

学内交流促進を目的とした イントラ SNS における友人推薦手法の検討

韓超[†] 小林智也[†] 西本一志[†]

学内 SNS は、学内の人的交流を促進するためのプラットフォームとして有用である。本稿では、学内交流をさらに活性化することを目的として、新たな友人推薦手法を提案する。従来の SNS における友人推薦は、「友人の友人」という関係を用いるものが多い。しかしながら、友人の友人が自分の友人となる特徴を必ずしも持たない。本提案手法では、自分の友人と類似した人物を検索し、これを友人候補として推薦する。具体的には、自分の友人リストに登録されている各ユーザのプロフィール情報や日記へのコメント情報をもとにして、自分の友人が持つ共通の特徴を求める。この共通の特徴と、まだ友人リストに登録されていない各ユーザを比較し、類似性の高いユーザを新たな友人となる候補として推薦する。学内イントラ SNS を構築・運用して得られたデータを用いた実験を実施した結果、提案手法に一定の有効性が認められた。

Introducing new friends using intra-SNS to encourage intercommunication in an institute

Chao Han[†] Tomoya Kobayashi[†] Kazushi Nisimoto[†]

Intra-SNS is a useful platform to encourage intercommunication in an institute and an organization. This paper proposes a new method for recommending new friends to vitalize the intercommunication much more. Most of the ordinary methods for recommending new friends in the SNSs exploited “friends’ friends” relations. However, someone’s friend’s friend does not always have features that his/her friends commonly have. Our method recommends people who have similar features with his/her friends as friend-candidates. Concretely, based on profile data and comments to other users’ diary articles of users who are registered to his/her friend-list, we extract common features that all of his/her friends have. Then, we calculate similarity between the common features and each other user. Finally, people who have high similarity are recommended as friend-candidates. We installed intra-SNS into our institute and obtained data for estimating our proposing method. As a result, we confirmed basic effectiveness of our method.

1. はじめに

SNS (ソーシャルネットワーキングサービス) は、社会的ネットワークをインターネット上で構築するサービスのことである。日本では 2002 年頃から登場し、2004 年 2 月の「mixi」[1]、2006 年 2 月の「モバゲータウン」[2]のサービス開始をきっかけに、利用者が著しく増加している。この他の大規模な SNS としては、世界最大会員数を持つ Facebook[3]や中国の RenRen[4]などが挙げられる。地域振興を目的とした地域密着型の SNS も急増している[5]。大手企業各社や大学でも、社内や学内でのコミュニケーションの活性化や情報の地域間格差を解消するために SNS を導入しているケースも見られる[6]。このように、人と人とのコミュニケーションを促進させるツールとして SNS は一般化している[7]。これらの既存 SNS は、いずれも現実世界での人間関係を支えるとともに、新しい人間関係の構築も支援している。しかし、これらの 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[8]。

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。例えば、筆者らが所属している北陸先端科学技術大学院大学 (以下では 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。組織の壁を超えた交流により、普段自分の周りに存在していない情報や知識が獲得され、より革新的かつ効率的な知識創造を実現できる。しかしながら、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。

そこで本稿では、学内の人的交流を促進するため、イントラ SNS 上における新たな人間関係構築を支援する友人推薦手法を提案する。従来の友人推薦手法では、「友人の友人」という関係を用いるのが一般的であった。しかし、友人の友人は、自分の友人となる特徴を必ずしも持たない。共通の友人を有するという以外、全く共通性の無い人物である場合も多々あると思われる。本研究で提案する手法は、「自分の友人と似た特徴を持っている人物」を検索し、その人を友人候補として推薦するというものである。具体的には、自分の友人リストに登録されている各ユーザのプロフィール情報や日記へのコメント情報をもとにして、自分の友人が持つ共通の特徴を求める。この共通の特徴と、まだ友人リストに登録されていない各ユーザを比較し、類似性の高い

