

## 中高生向けソーシャルメディアにおける人間関係推定方式 に関する一考察

本庄勝<sup>†1</sup> 田上敦士<sup>†1</sup> 橋本真幸<sup>†1</sup> 黒川雅幸<sup>†2</sup>  
三島浩路<sup>†3</sup> 吉田俊和<sup>†4</sup> 長谷川亨<sup>†1</sup>

中高生の間で、プロフィールやマイリンク、ブログ、ゲスブと呼ばれるサービスを提供するソーシャルメディア（中高生向けソーシャルメディアと呼ぶ）の利用が広がっている。これらのサービスは、コミュニケーションツールとして円滑な人間関係構築に利用される一方で、特定の相手に対する無視や仲間外れといったネットいじめにも利用されることもあり、安心して利用できる中高生向けソーシャルメディアが求められている。我々はこれまでに、中高生の間で発生するネットいじめを自動で検出するためのフレームワークについて検討を進めてきた。本論文では、人間関係の推定に焦点を当て、教育現場の協力によって得られたソシオメトリのデータを用いて、中高生の二者が親密さを確認するために行う、特徴的な相互行為について分析調査を行った。また機械学習を用いた二者間の親密さの推定精度についても評価したので報告する。

### A Study on Link Prediction Method for Estimating Intimate Friends on Teen's Social Media

MASARU HONJO<sup>†1</sup> ATSUSHI TAGAMI<sup>†1</sup> MASAYUKI HASHIMOTO<sup>†1</sup>  
MASAYUKI KUROKAWA<sup>†2</sup>  
KOJI MISHIMA<sup>†3</sup> TOSHIKAZU YOSHIDA<sup>†4</sup> TORU HASEGAWA<sup>†1</sup>

Recently, the use of social media which provides services such as profile, friend list, blog or guestbook (referred as teen's social media) prevails among high school or junior high school students in Japan. They utilize teen's social media aiming at maintaining intimate relationship with their friends, while they use it as a tool to attack their friends, namely causing cyber school bullying. In order to mitigate this problem, we have developed a framework which detects cyber school bullying among high or junior high school students. In this paper, we focus on the estimation of the intimate relationship and report an investigation results which analyzed mutual interactions among intimate friends on teen's social media, using sociometric data which was acquired through the cooperation with educators. Furthermore we evaluated the accuracy of estimation of intimate relationship using a typical machine learning engine.

#### 1. はじめに

ネット接続可能な携帯電話の普及に伴い、中高生の間で発生するネットを用いたいじめ（ネットいじめ）が増加している。特にプロフィールやマイリンク、ブログ、ゲスブと呼ばれるサービスを提供するソーシャルメディア（中高生向けソーシャルメディアと呼ぶ）は、学校内の友達や、仲良しの友達など、対面で一定の関係のある相手とのコミュニケーションに利用される一方で、仲間外れや無視といったネットいじめにも利用されることがあり[1]、社会的な問題となっている。

ネットいじめの本質については、匿名性など、ネット特有の性質と言われることが多いものの、学術的には確証が得られていないのが現状である。これに対して、我々の調査では、中高生の間で発生するネットいじめは、学校等対

面でのいじめの延長であること、すなわちネットいじめは対面での人間関係によって引き起こされていることなどが判明した[1]。特に女子中高生の間では、親密な関係である集団の中で、互いにネット上の相手を理解した上で、意図的な仲間外れや無視といった人間関係を操作する攻撃が観測された。すなわち、匿名性といったネットの特有の性質とは関係なく、相手を認識した上で攻撃行為を行っている。Crick & Grotpeter らは、対面で発生するこのような攻撃を関係性攻撃 (Relational aggression) と呼んでいる[2]。また、関係性攻撃によるネットいじめにおいては、対面でのいじめと同様、教師等が生徒たちの人間関係を把握し、適切に介入することがその解決に有効であることも判明している[3]。すなわち、ネット上での人間関係を教師等が把握し、その対処を適切に行うことが、中高生の間で発生するネットいじめの問題解決に有用である。

中高生の間で発生するネットいじめの問題の解決に資するため、我々はネットいじめを自動で検出するためのフレームワークの検討を進めてきた[1][4]。本フレームワークでは、既存のキーワード検出など自然言語処理を用いた手法では発見しづらい、仲間外れや無視といった関係性攻

†1 (株) KDDI 研究所  
KDDI R&D Laboratories Inc.

