

分散型データベースシステムの検索方法及び性能評価

6W-8

広瀬直美 勝本道哲 柴田義孝

東洋大学工学部情報工学科

1 はじめに

地場産業としての繊維工業分野においては、その地域の特色を活かした伝統的な織物、生地、柄などの画像データを、その地域毎にデータベース化し、それらを有効活用しようという動きがあり、自律分散したデータベースを統合し、透過的にアクセスできるような情報検索システムが望まれる。そこで著者らは、テキストと画像より構成された、複数の分散データベースを広域ネットワークにより相互接続し、クライアント-エージェント-サーバモデルによるダイナミックハイパーメディアシステム (Dynamic Hypermedia System: 以下 DHS) によるデザイン画像データベースを構築した [1,2]。本稿では、エージェント-サーバ間の検索プロトコルとして、Burst Filtering 法 (以下 BF 法) と Selective OID 法 (以下 S-OID 法) の2つを提案し、プロトタイプシステムにおいて性能評価を行ったので報告する。

2 デザイン画像データベースシステム

本システムは、DHS を基本として、複数のクライアントとマルチメディアデータベース (以下 MDB)、及び1つの知識エージェントより構成される (図1)。本システムは、埼玉県の地場産業である繊維工業のデザイン画像データベースシステムであり、デザイナーのためのデザイン支援システムである。これまで各地域で構築、管理されていたデータベースを統合し、分散型データベースとして管理する方式を考えている。本システムには、以下のような特徴がある。

- 1). 地域毎に独立的に分散管理されている。
- 2). 全ての MDB が同じデータモデルを用いた homogeneous database である。
- 3). 全てのオブジェクトが同一のスキーマを持っている定型である。
- 4). 従来のキーワード検索に加えて、ユーザモデルによって個々のユーザの感性を反映した感性検索が可能である [3]。

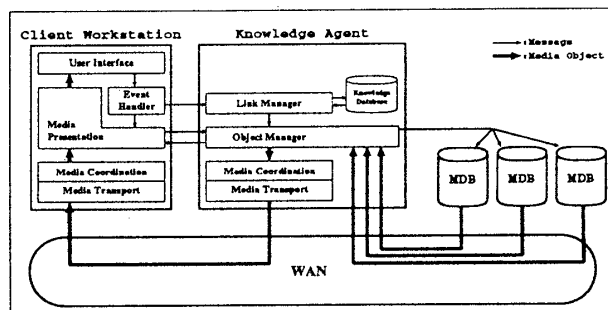


図1: ダイナミックハイパーメディアシステム

格納されている各オブジェクト (生地に関する画像データとテキストデータを統合したもの) には、オブジェクト識別子 (以下 OID) がつけられており、システムはこの OID により、オブジェクトを一意に決定することが出来る。MDB に格納されているオブジェクトには、重複されて格納されているものがあるので、この重複を取り除いてからユーザに提供する必要がある。

3 検索プロトコル

上記のようなデザイン画像データベースにおけるエージェント-サーバ間の検索プロトコルとして、以下の2つの方法を提案する。

1) Burst Filtering 法

この方法は、クライアントから発行された User Query (以下 UQ) を、知識エージェントにおいてユーザモデルを用いて Agent Query (以下 AQ) に変換した後、各 MDB にその AQ を発行する。AQ を受けとった MDB は、その AQ に合致するオブジェクト全てを知識エージェントに転送する。知識エージェントでは各データ中の OID をチェックし、重複を除去してクライアントへ転送する (図2)。

2) Selective OID 法

この方法は、クライアントから発行された UQ を、知識エージェントにおいてユーザモデルを用いて AQ に変換し、各 MDB に AQ を発行するまでは、上記の BF 法と同じである。知識エージェントからの AQ を受けとった MDB は、OID のみを転送する。知識エージェントでは、その OID をチェックし、重複を調べ、重複が除去できたら、各 OID を検索した MDB に、選択された OID による Query (OID Query: 以下 OQ) を発行する (図3)。

