

電子メールを基にした情報共有システムに関する考察

神堀 真也[†] 相田 仁[†]

[†]東京大学大学院 新領域創成科学研究科 基盤情報学専攻

電子メールシステムは、我々が社会的活動を行う上で不可欠なコミュニケーションツールである。しかし、通信手段としての信頼性や情報処理・共有の支援という観点から見ると、電子メールは十分な機能を備えていない。

そこで我々は、特に企業や大学などの組織におけるコミュニケーションの円滑化と情報処理の効率化を目的とし、それを実現する電子メールを基にした情報共有システムの検討を行っている。

本稿では、送信者が電子メールにメタ情報を付加することにより、受信者の情報処理を支援する仕組みについて検討する。送信者がメールにタグや情報間の関係性をメタ情報として付加することで、受信者側の MUA で情報の話題分類や重要度評価が可能となり、受信者の情報処理支援につながると考えられる。

A Study for Information Sharing System based on E-mail

Shinya KAMBORI[†]

Hitoshi AIDA[†]

[†]Department of Frontier Informatics, Graduate School of Frontier Sciences, The University of Tokyo.

E-mail system has become a necessary communication tool in our social activity. However, it is not equipped with functions for reliability or assistance for user's information processing.

We aim at implementing an information-sharing system based on e-mail in order to assist communication and information processing in an organization such as a company or a university.

In this paper, we discuss about a system to assist receivers in processing information by adding meta-information to messages. The sender describes tags and relationship between messages as meta-information. This enables the receiver's MUA to classify the information by its topic and evaluate importance of the information.

1 はじめに

電子メールは、すでに通信手段の一つとして広く社会に受け入れられている。個人的な連絡のみならず、会社での業務連絡や顧客とのやり取りなど様々な用途にメールが利用されており、今後もメールは重要な役割を担い続けると考えられる。しかしながら、通信手段としての信頼性や情報処理・共有の支援という観点から

見ると、電子メールは十分な機能を備えていない。

そもそも、電子メールは通信速度や品質が保証されないインターネットを活用するために作られたコミュニケーションツールであり、作られた当時は“メールが送信できること”が最も重視されていた。したがって、電子メールの信頼性やセキュリティに関わる機能はあま

り備わっていない。そして、電子メールのコミュニケーションツールとしての重要性が格段に増した現在においても、電子メールの仕組みや機能そのものはほぼ変わっていない。

我々は、電子メールを基にした効率的な情報共有システムを実現することを目指しており、コミュニケーションツールとしての信頼性の向上と情報処理支援の観点から検討を行っている。本稿では、情報処理支援に着目し、送信者が電子メールにメタ情報を付加することで受信者の情報処理を支援する仕組みについて検討する。

2 電子メールの基礎技術

電子メールが送信者から受信者に配送される過程は、メールクライアントからメールサーバ、またはメールサーバからメールサーバへのメールの配送と、メールサーバ上のメールの受信の2つに大別できる(Fig. 2.1)。そのためのプロトコルとして、前者には SMTP が、後者には POP3 や IMAP が利用される。

メールを送信する際は、まずメールクライアントから SMTP によりメールサーバにメールが送信される。メールを受信したメールサーバは、メールの宛先から次に配送すべきメールサーバを判断し、そのメールサーバに対して SMTP で送信する。メールが、宛先のメールサーバに到着すると、メールサーバはユーザのメールボックスに保存する。受信者は、POP や IMAP を用いてメールサーバにアクセスすることで、自分のメールボックスに保存されているメールを読むことができる。

