V-1 計算知能分野

1. メンバ

教授 北川博之

講師 天笠俊之,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一的に扱うための 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や素粒子宇宙研究部門と連携して、計算科学の各分野における応用的な研究を推進した。

3. 研究成果

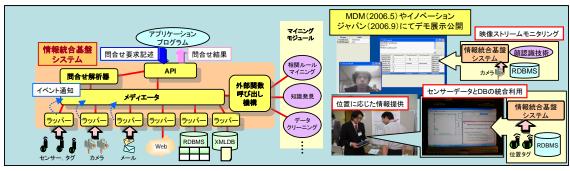
【1】情報統合基盤技術

(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有するが, 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

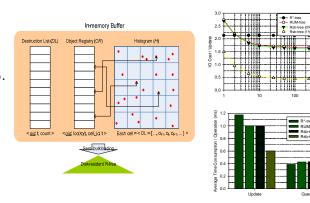
対応するための DBMS には、高速データ挿入機能、データ解析機能、連続的問合せ機能等が要求される。これらの機能を実現するセンシングデータベース基盤として KRAFT を開発した。 UPS 付きマシンのメモリを永続的記憶装置とみなすことで実現する高速データ挿入機能、時系列データに対する類似検索関数、FFT 関数、集約関数等の解析機能が特徴である。

(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。

(4) 多次元ストリーム用高性能索引機構(北川)

ストリームデータの中には、移動体の位置情報等、多次元データが存在する。多次元データでは、空間的検索を支援する必要がある。例えば、距離検索や近傍検索等である。多次元データに対する索引機構としては、R木等の空間索引が従来用いられてきたが、従来の空間索引は静的なデータを対象としており、オブジェクトの現在値が頻繁に変化する動的な環境下では、極めて性能が劣化



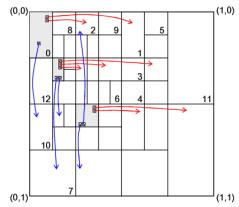
することが知られている。本研究では、R 木を元に、主記憶と二次記憶を有機的に利用することで、この問題に対応するための新たな索引機構 Rsb 木(R-tree with Semi-Bulkloading)を開発した。

(5) P2P ネットワーク基盤技術(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「オーバレイネットワーク」の基盤技術とし

て近年注目されている.

我々は分散ハッシュ表(DHT; Distributed Hash Table)と呼ばれる P2P に着目し、その効率化に取り組んでいる. 具体的には、CAN (Content Addressable Network)と呼ばれるプロトコルを取り上げ、その負荷分散やアクセス効率化を図った. 従来の CAN では、特定のエントリポイントに対してピアの加入・離脱が頻発すると、ピア間の負荷分散に偏りが発生し、それに伴い検索効率が低下することが知られていた.



この問題に対処するために、分散かつ自律的な負荷分散機構を取り入れるとともに、検索処理の効率 化のための長距離リンクを加えたRCAN を提案した。これにより、上記の問題が解消されることを実験に よるシミュレーションで示した。

【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)

多様な情報源に対する外れ値検出, 時系列文書クラスタリング, 文書のトピック分析, 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.

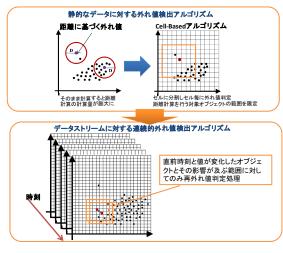
(1) 外れ値検出(北川)

外れ値とは通常の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。実世界には多様な情報源が存在するため、それぞれの情報源に対して固有の特徴を考慮した外れ値を定義し検出を行うことが重要である。

多くの外れ値検出技術は、数値属性値を主体としたデータを対象としており、トランザクションデータからの外れ値検出については既存の研究は極めて少ない、POS データで知られるトランザクションデータから、他とは逸脱した動きを見せるトランザクションを検出出来れば、データ解析上、大きな助けとなることが期待できる。そこで我々は、トランザクションデータにおけるアイテム間の相関性に着目し、本来ならばトランザクション中のアイテム集合と共起するはずのアイテムが、数多く出現していないトランザクションを、外れ値として検出する手法を開発した。外れ値度は、高い確信度を持つ相関ルールを利用して計算される設計となっている。WAIM2008 国際会議における本研究成果発表で、成田和世氏が Hongiun Lu Best Student Paper Award を受賞した。

一方, データストリームとして提供されるデータ量の増加により, データストリームに対するデータマイニングが重要となっている. データストリームに対する処理では, 時々刻々と到着するデータに対して連続的にモニタリングすることが求められる. そこで本研究ではデータストリームに対する連続的な外れ値検出手法を開発した.

