

知識ベースシステムを用いた分散型メール配送システム MILD

上田 宏高[†] 門林 理恵子^{†,††} 萩野 浩明[†]
塚本 昌彦[†] 西尾 章治郎[†]

本稿では、分散型知識ベースシステムを用いたメール配送システムの構築について述べる。筆者らの研究グループでは、グループ通信をサポートする従来のシステムでは不完全であった送受信双方の選別要求を満たすことを目的として、推論機構によるグループの動的構成手法を用いたメール配送システム MILD (MaIL Distribution system) を開発し、2年間にわたり運用を進めてきた。MILD は指定された条件を満たすユーザからなるグループを知識ベースシステムを用いて動的に構成し、電子メールを配送するシステムである。しかし、運用を進めるうちに、知識管理の困難化、配送遅延の増大等 MILD が集中管理型システムであることに起因する問題点が明らかになってきた。今回、筆者らはこれらの問題点を解決するため、MILD の分散化を行った。システムおよび知識を分散化することによって、MILD は、広範囲にわたるユーザが柔軟なグループ通信を容易に利用できるシステムとなった。

A Decentralized Mail Distribution System Using Knowledge-base Systems

HIROTAKA UEDA,[†] RIEKO KADOBAYASHI,^{†,††} HIROAKI HAGINO,[†]
MASAHIKO TSUKAMOTO[†] and SHOJIRO NISHIO[†]

In this paper, we present the design and implementation of a decentralized mail distribution system using knowledge-base systems. In these two years, we have developed and maintained a system for distributing mails using a reasoning-based dynamic group construction method, called MILD (MaIL Distribution system). However, in a practical use, various problems have identified to be serious. One of them is the difficulty in managing large size of knowledge-base, and another is the increase of the delay time for mail distribution caused by the long reasoning time. We consider that the centralized management of MILD is the main reason for such problems. Therefore, we designed and implemented a new version of MILD, which we call MILD3, where servers and knowledge-bases are decentralized onto several hosts. In MILD3, each knowledge-base is managed independently by each organization, and it can be used in an integrated way for mail distribution. Furthermore, because of decrease of the reasoning time, it becomes possible to distribute mails in a shorter time than the previous MILDs. As the result, the new version of the MILD can provide a more flexible group communication for a large number of its users.

1. はじめに

コンピュータネットワークは、通信メディアを用いた特定および不特定の集団におけるコミュニケーション、いわゆるグループ通信を提供できるという点で、従来の電話や手紙、TVや新聞といったメディアと比べ

て優れた利便性を持っている。コンピュータは個人の作業を支援するだけでなく、複数の人間の協調作業を支援する目的でも用いられ、しかもそのような共同作業は、組織的に行われるフォーマルなものだけではなく、個人で行われるインフォーマルなコミュニケーションまでもが対象となっている。また、グループ通信におけるグループの多様化、増加が進んでおり、近年のコンピュータネットワークの普及にともなうユーザ数および情報量の急激な増加によって、コミュニケーションを行う相手や、やりとりする情報の選別が困難になっている。このように、ネットワーク上では多数の人による多様なコミュニケーションを支援する必要

[†] 大阪大学大学院工学研究科情報システム工学専攻
Department of Information Systems Engineering,
Graduate School of Engineering, Osaka University

^{††} 株式会社エイ・ティ・アール知能映像通信研究所
ATR Media Integration & Communications Research
Labs.

性が増してきており、これまで様々な研究がなされている^{2),4),7)}。

一般にグループ通信においては、2つの基本的な要求が考えられる。1つはネットワーク上の膨大な情報から自分に必要なものだけを得たいという、受信側から見た情報選別要求であり、もう1つは、メッセージを分配するグループを目的や状況に応じて柔軟に構成したいとする、送信側から見た配送先の選別要求である。

従来のコンピュータネットワーク上のグループ通信環境をサポートする代表的なシステムであるメーリングリストは、それぞれがある一定のグループと対応しており、それぞれに対応したメールアドレスを持つ。メーリングリスト宛に出された電子メールはそのメーリングリストの参加者全員に配送される。ユーザは参加するメーリングリストを選択することによって自分の得たい情報を得ることができるが、参加するメーリングリストの単位でしか情報を選別することができず、受信側から見た情報選別要求は完全には満たされていない。

一方、ユーザはメールを配送するメーリングリストを選択することによって自分の送りたい相手にメールを配送することができる。しかし、メーリングリストのグループは固定的であり、ある2つのグループにまたがるメンバにだけメールを配送したり、あるグループから任意のメンバを除いたグループに対してメール配送を行ったりすることはできない。このような場合には、一般に、新たにメーリングリストを作成するしか方法がない。このように、メーリングリストを用いたシステムではコミュニケーションの相手の選別の柔軟性が欠け、送信側から見た情報選別もまた不完全である。

情報の選別については情報フィルタリングシステムに関する研究⁶⁾が盛んに行われているが、これまでの研究は主に受信側の情報選別要求を満たすことが目的であり、送信側の配送先選別要求については十分に考慮されてこなかった。一般に、グループ通信においては、送信側においても、コミュニケーションの対象をまったく限定せずに情報を発信したい場合は稀であり、むしろ、配送先を限定したい場合が多いので、このことを考慮しないのは不十分である。また、送信側からの情報配送先の選別はトラフィックの削減という点において、受信側での情報選別に対して明らかな優位性を持つ。グループ通信を行ううえでこの2つの選別を考慮することが重要である。

このような観点に基づき、筆者らの研究グループでは、これまでにグループ通信において受信側および送信側の立場から目的や状況に応じた柔軟な情報の選別を可能とすることを目的として、電子メールを用い

