



メール要点抽出&転送システム／早解メール™

奥村 晃弘
野崎 正典

野中 雅人
奥村 幸治

濱口 佳孝
清水 泰志

インターネットの普及と共に、電子メール（以下、メールと省略）を用いた業務連絡スタイルが企業の規模や業種を問わず浸透してきている。いまやメールはオフィスで利用されるだけでなく、携帯電話などを用いて外出先からも送受信されるようになった。しかしメールの普及とともに利用者の負担も増加しており、「メールの処理に時間を要する」、「重要なメールが埋没する」などの不便が目立つようになってきている。さらに現在行われている携帯電話へのメール転送サービスを利用した場合、「長いメールが分割転送され読みにくい」、「メールの必要性に関らず最初から全文が転送され、通信料金の負担が増える」などの問題も生じている。

本文で解説する早解メール™*1) では、メールが持つ文章の意図を推定することにより、本来の主旨に合致した要点だけの文章を、指定したメールアドレスへ転送することを実現した。要点だけのメールにすることで、利用者は携帯電話のようなハードウェア的に制約のある環境でも、手軽にメールの概要を把握することができる。また最初に要点を見て関心のあるメールのみ全文を取り寄せることになるため、メールを読む時間が短縮されるだけでなく、通信料金の削減も可能となる。

早解メール™概要

早解メール™は、アプライアンスサーバの形でイントラネット上に設置することを想定している。図1は早解メール™サーバを利用したメールシステムの構成例である。サーバは収容するユーザ数に応じて、タワータイプとラックマウントタイプの2タイプから選択することができる。以下、図1を用いて社外からのメールが携帯電話に転送される様子を説明する。

社外から送信されたメールは、イントラネット上の既存のメールサーバで受信される。・・・(1)

ユーザが社内に在席している場合は、社内のPCからPOP3 (Post Office Protocol3) などのプロトコルを用いてメールを受信する。ユーザが外出している場合は、既存メールサーバから早解メール™サーバへSMTP (Simple Mail Transfer Protocol) もしくはPOP3によりメール配信を行う。・・・(2)

早解メール™サーバでは、最初に差出人アドレスなどから転送するメールをフィルタリングする。次いで、指定文字数分だけ要点を抽出したメールを生成し、転送先の携帯電話に送信する。・・・(3)

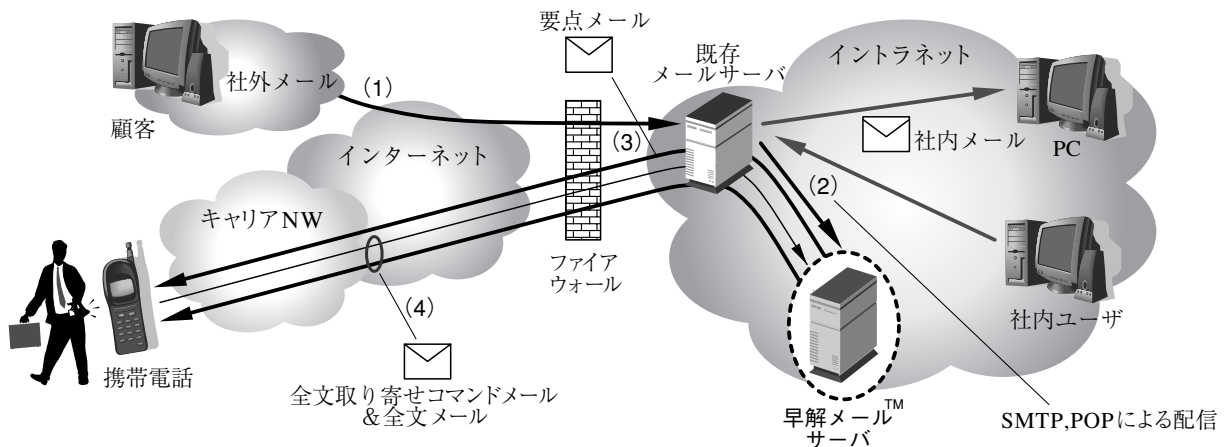


図1 早解メール™サーバを用いたメールシステム構成例

*1) 早解メール™は、沖電気工業(株)の商標。

元のメール全文や送付先リストなどは、携帯電話から早解メール™サーバにコマンドを記述したメールを送信することで取得できる。・・・(4)