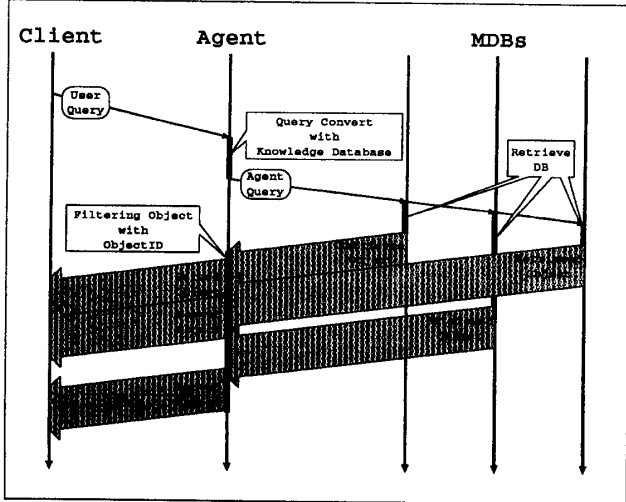


図 2: BurstFiltering 法

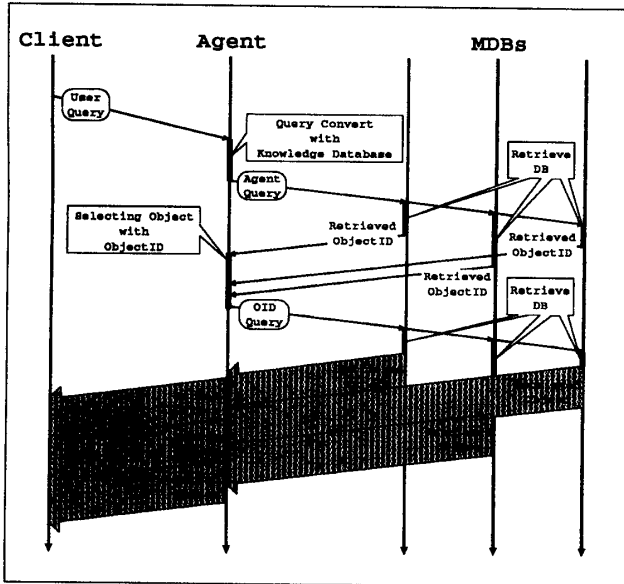


図 3: SelectiveOID 法

4 性能評価

本稿では、提案した2つのプロトコルの有効性を確かめるため、以下のようなプロトタイプを用いて性能評価を行った(図4)。クライアント、知識エージェント各1台、MDB2台で、関係モデルを用いて、同じSQL言語、及びスキーマ構造を持った登録件数100件のデザイン画像が格納されている。ネットワークはEthernetを用いた。評価は、データベースサーバが1台の場合と2台の場合で比較し、ユーザが検索条件を設定し検索ボタンを押してから、画像が表示し終るまでの時間を測定した。グラフ中の0~100%の数字は、検索されたオブジェクトの重複率を示している。以下に示すグラフ(図5)は、BF法における評価結果である。これを見ると、重複率が50%までの場合は、データベースサーバ2台の時の方が、1台の時よりも性能

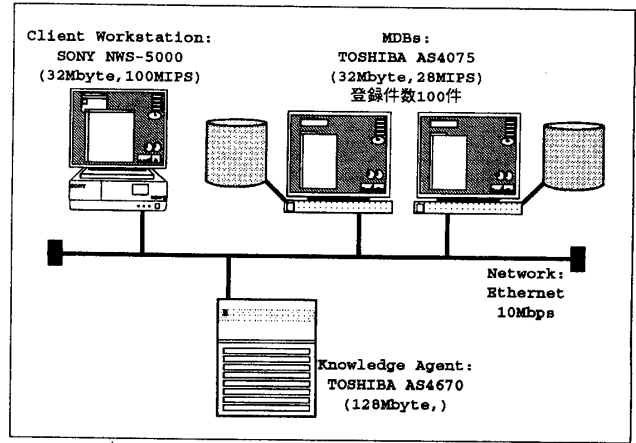


図 4: 性能評価プロトタイプ

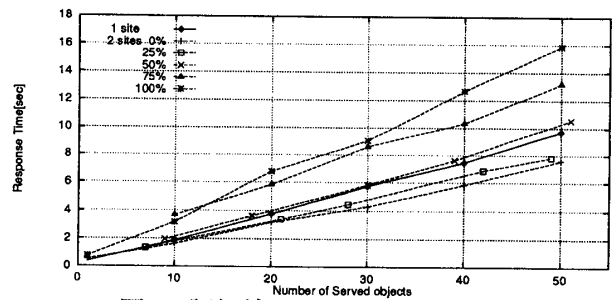


図 5: 評価結果:Burst Filtering 法

の向上が見られるが、それを越えると、性能は急に悪くなっている。

5 まとめ

本稿では、エージェント-サーバ間の検索プロトコルとして、BF法とS-OID法の2つを提案し、性能評価を行った。その結果、BF法では、オブジェクトの重複率が50%以下の場合には、MDBが2台の方が1台の場合よりも性能が向上することが判明したが、それ以上の重複率になると性能は悪かったので、分散の効果は見られなかった。今後の課題としては、MDBが2台以上の場合、及びデータベースの分割の割合、重複の割合等をパラメータとした評価を行う予定である。

参考文献

- [1] M.Katsumoto, Y.Shibata, "Dynamic Hypermedia System Using Knowledge Agent for Multimedia Information Network", JWCC-8, C2-2-1~C2-2-8, Dec.1993
- [2] 勝本, 入江, 広瀬, 瀬田, 柴田, "ISDNによる広域デザイン画像データベースの構築と性能評価", 情報処理学会マルチメディア通信と分散処理 64-9, 1994.3.3
- [3] 勝本, 入江, 柴田, "デザイン画像データベースシステムにおけるヒューマンインタフェースの研究", 情報処理学会第48回全国大会, 2E-6, Mar.1994