

グリッドコンピューティングによる実世界情報マイニングの提案

久保 類 真鍋 義文 盛合 敏

日本電信電話株式会社 NTT サイバースペース研究所

要 旨

ユビキタスコンピューティング環境で蓄積できる、実世界から発せられる情報を利用するために、遍在する様々な物の単位で情報を解釈するモデルを提案する。具体的には、実世界から発せられ蓄積された情報に対して、その情報を利用するために有用なメタ情報を遍在する様々な物の単位で付与し、それらを組み合わせることで有用な情報を抽出することを目指す。提案モデルの有効性を示すために、偏在する物の単位で登録された処理を実行するためのアーキテクチャを、グリッドコンピューティングを用いて設計しプロトタイプシステムを実装した。プロトタイプシステムの上で動くサンプルアプリケーションの作成により、提案モデルの有効性を示す。

Real World Data Mining using Grid Computing

Rui Kubo Yoshifumi Manabe Satoshi Moriai

NTT Cyber Space Laboratories, NTT Corporation

Abstract

This paper proposes a model to interpret information by each artifact in order to use information generated by various sensors in ubiquitous environment. Meta information is annotated to the information generated by the sensors for each ubiquitous artifact. Useful information can be derived by combining these meta information. We designed an architecture to execute programs registered for each artifact using grid computing. We implemented a prototype system and sample application programs. Finally, we show the effectiveness of the proposed model.

1 はじめに

1.1 背景

近年、ストレージ技術の急速な発展により、膨大な情報を蓄積することが出来るようになってきた。同時に、ブロードバンドの普及によりユーザがそれらの膨大な情報を利用する環境が整いつつある。

ネットワーク上では様々な情報がやりとりされ、それらの情報を蓄積する事が出来る。例えば、Web ページや電子掲示板、電子メールやテレビ電話などでは、言語や画像や音声や映像などの情報がやりと

りされ蓄積される。従って、蓄積された膨大な情報から、有用な情報を抽出する情報マイニング技術が非常に重要になりつつある。

近年、注目される新しい形の情報として、実世界から発せられる情報が挙げられる。ユビキタスコンピューティング環境では、我々の実世界での生活の中にコンピュータや無線タグやセンサなどのデバイスが遍在する世界が想定される。この世界では、従来は容易に得ることが出来なかった、人や物の位置、状態、環境などの情報を得られるようになることが期待される。

これに伴い、ネットワーク上では今まで以上に様々な形の情報がやりとりされ、蓄積される情報はより膨大になる。我々は、これらの膨大な情報を想定し、それを利用するための基盤となる技術の研究開発に取り組んでいる。

1.2 情報の蓄積とその解釈

蓄積された情報を利用するために、その情報に対する解釈を与える情報である、メタ情報は非常に有用である。例えば、音楽や映像を検索するために、それらのタイトルや作者名などからなるメタ情報は非常に有用である。また、Web ページであればページに付与されるハイパーリンク構造や更新履歴やキーワード情報は、検索エンジンの検索精度の大幅な向上や情報の集約による Web ページの閲覧性の大幅な向上を可能にした。また、メールや電子掲示板であればその引用関係に注目することで可視性が向上する。

Web ページやメールなどでは、人が意識的に情報を発信し表現をする。そのため、ハイパーリンク構造や引用関係などのメタ情報を、情報の発信者自身が人手で付与することが出来る。つまり、情報の発信者自身がその情報に対する解釈を与えることが出来る。

一方、実世界から発せられユビキタスコンピューティング環境で蓄積される情報は、人が無意識の内に蓄積される特徴があると考えられる。例えば、物や環境から発せられる位置や状態などに関する情報は、人を介さず蓄積される。また、人から発せられる情報であっても、人の体に付けることが出来るカメラや脈拍計などのウェアラブルセンサから発せられる情報は、人が意識しない形で蓄積することが出来る。そのため、これらの実世界から発せられ蓄積される情報に対して、その解釈を人手で与えることは非常に困難である。

本稿では、このような実世界から発せられユビキタスコンピューティング環境で蓄積される情報から、有用な情報を抽出するための、実世界情報マイニングのモデルを提案する。以降、第 2 章では提案モデルについて詳細を述べる。その後、第 3 章では提案モデルを実現するためのグリッドコンピューティングを用いたアーキテクチャについて、第 4 章ではそのプロトタイプシステムについて述べ、第 5 章では関連研究、第 6 章ではまとめと今後の課題を述べる。