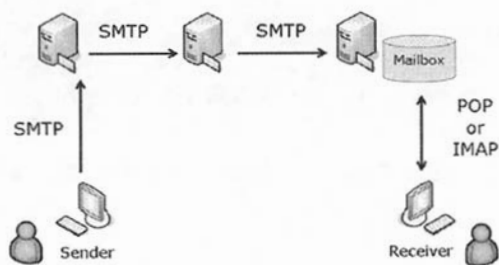


Fig. 2.1: メール転送の仕組み

また、電子メールのメッセージフォーマットは RFC2822[1]で定義されており、メールヘッダとメール本文からなっている。ヘッダは、送信元アドレスを示す From フィールド、宛先アドレスを示す To フィールドなどが含まれており、ヘッダと本文とは空行で区切られている(Fig. 2.2)。

```
Return-Path: <alice@example.com >
Received: from mail.example.jp
Received: from mail.example.com
Message-ID: <xyz1234@example.com>
Date: Wed, 06 Feb 2008 09:13:53 +0900
From: <alice@example.com >
Reply-To: alice@example.com
User-Agent: Mozilla/5.0
MIME-Version: 1.0
To: bob@example.jp
Subject: Hello!

Hello Bob.
How have you been?
Alice.
```

Fig. 2.2: メッセージフォーマット

3 MUA に求められる情報処理支援機能

電子メールが重要な通信手段として様々な場面で利用されるようになった結果、ユーザは大量のメールを受信しそれらを処理しなければならない状況に置かれている。そこで、受信したメール、すなわちユーザが受け取った情報を効率よく処理することを支援する機能が

MUA (Mail User Agent) に求められている。ここでは、ユーザの情報処理を支援するためにはどのような機能が必要か議論する。

話題別分類

我々は、電子メールを利用して様々な情報を送受信している。このように異なる話題や重要度の情報が混在している状況では、情報の受信者はまず話題別に情報を分類しなければならない。情報を話題別に分類することで、ユーザに対して情報をその分類ごとに提示することや、ユーザが参照している情報に関連する情報を参考として提示することが可能になる。

また、話題別の分類と類似しているが、情報の関係性の分析も有用である。1つの話題について複数のメールが当事者間でやり取りされる場合、それらは以前に送信されたメールとの間に、返信、補足、訂正等の関係性を持っていることが多い。これらの関係性を MUA が判断できれば、ユーザへの情報提示に活用することができる。

重要度の評価

電子メールにより伝達される情報には、受信者にとって非常に重要なものから無意味なものまで様々である。重要な情報と需要でない情報が混在していると、重要な情報がそのほかの情報に埋れてしまい、受信者がそれを見逃してしまう危険性がある。また、重要でない情報の処理にも受信者の労力が割かれてしまう。特に、現在多くの電子メールユーザを悩ませている問題のひとつに、本人が望んでいないのに送られてくる迷惑メールの氾濫がある。これも、迷惑メールが不要な情報であること、つまり電子メールの重要度が適切に評価されていないのが原因であると考えることができる。情報の重要度を評価することで、ユーザに対して情報を

重要度の高い順に提示することができ、ユーザが重要度の高い情報から処理することを支援することが可能になる。また、重要でない情報を選別することで、ユーザの情報処理にかかる負荷を軽減することにもつながる。

メールベースのアプリケーションへの拡張性

アンケート調査、スケジュール調整など、電子メールでのコミュニケーションの中には、ある程度決まった型のコミュニケーションが存在する。このような定型のコミュニケーションについては、特に最終的に情報を収集するユーザに対して収集内容の分析するような電子メールを基にしたアプリケーションが容易に構築できると有用である。

4 MUA の機能の現状と課題

既存の MUA にも、情報処理を支援するための機能がいくつか存在する。そこで、既存の機能でどこまで可能であるか、そしてどのような機能が不足しているかについて考察を行う。

話題別分類

情報を話題別に分類するには、複数のメールアドレスの使い分けや件名による話題の表示、送信元アドレスに基づく振り分けを行っている。

最近では、一人で複数のメールアドレスを所有していることは珍しくない。特に、所属する会社や大学で発行されるアドレスや、利用している ISP のサービスの一つとして発行されるアドレス、Gmail や Yahoo!メールなどのフリーメール等、我々は多くのメールアドレスを所有し、それぞれを用件に応じて使い分けしている。例えば、同じ人にメールを送る場合でも、仕事に関する連絡は会社のメールアドレスである `alice@xxxx.co.jp` 宛に送るが、プライベートな

