

イントラ SNS における友人リストとの類似性に基づく友人推薦手法

韓 超[†]

小林 智也[†]

西本 一志[†]

近年ネット上のコミュニケーションツールとして SNS（ソーシャルネットワーキングサービス）が注目されている。SNS 上での人間関係は、基本的に現実世界のものを再現する程度に留っており、これまで面識がなかった人物と新たに関係を構築することが難しい。本稿では、SNS に登録された各ユーザのプロフィール情報や日記へのコメント情報などを利用し、あるユーザが登録している友人リストとの類似性に基づき、新しい友人を紹介する手法を提案する。学内イントラ SNS を構築・運用して得られたデータを用いて、提案手法の初期的検証を実施した。

A method for introducing new friends in an intra-SNS based on the similarities with the friend list

CHAO HAN[†]

TOMOYA KOBAYASHI[†]

KAZUSHI NISHIMOTO[†]

In recent years, SNS (Social Networking Service) has attracted more attention as a communication tool. Human relations in the SNS are usually a subset of those in the real world; it is difficult to construct new relations with unacquainted people. This paper proposes a novel method for introducing new friends based on the similarities with users' friend lists using personal profile data and comments to diaries. We conducted user studies using an intra SNS installed to our institute and evaluated the effectiveness of our proposed method.

1. はじめに

近年、SNS（ソーシャルネットワーキングサービス）や blog などの新しいネットワークコミュニケーションツールの利用が増加している。SNS は日常的なコミュニケーションの支援を目的として、コミュニケーション主体である個人の存在を明示化し、個人間の情報流通を実現するためのシステムであると定義できる[1]。大規模な SNS として、世界最大会員数を持つ Facebook¹、日本の mixi²や中国の RenRen³などが挙げられる。しかし、これらの既存 SNS は、新しい人間関係の構築を支援する機能がまだ十分ではなく、基本的に現実世界の人間関係を SNS 上に再現する程度に留っており、これまで面識がなかった人物と新たに関係を構築することが難しい[2]。

SNS を用いた新たな人間関係の構築は、特にある組織の内部で運用されているイントラ SNS において重要であると考えられる。たとえば筆者らが所属している北陸先端科学技術大学院大学（以下は JAIST と略する）には3つの研究科がある。研究科を跨いだ課外活動が少なく、一緒に受講できる授業もほとんどないの

で、研究科を超えて新しい人間関係を構築することが難しい。筆者らは、普段自分の周りに存在していない情報や知識を獲得することによって、革新的かつ効率的な知識創造ができると考える。だが、前述のような現状では、新たな人的交流と、その結果としての革新的な知識創造は生じにくい。

本稿では、各ユーザが登録した友人リストに基づき、各ユーザのプロファイル情報及び日記へのコメント付け情報を用いて、友人リストに登録されていないユーザの中から友達として最適と思われるユーザを推薦する手法を提案する。JAIST 内に構築したイントラ SNS をテストベッドとして、ユーザスタディにより、提案手法の有効性を評価する。

2. 先行研究

SNS 上での新しい人間関係構築を支援する研究として以下の2つが挙げられる。鳥谷部らが提案した TTT システムは、ユーザの近況を個々にデータとして蓄積し、知人に限定して公開する。さらに、ユーザに対しフィードバックを促しコミュニケーションを活性化させることで、継続的に多数の知人と近況を共有することを支援する。さらに、友達との会話に別の友達を呼ぶ仕組みを用い、初対面のユーザ間のコミュニケーションを共通の知人が仲介することによって信頼感のある出会いを支援している[3]。春日らは、SNS

[†] 北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

¹ <http://www.facebook.com/>

² <http://mixi.jp/>

³ <http://www.renren.com/>

に「自分」「友達（仲介者）」「友達の友達」まで参加できる仲介型のチャットを組み込んだ新しい SNS を提案した[4].