2 実世界情報マイニング

2.1 概要

Web ページやメールのような情報の発信者が意識的に情報を発信する場合には、情報の発信者自身がその情報に対する解釈を与え、それを利用するモデルが成立した。しかし、1.2 章で述べたように、実世界に関する情報は人が無意識の内に蓄積される情報であり、なおかつ、それらの情報が人から発せられるとも限らないため、従来のモデルを適用することが出来ない。そのため、実世界から発せられる情報を利用するためのモデルを新たに考える必要がある。

ユビキタスコンピューティング環境で蓄積される様々な情報を、実世界に偏在する物の単位で解釈するモデルを提案する。具体的には、実世界から発せられる情報を、その情報に関連する様々な物の単位で取得し、各々の物の単位でメタ情報を付与し、それらを組み合わせる。

各々の物の単位でのメタ情報の付与と、そのメタ情報を組み合わせることで利用することについて順に述べる。

2.2 メタ情報の付与

メタ情報とは対象とする情報に対して機械可読な解釈情報を表現する情報である。提案モデルでは、人や物や環境などから発せられる情報をメタ情報を付与する対象とする。メタ情報を付与するための解釈の単位は偏在する様々な物である。

例 1 部屋の温度を取り上げる。元来、温度に対する解釈は様々であり、人同士でも暑がりの人や寒がりの人、物の中でも要冷蔵や常温保存の物がある。提案モデルでは、部屋の温度の情報を取得し、各々の物の単位でそれを解釈する。例えば、冷凍保存すべき商品を主体として、部屋の温度は 24 度という情報を取得した場合には、「暑い」ことを表すメタ情報を付与する。また、同じ状況で、温かい飲み物が入ったペットボトルを主体とした場合には、「寒い」ことを表すメタ情報を付与する。このように、実世界から発せられる情報を偏在する様々な物の単位で解釈する。

例 2 商店へ輸送中の商品の場合、各商品の単位で輸送状態の善し悪しを求め、メタ情報を付与す

る。この場合は輸送中の環境に関する情報を取得する。これは、倉庫の温度や湿度、輸送されているトラックの振動などの複数の情報が考えられる。これらの情報を取得し、各商品の単位で輸送状態の善し悪しを判定する。このように、物の単位でメタ情報を求めるために取得する情報は、一種類に限られず必要に応じて複数種類取得する。

例3 家庭で使われているコップの単位で「飲まれている」ことを表すメタ情報を求めるために、コップの傾きに関する情報を求める。コップに傾きセンサが付いている場合にはそこから情報を取得する。しかし、対象としているコップに傾きセンサが付いているとは限らない。そのような場合には、コップが映っている映像情報を取得して、そこから傾きを求める手段が考えられる。このように、物の単位でメタ情報を求めるために、有効な情報が複数種類あれば、取得の可能性や処理の容易さなどの観点から最適な情報を取得する。

2.3 メタ情報の組み合わせと利用

提案モデルでは、実世界から発せられる情報を利用するために有用なメタ情報を、監視や管理などの様な特定の目的に着目してトップダウンで定義するのではなく、物に着目して定義しボトムアップで有用な情報を抽出する。これにより、広く有用な情報を抽出する可能性を拡げ、特定の目的に着目していた場合には想定できなかった情報を得られるようにする事を目指す。

具体的には、様々な物の単位で付与されたメタ情報から、いくつかの観点で組み合わせを求める。これは、様々な物に付与されたメタ情報の集合の中から、情報の利用者の観点に基づいた有用な部分集合を求める処理に対応する。メタ情報の集合とは、物理的な空間や時間軸で指定される範囲内に存在する物に付与されたメタ情報の集合である。例えば、ある部屋に存在する全ての物に最近1週間の間で付与されたメタ情報の集合、などのことである。このようなメタ情報の集合から、観点に基づいた部分集合を求める。

