

情報検索を支援する電子メール・エージェントの開発

5 L - 9

ワサカ・ヴィスーティヴィセツト、早川栄一、並木美太郎、高橋延匡
(東京農工大学)

1. はじめに

ネットワークの普及により、電子メールを利用する局面が増えている。ユーザは、限られた時間で多くのメールの処理を行わなければならない。このことから、(1) メールが多い場合、重要なメールまたは興味のあるメールを優先的に読みたい、(2) 過去に受け取ったメールから必要な情報を探したいという要求がある。

これに対して、我々は電子メールの中の情報整理、検索を行うメールエージェント[1]を作成することで、これらの支援を行う。本稿では、このシステムの設計と実現について述べる。

2. 本エージェントの特徴

まず、電子メールの利用形態の調査のために、著者が所属する学科で電子メール利用者 38 人に対してアンケートを行った。その結果、電子メールを利用する上で、メールの検索、優先度づけ、およびアドレスの整理の 3 点がもっとも重要であると考え、エージェントに次の特徴を持たせた。

(1) メール検索機能

複数のメールから、必要な情報を検索して提供する検索機能。

(2) フィルタリング機能

ユーザが設定した重要度に沿って、到着したメールの提示順序を変更するフィルタリング機能。

(3) アドレス登録機能

メールヘッダおよび本文内にある送り元のユーザの情報を抽出し、アドレス帳へ登録する機能。

さらに、これらを利用しやすくするために、GUI を備えた統合環境を提供する。

3. 検索機能

ユーザが本当に欲しいメールを取り出すために複数のキーワードで検索することが望ましい。よって、ユーザが一度に複数のキーワードを入力

できるようなキー入力エントリを用意した。また、多様な検索条件に対応できるように、正規表現をサポートする入力エントリも用意した。

現在の電子メールの使用状況についてのアンケートの結果から、もっとも使われている検索方法は受信した年月日と送り元で検索することである。よって、受信した年月日と送り元も指定できるようにそれぞれの入力エントリを用意した。ヘッダを含めて検索するかどうかはユーザに指定してもらう。

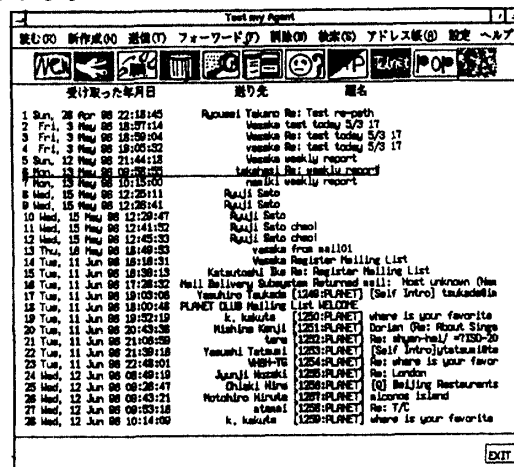


図1: メールの一覧を表示するメインウィンドウ

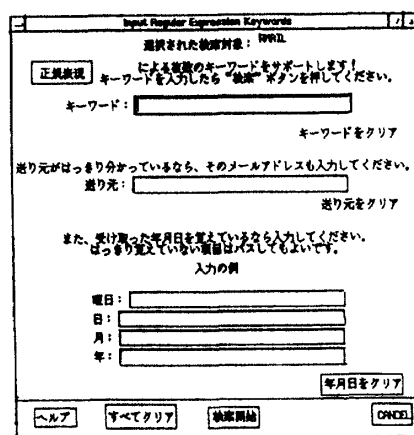


図2: 検索キーワードを入力するウィンドウ

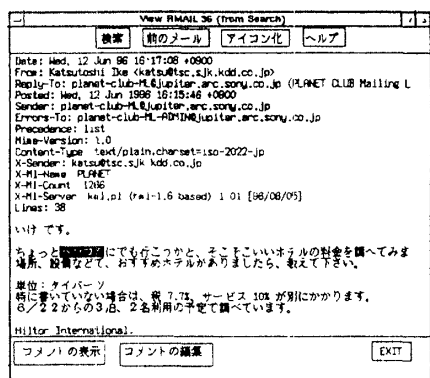


図3：検索結果のメール

(マッチされたキーワードのところが点滅)

ほかに、特にメーリングリストで見かけるような、もらったメールを引用し、それに対するコメントを返信するようなメールがたくさんある。このような関連のあるメールを取り出したい要求があった。本システムでは、このような関連するメールを取り出し、次のルールで関連性の高いものと低いものを決め、順位を付ける。

- (1) サブジェクトのところに調べたいメールの題名を返信したという印 (Re:) があるものはないものより関連性が高い。
- (2) 調べたいメールに対し、もっとも引用した時間の差が少ないものほど関連性が高い。
- (3) 引用した行数が多いほど関連性が高い。

4. 電子メールフィルタリング

ユーザが読むべき重要なメールとユーザが興味あるメールを取り出す。どのメールが重要であるか、どのメールがユーザの興味あるかを定めるための検索キーワードを次のように定めた。

- (1) 重要なメールなら、“重要”、“Urgent”、“至急”、“早く”などのキーワードが含まれていると仮定する。
- (2) 興味のあるメールを取り出すには、検索のキーワードとしてユーザから直接に興味のあるものを登録してもらう。

フィルタリングの方法としては、登録されたキーワードを検索し、それらのキーワードとマッチした回数で見つかったメールの優先順位を決める。フィルタリングの結果は、その優先順位でメールの一覧に表示する。

5. アドレス帳登録機能

大量のアドレスデータだと、ユーザが手作業で全部登録するのは手間がかかるので、本システムでは受信したメールのヘッダの“From”フィールドから送り先のメールアドレスを収集するように設計した。

メールの最後に付いているシグネチャには、所属、電話番号、ホームページアドレスなど相手についての情報も含まれていることが多い。よって、これらを有効に活用するためにシグネチャの解析を行い、アドレス帳に登録する。

シグネチャの解析法としては、メールの最後から10行以内のものであり、かつ‘Tel’、‘Fax’、‘http:’、などのよくシグネチャに書かれているキーワードをもとに検索する。

6. 実現

本システムは、X Window上で、C言語およびTcl/Tkを用いて実現した。実現規模は、全部あわせて4,000行程度である。現在、メール検索機能が実現されている。

7. おわりに

情報検索を支援する電子メールエージェントシステムを作成した。これにより、多量なメールを受け取るユーザに対する支援が可能になった。

現在、本システムでは、マッチされたキーワードの関数でメールの検索結果の優先順位を付ける。だが、複数のキーワードの場合、入力されたキーワードはそれぞれの重要度を持っているので、マッチしたキーの回数だけでは完全なユーザが欲しいメールの順位を取れない。キーワードの重要度を定める最適な方法を、求めることが今後の課題である。

参考文献

- [1] Pattie Maes, “Agents that Reduce Work and Information Overload”, Communications of the ACM, July 1994, Vol.37, No.7, pp.31-40