

電子メール宛先情報  
提供サービスシステムの開発

2N-5

楊 林\*

(\*筑波大学理工学研究科

海老原 義彦\*\*

\*\*筑波大学電子情報工学系)

1.はじめに

近年の計算機システムの発展にともない、ユーザはより良いサービス提供を要求している。一方、ネットワークに分散した計算機資源を効率良く利用でき、かつ各オブジェクトを一貫して管理できる分散OSの開発が必要となってきた。本研究の目標としては、電子メールシステムのサービスを充実するために、ユーザの宛先に対する問い合わせ機能を提供すること、かつその宛先属性情報を知識ベースとして管理する機能を提供することである。従来までのDB構成方法と違い、オブジェクトに向いて、分散環境に名前を管理するという特徴がある。ここでは、分散環境に適した電子メールシステムの宛先管理の実現方法と、知的マンマシンインタフェースを構成するために、入力文字系列と検索項目との柔軟性のある整合アルゴリズムについて述べる。

2.分散環境下のデータベース

分散環境下で、システムの透過性を守るために、システム全体に各オブジェクトが唯一のnameを持たなければならない。電子メールシステムには、互いに通信できるすべてのユーザが存在しているので、どのようにユーザの名前を管理するかは重要な課題になってくる。すべての名前を一つのデータベースに収集してユーザの問い合わせに答える方法はあるが、登録しているユーザの数が増えると、データベースはyellow pagesのように大きくなって、ユーザの要求にうまく対応できない。逆に、各ユーザの名前をホスト別に登録し処理するのは応答時間がかかると考えられる。ここでは、折衷的な案を提案した。つまりホスト別にユーザ名前のデータベースを構築し、リモートへの問い合わせを同報通信によって処理し、有用な情報をローカルデータベースに蓄積することである。

名前管理には、三つのレイヤの階層的なdirectoryを設けて管理する。global directory layer, regional directory layerとlocal directory layer. 対応してディレクトリサーバ、オブジェクトサーバとファイルサーバがある。各オブジェクトに完全な名前を用いる(図1参照)。ATENAというオブジェクトがregional layerに格納される。ATENAサーバごとにネットワーク上の各ホストの全名を設けて管理する。一旦ローカルデータベースに対する検索に失敗すると、同報通信によって他のホストへ調べに行く。システム構成を図2に示す。

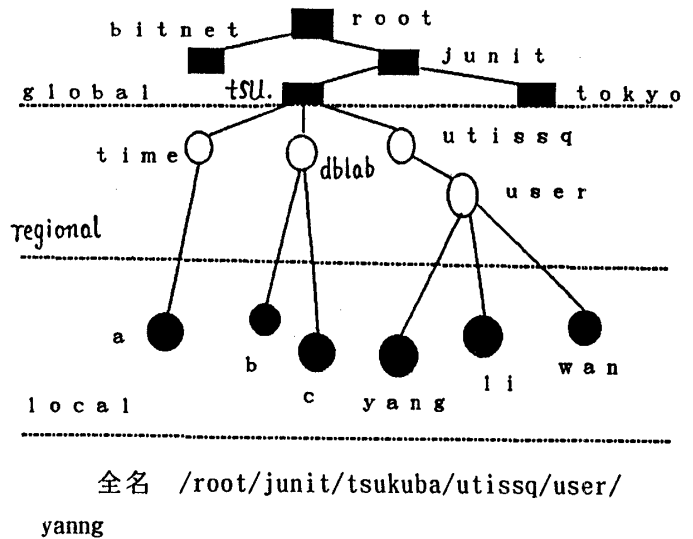


図1 分散環境下の名前管理

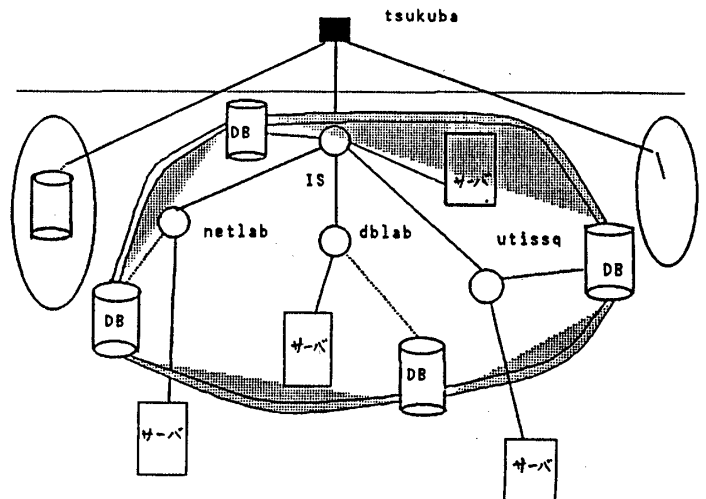


図2 system configuration

Name Guide Server in Electronic Mail Sys.

