

放送型データ受信のためのアクティブデータベースについて

寺田 努[†] ソムヌック サグアントラクーン[†] 塚本 昌彦[†] 西尾 章治郎[†]
三浦 康史[‡] 松浦 聡[‡] 今中 武[‡]

[†]大阪大学大学院工学研究科情報システム工学専攻

[‡]松下電器産業株式会社研究本部中央研究所

あらまし: 近年, 放送衛星や通信衛星の新たな打ち上げや新サービスの開始により, 衛星を用いてデータを放送するシステムに対する注目が高まっている. このような放送型システムが普及することで放送量, サービスの種類は共に飛躍的に増加する. そのため, 一般ユーザが大量に放送されるデータから必要な情報を効率的に抽出し, 格納, 再利用できるようなシステムが望まれている. このような要求に対し, 筆者らはアクティブデータベースを放送型システムに適応させたスーパーアクティブデータベース (SADB: Super Active DataBase system) を提案し, プロトタイプの開発を行ってきた. アクティブデータベースはシステムの動作が容易に記述できるため, あらゆるタイプのデータ受信基盤として機能する汎用性の高いシステムが構築できる. 本稿では, SADB の動作管理及びデータ処理を放送型システムに適応させ, 従来のアクティブデータベースと比較して, 大量に送られてくるデータを高速に処理できるようにした. 本稿で提案する SADB を利用することで, 送信側, 受信側共に柔軟なシステムを構築できるようになる.

On Active Databases for Broadcasting Data

Tsutomu TERADA[†] Somnuk SANGUANTRAKUL[†] Masahiko TSUKAMOTO[†] Shojiro NISHIO[†]
Kouji MIURA[‡] Satoshi MATSUURA[‡] Takeshi IMANAKA[‡]

[†]Department of Information Systems Engineering, Graduate School of Engineering, Osaka University

[‡]Central Research Laboratories, Corporate Research Division, Matsushita Electric Industrial Co., Ltd.

Abstract: In recent years, many broadcast and communication satellites have been launched to start various new services which attract a great deal of attention on data broadcasting system using satellites. In such systems, the volume of data to be broadcasted and the number of services to be provided will increase significantly. This results in a growing demand on a system that can store a large volume of data, and provide users with the ability to find their necessary information in an efficient way. Highly motivated by such requirements, we have proposed the Super Active Database System (SADB) which is a data receiving system based on the active database and have implemented a prototype SADB. The use of the active database has significantly increased the flexibility in describing the system behavior. Therefore, the system can receive and process many types of data simultaneously. In this paper, we modify the operation management and the data arrangement of the SADB to accommodate to broadcast systems. As a result, the SADB can process a great deal of data forwarded at higher speed than the conventional active database. Using the new SADB proposed in this paper, we can build various flexible systems for both sender and receiver sides.

1 はじめに

近年、放送衛星や通信衛星の新たな打ち上げ、デジタル放送の開始などにより、衛星を用いた放送型システムに対する注目が高まっている。今後新たに衛星が打ち上げられることで、放送されるコンテンツの種類、量ともに大幅に増加することが予測できる。また、デジタル放送を用いる事で、一般の番組だけを放送するのではなく、コンピュータで用いられるデータなどのデジタル情報を放送するといった新たなサービスも多数提供されるようになると考えられる。

そのため、受信側にはさまざまな種類のデータが大量に放送されることになり、ユーザがこの大量の受信データを処理し、効率よく利用することは非常に困難である。

これまで、放送型システムに関する研究はサーバ側のデータ放送戦略に関するものについて多く行われている[2][3][4]。しかし、大量のデータが放送される環境においては、データ放送戦略と共に、いったんデータを蓄積し、必要な物を再利用することの重要性が非常に高い。このような観点にもとづき、筆者らはアクティブデータベースを放送型システムに適応させることにより、大量の放送型データを効率よく受信し、ユーザが蓄積された情報に手軽にアクセスできるようなシステムについて考察してきた[8]。提案するアクティブデータベースは、簡単な言語を用いて動作を記述でき、多種のデータを受信可能な放送型データ受信基盤として有効に動作する。また、今回実現したデータベースシステムは、動作記述やデータ処理部分を改良してより放送型データに適応させたため、大量のデータを高速に処理できる。さらに、本稿ではいくつかの応用システムを構築することで、従来困難であったさまざまなサービスが容易に実現できることを示す。