†2 福岡教育大学  
Faculty of Education, Fukuoka University of Education

†3 中部大学現代教育学部  
College of Contemporary Education, Chubu University

†4 名古屋大学大学院教育発達科学研究科  
Graduate School of Education and Human Development, Nagoya University

撃によるネットいじめを対象としており、以下の二段階の処理で構成される。(1)マイリンクやプロフ等での友達登録といった友人関係を表す Contact Network と、ブログ等での書き込み、応答といった活動を表す Activity Network という二つの人間関係グラフ(ソーシャルグラフ)を抽出する。(2) ソーシャルグラフの構造的な変化からネットいじめの予兆や発生を検出する。

我々はこれまで、実際の教育現場の教師等に協力を仰ぎ、第一段階までを試作したツールを用いて、実践的な研究を進めてきた[5]。その中で、ネット上での人間関係は教師や親にとって把握しにくく、ネット上での友人関係を表す Contact Network の可視化のみでも十分有用であることが分かった。その一方で、Contact Network だけではトラブルの有無等が捉えにくく、生徒間のやり取りやテキストの内容を見るなど、教師等による分析作業のコストが問題となった。これに対応するためには、ネット上での行動を表す Activity Network から、中高生同士の「親密さ」といった理解しやすい尺度の抽出が必要である。また第二段階に進むためには、仲間外れや無視といった攻撃の構造を理解する必要があり、中高生同士の「親密さ」の定量的な評価が必要となっていた。

そこで本論文では、中高生向けソーシャルメディアにおける中高生の「親密さ」を推定する手法を確立することを目的として、二者間の特徴的な相互行為に関する分析調査を行ったのでその報告を行う。榎本[6]によれば、青年期における友人関係の中には、相互理解活動、親密確認活動、共有活動、閉鎖的活動、といった活動が見られるといわれている。実際に、中高生においては、15分ルールや好意の返報など、親密さを確認するための特徴的な相互行為があることも一般に知られている。そこで、中高生向けソーシャルメディアの各サービスにおける行為を規定し、親密さ

を推定するための二者間の相互行為について定量的に評価した。親密さの変化は、Activity Network に表出することがこれまでの事例調査で分かっている[1]。そこでそれを検証する上で、本稿では、Contact Network, Activity Network それぞれの相互行為を評価した。Preferential attachment や Jaccard 係数など、社会ネットワーク分析で用いられる指標についても同様に評価した。さらに、各種の機械学習アルゴリズムを用いて、親密さの推定の性能も評価した。

本分析調査に先立ち、我々は、ある実証実験校において、生徒の学校での親密さの関係を表すデータ(ソシオメトリと呼ぶ)を得た。親密さの評価値はこのソシオメトリのデータを用いる。一般に、このようなデータを得ることは困難であり、本分析調査は他に類のない取り組みである。

本論文の構成は以下のとおりである。まず2章では、我々が対象とする中高生向けソーシャルメディアの概要を説明する。そして、3章では、今回の分析調査で対象とする親密さを確認するための特徴的な相互行為や一般的なネットワークの指標について説明を行い、4章でその分析結果を説明する。また5章においては、考察を行うと共に各種機械学習アルゴリズムを用いた性能評価について述べる。

## 2. 中高生向けソーシャルメディアの概要

### 2.1 中高生向けソーシャルメディアの概要

中高生向けソーシャルメディアとは、mixi[7]やfacebook[8]といった一般的なソーシャルメディアで提供される各種サービスと同等のサービスを中高生に対して提供するソーシャルメディアである。利用形態の特徴として、中高生等は好み等によって、サービスごとにサービスプロバイダを使い分け、それらをハイパーリンクでつなぎ合わせることで、あたかも独自のサイトを構築している。このため、サービスプロバイダによって異なる名前を使うなど、



図 1 中高生向けソーシャルメディアで提供される主なサービス(サイト)。

教師や親など第三者には全体を把握することが困難である。

中高生向けソーシャルメディアが提供するサービスは、主にホームページ、プロフィール、ブログ、リアル、ゲスブ、マイリンク、の6つのタイプに分類できる。プロフィール、マイリンク、ブログ、ゲスブの例を図1に示す。なおプロフィールやマイリンク、ブログ、ゲスブ等の名称は通例の呼び名であり、それぞれのサービスもしくは、サービスによって作成されたサイト(Web ページ群)を示すが、本稿ではサイトの名称としてこれらを用いる。

#### ホームページ

ホームページを作成できるサイトである。一般的には、トップページとして利用されることが多く、形式もフリーフォーマットであることが多い。中高生等にはプロフィールと同様、自己紹介や、親密な相手を紹介することを目的として利用されることが多い。