たメッセージ分配システム MILD (MaIL Distribution system)⁹⁾および MILD2³⁾を開発し、2年以上にわたって研究室で運用してきた。本論文の以下の部分では両システムをまとめて MILD とよぶ。MILD ではユーザ管理に推論機構を持つ知識ベースシステムを用いる。ここでいう知識ベースシステムとは、推論機構とデータベースを組み合わせたものである。知識ベースシステムの提供する推論機構により、明示的に記述していないが実際には成立しているような関係を演繹により導くことができるため、宛先の指定を直観的かつ簡潔に記述できる。さらに、推論結果に対する演算を可能にすることにより、きめ細かなコミュニケーション対象の指定が可能になる。また、ユーザは自分に関する知識を記述することで、グループ生成に自分の意図を反映させることができる。その結果、MILD では送受信双方の意図を反映し、柔軟なグループ通信を提供することができる。

しかし、MILD の運用を進めるに従い、MILD が集中管理システムであることに起因する様々な問題点が明らかになってきた。現在運用中の MILD の知識ベースには約 50 名のユーザに関する約 500 の知識が存在する。しかし、現行以上の知識の追加は、推論に要する遅延時間とサーバに集中する負荷を考慮すると、非実用的である。また知識を集中管理することは知識ベースの管理者に多大な負担をかける。

そこで筆者らは、MILD サーバと知識ベースを分散化することにより、負荷の軽減と知識管理の効率化を実現する MILD3 の設計および実装を行った。以下、2章で従来のメッセージ分配システム MILD について概説し、3章では MILD の分散化について論ずる。4章では MILD3 に対する考察を行い、5章では本稿のまとめと今後の研究課題について述べる。

2. MILD の概要

MILD は、ユーザが指定した条件を満たすメンバからなるグループを知識ベースシステムを用いて動的に構成し、そのグループに対して電子メールを配送するシステムである。本章では MILD の概略について述べる。

2.1 知識ベースシステム DOT

MILD では知識ベースシステム DOT (Deductive Object-oriented Term representation)¹⁰⁾を用いてユーザに関する情報や、一般的な概念に関する情報の管理を行っている。DOT とは、ドット記法と IS-A 関係を用いて知識を表現し、ドット記法によって拡張された IS-A 関係の順序性と継承関係をもとに推論を行

う知識ベースシステムである。DOT を採用することにより、簡潔な表現を用いて様々な観点からの情報を統一的なモデルで扱うことが可能になる。ここで DOT の概要を述べる。

DOT では、知識は DOT 式間の IS-A 関係を用いて表現される。DOT 式、および IS-A 関係は次の形式で表現される。

- DOT 式: $A.L_1 \dots L_n$ ($n \geq 0$) を DOT 式とよび、 A をオブジェクト名、 L_i ($1 \leq i \leq n$) をラベルとよぶ。
- IS-A 関係: R, S を DOT 式とすると、 R IS-A S という関係を IS-A 関係とよぶ。

DOT では、IS-A 関係で表現される知識に対し、IS-A 関係の一方を未知数とした問合せを行うことができる。以下に MILD で利用している問合せを示す。ここで、 R を DOT 式、 L_1, \dots, L_n ($n \geq 1$) をラベルとする。

- a. 何(だれ)が R であるか?
 - b. 何(だれ)の L_1 の \dots の L_n が R であるか?
- 問合せの解は、以下の推論規則に基づいて導出される。ここで、 R, S, T を DOT 式、 L をラベルとする。
- i. X IS-A X
 - ii. X IS-A Y, Y IS-A Z ならば X IS-A Z
 - iii. X IS-A Y ならば $X.L$ IS-A $Y.L$

問合せの解は、正規表現で表され、それを求めるアルゴリズムが示されている¹⁰⁾。そのため、複数の問合せの解に対し、和や積等の集合演算を行うことも可能である。

2.2 メール配送システム MILD

知識は図 1 に示すように DOT 式間の IS-A 関係を用いて表現される。各ユーザは、興味のある話題に関するキーワードや、自分が所属するグループ、あるいは自分の住所や出身等、メールの配送先を決める要素となりうる様々な情報を MILD の持つ知識ベースに登録しておく。MILD 管理者は、「大阪は関西である」等の一般的な概念に関する知識を登録するとともに各知識の整合性をチェックする。

MILD は推論機構およびアドレス検索機構から構成される。ユーザは、図 2 に示すように Subject 欄に「[]」で括って宛先を決定する条件式を記述したメールを MILD サーバに送信する。MILD サーバは、メールを受信すると条件式を取り出して、推論機構に渡す。推論機構は、知識ベースへの問合せ処理を行い、問合せに対する答をアドレス検索機構に渡す。アドレス検索機構では、入力された演算結果と表 1 に示すアドレステーブルの DOT 式を比較し、推論結果に DOT 式

春本 IS-A 西尾研.助手
 春本 IS-A DB 移動チーム
 春本.趣味 IS-A ゴルフ
 かに IS-A 西尾研.D2
 かに IS-A 透明人間グループ
 かに IS-A 透明人間グループ.リーダー
 かに.趣味 IS-A バドミントン
 相坂 IS-A 西尾研.M1
 相坂 IS-A DB 圧縮チーム
 相坂.趣味 IS-A 読書
 相坂.タイプ IS-A 酒井美紀
 秋山 IS-A 西尾研.M1
 秋山 IS-A DB 移動チーム
 秋山.友人 IS-A 相坂
 秋山.趣味 IS-A 水泳
 西尾研.D2 IS-A 西尾研.学生
 西尾研.M1 IS-A 西尾研.学生
 西尾研.学生 IS-A 西尾研
 西尾研.助手 IS-A 西尾研
 DB 移動チーム IS-A DB グループ
 DB 圧縮チーム IS-A DB グループ
 水泳 IS-A スポーツ
 バドミントン IS-A スポーツ
 ゴルフ IS-A スポーツ

図 1 MILD の知識例

Fig. 1 An example of user knowledge used in MILD.