同様にコマンドメールを使用して、元のメールに対しあたかも会社から送信されたような返信メールを出すことも可能である。

早解メール™サーバの主要機能

早解メール™サーバは基本機能となるメール本文からの要点抽出&転送機能、メール全文の取り寄せ機能、メール差出人への返信機能に加え、メール転送に関する設定機能を備えている。

メールの利用環境は、使用する携帯電話や契約している携帯電話会社、メールの利用頻度や内容、セキュリティポリシーなどユーザによってさまざまである。早解メール™サーバでは、ユーザはクライアントPCからWebブラウザを通じてそれぞれのメール環境に合わせた設定ができるようになっている。

表1に早解メール™サーバの主要機能一覧を示す。

ユーザは転送する要点文の文字数や全文取り寄せ時のメール文字数を指定できる。これにより携帯電話会社のメールサービスの文字数に合わせたメール転送が実現される。また、本文中の改行を削除して文を詰めて表示する、メールの題名を本文に挿入するといったメール表示形態の変更もできる。転送制御機能としては、差出人アドレスやサブジェクト・本文に含まれる文字列から転送するメールを制限するフィルタリング機能、携帯電話へのメール転送を一時的に停止する機能を備えている。一時停止はコマンドメールで制御することもできる。

管理者向けの機能としては、ユーザのアカウント作成/削除などのユーザ管理機能に加え、安易に社外へのメール転送が行われないように転送可能なドメインの指定やコマンドメールの利用を制限する機能が用意されている。

表1 早解メール™の主な機能

ユーザ向け機能	
要点の抽出転送機能	レイアウト変更機能 (改行除去, 題名の本文挿入など)
メール全文取り寄せ機能 (一括/分割パートごと)	抽出文字数設定機能
全文の分割機能	全文分割文字数設定機能
返信機能	Webによる設定変更機能
転送ON/OFF機能 (転送先からも可能)	POPサーバ設定機能
フィルタリング機能	
管理者向け機能	
転送可能ドメイン設定機能 (ユーザごと指定可能)	差出人メールアドレス (送信時) の書き換え機能
コマンドメールの制限機能 (ユーザごと指定可能)	

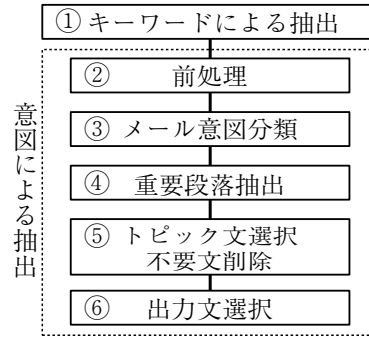


図2 要点抽出の流れ

メールからの要点抽出方法

早解メール™における要点抽出の流れを図2に示す。早解メール™ではキーワードによる抽出(①)と意図による抽出(②~⑥)の2種類の抽出方法を併用している。

意図による抽出とは、メール本文の表層表現から書き手の意図を分類し、その分類結果に基づいて要点を抽出する独自の要点抽出アルゴリズム^{1) 2)}である。以下、それぞれの抽出方法について説明する。

(1) キーワードによる抽出

キーワードによる抽出は、メール本文で「箇条書き」になっている部分から要点を抽出するものである。具体的には、「日時」、「場所」など「箇条書き」の見出しとして頻繁に利用されるキーワードが登録されており、このキーワードの後に「:」(コロン)や空白、タブなどが記述されている場合にその行を抽出する。

(2) 意図による抽出

キーワードによる抽出が「箇条書き」を対象とするのに対して、意図による抽出は一般の文章を対象とする。

「②前処理」では、メール本文の文書構造を解析し、文と段落に分割する。早解メール™では、空行や「----」のような記号だけの行があると段落の切れ目と判断する。また、行頭に「>」などの引用記号がある行を引用段落とする処理も行う。処理過程の例を図3、図4に示す。

メール本文

伊藤様、開発部の山田です。

>出退勤管理システムに関する質問です。
 >個人カレンダー更新してからメニューに戻ると、
 >出退勤情報更新を選択できなくなることがあります。
 >この場合一度終了しないといけないのでしょうか？

いつもお世話になっております。
 お問い合わせの件ですが、ご指摘のとおり一部不具合がありましたことを開発部でも確認致しました。
 現在、修正中ですので、もうしばらくお待ち下さい。