#### プロフィール

自身の個人属性情報を公開できるサイトである。前略プロフィール[9]などの様に、あらかじめ決められた項目に対して、自身を表現する形式のものもある。自己紹介だけでなく、親密な相手を紹介することも利用されることがある。中高生等は、親密な相手を紹介する際、「イツメン」「双子」「相方」といった表現を用いることもある。

#### マイリンク

友達のリストを公開できるサイトである。公開する際、相手の承認が必要な場合が多い。中高生等には、ホームページやプロフィールと同様親密な関係の相手を公開することに加え、学級の友人や部活動の友人など、一定の関係のある相手も紹介することを目的として利用されることが多い。

#### ブログ

自身が記事を投稿したり、他人が足跡やコメントを投稿したりできるサイトである。一般的に言われるブログと同等であり、自身のブログで記事を投稿したり、親密な相手のブログの記事には、コメントを投稿したり、足跡(閲覧履歴)を残したりするために利用される。

#### リアル

個人の気持ちや状況を短い文章で投稿できるサイトである。

ブログと異なり、足跡を残す機能はないことが多い。Twitter[10]など一般的に言われるマイクロブログに近いが、中高生等の利用目的は若干異なる。リアルは今の感情を不特定他者に公開するために用いられるが、実際には、特定の相手に向けられており、周りの友人らは、周りの空気を読むために利用する[11]。親密な相手であっても、コメントを投稿することはブログに比べ少ない。文章が短いため、短絡的な書き込みになることが多く、トラブルが発生することも多い。

#### ゲスブ

他人が訪問履歴として記事を投稿できるサイトである。ゲストブックを略してこのように呼ばれる。個人用の掲示板である。自分のゲスブに投稿するのではなく、他者のゲスブに記事を投稿する。親密な相手とのやり取りにも利用されるが、出会い目的等で利用されることの方が多い。お互いが相手のゲスブに記事を投稿するため、会話の内容は、双方のゲスブに表出する。

これらのサイトは、自己紹介や友人紹介といった情報発信を目的とし、あまり更新されないサイト(ホームページ、プロフィール、マイリンク)と、日記投稿やそれに対するコメント投稿など、友人との交流を目的とし、頻繁に更新されるサイト(ブログ、リアル、ゲスブ)に分類できる。本稿では前者を静的なサイト、後者を動的なサイトと呼ぶ。静的なサイト、及び動的なサイトの情報をもとに生成されるソーシャルグラフがそれぞれ、Contact Network, Activity Networkとなる。

### 2.2 中高生向けソーシャルメディアを用いた行為

本稿では、中高生向けソーシャルメディアの各サイトについて、以下の行為を観測対象とした。

#### 1) 静的なサイト

- ①自分のサイトに相手のサイトへのリンクを公開する

#### 2) 動的なサイト

- ①自分のサイトに記事を投稿する
- ②相手のサイトに記事を投稿する
- ③相手のサイトにコメントを投稿する
- ④相手のサイトに足跡を投稿する(足跡を残す)

これらは、サイトのタイプによって、傾向の違いがある。

表1 中高生向けソーシャルメディアのサービス(サイト)で可能な行為。

行為	$m$	ホームページ	プロフィール	マイリンク	ブログ	リアル	ゲスブ
$friend(v_i \rightarrow v_j, v_i, m)$	$v_i$ が管理するタイプ $m$ のサイトにおいて、 $v_i$ が $v_j$ のサイトへのリンクを登録する行為	○	○	○	—	—	—
$article(v_i \rightarrow \emptyset, v_i, m)$	$v_i$ が管理するタイプ $m$ のサイトにおいて、 $v_i$ が(不特定他者に向けて)記事を投稿する行為	—	—	—	○	○	—
$article(v_i \rightarrow v_j, v_j, m)$	$v_i$ が管理するタイプ $m$ のサイトにおいて、 $v_i$ が $v_j$ に向けて記事を投稿する行為	—	—	—	—	—	○
$comment(v_i \rightarrow v_j, v_j, m)$	$v_i$ が管理するタイプ $m$ のサイトにおいて、 $v_i$ が $v_j$ に向けてコメントを投稿する行為	—	—	—	○	○	—
$footprint(v_i \rightarrow v_j, v_j, m)$	$v_i$ が管理するタイプ $m$ のサイトにおいて、 $v_i$ が $v_j$ に向けて足跡を投稿する行為	—	—	—	○	—	—