連絡は一般のメールサービスで発行された `alice@yyyyy.com` 宛に送る等、送信者はメールの話題に応じて宛先アドレスを選択している。これにより、仕事の用件とプライベートな用件が混在することを防いでいる。

また、ほとんどの MUA にはメッセージフィルタリング機能が備わっている。これにより、メール受信時に送信元アドレス、件名、本文等を解析し、あらかじめ定めたフィルタリングルールに基づいてメールを分類することができる。メッセージフィルタリング機能を利用することで、メールマガジンのような話題があらかじめ決まっているものについては、送信元アドレスから分類することが可能である。また、メールアドレスのドメイン部分から送信者の属する組織を判別し、それを基に話題を推定することも可能である。他にも、件名や本文に特定の単語が含まれているか否かで話題を推定し分類することも行われている。

しかしながら、これらの機能によりメールの話題による分類が完全に行われているわけではない。まず、自分の所有するメールアドレスの使い分けに関するルールがメールの送信者と共有されているとは限らないので、受信者の意図する話題とは異なるメールがまぎれる可能性がある。そもそも、ユーザが所有するメールアドレスの数よりも、電子メールによりやり取りされる話題の分類の方が一般に多いため、送信先となるメールアドレスから判断しても、十分な分類ができていないのが現状である。また、件名や本文に含まれる特定の単語を基に話題を推定しフィルタリングを行う方法についても、必ずしも正確であるとはいえない。

重要度の評価

重要度の評価については、既存の仕組みとして迷惑メールフィルタリングが挙げられる。こ

れは、重要度の高いものの抽出とは逆のアプローチではあるが、迷惑メールのような受信者にとって重要ではないメールを低く評価する仕組みといえる。迷惑メールフィルタリングの方式としては、現在以下のものがよく用いられている。

Keyword filters: 指定した文字列がメールに含まれるか否かで判定する。

Heuristic filters: 迷惑メールの送信元や文面などの特徴をルールとして膨大なデータベースに集め、それを元に判定する。

Bayesian filters: ユーザによる判定を先例として学習し、それを元に判定を行う。

以上に加えて、送信元サーバのアドレスの利用し、あらかじめ拒否するアドレスを設定しておくブラックリスト方式や、許可するアドレスを設定しておくホワイトリスト方式などがある。このほかにも、迷惑メールをフィルタリングする仕組みについては様々な手法が検討されている[2-5]。

しかしながら、迷惑メールフィルタリングでは、迷惑メールか否かの判定しか行われないため、重要度として考えた場合、2段階にしか分類されておらず不十分である。

また、Outlook や Thunderbird など一部の MUA では、送信者がメールの重要度を指定する機能を備えている。これは、メールヘッダに X-Priority 等のフィールドを用意し、そこにメールの重要度を表す値を記述する仕組みである。しかしながら、送信者が考えるメールの重要度と、受信者が考えるメールの重要度は必ずしも一致しない。そのため、この機能はあまり有効に利用されてはいない。

メールベースのアプリケーションへの拡張性

既存の電子メールシステムは、基本的にテキストの送受信である。メーリングリストを用いて特定のユーザ間で情報を共有することは盛んに行われているが、アンケート調査、スケジュール調整など、電子メールによる定型のコミュニケーションを支援する仕組みについては、一般に普及しているものは存在しない。

5 送信者によるメタ情報の付加

MUA によるユーザの情報処理支援を実現するには、情報の話題による分類と重要度の評価を行い、それに基づいてユーザに提示する技術が必要である(Fig. 5.1)。しかしながら現状では、MUA がそれらを判断するための十分な情報がメールに含まれていない。

