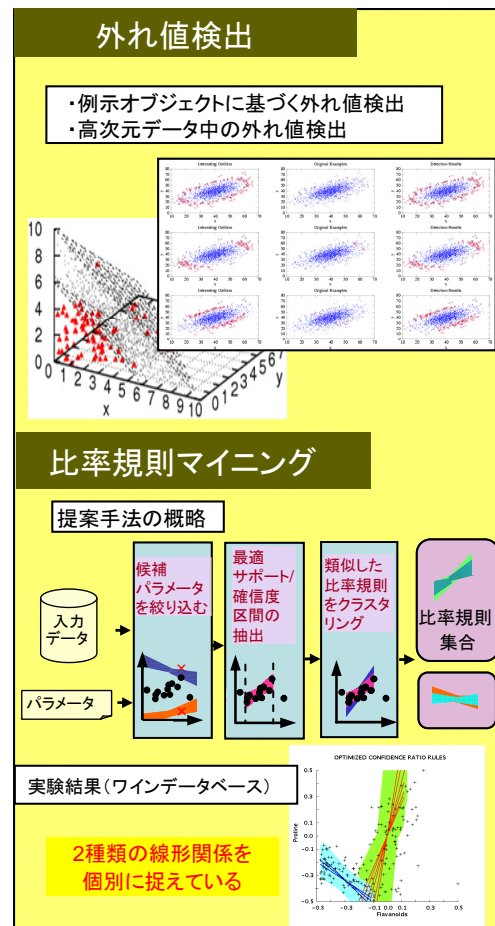
【2】データマイニング・知識発見技術

(主な研究費：科研費特定領域研究，科研費基盤研究 A，科研費若手研究 B，JST CREST)

外れ値検出，比率規則マイニング，文書情報源からの情報抽出，時系列文書クラスタリング，文書のトピック分析，話題構造マイニング，移動体オブジェクトに対する移動統計量抽出等の種々のデータマイニング・知識発見技術の研究開発を進めた。以下では、主なものについて述べる。

(1) 外れ値検出 (北川)

外れ値とは通常データから大きく例外的なデータのことであり、異常検出や興味あるデータの発見等に有用である。代表的な外れ値検出方法として、オブジェクト間の距離や分布密度に基づく検出手法がある。しかし、これらの手法を利用する上では、適切なパラメタの設定が容易でないという問題点や、高次元データで



2種類の線形関係を個別に捉えている

は距離や密度に基づくオブジェクトの特徴付けが困難になるという問題点がある。これらの問題点に対応するため、利用者が与える外れ値データのサンプルを用い、適切な低次元の部分空間を検出して外れ値検出を行う手法を開発した。

また、ストリーム系データに対する差分計算に基づく効率的な連続的外れ値検出アルゴリズムを開発した。この研究成果により、北川教授と石田梢氏が DEWS2008 において優秀論文賞を受賞した。

一方、これまでの外れ値検出は主に数値データを対象としており、離散値を対象とするものはほとんど提案されていないのが現状である。離散値を主体とするレコード型データやトランザクションデータを対象とする外れ値検出手法として、相関ルールに基づく手法を開発した。WAIM2008 国際会議における本研究成果発表で、成田和世氏が Hongjun Lu Best Student Paper Award を受賞した。

(2) 比率規則マイニング (北川)

比率規則は欠損値の補完、予測、外れ値検出など多種の応用が可能であるため、その抽出は重要な技術課題である。既存の比率規則抽出手法として代表的な主成分分析を用いた手法では、複数の線形関係が混在するような状況において、個々の線形関係を区別できない、部分的に成り立つ線形関係を捉えることもできないといった問題がある。

これらの問題を解決した新たな比率規則マイニング手法の研究開発を行った。新たな手法では、相関規則マイニングに準じてサポートと確信度を定義し、それらに基づいた比率規則マイニングを実現している。本研究成果により、濱本雅史氏が DEWS2007 において優秀プレゼンテーション賞を受賞した。

(3) 文書情報源からの情報抽出 (北川)

Web 上には膨大な情報が存在し、かつ拡大を続けている。しかし、情報源の異種性等により、膨大な情報の有効な利用は必ずしもなされていないというのが現状である。Web や各種文書データから有用な情報を得るための手法として、情報抽出に関する研究がこれまで行われてきたが、膨大な情報源から、利用者が必要とする情報をいかに選択的に抽出するかが重要な問題となっている。本研究では、利用者からの例示レコードおよび提示されたサンプルデータベースの情報をもとに、関連の深いレコードデータを中心的に抽出する手法を開発した。抽出レコードデータの評価に基づいて抽出対象の文書情報源を選別することで、情報抽出における精度と抽出数の向上に役立てる点に特徴がある。

(4) 文書のトピック解析 (北川)