以下、2章では放送型システムの特徴について述べ、3章ではアクティブデータベースと放送型システムへの拡張について述べる。4章では、システムの実装について述べ、5章では本稿で提案するシステムを用いた応用システムについて述べる。6章で

評価および他システムとの比較を行ない、最後に7章でまとめと今後の課題について述べる。

2 放送型システム

一般に、放送型システムは以下の特徴を持つ。

- 全ユーザが同じ内容を受信する
放送型通信であるため、同じ衛星からの放送であればすべてのユーザが同じ情報を受信する。
- 広帯域である
衛星を用いることにより、大量のデータを送信することができる(一般に数百 kbps~数 Mbps)。
- 上りと下りの帯域幅が平衡でない
放送の帯域(下り回線)はたいへん広いが、上り回線は存在しないか、存在しても通常は電話回線などの狭い帯域しかもない。

このように、放送型システムではユーザごとに受信内容を変化させることができないため、受信される可能性のあるデータをすべて送信していることになる。したがって、通常の番組放送について考えると、放送される情報量が増えるためにチャンネル数が膨大になり、番組内容もより細分化されて、あるジャンルに特化された番組が増加することになる。このような状況においては、一般のユーザがチャンネル表を片手に膨大な情報を処理することは非常に困難であり、放送されるデータを効率よく処理できるデータ受信基盤に対する要求が高まっている。

また、データ放送は、アンケートや申込書のフォームなどのように一度利用すればそのデータを受信側に蓄積する必要がないものだけでなく、番組表などのように期限付きであるが使い捨てでないものや、天気予報などのように同じコンテンツであるが内容だけは頻繁に更新されるものなどさまざまな種類が考えられる。一度利用すれば蓄積する必要のないデータはクライアントのデータベースから即座に削除されるべきであるし、番組表は参照されても期限までは残しておくべきである。このように、放送型データ受信システムはデータの種別によってそのデータの扱いを変更できることが必要になる。

3 アクティブデータベースの拡張

本章では、筆者らが提案するスーパーアクティブデータベース (SADB: Super Active Database System) の概要について述べる。まず、拡張の元となるアクティブデータベースについて説明し、次に拡張の方針について詳しく述べる。

3.1 アクティブデータベース

アクティブデータベースとは従来のデータベースと異なり、データベース内部で起こる事象を監視し、あらかじめ定義された条件に適応する事象の発生に反応して自動的に更新等の操作を行なうデータベースシステムである [5]。

アクティブデータベースはイベント駆動型データベースと呼ばれ、その動作は発生する事象(イベント)、実行させるための条件(コンディション)、イベントによって発火する操作(アクション)の3つの組からなるECAルールによって記述される。

アクティブデータベースを放送型データ受信に利用することで、以下のような利点が生まれる。

- アクティブデータベースはイベント駆動型であるため、データの到着や参照などデータに対する動作を基本としたシステムが構築できる。したがって、必要なデータが到着したときに、そのデータに対する処理として要求が記述できる。システムの機能はこのような要求の集合として表現されることになる。
- アクティブデータベースの機能はすべてECAルールの組で表現されるため、その内容となるECAルールを部分的に変更することで機能のカスタマイズが可能である。
- データベース内容の変化を検出することができるため、データの期限切れや時間によるデータの価値の変化などに対する処理が自由に記述できる。
- データの参照をイベントとして検出できるため、たとえば‘有料データのアクセスによる課金’といった要求が簡単に記述できる。

