

情報爆発に対応する高度にスケーラブルなモニタリングアーキテクチャ

研究代表者	中島 達夫	早稲田大学・理工学術院・教授
研究分担者	村岡 洋一	早稲田大学・理工学術院・教授
	後藤 滋樹	早稲田大学・理工学術院・教授
	山名 早人	早稲田大学・理工学術院・教授
	甲藤 二郎	早稲田大学・理工学術院・教授
	追川 修一	筑波大学・システム情報工学 研究科・助教授

1. 研究概要

研究目的：

大規模な分散システムを安定して動作させるためにはシステムが自分自身が置かれた状況を理解することを可能とする必要がある。本研究では、そのためのインフラストラクチャとして様々な実時間に生成された情報を収集、分析することを可能とするためのモニタリングアーキテクチャに関する研究をおこなう。

本年度の研究成果の概要：

本年度の目的は情報収集のための OS アーキテクチャ、各ノードが取得した莫大な情報を収集し分析するための分散ミドルウェアの2つのサブシステムが満たすべき要求を明らかにするための検討をおこない、次年度以降へのプロトタイプを試作への知見を得ることでできた。

2. モニタリングアーキテクチャ

情報爆発に対処するため、次世代の IT サービスは、今までよりはるかに高度な信頼性、セキュリティが必要とされている。そのような新しい要求を実現するためには、システムの内状態やそのシステムが置かれた実世界の状態をリアルタイムモニタリングし、集めた情報から有用な情報を抽出することが必要不可欠となる。システムの現在の状態を様々な面から把握することにより、システムが異常な状態にならないように正しく制御することが可能となる。従来の分散システムと異なり、多様なシステムに関する情報をリアルタイムに観察することにより、システムの振る舞いを動的に現在の状況に応じて適応可能とすることや、システムが異常な状態になるのを未然に防ぐことが可能となる。本研究では、システムモニタリングに適した OS アーキテクチャの開発と収集した情報を分析可能とするミドルウェアの開発をおこなう。

モニタリングアーキテクチャはシステム全体の信頼性やセキュリティの向上を目指す。本

研究により開発したシステムを他の研究により開発したシステムと統合することにより、システム全体の安全性の向上が可能となる。

本研究の最大の特徴は、システム内の情報とセンサー等により取得した実世界情報の両方を統一的扱うことを可能とするモニタリングアーキテクチャを開発することである。従来のモニタリングアーキテクチャは、システム全体の負荷をバランスすることを最も重要な課題としているが、ハードウェアの進歩がシステム全体の効率的利用より信頼性やセキュリティの向上の方がより重要な課題となってきた。

本プロジェクトにより開発するモニタリングアーキテクチャは、システム全体のプログラムの振る舞いをモニタリングすることによるセキュリティの向上、現在のシステムの状態を利用することによるアプリケーションの自動チューニング、実世界モニタリングによる物理的セキュリティの向上を目指す。このような大規模なモニタリングアーキテクチャの開発は世界に例がなく、本システムの開発により、従来のシステムの信頼性、セキュリティを格段に向上することが可能となる。

本研究課題では以下のサブシステムに分類し研究をおこなっているため、それぞれのサブシステム毎に研究経過を報告する。

S1) 情報収集のための OS アーキテクチャ

S1-1) ローカルノードに関する詳細な情報取得を可能とするマイクロカーネル

S1-2) 取得した情報の分析、フィルタリングをおこなうためのミドルウェア

S2) 各ノードが取得した莫大な情報を収集し分析するための分散ミドルウェア

S2-1) 分散システム情報収集分析ミドルウェア

S2-2) 自己組織型セキュリティミドルウェア

本年度の目的は以上に述べたサブシステム群が満たすべき要求を明らかにするための検討をおこない、次年度以降へのプロトタイプを試作への知見を得ることである。以下の節では各サブシステム毎に本年度の成果報告に関して報告する。

3. ローカルノードに関する詳細な情報取得を可能とするマイクロカーネル

本研究では、ローカルノードに関する詳細な情報取得を可能とするマイクロカーネルとして、分散処理、ネットワーク処理に適した機能を持つ汎用 OS と、実時間処理に適した機能を持つ実時間 OS (RTOS) の共存を可能にするマイクロカーネルの研究を行っている。