大規模な文書群に含まれる主要なトピックとその相互関係をグラフ解析的なアプローチで抽出する手法を開発した。この研究成果により、北川教授と戸田浩之氏が DEWS2007 において優秀論文賞を受賞した。

(5) 時系列文書クラスタリング (北川)

インターネット上では、大量の文書情報が日々流通しているが、あまりにその量が膨大であるため、情報の集約技術が求められる。本研究では、ニュース

記事等のように発行された日時などの情報が付加された時系列文書の集約技術に焦点を当て、その新規性に注目したクラスタリング手法を研究開発した。

(6) 移動軌跡データ管理 (北川)

これまで数年モバイル環境における移動体の移動軌跡データを効率的に格納管理する手法を開発してきたが、本研究に関する成果発表論文で、北川教授と名古屋大学の石川教授 (2006年3月まで筑波大学計算科学研究センター所属) が平成19年度電子情報通信学会論文賞を受賞した。

【3】XML・Webプログラミング

(主な研究費：科研費特定領域研究, JST CREST, 科研費若手研究B)

XML (Extensible Markup Language) は、データ記述のためのメタ言語であり、クリアテキストによって複雑なデータ構造を記述することができる。XMLはネットワーク環境におけるデータ相互利用のための標準的なデータ記述フォーマットとして広く認知され、多くの分野で利用されている。XML形式で生成、蓄積されるデータ量は爆発的に増加しており、今後もその傾向は継続することが予想される。このため、XML形式で記述された大量の情報資源の効率的な蓄積および利活用を目的として、種々の研究を行った。

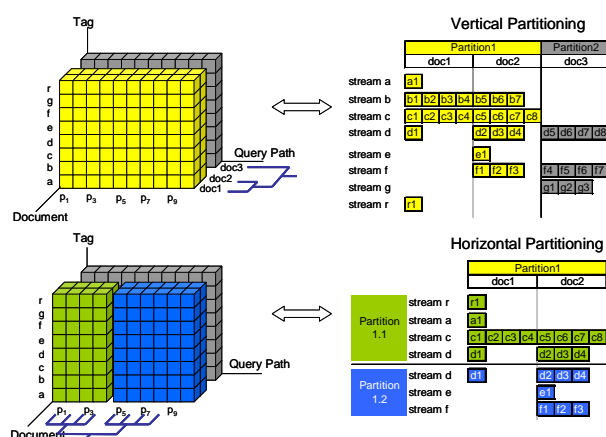
(1) XMLデータに対するOLAP (天笠, 北川)

XMLデータに対して必要な情報を獲得するための処理としては、検索が一般的である。しかしながら、XMLの応用範囲が広がるにつれ、検索処理だけではなく、より複雑な分析処理と知識発見を可能にする対話的分析処理 (OLAP; Online Analytical Processing) のサポートが重要になる。本研究では、XMLデータの分析処理を可能にするXML-OLAP技術の研究開発を行っている。

OLAPでは、データを多数の属性からなる仮想的な多次元キューブにとらえ、キューブに対して演算を適用することによって分析を行う。今年度は、XML-OLAPにおいて重要な役割を持つ TOPOLOGICAL ROLLUP 演算に着目した。これは、XMLの特徴である木構造を利用し、葉ノードから根ノードに向かって集約計算を繰り返し行う演算であり、XMLの木構造におけるさまざまなレベルでの集約値を使った解析を可能とする。本研究では、この TOPOLOGICAL ROLLUP 演算を高速に実行するためのいくつかのアルゴリズムを考案し、その特質を実験により評価した。この研究により、Chantola Kit氏が「日本データベース学会・情報処理学会データベースシステム研究会・電子情報通信学会データ工学研究会優秀若手研究者賞」を受賞した。

(2) 大規模XML検索の並列処理 (天笠, 北川)

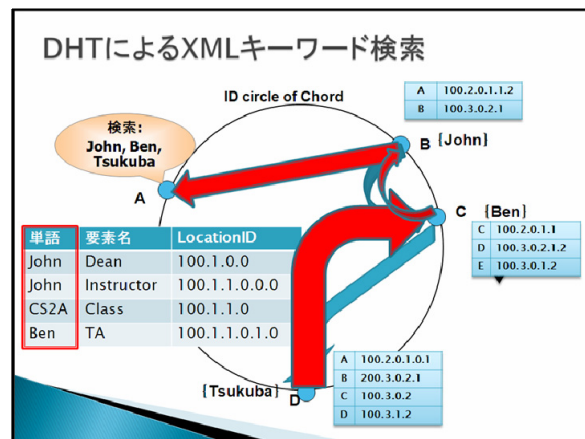
XMLデータの大規模化に伴い、数百ギガバイト、あるいは数テラバイトのXMLデータを効率的に扱うための手法も今後必要となる。しかしながら、XMLは本質的に木構造であるため、その処理には多大なコストを要することが問題となる。これを