本研究では、各時刻におけるデータ分布が直前の時刻のものと類似していることが多いというデータストリームの特徴に着目して、差分処理を行い効率的な外れ値検出を行う、提案アルゴリズムでは、連続的に外れ値検出を行う際に、処理を最小限に抑えるため、直前時刻と比較して変化が生じたオブジェクトとそのオブジェクトが外れ値の判定に影響を与える範囲に限定して処理を行う。これにより無駄な処理を省き、効率化を図った。実データ、人工データを用いた実験により、既存のCell-Basedアルゴリズムをスナップショット毎に適



用する手法より効率的であることを確認した.この研究成果により、北川教授および石田梢氏が DEWS2008 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)

文書クラスタリングは、文書集合からのトピック抽出や文書分類など、さまざまな応用が考えられている。しかしながら、文書の時間的な変化を考慮したクラスタリングおよびその結果の提示方法については、今後の研究の進展が期待されている。本研究では、大量のクラスタのストリームにおけるトピックの進展について問合せする枠組みを構築した。特に、時系列文書のクラスタリング結果に対する変遷パターンを問い合わせることができる問合せ言語を提案するとともに、その処理方式について検討した。

【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.

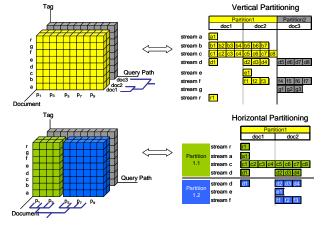
(1) XML データに対する OLAP(天笠, 北川)

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.我々は、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ととらえ、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. 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.また、Chantola Kit 氏、天笠講師、北川教授が、iiWAS2008 Best Student Paper Award 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、天笠講師、北川教授は iDB2008 Excelent Paper Award および iiWAS2008 Best Student 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)

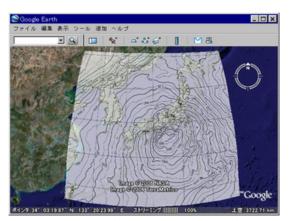
XML の普及に伴い、P2P ネットワークによって構築されたオーバレイネットワークにおいても、XML 形式で記述されたデータあるいはメタデータが今後増えることが予想される。このため、我々は P2P ネットワーク、特に分散ハッシュ表(DHT)における XML データの格納と検索手法について研究を進めている。本年度は、昨年度までの研究成果を受け、分散ハッシュ表(DHT)において、XPath とキーワードの組み合わせによる検索を可能にするための手法を開発した。

【4】科学分野におけるデータベース応用

(主な研究費: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://gpvjma.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、週間アンサンブル、月間



アンサンブル,季間アンサンブルの 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) 格子 QCD アンサンブル XML のファセット検索(天笠, 北川)

ILDG (International Lattice Data Grid)では、格子 QCD 配意データのメタデータとして XML が用いられている。世界中の地域グリッドで公開されている配意データを検索するため、利用性の高いインタフェースの開発が望まれていた。このため我々は、QCDmlのためのファセット検索インタフェースを設計、実装を行った。ファセット検索とは、検索対象オブジェクトの集合を効率よく探索するための手法である。オブジェクトは、あらかじめファセットと呼ばれるいくつかの独立した



カテゴリ毎に分類されている。各カテゴリ(ファセット)において、オブジェクトは着目する属性の値毎にグルーピングされており、その値がリスト表示されている。利用者はファセットに含まれる具体的な値を選択することで、オブジェクトの絞り込みを行い、探索を行う。 XML データに対してファセット検索を適用するため、 XML は半構造性を考慮したファセットの抽出および QCDml におけるファセットの検討を行い、実際にシステムを構築した。