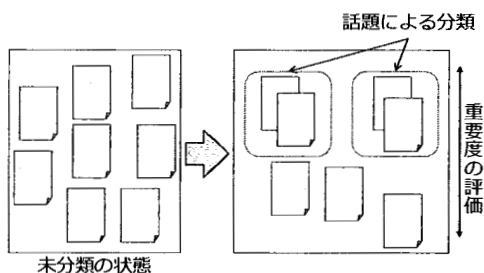


Fig. 5.1: MUA による情報整理支援のイメージ

そこで、MUA がこれらの情報処理支援を実現するため、送信者による電子メールへのメタ情報の付加を提案する。

メタ情報としては、メールの内容を示すタグと、情報の関係性を考える。

5.1 メールへのタグ付け

タグ付けによる情報の分類は、ブログや写真共有サイトなどにおける対象の分類手法として広く利用されている。1つの記事に対して複数のタグを付けることにより、複数の視点から

の情報分類が可能である。

既に、Gmail や Thunderbird 2 では、メールに対してタグ付けを行う機能がある。これは、受信したメールに対して、受信者が内容等に応じてタグ付けをするものである。これに対して、本研究で考えるタグ付けは、送信者によるものである。これは、受信者によるタグ付けと比較して、以下の特徴が挙げられる。

受信者がタグ付けを行う場合、一定のルールに基づいて自動的にタグ付けを行う場合と、受信者自身が手動でタグ付けを行う場合とがあるが、いずれの場合も送信者の意図とは異なるタグを付ける可能性がある。また、手動で行う場合について言えば、一度受信者がメールの内容を把握する必要があり、後から目的のメールを探しやすくする意味合いの方が大きく、メール処理の効率化という点では、効果は限定的である。

送信者がメールに対してタグ付けを行うことで、話題や重要度を正確に反映させることができる。

5.2 タグと送信者との関連付け

送信者によるタグ付けに関する課題も考えられる。それは、異なる意味を持つタグの名称の重複である。受信者が自分宛のメールに対してタグ付けを行う場合はタグ名の重複はないが、送信者がタグ付けを行う場合、異なる送信者が異なる意味を持つ同名のタグを付ける可能性がある。

そこで、タグの定義をしたユーザのメールアドレスとそのタグと関連付ける。

たとえば、alice@example.com というアドレスを持つユーザが「会議」というタグを定義すると、そのタグはalice@example.com:会議とする。これにより、bob@example.jp というアドレスを持つユーザがタグに同じ「会議」という

名称を用いても、bob@example.jp:会議となり識別することが可能である。

5.3 情報の関係性

ある話題について複数回のメールの送受信が行われる場合、これらのメールは以前に送受信されたメールに対する返信、補足、修正、再送などの関係性を持っている。これらの情報の関係性を、メタ情報としてメールに付加することで、MUA がメールの話題の分類や重要度の評価を行う際に活用できる。

6 メタ情報の記述法

メタ情報の記述法としては、2通り考えられる。

1つ目は、既存のメールヘッダの仕様に従い、タグ情報や、情報の関係性を示すフィールドを定義し、そこにメタ情報を記述する方法である。メールヘッダの仕様では、既に定義されているフィールドのほかに、“X-”で始まるフィールドを定義し自由に利用することができる[6]。そこで、例えば“X-MESSAGE-TAGS”や“X-MESSAGE-RELATIONS”のようなフィールドを新たに定義し、タグ情報や情報の関係性を記述することで、メールにメタ情報を付加することができる。

2つ目は、XMLを採用し、タグ情報や情報の関係性のほか、メール本文もあわせてXMLの要素として記述する手法である。

この2つの手法について、それぞれ長所と短所を考察する(Fig. 6.1)。

メールヘッダへの記述

メールヘッダにメタ情報を記述する場合の長所として、既存のメッセージフォーマットの枠組みの中で実現できる点が挙げられる。“X-”で始まるフィールドを新たに定義することは、