背景に、PC クラスタを用いた大規模 XML データの並列処理方式について研究を進めている。本年度は、XML 問合せ処理における重要なアルゴリズムである“Holistic Twig Join”に着目し、これをPC クラスタ上で分散・並列実行する手法をそれぞれ検討した。Holistic Twig Join の分散処理のため、XML データおよび問合せワークロードから得られる情報に基づき XML を分割する。これをクラスタノードに配置し、問合せを実行する。特定のノードに処理が集中するなど、処理負荷が不均衡になった場合には、過負荷のノードが実行している処理を、他ノードに分散することで、システム全体の性能を保つことができる。この研究発表により、Machdi Imam 氏は iDB2008 Excelent Paper Award を受賞した。

(3) P2P ネットワークにおける XML 探索 (天笠, 北川)

インターネットを基盤としたデータの相互運用がグローバル化する一方で、インターネット上に特定の組織やグループに閉じた専用のネットワークを構築したいという要求は、情報保全や情報流通の効率化の観点から急速に高まっている。ピアツーピアネットワーク (P2P) は、特定の応用向きのネットワークを柔軟に構築できる「オーバーレイネットワーク」の基盤技術として近年注目されており、P2P ネットワークを利用した科学データ等の流通に関する研究開発も行われている。その際、標準



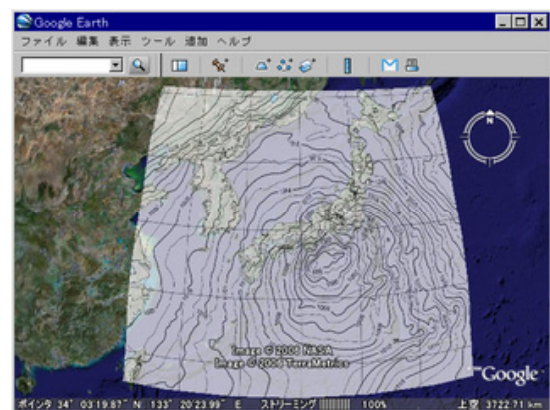
データフォーマットとして XML の利用が想定されることより、P2P ネットワークを利用した XML データの格納および検索手法に関する研究開発を行っている。本年度は、昨年度までの研究成果を継承し、分散ハッシュ表 (DHT) における XML キーワード検索を可能にするための手法を開発した。

【4】気象データベース

(主な研究費：科研費萌芽研究，科研費若手研究 B)

(1) GPV/JMA アーカイブ (天笠, 北川)

地球生物環境研究部門と共同で、気象庁気象予報データベース「GPV/JMA アーカイブ」(<http://gpv.jma.ccs.hpcc.jp>)の開発、および管理、運用を行っている。GPV/JMA アーカイブは、気象庁が公開している気象予報グリッドデータ (GPV データ) を蓄積するとともに、外部登録ユーザへのデータを提供することを目的としている。GPV/JMA アーカイブで提供しているデータは、全球モデル、メソスケールモデル、リージョナルスケールモデル、週間ア



ンサンプル, 月間アンサンプル, 季間アンサンプルの 6 種類である. さらに, これらのグリッドデータに加えて, 数値データを元に作図した天気図を公開するとともに, 天気図の閲覧性の向上するため, GoogleEarth 上に天気図をマップするための KML ファイルの公開サービスも行っている (エラー! 参照元が見つかりません).

(2) 遠隔科学データベースアクセスツール MDFS (天笠, 北川)

今日, 膨大な科学データがインターネット上に公開されるようになった. 各科学分野において, それらの情報資源を活用することは, 研究を遂行する上で欠かすことができない. しかしながら, それらの多くは Web ブラウザによる対話操作による検索およびデータのダウンロードを必要としているため, 研究者の負担となっている. 本研究では, インターネット上に公開されているデータ (Web データベース) へのアクセスを容易にするため, MDFS なるシステムを開発した. MDFS は, Linux のカーネルモジュールである FUSE を利用し, Web データベースを, それがあたかもローカルなストレージに格納されているかのように扱うことを可能にする. 科学者は, 通常の研究業務に利用しているプログラムを一切変更することなく, ネットワーク越しにデータを直接操作することができる.

(3) 気圧配置図の自動分類 (川島, 北川)

気圧配置は気象学において 15 種類に分類されている. 気象学研究者は, 西高東低冬型や南高北低夏型などの, ある特徴をもつ気圧配置の事例を多数必要とすることがある. 過去の気圧配置データは膨大に蓄積されているが, 各データが研究者にとって必要な気圧配置であるかを判別するには, 目視以外の手法は存在しないのが現状である. 我々は地球生物環境研究部門と共同で, この気圧配置の一つである西高東低冬型を Support Vector Machine (SVM) を用いて分類する手法を開発した.