4. 研究業績

<学術雑誌論文>

- 1. Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Cube-Based Analysis for Maintaining XML Data Partition for Holistic Twig Joins", 日本データベース学会論文誌, Vol. 7, No. 1, pp. 121-126, June 2008.
- 2. 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処理システムの提案", 日本データベース学会論文誌, Vol. 7, No. 1, pp. 157-162, June 2008.
- 3. 澤菜津美, 森嶋厚行, 杉本重雄, 北川博之, "HTMLラッパ自動構築手法の提案",日本データベース学会論文誌, Vol. 7, No. 1, pp. 263-268, June 2008.
- 4. 成田和世,北川博之, "カテゴリ型レコードデータからの属性値間の相関性を利用した外れ値 検出手法", 情報処理学会論文誌:データベース, (TOD39). Vol. 1, No. 2, pp. 38--53. September, 2008.
- 5. Sophoin Khy, Yoshiharu Ishikawa, and Hiroyuki Kitagawa, "Querying Topic Evolution in Time Series Document Clusters", 日本データベース学会論文誌, Vol. 7, No. 3, pp. 7-12, December 2008.

- 6. Djelloul Boukhelef and Hiroyuki Kitagawa, "Efficient Routing in Multi-ring Content Addressable Network",日本データベース学会論文誌, Vol. 7, No. 3, pp. 19-24, December 2008.
- 7. 李暁晨, 天笠俊之, 北川博之, "HTML 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", 日本データベース学会論文誌, Vol. 7, No. 3, pp. 67-72, December 2008.
- 8. 高橋公海, 森嶋厚行, 松本亜季子, 杉本重雄, 北川博之, "Web コンテンツ管理のための一貫性制約発見支援", 日本データベース学会論文誌, Vol. 7, No. 3, pp. 25-30, December 2008.
- 9. MoonBae Song and Hiroyuki Kitagawa, "Managing Frequent Updates in R-trees for Update-intensive Applications", IEEE Transactions on Knowledge and Data Engineering, Vol. 99, No. 1. DOI#10.1109/TKDE.2008.225.
- Hiroki Kimura, Hideyuki Kawashima, Hiroyuki Kusaka and Hiroyuki Kitagawa, "Applying a Machine Learning Technique to Classification of Japanese Pressure Patterns", Data Science Journal, Vol. 8, S59-S67, March 2009.

<国際会議発表論文>

- Atsuyuki Morishima, Akiyoshi Nakamizo, Toshinari Iida, Shigeo Sugimoto, Hiroyuki Kitagawa, "PageChaser: A Tool for the Automatic Correction of Broken Web Links", The 24th International Conference on Data Engineering (ICDE 2008), pp. 1486-1488, Cancun, Mexico, April 7-12, 2008.
- 2. Kazuyo Narita and Hiroyuki Kitagawa, "Detecting Outliers in Categorical Record Databases Based on Attribute Associations", Proc. 10th Asia-Pacific Web Conference International Conference (APWeb 2008), LNCS 4976, pp. 111-123, Shenyang, China, April 26-28 2008.
- 3. Jiajun Gu and Hiroyuki Kitagawa, "Extending Keyword Search to Metadata on Relational Databases", Proc. The first International Workshop on Information-explosion and Next Generation Search (INGS 2008), pp. 97-103, Shenyang, China, April 26-27 2008.
- 4. Ryo Sato, Hideyuki Kawashima and Hiroyuki Kitagawa, "The Integration of Data Streams with Probabilities and a Relational Database using Bayesian Networks", Proc. International Workshop on Sensor Network Technologies for Information Explosion Era (SeNTIE 2008), pp. 114–121, Beijing, China, April 27, 2008.
- 5. Zhitao Shen, Hideyuki Kawashima and Hiroyuki Kitagawa, "Lineage-based Probabilistic Event Stream Processing", Proc. International Workshop on Sensor Network Technologies for Information Explosion Era (SeNTIE 2008), pp. 106-113, Beijing, China, April 27, 2008.
- 6. Koichi Inoue, Dai Akashi, Michihiro Koibuchi, Hideyuki Kawashima and Hiroaki Nishi, "Semantic router using data stream to enrich services", Proc. International Conference on Future Internet Technologies (CFI08), Seoul, Korea, June 18–20, 2008. pp. 20–23.
- 7. Keiichirou Ui, Toshiyuki Amagasa, and Hiroyuki Kitagawa "A FUSE-Based Tool for Accessing Meteorological Data in Remote Servers." 20th International Confarence on Sientific and Statistical Database Management (SSDBM'08), pp.592-597, Hong Kong, China, July 9-11,

2008.