To: mildserver@nishirolab.osaka-u.ac.jp
 Subject: ソフトボール大会のお知らせ [西尾研の学生かつ趣味がスポーツ]
 ソムヌックです。
 今週の土曜日に情シスのソフトボール大会があります。

図 2 MILD を利用したメールの例

Fig. 2 An example of MILD mail.

表 1 アドレステーブルの例

Table 1 An example of user address table.

DOT 式	電子メールアドレス
春本	harumoto@nishirolab.ac.jp
かに	kani@nishirolab.ac.jp
相坂	aisaka@nishirolab.ac.jp
秋山	akiyan@nishirolab.ac.jp

が含まれるとき、DOT 式に対応するメールアドレスを結果として返す。MILD は受けとった答からアドレスの集合を生成し、それを宛先としてメールを送信する。図 2 に示したメールは図 1 の知識により、「かに」および「秋山」に配送されることになる。

このようにユーザの宛先指定から動的にグループを生成することによって、送信側の配送先選別要求を実現できる。また、各ユーザに関する知識を知識ベースに蓄え、それをグループ生成に反映させることにより受信側の情報選別要求も満たすといえる。また、筆者

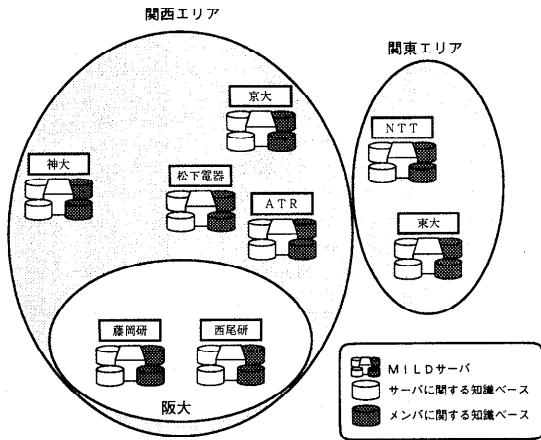


図3 MILD3サーバの配置例

Fig. 3 An example of MILD3 server arrangement.

らは MILD に対してメールの分類検索機構を付加しており¹⁾, MILD は関連研究^{5),7)}に比べて送受信双方の選別要求を満たす柔軟なグループ通信を提供できる。

3. MILD の分散化

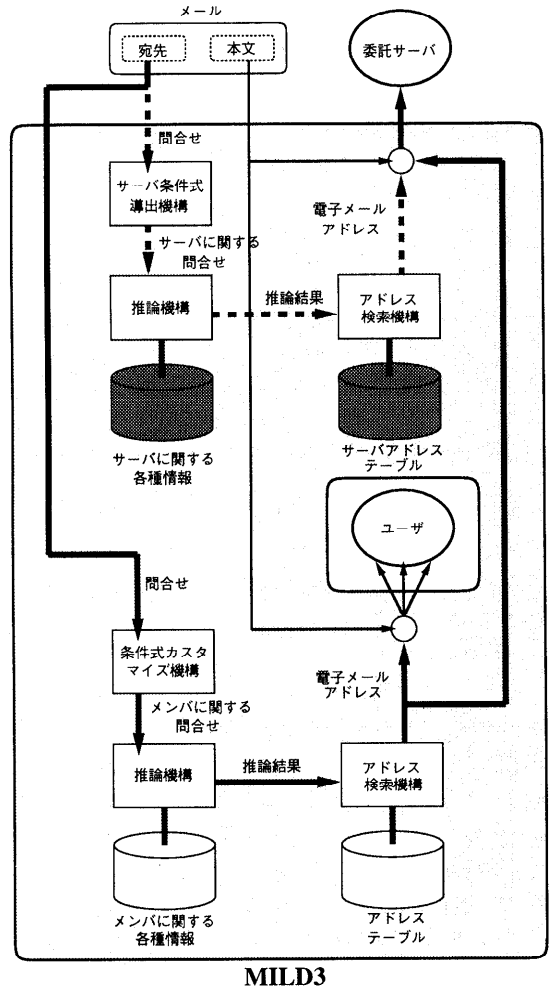
MILD が集中管理型システムであることによる前述の問題を解決するため、筆者らは MILD の分散化を行った。本章では、MILD を分散化した MILD3 について概略を述べる。

3.1 サーバ構成

MILD サーバは知識ベースと組になって、各組織、部所、グループ等 1 単位につき原則として 1 つ置くことを前提とする。MILD を分散化した MILD3 サーバの配置の一例を図 3 に示す。ここでは、簡単のために MILD サーバの数を減らしている。本来は各大学の研究室、各企業の部署程度の規模で MILD サーバを配置することを想定している。

各 MILD3 サーバの構成を図 4 に示す。各 MILD サーバはすべて同等の構成である。図中の点線の矢印は、最初にユーザからメールをもらったサーバが、分散配送処理を行うサーバを決定する過程を示す。その一連の処理には、サーバ条件式導出機構、分散された MILD サーバに関する知識ベース、および、そのアドレステーブルが必要である。また、条件式カスタマイズ機構により個々のサーバに適した条件式を導出する。各サーバで導出された条件式をもとにユーザへのメール配送が行われる。

メンバに関する知識ベースは図 1 に示した従来の MILD のものとまったく同一である。各 MILD サーバは自サーバに所属するメンバに関する情報を知識ベースに管理する。そのため、各サーバのメンバに関する



MILD3

図4 MILD3の構成

Fig. 4 The architecture of MILD3.