4. 研究業績

(1) 学術雑誌論文

1. 澤菜津美, 森嶋厚行, 飯田敏成, 杉本重雄, 北川博之, 「Web ページ移動先発見のための効率的なクロウリング手法」, 情報処理学会論文誌: データベース. Vol.48, No. SIG 11(TOD34), pp.149-161, 2007 年 6 月.
2. 辻良繁, 川島英之, 「CC-Optimizer: キャッシュを考慮した問合せ最適化器」, 日本データベース学会 Letters, Vol. 6, No. 1, pp. 45-48, 2007 年 6 月.
3. 濱本雅史, 北川博之, 「対称比率規則の抽出手法」, 日本データベース学会 Letters, Vol. 6, No. 1, pp. 73-76, 2007 年 6 月.
4. 呉俊輝, 天笠俊之, 北川博之, 「構造型 P2P ネットワークにおける負荷分散を考慮した XML データ処理」, 日本データベース学会 Letters, Vol. 6, No. 1, pp. 93-96, 2007 年 6 月.
5. 広瀬健志郎, 川島英之, 佐竹聡, 今井倫太, 「異種ロボット間でのジェスチ

- ャ情報の共有化」, 日本データベース学会 Letters, Vol. 6, No. 1, pp. 125-128, 2007年6月.
6. 張建偉, 石川佳治, 北川博之, 「トピックを考慮した大規模文書情報源からのレコード抽出」, 情報処理学会論文誌: データベース, Vol. 48, No. SIG14 (TOD35), pp. 107-123, 2007年9月.
 7. 山田真一, 渡辺陽介, 北川博之, 天笠俊之, 「データストリーム管理システム Harmonica の設計と実装」, 情報処理学会論文誌: データベース, Vol. 48, No. SIG14 (TOD35), pp. 91-106, 2007年9月.
 8. 渡辺陽介, 北川博之, 「分散ストリーム処理環境における持続型問合せ処理方式」, 日本データベース学会 Letters, Vol. 6, No. 2, pp. 41-44, 2007年9月.
 9. 佐竹聡, 川島英之, 今井倫太, 安西祐一郎, 「Brownie: カメラ上に指定された過去のランドマーク情報に基づく実世界探し物検索システム」, 知能と情報(日本知能情報ファジィ学会誌), Vol. 19, No. 5, pp. 556-569, 2007年10月.
 10. Sophoin Khy, Yoshiharu Ishikawa, and Hiroyuki Kitagawa, "A Novelty-based Clustering Method for On-line Documents", World Wide Web Journal, Vol. 11, No. 1, pp. 1-37, March 2008.
 11. 渡辺陽介, 秋山亮, 大喜恒甫, 北川博之, 「映像ストリームのための映像情報統合基盤システムの提案」, 日本データベース学会 Letters, Vol. 6, No. 4, pp. 13-16, 2008年3月.
- (2) 国際会議発表論文
1. Chantola Kit, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "OLAP Query Processing for XML Data in RDBMS", Proc. 3rd IEEE International Workshop on Databases for Next-Generation Researchers, Istanbul, Turkey, April 2007.
 2. Yuan Li and Hiroyuki Kitagawa, "DB-Outlier Detection by Example in High Dimensional Datasets", Proc. 3rd IEEE International Workshop on Databases for Next-Generation Researchers, Istanbul, Turkey, April 2007.
 3. Takeshi Kanda, Yutaka Yanagisawa, Takuya Maekawa, Michita Imai, Hideyuki Kawashima and Takeshi Okadome, "A Distributed Inference System on Sensor Nodes using Neighbor's Context Data", Proc. 3rd IEEE International Workshop on Databases for Next-Generation Researchers, Istanbul, Turkey, April 2007.
 4. Manabu Nakamura, Hideyuki Kawashima, Satoru Satake, and Michita Imai, "Mana: A Real World Oriented Query Processing System Supporting Control of Sensor Characteristics", Proc. International Workshop on SensorWebs, Databases and Mining in Networked Sensing Systems (SWDMNSS), Braunschweig, Germany, June 2007.

pp. 23--30.

