

K-025

電子メールマップインタフェースを用いた電子メール閲覧システムの提案 Implementing of an E-Mail Consult System based on E-Mail Mapping

平岡佑介[†] 大園 忠親[†] 伊藤孝行[†] 新谷 虎松[†]

Yusuke Hiraoka, Tadachika Ozono, Takayuki Ito, Toramatsu Shintani

1 はじめに

電子メールは連絡手段だけでなく知識及び情報の収集に利用される [1]。我々は、研究室などのコミュニティ内において電子メールの共有を行ない、知的活動を活性化する目的で、電子メール閲覧支援システム Wisdom Mail を開発している。図 1 に Wisdom Mail のスナップショットを示す。本システムは LAN 内で Wisdom Mail を利用しているユーザ間で電子メール共有を行なう。図中の (1) で示した部分に他ユーザの共有電子メールのフォルダがリスト形式で示されている。電子メールの所有者は共有を行なうかの設定をフォルダ又は電子メール毎に行なうことができ、表示されている各リストを選択することで、他ユーザの所持する電子メールを閲覧できる。しかし、共有された電子メールは多量かつ大量であり、ユーザにとって興味のある電子メールを探し、閲覧することは困難である。本問題点を緩和する目的で、電子メールマップの表示を行なうシステムの開発を行なった。本電子メールマップの機能として電子メールの類似度に基づいたマップ上へ電子メールの配置及びマップマークがある。生成される電子メールマップは互いに類似する電子メールがより近くに配置されるという特徴がある。また、電子メールマップ上で特定の電子メールをマップマークとして目立たせることができる。以上の特徴から、共有されている電子メールからある電子メールに類似した電子メールを収集し整理または閲覧を行うことを電子メールマップ上のマップマーク近隣の電子メールを選択するという直感的な操作で行うことができる。

2 電子メールマップ

2.1 Self Organizing Map

本システムでは Self Organizing Map (以下 SOM) [2] を用いて電子メールマップの生成を行なう。SOM は教師なし競争強化学習及び近傍学習を用いることで、ある分布に従う n 次元ベクトルで表現された入力データの特徴を抽出し、その分布を近似した特徴マップの生成を行う。SOM は入力層と 2 次元平面マップ上にノードを格子上に配置された出力層の 2 層からなり、入力されたデータが出力ノードの 1 つに出力される。各ノードは 2 次元平面上に表記され、類似した特徴を持つデータはマップ上の近い位置に出力されるという特徴がある。

2.2 電子メールマップ生成

図 2 に SOM に基づく電子メールマップ生成手法の流れを示す。前処理として電子メールの本文中から署名等のマップ生成処理を行う上で必要でない部分の削除を行う。次に各電子メールを特徴ベクトルに変換する。本電子メールマップ生成手法では、電子メールに

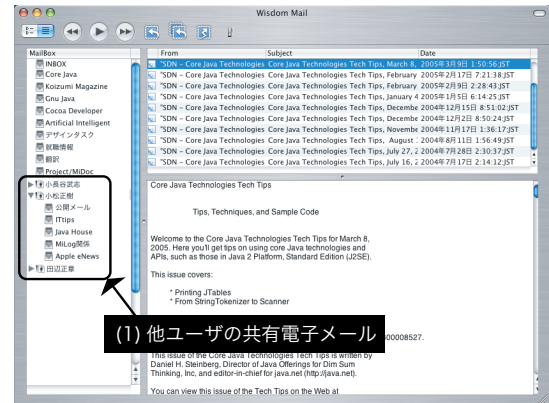


図 1: Wisdom Mail のスナップショット

はメーリングリスト単位及びスレッド単位でまとまりがあるというヒューリスティックに基づき、マップの生成を行う。具体的には、対象とする電子メールの属するメーリングリストのベクトル及びスレッドのベクトルを生成し、2つのベクトルを組み合わせたハイブリット型特徴ベクトルを用いる。電子メール M のメーリングリストベクトル $Vector_{list}(M)$ 及びスレッドベクトル $Vector_{thread}(M)$ を以下に示す。

$$Vector_{list}(M) = (l_1, l_2, \dots, l_r) \quad (1)$$

$$Vector_{thread}(M) = (k_1, k_2, \dots, k_s) \quad (2)$$

$Vector_{list}(M)$ 及び $Vector_{thread}(M)$ の生成はそれぞれ対象とするメーリングリスト及びスレッド中に出現する単語に基づいて生成する。以下に具体的な $Vector_{list}$ の生成の流れについて述べる。

ステップ 1. 各メーリングリスト本文又はメーリングリストの主題に単語 w が出現するスレッドの数を全て調べる。
ステップ 2. 単語を出現するスレッド数の多い順に並べ替え、上位 m 個の単語 (w_1, \dots, w_m) を特徴ベクトルの要素とする。
ステップ 3. メーリングリストの特徴ベクトルの s 番目の要素 l_r を次式で求める。

$$l_r = \begin{cases} 1 & \text{メーリングリストの主題に単語 } w_r \text{ を含む} \\ 0.5 & \text{メーリングリスト中の電子メールの本文にのみ単語 } w_r \text{ を含む} \\ 0 & \text{それ以外} \end{cases}$$