ローカルノードに関する詳細な情報の取得には実時間処理が必要とされ、そのための機能を提供する RTOS が使用できることが望ましい。例えば、RTOS ではタスクの優先度に従ったスケジューリングが行われるため、重要な情報収集がそれほど重要でない情報の収集に阻

害される場合を出来る限り排除することができる。また、定期的に情報を収集するためには、RTOS の提供する周期タスクの機能を用いることで、精確な時間に情報の収集が可能になる。この時間の正確さも優先度に従った順になり、重要な情報の収集のタスクはより精確な時間に実行可能になる。

一方、取得したローカルノードに関する情報を他ノードやサーバ等に提供するため、また情報をもとに協調した処理を行うためには、分散処理、ネットワーク処理に適した機能を持つ汎用 OS が必要となる。RTOS もネットワーク処理が可能なものもあるが、多くの場合その機能は制限されたものであり、広域環境における相互接続性、分散処理を行ううえで必要となるミドルウェアの機能といった面から、汎用 OS の使用が望ましい。

そこで本研究では汎用 OS と RTOS の共存を可能にするマイクロカーネルの研究を行っている。RTOS としては μ ITRON に準拠したオープンソース実装である TOPPERS/JSP を用いる。また、汎用 OS としては近年組み込み用途にも使用されている Linux を用いる。 μ ITRON は非常に軽量な RTOS であるため、タスクの優先順位に基づいたスケジューリングの制御がしやすく、そのため重要なタスクに対し高い時間精度を提供することができる。その反面、アプリケーションと RTOS カーネル間の保護を提供しないため、1 つアプリケーションの故障が他のアプリケーションの実行に影響を与えてしまい、最悪同じ RTOS 上の全てのアプリケーションの実行に障害を引き起こしてしまう。そのため本アーキテクチャでは、複数の RTOS インスタンスの実行を可能にする。マイクロカーネルは、それぞれの RTOS インスタンスに保護ドメインを提供するため、1 つの RTOS インスタンスの障害をそのインスタンス内に封じ込めることが可能であり、障害の伝播を防ぐことができる。

複数 RTOS と Linux を同一の仮想アドレス空間上に配置するのが、本アーキテクチャの特徴である。これまで、RTOS、Linux のそれぞれに独立した仮想アドレス空間を提供することが多かった。独立した仮想アドレス空間を提供することで、マイクロカーネルの機能としては RTOS、Linux を区別する必要がないためである。しかしながら、独立した仮想アドレス空間は、RTOS、Linux の実行の切り替え時に、仮想アドレス空間の切り替えを伴う。仮想アドレス空間の切り替えは、高価なサーバ用プロセッサで提供されるタグ付 TLB のような機能が提供されない場合、非常に高価なものになってしまう。同一の仮想アドレス空間上に配置することで、そのようなオーバーヘッドを避けることができ、RTOS への切り替えを高速化することができる。

現在、本アーキテクチャに基づいたマイクロカーネルのプロトタイプ実装を進めている。RTOS としては μ ITRON に準拠したオープンソース実装である TOPPERS/JSP を用い、また汎用 OS としては近年組み込み用途にも使用されている Linux を用いている。マイクロカーネルは、Mach や L4 といった計算資源を抽象化して提供するこれまでのマイクロカーネルとは異なり、ハードウェアの保護は行いながらできるだけそのままのかたちで提供することにより、マイクロカーネル上で Linux を動作させるコストをできるだけ小さくすることを目指している。これは、Mach や L4 上での汎用 OS サポートのコストの大きさが実用面で

の大きな足かせとなり、限定的な利用に留まってしまったためである。ハードウェアの保護は行いながらできるだけそのままのかたちで提供するタイプのマイクロカーネルの例としては ExoKernel がある。しかしながら、ExoKernel 上では μ ITRON のような保護ドメインを持たない OS (LibOS) はサポートされたが、汎用 OS はサポートされなかった。

