

コミュニティシステムのためのパーソナルネットワークの 利用とその分析

Application and Analysis of Personal Networks for a Community System

濱崎雅弘^{*1*2}
Masahiro HAMASAKI

武田英明^{*2*1}
Hideaki TAKEDA

大向一輝^{*1*2}
Ikki Ohmukai

市瀬龍太郎^{*2*1}
Ryutarō ICHISE

^{*1}総合研究大学院大学
The Graduate University for Advanced Studies

^{*2}国立情報学研究所
National Institute of Informatics

In this paper, we discuss importance and utilization of personal network in a community system through the result of management and analyzing of our scheduling support system for an academic conference. One of the features of our system is the sharing model to support information exchanging among participants and information discovery. Our system realizes to generate participant's personal networks and provide information sharing control using it. We managed this system in the academic conference called JSAI2003. We obtained 276 users and their personal networks. As a result of analysis of the personal networks, we found that personal networks can promote information exchanging among participants since personal network showed existence of participants to the others and personal networks can support information discovery since recommendation using personal networks can provide appropriate information to participants.

1. はじめに

本発表では、コミュニティシステムにおけるパーソナルネットワークの有用性および重要性について、実際に運用したコミュニティシステムの利用ログの分析結果を元に述べる。

本論文では、2003年度人工知能学会全国大会にて運用したパーソナルネットワークを利用したコミュニティシステムの構築およびその運用結果を分析して、コミュニティシステムにおけるパーソナルネットワークの重要性および有用性について論じる。

学術会議に参加する目的は、自身の興味に合った発表を聴講することや関心の合う他の参加者と交流を行うことと考えられる。そのためには興味のある発表や参加者を見つけることが重要になるが、ある程度の規模の学術会議になると発表件数も参加者数も数百となり、その中から自身が興味のある発表・参加者を見つけ出すのは困難である。

本研究では、これらの問題を解決するためのシステムとしてパーソナルネットワークを用いたコミュニティシステムを提案・開発し、問題解決に向けて実際に運用した。パーソナルネットワークを用いたコミュニティシステムとは、参加者間のインタラクションを促進させるために知り合い関係という新しい情報源を取り込んだ情報共有システムを指す。

提案するコミュニティシステムでは、パーソナルネットワークを情報へのアクセス経路として、また、情報推薦のための情報源として用いた。本論文では運用結果を元に、それらパーソナルネットワークを用いたサービスについて分析を行う。

2. パーソナルネットワークを用いたコミュニティシステム

2.1 システム概要

本システムは学術会議のオンラインプログラムをベースとした、動的に Web コンテンツを提供するシステムである。シ

ステムは著者と発表論文とセッション、さらにスケジュール表の 4 種類の HTML ページを動的に生成する。HTML ページ間はデータベースに格納された関係に基づいてリンクが張られており、利用者は生成された HTML ページを自由に閲覧することができ、さらに新しい関係を追加することもできる。この利用者によって追加された関係を元に個人用スケジュール表が変化していく。

システムは MySQL データベースと Perl で記述されたプログラムにより構成される。データは全て MySQL で管理され利用者からのアクセスは CGI プログラムが受け付ける。利用者は Web ブラウザを用いてシステムにアクセスする。スケジュール支援システムとしてのシステム詳細および運用結果については [濱崎 04] にて述べているので、ここでは割愛する。

2.2 リンク追加機能

基本機能として、個人用スケジュール表と友人リストの作成機能を提供した。利用者は興味のある発表論文を会議スケジュールや論文一覧から発見して、自分のスケジュールに加えることができる。すると自分のスケジュール表が更新される。同様に知っている人を発見して知り合いであると指示すると、自分および相手の友人リストに追加され、各々のページに表示される。これらはそれぞれ Check リンク、Know リンクを追加することによって行われる。

追加されたリンクは一種のプライバシー情報であるので、アクセス制御が可能であることが望ましい。本システムでは生成されたネットワークを用いてアクセスコントロールを行う。具体的には、自分が友人として登録されている場合のみ、その人の詳細情報にアクセスでき、自分がある論文をスケジュールに取り込んでいるときのみ、その論文に興味を持った他の人たちを知ることができる。この方法は、アクセス制御とコミュニティ形成の支援を同時に行うことができる。

2.3 推薦機能

リンク追加による個人用ページの生成に加えて、利用者が追加したリンクを利用した情報推薦サービスを行った。この推薦サービスでは、発表論文だけでなく人の推薦も行われる。利

連絡先: 濱崎雅弘, 国立情報学研究所 大学院, 〒 101 東京都千代田区一ツ橋 2-1-2, 03-3630-2681, 03-3630-2681, hamasaki@grad.nii.ac.jp