ことが予想される。友人を紹介する際の相手側の承認の有無や、ブログやリアルでのコメントを投稿する際の敷居の差異、などである。このため、表 1 の様にデータを定義した。ここで  $v_i, v_j$  は、サイトの管理者、すなわち、中高生であり、 $m$  はサイトのタイプである。この表では、一般的な中高生向けソーシャルメディアで可能な、行為の可否を記載した。リアルでの足跡の可否など、若干の差異はあるが、傾向はどのサービスプロバイダもほぼ同様と考えてよい。

### 3. 分析

本稿では、ネットいじめを特定するために二者間の「親密さ」に関する分析を行った。Contact Network, Activity Network のそれぞれについて指標を検討した。

#### 3.1 指標の選定

##### 3.1.1 二者間の相互行為に関する指標

今回の分析では、これまでの調査の中で得られた知見をもとに、以下の視点で指標を選定した。

##### 1) 静的なサイト：Contact Network

・**双方向性 (Reciprocity) : 互いに相手のサイトへのリンクを公開しているか**

親密な二者であれば、相手のサイトへリンクを互いに公開していると考えられる。我々が過去に調査した仲間集団内で起きたネットいじめの事例においても、仲間集団内の生徒は互いに相手のサイトへのリンクを公開していた[1]。片方だけが公開しているのは好意の返報からも不自然と考えられる。但しサービスプロバイダによっては必ず双方向になるケースもあるので注意する必要がある。またマイリンクは、親密ではない普通の相手へのサイトも公開しているケースもあるので、マイリンクではなく、ホームペやプロフを観測の方が意味のある結果が得られる可能性はある。

##### 2) 動的なサイト：Activity Network

・**応答率 (Response Ratio) : 相手の投稿した記事に対して、どの程度コメント (もしくは足跡) を投稿しているか**

ブログやリアルの内容は、親密な関係の間柄の中ではシェアされていることが多く、ブログについては、コメントや足跡を投稿することが多い。特にコメントは、内容ではなく、投稿することに意味があることが考えられる。そこでブログやリアルで投稿された記事の件数に対して、どの程度応答しているかを観測することは、有意であると考えられる。

・**即応率 (Quick Response Ratio) : 相手の投稿した記事に対して、一定時間内にどの程度コメント (もしくは足跡) を投稿しているか**

応答率に関連するが、どの程度早く返答するかも中高生の中では親密さを確認するための重要な要素と考えられる。実際に、あるブログの記事の投稿に対して、数分内でコメントや足跡を投稿する例は、これまでの調査でも見受けられた。ブログやリアルで投稿された記事に対して、一定時間内にコメントや足跡を投稿する頻度を観測することは、有意であると考えられる。

・**投稿比率 (Article Ratio) : 互いに、相手のサイトに記事をどの程度投稿しているか**

ゲスブにおいては、一般にコメントや足跡といった概念がないため、互いのサイトに記事を投稿する。出会い目的等で利用されることも多いが、親密な間柄であっても、ゲスブを利用することもある。この場合は、短時間で大量の記事を投稿することが多い。通称「盛る」と呼ばれる行為だが、互いのゲスブに 100 件以上の記事を投稿することもある。内容でなく投稿のやり取りに親密さを確認する意味があると考えられる。そこで、両者の投稿の比率を観測することが有意であると我々は考える。

今回選定した指標を表 2 にまとめた。なお  $comment_T()$ 、 $footprint_T()$  はそれぞれ、記事に対して  $T$

表 2 親密な二者間の相互行為に関する指標。

Feature	Value
Contact Network	
Reciprocity	$(friend(v_i \rightarrow v_j, v_i, m) + friend(v_j \rightarrow v_i, v_j, m)) / 2$
Activity Network	
Response Ratio (comment)	$\min\left(\frac{comment(v_i \rightarrow v_j, v_j, m)}{article(v_j \rightarrow \emptyset, v_j, m)}, \frac{comment(v_j \rightarrow v_i, v_i, m)}{article(v_i \rightarrow \emptyset, v_i, m)}\right)$
Response Ratio (footprint)	$\min\left(\frac{footprint(v_i \rightarrow v_j, v_j, m)}{article(v_j \rightarrow \emptyset, v_j, m)}, \frac{footprint(v_j \rightarrow v_i, v_i, m)}{article(v_i \rightarrow \emptyset, v_i, m)}\right)$
Quick Response Ratio (comment)	$\min\left(\frac{comment_T(v_i \rightarrow v_j, v_j, m)}{article(v_j \rightarrow \emptyset, v_j, m)}, \frac{comment_T(v_j \rightarrow v_i, v_i, m)}{article(v_i \rightarrow \emptyset, v_i, m)}\right)$
Quick Response Ratio (footprint)	$\min\left(\frac{footprint_T(v_i \rightarrow v_j, v_j, m)}{article(v_j \rightarrow \emptyset, v_j, m)}, \frac{footprint_T(v_j \rightarrow v_i, v_i, m)}{article(v_i \rightarrow \emptyset, v_i, m)}\right)$
Article Ratio	$\frac{article(v_i \rightarrow v_j, v_j, m) + article(v_j \rightarrow v_i, v_i, m)}{2 \times \max(article(v_i \rightarrow v_j, v_j, m), article(v_j \rightarrow v_i, v_i, m))}$