プロトタイプ実装の完了、評価が当面の課題であり、1CPU システムでの実装、評価後にマルチプロセッサシステムのサポートを検討する予定である。

4. 取得した情報の分析、フィルタリングをおこなうためのミドルウェア

本年度は取得した情報の分析、フィルタリングをおこなうためのミドルウェアの実現に向けて、以下の 2 つの基礎的な研究をおこなった。

- 1) センサー情報を収集分析するためのミドルウェアであるシナモンの構築
- 2) 汎用ロギングサービスの基本設計

ユーザが身に着ける日常品であるヘッドホン、腕時計、指輪、ベルトなどにセンサーネットワークを埋め込みユーザの行動を認識するコンテキストウェアアプリケーションの構築を支援するミドルウェアであるシナモンを実際に実装し、有効性を示すアプリケーションを作成した。

シナモンは、データを抽象化する複数のワーカとワーカ使用する入出力データを共有するための共有データスペースから構成されている。各ワーカは様々なデータ分析するためのアルゴリズムを含んでおり、共有データスペース上にある適当なデータを取り出し、データから高度のコンテキストデータを抽出し、それを共有データスペースに書き込む。複数のワーカはそれぞれ並行に動作し、センサーデバイスにより書き込まれたセンサーデータをリアルタイムに高い抽象度のコンテキストデータに変換していく。例えば、低レベルのワーカは、センサーを埋め込んだ日常物を身に着けたユーザの頭、腕、体等の動きを認識する。高レベルのワーカはそれらのワーカが共有データスペースに書き込んだデータを読み込み、歩いているとか本を読んでいるなどの高レベルのコンテキストを抽出する。

我々の研究グループでは、実際にノキア東京リサーチセンターと共同で開発したセンサーネットワーク Cookie をヘッドホン、腕時計、指輪、ベルトなどの人工物に埋め込め込んだ。Cookie は Bluetooth インタフェースをサポートしているので bluetooth をサポートしている携帯端末にデータを取り込むことが可能である。

現状では、携帯端末として Cookie と同様にノキア東京リサーチセンターと共同で開発した Muffin とノキアの Linux を使用した携帯端末である N770 上にシナモンを実装した。また、いくつかのワーカを実際に実装し、有効性を実証するための簡単なアプリケーションを作成した。本アプリケーションは現在のユーザの行動をモニタリングし、ユーザが何もする

ことがないことがわかると適当な情報を提供する。本アプリケーションでは、ユーザが、「寝ている」、「歩いている」、「立っている」、「座っている」、「本を読んでいる」、「端末を見ている」、「何もしていない」のコンテキストを認識し、「何もしていない」か「端末を見ている」時にコンテンツの提供をおこなう。

シナモンの開発から、汎用的な分散システム上のモニタリングシステムを実現する際に多くの知見を得られた。アプリケーションを容易に構築するためには、汎用的なコンテキストの表現が重要となる。実際に、多くの研究では、標準的にコンテキストを表現するためのコンテキストオントロジーや表現フレームワークが開発されている。それらの研究はオントロジーで表現された情報を容易に取得できることを前提に、コンテキストオントロジーや表現フレームワークの完全性を議論しているものが多い。しかし、我々の研究の結果からまずは、どのようなアプリケーションが存在し、それらのアプリケーションがどのようなコンテキストオントロジーを必要とするかを認識することが重要である。また、それらのコンテキストオントロジーをどのように抽出するかを検討する必要がある。その際、アプリケーションからの要求と実際に取得可能なセンサー情報を考慮して適切なコンテキストオントロジーを選択することが重要であることが理解できた。

モニタリングアーキテクチャは実世界に関する情報やシステム自体に関する様々な実時間情報を収集、分析する必要がある。また、これらの情報を利用するアプリケーションも多様である。そのため、次のステップとしては、収集可能な情報が何であるか、また、本モニタリングシステムを利用するアプリケーションは何が存在するかを明確にすることが重要であると思われる。