知識ベースは各組織の管理者が独立して行うことができ、従来の集中管理に比べ、管理者の負担は少ない。メンバに関する知識ベースは各 MILD サーバで自由に構築できるが、個人名、役職名、組織名等外部から参照される可能性のある知識については必ず記述する。将来的には一般的な知識については、各 MILD サーバで統一することを検討している。図 3 のサーバ配置情報は図 5 のように各 MILD サーバに保持される。

図 5 に示すサーバに関する知識^{*}は、MILD サーバ

* たとえば「京大 IS-A 関西」という知識は厳密には IS-A 関係としては適切でなく、「京大 IS-A 関西. 大学」等とするのが本来正しい。ここでの「関西」は「関西にあるもの」を表すと解釈する必要がある。「京大 IS-A 関西. 大学」とした場合には「関東以外の大学」という条件式では京大は導出されないため、今回のシステムでは知識としては不適切となる。ただ、このような表記は使用者に誤解を招くおそれがあるため、自然な知識表現が行えるよう今後の見直しが必要である。

西尾研 IS-A 情報システム工学科
 藤岡研 IS-A 情報システム工学科
 情報システム工学科 IS-A 阪大
 阪大 IS-A 関西
 阪大 IS-A 大学
 京大 IS-A 関西
 京大 IS-A 大学
 ……
 ATR IS-A 関西
 ATR IS-A 企業
 NTT IS-A 関東
 NTT IS-A 企業
 関西 IS-A 日本
 関東 IS-A 日本

図5 サーバ知識例

Fig. 5 An example of knowledge about MILD servers.

名と、各 MILD サーバが属するドメイン名からなる。MILD サーバ名は「西尾研」、「東大」等、表 2 のサーバアドレステーブルに記載されている名前であり、その他の「関西」、「大学」等は、MILD サーバが属するドメイン名である。

表 2 に示すアドレステーブルには、すべての MILD サーバのアドレスを優先度順に記述する。優先度は各 MILD サーバで独自に設定でき、使用頻度の高い順や、通信コストが低い順に並べることにより、サーバ間のメールのフォワードによるロスが少なくなり、メール配送が高速になる。自サーバの優先度は一番高くする。表 2 の場合、自サーバ（西尾研）が一番優先度が高く、次に阪大内部の研究室の優先度が高く設定されている。

最初にメールを受けとった MILD サーバが、メールを配送すべき組織を決定するため、関連サーバ間でサーバに関する知識の完全な整合性がなくてもメール配送は行われる。ただしユーザの誤解を招かないように、サーバに関する知識中で使用するドメイン名等は関連サーバで統一しておくことが望ましい。新たにサーバを追加したり、サーバ知識を更新するときはコマンドメールを用いていっせいに更新するため、一貫性は保持される。ただし、将来的にサーバ数が増加し、サーバ知識が膨大になった場合には、一貫性の保持が困難になることが予想される。大規模環境における一貫性の保持等の問題点の抽出、およびその解決は今後の課題の 1 つである。

3.2 分散配送処理

MILD3 を利用したメールの一例を図 6 に示す。メールの書き方は従来の MILD と同一であり、MILD が分散されていることをユーザが意識することはない。ユーザは自分の属する組織の MILD サーバに条件式

表 2 サーバアドレステーブル例

Table 2 An example of address table of MILD servers.

DOT 式	MILD サーバのアドレス
西尾研	mildserver@nishiolab.osaka-u.ac.jp
藤岡研	mildserver@fujiokalab.osaka-u.ac.jp
京大	mildserver@hogehogelab.kyoto-u.ac.jp
神大	mildserver@nyoronyorolab.koube-u.ac.jp
ATR	mildserver@atr.co.jp
松下電器	mildserver@panasonic.co.jp
東大	mildserver@foobarlab.tokyo-u.ac.jp
NTT	mildserver@ntt.co.jp

To: mildserver@nishiolab.osaka-u.ac.jp
 Subject: お知らせ [ATR の 門林 さんと
 関西の大学の学生かつ趣味が映画観賞]
 ソムヌック@阪大です。
 来る 6 月 1 日にタイの映画の上映会を開催します。

図 6 MILD3 を利用したメールの例
 Fig. 6 An example of MILD3 mail.

を書いたメールを出す。この例の場合、西尾研の学生が西尾研の MILD サーバにメールを出している。

最初にユーザからのメールを直接受けとった MILD サーバをローカルサーバと呼ぶ。この例では西尾研の MILD サーバがローカルサーバであるが、ローカルサーバは条件式カスタマイズ機構を通して、「ATR の門林さんと関西の大学の学生かつ趣味が映画観賞」という条件式から、サーバ決定に関与しない条件式の削除等を行い、「ATR と関西かつ大学」というサーバ条件式を導出する。次に推論機構により図 5 に示したサーバに関する知識を用いて、ATR と関西の大学の MILD サーバ（委託サーバとよぶ）を導出する。メールはこの委託サーバによって配送される。ここでは、西尾研、藤岡研、京大、神大、ATR の MILD サーバが委託サーバとなる。

また、ユーザが意図しない広範囲のユーザに対してメールが配送されるのを防ぐために、「白井さん」等のサーバの指定を含まない条件式を持つメールは各 MILD サーバで任意に指定できるデフォルトの委託サーバにのみ配送される。デフォルトの委託サーバをその MILD サーバの属する組織内等にする事で、ユーザの意図にそった配送が可能になる。ちなみに、日本中の白井さんに配送したい場合には「日本の白井さん」というように明確に配送範囲を指定する必要がある。

