

電子メール通信履歴を用いた人脈活用システム

5M-6

吉府研治 谷川哲司 宮下敏昭
NECヒューマンメディア研究所

1 はじめに

近年のインターネットの普及により、電子メール、ニュース、電子掲示板などネットワークを媒体としたコミュニケーションが活発に行われている。これらのコミュニケーション量が多くなるにつれて、ただ情報をやりとりするだけでなく、やりとりされた情報を様々な観点で整理して再利用したいというニーズが増えてきている。本稿では、送受信した電子メールの内容及び通信履歴から自分と相手との関係の強さや相手の人物属性のような、いわゆる人脈情報を自動的に取得する方式および人脈情報の視覚化方式を提案する。

2 人脈情報活用に向けた考察

2.1 人脈情報モデル

人脈情報とは、自分と関連のある人物に関する情報で、次の各情報から構成される。

関係分類情報：自分と相手がどのような関係か分類した情報。上司、同僚、友人などがある。

関係強度情報：自分と相手がどの程度関係が強いかわを示す情報。相対評価で表す。

属性情報：相手の属性をあらわす情報。所属、氏名、電話番号、URLなどがある。

2.2 従来システム

上記の各情報のうち、関係強度情報及び属性情報を得るために、送受信した電子メールを利用するのは有効な手段である。電子メールでは、差出人と宛先および差出人の署名情報が記載されているため、上記情報が比較的容易に取得できる。

従来システムとして、受信した電子メールから人脈情報を取得し、管理するシステムがある[1]。しか

し、これらは電子メールの本文を参照せず、授受関係(From と To)だけから関係強度情報だけを獲得している。より精度の高い関係強度情報及び属性情報を得るには相手との社会的関係及び属性情報を手作業で設定する必要がある。

2.3 アプローチ

提案する人脈活用システムは、送受信した電子メールから精度の高い人脈情報を自動取得する。そのため、本システムでは、宛先の記述形態および本文の宛先情報・署名情報を解析する手段を設け、差出人が考える宛先に対する関係強度を反映させた人脈情報を生成させる。

3 人脈活用システム

3.1 システム概要

人脈活用システムは、以下のモジュールから構成される。

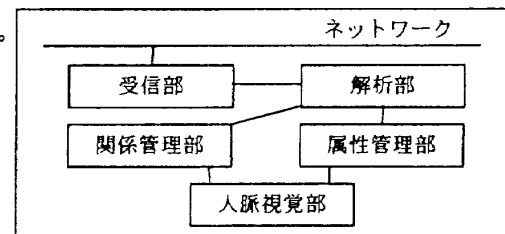


図1 モジュール構成

受信部は、相手から来た電子メールを受信し、フォルダに蓄積する。解析部は、電子メールのヘッダ記述形態・本文における宛先記述形態・本文下部の署名記述形態を解析し、差出人と宛先の関係付けを行い、署名情報を抽出する。関係管理部は、上記関係付けを集計し人脈強度を数量化する。属性管理部は、相手の署名情報を複数管理する。人脈視覚部は、数量化された人脈強度を基に人脈をノードリンク形式で表示する。

3.2 人脈情報解析方式(解析部)

関係強度情報算出の基本単位となる差出人と宛先の関係を以下のように記号化する。

差出人(アドレス)ー宛先(アドレス)：関係記号

Email-Based Personal Connection Analysis System.

Kenji Yoshifu, Tetsuji Tanigawa
Human Media Res.Labs., NEC Corporation
8916-47, Takayama-Cho, Ikoma, Nara, 630-0101,
Japan

(1) ヘッダ情報解析

宛先記述場所(To, Cc)及び宛先記述順位から、宛先記述場所に記載されたアドレスと差出人アドレスの関係を記号化する。例えば、To:yoshifu,tanaka Cc:yamada From: tanigawa の場合、tanigawa-yoshifu:To1-2(記述場所 To 記述順位 1-記述総数 2)、tanigawa-tanaka : To2-2、tanigawa-yamada : Cc1-1となる。