次の研究は汎用ロギングサービスの基本的な設計である。ロギングサービスは実世界で発生した様々な情報やシステム自体に関する情報を蓄積する。従来のロギングサービスは、各サブシステムが必要とする情報のみを収集していた。ロギングしたデータを分析することにより、過去に何が発生したかを分析することが可能となり、また、将来の障害を予測することを可能とする。また、コンテキスト情報をロギングすることによりユーザの好みや行動パターンを予測することも可能となる。従来のロギングサービスはサブシステム毎にデータを管理していたため、それらの情報を統合化し、より高度な予測や分析をおこなうことが困難であった。より多くの情報を収集することによりより高度な予測や分析をおこなうことが可能となり、システム全体の安全性を向上したり、新しいアプリケーション分野を開拓する可能性が生じる。

本研究では、システムが収集したあらゆるデータを統合化されたデータベースに格納し、様々な高度な予測や分析をおこなうことを可能とする。現状では、ロギングサービスが使用するデータ格納方式と情報の表現形式の検討をおこなっている。本システムと1)で開発したシナモンを統合することによりローカルシステム内の様々な情報を収集し、分析するためのインフラストラクチャとして利用することが可能となる。

今後の展望

来年度は以下の3点の研究を中心におこなう。

- 1) 汎用ロギングサービスの構築とシナモンとの統合
- 2) 1つのアドミニストレーションドメイン内の情報収集に関する検討
- 3) 情報収集の際のプライバシーの問題の検討

1)では、本年度におこなった研究を発展してモニタリングシステムの基盤として利用可能なシステムを構築することである。2)では、1)で開発したシステムを拡張し、より広範なエリアやシステムに関する情報を取得するための手段に関する検討をおこなう。また、3)では、実時間の実世界情報を収集する際に問題となるプライバシーの問題に関して検討をおこなう。

5. 分散システム情報収集分析ミドルウェアの研究

分散システム情報収集分析ミドルウェアの研究として、本年度は、情報収集分析のためのアーキテクチャについての課題に取り組み、汎用的な情報収集分析のスキームを提案した。

まず、情報収集分析のためのアーキテクチャとして、従来の「時間情報を含むシーケンシャルパターンマイニング」の一般化を目指した。従来、抽出シーケンス中のアイテム発生間隔は、アイテム個数 (Gap) や発生時間間隔で表現されていた (図1)。Gap はアイテム数により間隔を表現するため、時間間隔のほうがアイテム発生間隔を正確に表現できるという特徴がある。しかし、これまでの手法は、時間間隔の違うシーケンスを区別して同時にカウントできない問題や時間間隔を等間隔でしかアイテム化できないという問題点を持つ。

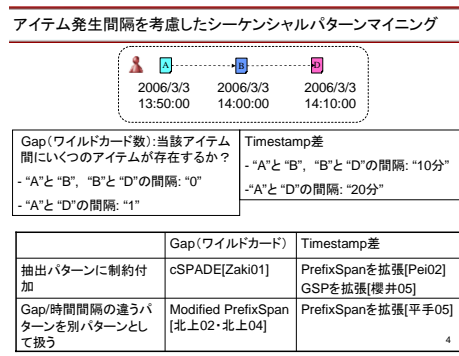


図1 従来の時間情報を含むシーケンシャルパターンマイニング手法

そこで、我々は、(1)アイテム化関数を導入することにより時間間隔のアイテム化プロセス

を一般化し、(2)アイテム化した時間間隔による制約を設けたシーケンシャルパターンマイニングを提案した。提案手法では図2に示すように時間間隔付シーケンスを用いる。

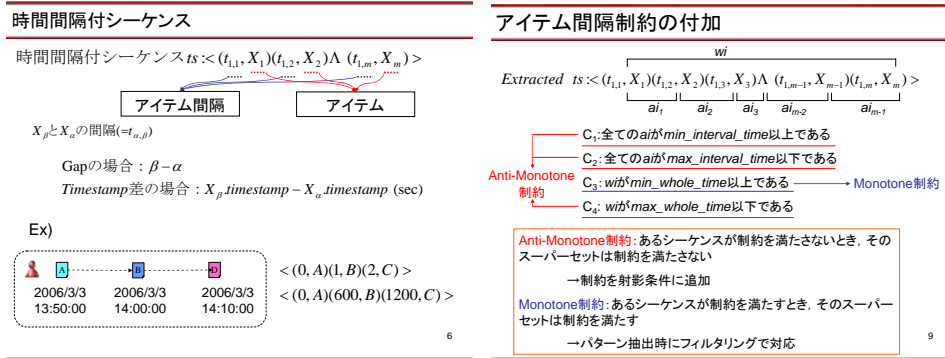


図2 時間間隔付シーケンスと制約付加 (提案手法)

