

# 検索対象の情報保有力を考慮した情報検索エージェント方式の提案

金子 祐真<sup>†</sup> 長橋 和哉<sup>†</sup> 小泉 寿男<sup>†</sup>

東京電機大学理工学部情報システム工学科<sup>†</sup>

東京電機大学大学院理工学研究科<sup>‡</sup>

## 1. はじめに

本稿では、サーバを必要としない非集中型システムにおける情報検索エージェント方式を提案する。ユーザの知り合いというグループでの情報の共有と他のユーザのグループとの円滑な共有を目的とする。グループはユーザの集合と捉え、ユーザの持つ情報をグループの情報とする。検索はグループにまたがるユーザを利用し、グループからグループへと検索を伝播する。検索の際にユーザが欲する情報を誰が持っているかを特定し、情報を保有するユーザとそのユーザのグループの情報を情報保有力とした。情報保有力の高いユーザを重点的に検索することによって効果的な検索を目指す。

## 2. 情報検索エージェント方式

### 2.1. 本方式の流れ

本研究は、面識の無いユーザ間での電子文書の検索、取得を行うと想定し、電子文書の検索を目的とした。全体の流れとして、まず検索要求をユーザが送信する。検索要求の送信の際に、検索要求ユーザは検索を行うエージェントに対し検索条件である検索キーワードと取得する Indicate キーワードの数を決定する。受信したエージェントは自分の持つ電子文書の検索を行う。さらにエージェントが持つリストに記載されているエージェントへと検索要求を送信する。これを検索要求ユーザの定めた回数繰り返す。検索を行ったエージェントは検索結果を検索要求ユーザに送信する。検索結果を受信したエージェントは検索結果から検索キーワードを持つ電子文書の分布を測定し、次回の検索において利用する。

### 2.2. 検索キーワード伝達

ユーザが検索要求を出すと、エージェントは自分が持つ他のユーザの IP アドレスが記載されたリストを参照し、他のユーザのエージェントへ検索要求を伝播する。伝播する際に、伝播が重ならないよう、また情報の場所を特定する為に、グループに番号をつけ、階層構造になるように次に示す方法でキーワードの伝播を行う。

要求の伝播と階層構造は図 1 のように形成される。

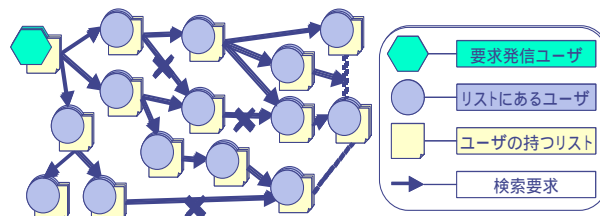


図 1 検索要求の伝播概念図

階層構造を形成する為のルールを以下に記述する。検索要求を受信したエージェントは自分の持つリストにある他のユーザのエージェントへと検索要求を伝播する。

検索要求を発信したエージェントは親となり発信先のエージェントを子としたグループを形成する。

検索要求を受信した子のエージェントは自分がリストを持っていれば検索要求を発信し、親となる。

検索要求を発信した結果、すでに親子のグループを形成している場合は検索を依頼せず、検索中はリストから削除しておく。

各エージェントの持つ検索要求と階層データは検索が終了するまで保存する。要求の伝播によって形成される階層構造とグループを図 2 に示す。

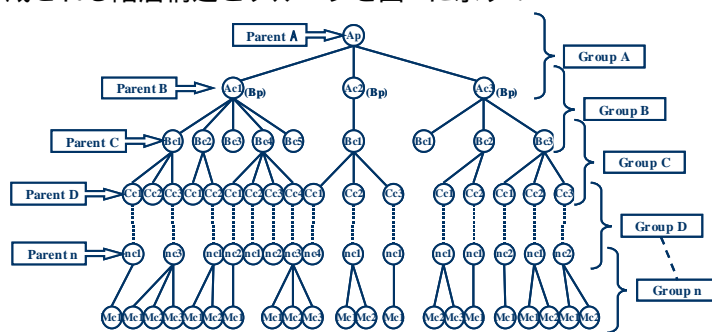


図 2 形成された階層構造とグループ

### 2.3. キーワードによる電子文書の検索

#### (1). 電子文書の検索

要求を受け取ったエージェントは与えられたキーワードを使い、自分の持つ電子文書を検索する。

#### (2). 電子文書内の I キーワード (Indicate キーワード) の抽出

I キーワードとは検索キーワードを持つ電子文書内から抽出したキーワードとして有効な語を示す。特定した電子文書から、I キーワードを抽出する

A proposal of information retrieval agent method considering amount of information in retrieval targets

<sup>†</sup>Yushin Kaneko, <sup>†</sup>Hisao KOIZUMI, <sup>‡</sup>Kazuya NAGAHASHI