観点の例を次に挙げる。

よくあることを知りたい 繰り返し付与されるメタ

情報の頻出パターンを求める。例えば、ドアが開くと電気がつくことや、コップにコーヒーが注がれると会議が始まることなどを発見したい場合に適する。

特定の人や物について知りたい 特定の人や物を中心にして、それらと接点のある物に付与されたメタ情報を求める。特定の人や物について振り返ったり、監視をしたい場合に適する。

全体の流れを把握したい 全体の流れに変化がある時には、多くの物にメタ情報が付与されると仮定し、そのメタ情報を求める。家や建物など、広い空間に対して長期的な流れを発見したい場合に適する。

なぜなのか知りたい その集合を特徴付けるメタ情報の特異なパターンを、他の集合と比較することで求める。うまくいった時に、なぜうまくいったのかを発見したい場合などに適する。

このように、提案モデルでは、実世界から発せられる情報を実世界に遍在する様々な物の単位で解釈する。具体的には、実世界から発せられ蓄積される情報に対して、その情報を利用するために有用なメタ情報を、遍在する様々な物の単位で付与し、それらを組み合わせることで、有用な情報を抽出することを目指す。

3 グリッドコンピューティングによる実世界情報マイニング

3.1 アーキテクチャ

提案モデルを実現するために必要な、遍在する様々な物の単位で、実世界から発せられた情報を解釈するためのアーキテクチャについて述べる。具体的には、実世界に遍在する物に対してプログラムを登録する機能と、登録されたプログラムを実行する機能を実現する。

ユビキタスコンピューティング環境を想定し、実世界に偏在する物をネットワーク側から識別可能な状況を想定する。例えば、物にはID(識別子)情報が記録されている無線タグが付与され、無線タグリーダを介して各々の物を識別できる状況である。

プログラムを登録する機能とは、各々の物に対して、メタ情報を求めたり、そのメタ情報を組み合わ

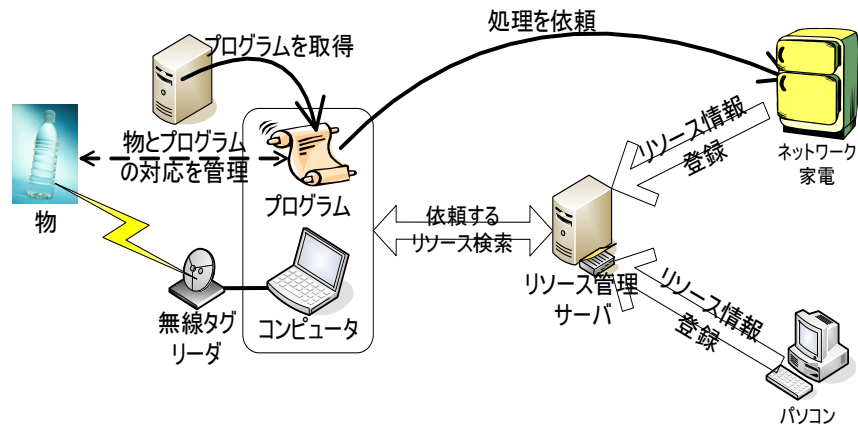


図 1: プログラムの実行を依頼する流れ

せたりするプログラムを登録する処理である。プログラムは、メタ情報を求めるために必要なセンサ情報を取得し、メタ情報を求め、その情報を蓄積する。また、メタ情報を組み合わせるために、他の物に付与されているメタ情報を取得し、それらから新たな情報を求めたり、可視化や何らかのアクションを起こしたりする。

また、物に対してプログラムを登録する人は特定の人に限られない。例えば、対象としている物の製造者や組み立て業者が、物の特性や特徴をよく知る立場から、メタ情報を求めるプログラムを登録したり、また、物の所有者や利用者が、それぞれの立場からメタ情報を利用するプログラムを登録することを可能にする。このように、遍在する物に対して、特定の目的だけではなく様々な人が様々な立場と観点からメタ情報を付与したり、それを利用することを可能にすることは、提案モデルを実現するための重要な仕組みである。

登録されたプログラムを実行する機能では、実際にメタ情報を求めたり、それを利用するプログラムを実行する。ユビキタスコンピューティング環境で想定されるような多くの物は、物自身でプログラムを実行する機能を持たない。そのため、プログラムを実行可能な場所で、そのプログラムを物に代わって実行する事を可能にする。