5. Satoru Satake, Hideyuki Kawashima, and Michita Imai, "Brownie: Searching Concealed Real World Artifacts", Proc. 4th International Conference on Networked Sensing Systems (INSS'07), Braunschweig, Germany, June 2007. pp. 159--162.
6. Hideyuki Kawashima, "KRAFT: A Real-Time Active DBMS for Signal Streams", Proc. 4th International Conference on Networked Sensing Systems (INSS'07), Braunschweig, Germany, June 2007. pp. 163--166.
7. Jianwei Zhang, Yoshiharu Ishikawa, and Hiroyuki Kitagawa, "Record Extraction Based on User Feedback and Document Selection", Proc. Joint Conference of 9th Asia-Pacific Web Conference and 8th International Conference on Web-Age Information Management (APWeb/WAIM 2007), LNCS 4505, pp. 574-585, HuangShan (Yellow Mountains), China, June 2007.
8. Toshiyuki Amagasa, Chunhui Wu, and Hiroyuki Kitagawa, "Retrieving Arbitrary XML Fragments from Structured Peer-to-Peer Networks", Proc. Joint Conference of 9th Asia-Pacific Web Conference and 8th International Conference on Web-Age Information Management (APWeb/WAIM 2007), LNCS 4505, pp. 317-328, HuangShan (Yellow Mountains), China, June 2007.
9. Masafumi Hamamoto and Hiroyuki Kitagawa, "Locality-aware Ratio Rule Mining", Proc. 4th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), Vol. 2, pp. 693-698, Haikou, China, August 2007.
10. Yousuke Watanabe, Shinichi Yamada, Hiroyuki Kitagawa, and Toshiyuki Amagasa, "Integrating a Stream Processing Engine and Databases for Persistent Streaming Data Management", Proc. 18th Int'l Conference on Database and Expert Systems Applications (DEXA2007), LNCS 4653, pp. 414-423, Regensburg, Germany, Sept. 2007.
11. Toshiyuki Amagasa, Lianzi Wen, and Hiroyuki Kitagawa, "Proximity Search of XML Data using Ontology and XPath Edit Similarity", Proc. 18th Int'l Conference on Database and Expert Systems Applications (DEXA2007), LNCS 4653, pp. 298-307, Regensburg, Germany, Sept. 2007.
12. Toshiyuki Amagasa, Kentarou Kido, and Hiroyuki Kitagawa, "Querying XML Data using PC Cluster System", Proc. 2nd International Workshop on XML Data Management Tools and Techniques (XANTEC'07) in conjunction with DEXA 2007, pp. 5-9, Regensburg, Germany, Sept. 2007.

13. Hiroyuki Kitagawa and Yousuke Watanabe, "Stream Data Management Based on Integration of a Stream Processing Engine and Databases", Proc. IFIP International Conference on Network and Parallel Computing Workshops, pp. 18-22, Dalian, China, Sept. 2007. (Invited paper).
14. Toshiyuki Amagasa, Hiroyuki Kitagawa, and Tatsuya Komano, "Constructing a Web Service System for Large-scale Meteorological Grid Data", 3rd IEEE Int'l Conf. on e-Science and Grid Computing (e-Science 2007), pp. 118-124, Bangalore, India, December 10-13, 2007.
15. Natsumi Sawa, Atsuyuki Morishima, Shigeo Sugimoto, and Hiroyuki Kitagawa, "Wraplet: Wrapping Your Web Contents with a Lightweight Language", Proc. of the 3rd IEEE International Conference on Signal-Image Technology and Internet-based Systems (SITIS'2007), Shanghai, China, December 16-19, 2007, 8 pages.
16. Hiroyuki Toda, Hiroyuki Kitagawa, Ko Fujimura, and Ryoji Kataoka, "Topic Structure Mining using Temporal Co-occurrence", Proc. 2nd International Conference on Ubiquitous Information Management and Communication (ICUIMC 2008), pp. 253-258, Suwon, Korea, January. 2008.
17. Yousuke Watanabe, Ryo Akiyama, Kousuke Ohki and Hiroyuki Kitagawa, "A Video Stream Management System for Heterogeneous Information Integration Environments", Proc. 2nd International Conference on Ubiquitous Information Management and Communication (ICUIMC 2008), pp. 219-224, Suwon, Korea, January, 2008.
18. Yuan Li and Hiroyuki Kitagawa, "Example-Based Robust DB-Outlier Detection for High Dimensional Data", Proc. 13th International Conference on Database Systems for Advanced Applications (DASFAA 2008), LNCS 4947, pp. 330-347, New Delhi, India, March 19-21 2008.
19. Lianzi Wen, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "An Approach for XML Similarity Join using Tree Serialization", Proc. 13th International Conference on Database Systems for Advanced Applications (DASFAA 2008), LNCS 4947, pp. 562-570, New Delhi, India, March 19-21 2008.

(3) 学会発表

(A) 招待講演

1. 川島英之, 「ユビキタス・センサネットワークのデータベース技術」, 電子情報通信学会ユビキタスセンサネットワーク研究会(USN), 2007年5月24日.

2. 川島英之, 「情報爆発時代のデータベースーセンサネットワーク技術がもたらすデータベース技術の新展開と応用ー」, FIT. 2007年9月5日.
3. Hiroyuki Kitagawa, "Stream Data Management Based on Integration of a Stream Processing Engine and Databases", IEEE International Workshop on Frontier of Data-Driven Engineering (FDDE 2007), Dalian, China, Sept. 16, 2007.
4. 川島英之, 「ユビキタスセンサネットワーク時代の新しいデータベース技術」, 電子情報通信学会北海道支部講習会「生活に融け込むコンピューティング」. 2007年11月16日.