提案手法を強震ネット(<http://www.k-net.bosai.go.jp>)から生成した1996年5月~2003年12月の日本で発生した地震データに適用した結果、抽出パターン数の削減および実行時間の短縮を図ることができることを確認した。



図3 実験結果

次に、情報収集分析における新しい分析を実現するための「面情報による検索」について提案、評価した。これまでの検索では、Queryを入力しQueryに基づいた検索を実現するのが一般的である。しかし、例えば「人」や「もの」という「面情報」をQueryとして入力することにより、「人」や「もの」が繋がった空間を検索・分析することが可能になると考える(図4)。

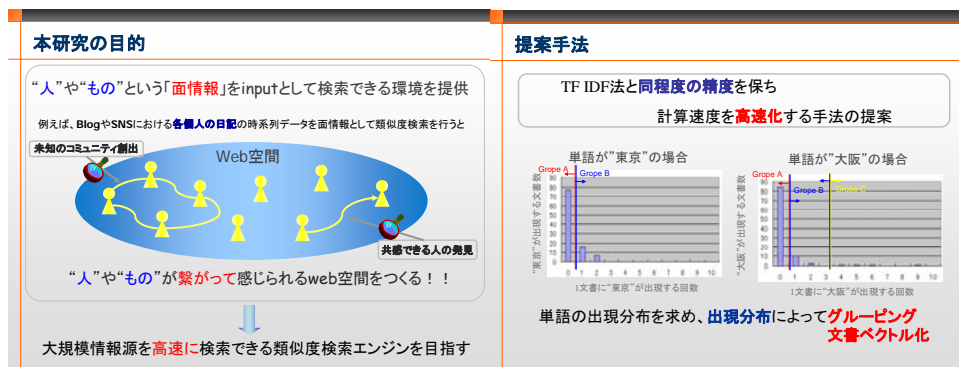


図4 提案手法(面情報検索)

試作では、基本的な情報検索手法である TF/IDF 法との比較を行い、図5に示すように上位 10%の検索結果において TF/IDF 法に比較して 86%の再現率を達成した。また、実行速度では大幅な高速化 (3.3 倍) を達成した。

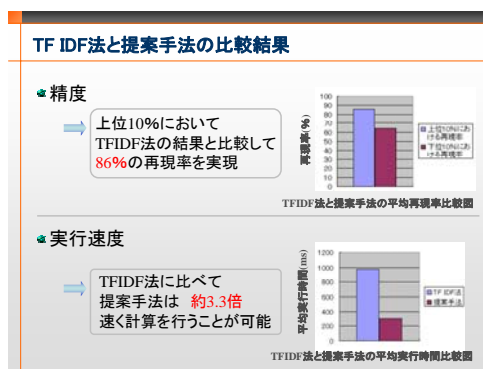


図5 実験結果

6. 自己組織型セキュリティミドルウェア

本年度は自己組織型セキュリティミドルウェアの実現に向けた第一歩として、まずネットワーク構成技術の面から、障害耐性を有する自律分散的な経路制御プロトコルの評価実験を行った。さらに、センサーによるデータ収集時に発生する各種のノード障害の影響と回避策に関する検討を開始した。またネットワーク測定技術の面から、測定性能を向上するための要素技術を検討した。さらに、ネットワーク測定の数値評価とユーザから見た通信品質(QoS)の関連を実証する実験を行った。ユーザの閲覧行動を分析するための測定法について検討を開始した。