これらは共通の友人によって、新たな人間関係を構築することを目指している。しかし、ある人と自分とが共通の友人を持っているということだけでは、その人と自分が友人になりたいと思うには十分とは言えない。例えばその人と自分とが何らかの共通点を持っているなどの、更なる関係性が必要と考えられる。よって単純に共通の友人を介する手段は、新たな友人の発見手段としては十分なものではないと考える。

谷川らは、筆者らの研究と同様、推薦を受けるユーザ周辺のリンク構造とコンテンツ（友達紹介文）を用いて新たな友達を紹介する方法を提案している[5][6]. ただしこの手法で推薦される候補は、友達の友達に限定されている。後述する筆者らの提案手法は、嗜好などの個人内情報や日記へのコメント数から、自分の友人と似ている人を探す手法であり、推薦候補が友達の友達に限定されずそれ以外の人達も推薦候補となる。

3. 提案手法

3.1 提案手法の概要

本研究の基本的な発想は、すでに自分の友人である人と似ている人なら友人になれる可能性が高いだろうというものである。ただし特定の友人 1 人だけを見ても、その友人が持つ属性のどれが重要かを判断することは一般に難しい。そこで我々は、自分の友人を集合的に扱い（すなわち、個々の友人ではなく、友人リストを対象とする）、友人達が持つ属性を合成することにより、自分の友人になりやすい人が持つ属性を浮き彫りにできると考えた。

SNS には利用できる情報が 2 種類ある。個人内情報と個人間情報である。個人内情報とは SNS に登録した個人のプロフィール情報などの、個人に閉じた情報を指す。個人間情報とは、日記へのコメントなどの SNS 上での複数のユーザ間の関係に基づく情報を指す。本研究ではこの 2 種類の情報をそれぞれ利用し、個人内情報を利用する推薦方法と個人間情報を利用する推薦方法を提案する。

3.2 個人内情報を利用する推薦方法

個人内情報を利用する方法では、ユーザ U_p が最初にプロフィール情報として登録した、性別や出身地、趣味等の属性情報を用い、ある属性 a_i の要素値 δ_i を、その属性を持つ場合は 1、持たない場合は 0 として、各ユーザの特徴を表現するプロフィールベクトル

$\vec{A}^p = \{\delta_1, \delta_2, \dots, \delta_i, \dots, \delta_m\}$ (m は属性の総数) を生成する。ユーザ U_p に新たな友人候補を推薦する場合は、 U_p がすでに自分の友人リストに友人として登録している友人 F_k^p 全員のプロフィールベクトルを合成して（合成方法は後述する）、合成友人プロフィールベクトル $\vec{A}^{PF}$ を生成する。この合成友人プロフィールベクトルと、自分の友人リストに登録されていない人 NF_j （このような人を、以下では「非友人」と呼ぶ）のプロフィールベクトル $\vec{A}^{NF_j}$ （以下、非友人プロフィールベクトル）との余弦を求め、値が大きい NF_j を新たな友人候補として推薦する。

合成友人プロフィールベクトル $\vec{A}^{PF}$ の求め方を説明する。全ユーザの人数を N 、そのうちある属性 a_i を有している人数の合計を n_i 、新たな友人を紹介したいユーザ U_p の自分の友人リストに登録されている友人の数を N^{PF} 、 U_p の友人の中で属性 a_i を有している人数を n_i^{PF} とするとき、合成友人プロフィールベクトル $\vec{A}^{PF}$ における属性 a_i の重み A_i^{PF} を次式で定義する。

$$A_i^{PF} = n_i^{PF} - \left(\frac{n_i}{N} \cdot N^{PF}\right) \quad (1)$$

この式の右辺第 2 項は、全ユーザから任意の N^{PF} 人のユーザを取り出したとき、属性 a_i を持つ人数の期待値である。したがって、式(1)は属性 a_i を有する U_p の友人の数と期待値との差であり、 U_p の友人がある属性を平均よりも多く有する場合には正の値を、平均よりも少ない場合は負の値を取る。