(B) その他学会発表

1. 文連子, 天笠俊之, 北川博之, 「木直列化を用いた XML データの類似結合」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 91-96.
2. 濱本雅史, 北川博之, (学生発表奨励賞受賞), 「比率規則マイニングシステムの開発と評価」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 177-182
3. 渡辺陽介, 北川博之, 「仮想マシン技術を用いた持続型ストリーム処理環境の評価」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 339-344.
4. 稲守孝之, 渡辺陽介, 北川博之, 天笠俊之, 川島英之, 「分散ストリーム処理環境におけるアプリケーション配置最適化手法」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 345-350.
5. 川島英之, 北川博之, 「確率推論センサデータベースシステムの設計」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 351-356.
6. 辻良繁, 川島英之, (学生発表奨励賞受賞) 「CC-Optimizer のパフォーマンスカウンタによる微視的評価」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 357-362.
7. 成田和世, 北川博之, (学生発表奨励賞受賞) 「トランザクションデータベースに対する高確信度の相関ルールを用いた外れ値検出手法」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 399-404.
8. Sophoin Khy, Yoshiharu Ishikawa, Hiroyuki Kitagawa, 「A Document Clustering Method Focusing on User Specified Time of Interest」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007年7月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 411-416.
9. Yuan Li, Hiroyuki Kitagawa, (学生発表奨励賞受賞), 「Detecting

10. 澤菜津美, 森嶋厚行, 杉本重雄, 北川博之, (学生発表奨励賞受賞), 「非定型 Web コンテンツ管理のための軽量ラッピング言語」, 夏のデータベースワークショップ 2007 (DBWS 2007), 2007 年 7 月. 電子情報通信学会技術研究報告 Vol. 107, No. 131, pp. 527-532.
11. 濱本雅史, 北川博之, 「多次元データに対する比率規則抽出」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
12. 成田和世, 北川博之, 「カテゴリ型レコードデータに対する関数従属性を用いた外れ値検出」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
13. Sophoin Khy, Yoshiharu Ishikawa and Hiroyuki Kitagawa , "Summarizing Document Cluster Transitions Using Descriptive Sentences", 電子情報通信学会第 19 回データ工学ワークショップ (DEWS2008), 2008 年 3 月 9 日～11 日.
14. Chantola Kit, Toshiyuki Amagasa and Hiroyuki Kitagawa , "An Efficient Structure-based Grouping for XML-OLAP", 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
15. Imam Machdi, Toshiyuki Amagasa and Hiroyuki Kitagawa , "A Multidimensional Data Structure for Maintaining XML Data Partitions", 電子情報通信学会第 19 回データ工学ワークショップ (DEWS2008), 2008 年 3 月 9 日～11 日.
16. Lianzi Wen, Toshiyuki Amagasa and Hiroyuki Kitagawa , "A Text-based Similarity Join for XML data", 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
17. Yuan Li and Hiroyuki Kitagawa, "A Robust Method for Detecting DB-Outliers from High Dimensional Datasets", 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
18. 稲守孝之, 渡辺陽介, 北川博之, 天笠俊之, 川島英之, 「広域分散ストリーム処理環境における演算の配置最適化手法の評価」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
19. Zhitao Shen, Hideyuki Kawashima and Hiroyuki Kitagawa , "Probabilistic Event Stream Processing with Lineage", 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日～11 日.
20. Jiajun Gu and Hiroyuki Kitagawa, "Extending Keyword Search to

- Metadata in Relational Database", 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
21. 李曉晨, 天笠俊之, 北川博之, 「構造型 P2P ネットワークにおける転置リストを利用した XML 文書検索」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 22. 柿木健, 北川博之, 「SaaS 環境の企業間データ仮想統合を志向した問合せ言語に関する研究」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 23. 大喜恒甫, 渡辺陽介, 秋山亮, 北川博之, 天笠俊之, 川島英之, 「ストリーム処理における情報源の動的選択機能の評価」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 24. 石田梢, 北川博之, (優秀論文賞受賞), 「時系列データに対する効果的な外れ値検出手法」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 25. 佐藤亮, 川島英之, 北川博之, 「ベイジアンネットワークを用いた確率的データストリームモデルの提案」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 26. 高橋翼, 北川博之, 「ソーシャルブックマークを利用したユーザ嗜好に基づくページの評価」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 27. 宇井敬一郎, 天笠俊之, 北川博之, 「FUSE による遠隔気象データアクセスツール」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 28. 木村広希, 川島英之, 北川博之, 「気象データにおける特徴的気圧配置の自動分類」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 29. 澤菜津美, 森嶋厚行, 杉本重雄, 北川博之, 「情報統合利用を目的とした HTML ページのラッピング支援」, 電子情報通信学会第 19 回データ工学ワークショップ(DEWS2008), 2008 年 3 月 9 日~11 日.
 30. MoonBae Song, Hiroyuki Kitagawa, "Managing Frequent Updates in R-trees by Semi-Bulkloading", 情報処理学会 第 70 回全国大会講演論文集(1), pp.439-440. 筑波大学, 2008 年 3 月 13 日~15 日.
 31. 濱本雅史, 北川博之, 「多次元比率規則の抽出方法」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.391-392. 筑波大学, 2008 年 3 月 13 日~15 日.
 32. 成田和世, 北川博之, 「アイテム間の相関性を利用したトランザクションデータに対する外れ値検出手法」, 情報処理学会 第 70 回全国大会講演論文集(5), pp.41-42. 筑波大学, 2008 年 3 月 13 日~15 日.
 33. Sophoin Khy, Yoshiharu Ishikawa, Hiroyuki Kitagawa, "Generating High Level Descriptions from Document Cluster Transitions", 情報処