- 8. 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.
- 9. Tsubasa Takahashi and Hiroyuki Kitagawa, "S-BITS: Social-Bookmarking Induced Topic Search", Proc. 9th International Conference On Web-Age Information Management (WAIM 2008), pp. 25-30, Zhangjiajie, China, July 20-22 2008.
- 10. Hideyuki Kawashima, Ryo Sato and Hiroyuki Kitagawa, "Models and Issues on Probabilistic Data Stream Systems with Bayesian Networks", Proc. 2nd International Workshop on SensorWebs, Databases and Mining in Networked Sensing Systems (SWDMNSS 2008), pp. 157–160, Turku, Finland, July 28 August 1, 2008.
- 11. Xiaochen Li, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Searching XML Documents by Keywords in Structured P2P Networks", 3rd International Workshop on XML Data Management Tools and Techniques (XANTEC 2008), pp. 274–278, Turin, Italy, Sept 1 5, 2008.
- 12. Kozue Ishida and Hiroyuki Kitagawa, "Detecting Current Outliers: Continuous Outlier Detection over Time-Series Data Streams", 19th International Conference on Database and Expert Systems Applications (DEXA 2008), pp. 255-268, Turin, Italy, Sept 1 5, 2008.
- 13. Kosuke Ohki, Yousuke Watanabe, and Hiroyuki Kitagawa, "Dynamic Source Selection to Handle Changes of User's Interest in Continuous Query", 16th International Conference on COOPERATIVE INFORMATION SYSTEMS, LNCS 5333, pp. 6-7 Monterrey, Mexico, Nov 12 14, 2008.
- 14. Boukhelef Djelloul and Hiroyuki Kitagawa, "Multi-ring infrastructure for Content Addressable Networks", 16th International Conference on COOPERATIVE INFORMATION SYSTEMS (CoopIS 2008), pp. 193–211, Monterrey, Mexico, Nov 12 14, 2008. LNCS 5331 Springer 2008.
- 15. Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "GMX: An XML Data Partitioning Scheme for Holistic Twig Joins", Proc. 10th International Conference on Information Integration and Web-based Applications and Services (iiWAS 2008), pp. 137-146. Linz, Austria, Nov 24 - 26, 2008.
- 16. Chantola Kit, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Algorithms for Efficient Structure-based Grouping in XML-OLAP", Proc. 10th International Conference on Information Integration and Web-based Applications and Services (iiWAS 2008), pp. 170-177, Linz, Austria, Nov 24 26, 2008.
- 17. Sophoin Khy, Yoshiharu Ishikawa, and Hiroyuki Kitagawa, "A Query Language and Its Processing for Time-Series Document Clusters", Proc. 11th International Conference on Asia-Pacific Digital Library (ICADL 2008), pp. 82–92, Bali, Indonesia, Dec 2 5, 2008.
- 18. Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "XML Data Partitioning Strategies to Improve Parallelism in Parallel Holistic Twig Joins", Proc. 3rd International Conference on Ubiquitous Information Management and Communication (ICUIMC-09), pp. 501-510, Suwon,

- Korea, Jan 15 16, 2009.
- 19. Djelloul Boukhelef and Hiroyuki Kitagawa, "Dynamic Load Balancing in RCAN Content Addressable Network", Proc. 3rd International Conference on Ubiquitous Information Management and Communication (ICUIMC-09), pp. 98-106, Suwon, Korea, Jan 15 16, 2009.
- 20. Kosuke Ohki, Yousuke Watanabe, and Hiroyuki Kitagawa, "Evaluation of a Framework for Dynamic Source Selection in Stream Processing", Proc. International Workshop on Data Management for Information Explosion in Wireless Networks (DMIEW 2009), pp. 1106-1111, Fukuoka, Japan, March 16 19, 2009.
- 21. Kenshiro Hirose, Hideyuki Kawashima, Satoru Satake and Michita Imai, "Sharing Gesture Contents among Heterogeneous Robots", Proc. International Workshop on Data Management for Information Explosion in Wireless Networks (DMIEW 2009), pp. 1076–1081, Fukuoka, Japan, March 16 19, 2009.