また、前記の不具合を回避する方法がありますので、お伝えしておきます。
 一度、他の月のデータを表示させれば、フラグがクリアされます。
 フラグがクリアされれば、また出退勤情報更新を選択できるようになるはずですが。

色々ご迷惑をお掛けして、申し訳ありません。
 今後ともよろしくお願い致します。

要点抽出結果

お問い合わせの件ですが、ご指摘のとおり一部不具合がありましたことを開発部でも確認致しました。現在、修正中ですので、もうしばらくお待ち下さい。

処理過程

段落	引用	文	内容	意図分類	マーク
1		1	伊藤様、開発部の山田です。		
2	>	2	出退勤管理システムに関する質問です。	回答	
		3	個人カレンダー更新してからメニューに戻ると、出退勤情報更新を選択できなくなることがあります。		
		4	この場合一度終了しないといけないのでしょうか？	回答	
3	重要段落	5	いつもお世話になっております。		×
		6	お問い合わせの件ですが、ご指摘のとおり一部不具合がありましたことを開発部でも確認致しました。	回答	○
		7	現在、修正中ですので、もうしばらくお待ち下さい。	依頼	
4		8	また、前記の不具合を回避する方法がありますので、お伝えしておきます。	報告	
		9	一度、他の月のデータを表示させれば、フラグがクリアされます。		
		10	フラグがクリアされれば、また出退勤情報更新を選択できるようになるはずですが。		
5		11	色々ご迷惑をお掛けして、申し訳ありません。		×
		12	今後ともよろしくお願い致します。		×



メール全体の意図 回答

図3 要点抽出の処理過程の例 (その1)

メール本文

伊藤様、開発部の山田です。

>出退勤管理システムに関する質問です。
 >個人カレンダー更新してからメニューに戻ると、
 >出退勤情報更新を選択できなくなることがあります。
 >この場合一度終了しないといけないのでしょうか？

いつもお世話になっております。
 お問い合わせの件ですが、ご指摘の不具合を開発部で確認することができませんでした。
 現在、調査中ですので、もうしばらくお待ち下さい。

また、不具合が発生する条件を絞りこみたいので幾つか教えてください。
 上司設定や承認ルート設定はどのようになっていますか？
 ユーザカスタマイズでなにか特別な設定をされていますか？

色々ご迷惑をお掛けして、申し訳ありません。今後ともよろしくお願い致します。

要点抽出結果

また、不具合が発生する条件を絞りこみたいので幾つか教えてください。上司設定や承認ルート設定はどのようになっていますか？ユーザカスタマイズでなにか特別な設定をされていますか？

処理過程

段落	引用	文	内容	意図分類	マーク
1		1	伊藤様、開発部の山田です。		
2	>	2	出退勤管理システムに関する質問です。	回答	
		3	個人カレンダー更新してからメニューに戻ると、出退勤情報更新を選択できなくなることがあります。		
		4	この場合一度終了しないといけないのでしょうか？	回答	
3	重要段落	5	いつもお世話になっております。		×
		6	お問い合わせの件ですが、ご指摘の不具合を開発部で確認することができませんでした。		○
		7	現在、調査中ですので、もうしばらくお待ち下さい。	依頼	
4		8	また、不具合が発生する条件を絞りこみたいので幾つか教えてください。	質問	
		9	上司設定や承認ルート設定はどのようになっていますか？	質問	
		10	ユーザカスタマイズでなにか特別な設定をされていますか？	質問	
5		11	色々ご迷惑をお掛けして、申し訳ありません。		×
		12	今後ともよろしくお願い致します。		×



メール全体の意図 質問

図4 要点抽出の処理過程の例 (その2)

次に、「③メール意図分類」で、各文の意図を推定し、各文の意図の重みを合計することで、メール全体を6種類の意図（連絡、報告、指示、依頼、質問、回答）に分類する。一般に日本語では意図は文末などに特徴的表現として表れる。早解メール™では意図とさまざまな特徴表現を対応付けた辞書を搭載している。この辞書を利用して各文の意図を推定する。辞書の内容例を表2に、それを使った各文の意図の分類結果を図3の分類結果の列に示す。この例ではメール全体の意図は「回答」となる。