また、システムの動作記述にECAルールを用いることの利点として以下の点が挙げられる。

- ECAルールによって要求を記述するため、そのECAルールを変更するだけで要求を柔軟に記述できる。
- ECAルールを放送することによって、放送側からある程度データの処理を制御できる。
- ECAルールはテキストで表現可能なので、システム間で機能の移動が簡単である。

3.2 スーパーアクティブデータベース

アクティブデータベースを放送型データ受信基盤として利用するために、既存のアクティブデータベースの概念を放送型データにあわせて拡張したスーパーアクティブデータベースシステム (Super Active Database System: SADB) を提案する。SADBは通常のアクティブデータベースの機能に加え、以下の機能を持っている。

- 放送されるデータを受信できる。
次々と放送されるデータを受信し、格納することができる。受信したデータは格納する前にECAルールによる格納判定をすることによって、データの取捨選択ができる。
- 各データベース間でデータやパケットのやり取りができる。
それぞれのシステム間でデータベース内のデータ等をパケットとしてやり取りできる。たとえば、あるシステムで最適化された木の構造を他のシステムに渡したり、必要とするデータがあったときに、放送されるまで待つのではなく、他のシステムから取ってくるのが可能になる。
- ECAルールを受信することができる。
送信元がECAルールを送信できることにより、放送するデータがある条件を満たしたときに削除するルールを送信したり、データが参照されたことを確認するECAルールを送信したりできるようになる。

- ECA ルールのグループ化や実行停止ができる。

ECA ルールをまとめてグループ化し、グループ単位で取り扱うことができることにより、機能の追加や削除が容易にできるようになる。また、実行や停止ができることにより、ユーザーごとにルール群を用意しておいて利用中のユーザーに関するルールのみ実行可能にするなど、場合に応じて実行する ECA ルールを変化させることができる。

- タイマ処理が可能である。

タイマ処理を可能にすることにより、一定時間ごとに自動的にデータベースの最適化をしたり、データの新鮮さを表わすパラメータの減少処理などを行うことができる。また、有効期限付きのデータに対する処理を行うことができる。

同様にアクティブデータベースを拡張したシステムとしてアクティブモバイルデータベース (Active Mobile Database System: AMDS) が提案されている [6]。AMDS は移動体計算環境の基盤システムとして提案しているアクティブデータベースであり、バケット到着の検出や、ECA ルールの送受信など SADB と同様の機能を持つ。AMDS が移動体の動作を把握し、移動体からのデータ収集を行なうことを主な目的としているのに対し、SADB は、放送型データの受信基盤として、データを受信、格納することを目的としている点が異なっている。

4 システムの実装

本章では、前章で述べたスーパーアクティブデータベースの実装およびルールシステムの改良によるパフォーマンスの向上について述べる。

4.1 システムの構成

スーパーアクティブデータベースのシステム構成を図1に示す。データ受信部では、送信されてくる放送データを受信し、データの受信イベントを発生させる役目をもつ。DB 管理部では、さまざまなデータベースに対する要求を受け、実際にデータベースの操作を行い、データベース操作イベントを発生さ

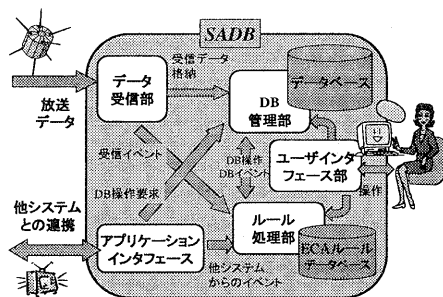


図 1: SADB のシステム構成

せる。ルール処理部では、システム上に起こるイベントを検出し、ECA ルールデータベースを検索してコンディションの評価を行ない、ルールを実行する。アプリケーションインタフェース部では、オンラインでつながる他のスーパーアクティブデータベースやアプリケーションとの通信を行なう。ユーザーインタフェース部では、ユーザーにグラフィカルなインタフェースを提供し、ECA ルールの入力や、データベース操作などを容易に行なえる環境を提供する。