用者はブラウジングだけでなく、推薦によっても興味ある発表論文や人の発見を行うことができる。

本サービスでは、人と論文の 2 種類のリソースを推薦した。さらに推薦には 2 種類の方法を用いた。一つは GroupLens[Konstan 97] 等の協調フィルタリング的手法である。これは同じような関心を持っている人のデータを参考に推薦を行う。論文推薦の場合、利用者が Check リンクの張った論文に対して、同様に Check リンクを張っている人たちがどの論文に Check リンクを張っているか、という情報を利用する。もう一つはパーソナルネットワークを利用する手法であり、知り合いが持つデータを参考に推薦を行う。具体的には Know リンクを張っているまたは張られている相手を持っているリンクを用いる。本論文では、前者の Check リンクを用いた推薦を関心類似型推薦、後者の Know リンクを用いた推薦をパーソナルネットワーク型推薦と呼ぶ。

このように、2 種類のリソースを 2 種類の手法で、合計 4 種類の情報推薦を行った。

推薦するリソースの見つけ方について説明する。人 h_0 が論文 p に対して Check リンクを張っている場合は $Check(h_0, p) = 1$ 、そうでない場合は 0 を取る $Check$ 関数、人 h_0 が人 h_1 に対して Know リンクを張っている場合は $Know(h_0, h_1) = 1$ 、そうでない場合は 0 を取る $Know$ 関数を定義する。さらに $Know(h_0, h_1) = 1$ または $Know(h_1, h_0) = 1$ の時にのみ $Relate(h_0, h_1) = 1$ となる $Relate$ 関数を定義する。

$V_{hc}(h_0, h_x)$ と $V_{pc}(h_0, p_x)$ は人 h_0 に対する関心類似型推薦における、人 h_x および論文 p_x の推薦度をそれぞれ計算する。

$$V_{hc}(h_0, h_x) = \sum_{\{p|Check(h_0, p_x)=1\}} Check(h_x, p_x)$$

$$V_{pc}(h_0, p_x) = \sum_{\{h|V_{hc}(h_x, h_0)>0\}} Check(h_x, p_x)$$

$V_{hk}(h_0, h_x)$ と $V_{pk}(h_0, p_x)$ は人 h_0 に対するパーソナルネットワーク型推薦における、人 h_x および論文 p_x の推薦度をそれぞれ計算する。

$$V_{hk}(h_0, h_x) = \sum_{\{h_k|Relate(h_0, h_k)=1\}} Relate(h_k, h_x)$$

$$V_{pk}(h_0, p_x) = \sum_{\{h_k|Relate(h_0, h_k)=1\}} Check(h_k, p_x)$$

今回の推薦サービスでは、それぞれ 4 種類の推薦ごとに推薦度 V が高い上位 5 つを取り出し、そのうち推薦度 V が閾値 (今回は 2) 以上のものを推薦情報として利用者に提示した。

3. 分析結果

提案するコミュニティシステムを 2003 年 4 月 30 日から、人工知能学会全国大会開催の 6 月 23~27 日を挟んで 6 月 30 日までの 2ヶ月間運用した。本章では、この間の利用ログの分析結果を述べる。

3.1 利用者によるリンク構造の利用の分析

本節では、利用者のページの移動パターンから、人のネットワークがどの様に活用されたかを示す。

HTTP サーバのアクセスログから利用者の移動パターンを解析した*1。図 1 は各種ページ間の利用者による移動回数のうち値が十分大きいもののみを示している。

図から、スケジュールからセッション、そして論文のページへと移動するのが大きな流れとなっていることがわかる。それぞれ 2~3000 件程度の移動履歴が残っていた。スケジュールの中からセッションを選び、セッションの中から論文を選ぶという流れは典型的なオンラインプログラムの閲覧の仕方であるといえる。論文のページから人のページには 1000 件近くの移動履歴があり、逆も同様であった。さらに人のページから人のページへの移動も 1300 件程度あることがわかった。