Lin Yang \*, Yoshihiko Ebihara\*\*

\*, \*\* Univ. of Tsukuba

### 3、電子メール宛先サービス

電子メールアドレスを知らないと、相手にメールを送り出すことはできない。また、相手のアドレスをよく覚えていないと、発信不能、あるいは誤り転送になる場合もある。本研究で開発している分散環境下の名前管理の一つのプロトタイプであるATENAシステムは、ユーザの問い合わせに対する検索を行い、該当相手の登録したネットワークアドレスを表示する。宛先問い合わせ処理の様子を図3に示す。

```
%atena
**** ATENA START ****
Wed Nov 1 19:00:25 jst 1989
name? k. 楊 林
所属? rikougaku
学年、役職? 1
----- REPLY -----
yanng@utissq. is. tsukuba. ac. jp
yang@gama. tsukuba. ac. jp

**** ATENA END ****
%
```

図3 ユーザインタフェース

ユーザはコマンド"atena"で問い合わせを行う。システム管理人はコマンド"atena\_db"でデータベースに対する初期化、定期的変更などが行う。

同姓同名などの検索では、ユーザの属性情報を基に会話処理でデータベースの検索を行う。一部分の情報はUNIXの/etc/passwd というfileから得られ、また、ほかの属性情報はユーザとの会話処理過程で得られる。つまりシステムは学習機能を持つ。(host\_name, uid, gid)は主キーである。実際には属性の組み合わせで検索することも多い。該当ユーザの数が複数個ある場合が考えられる。これらのユーザの宛先はすべて表示されて、ユーザとの会話処理で定まる。

### 4、柔軟な検索

宛先情報データベースに対する検索キーは英文字

と漢字の系列で、システムはユーザとの会話により検索を行う。人間同士のコミュニケーションであれば、多少の不正確な情報であっても知識処理により、柔軟に対処し、コミュニケーションが円滑に進められる。ある程度の柔軟性をシステムに与えるため、曖昧な入力キーにも対処できるような文字列間の距離を測定し、距離に基づいて検索を行う。ここで用いた検索アルゴリズムを改良型非線形伸縮アルゴリズムと呼ぶ。入力文字系列と検索項目の属性値の間の距離を図4のように定義する。

記号系列 $A=a_1a_2\dots a_n$  と  $B=b_1\dots b_n$  との間の距離は  $D(A, B)=d(m, n)$  と定義される。

$$d(i, j) = \min\{d(i-1, j-1)+p, d(i, j-1)+q, d(i-1, j)+r\} + \delta,$$

ただし、 $d(1, 1)=p$ 。

$\delta, p, q, r$ は下のルールによって定まる。

rule0:  $\delta = 2$ ;  $\delta = 0$  :  $(\text{abs}(i, j) > 2)$

rule1:  $a_i a_{i+1} \rightarrow b_j b_{j+1}$ :  $(a_i = b_j) \& (a_{i+1} = b_{j+1}) p = 0$

rule2:  $a_i \rightarrow b_j$ :  $(a_i = b_j) p = 0$ ;  $(a_i \neq b_j) p = 1$

rule3:  $a_i \rightarrow \varepsilon$ :  $(a_i = b_j) q = 0$ ;  $(a_i \neq b_j) q = 1$

rule4:  $\varepsilon \rightarrow b_j$ :  $(a_i = b_j) r = 0$ ;  $(a_i \neq b_j) r = 1$

図4 距離の定義

### 5、おわりに

本論文では、分散環境における統一した名前管理の方法を提案し、電子メール宛先情報を提供するプロトタイプシステムを開発した。現在、システムの分析と設計が終わり、電子メールシステムの宛先情報を管理する実験環境が整った。今後は、システムパフォーマンスを評価し、より有効なアルゴリズムを開発する予定である。

参考文献:

- [1]. 池田克夫: "知的インタフェース", 電子通信学会誌, Vol. 69, No. 11, PP. 1160-1166 (1985).
- [2]. David R. Cheriton, Timothy P. Mann, "A Decentralized Naming Facility", Report No. STAN-CS-86-1098.