前述のとおり既存のRFCに従ったものである。よって、メタ情報が付加されたメールを従来のMUAで受信したとしても、通常のメールとして扱える。

その一方、メール単位でしかタグや情報の関係性を記述することができないという短所がある。一般に、1通のメールに複数の話題について書かれることがある。この場合、メールヘッダ内にメタ情報を記述すると、メールのどの部分に対する情報であるかを明示することができない。したがって、情報の分類があいまいになる可能性がある。

メール本文のXML化

メール本文もあわせてXMLで記述した場合、メールの特定の部分についてタグ付けや関係性の記述を行うことができる。したがって、メールヘッダに記述する場合と比較して、より詳細な情報を得ることができる。

一方で、従来のMUAは本文がXMLで書かれたメールには対応していない。そのため、受信者のMUAを意識してメールを作成する必要が生じる。

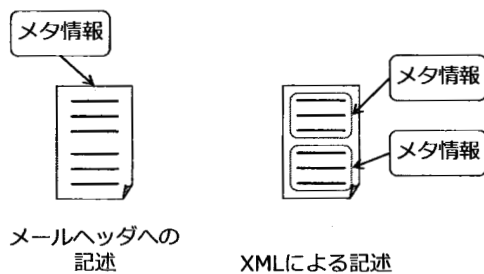


Fig. 6.1: メタ情報の記述法の比較

7 まとめ

本稿では、MUAによる情報の話題分類と重要度評価を実現するため、送信者がメールに対してメタ情報を付加することを提案し考察を

行った。メタ情報として、情報の内容を示すタグと、情報間の関連性とを想定し、それらを活用して受信者の MUA で情報の分類と重要度評価を行うことで、ユーザの情報処理を支援することが可能になると考えられる。

メタ情報の記述法としては、メールヘッダへの記述とメール本文の XML を挙げ、両者の長所と短所を比較した。従来の MUA との互換性を考慮するとメールヘッダへの記述の方が容易であるが、記述できる情報の詳細さや拡張性を考慮すると、XML を用いた方が有用性が高い。

今後は、XML によるメタ情報の記述法や、メタ情報に基づいた情報分類や重要度評価を行うシステムについて検討していく。

参考文献

- [1] P. Resnick, "Internet Message Format", IETF RFC 2822, Apr. 2001.
- [2] Muhammad N. Marsono, M. Watheq El-Kharashi, Fayez Gebali, "Rejecting Spam during SMTP Sessions," IEEE PACRIM'07, 2007.
- [3] Park So Young, Kim Jeong Tae, Kang Shin Gak, "Proposal of a New Effective Spam Mail Regulation," IEEE ICACT 2005, 2005.
- [4] Muhammad N. Marsono, M. Watheq El-Kharashi, Fayez Gebali, Sudhakar Ganti, "Distributed Layer-3 E-mail Classification for Spam Control," IEEE CCECE/CCGEI, May 2006.
- [5] 松原義継, "偽の送信者メールアドレスを持つメールの配送を防止するフィルタ," 情報処理学会論文誌, Vol.47, No.4, Apr. 2006.
- [6] R. Fajman, "An Extensible Message Format for Message Disposition Notifications", IETF

RFC2298, Mar. 1998.

- [7] 大向一輝, 武田英明, "人間関係ネットワークに基づく情報フィルタリングを用いた強調的タスクスケジューラ", 電子情報通信学会論文誌, Vol.J87-D-I, No.11, pp.1020-1029, Nov. 2004.
- [8] 永田周一, 安村通晃, "Enzin: 情報の公開範囲を手軽に変更できるコミュニケーションツール", 第13回インタラクティブシステムとソフトウェアに関するワークショップ, 2005.
- [9] 高林哲, 増井俊之, "QuickML: 手軽なグループコミュニケーションツール", 情報処理学会論文誌, Vol.44, No.11, Nov. 2003.
- [10] 江渡浩一郎, 高林哲, 増井俊之, "qwikWeb: メーリングリストと Wiki を統合したコミュニケーションシステム", インタラクシオン 2005, Feb. 2005.