[†] 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

ユーザを新たな友人となる候補として推薦する。JAIST 内に構築したイントラ SNS をテストベッドとしてデータを収集し、提案手法の有効性を検証した。

2. 関連研究

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 3 つが挙げられる。鳥谷部らは、知人や知人が紹介する知人と近況を共有し、コミュニケーションを行うシステム「友達の友達は友達」(TTT) システムを提案した[9]。春日らは、SNS 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[10]。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[11][12]。ただしこの手法で推薦される友人候補は、友達の友達に限定されている。

このように、従来の手法は「友人の友人」という関係を用いるものがほとんどであった。これは、従来研究では一般向けの公衆 SNS を対象としていたためと考えられる。公衆 SNS の場合、全く無関係な人を友人候補として推薦されても、信用することが難しいが、自分が知っている友人の友人であれば、一応信頼に値するとみなせる。これに対し、本研究ではイントラ SNS を対象としているので、推薦される相手は、たとえ知らない人物であっても自分と同じ組織に属している。ゆえに、誰が推薦されても、その人物に一定の信頼を置くことができる。このため、友人の友人に限定しない提案手法によって友人候補を推薦しても、特に問題は無いと考える。

3. 提案手法

提案手法の基本的な発想は、すでに自分の友人である人と似ている人物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である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共通して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

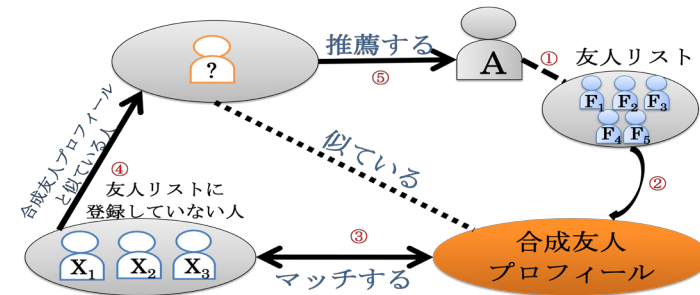


図1 個人内情報を利用する推薦方法

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。以下にこの 2 種類の推薦方法についてそれぞれ具体的に述べる。

3.1 個人内情報を利用する推薦方法

個人内情報を利用する推薦方法の概要を図 1 に示す例で説明する。ユーザ A の友人リストには F1 から F5 までの 5 人が登録されているとする。この 5 人の友人のプロフィール情報をすべて合成して、合成友人プロフィール情報を求める(合成方法は後述)。次いで、友人リストに登録していないユーザ(以下では非友人と呼ぶ。この例では X1~X3)の中で、A の合成友人プロフィール情報と似ている個人内情報を持っているユーザを探す。最終的に、X1~X3 の中で最も類似度が高いユーザを将来的に友人になる可能性がある人物として A に推薦する。以下、具体的な処理について説明する。

まず、各ユーザのプロフィールベクトルを求める。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、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。

次いで、あるユーザの友人リストに含まれるユーザ群のプロフィールベクトル群を、全ユーザのプロフィールベクトル群と比較して重み付けしたものを合成することにより、合成友人プロフィールベクトルを求める。全ユーザの人数を N 、その内のある属性 a_i を有している人数の合計を n_i とする。新たな友人を推薦する対象者を U_p とするとき、 U_p 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。このとき、合成友人プロフィールベクトル

$\overline{\mathbf{A}}^{\text{PF}} = \{w_1^{\text{PF}}, w_2^{\text{PF}}, \dots, w_i^{\text{PF}}, \dots, w_m^{\text{PF}}\}$ における属性 a_i の重み w_i^{PF} を次式で定義する:

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

式(1)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。

一方、友人リストには含まれていないユーザ（これを非友人と呼ぶ）のプロフィールベクトルを、全ユーザのプロフィールベクトル群と比較して重み付けすることにより、非友人プロフィールベクトルを求める。非友人プロフィールベクトル