非友人プロフィールベクトル $\vec{A}^{NF_j}$ の属性 a_i の重み $A_i^{NF_j}$ を式(2)で定義する。

$$A_i^{NF_j} = \begin{cases} 1 - \frac{n_i}{N}, & NF_j \text{ が属性 } a_i \text{ を有する場合} \\ 0, & NF_j \text{ が属性 } a_i \text{ を有しない場合} \end{cases} \quad (2)$$

つまり、全体的に属性 a_i を有している人数 n_i が少ないほど、 NF_j がこの属性 a_i を有する場合に重みが大きくなる。

3.3 個人間情報を利用する推薦方法

個人間情報を利用する方法では、ユーザ U_p が自分以外の全ユーザ U_j ($j \neq p$) の日記にコメントした数 c_j^p を要素値とするコメントベクトル $\vec{C}^p = \{c_1^p, c_2^p, \dots, c_p^p = 0, \dots, c_N^p\}$ を生成する。ユーザ U_p に新たな友人候補を推薦する場合は、 U_p がすでに自分の友人リストに友人として登録している友人 F_k^p 全員のコメントベクトルを合成して（合成方法は後述する）、合成友人コメントベクトル $\vec{C}^{PF}$ を生成する。この合成友人コメントベクトルと、非友人 NF_j のコメントベクトル $\vec{C}^{NF_j}$ （以下、非友人コメントベクトル）との余弦を求め、値が大きい NF_j を新たな友人候補として推薦する。

合成友人コメントベクトル $\vec{C}^{PF}$ の各要素の重み C_i^{PF} を

式(3)で定義する.

$$C_i^{PF} = \frac{1+S_i^{all}}{1+S_i^{PF}} \cdot \frac{\overline{C}_i^{PF}}{\overline{C}_i^{all}} \quad (3)$$

ここに, S_i^{all} は全ユーザを対象とした場合の U_i へのコメント数の標準偏差, S_i^{PF} は U_p の友人集合 PF を対象とした場合の U_i へのコメント数の標準偏差, $\overline{C}_i^{PF}$ は U_p の友達集合 PF を対象とした場合の U_i へのコメント数の平均, $\overline{C}_i^{all}$ は全ユーザを対象とした場合の U_i へのコメント数の平均である. 右辺の第 2 項は, PF に属する人が全体平均よりも多く U_i に対してコメントしている場合, U_i は PF にとって重要度が高い人であると考え, 重みを大きくする項である. ただし, 平均値が同じでも, PF の友人達が全員 U_i に対してコメントしている場合と, 特定の誰かだけがコメントしている場合とを比べると, 全員がコメントしている場合の重要度を高くすべきであると考え. ただし, PF の友人達が全員 U_i に対してコメントしている場合でも, 全ユーザもやはり U_i に対してコメントしている場合は, U_i は PF にとってのみ重要なわけではないので, 重要度を下げ必要がある. 右辺の第 1 項は, この考えに基づく重要度の補正項である.

非友人 NF_j のコメントベクトル $\overline{C}^{NFj}$ の各要素の重み C_i^{NFj} を式(4)で定義する.

$$C_i^{NFj} = \sqrt{\frac{C_i^j}{\overline{C}_i^j} \cdot \frac{C_i^j}{\overline{C}_i^{all}}} \quad (4)$$

ここに, C_i^j はユーザ U_j の U_i へのコメント数, $\overline{C}_i^j$ は U_j が全ユーザにつけたコメントの平均数, $\overline{C}_i^{all}$ は全ユーザが U_i につけたコメントの平均数である. つまり, この式は, U_j が自分の平均コメント数よりも U_i に対して多くのコメントを付け, かつ全ユーザの平均コメント数よりも U_j が U_i に対して多くのコメントを付けた場合に重みを大きくする.