以下に具体的に述べる。

障害耐性を有する自律分散的な経路制御プロトコルの評価実験では、IP 網上に P2P 型（オーバーレイ型）の配信木を構成して、マルチメディア情報の一対多リアルタイム配信実験を行った。この方式は、IP マルチキャストルータの制約を受けない柔軟性を有するが、一方で、配信木を構成するノードがセッション離脱する確率が高く、そのたびにリアルタイム配信が中断する。この問題を回避するために、ノード離脱が発生する前に各ノードが予備経路を計算しておき、離脱発生時には即座に予備経路に切替える経路制御プロトコルを提案し、シミュレーション、理論、ソフトウェア実装の三面から有効性を実証した。その際、帯域に余裕のあるノードでは予備経路用の帯域を確保しておくことで、離脱後に経路再構築を行う手法に対しては 2 倍以上の経路回復時間の削減を、既存の予備経路手法に対しては 2 倍以上のオーバーヘッド削減を実現した。これによって、提案方式の障害耐性とノード数に対するスケールビリティを実証した。

センサーによるデータ収集時に発生する各種のノード障害の影響と回避策の検討に関しては、まず、ノート PC に、AODV と呼ばれるオープンソースの無線アドホックネットワーク用ルーティングプロトコルを実装し、さらに独自のマルチパス拡張を行うことで、上記と同様に、ノード障害、もしくは経路障害時には予備経路に切替える方式を提案した。そして、シミュレーション評価を併用しつつ、実際のマルチメディア情報のリアルタイム配信実験を行い、既存の AODV 方式に対してセッション中断時間を 1/2 にし、かつ約 20%のスループットの改善を実現した。また、マルチメディア配信実験とは別に、自立走行型のロボットを利用した位置推定アルゴリズムに対するノード障害の影響の評価と、逆にノード障害にロバストな位置推定アルゴリズムの検討を進めている。現在は、MATLAB を用いた評価実験と、実機を用いた測定実験を並行して進めている。

ネットワーク測定は、セキュリティミドルウェアにおいて自己モニタ機能を実現するために必要である。測定は専用のハードウェアを用いる方法と、ソフトウェアによる方法がある。ハードウェアは性能の面で有利であるが、コストが高い。自己モニタのためには多数の測定点を設ける必要がある。そこでソフトウェアによる測定機能の性能を向上する研究を行った。まず、ソフトウェアによる測定性能を規定している箇所を特定するために、実際に稼動しているシステムの動作を解析した。その結果、測定データをディスクに書き込む部分がボトルネックであることが判明した。そこで、ディスクへの書き込みの回数を低減するために、測定データを圧縮する方法を導入する。さらに圧縮処理のためのバッファを設けて、ディスクへの書き込み回数を大幅に低減する。各種の圧縮アルゴリズムを実証的に評価して、測定可能な通信速度を 1.8 倍にすることに成功した。またデータを約 1/5 に圧縮することができた。

ネットワーク測定の結果は一般的に膨大なデータとして表現される。多量のデータを蓄積しただけでは、結果を簡単に把握することが難しい。本研究では、ネットワークの通信品質を人間の感覚に近いパラメータで表現することを考えた。その一例として IP 電話の通話品質に用いられる基準(R 値, MOS 値)に注目した。ここで具体的に考察したのは、通信品質の

パラメータを用いて、ネットワークの優先制御を行う方法である。通信品質の劣化しやすい無線 LAN を用いて実験を行い、標準規格である IEEE 802.11e の EDCA パラメータを通常の設定から最適値に変更することによって通話品質が向上することを実証した。このようにネットワークのパラメータを動的に制御する機能が自己組織化のために必要である。

広域のネットワーク測定の結果を情報共有するために、新しい方法が必要である。従来の方法を単純に適用して、データを集中管理すると規模が拡大するにつれて管理が困難となる。そこで P2P 方式によるデータ管理を検討する。本年度はネットワークの利用者間でデータの共有をするための P2P データ管理方式の検討を行った。

今後の展望

今後の展望は下記のとおりである。

- 1) 有線、無線共に、ノード障害に耐性を有するリアルタイム配信方式に関しては一通りの見通しを付けたが、位置推定などの具体的なアルゴリズムへの影響に関しては十分な検討が行えているとは言い難く、次年度はアルゴリズム評価に重点を置いた検討を進め、かつ配信方式へのフィードバックを試みる。
- 2) ネットワークの測定に関しては、本年度に確立したソフトウェアによる多地点の観測の他に、経路制御の要点であるルータにおける測定が重要である。次年度はルータにおける flow data の解析を行い、ネットワークの動作を把握できるように検討を進める。