時間内に投稿されたコメント、足跡の件数を表す。

### 3.1.2 一般的なネットワーク指標

社会ネットワーク分析における一般的な指標についても評価した。今回選定した指標を表 3 にまとめる。ここでは紙面の都合上、詳細は割愛する。指標の選定においては、文献[12][13][14][15][16][17][18][19]等を参照した。今回の分析調査では、マイリンクのサイトの情報によって構成される Contact Network のソーシャルグラフについてのみ評価する。ソーシャルグラフ及び、表におけるデータの定義は以下の通りである。

- ・ソーシャルグラフ :  $G = (V, E)$
- ・ノード (管理者) :  $v_i \in V, (i = 1, 2, \dots, |V|)$
- ・エッジ (管理者間のハイパーリンク) :  $(v_i, v_j) \in E$
- ・隣接ノード :  $\Gamma(v_i) := \{v_j | (v_i, v_j) \in E \text{ or } (v_j, v_i) \in E\}$   
 $\Gamma^+(v_i) := \Gamma(v_i) \cup \{v_i\}$
- ・サブグラフ :

$$\begin{aligned} nhs\text{ubgraph}(v_i, v_j) &= \{(v_k, v_l) \in E | v_k, v_l \in \Gamma(v_i) \cup \Gamma(v_j)\} \\ nhs\text{ubgraph}^+(v_i, v_j) &= \{(v_k, v_l) \in E | v_k, v_l \in \Gamma^+(v_i) \cup \Gamma^+(v_j)\} \\ inners\text{ubgraph}(v_i, v_j) &= \{(v_k, v_l) \in E | (v_k \in \Gamma(v_i) \text{ and } v_l \in \Gamma(v_j)) \text{ or } (v_k \in \Gamma(v_j) \text{ and } v_l \in \Gamma(v_i))\} \end{aligned}$$

## 3.2 分析のセットアップ

### 3.2.1 サンプルデータの作成

本稿では、親密な二者が行う特徴的な相互行為を評価するため、教育現場に協力を仰ぎ、サンプルデータを作成した。作成手順は以下のとおりである。

まずある高等学校の1年生、2年生の全生徒を対象として、2011年11月下旬にアンケート調査を実施した。学校適応

表 3 一般的なネットワーク指標。

Feature	Value
Contact Network	
Common Friends	$ \Gamma(v_i) \cap \Gamma(v_j) $
Total Friends	$ \Gamma(v_i) \cup \Gamma(v_j) $
Jaccard's Coeff	$\frac{ \Gamma(v_i) \cap \Gamma(v_j) }{ \Gamma(v_i) \cup \Gamma(v_j) }$
Preferential Attachment	$ \Gamma(v_i)   \Gamma(v_j) $
Friends Measure	$\sum_{x \in \Gamma(v_i)} \sum_{y \in \Gamma(v_j)} \delta(x, y)$
Adamic and Adar	$\sum_{\{x   x \in \Gamma(v_i) \cap \Gamma(v_j)\}} \frac{1}{\log  \Gamma(x) }$
Subgraph Edge Number	$ nhs\text{ubgraph}(v_i, v_j) $
Subgraph Edge Number+	$ nhs\text{ubgraph}^+(v_i, v_j) $
Inner Subgraph	$ inners\text{ubgraph}(v_i, v_j) $

感や友人関係の適応等に関する質問に加え、親密な相手に関する質問を行った。「ふだん、一緒に過ごすことが多い友人はだれですか?」といった質問を行い、上位3名の名前を記入してもらった。その結果からソシオメトリのデータを作成した。

次にアンケート実施月における、中高生向けソーシャルメディアにおける行為の記録を我々が開発したツールの通信ログから抽出した。具体的には、ホームペ・プロフ・マイリンクについては、2011年11月下旬の特定の日に観測された行為を抽出した。ブログ・リアル・ゲスブについては、2011年11月の1か月間に発生した行為を抽出した。同様にひと月前後における行為についても抽出した。なお、ツールでは、実証実験協力校の生徒とされる管理者のサイトを中心に、2週間に1回程度の頻度で、対象期間を含む期間の間、定期的に収集した。