4. 実験

まずデータを収集するために, 株式会社手嶋屋が中心に開発しているオープンソースの OpenPNE 3.0 を使用して学内イントラ SNS を構築した. OpenPNE は, 日記, アルバム, コミュニティ, メッセージ, 友達招待, システム表示言語の変更などの機能を持っている.

4.1 データ収集実験

2010 年 7 月 9 日に JAIST の学生全体に本 SNS を公開し, 現在も稼働中である. オープン制で, JAIST の学生ならだれでも参加できるように設定した. また,

日常の使用に耐える SNS を目指すために, 上記の OpenPNE が持っている機能以外に, つぶやき, 天気予報, RSS リーダー, 学校内で良く使うリンクとの連携機能を付加した.

OpenPNE には様々なプロフィール項目が登録可能だが, 本研究では以下の 10 個のプロフィール項目を用意した.

- 性別
- 血液型
- 研究科
- 専攻
- 職歴のありなし
- 出身地
- 趣味
- 就職希望業界
- 出身大学
- 研究室

この中で, 単一選択項目は性別, 血液型, 研究科, 専攻, 職歴のありなしと出身地である. 研究科と専攻は JAIST にある 3 つの研究科と 9 つの専攻に設定し, 出身地は日本全国の県と JAIST 在籍留学生の国籍を設定した. 多数選択項目は趣味と就職希望業界である. 趣味の選択項目は Wikipedia と mixi の趣味一覧を参考に 61 項目を設定した. 就職希望業界の選択項目は, 就職サイトのリクナビを参考に 25 項目を設定した. また, 出身大学と研究室はフリーフォーマットの入力項目である.

4.2 データ収集結果

本稿では 2010 年 12 月 8 日時点のデータを使用する. 以下の分析では本稿第 1 著者を除いたデータを用いる.

ユーザ登録者総数 N は 103 人 (男性 71 人, 女性 32 人), 友人リスト登録総数 363 人, 平均友人数は 3.52 人, 平均友人数以下の人の割合は 65.4% である. 友人リスト登録が 0 人の人数は 37 人である. 本研究では, 推薦を受ける者 (被推薦対象者) は, 友人 0 人のユーザを除いた 66 人のユーザである.

フリーフォーマットの入力項目は, 表記揺れ等を手動で統一した結果, 出身大学は 76 個, 研究室は 39 個であった. 収集した日記数は 235 個, 日記へのコメント総数 301 個, コメントを付けた人は 43 人であった.

4.3 推薦結果

3.2 で提案した個人内情報を利用するアルゴリズムを C# で実装し, 友人候補を算出した. また, 3.3 で提案した個人間情報を利用するアルゴリズムを用いた友人候補は, Microsoft Excel を用いて算出した.

評価では, 本来ならば友人候補をその被推薦対象者に紹介し, 友人となれるかどうかを判定してもらう必要がある. しかし, 全く知らない人を実際に友人とできるかどうかの判定には長い時間がかかるため, この方法は現実的ではない. そこで本研究では, 現実世界では友人であるが, 自分の友人リストには登録されて

いない人が友人候補として推薦されるかどうかをもとに評価を行う。ただし、被験者に自分の友人を列挙してもらうのは、倫理的に問題がある。このため本研究では、被験者と同じ研究室に所属する人は友人であることを仮定し、被推薦対象者と同一研究室に所属する者が友人候補となるかどうかを判定基準とする。

4.3.1 個人内情報を利用する推薦結果

3.2 のアルゴリズムで算出した余弦値をもとに、新たな友人の推薦を受ける 66 人それぞれについて、友人候補を 5 人求めた。次いで、各被推薦対象者に、5 人の候補のうちで自分と同じ研究室に所属する者をリストアップしてもらった。この結果、推薦候補の上位 5 位の中に平均 1 人は被推薦対象者と同一の研究室メンバーがいるのがわかった。