3.2.1 通常モードによる配送処理

図 7 に示すように、通常、メールはすべての委託

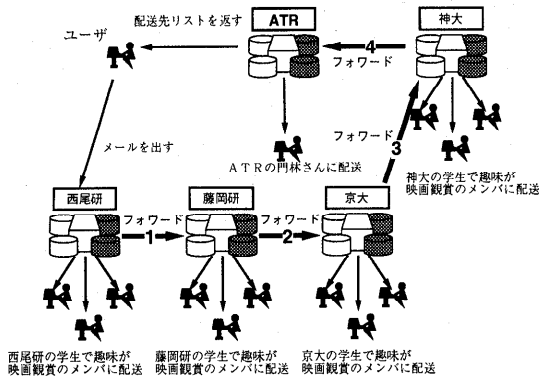


図7 通常モードのメール配送

Fig. 7 An example of ordinary mode mail delivery.

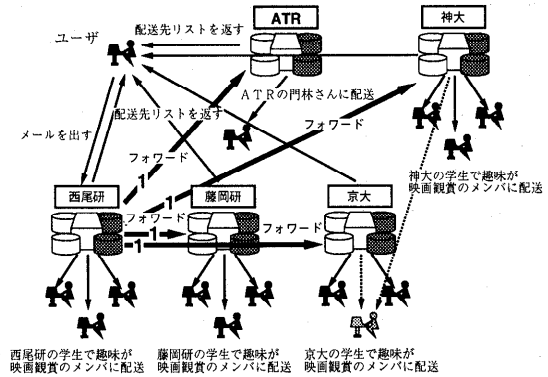


図8 速達モードのメール配送

Fig. 8 An example of express mode mail delivery.

サーバを優先度順にまわすことによって、条件を満たすすべてのユーザに配送される。メールはサーバ間を图中的矢印についた番号順にフォワードされる。各委託サーバは条件式カスタマイズ機構により、条件式を各サーバのメンバに関する知識に適合するようにカスタマイズした後、推論を行い、配送先を決定する。たとえば、西尾研のMILDサーバは「ATRの門林さんと関西の大学の学生かつ趣味が映画観賞」という条件式から「西尾研の学生かつ趣味が映画観賞」という条件式を生成し、この条件式を満たすユーザに対してメールを配送する。同様にATRのMILDサーバは「ATRの門林さん」という条件式を満たすユーザに対しメールを配送する。条件のカスタマイズを行うことにより、各MILDサーバのユーザに関する知識に手を加えることなしに知識ベースを分散化することが可能になった。また、サーバ間のフォワードの際、配送済みメンバのアドレスをヘッダに付加することで、同一メールの同一ユーザへの重複配送を回避する。送信者は最後の委託サーバから1通にまとめられた配送先リストを受けとる。

3.2.2 速達モードによる配送処理

すべての委託サーバを順にまわると、サーバ数が多い場合、配送遅延が大きくなる。そこでMILD3は図8に示す速達モードを提供する。速達モードでは上述の重複配送回避処理を行わず、すべての委託サーバに图中的矢印のようにメールをいっせいにフォワードし、配送処理を並列化することにより、高速なメール配送を提供する。そのかわり、重複して同一のメールを受けとるメンバが発生する可能性がある。

たとえば、京大生であり神大の特別研究生であるようなメンバが存在するとそのメンバは京大と神大の両方のMILDサーバから同じメールを受けとることに

なる(図8で点線で示した矢印)。また、送信者はすべての委託サーバから個別に配送先リストを返されることになる。送信者は、配送するメールの性格によって2つのモードを使い分けることができる。

速達モードを使用すると、関連サーバに一斉配送するローカルサーバに負荷が集中することが懸念されるが、一般的なテキストが主体のメールの配送を想定すると、関連サーバ数が数百程度ならば、システムに過大な負荷をかけることなく一斉配送することが可能である。実際、メーリングリスト等で、この程度の一斉配送は珍しくない。

3.2.3 分散配送処理が不可能な場合

「京大の学生かつNTTの研修生」といった複数のサーバの情報を統合しないと配送先が決定できない場合もある。そうした条件式の場合は各委託サーバはアドレスの導出のみを行い、実際の配送を行わず、そのアドレスリストをローカルサーバに返す。ローカルサーバは、それらのアドレスリストに対して集合演算を行い、配送先を決定し、配送を行う。西尾研の学生が「京大の学生かつNTTの研修生」にメールを送った場合を図9に示す。图中的各矢印の数字は、メールの流れを表す。メールを受けとった西尾研のMILDサーバは京大とNTTのMILDサーバに対し、問合せを行う。京大のMILDサーバは「京大の学生」のアドレスリストをローカルサーバに返し、NTTのMILDサーバは「NTTの研修生」のアドレスリストをローカルサーバに返す。ローカルサーバは返ってきたアドレスリストに対して積演算を行い、「京大の学生かつNTTの研修生」のアドレスを導出し、メールを配送する。

3.2.4 VERIFYモード

MILD3は、配送を行わず、配送先リストだけを送

信者に返す VERIFY モードを提供する。ユーザはこの機能を用いることによって、導出結果を確認でき、ユーザの意図と一致しない場合には、導出結果を参考に条件式を変更する、あるいは、導出結果を編集して To 欄に記述する方法で利用できる。

3.3 実装

MILD3 の実装は SunOS4.1.3 上で、Perl および DOT-Shell⁸⁾を用いて行った。MILD3 ではメッセージの授受は sendmail を介して行う。そのため、ファイアウォールを越えて本システムを運用することが可能である。