3.2 グリッドコンピューティング

各々の物に登録されているプログラムを実行するために、グリッドコンピューティング [7] に注目を

した。

グリッドコンピューティングとは、パワーグリッド（高圧送電線網）に由来する言葉である。コンセントを通していつでも必要なだけ電力を利用できるように、ネットワークを通して、それを意識することなくコンピューティングリソースを利用しようとする技術である。近年話題の遊休PCを使ったPCグリッドやビジネスグリッドは、その一適用例である。

提案モデルを実現するために、グリッドコンピューティングを利用した。具体的には、ユビキタスコンピューティング環境で遍在するコンピュータや無線タグやセンサなどのデバイスが、グリッドコンピューティングの各ノードを構成し、各々のリソースを相互に利用可能にする。

無線タグやセンサは実世界から発せられる情報を捉える役割を担う。具体的には、人や物や環境を識別したり、映像や音や温度や位置などの情報を取得する。

物は無線タグやセンサが捉えた情報を解釈する役割を担う。具体的には、物の単位でメタ情報を付与したり、そのメタ情報を組み合わせるために、物に対して登録されたプログラムを適切なタイミングで実行する。多くの物は、対応付けられたプログラムを自ら実行する機能を持たないため、プログラムの実行を他のコンピュータへ依頼する(図1)。

依頼先は、身近に遍在するようなネットワーク家電であったり、将来実現されれば小型コンピュータが組み込まれた机や壁などである。また、場合によっては、身近とは逆に、遠くにあるサーバへ実行を依頼することも考えられる。このような依頼先の

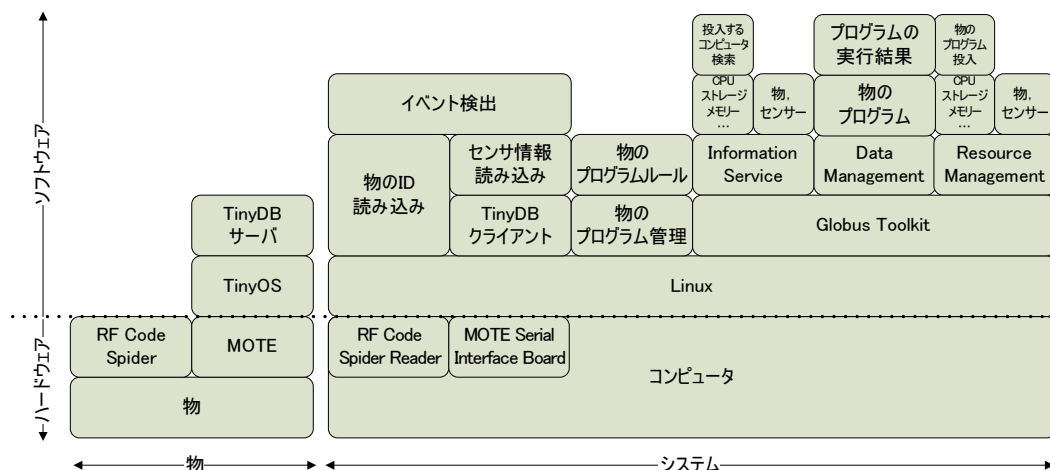


図 2: プロトタイプシステムの機能構成図

選択は、セキュリティやプログラムの規模などの観点から状況に応じて行う。

4 プロトタイプシステム

4.1 概要

提案モデルを実現するために、グリッドコンピューティングを使い、各々の物に登録されているプログラムを実行するプロトタイプシステムの実装を行った(図2)。

4.2 情報の蓄積

実世界から発せられる情報を蓄積する環境を構築した。具体的には、実世界に無線タグとセンサを遍在させた。

物を識別可能にするために、アクティブ型の無線タグを物に付与した。具体的には、RF Code Spider[2]を使用した。アクティブ型の無線タグは、ID(識別子)情報を逐次発信し続けるため、パッシブ型の無線タグと比べた場合に、無線タグリーダとの間で長距離の通信をすることが可能である。環境にも依存するが、無線タグリーダとの間で5~6m程度の距離が離れていても通信をすることが可能である。電波の読み取り範囲内にある無線タグリーダを介して、無線タグの発信するID情報を取得、蓄積する。