Department of Computers and Systems Engineering, Tokyo

Denki University

方法を以下に示す。

1. 特定した電子文書内における単語を特定する。
2. 特定した単語の出現頻度を測定する。
3. 出現頻度(7割~9割)の高い単語を抽出する。
4. 3の単語群の中から名詞であるものを1キーワードとする。

#### 2.4. 検索結果の伝播

検索要求を取得したエージェントは検索結果を要求発信ユーザの解析エージェントに送信する。検索結果は取得したキーワードを含む電子文書、その電子文書の内容から取得したIキーワードである。検索結果の送信手順は子のエージェントが親のエージェントに取得したIキーワードを送り返す。親のエージェントは子のエージェントの検索結果として送信されたIキーワードを集計する。親のエージェントはリストに載っている子のエージェント数を規定数とし、集計数が規定数を満たしていない場合はさらに自分の親のエージェントへとIキーワードを返す。規定数を満たしたエージェントは検索要求ユーザへと検索結果を送信する。

#### 2.5. 検索結果解析

検索要求ユーザは受信した検索結果を集計し、与えたキーワードに対してどのようなIキーワードを持つ電子文書があるのかを特定する。エージェントは検索において最適なIキーワードを内包した電子文書を取得する。エージェントは集計した検索結果をもとにIキーワードを送信したエージェントのIPアドレスとIキーワードの分布を解析し、親と子をつなぐ一つのグループとしてデータに関連性があるかどうか検索結果から特定する。ユーザが要求したキーワードを含む電子文書を発見したエージェントのIPアドレスやそのグループ持つ情報保有力とその情報保有力の散らばりを情報として今後の検索に対して利用していく。

#### 2.6. エージェントの構成と機能

本研究では、ユーザの検索能力に関わらず最適な検索が可能なシステムとしてマルチエージェントを利用した。エージェントという自律的なプログラムを利用することによってユーザの語彙力によらない探索・検索・解析を行う。

##### (1) 構成

- 本方式は、4つのエージェントから構成される。
- キーワードをユーザから取得
- キーワードを伝達(過去の検索履歴がある場合は考慮する)
- 電子文書の検索, キーワードの取得
- Iキーワードの集計, 解析
- 各エージェント間の関係を図3に示す。

##### (2) 各エージェントの機能

- ・カテゴリエージェント  
カテゴリエージェントは検索要求ユーザの入力したキーワードの管理を行うエージェントである。

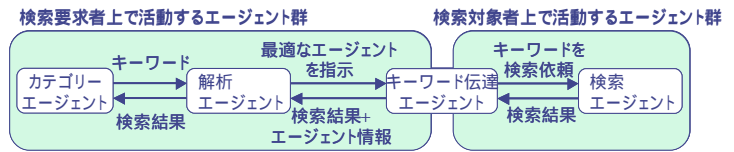


図3 エージェント間の関係と検索データの流れ

- ・解析エージェント  
解析エージェントは検索結果からどのようなIキーワードを含む電子文書をどのユーザが持っているのかを解析するエージェントである。
- ・キーワード伝達エージェント  
キーワードの伝達と検索の為に階層構造を作成するエージェントである。
- ・検索エージェント  
電子文書を検索し、電子文書からIキーワードを取得するエージェントである。

#### 3. 検索シミュレーションの構築

エージェントの動作、検索の流れ、情報の取得状況、検索結果の解析を行うシミュレーションプログラムを構築する。シミュレーションプログラムの開発はWindows2000上で、Java(JDK1.3)、XMLパーサとしてJDOMを用いて行った。

構築するシミュレーションプログラムのクラス図を図4に示す。

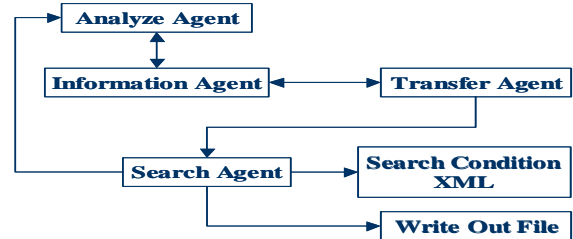


図4 シミュレーションプログラムのクラス図

各クラスは各エージェントの機能に対応している。

#### 4. まとめと今後の課題

サーバレス環境における情報の共有とその利用方法を提案した。シミュレーションを行いパラメータを変更し様々な環境においての動作を確認中である。

##### 参考文献

- [1] 長橋和哉, 髭白茂男, 北島聡史, 金子祐真, 塩澤秀和, 小泉寿男” 分散環境におけるサーバレスコミュニケーション方式と検索システム方式の提案”, Dicom2002 シンポジウム論文集 pp33-36, July, 2002
- [2] Hélène ARFAOUI, Kinji MORI” Autonomous Navigation Architecture for Load Balancing User Demands in Distributed Information Systems”, Vol.E84-B No.10 pp.2740-2748 2001/10