4. 考察

本章では MILD サーバおよび知識ベースの分散化により得られた成果をまとめ、システムに関する考察を行う。

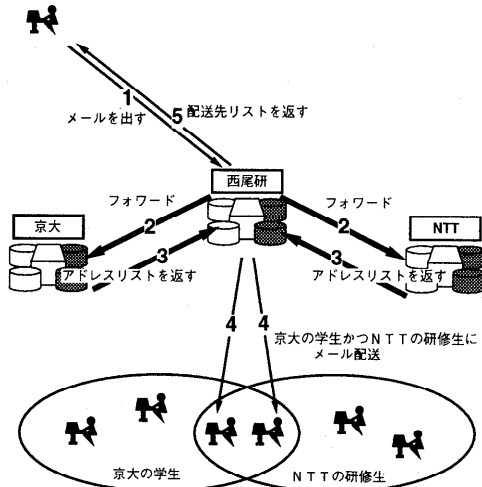


図9 分散配送処理が不可能な場合のメール処理例

Fig. 9 A mail processing example of the cases when the decentralized processing is unavailable.

4.1 処理速度の向上について

筆者らの研究室において、従来の MILD は知識ベースに約 50 名分のアドレスとそれに関する約 500 の知識を持っていた。そのメンバを 3 つのグループに分け、アドレスと知識を 3 つのサーバに分割したところ、処理速度の向上が見られた。図 10 にその結果を示す。ここでは、VERIFY モードで MILD にメールを出してから、配送先リストが返ってくるまでの時間を処理時間として測定した。従来の MILD は 1 つのサーバで集中処理を行う。それに対し、MILD3 は 3 つのサーバで分散して処理を行う。黒い部分は全体の処理時間に占める推論時間を表しており、MILD3 の場合は 3 つのサーバの推論時間の合計を表している。推論時間以外の部分はメールのフォワードに要する時間である。

筆者らの研究室では、これまで、研究室のメンバ全体を表す「西尾研」という条件式が最も多く使われている。また「西尾研」という条件式は最も継承が多く、推論に時間のかかる条件式である。その条件式において、MILD3 は 3 つのサーバをフォワードするロスタイムがあるにもかかわらず、従来の MILD に比べて半分以下の時間で処理を完了している。推論時間だけを見れば、MILD3 の 3 つのサーバの推論時間の合計は従来の MILD の推論時間の 7 分の 1 程度である。このことから知識量が推論時間に大きく影響することが分かる。また、「春本さん」等のような個人名だけからなる条件式は、結果的に推論の必要がないものである。従来の MILD のほうが有利であるものと予測していたが、両方の処理時間はほとんど変わらないことが確認された。

ただし、今回の測定では分散された MILD サーバ間は LAN で結ばれ、メールのフォワードに時間がほとんどかからない状況であることも考慮しなければならない。各 MILD サーバ間の帯域が狭い場合、MILD3 における各サーバ間のフォワードに要する時間が大きく

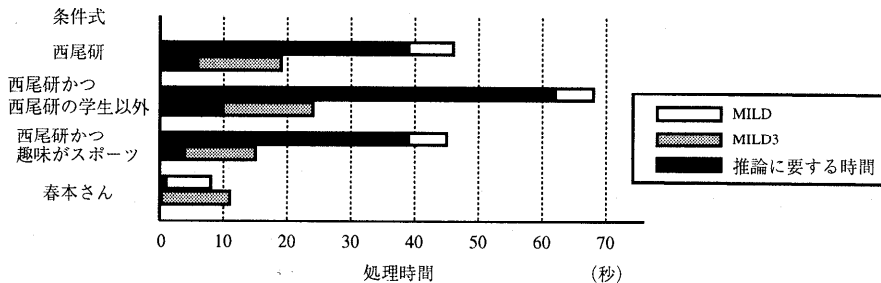


図 10 従来の MILD と MILD3 の処理時間比較

Fig. 10 Comparison of processing time of MILD and MILD3.

なり、全体の配送遅延が大きくなることが予想される。しかし、それでもトータルの時間は従来の MILD と同程度、あるいはより速くなることが期待できる。なぜなら、全体の知識量が増えれば増えるほど従来の MILD の推論に要する遅延時間は大きくなるが、MILD3 においては、サーバを分割することによって、推論に要する遅延を小さくできるためである。さらに、MILD3 においては、組織内の配送は組織内のサーバでのみ行われるため、広域を 1 つのサーバで集中管理する場合よりも高速な配送が行われる。また、広範囲にわたる配送においても、MILD3 の速達モードを使用すれば、サーバ間のフォワードに要する遅延が少なくなり、さらに高速な配送を行える。

知識量の増加により急激に推論時間が増大することは、重要な問題である。現在の DOT の実装ではデータ構造や探索アルゴリズムが単純であるため、データ構造の見直しやハッシュ探索の導入といった最適化により、さらに処理時間を減少することが可能である。しかし、ここに示したように、知識ベースの分割により、推論時間を大きく減少させることが可能であるので、知識に何らかのインデックスを付与することも有効であると考えられる。

4.2 MILD の利点

通常のメーリングリストの管理でも、規模が大きくなる場合等には、アドレスのリストには各組織やグループ単位の代表アドレスのみを登録しておいて、各組織内で、その代表アドレスに個別のメンバの登録をする管理方式を採用して、管理コストの分散や、トラフィックの軽減を図っていることが多い。今回 MILD においても同様の方法を用いて、負荷の分散を図ったが、既存のシステムや従来の MILD に比べて次にあげるような利点を持つ。