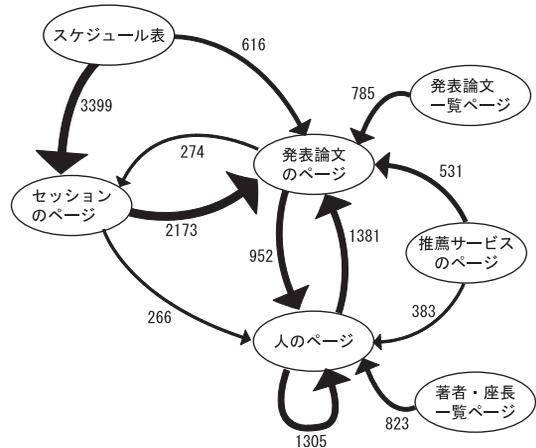


図 1: アクセス遷移図

セッションから論文への一般的な移動パターンの半分に相当する回数のアクセスが、人のページを介して行われた。また、人から人への移動もそれに相当する回数行われた。前者は論文 (のページ) へのリンクを持った人 (のページ) というコンテンツが利用者に多く利用されたことを示し、後者は人と人のネットワークが利用者に多く利用されたことを示している。これらの結果は、本システムの目的であった人のコンテンツ化およびネットワーク化による利用者の情報収集支援がなされたことを示している。

3.2 生成されたリンク構造の分析

本節では、利用者により追加されたリンクによって、どのようなネットワークが作られたのかを分析する。

人は Check リンクと Know リンクの始点、および Know リンクの終点となることができる。同様に、論文は Check リンクの終点になることができる。表 1 は Know リンクを持っている (追加したことがある) 人の数と、少なくとも 1 つ以上 Know リンク又は被 Know リンクを持っている人の数を示している。

Check リンクも Know リンクも共に、リンク追加を行った人は登録されている人に対して 3 分の 1 にも満たなかった。しかし論文では 95% が、人では 49% が、それぞれ Check リンクまたは Know リンクのネットワークに接続された。利用者がどこにリンクを張るかは自由であるが、結果的にこのように広範囲にわたる新しいネットワークを作ることができた。利用

*1 この手法では Web ブラウザで「戻る」ボタンを押して移動した場合は追跡できない。また、ブックマークやメールから URL を直接指定した場合にも追跡できない。このため図 1 の各ノードの入出力の和は 0 にならない

表 1: Know リンクとリソース

	持っている人	全体に占める割合
Know リンク	99 人	0.18
被 Know リンク	260 人	0.47

者によって追加されたリンクで作られたこのネットワークは新しい情報探索のルートを提供することになり、情報流通を促す機能を果たす。これはシステムの有用性を示す結果であるといえる。

本システムでは、人と人をつなぐリンクには Know リンクと共著関係との二種類がある。図 2 はそれぞれ Know リンクのネットワーク(図 2-a)と共著関係のネットワーク(図 2-b)を示している。Know リンクネットワークは 5 つのクラスタ(各クラスタに含まれるノード数はそれぞれ 262, 4, 3, 2, 2)から形成されているのに対し、共著関係で作られた人のネットワークは 73 のクラスタから形成されていた*2。



a) Know リンクネットワーク b) 共著ネットワーク

図 2: 2 種類のパーソナルネットワーク

Know リンクネットワークと共著関係ネットワークとで共有しているリンクは 135 個であり、Know リンクの約半数は共著関係からは見つけられないものであったことがわかる。Know リンクと共著関係を合わせたネットワークでは、クラスタ数は 19 個になった*3。

共著関係は人のネットワークにおける重要なリンクであるが、それだけでは断片的なネットワークを構成してしまい、Know リンクの追加によって初めて大きなパーソナルネットワークを構築することができた。Know リンクは、アクセス経路としてのパーソナルネットワークにとって重要であったと考えられる。

3.3 パーソナルネットワークを用いた推薦の分析

システムの推薦サービスによって推薦されたリソースが、利用者に受け入れられたかどうかを調べた。推薦は 2.3 節で説明した手法で計算した推薦度の高いリソースを列挙することで行われる。利用者は推薦リソースの名前をクリックすることで推薦リソースのページへ移動することができる。そこでリンク追加ボタンをクリックすれば、Check/Know リンクが生成される。推薦リソースのページにアクセスしたということは、そのタイトルを見てある程度の関心を持ったことであり、推薦結果の弱い受理と言える。さらにその推薦リソースに対してリ

ンクを生成した場合は、推薦されたリソースに対して強い関心があったといえ、これは推薦結果の強い受理と言える。

この方法は Cosley らの論文 [Cosley 02] で”Click”と”Buy”という指標に相当する。図 2 に結果を示す。なお、推薦数は正確に集計できなかったので参考値として幅を持って示した*4

Cosley は推薦においては多様なアイテムの提示が重要であるとして、ResearchIndex*5を対象とした推薦システム REFFEREE では 9 種類の推薦を調査している [Cosley 02]。9 種類のうちアクセス率の高かった上位 5 つの推薦手法のアクセス率は 2.00 ~ 4.63%、追加率(Cosley のシステムではファイルをダウンロードした割合)は 0.63 ~ 2.15%、追加率/アクセス率は 26.2 ~ 32.0% であった。この結果から、パーソナルネットワーク型推薦は複数の推薦手法の一つとして用いるには十分な性能があることがわかる。なお、追加率/アクセス率が REFFEREE と比較して悪いのは、ファイルダウンロードよりもスケジュール追加の方が心理的負荷が高いためではないかと思われる。