$\overline{\mathbf{A}}^{\text{NFj}} = \{w_1^{\text{NFj}}, w_2^{\text{NFj}}, \dots, w_i^{\text{NFj}}, \dots, w_m^{\text{NFj}}\}$ の属性 a_i の重み w_i^{NFj} を式(2)で定義する。

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

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。

こうして得られた合成友人プロフィールベクトル $\overline{\mathbf{A}}^{\text{PF}}$ と、非友人プロフィールベクトル $\overline{\mathbf{A}}^{\text{NFj}}$ との余弦値を求め、余弦値がより 1.0 に近い非友人プロフィールベクトルを持つ非友人を、新規友人候補として選出し、 U_p に推薦する。

3.2 個人間情報を利用する推薦方法

個人間情報を利用する推薦方法の概要を図2の例で説明する。ユーザ A の友人リストには F_1 から F_5 までの 5 人が登録されているとする。この 5 人の友人がそれぞれ他者の日記に対してつけたコメントの情報を合成して、合成友人コメント情報を生成する（合成方法は後述）。次いで、非友人（ここでは X_1, X_2 ）の中で A の合成友人コメント情報と似ているコメント情報を持っているユーザを探す。この結果類似度が高い非友人を将来的に友人になる可能性がある人として A に推薦する。以下、具体的な処理について説明する。

まず、各ユーザのコメントベクトルを求める。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

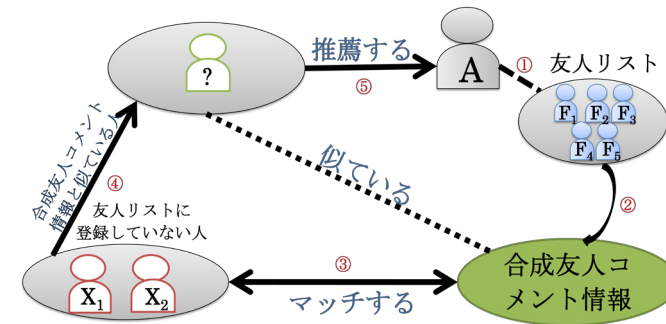


図2 個人間情報を利用する推薦方法

メントベクトル $\overline{\mathbf{C}}^p = \{c_1^p, c_2^p, \dots, c_i^p, \dots, c_N^p = 0, \dots, c_N^p\}$ を生成する。

次いで、あるユーザの友人リストに含まれるユーザ群のコメントベクトル群を、全ユーザのコメントベクトル群と比較して重み付けすることにより、合成友人コメントベクトル $\overline{\mathbf{C}}^{\text{PF}} = \{c_1^{\text{PF}}, c_2^{\text{PF}}, \dots, c_i^{\text{PF}}, \dots, c_N^{\text{PF}}\}$ を求める。 $\overline{\mathbf{C}}^{\text{PF}}$ の各要素の重み C_i^{PF} は、式(3)で定義する。

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

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、 $\overline{C}_i^{\text{PF}}$ 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、 $\overline{C}_i^{\text{all}}$ 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るべきであると考え。さらに、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、重要度を下げる必要がある。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。一方、非友人のコメントベクトルを全ユーザのコメントベクトルと比較して重み付

けすることにより、非友人コメントベクトルを求める。非友人 NF_j のコメントベクトル

$\overline{C}^{NFj} = \{c_1^{NFj}, c_2^{NFj}, \dots, c_i^{NFj}, \dots, c_N^{NFj} = 0, \dots, c_N^{NFj}\}$ の各要素の重み c_i^{NFj} を式(4)で定義する。

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

ここに、 c_i^j はユーザ U_j の U_i へのコメント数、 \bar{C} は U_j が全ユーザにつけたコメントの平均数、 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。

こうして得られた合成友人コメントベクトルと非友人コメントベクトルとの余弦値を求め、余弦値がより1.0に近い非友人コメントベクトルを持つ非友人を新規友人候補として選出し、 U_p に推薦する。

4. 個人内情報と個人間情報のデータ収集

提案手法を評価するデータを収集するために、(株)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。本システムは2010年7月9日にJAISTの学生全体に公開し、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。