「④重要段落抽出」ではメール全体の意図の根拠となった重要段落を抽出する。図3の場合、メール全体の意図は「回答」なので、意図が「回答」となる第6文のある第3段落を重要段落とする。なお第2段落も「回答」となる文を含むが、引用段落なので重要段落とはしない。

「⑤トピック文選択／不要文削除」では、「～については」のような話題提示表現に基づいて、話題を提示している文をトピック文としてマークし、挨拶や常套句などを不要文として削除する。処理例を図3のマークの列に記載する。ここで、「○」はトピック文を、「×」は不要文を表す。

最後に、「⑥出力文選択」で指定文字数に収まるように調整して出力する。重要段落中のトピック文を最優先に、次に同じ重要段落の一般文という順序で抽出する。

早解メール™は、メール全体の意図に注目し、その意図に沿った部分を要点として抽出するところに特長がある。このため、単純に重要表現のパターンを設定している場合とは異なり、メール全体の内容に合わせて抽出箇所がダイナミックに変わる。例えば、図4では図3とほとんど同じ構成のメールを処理しているが、それぞれ違う段落から要点を抽出している。図4ではメール全体の意図が「質問」に分類され、よりメールの主旨に近いと類推される文が選ばれる。一方、意図による抽出方法は、友達どうしの会話などでよく見られる複数の用件を含むメールや用件のはっきりしないメールにはあまり適していない。良好な結果が得られるのは図3に挙げたような用件が比較的明確なビジネスメールである。

表2 早解メール™の意図分類辞書（内容例）

適用条件	分類パターン	(意図；重み)
通常文	～ 質問です。	(質問；3)
引用文	～ 質問です。	(回答；3)
通常文	～ か？	(質問；2)
引用文	～ か？	(回答；2)
通常文	～ 指摘のとおり～	(回答；2)
通常文	～ お伝えし～	(連絡；1)，(報告；3)
通常文	～ 下さい。	(指示；1)，(依頼；1)

今後の展開

(1) セキュリティの強化

社外へのメール転送では盗聴に気をつける必要がある。現在、早解メール™では各々のユーザのセキュリティポリシーに合わせて個別に対応しているが、セキュリティ強化策として全文の参照をSSLに対応したWebメールにすることを検討中である。

(2) メール閲覧における操作性の向上

現在、我々はメールの意図や本文に含まれる日時情報などを用いて、グループ別にメールを提示するシステムを開発中である。このシステムでは、「質問」に関するメールだけを閲覧する、今日の日付が含まれるメールだけを閲覧するといったことが可能になり、既読メールの確認が容易になると期待できる。特にハードウェアに制約の多い携帯電話では効果的な機能になると思われる。

以上、メール要点抽出&転送システム／早解メール™が提供する機能と実現技術について説明した。早解メール™ではメールの要点だけを転送することにより、現在の携帯電話向けメール転送サービスに比べ、使いやすさや通信料金の面で優れている。

今後はメール要点抽出技術をコアに、よりビジネスシーンで使いやすいシステムに発展させていきたい。また、携帯端末を使った業務支援ソリューションやインターネットサービスプロバイダ向けのサービスなどの新しい提供形態も検討していきたい。

参考文献

- 1) 福本， 樹井：情報抽出技術の電子メールへの応用，知識発見のための自然言語処理シンポジウム，オンライン論文，電子情報通信学会，1999年11月
- 2) 福本， 樹井：沖電気研究開発181号“メールロボ：インターネットメールからの情報抽出”，Vol.66 No.2，pp.55-58，1999年10月

筆者紹介

奥村晃弘：Akihiro Okumura.研究開発本部 情報サービスシステムラボラトリ

野中雅人：Masato Nonaka.研究開発本部 情報サービスシステムラボラトリ

濱口佳孝：Yoshitaka Hamaguchi.研究開発本部 情報サービスシステムラボラトリ

野崎正典：Masanori Nozaki.ネットワークシステムカンパニー ネットワークインキュベーション本部 ネットワークサーバ事業推進部

奥村幸治：Kouji Okumura.ネットワークシステムカンパニー ネットワークインキュベーション本部 ネットワークサーバ事業推進部

清水泰志：Yasushi Shimizu.ネットワークシステムカンパニー ネットワークインキュベーション本部 ネットワークサーバ事業推進部