<学会発表>

(A)招待講演

1. Hiroyuki Kitagawa, "StreamSpinner Project: Stream-based Real World Data Management", 基調講演, The 4th Korea-Japan Database Workshop (KJDB 2008), Fukushima, Japan, Sept. 19-21, 2008.

(B)その他学会発表

- 1. Chantola Kit, Toshiyuki Amagasa, Hiroyuki Kitagawa, "Algorithms for Efficient Structure-based Grouping in XML-OLAP", 情報処理学会研究報告 Vol.2008, No.56, pp.137-137. (jDB08), 2008年6月19日~20日.
- 2. 李暁晨, 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", iDB フォーラム 2008. 2008 年 9 月 21 日~23 日, 情報処理学会研究報告 Vol. 2008, No. 88, pp. 295-300.
- 3. Sophoin Khy, Yoshiharu Ishikawa, and Hiroyuki Kitagawa, "Querying Time-Series Document Clusters", iDBフォーラム 2008, 2008年9月21日~23日, 情報処理学会研究報告 Vol. 2008, No. 88, pp. 61-66.
- 4. 柿木 健, 北川博之, 天笠俊之, "レコードデータに対するリッジ問合せ", iDB フォーラム 2008. 2008 年 9 月 21 日~23 日, 情報処理学会研究報告 Vol. 2008, No. 88, pp. 271-276.
- 5. 高橋公海, 森嶋厚行, 杉本重雄, 北川博之, "Web コンテンツー貫性管理のための制約発見支援", iDB フォーラム 2008. 2008 年 9 月 21 日~23 日, 情報処理学会研究報告 Vol. 2008, No. 88, pp. 127-132.
- 6. Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "An Algorithm for Parallel Holistic Twig Joins on a PC Cluster", iDB フォーラム 2008, 2008 年 9 月 21 日~23 日,情報処理学会 研究報告 Vol. 2008, No. 88, pp. 1-6.
- 7. Djelloul Boukhelef and Hiroyuki Kitagawa, "RCAN: A Multi-ring Content Addressable Network", iDBフォーラム2008, 2008年9月21日~23日, 情報処理学会研究報告 Vol. 2008, No. 88, pp. 91-96.
- 8. 大喜恒甫, 渡辺陽介, 秋山亮, 北川博之, 天笠俊之, 川島英之, "対象情報源の動的変化を

- 考慮した分散ストリーム処理最適化手法の提案", iDB フォーラム 2008. 2008 年 9 月 21 日 ~ 23 日, 情報処理学会研究報告 Vol. 2008, No. 88, pp. 265-270.
- 9. Chantola Kit, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Evaluating Algorithms for TOPOLOGICAL ROLLUP in XML-OLAP", iDB フォーラム 2008, 2008 年 9 月 21 日~23 日,電子情報通信学会技術研究報告, Vol. 108, No. 211, pp. 17-18.
- 10. Yi Wang, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Ontology-based Integration for Heterogeneous River Data", iDB フォーラム 2008, 2008 年 9 月 21 日~23 日, 電子情報通信 学会技術研究報告, Vol. 108, No. 211, pp. 21-22.
- Hiroki Kimura, Hideyuki Kawashima, Hiroyuki Kusaka, and Hiroyuki Kitagawa, "Detection of a Specific Pressure Pattern around Japan by Support Vector Machine", Fifty Years after IGY –Modern Information Technologies and Earth and Solar Sciences-, Tsukuba, Japan, November 11–13, 2008.
- 12. Keiichiro Ui, Toshiyuki Amagasa, Hiroyuki Kitagawa and Takeshi Horinouchi "A FUSE-based access tool for scientific data on the Web", Fifty Years after IGY -Modern Information Technologies and Earth and Solar Sciences-, Tsukuba, Japan, November 11-13, 2008.
- 13. 大喜恒甫,山口卓郎,渡辺陽介,北川博之,川島英之,塩川浩昭, "異種データストリーム統合基盤: StreamSpinner", Web とデータベースに関するフォーラム(WebDB Forum 2008), 2008年12月1日~2日,電子情報通信学会技術研究報告, Vol. 108, No. 329, DE2008-75, p. 55.
- 14. 川島英之, 北川博之, 寺島裕貴, "ストリーム処理エンジンにおける効率的な来歴管理", 第 1 回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日~10 日.
- 15. 永元芳幸, 天笠俊之, 北川博之, "注釈を利用した XML データのためのトレーサビリティ機構の提案", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年3月8日~10日.
- 16. 大喜恒甫,渡辺陽介,北川博之,川島英之,"分散ストリーム処理における対象情報源の動的変化を考慮した問合せ最適化手法の評価",第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年3月8日~10日.
- 17. 柿木健, 北川博之, 天笠俊之, "レコードデータに対するリッジ問合せの性能評価", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009年3月8日~10日.
- 18. Jiajun Gu, Hiroyuki Kitagawa, "Metadata-aware Keyword Search in Relational Databases",第 1 回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日~10 日.
- 19. Zhitao Shen,Xin Li,Hideyuki Kawashima and Hiroyuki Kitagawa, "Pattern-based Window: A Novel Window Operator to Support Event Detection for Data Stream Processing",第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009年3月8日~10日.
- 20. 李暁晨, 天笠俊之, 北川博之, "DHT における XPath に基づく XML キーワード検索の性能評価", 第 1 回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日 ~10 日.
- 21. 佐藤亮, 川島英之, 北川博之, "データストリーム処理へのベイジアンネットワークの導入", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009年3月8日~10日.
- 22. 木村広希,川島英之,北川博之,"サポートベクターマシンを用いた気圧配置の自動分類", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日~10