4.1 学内イントラSNSの機能と設定

本SNSはOpenPNEが備えている日記、アルバム、コミュニティ、メッセージ、あしあと、友人招待、メンバー検索、システム表示言語の変更などSNSの一般的な機能を利用している。それに加えて、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を登録可能だが、個人内情報の類似度の算出には以下の10個のプロフィール項目を利用した。

- | | |
|-----------|----------|
| • 性別 | • 出身地 |
| • 血液型 | • 趣味 |
| • 研究科 | • 就職希望業界 |
| • 専攻 | • 出身大学 |
| • 職歴のありなし | • 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。複数選択項目は、「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

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356人、平均友人数は3.46人、平均友人数以下の人の割合は67.0%である。そのうち、友人リスト登録が0人の人数は37人であり、全体での割合は35.9%、友人リスト登録が1人の人数は19人であり、全体での割合は18.4%であった。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人のユーザとした。

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、出身大学数は76、研究室数は39であった。収集した日記数は235、日記へのコメント総数301、コメントを付けた人は43人であった。この43人を、個人間情報を利用する推薦方法において推薦される者(推薦候補者)とする。

5. 評価実験

3.1節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、3.2節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、本来ならば選出された友人候補をその被推薦対象者に紹介し、友人とさ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、本研究では以下の2つの近似的な基準に基づく評価を行った。

5.1 評価1—友人リストに登録されている友人の類似度と非友人の類似度の比較

5.1.1 評価手法

本評価方法では、すでに友人リストに登録されている友人を順番に1人ずつ友人リストから除外し、この除外された友人と、友人リストの残りの友人群との類似度(以下では友人類似度と呼ぶ)が、非友人に対する類似度(以下では非友人類似度と呼ぶ)よりも全体的に高くなるかどうかを評価する。

評価は、以下の手順で実施する。

1. 非友人類似度の導出

- A) 3.1 節で説明した方法で、友人リストを持つ各ユーザについて合成友人プロフィールベクトルを求める。
- B) 求めた合成友人プロフィールベクトルと、すべての非友人の非友人プロフィールベクトルとの余弦値（すなわち非友人類似度）を求める。
- C) 友人リストを持つすべてのユーザについて非友人類似度を求め、プロフィールベクトルに基づく非友人類似度集合を得る。

なお、コメントベクトルについても 3.2 節で説明した方法に基づき同様に処理し、コメントベクトルに基づく非友人類似度集合を得る。

2. 友人類似度の導出

- A) 友人リストを持つ各ユーザについて、
 - (ア) 友人リストに含まれる友人 1 名を友人リストから除外する。
 - (イ) 除外されなかった友人達について、3.1 節で説明した方法で合成友人プロフィールベクトルを求める。
 - (ウ) 除外した友人と、(イ) で求めた合成友人プロフィールベクトルとの余弦値（すなわち友人類似度）を求める。
 - (エ) 以上の処理を、友人リストに含まれるすべての友人を順番に除外して実施する。
- B) 友人リストを持つすべてのユーザについて上記手順で友人類似度を求め、プロフィールベクトルに基づく友人類似度集合を得る。

なお、コメントベクトルについても 3.2 節で説明した方法に基づき同様に処理し、コメントベクトルに基づく友人類似度集合を得る。

3. 友人類似度と非友人類似度との比較

- A) 非友人類似度集合に含まれるすべての非友人類似度の平均値と分散を求める。
- B) 友人類似度集合に含まれるすべての友人類似度の平均値と分散を求める。
- C) 上記の結果に対して t 検定を施し、友人類似度の平均値と非友人類似度の平均値に有意差があるかどうかを検定する。
- D) 検定の結果、友人類似度の平均値が非友人類似度の平均値よりも有意に高い値をとるならば、提案手法はすでに友人である者（除外された友人）に対して全体として高い類似度を与えることになるので、提案手法の基本的