4.3.2 個人間情報を利用する推薦結果

他者の日記に対してコメントを付けた 43 人を推薦候補として、3.3 のアルゴリズムで算出した余弦値をもとに、被推薦候補者 66 人それぞれについて友人候補を 5 人求めた。次いで、各被推薦対象者に、5 人の候補のうちで自分と同じ研究室に所属する者をリストアップしてもらった。この結果、推薦候補の上位 5 人の中に平均 0.72 人（1 人未満）は被推薦対象者と同一の研究室メンバーがいるのが分かった。また、推薦候補に偏りがあり、推薦候補の 1 人が頻繁に推薦されてしまった。

5. 考察

自分の友人リストに登録されていない同一研究室メンバーが推薦されるかどうかを基準としているため、提案手法の有効性を厳密に評価することはできない。しかし、個人内情報と個人間情報のいずれを利用する推薦の方法においても、各被推薦対象者が友人リストに登録していない同一研究室所属者が上位 5 位以内に選ばれたことから、提案手法について一定の有効性が示唆されたと考える。ただし、その人数は平均 1 人かそれ以下と少ないため、提案手法の有効性を示すためには、さらなる検証が必要である。これは今後の課題とした。また、個人間情報を利用する方法の場合、特定の 1 人が頻繁に推薦された。これは、他者の日記にコメントをつける人が限られ、全体としてコメント数も少ないことが原因と考えられる。今後、コメント数を増やすことを検討したいと考えている。

6. まとめ

本稿では、SNS 上で得られる個人内情報と個人間

情報を用いて、新たな友人を紹介する方法を提案した。従来手法が共通の友達を介して新しい人間関係を構築するものであったのに対し、提案手法は自分の友達と似ている人を探して推薦する点が特徴である。OpenPNE を利用した SNS を筆者らが所属している大学院で運用し、2010 年 7 月から 12 月までに収集したデータを用いて推薦アルゴリズムを評価した。この結果、提案手法に関する一定の有効性が示唆された。

今後は、さらなる評価検証を進めたい。ユーザの友人リストから一部の友人を抜いたものを用いて推薦候補者を抽出し、その中に先に抜いた友人が含まれるかどうかを調査する予定である。また、個人間情報を利用する推薦手法の結果から見れば、コメントを増やす必要がある。さらに、Mislove らは、友人数が多いユーザほど友人数が少ないユーザと繋がる傾向があることを明らかにしている[7]。この知見に基づき、友人数が 0 人のユーザに、友人が多い人を推薦する方法を新たに考案したい。

参 考 文 献

- 1) 大向一輝：SNS の現在と展望—コミュニケーションツールから情報流通の基盤へ、IPSJ Magazine Vol.47 No.9 pp.993-1000, (2006).
- 2) 嶋田陽介, 加藤貴之, 廣嶋拓也, 豊田陽一, 萩野達也：共通の趣向を持つ利用者を発見するソーシャルネットワークキングシステム, 情報処理学会第 67 回全国大会, (2005).
- 3) 鳥谷部桜, 原口雄一郎, 村田大, 稲蔭正彦：TTT：人が仲介するコミュニケーション支援システム, 情報処理学会研究報告 GN, pp.43-47, (2004).
- 4) 春日章宏, 三枝優一, 古井陽之助, 速水治夫：SNS でのチャットによる友達の輪拡大支援システムの提案, 情報処理学会研究報告 GN, pp.61-66, (2007).
- 5) 谷川恭平, 大坪正範, 土方嘉徳, 西田正吾：SNS における友達推薦のための基礎検討, 第 50 回システム制御情報学会研究発表講演会, pp.363-364, (2006).
- 6) 谷川恭平, 大坪正典, 土方嘉徳, 西田正吾：FriendRank：SNS における友達推薦システム, 人工知能学会第 20 回全国大会(JSAI2006), 3D4-1, (2006).
- 7) Mislove,A., Marcon,M., Krishna, Gummadi,P., Druschel,P. and Bhattacharjee,B.: Measurement and Analysis of Online Social Networks. IMC Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, pp.29-42, (2007).