(2) 本文情報解析

・差出人名抽出

本文先頭から数行以内で「○○と申します」などの自分を名乗る表現がある部分を抽出し、上記○○を差出人名とする。所属を表す単語「@」、「の」などが入っている場合は所属名も抽出する。抽出内容は人物属性の氏名、所属として保存される。

・宛先名抽出

接頭語 (>のような宛先を示す記号)、接尾語 (様、さん、課長等のような敬称、役職を表す語) を抽出し、接頭語に続く単語及び接尾語の前の単語を宛先人名とする。引用部分(○○は書きました等の前置き表現や○○>など引用記号に宛先名をつけている表現)を抽出し、上記○○を宛先人名とする。

・関係強度情報

上記で抽出された宛先名と差出人のアドレスとの関係を記号化する。本文が「吉府@NECです。...>田中殿」と記述されている場合、吉府-田中: H(Honbunの頭文字)1-1となる。

(3) 署名情報解析

本文の最下行から数行以内でセパレータ記号(-, *)が連続している行を検索し、その行(複数存在すれば最上位行)から最終行までを署名情報とする。セパレータ記号が連続して存在していない場合、最終行から上数行までを署名情報とする。この署名情報から個人属性情報である氏名、住所、電話番号、URLなどを抽出する。抽出方法は各属性のキーワードである〒, TEL, URL, http などに続く値をそれぞれの属性値とする。氏名、所属などは名前、会社名DBを参照しながら決定する。

3.3 人脈強度算出方式(関係管理部)

人脈情報抽出の対象となるメール文書を選択し、

それらに登場する差出人と宛先の関係記号の数をすべて集計する。表1に差出人Aと宛先Bの関係記号集計例について示す。

表1 関係記号集計例(差出人A-宛先B)

| 番号 i | 記号 k_i | 集計数 $R(k_i)$ | 重み $W(k_i)$ | $R(k_i) \times W(k_i)$ |
|--------|-------------|--------------|-------------|------------------------|
| 1 | To1-1 | 15 | 1.0 | 15 |
| 2 | To1-2 | 10 | 0.8 | 8 |
| ... | ... | ... | ... | ... |
| n | H n - n | 3 | 0.5 | 1.5 |

集計した各関係記号の数を $R(k_i)$ 、各関係記号の重みを $W(k_i)$ ※とすると、人脈強度値 V は以下のようにして求められる。

$$V = \sum_{i=1}^n (R(k_i) \times W(k_i))$$

※ $W(k_i)$ は任意に設定可能。初期値は To1-k > Cc1-k, H1-k > H2-k のように設定される。

3.4 人脈情報視覚化方式(人脈視覚部)

人脈強度値を基に人脈マップを生成する(図2)。

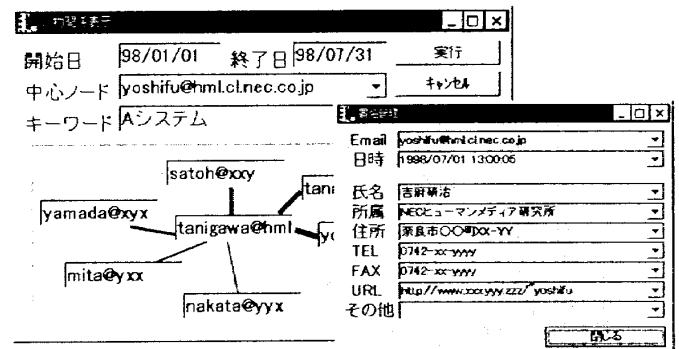


図2 人脈マップと署名情報

人脈マップは、受信者・発信者をノードで表現し、ノード間をリンクで表現する。人脈強度値が高いほどノード間の距離を短く、又はリンクを太く表現する。ノード選択時には署名情報から得られた各属性情報を表示する。

4 おわりに

本稿では、電子メール通信履歴を用いた人脈活用システムについて提案した。本システムにより、受信した複数の電子メールを利用して、精度の高い人脈情報の整理を自動的に行うことができる。

参考文献

- [1] 相曾ほか, 協調エージェントを用いた人脈活用支援システムの構築, 信学技報, OFS96-15(1996)