4.2.1 従来から持つ MILD の優位性

メーリングリストにおいては、グループを中心としてとらえているため、グループとメンバの関係が固定的であり、グループ構成における柔軟性に欠ける。一方、MILD は個人やその所属等の情報を知識ベースシステムを用いて管理し、電子メールの宛先のグループは複合的な条件を用いて推論を行うことにより動的に構成されるため、コミュニケーションの相手をきめ細かく指定できる。また、推論機構として DOT を採用しているため、知識表現が簡潔で理解しやすい形式であるため、管理に必要な情報を知識として表現したり、記述された知識から内容を読み取ることが容易である。これによって、知識の追加や変更等が容易に行える。

4.2.2 各組織ごとの独立した知識管理

知識ベースの分散化により、知識管理が各組織で独立して行えるようになり、管理者の負担が減少した。同時に、サーバにかかる負荷の集中も回避できた。また、各 MILD サーバで条件式のカスタマイズを行うことによって、サーバを分散化する際にメンバに関する知識ベースに手を加える必要がなく、分散されていることを意識せずに知識を記述できる。

4.2.3 広範囲にわたる直観的な宛先指定

電子メールをある個人に送りたい場合、通常、何らかの方法でその人のメールアドレスを調べる必要がある。しかし、MILD では宛先が直観的に記述できるため、個人の属する組織名さえ分かれば配送が可能である。たとえば、「NTT データの技術開発本部の白井さん」とすることによって、ユーザは白井さんのメールアドレスを知らなくても、白井さんに対してメールを送ることができる。電子メールアドレスは一般に人間にとって分かりにくいものであるが、MILD を用いて各組織が自組織の MILD サーバに自組織のメンバに関する知識を管理しておくことで、電子メールの宛先指定に関する利便性を高めることができる。

MILD では個人だけでなく、ある特徴を持ったグループを動的に構成し、コミュニケーションの対象をきめ細かく指定できる。分散化により広範囲にわたる組織に属するユーザに対して、膨大な知識を統合し配送先を決定することによって、柔軟なグループ通信を提供することが可能になった。たとえば、「阪大以外の関西の大学の学生かつ情報処理学会の会員」や「(京大と阪大)の教授かつ趣味がゴルフ」といった広範囲の組織にわたる柔軟な宛先指定が可能である。しかもユーザは、MILD サーバが分散されていることを意識せずに、従来と同様の直観的な記法で宛先を指定できる。

4.2.4 効率的な配送処理

サーバの分散化を行っても、従来どおり、ローカルな配送はローカルなサーバ内でのみ行われるため、効率的な配送が行える。一方、広範囲にわたる配送においては、複数の委託サーバの連係によって配送処理がなされるため、各サーバにかかる負荷は小さく、実用的な時間内で配送が行われる。

4.3 分散化による弊害について

本研究の目的として各 MILD サーバで独立して知識を管理しても全体として配送が適切に行われるような分散化を行いたいということがあげられる。そのため、各 MILD サーバの扱う知識には手を加えず、配送条件式を各 MILD サーバの知識に合わせてカスタマ

イズすることで分散しておかれた知識を統合する工夫を行った。このことにより、各 MILD サーバには非常に大きな自由度が与えられたが、知識の分散化によって様々な弊害が生じることが懸念される。

まず第一にあげられるのは知識の一貫性や矛盾に關する問題である。分散されたユーザに関する知識は各 MILD サーバに固有なものであり、組織ごとにある程度独立して管理できる。一方、抽象的な概念の表現は運用の中で統一を図っていく必要がある。そのため各 MILD サーバ管理者が自サーバのユーザに関する知識に矛盾や誤りがないか、抽象的な概念の表現は指針に沿っているか管理することで、全体としての動作保証が得られる。ある程度の語句の相違は知識による吸収が可能であり、同意語辞書を導入することによって整合性の問題は軽減できる。また、MILD が推論機構として採用している DOT は、IS-A 関係が矛盾したり、解の集合が無限になったりしても、必ず有限の時間で有限の表現の解が得られることが保証されている¹⁰⁾。そのため、各 MILD サーバの管理情報の内容や宛先の指定がどのようなものであっても、システムに致命的な欠陥が生じることはない。

次にあげられるのは全体の知識の把握が困難になったことである。各 MILD サーバの有する知識は Web 上のホームページ等で確認することができるが、関連サーバの知識をすべて確認することは現実的に困難である。このため、たとえば、「ATR の門林さん」というメールを出す場合を考えてみても、ATR のサーバがこの宛先に該当する知識を有しているという保証はない。こういったメールの配送を保証するためには、他のサーバから使用されるかもしれない知識（名前、所属、役職等）に関しては、最低限記述するという指針が必要である。また、将来的には、プライバシーにかかわるようなローカルでのみ参照される知識と、外部から参照されてもかまわない知識に分割して、プライバシーの保護を図ることを検討している。

また、配送洩れがないか配送前に確認することが、非常に困難になることが懸念される。配送先が広範囲に広がれば広がるほど VERIFY モードの返答が返ってくるまで長い時間がかかることが想定されるからである。ただし、配送先が確定していて、確実に配送されなければならないメールの配送は通常メーリングリストを用いて行われると想定している。これは、現時点では MILD においては、広範囲に分散された知識の把握が困難であり、VERIFY モードを用いなければ、配送先の確認ができないという理由による。もちろん、全体の知識を把握しやすい環境下においては MILD

はメーリングリストにはできないきめ細かく正確な配送指定が可能であり、配送洩れも容易に防ぐことができる。

MILD は閉じられた環境においては非常に有効であることが過去 2 年間にわたる試験運用から確認されている。分散化によって、MILD はより広範囲のユーザに対応できるようになったが、基本的にはグループウェアの延長上にあるものである。システムの構成上、知らない人の知識を用いてメールを配送することは非常に困難であり、不特定のメンバに対してメールを配送する場合にはまだ多くの制限が存在する。筆者らはある程度限られた現実的な範囲内での知識の運用を想定しており、そのような環境下においては MILD は効果的に利用できる。