センサ情報を蓄積するために、無線通信が可能なセンサを使用した。具体的には、MOTE[10]を使用

した。MOTEを使用することで音、光、温度、加速度、磁気に関するセンサ情報を取得することが可能になる。また、センサ情報を取得するために、MOTEで扱うセンサ情報を管理するミドルウェアのTinyDB[3]を使用した。

4.3 グリッドコンピューティング

グリッドコンピューティング環境を構築するためにGlobus Toolkit[1]を使用した。Globus Toolkitはグリッドコンピューティングを実現するためのミドルウェアである。

Globus Toolkitの持つ機能を元に、グリッドコンピューティングの実現に必要な機能と、提案モデルを実現するために必要な機能を整理する。

Globus Toolkitは、大きく3つの機能群から構成される。

Information Services コンピューティングリソースの情報を管理する機能

Data Management ファイルやストレージを管理をする機能

Resource Management コンピューティングリソースを利用する機能

Information Servicesは、グリッドコンピューティングを構成するコンピュータのCPU、ストレージ、メモリー、ネットワーク、OSに関する情報をディレクトリサーバで管理する。通常のグリッドコンピュー

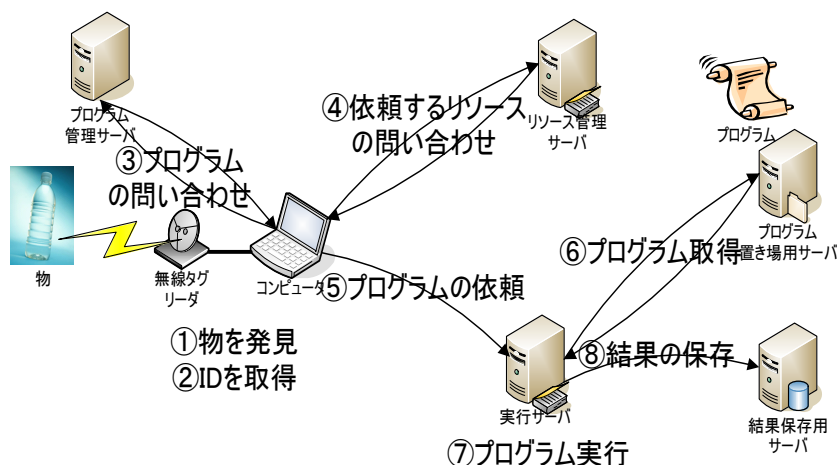


図 3: プロトタイプシステムの処理の流れ

ティングでは、プログラムの実行を依頼する時に、依頼に適したコンピュータを検索するためにディレクトリサーバへ問い合わせを行う。プロトタイプシステムでは、通常のコンピューティングリソースに加え、物とセンサに関する情報を管理するように拡張した。これらに基づき、プログラムの実行を依頼するのに適したコンピュータを検索する時に加えて、プログラムの実行に必要な物やセンサを検索する時にも、ディレクトリサーバへ問い合わせを行う。

Data Management は、グリッドコンピューティングを構成するコンピュータ間でのファイルのコピーや、分散ストレージを容易に実現するための仕組みを提供する。プロトタイプシステムでは、Globus Toolkit が提供する機能をそのまま使用した。

Resource Management は、グリッドコンピューティングを構成するコンピュータの持つリソースを相互に利用する機能を提供する。具体的には、プログラムの実行を依頼する機能を提供する。プロトタイプシステムでは、プログラムの実行の依頼に加え、物やセンサと情報のやりとりをする機能を提供する。

4.4 サンプルアプリケーション

プロトタイプシステムの上で動く、ペットボトルの単位で「飲まれている」ことを表すメタ情報を付与するアプリケーションを実装した。

アプリケーションを動かすために、ペットボトルへ付与した無線タグの ID と、ペットボトルへ対応付けるプログラムの実行情報（プログラムの保存場

所、プログラムの実行結果の保存場所）を登録した。実行したプログラムは、0.5 秒ごとにペットボトルの加速度センサの情報を取得し、その変化から「飲まれている」「飲まれていない」の情報を求める。