4.2 ルールシステムの改良

筆者らはこれまでに、前節に述べたシステム構成でプロトタイプシステムを構築した。しかし、構築したプロトタイプシステムでは、実際の環境を想定した場合、ルール評価のプロセスがボトルネックとなり、実用に耐え得る処理速度が得られていなかった。実際の環境では、短時間に大量のデータが到着するが、データ到着ごとにイベントが発生するため、それぞれのイベントに対していちいちルールデータベースを検索し、マッチするイベントを探していたことがパフォーマンスが悪い主要因であった。

そこで本稿では、イベントの発生頻度に注目する。衛星放送をターゲットとした本システムでは、大量のデータが短時間に到着するため、「RECEIVE」イベントが非常に頻繁に、また継続的に発生する。逆に、「SELECT」等のユーザーやアプリケーションからの要求によって起こるようなイベントは、RECEIVE イ

イベントに比べて非常に少ない頻度で起こると考えられる。送られてきたデータを格納するときに起こる‘INSERT’イベントも、受信したデータの大半が不要なデータであることを考えると、RECEIVE イベントほど頻繁に発生しないことが予想できる。

つまり、RECEIVE イベントを処理する際の、ルールデータベースの検索やコンディション判定の処理がシステムのパフォーマンスのボトルネックになっていると考えられる。したがって、このボトルネックを解消するために、RECEIVE イベントに対する一次フィルタリング処理をデータ受信部において行なう。

ユーザおよびアプリケーションがRECEIVE イベントに対するECAルールを記述したとき、そのルールは一般のECAルールデータベースには格納されず、中間言語に分解されてデータ受信部が持つルールリストに登録する。さらにそのコンディションがヘッダ比較などの単純なものであった場合、そのルールはヘッダ比較リストに登録される。あるデータが到着したとき、データ受信部ではまずヘッダ比較リストを参照して適合するものが無いか調べる。その後自分が持つルールリストを検索する。こうすることで、イベントとしてルール処理部にデータを渡し、ECAルールデータベースを用いて処理を行なう手間が省け、効率のよい処理が実現できる。

5 応用システム

本章ではスーパーアクティブデータベースを用いた応用システムについて述べ、さまざまなサービスが容易に提供できることを示す。まず、データ受信側の応用システムとして、木構造を持つ情報フィルタリングシステムとスーパーアクティブデータベースを統合した放送データ受信、閲覧システムであるアクティブ情報ストアについて述べる。次に、放送側の応用システムとして、ECAルールを用いてユーザに自動的に情報を配信するオーサリングシステムについて述べる。

5.1 アクティブ情報ストア

SADBを用いることで、放送されるデータを柔軟に利用できる基盤ができた。そこで筆者らはSADBを基盤として利用し、ユーザ適応型のフィルタリングシステムを用いることで、必要な情報を容易に取得、閲覧できるアクティブ情報ストアを提案し、プロトタイプの実装を行った[7][8]。

アクティブ情報ストアは、データ受信基盤にSADB、木構造を持つ情報フィルタリングシステムを用いることにより、放送型データを受信、格納し、ユーザに対して木構造型のインタフェースを用いて容易なデータアクセスを提供する。

5.1.1 情報フィルタリングシステム

アクティブ情報ストアはユーザインタフェースとして木構造を持つビューを用いる。木の各節点はデータ分類の概念をあらわし、葉に向かうにつれてより詳細な概念になる。たとえばある節点‘ニュース’があるとすると、その葉節点として‘政治’や‘スポーツ’が考えられる。