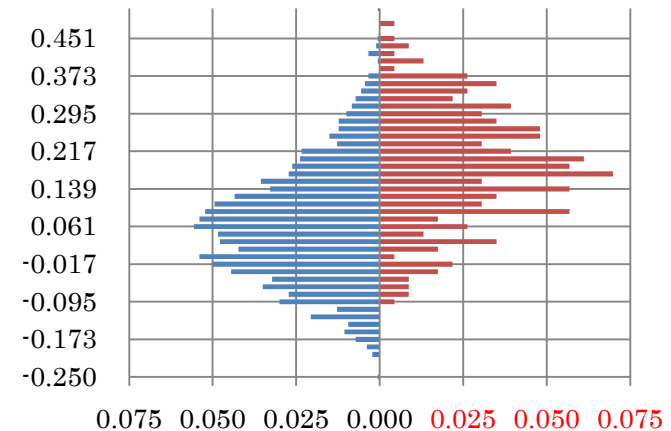


図 3 個人内情報から求めた非友人類似度と友人類似度の正規化度数分布．青色のグラフは非友人類似度の度数分布，赤色のグラフは友人類似度の度数分布である．

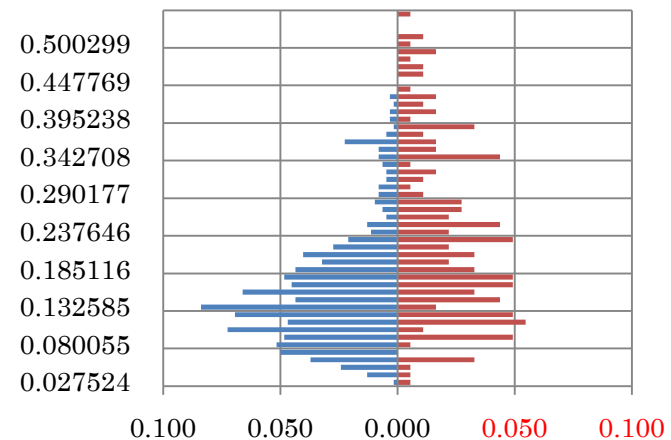


図 4 個人間情報から求めた非友人類似度と友人類似度の正規化度数分布．青色のグラフは非友人類似度の度数分布，赤色のグラフは友人類似度の度数分布である．

な有効性が示される。

5.1.2 評価結果

友人リストに登録されている友人数が少ない場合、プロフィールベクトルおよびコメントベクトルの合成による友人リストの特徴抽出がうまく機能しないと思われるため、本評価では友人リストに5人以上登録している20名のユーザを対象として評価値を求めることとした。

3.1節で説明した個人内情報（プロフィールベクトル）を用いて求めた非友人類似度と友人類似度の正規化度数分布を図3に示す。左側の青色のグラフは非友人類似度の度数分布で、右側の赤色のグラフは友人類似度の度数分布である。非友人類似度と友人類似度のそれぞれの平均値の差をt検定によって検定した結果、 $t(2025)=14.751$, $p < .01$ となり、友人類似度の平均が非友人類似度の平均よりも有意に高いことが示された。このことから、個人内情報を利用する提案手法は、全体的な傾向として有効であることが示された。

3.2節で説明した個人間情報（コメントベクトル）を用いて求めた非友人類似度と友人類似度の正規化度数分布を図4に示す。左側の青色のグラフは非友人類似度の度数分布で、右側の赤色のグラフは友人類似度の度数分布である。非友人類似度と友人類似度のそれぞれの平均値の差をt検定によって検定した結果、 $t(801)=10.236$, $p < .01$ となり、友人類似度の平均が非友人類似度の平均よりも有意に高いことが示された。このことから、個人間情報を利用する提案手法も、全体的な傾向として有効であることが示された。