最後に、ソシオメトリのデータと通信ログのデータについて、マッピング作業を行い、サンプルデータを作成した。

### 3.2.2 サンプルデータの基本情報

アンケート調査で有効な回答をした生徒の中から、ペアを抽出した。ペアの抽出に当たっては、同一学級のペアに限定した。結果を表 4 に示す。互いに第一選択したペアは計 77 組だった。全員大好きなど、相手が特定できない回答をしている生徒を含むペアは除去した。括弧内の数値は、二人の生徒がともに、マッピングできた (両者がサイトを保有していた) ペアの数を表す。

今回、いずれかのペアに属し、サイトを 1 つ以上保有し

表 4 ソシオメトリのデータ。

ペア	ペア数
第一選択/第一選択	77(11)
第一選択/第二・三選択	97(13)
第二・三選択/第二・三選択	123(19)
第一選択/非選択	89(16)
第二・三選択/非選択	285(36)
非選択/非選択	6,658(681)

表 5 ホームペ・プロフ・マイリンクの平均友達公開数。

	10月	11月	12月
ホームペ	8.4 (5.6)	8.2 (5.5)	7.4 (5.4)
プロフ	0.0 (0.0)	0.0 (0.0)	0 (0)
マイリンク	37.3 (12.6)	48.8 (15.4)	48.0 (15.2)

表 6 ブログ・リアル・ゲスブの平均投稿数。

	10月	11月	12月
ブログ (記事)	2.5	2.07	2.35
ブログ (コメント)	17.32 (6.33)	17.35 (5.80)	20.18 (6.68)
ブログ (足跡)	95.80 (29.82)	87.73 (26.0)	104.48 (31.28)
リアル (記事)	3.42	3.21	5.93
リアル (コメント)	0.89 (0.22)	0.75 (0.12)	2.04 (0.41)
ゲスブ (記事)	14.24 (2.70)	13.4 (3.02)	15.67 (2.99)

ていた生徒の数は、168名だった。対象生徒168名について、11月のホームペ・プロフ・マイリンクの平均公開数について調べた結果、それぞれ、8.2, 0.0, 48.8, であった(表5)。括弧内の数値は、対象生徒に限った場合の数である。なお今回調査したマイリンクは相手の承認が必要であり、承認されると双方向にサイトへのリンクが公開される。11月にマイリンクの数が増えているのは、学内行事(学校祭)があったのが要因の一つと考えられる。

次にブログ・リアル・ゲスブの平均投稿・コメント・足跡件数を調べたところ、11月では、管理者あたり、2.07件のブログの記事を投稿しており、17.35件のコメントや87.73件の足跡を投稿していた。その他の数値も含め結果を表6に示す。括弧内の数値は、対象生徒に限定した時の数値を表す。大きな差はないものの、12月に件数が増加しているのは冬季休業期間(冬休み)が影響していると考えられる。

## 4. 結果

### 4.1 分析結果

今回は、表4の結果をもとに、双方がサイトを保有し、且ついずれかが相手を選択していた親密ペア(95組)と、いずれも選択をしていなかった非親密ペア(681組)の二つに分けて各指標の平均値、標準偏差を評価した。11月の結果を表7に示す。即応率については $T=60[\text{min}]$ とした。一般的なネットワーク指標については、対象生徒に隣接する生徒も含めたサンプルデータを使用し、対象生徒についての各指標を評価した。

## 5. 考察

今回の分析調査から以下のことが分かる。まず、Contact Networkの双方向性については、マイリンクに大きな差が得られた。親密ペアの平均値は、0.71であり、標準偏差が0.44と数値は高いものの、親密な二者は、互いにサイトへのリンクを公開している傾向が高い。但し、非親密ペアについても平均が0.31と高いスコアになっている。この理由としては、非親密ペアも同一学級の生徒であり、対面で一定の人間関係があったからだと推察される。学級に関する制約を除けば、この差はより大きくなるだろう。

今回の調査では、殆どの生徒が共通のサービスプロバイダのマイリンクを使用していた。考えられることは、共通のサービスプロバイダのマイリンクを使い、互いの人間関係を構築する一方で、他のサービスプロバイダでリアルなサイト構築し、コミュニケーションするための複数のチャンネルを形成していたと考えられる。はやりや廃りがあるものの、共通とするサービスプロバイダは学校や学年ごとに存在する可能性がある。プロフについては、今回の分析調査では、友達のサイトへのリンクを公開している生徒は0であり、一部の生徒がホームペでリンクを公開していた。調査したサービスプロバイダではホームペがトップページ扱いであり、プロフはリンクをたどらないと閲覧することができない。そこでトップページであるホームペを用いて親密さを周りにアピールしていたと推察される。Contact Networkについては、我々は当初、親密ペアのマイリンクがほぼ1に近い値を持つことを予想していた。これまでの

表7 各指標の基本統計.