次に、関心類似型推薦とパーソナルネットワーク型推薦との推薦結果について比較分析をした。図 3 は両手法で推薦された論文のうち、利用者がアクセスした論文および Check リンクを追加した論文が持つ、被 Check リンク数を示している。

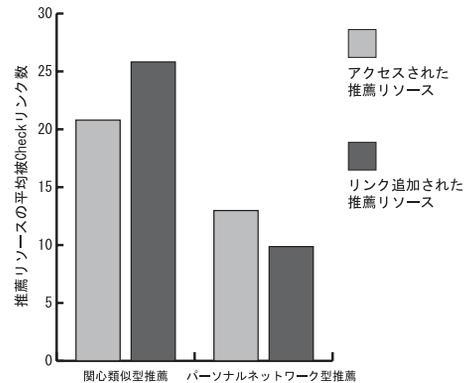


図 3: 推薦手法と被リンク数の関係

グラフを見ると、アクセスされた推薦コンテンツが持つリンク数は関心類似型推薦の方が多い。興味深いことにリンク追加の場合はさらにこの傾向が強化され、一方は増え、一方は減るという結果になっている。

関心類似型推薦では全体で人気のあるリソースが推薦されがちであるのに対し、パーソナルネットワーク型推薦では自分の周りで人気のあるリソースが推薦される。それらは必ずしも全体においても人気のあるリソースであるとは限らない。そして推薦を受けた側もその傾向を上手く使い分けられていることがうかがえる。

パーソナルネットワーク型推薦は、パーソナルネットワークによって推薦時に用いる情報が絞り込まれる。そのために利用者が現時点で持っている情報と似たものを見つける傾向よりはむしろ、利用者の興味に近いが見つけ損ねたような情報を見つけ出す傾向があると考えられる。この点が図 3 に現れたような、全体的に人気のないものを受け入れるという結果を引き出

*2 ノード数は 462, リンク数は 885。

*3 共著ネットワークには Know リンクネットワークにはないノードがあるため、Know リンクネットワーク単体よりもクラスタ数が多くなっている。

*4 今回のシステムは Check リンク, Know リンクの変化に応じて推薦アイテムが適時変化するため、いくつかの異なるアイテムを推薦したかを集計するのは難しい。今回は推薦ページのアクセス回数を上限、推薦ページにアクセスしたユーザ数 (IP 数) を下限として、その値に 5(最大推薦アイテム数) をかけたものを用いる。

*5 <http://citeseer.ist.psu.edu/>

表 2: 論文推薦に対するアクセス数とリンク追加数

	アクセス数(アクセス率)	リンク追加数(リンク追加率)	追加数/アクセス数
関心類似型推薦	347(6.6~26.4%)	72(1.4~5.6%)	20.7%
PN 型推薦	210(2.1~16.4%)	23(0.3~1.2%)	10.9%

PN 型推薦: パーソナルネットワーク型推薦

したのだと思われる。

推薦システムは単に利用者が知るべき情報を発見し提示するよりは、利用者の意思決定(この場合はどの様な論文を読むかということ)を支援するものであると考える [Cosley 02]。そのためには、理由を示しつつ様々な推薦をする事は有意義であるといえる。今回のパーソナルネットワーク型推薦は、この面から見ると図 3 が示すように通常の関心類似型推薦とは違ったアイテムの推薦が行え、かつそれがある程度受け入れられたという点で、新しい選択肢として価値があると思われる。

3.4 パーソナルネットワークの推薦能力の考察

パーソナルネットワーク型推薦の推薦能力について今回の運用結果に基づいて考察する。

利用者 h_0 が Know リンクを張っている先の人たちを $H_k = \{h | Relate(h_0, h) = 1\}$ とする。 h_0 が Check リンクを張っている論文を $P_0 = \{p | Check(h_0, p) = 1\}$ 、 H_k が Check リンクを張っている論文を $P_k = \{p | Check(h, p) = 1, h \in H_k\}$ とする。このとき $P_0 \cap P_k$ が利用者 h_0 とその知り合い H_k が共に Check リンクを張っている論文になる(図 4)。

今回の結果では、各個人が Check リンクを張った論文に対して、自分が Know リンクを張っている人のうち少なくとも 1 人が Check リンクを張っている割合は $\frac{|P_0 \cap P_k|}{|P_0|} = 0.67$ であった。