目.

- 23. 高木崇, 川島英之, 天笠俊之, 北川博之, "イベント検知に基づく衛星画像とコンテンツの統合", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年3月8日 ~10日.
- 24. 高橋翼, 北川博之, "ソーシャルブックマークにおける情報の活性度を考慮した Web ページのランキング", 第 1 回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日~10 日.
- 25. 山口卓郎, 渡辺陽介, 北川博之, "センサノード上で動作する汎用データ管理基盤の開発", 第 1 回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日~10 日.
- 26. 駒水孝裕, 天笠俊之, 北川博之, "XML データに対するファセットナビゲーションのためのフレームワークFoX の提案", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年3月8日~10日.
- 27. 塩川浩昭, 渡辺陽介, 北川博之, 川島英之, "分散ストリーム処理システムにおける高信頼化手法の提案", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009年3月8日~10日.
- 28. 寺島慎太郎, 天笠俊之, 北川博之, "木直列化に基づく XML データにおける木構造の統合", 第 1 回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年 3 月 8 日~10 日.
- 29. 渡邊桂太,高橋翼,北川博之,"ソーシャルブックマークにおけるユーザ間の類似度を考慮したスパマー検出",第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009),2009年3月8日~10日.
- 30. 高橋公海, 森嶋厚行, 杉本重雄, 北川博之, "Web ページを対象とした包含従属性の効率的な発見手法", 第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009年3月8日~10日.
- 31. 渡辺知恵美, 新井裕子(お茶の水女子大), 天笠俊之, "ブルームフィルタを用いたプライバシ 保護検索における攻撃モデルとデータ撹乱法の一検討", 第1回データ工学と情報マネジメン トに関するフォーラム(DEIM2009), 2009 年3月8日~10日.

く受賞>

- 1. 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.
- 2. 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」
- 3. 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」
- 4. iDB フォーラム 2008 学生奨励賞:
 - A) 李暁晨, 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 コンテンツー貫性管理のための制約発見支援」
- 5. iiWAS2008 最優秀学生論文賞 (iiWAS2008 Best Student Paper Award)
 - A) Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "GMX: An XML Data Partitioning Scheme for Holistic Twig Joins", Proc. 10th International Conference on Information Integration and Web-based Applications and Services (iiWAS 2008), pp. 137-146. Linz, Austria, Nov 24 26, 2008.
 - B) Chantola Kit, Toshiyuki Amagasa, and Hiroyuki Kitagawa, "Algorithms for Efficient Structure-based Grouping in XML-OLAP", Proc. 10th International Conference on Information Integration and Web-based Applications and Services (iiWAS 2008), pp. 170-177, Linz, Austria, Nov 24 - 26, 2008.
- 6. DEIM フォーラム 2009 優秀インタラクティブ賞:高橋公海,森嶋厚行,杉本重雄,北川博之,「Web ページを対象とした包含従属性の効率的な発見手法」,第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), 2009 年3月8日~10日.

5. その他成果デモ展示

1. 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.