5. ま と め

本論文では知識ベースを用いたメール配送システム MILD が単一のサーバからなる集中管理システムであったことに起因する問題点を解決するため、サーバおよび知識ベースの分散化を行った。これによって、知識管理が各組織ごとに行え、知識管理が容易なシステムになったといえる。また、それらの膨大な知識を統合してメールを配送することが可能となった。さらに、その膨大な知識を扱う場合でもサーバおよび知識ベースを分割することにより、サーバにかかる負荷を分散し、高速に推論が行え、メールを配送できるようになった。今後の課題としては、前述したような各サーバ知識の一貫性の確保、プライバシーの保護等のほかに、知識のインデックス化、サーバの階層化等の手法による推論の高速化、ヘッダの暗号化によるアドレスリストの漏洩防止等のセキュリティ対策、VERIFY モードに代わる配送先確認ツールの開発、より良い配送方式の研究等があげられる。

謝辞 末筆ながら、本研究を行うにあたって、有益なご助言をいただいた春本要助手をはじめとする西尾研究室の諸氏に感謝の意を表す。

参 考 文 献

- 1) 萩野浩明, 門林理恵子, 清 一隆, 塚本昌彦, 西尾章治郎: 推論機構を用いたメール分配システム MILD におけるメール検索機構, マルチメディア通信と分散処理ワークショップ論文集, pp.283-290 (1995).
- 2) 平岩真一: グループウェア管理支援システムの構築, 情報処理学会研究報告, マルチメディアと分散処理 63-21/グループウェア 5-21, pp.157-164 (1994).

- 3) 岩室元典, 門林理恵子, 塚本昌彦, 西尾章治郎: 配送先に関する知識ベースを用いたメッセージ配送システム, 電子情報通信学会第6回データ工学ワークショップ論文集, pp.135-142 (1995)
- 4) 松下 温: グループウェア実現のために, 情報処理, Vol.34, No.8, pp.984-993 (1993).
- 5) 溝口文雄, 柳田正博, 大和田勇人: 帰納学習を用いた電子メールエージェントの設計, 日本ソフトウェア科学会第12回大会論文集, pp.277-280 (1995).
- 6) 森田昌宏, 速水治夫: 情報フィルタリングシステム—情報洪水への処方箋, 情報処理, Vol.37, No.8, pp.751-758 (1996).
- 7) 岡田謙一, 市村 哲, 松浦宣彦: グループウェアにおけるコミュニケーション支援, 情報処理, Vol.34, No.8, pp.1028-1036 (1993).
- 8) 清 一隆, 塚本昌彦, 西尾章治郎: ドット記法とIS-A 関係を用いた推論システムの高速化に関する研究, 第8回人工知能学会全国大会論文集, pp.517-520 (1994).
- 9) 塚本昌彦, 岩室元典, 門林理恵子, 西尾章治郎: 推論機構によるグループの動的構成手法を用いたメール分配システム, 人工知能学会誌, Vol.11, No.5, pp.735-743 (1996).
- 10) Tsukamoto, M. and Nishio, S.: Inheritance Reasoning by Regular Sets in Knowledge-base with Dot Notation, *Deductive and Object-Oriented Databases*, Ling, T.W., Mendelzon, A.O. and Vieille, L. (Eds.), *Lecture Notes in Computer Science*, Vol.1013, pp.247-264, Springer-Verlag (1995).

(平成9年5月12日受付)

(平成9年9月10日採録)



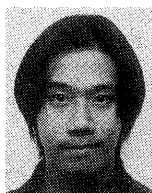
上田 宏高 (学生会員)

1997年大阪大学工学部情報システム工学科卒業。現在, 同大学大学院工学研究科情報システム工学専攻博士前期課程在籍。知識ベースに興味を持つ。



門林理恵子 (正会員)

1985年大阪大学文学部史学科卒業。1997年同大学大学院工学研究科博士後期課程修了。博士(工学)。ソフトウェア会社勤務を経て, 1990年シャープ(株)入社, 移動体通信プロトコルの研究等に従事。1997年8月より, (株)エイ・ティ・アール知能映像通信研究所奨励研究員。コミュニケーション支援技術の研究に従事。電子情報通信学会会員。



萩野 浩明 (学生会員)

1996年大阪大学工学部情報システム工学科卒業。現在, 同大学大学院工学研究科情報システム工学専攻博士前期課程在籍。モバイルコンピューティングに興味を持つ。



塚本 昌彦 (正会員)

1987年京都大学工学部数理工学科卒業。1989年同大学大学院工学研究科修士課程修了。同年, シャープ(株)入社。1995年大阪大学工学部情報システム工学教室講師, 1996年より同大学大学院工学研究科情報システム工学専攻助教授, 現在に至る。工学博士。時空間データベースおよびモバイルコンピューティングに興味を持つ。ACM, IEEE等7学会の会員。



西尾章治郎 (正会員)

1975年京都大学工学部数理工学科卒業。1980年同大学大学院工学研究科博士課程修了。工学博士。京都大学工学部助手, 大阪大学基礎工学部および情報処理教育センター助教授を経て, 1992年より大阪大学工学部情報システム工学教室教授となり, 現在に至る。この間, カナダ・ウォータールー大学, ビクトリア大学客員。データベース, 知識ベース, 分散システムの研究に従事。現在, Data & Knowledge Engineering, Data Mining and Knowledge Discovery, New Generation Computing等の論文誌編集委員。ACM, IEEE等7学会各会員。