また、各個人が持っている Check リンクの平均値は $\overline{|P_0|} = 12.4$ 、そして各個人が Know リンクを張っている先の人たちが Check リンクを張っている論文の数の平均値は $\overline{|P_k|} = 52.8^{*6}$ であった。

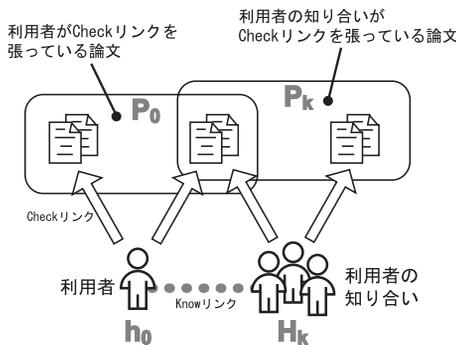


図 4: 利用者とその知り合いが持つ Check リンクの関係

H_k が Check リンクを張っている論文 B が、利用者 h_0 に Check リンクを張られている割合は、 $\frac{|P_0 \cap P_k|}{|P_k|} = \frac{|P_0 \cap P_k|}{|P_0|} \times \frac{|P_0|}{|P_k|} = \frac{12.4 \times 0.67}{52.8} = 0.16$ となる。ここから、パーソナルネットワーク型推薦によって推薦された論文 B が利用者 h_0 にとって Check リンクを張る対象となる(すなわち利用者 h_0 にとって

関心のある)論文であるかどうかの確率はだいたい 16%であることが予想できる。

同様に人を推薦する場合も考える。Check リンクを Know リンクに置き換えると、人推薦の場合になる。この場合、 P_0 に相当するのが H_k 、 P_k に相当するのが $H_{kk} = \{h | Relate(h', h) = 1, h' \in H_k\}$ である。自分が Know リンクを張った人に対して Know リンクを張っている割合は $\frac{|H_k \cap H_{kk}|}{|H_k|} = 0.76$ であった。各個人が持っている Know リンクの平均値は $\overline{|H_0|} = 8.5$ 、各個人が Know リンクを張っている先の人たちが Know リンクを張っている人の数の平均値は $\overline{|H_{kk}|} = 42.5$ であった。従って知り合いの知り合いが自分の知り合いであった確率は $\frac{|H_k \cap H_{kk}|}{|H_k|} = \frac{|H_k \cap H_{kk}|}{|H_k|} \times \frac{|H_k|}{|H_{kk}|} = \frac{8.5 \times 0.76}{42.5} = 0.15$ となり、だいたい 15%であることが予想できる。

4. まとめ

本論文では、2003 年度人工知能学会全国大会にて実際に運用し、得られた利用ログの分析結果から、コミュニティシステムにおけるパーソナルネットワークの有効性について検証した。提案システムでは、パーソナルネットワークを情報へのアクセス経路の一つとして、情報推薦を行う情報源の一つとして、それぞれ利用した。利用ログを分析した結果、知り合い関係によるパーソナルネットワークの構築は、アクセス経路の一つとしてのパーソナルネットワークに貢献することがわかった。また、推薦においては協調的フィルタリングとは異なる性質を持った推薦が行えることがわかった。

最近では Orkut^{*7}や Mixi^{*8}など多くの Social Networking サービスが提供され、パーソナルネットワークをオンライン上に構築することは珍しいことではなくなりつつある。しかしながら、これらはシステム内で閉じた物であり、ある程度以上の広がりを求めることは難しい。これは提案システムにも言えることである。今後は FOAF [Brickley] などを用いてパーソナルネットワークをシステムに閉じない形で保存し、活用することを考えていきたい。

参考文献

- [Brickley] Brickley, D., et al.: Friend of a Friend: <http://www.foaf-project.org/>
- [Cosley 02] Cosley, D.: REFEREE: An open framework for practical testing of recommender system using ResearchIndex, in *Proceedings of the 28th VLDB Conference* (2002)
- [Konstan 97] Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gorden, L. R., and Riedl, J.: GroupLens: Applying collaborative filtering to usenet news., *Communications of the ACM*, Vol. 40, No. 3, pp. 76-87 (1997)
- [濱崎 04] 濱崎 雅弘, 武田 英明, 大向 一輝, 市瀬 龍太郎: 2003 年度人工知能学会全国大会スケジューリング支援システムの開発と運用, 第 18 回人工知能学会全国大会論文集 (2004)

*7 <http://www.orkut.com/>

*8 <http://mixi.jp>

*6 重複があるのでこの値は $\overline{|H_k|} \times \overline{|P_0|}$ とは等しくならない。