5.2 評価2—被推薦対象者によるリストアップに基づく評価

5.2.1 評価方法

本評価方法では、被験者と同じ研究室に所属する人は友人であると仮定し、被推薦対象者と同一研究室に所属する者が友人候補として選出されるかどうかを判定基準とする。本来はこのような仮定はおかずに、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推薦候補として推薦されるかどうかを基準とすべきであろう。しかし、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ため、上記のような近似的基準を採用した。

本提案手法では、友人リストに友人を1人も登録していないユーザは被推薦対象者とできない。加えて、実名で登録していないと、同一研究室所属者をリストアップしてもらえないため、実名を登録していないユーザを評価対象者から除く。このため、ここでは友人を持っている実名で登録しているユーザ62人を評価1の評価対象者とする。さらに、友人リストに含まれる友人が全員誰にもコメントをつけていないようなユーザについては、個人間情報を用いた推薦を行えない。このようなユーザが1人いたため、個人間情報を用いる推薦手法に関しては、62人のうち61人に

対して評価を実施した。

評価は、以下の手順で実施する。

1. 各ユーザとそのユーザのすべての非友人（そのユーザの友人リストに登録されていない人々）について、3.1節あるいは3.2節で説明した方法で類似度を求める。
2. 各ユーザについて、1. で求めた類似度の値に基づき非友人をソートする。
3. 2. のソートの結果から、類似度が高い5人を推薦候補として選出する。
4. 各ユーザに対し、3. で選出した5人の推薦候補を提示し、その中から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う。
5. 推薦候補の中に何人の同一研究室所属者が含まれるかを、提案手法の有効性の指標とみなして評価する。

5.2.2 評価結果

3.1節で説明した個人内情報（プロフィールベクトル）を用いた推薦手法で選出された推薦候補者に関して、評価対象者62人に、推薦候補者のうちから自分と同じ研究室に所属している人をリストアップしてもらった。この結果、その総数は82人で、推薦候補の上位5位の中に平均1.32人の同一研究室メンバーがいることがわかった。この結果は、非友人集合からランダムに5人を選んだ場合に含まれる同一研究室所属者数の期待値よりも十分に大きい値であることを確認している。リストアップしてもらった結果を見ると、同じ研究室に所属している人を複数推薦されるユーザが多く、62人の内15人のユーザには同一研究室所属者が3人以上推薦されていた。

3.2節で説明した個人間情報（コメントベクトル）を用いた推薦手法で選出された推薦候補者に関して、評価対象者61人に、推薦候補者のうちから自分と同じ研究室に所属している人をリストアップしてもらった。この結果、その総数は46人で、推薦候補の上位5位の中に平均0.75人の同一研究室メンバーがいることがわかった。この結果は、非友人集合からランダムに5人を選んだ場合に含まれる同一研究室所属者数の期待値よりも多少大きい値であることを確認している。特定人物が多くの評価対象者に対する推薦候補上位者として選出される傾向が見られ、14人の評価対象者に対して1位として推薦される人物や、11人に対して1位として推薦される人物がいた。

6. 考察

6.1 評価方法1の結果に基づく考察

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のいずれにおいても、友人類似度の平均値が非友人類似度の平均値よりも1%水準で有意に高いとい

う結果が得られた。つまり、評価のために敢えて除外された友人は、友人候補として非友人よりも全般的に高く評価されるということである。このことから、提案手法の基本的な妥当性が裏付けられたと考えられる。

6.2 評価2の結果に基づく考察

本評価では、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、かつランダムに5人を選んだ場合よりも多い人数が選ばれたことから、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2人以下か1人以下と少ないため、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。

なお、個人内情報を利用する推薦方法では、リストアップしてもらった結果には、同じ研究室に所属している人を複数推薦される被推薦対象者が多く、62人の内15人の被推薦対象者には同一研究室所属者が3人以上推薦されていた。この15人の被推薦対象者の友人リストに登録されている友人のプロフィール情報を調査したところ、この15人のユーザはいずれも自分と同じ研究室の者だけを友人として登録していた。このため、これらのユーザに関しては「研究室」という属性の重みがとりわけ高くなり、同一研究室所属者が特に多数推薦される結果となったと思われる。この結果は、新規な友人の推薦という点では問題が残るが、本提案手法の信頼性を裏付けるものであると言えるだろう。