$Vector_{thread}$ も同様にスレッド中に出現単語に基づいて生成を行う。 $Vector_{list}$ 及び $Vector_{thread}$ の 2 つのベクトルを組み合わせたハイブリット型特徴ベクトル

[†] 名古屋工業大学 大学院工学研究科 情報工学専攻

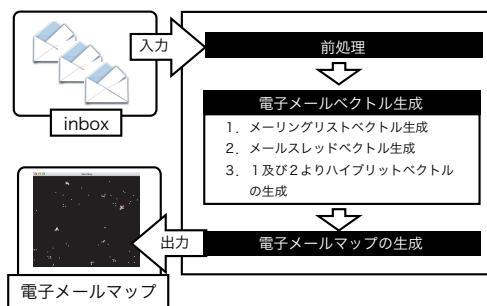


図 2: 電子メールマップ生成の流れ



図 3: 電子メールマップ例

Vector(M) を式 (3) に示す .

$$Vector(M) = (l_1, \dots, l_r, w \times k_1, \dots, w \times k_s) \quad (3)$$

ただし, w は重みであり, w を小さくすることでよりメーリングリスト毎のまとまりを重視した電子メールマップを生成する. 得られたハイブリッド型特徴ベクトルを対象に SOM を用い, マップ生成を行なう.

2.3 電子メールマップ例

図 3 に実際に電子メールマップを本システムによって表示した例を示す. 本マップでは各電子メールが白点で表されており, 類似する電子メールが多いほど白点の密度が高くなる. ユーザはマウスを用いてマップ中の領域を指定することで指定した領域に属する電子メールのリストを表示することができる. また, マップ上の目印としてマップマークを用いることができる. ユーザは興味のある電子メールを選択し, マップマークとして設定する. マップマークとして設定された電子メールはマップ中で赤点を用いて表示される. 本マップは互いに類似した電子メールが近くに配置されるため, マップマークの近隣の電子メールを選択する事で, 興味のある電子メールに類似する電子メールの閲覧を直感的に行なう事が可能になる.

3 実験

本システムで用いるハイブリッドベクトル中の要素が電子メールマップ生成に与える影響を調べるためにハイブリッドベクトル中の重み w を変化させてマップ生

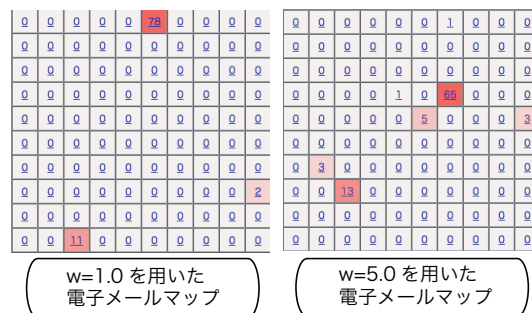


図 4: $w=1.0$ 及び $w=5.0$ の際の電子メールマップ

成を行った. マップ生成には Yahoo Groups¹中のメーリングリスト及び人工知能学会のメーリングリスト²を含む 9 種類のメーリングリストに送付された電子メールを対象として合計 129 通を用いた. 実験には 10×10 のマップを持つ SOM を用い, ハイブリッドベクトル中の $Vector_{list}$ 及び $Vector_{thread}$ の要素数を各々 300 としてマップ生成を行った. ハイブリッドベクトル中の重み w は 1.0 及び 5.0 を用いた.

図 3 に生成されたマップのスナップショットを示す. 分かりやすさのため, マップの各格子に出力されるメールスレッドの数を各格子の濃淡で示す. 色の濃い格子に多くのメールスレッドが出力されている. マップの各ノードに出力されたメールスレッドの数の平均は $w=1.0$ の時 30 スレッドで $w=5.0$ の時 13 スレッドであった. これは $w=1.0$ を用いたマップに比べて $w=5.0$ を用いたマップがスレッドの内容を重視して分類しているため, より詳細なマップ生成を行っているためと考えられる. $w=5.0$ のマップでは他メーリングリスト中の類似した内容の電子メールが近くに配置されることがあった. そのため, ユーザにとって類似した質問の電子メールや類似したイベント情報の記載された電子メールを調べる際に有効である. また, 重み w が小さいほど同一メーリングリストがマップ中で近隣に出力されるため, 対象とするメーリングリストに類似したメーリングリストに属する全電子メールを閲覧する目的で用いることができる.

4 まとめ

本稿では, ユーザの電子メール発見を支援する目的で, 電子メールマップのインタフェースを持つ電子メール閲覧システムの提案を行なった. 本システムは類似度に基づいた電子メールマップの生成及びマップマークの 2 つの機能を用いることでユーザが興味のある電子メールを探し出し, 閲覧を行なうことを支援する.

参考文献

- [1] Steve Whittaker, Candace Sidner: "Email overload: exploring personal information management of email", Human factors in computing systems(1996)
- [2] T.Kohonen, 徳高平蔵, 岸田悟, 藤村喜久郎:"自己組織化マップ", シュプリンガー・フェアラーク東京 (1996)

¹<http://groups.yahoo.com/>

²<http://www.ai-gakkai.or.jp/jsai/ml/>