- 理学会 第 70 回全国大会講演論文集(1), pp.199-200. 筑波大学, 2008 年 3 月 13 日～15 日.
34. Chantola Kit, Toshiyuki Amagasa, Hiroyuki Kitagawa, "An Algorithm for XML Cube Computation in XML-OLAP", 情報処理学会 第 70 回全国大会講演論文集(5), pp.25-26. 筑波大学, 2008 年 3 月 13 日～15 日.
 35. Imam Machidi, Toshiyuki Amagasa, Hiroyuki Kitagawa, "XML Data Partitioning for Parallel Holistic Twig Join Processing", 情報処理学会 第 70 回全国大会講演論文集(1), pp.393-394. 筑波大学, 2008 年 3 月 13 日～15 日.
 36. 稲守孝之, 渡辺陽介, 北川博之, 天笠俊之, 川島英之, 「分散ストリーム処理管理システム ORINOCO の評価」, 情報処理学会 第 70 回全国大会講演論文集(5), pp.235-236. 筑波大学, 2008 年 3 月 13 日～15 日.
 37. Lianzi Wen, Toshiyuki Amagasa, Hiroyuki Kitagawa, (学生奨励賞受賞), "An Improved Similarity Join for XML Data based on Text Similarity", 情報処理学会 第 70 回全国大会講演論文集(1), pp.599-600. 筑波大学, 2008 年 3 月 13 日～15 日.
 38. Yuan Li, Hiroyuki Kitagawa, (学生奨励賞受賞), "A Robust Method of Detecting DB-Outliers in High Dimensional Datasets", 情報処理学会 第 70 回全国大会講演論文集(1), pp.683-684. 筑波大学, 2008 年 3 月 13 日～15 日.
 39. 大喜恒甫, 渡辺陽介, 秋山亮, 北川博之, 天笠俊之, 川島英之, 「ストリーム処理における情報源の動的選択機能」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.697-698. 筑波大学, 2008 年 3 月 13 日～15 日.
 40. 王毅, 天笠俊之, 北川博之, 「関係表の垂直表現を利用した河川情報データの統合」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.583-584. 筑波大学, 2008 年 3 月 13 日～15 日.
 41. Jiajun Gu, Hiroyuki Kitagawa, (学生奨励賞受賞), "Keyword Search Including Metadata in Relational Databases", 情報処理学会 第 70 回全国大会講演論文集(1), pp.517-518. 筑波大学, 2008 年 3 月 13 日～15 日.
 42. 李曉晨, 天笠俊之, 北川博之 (学生奨励賞受賞), 「構造型 P2P ネットワークにおけるキーワードを用いた XML 文書検索」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.573-574. 筑波大学, 2008 年 3 月 13 日～15 日.
 43. Zhitao Shen, Hideyuki Kawashima, Hiroyuki Kitagawa, "Continuous Query over Uncertain Data Streams", 情報処理学会 第 70 回全国大会講演論文集(1), pp.691-692. 筑波大学, 2008 年 3 月 13 日～15 日.
 44. 木村広希, 川島英之, 北川博之, 「データマイニングによる気圧配置の分類」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.577-578. 筑波大学, 2008 年 3 月 13 日～15 日.

45. 佐藤亮, 川島英之, 北川博之, 「ベイジアンネットワークに対する効率的な更新・問合せ手法」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.443-444. 筑波大学, 2008 年 3 月 13 日~15 日.
46. 石田梢, 北川博之, 「時系列データに対する効果的な外れ値検出」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.681-682. 筑波大学, 2008 年 3 月 13 日~15 日.
47. 高橋翼, 北川博之 (学生奨励賞受賞), 「ソーシャルブックマークを利用したユーザ嗜好に基づくページの抽出」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.651-652. 筑波大学, 2008 年 3 月 13 日~15 日.
48. 宇井敬一郎, 天笠俊之, 北川博之 (学生奨励賞受賞), 「FUSE を利用した異種気象データの統合利用システム」, 情報処理学会 第 70 回全国大会講演論文集(1), pp.575-576. 筑波大学, 2008 年 3 月 13 日~15 日.
49. 高橋公海, 澤菜津美, 森嶋厚行, 杉本重雄, 北川博之, 「Web コンテンツ一貫性管理支援ツールの開発」, 情報処理学会 第 70 回全国大会講演論文集(5), pp.189-190. 筑波大学, 2008 年 3 月 13 日~15 日.