	全ペア		親密ペア		非親密ペア	
	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差
二者間の相互行為に関する指標						
Contact Network						
双方向性 (Reciprocity)						
ホームペ	0.00	0.05	0.02	0.10	0.00	0.03
プロフ	0.00	0.00	0.00	0.00	0.00	0.00
マイリンク	0.36	0.47	0.71	0.44	0.31	0.45
Activity Network						
応答率 (Response Ratio)						
ブログ (コメント)	0.06	0.21	0.17	0.30	0.05	0.19
ブログ (足跡)	0.20	0.36	0.35	0.41	0.18	0.34
リアル (コメント)	0.00	0.03	0.00	0.00	0.00	0.03
即応率 (Quick Response Ratio)						
ブログ (コメント)	0.05	0.18	0.13	0.26	0.04	0.16
ブログ (足跡)	0.14	0.29	0.26	0.36	0.13	0.28
リアル (コメント)	0.00	0.00	0.00	0.00	0.00	0.00
投稿比率 (Article Ratio)						
ゲスブ	0.03	0.13	0.09	0.22	0.02	0.11
一般的なネットワーク指標						
Contact Network (※マイリンクのみ)						
Common Friends	6.55	5.49	11.70	8.13	5.77	4.51
Total Friends	105.10	50.68	99.28	37.24	105.98	52.54
Jaccard's Coeff	0.07	0.05	0.12	0.07	0.06	0.04
Preferential Attachment	3115.94	3214.36	3097.48	2416.60	3118.75	3321.68
Friends Measure	189.05	153.77	286.53	181.51	174.20	143.87
Adamic and Adar	2.01	2.19	4.08	3.97	1.70	1.55
Subgraph Edge Number	311.07	169.74	335.83	164.80	307.30	170.75
Subgraph Edge Number+	418.23	187.85	399.13	168.80	421.14	191.19
Inner Subgraph	152.47	122.61	217.82	132.93	142.51	118.16

調査では、仲間集団の生徒は Contact Network でほぼクリークを形成していたからである。しかしながら、親密ペアの平均値は期待していた数値より低かった。一つの理由としては、親密ペアには、選択-非選択も含まれており、これらは一方的な好意を指していた可能性がある。

次に、Activity Network の応答率については、ブログのコメントが有意な差として現れた。非親密ペアでは平均が 0.05、標準偏差も 0.19 と低いスコアであることから、ブログにコメントを投稿することは、通常の関係では数居の高い行為であったと推察される。コメントと比較し、足跡では平均値が 0.18 であり、足跡は幾分数居が低い行為であると考えられる。即応率についてもほぼ同様の数値であった。つまり、中高生の中では、友達のブログの記事の投稿を常に意識し、投稿があった場合には、短時間の間に（授業中、就寝中などを除いて）、コメントもしくは足跡を投稿していたと考えられる。図 2 は、ブログのコメントについて応答時間の累積分布を調べた結果である。親密ペアについては 10 分以内にコメントを投稿しているのは全体の 5.5 割程度を占めていた。我々大人でも facebook 等で親しい相手が投稿した記事に対して足跡 (like ボタン) を投稿したりするが、このような緊張感のある使い方をしてる人はまれである。中高生の中にはこのような緊密な人間関係があることが本分析調査で確認することができた。リアルについては、これまでの実態どおり、記事の投稿は観測されたものの、コメントの投稿はほぼ 0 であった。今回規定した行為の範囲では、親密さに関わる行為としては有意な結果を得ることはできなかった。投稿比率については、親密ペアの平均が 0.09 であり、非親密ペアよりも高い数値がでていた。これは一部の親密ペアで「盛る」という行為が見られたのが理由として考えられる。一般にゲスブでは初対面で挨拶する際に使われることが多い。ゲスブで挨拶をして仲良くなった後、ブログなどにコメントや足跡を投稿する、といった具合である。調査月が 4 月や 5 月など新年度直後であれば、また異なる結果が得られる可能性がある。