7. 研究成果リスト

著書、論文

1. Shuichi Oikawa and Megumi Ito: “Experiences of Building Linux/RTOS Hybrid Operating Environments on Virtual Machine Monitors”, IJCSNS International Journal of Computer Science and Network Security, Vol.6, No.5A, pp146—152, 2006.
2. Shuichi Oikawa, Megumi Ito, and Tatsuo Nakajima: “Linux/RTOS Hybrid Operating Environment on Gandalf VMM”, The 2006 IFIP International Conference on Embedded and Ubiquitous Computing (EUC-06), pp287—296, 2006.
3. Kaori Fujinami and Tatsuo Nakajima: “Bazaar: A Middleware for Physical World Abstraction”, Journal of Mobile Multimedia, Vol. 2, No.2, pp124-145, 2006年6月

4. Tatsuo Nakajima, Ichiro Satoh: “A Software Infrastructure for Supporting Spontaneous and Personalized Interaction in Home Computing Environments”, Springer Personal and Ubiquitous Computing Journal, Vol.10, No.6, pp379 – 391, 2006年9月
5. Susanna Pirttikangas, Kaori Fujinami, and Tatsuo Nakajima: ”Feature Selection and Activity Recognition from Wearable Sensors”, The 3rd International Conference on Ubiquitous Computing Systems (UCS2006), Springer Lecture Notes in Computer Science 4239, pp516-527, 2006年10月
6. Hanaoka Kensuke, Takagi Ayako, and Nakajima Tatsuo: ”Software Infrastructure for Wearable Sensor Networks”, The 12th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, pp27-35, 2006年8月
7. Yu Hirate and Hayato Yamana: “Generalized Sequential Pattern Mining with Item Intervals”, Academy Publisher, Journal of Computers(JCP), Vol. 1, Issue 3, pp. 51-60 2006年
8. Yu Hirate and Hayato Yamana:” Web Structure in 2005”, WAW2006, Banff (2006.11.30), 2006年11月
9. 斎藤、村岡:「超高性能ヒューマノイドロボット用プラットフォーム WR-Xの開発」、日本機械学会ロボティクス・メカトロニクス講演会2006、2006年5月
10. 田代 崇, 上田高德, 堀 泰祐, 平手勇宇, 山名早人:”Web 上の文章を対象とした著作権違反自動検知システム”,日本データベース学会 Letters Vol.5, No.2, pp.25-28 (2006.9)
11. 平手勇宇, 相吉澤 明, 翁 松齡, 井奥雄一, 木戸冬子, 山名早人:”インターネットオークションにおける不正行為者の発見支援”,日本データベース学会 Letters, Vol.5, No.2, pp.77-80 (2006.9)
12. 清水 奨, 風間 一洋, 廣津 登志夫, 後藤 滋樹: “リアルタイム圧縮によるパケットキャプチャの高速化”, 情報処理学会論文誌: コンピューティングシステム, Vol.47, No.SIG 7 (ACS 14), pp183—193, 2006年5月
13. T.Kusumoto, J.Katto and S.Okubo: “Proactive Route Maintenance for Tree-Based

Application Layer Multicast and its Implementation”, IEICE Trans. on Comm., Vol.E89-D, No.12, pp2856—2866, 2006 年 12 月

14. K.Taniyama, T.Morii, S.Koizumi, K.Noguchi, Y. Kotani and J.Katto: “Experimental Evaluation of an On-demand Multipath Routing Protocol for Video Transmission in Mobile Ad hoc Networks”, Packet Video Workshop, pp145—150, 2006 年 4 月

受賞

1. 田代 崇：学生発表奨励賞，データ工学ワークショップ DBWS2006 (2006.7)
2. 平手勇宇：学生発表奨励賞，データ工学ワークショップ DBWS2006 (2006.7)

招待講演

1. Tatsuo Nakajima, “Building Intelligent Environments Using Smart Daily Objects and Personal Devices, Keynote Speech, International Conference on Computer and Information Technology, Sep. 2006.
2. 山名 早人：電気四学会関西支部専門講習会，Web ページの収集に基づく全世界 Web 構造の解析 (2006.10.6)