ある放送データを受信したとき、そのデータは1つまたは複数の節点に分類される。そのデータが実際に格納されるかどうかは、その節点の価値や更新頻度、そのジャンルに対するユーザの興味などを総合的に判断して決定される。データの分類、格納に利用される木の構造は、木の偏りやアクセスの変化などの要因に応じて動的に再構築されるためユーザに特化したデータの分類ができ、ユーザは常に木をたどる操作のみで必要な情報にアクセスできる。

5.1.2 システムの構成

アクティブ情報ストアのシステム構成を図2に示す。まず、放送データをアクティブデータベースによって受信し、データ内容やサイズ等の情報を情報フィルタリング部に通知する。情報フィルタリング部においては、そのデータの評価値を計算し、情報を受信するかどうかを判断する。

ユーザが情報を必要とするときは、ユーザが木構造の節点においてデータを指定すると、フィルタリング部がアクティブデータベース部に情報を要求し、

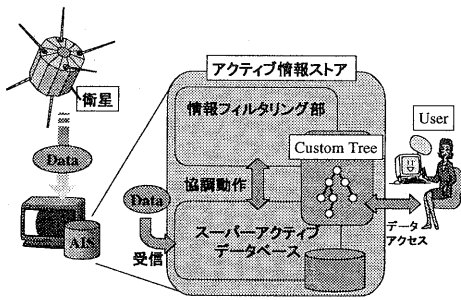


図 2: アクティブ情報ストア(AIS)の構成



図 3: システム稼働図

アクティブデータベース部が必要とされた情報を適切な形式で返し、それをユーザに提示する。

時間の流れにより情報の価値が変化するデータに関しては自動的に評価の値を更新し、期限付きデータの期限が切れた場合にはデータを削除する。

送信元からECAルールが送られてきた場合、受信判定を行い、適合するルールであった場合はシステムに格納する。

5.1.3 アクティブ情報ストアの実装

アクティブ情報ストアのプロトタイプとして、実際に衛星を利用した放送型データを受信するのではなく、WWWから情報を収集し、コンテンツを放送するサーバプログラムを開発し、放送されるデータを受信しユーザに提供するシステムを実装した。

システムの実行例を図3に示す。放送されるデータは、データタイプなどファイル情報を表すテキストや、HTML文書、画像ファイル、音声ファイルなど多岐にわたる。これらのデータを受信し、情報フィルタリング部においてジャンルごとに分類し、ユーザに提示する。ユーザがあるコンテンツを選択したときは、選択したコンテンツをブラウザ上に表示する。ブラウザとしてNetscape社のCommunicator4.03を用いた。

5.2 ECAルールオーサリングシステム

次に放送するコンテンツを作成する立場から考える応用システムであるECAルールオーサリングシステムについて説明する。

SADBの大きな特徴として、放送されるECAルールを受信できることが挙げられる。この機能を利用することで、ユーザにさまざまなサービスを提供できる。たとえば、アクティブ情報ストアのカスタムツリーにおいて、「アウトドア」に関するノードがあり、頻繁に参照されているが、「4WD車」に関するノードが無いユーザがいたとする。そのとき、「4WD車」に関するデータはツリーのノードに無いため受信されないが、両者は関連性が高いので、受信さえされれば見てもらえる可能性が高い。そこで、「もしユーザのツリーのノードに「アウトドア」があれば、「4WD車」のデータも受信する」というルールを配信しておけば、この問題は解決される。しかし、コンテンツ作成側で放送側の要求に合わせてECAルールを作成することは面倒である。したがって、ユーザからの情報や、コンテンツ制作側のターゲット要求、グローバルツリーの情報などをもとに、配信用のECAルールを自動作成する「ECAルールオーサリングシステム」を提供する。

オーサリングツールのイメージ図を図4に示す。オーサリングツールは、クライアントのSADBに置いたアンケート回収ECAルールや、ユーザのリクエスト、カスタムツリー情報などから得るユーザ情