プロトタイプシステムの上で、物の単位でメタ情報を付与する流れを示す（図 3）。

1. ペットボトル
 - (a) 無線タグリーダの読み取り範囲内に置かれる
2. プロトタイプシステム
 - (a) ペットボトルを発見する
 - (b) ペットボトルに登録されているプログラムの実行情報を取得する
 - (c) ディレクトリサーバへ問い合わせを行い、プログラムの実行依頼先を選択する
 - (d) 実行依頼先へプログラムの実行を依頼する
3. 実行依頼先のコンピュータ
 - (a) プログラムを保存場所から取得する
 - (b) プログラムを実行する
4. プログラム
 - (a) ディレクトリサーバへ問い合わせを行い、ペットボトルに対応する加速度センサの待ち受け情報を取得する
 - (b) センサへ問い合わせを行い、0.5 秒ごとに加速度センサの情報を要求する Query を送る
 - (c) 0.5 秒ごとに得られる加速度と、基準値となる加速度の差分を求め、閾値を超える（飲まれている）か、超えない（飲まれていない）かを出力する
5. 投入されたコンピュータ

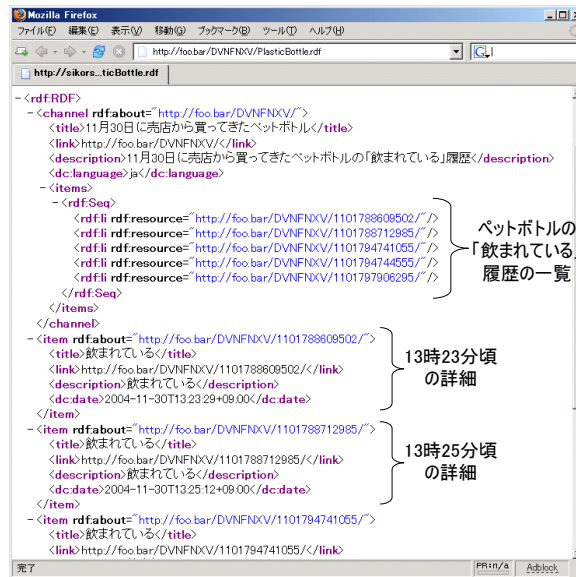


図 4: 物が求めたメタ情報の可視化の例

(a) 標準出力の保存場所へ出力を保存する (図 4)

プロトタイプシステムの上で動くサンプルアプリケーションを実装し、提案モデルを実現するために必要な、物に対応付けられたプログラムを実行できることを確認した。

5 関連研究

様々なメディアで表現される情報に対して、その情報に対する機械可読な解釈情報を与える事を目的とした、メタ情報を自動で付与しようとする研究は従来から数多く行われている [6, 9]。

近年は、センサなどのデバイス技術の発展により [5, 12]、実世界から発せられる情報を蓄積することが可能になってきた。そのため、これらの情報を対象に、メタ情報を付与し、それを利用しようとする研究が盛んに行われている。

Kern ら [8] は、ミーティングの最中の人の頭や胸にカメラやマイクを付け、それらから得られる音声、映像データを対象に、ミーティングに特化したメタ情報を付与している。人に付けられた加速度センサや位置センサの情報を使い、メタ情報を求め、その有効性を確認している。しかし、ミーティングに特化したメタ情報を付与することを想定しているため、広く有用な情報を抽出するためのメタ情報を付与する仕組みは考慮されていない。

Tsuchikawa ら [14] は、複数のウェアラブルセンサを使い、人-人や人-物の間のインタラクションを自動で抽出している。ここでのインタラクションとは、何かを見ていることを表す情報や、誰かとしゃべっていることを表す情報などで構成される。特定の目的に特化しない汎用的なメタ情報を付与するための基本的な枠組みを提供するが、人を主体としたメタ情報を付与することを想定しているため、人と直接のインタラクションがない物や環境に関するメタ情報を付与することが出来ないし、遍在する様々な物を主体としたメタ情報を付与する事が考慮されていない。

ユビキタスコンピューティングの分野では、センサやコンピュータを組み込んだ新たな人工物を使い、人のコンテキストを捉え [13]、人とコンピュータとの間のインタラクションの質的向上を目指した研究が盛んに行われている。

Beigl ら [4] は、コップに加速度センサや温度センサを付与し、これらのセンサ情報を赤外線通信を介してネットワーク側から取得可能にしている。Prante ら [11] は、壁や机やイスに、画面や接触センサやコンピュータを組み込んだミーティング環境を構築している。これらのような、センサやコンピュータが埋め込まれた物や部屋を構築する技術は、提案モデルで想定する環境を実現する上で重要である。