また、個人間情報を利用した推薦方法では、特定人物が多くの被推薦対象者に対する推薦候補上位者として選出される傾向が見られた。たとえば、14人の被推薦対象者に対して1位として推薦される人物や、11人に対して1位として推薦される人物がいた。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かったため、活発にコメントをつけていた一部のユーザの評価が特に高くなったことが原因だと考えられる。

7. おわりに

本稿では、学内で人的交流が生じにくいという問題を解決するために、学内SNSを利用して、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するものであったのに対し、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、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。提案手法は、友人リストに登録された友人のプロフィー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人として最適と思われるユーザを推薦する。

この手法の検証に使うデータを集めるため、OpenPNEを利用した学内イントラSNSを構築し、筆者らが所属している大学院で運用した。2010年7月から12月までの5カ月間、データの収集を行った。推薦アルゴリズムを検証するために、2種類の評価方法を利用して評価実験を行った。評価1では、すでに友人リストに登録されている人物を1人友人リストから除外し、この除外された友人に対する類似度が、非友人に対する類似度よりも高くなるかどうかを評価した。t検定により、除外された友人に対する類似度の平均が非友人に対する類似度の平均より有意に高いことがわかり、提案アルゴリズムの基本的妥当性が示された。評価2では、個人内情報および個人間情報を利用した候補者選出アルゴリズムによって選出された候補者上位5人の内で、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うことによって評価した。この結果、各被推薦対象者に対して、被推薦対象者と同一の研究室所属する者が平均1人程度推薦され、提案手法の一定の有効性が示唆された。

今後は、さらに長期間の運営を通じて、友人リストの登録人数と日記へのコメント数を増やすことで、提案手法のより詳細な妥当性検証を実施したい。実際に推薦を行った人物と友人関係を構築できるかどうかに関しても、長期的な調査を実施したいと考えている。

本稿では、友人を1人も登録していないユーザは被推薦対象者として扱わなかった。これは、提案手法が友人リストの存在を前提とした手法であるためである。しかしながら、友人リストを持たないユーザについては、友人リストを持つ各ユーザの友人リストとの類似度を計算し、最も高い類似度を持つ、友人リストを持つユーザを友人候補として推薦するという、逆の選出手法が考えられる。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[13]。この知見に基づけば、上述のような友人数が0人のユーザに、友人リストを有するユーザを紹介する方法は、有効なものと考えられる。今後は、このような手法の有効性についても検証を行いたい。

謝辞

本研究は、平成22年度国立情報学研究所公募型共同研究の支援を受けて実施されました。ここに謝意を表します。また、実験に協力してくださった、北陸先端科学技術大学院大学の多くの学生諸氏に感謝いたします。

参考文献

- 1) <http://mixi.jp/>
- 2) <http://yahoo-mbga.jp/>
- 3) <http://www.facebook.com/>

- 4) <http://www.renren.com/>
- 5) 岡本健志, 田中秀幸: 地域 SNS のユーザ同士のつながり方に着目したネットワーク分析, 日本社会情報学会学会誌, pp.45-55, 2009.
- 6) 加藤菜美絵, 小川祐樹, 諏訪博彦, 太田敏澄: 企業内 SNS 導入における有効性に関する調査研究, 日本社会情報学会学会誌, pp.19-32, 2009.
- 7) 伊藤大河, 山本利一: コミュニケーション能力育成を目指した SNS の効果的な活用 (情報機器の活用, 21 世紀の教育改革の行方を探る), 年会論文集 (25) 日本教育情報学会, pp.282-283, 2009.
- 8) 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ワーキングシステム, 情報処理学会第 67 回全国大会, 2005.
- 9) 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.
- 10) 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.
- 11) 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.
- 12) 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.
- 13) 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.