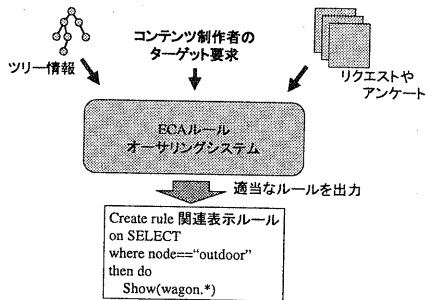


図 4: オーサリングシステムの概要

報と、コンテンツ制作側のターゲット要求、さらにグローバルツリーのなどから情報の関連性を見つけ、それに沿ったECAルールを出力する。

6 考察

本章では改良したシステムの評価および他のシステムとの比較を行なうことで、本システムについて考察する。

6.1 評価

4.2章で述べたように改良したルールシステムを実装し、性能を評価した。評価においては、5.1章で述べた応用システムアクティブ情報ストアを用いて、放送サーバから一般のHTMLデータおよびイメージデータを放送し、改良前のルール処理時間と改良後のルール処理時間を測定した。放送されるデータアイテムはそれぞれ一つまたは複数のHTMLデータおよびイメージデータを含む。また、フィルタリングによって2割の確率でデータが格納されるよう放送されるデータを作成した。また、不定期にデータベース参照などを行なうことで、RECEIVE イベント以外のイベントを発生させた。ECAルールを用いた一般のアプリケーションでは、通常10から20程度のルール数でシステムが動作している。フィルタリングにECAルールを用いる場合、フィルタリングが複雑になるにつれてECAルールの数も増加する。したがって、システム内で動作させるECA

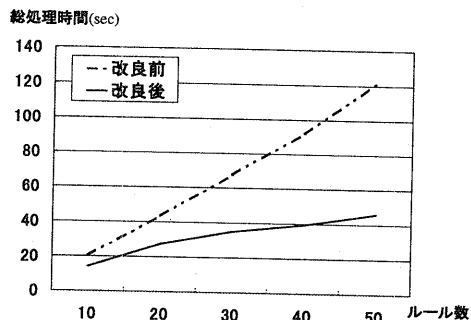


図 5: ECA ルール数に対するシステムの性能の変化

ルールの数を10から50まで変化させて、処理時間の変化を調べた。

図6.1に結果を示す。縦軸は、30データアイテム(1データアイテムのサイズは数Kbyteから数10Kbyte)を放送する間にルール処理にかかった時間を示し、横軸はルールの数を示す。改良前のシステムではECAルールの数が増加すればそれだけルール処理時間が増加しているが、改良後ではルールの増加に対して処理時間の増加が抑えられていることがわかる。また、ルール数が少ない場合でも、やはり改良後の処理時間が大幅に短縮されていることが分かる。

ルール数を固定してデータ数やデータサイズを変更した場合も同様のグラフが得られた。

6.2 他システムとの比較

本節では、すでに実用化されている放送型システムと、スーパーアクティブデータベースを用いたシステムとを比較する。

インターネット上でデータ(文字情報)を配信するシステムとしては、PointCast(ポイントキャスト社)[9]が挙げられる。インターネットを用いてデータを配信していることから、本稿で述べたアクティブ情報ストアのプロトタイプと比較すべきシステムと言える。PointCastは、コンテンツ作成者が作成したニュースなどの文字ベース情報をインターネッ

トを用いて配信し、専用のブラウザで閲覧するシステムである。ユーザは好きなジャンルを選択することで、その情報を受信することができる。

機能は簡潔で操作も軽快であるが、データの分類が1段階しか存在せず、コンテンツが増加したときに、ジャンル選択の数が非常に多くなってしまふ。アクティブ情報ストアではツリー構造を採用しているため、ジャンル数が多くなった場合に有効である。

PointCastでは、各データはどれも同様に取り扱いわれ、一定期間後削除される。したがって、興味のあるデータをまだ見ていなかったとしても削除されてしまう可能性がある。アクティブ情報ストアでは、データやジャンルにそれぞれ特性を持たせることができるため、データによって削除期間や重要度を設けることができる。また、大変興味があると判断されたデータはそれだけ保存期間も長くなるため、ユーザが見落とす可能性が低くなる。