(4) 受賞

1. 電子情報通信学会論文賞 : 石川佳治, 町田陽二, 北川博之, 「マルコフ連鎖モデルに基づく移動ヒストグラムの動的構築法」 電子情報通信学会論文誌 D, Vol. J90-D, No.2, pp.311-324, 2007 年 2 月.
2. WAIM2008 (The 9th International Conference on Web-Age Information Management) 最優秀学生論文賞 (Honjun Lu Best Student Paper Award) : Kazuyo Narita and Hiroyuki Kitagawa, "Outlier Detection for Transaction Databases using Association Rules", Proc. 9th International Conference on Web-Age Information Management (WAIM 2008), pp. 373-380, Zhangjiajie, China, July 20-22 2008.
3. DEWS2007 優秀論文賞 : 戸田浩之, 北川博之, 藤村考, 片岡良治 (優秀論文賞受賞), 「時間的近さを考慮した話題構造マイニング」 電子情報通信学会第 18 回データ工学ワークショップ (DEWS2007), 2007 年 2 月.
4. DEWS2007 優秀プレゼンテーション賞 : 濱本雅史, 北川博之 「対称比率規則の抽出手法」, 電子情報通信学会第 18 回データ工学ワークショップ (DEWS2007), 2007 年 2 月.
5. DEWS2008 優秀論文賞 : 石田梢, 北川博之, 「時系列データに対する効果的な外れ値検出手法」, 電子情報通信学会第 19 回データ工学ワークショップ (DEWS2008), 2008 年 3 月 9 日~11 日.
6. DBWS2007 学生発表奨励賞 :
 - A) 辻良繁, 川島英之, 「CC-Optimizer のパフォーマンスカウンタによる微視的評価」
 - B) 濱本雅史, 北川博之, 「比率規則マイニングシステムの開発と評価」
 - C) 成田和世, 北川博之, 「トランザクションデータベースに対する高確信度の相関ルールを用いた外れ値検出手法」

- D) Yuan Li, Hiroyuki Kitagawa, 「Detecting Outliers in High Dimensional Datasets with Examples」
- E) 澤菜津美, 森嶋厚行, 杉本重雄, 北川博之, 「非定型 Web コンテンツ管理のための軽量ラッピング言語」
- 7. 情報処理学会第 70 回全国大会学生奨励賞 :
 - A) Lianzi Wen, 天笠俊之, 北川博之 「An Improved Similarity Join for XML Databased on Text Similarity」
 - B) Yuan Li, 北川博之, 「A Robust Method of Detecting DB-Outliers in High Dimensional Datasets」
 - C) 顧 佳駿, 北川博之, 「Keyword Search Including Metadata in Relational Databases」
 - D) 李 曉晨, 天笠俊之, 北川博之 「構造型 P2P ネットワークにおけるキーワードを用いた XML 文書検索」
 - E) 宇井 敬一郎, 天笠俊之, 北川博之 「FUSE を利用した異種気象データの統合利用システム」
 - F) 高橋 翼, 北川博之, 「ソーシャルブックマークを利用したユーザ嗜好に基づくページの抽出」
- 8. 日本データベース学会・情報処理学会データベースシステム研究会・電子情報通信学会データ工学研究会優秀若手研究者賞 : Chantola Kit, 「Algorithms for Efficient Structure-based Grouping in XML-OLAP」
- 9. iDB フォーラム 2008 優秀論文賞 :
 - A) Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, 「An Algorithm for Parallel Holistic Twig Joins on a PC Cluster」
 - B) Djelloul Boukhelef and Hiroyuki Kitagawa, 「RCAN: A Multi-ring Content Addressable Network」
- 10. iDB フォーラム 2008 学生奨励賞 :
 - A) 李曉晨, 天笠俊之, 北川博之, 「構造型 P2P ネットワークにおけるキーワードを含む XPath による XML 文書検索」
 - B) 高橋公海, 森嶋厚行, 杉本重雄, 北川博之, 「Web コンテンツ一貫性管理のための制約発見支援」

(5) その他成果デモ展示

- 1. イノベーションジャパン 2007 (主催 : JST, NEDO) : 「大規模センサーデータのための連合型情報基盤システム」, 2007 年 9 月 12~14 日, 東京国際フォーラム.
- 2. イノベーションジャパン 2008 (主催 : JST, NEDO) : 「大規模センサーデータ処理のためのデータストリーム管理基盤」, 2008 年 9 月 16~18 日, 東京国際フォーラム.