6 おわりに

ユビキタスコンピューティング環境を想定し、実世界から発せられる様々な情報を利用するために、Web ページやメールなどを対象にしている従来型のモデルは適用できないことを示し、偏在する様々な物の単位で情報を解釈するモデルを提案した。提案モデルの有効性を示すために、偏在する物の単位で登録された処理を実行するためのアーキテクチャを、グリッドコンピューティングを用いて設計しプロトタイプシステムを実装した。プロトタイプシステムの上で動く、サンプルアプリケーションを作成し、提案モデルの有効性を確認した。

今後は、グリッドコンピューティングによるアーキテクチャの機能拡充を行い、物に対応付けるプログラムの作成をより容易にする。具体的には、状況に応じた最適なセンサ情報の取得を支援する仕組みや、物や実行するプログラムのポリシーに応じて、プログラムの実行の最適な依頼先を選択できるようにする。また、物とプログラムの対応付けや、プログラムが求めたメタ情報を蓄積、可視化するためのインタフェースを規定する。

参考文献

- [1] Globus Toolkit ,
<http://www-unix.globus.org/toolkit/>.
- [2] RF Code Spider , <http://www.rfcode.com/>.
- [3] TinyDB ,
<http://telegraph.cs.berkeley.edu/tinydb/>.
- [4] Beigl, M., Gellersen, H.-W. and Schmidt, A.: MediaCups: Experience with Design and Use of Computer Augmented Everyday Artefacts, *Computer Networks*, Vol. 35, No. 4, pp. 401–409 (2001).
- [5] Beigl, M., Krohn, A., Zimmer, T. and Decker, C.: Typical Sensors needed in Ubiquitous and Pervasive Computing, in *First International Workshop on Networked Sensing Systems*, pp. 153–158 (2004).
- [6] Elizabeth, S., Bates, R., Taylor, P., Stolcke, A., Jurafsky, D., Ries, K., Coccaro, N., Martin, R., Meteer, M. and Ess-Dykema, C. V.: Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech, *Language and Speech*, Vol. 41, pp. 443–492 (1998).
- [7] Foster, I., Kesselman, C. and Tuecke, S.: The Anatomy of the Grid: Enabling Scalable Virtual Organization, *The International Journal of High Performance Computing Applications*, Vol. 15, No. 3, pp. 200–222 (2001).
- [8] Kern, N., Schiele, B., Junker, H., Lukowicz, P., Troster, G. and Schmidt, A.: Context Annotation for a Live Life Recording, in *Proceedings of Pervasive 2004 Workshop on Memory and Sharing of Experiences*, pp. 9–14 (2004).
- [9] Kipp, M.: The Neural Path to Dialogue Acts, in *Proceedings of the 13th European Conference on Artificial Intelligence*, pp. 175–179 (1998).
- [10] Levis, P., Madden, S., Gay, D., Polastre, J., Szewczyk, R., Woo, A., Brewer, E. and Culler, D.: The Emergence of Networking Abstractions and Techniques in TinyOS, in *Proceedings of the First USENIX/ACM Symposium on Networked Systems Design and Implementation*, pp. 1–14 (2004).
- [11] Prante, T., Streit, N. A. and Tandler, P.: Roomware: Computers Disappear and Interaction Evolves, *IEEE Computer*, Vol. 37, No. 12, pp. 47–54 (2004).
- [12] Schilit, B. N. and Sengupta, U.: Device Ensembles, *IEEE Computer*, Vol. 37, No. 12, pp. 56–64 (2004).
- [13] Schmidt, A., Beigl, M. and Gellersen, H.-W.: There is more to Context than Location, *Computers & Graphics*, Vol. 23, No. 6, pp. 893–901 (1999).
- [14] Tsuchikawa, M., Iwasawa, S., Ito, S., Nakahara, A., Sumi, Y., Mase, K., Kogure, K. and Hagita, N.: Experience-sharing System using Ubiquitous Sensing Environments, in *Proceedings of the 2nd International Symposium on Ubiquitous Computing Systems*, pp. 49–56 (2004).