次に地上波テレビ放送を多重化することでHTMLデータ配信を行なうサービスであるADAMS(テレビ朝日)[10]と比較する。ADAMSは、コンピュータ上に受信ボードを搭載することで、HTMLデータが受信できるサービスである。このサービスは、文字放送と同様に、地上波放送の帯域を利用したもので、ニュースや天気などの情報をユーザに提供する。HTML形式で配信され、リンクをたどるときはインターネットを利用するなど、アクティブ情報ストアと共通の部分が見られる。ただ、ADAMSは放送されるデータの取捨選択を行わず、すべてのコンテンツをハードディスク上に保持するため、放送されるコンテンツが制限される。アクティブ情報ストアでは、フィルタリングを行なうため、大量のコンテンツにも対応できる。

また、両システム共に、放送されたどのデータにアクセスが多かったかなどの情報を得る手段に乏しい。アクティブ情報ストアではECAルールを用いることで、そのような情報を取得でき、それをコンテンツ作成に利用できる。

7 まとめと今後の課題

本稿では、放送型データ受信基盤として拡張したアクティブデータベースの利用について述べた。ア

クティブデータベースを用いることにより、対応するECAルールを記述しておくだけでさまざまな種類のデータを処理し、ユーザはデータの種別を意識することなく情報を利用できるようになった。さらに、本システムを利用した応用システムについて述べ、本システムを用いてアプリケーションを構築することで、今まで提供が困難だったさまざまなサービスが容易に提供できることを示した。

今後は、さらにデータ配置やイベント処理機構を最適化することで、実際の衛星放送の速度に充分対応できるようにシステムを改良する予定である。また、さまざまなアプリケーションを構築し、本システムが放送環境において有効であることを示していく予定である。

参考文献

- [1] 原島 博: 多チャンネル時代のコンテンツ製作, 日刊工業新聞社 (1997).
- [2] 箱守聰, 田辺雅則, 石川裕治, 井上潮: 放送型通信/オンデマンド型通信を統合した情報提供システム, 情処研報, Vol. 34, No. 8, pp. 55-60 (1997).
- [3] 井上潮, 石川裕治, 田辺雅則, 箱守聰: 放送型情報提供システムにおけるキャッシュ管理方式, 信学技報, Vol.97, No.160, pp. 127-132 (1997).
- [4] 石川裕治, 田辺雅則, 箱守聰, 井上潮: ハイパーテキスト間のデータ共有を考慮した放送型情報提供方式, 信学技報, Vol.97, No.160, pp. 121-126 (1997).
- [5] G.Lohman, et al.: Extensions to starburst: object, types, functions, and rules, *Communications of the ACM*, Vol. 34, No. 10, pp. 94-109 (1991).
- [6] 村瀬亨, 塚本昌彦, 西尾章治郎: アクティブデータベースシステムによる移動体計算環境におけるデータ統合, 電子情報通信学会データ工学研究会, Vol.95, No.287, pp. 41-48 (1995).
- [7] S.Sanguantrakul, 寺田努, 塚本昌彦, 西尾章治郎, 三浦康史, 松浦聰, 今中武: 放送型データのユーザ適応型分類・選択手法, 情処研報DPS, Vol.97, No.104, pp. 249-254 (1997).
- [8] 寺田努, S.Sanguantrakul, 塚本昌彦, 西尾章治郎, 三浦康史, 松浦聰, 今中武: アクティブデータベースを用いた放送型データ格納方式, 情処研報DPS, Vol.97, No.104, pp. 243-248 (1997).
- [9] PointCast Japan ホームページ:
<http://www.pointcast.ne.jp/>
- [10] ADAMS Homepage:
<http://www.tv-asahidata.com/>