一般的なネットワーク指標については、Common friends や Adamic and Adar に大きな差が出る結果となった。親密

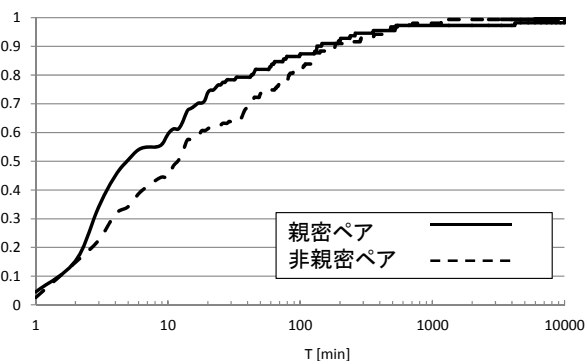


図 2 ブログのコメント投稿までの応答時間の累積分布。

ペアにおいて、Common friends が高い数値となった理由としては、親密な友達とは共通の部活に所属していたり、趣味などが共通であったりしていたことが考えられる。Adamic and Adar は二人の共通の友達について、それらが友達をあまりもっていない、すなわち閉鎖的な人間関係であると、高い数値をとる。親密な関係の二者には上記の傾向があることが分かった。

### 5.1 機械学習による推定

推定精度の傾向を把握するため、親密ペア、非親密ペアをそれぞれ 95 組抽出し機械学習による推定を行った。以下 3 つの指標セットを用意した。

- Mutual behavior features subset (二者間の相互行為に関する指標)
- Network features subset (ネットワークに関する指標)
- All features subset (全ての指標)

機械学習には WEKA[20]を使用した。WEKA とは、Java で書かれた機械学習ソフトウェアで、ニュージーランドの Waikato 大学において開発が進められている。各種の機械学習エンジンが利用可能であり、今回は、C4.5 (決定木)、Naïve-Bayes, kNN (k-nearest neighbor algorithm)、ANN (Artificial neural network)、SVM (Support vector machine)、の 5 つを使用した。パラメータはデフォルトの値とした。

機械学習では、10-fold の交差検証で行った。推定結果を図 3 に示す。All features subset での SVM、Naïve-bayes において最も高い推定精度 (66.8) が得られた。また Mutual behavior features subset と Network features subset では、全体的に後者の指標セットの方が高い推定精度となった。

我々の目標は、仲間集団を見つけるために、膨大な数のペアの中から親密なペアを探すことである。すなわち、親密ペアの再現率 (Recall) が高いほど良く、精度 (Precision) はある程度高ければよい。SVM について、再現率、精度を求めた結果を表 8 に示す。今回の結果では、再現率は全体的に低いものの (0.5 前後)、SVM の Mutual behavior features subset が最も高い (0.642) 結果となった。

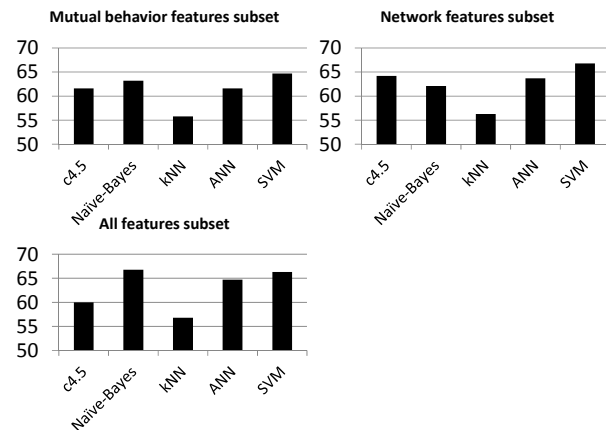


図 3 機械学習による推定結果。

表 8 機械学習における再現率と精度.

SVM		Precision	Recall	F-Measure
All features subset	親密ペア	<b>0.696</b>	<b>0.579</b>	<b>0.632</b>
	非親密ペア	0.64	0.747	0.689
Mutual behavior features subset	親密ペア	<b>0.649</b>	<b>0.642</b>	<b>0.646</b>
	非親密ペア	0.646	0.653	0.649
Network features subset	親密ペア	<b>0.786</b>	<b>0.463</b>	<b>0.583</b>
	非親密ペア	0.619	0.874	0.725

## 6. おわりに

本論文では、プロフやマイリンク、ブログ、ゲスブと呼ばれるサービスを提供する中高生向けソーシャルメディアを対象として、中高生の親密さに関する特徴的な相互行為について分析調査を行った。好意の返報や 15 分ルールなど、発達段階の時期における特徴的な行為が定性的に述べられることが多いが、本稿では、中高生向けソーシャルメディアの行為を規定し、それらを定量的に評価した。ホームペ・プロフ・マイリンクについては、双方向性、ブログ・リアル・ゲスブについては、応答率、即応率、投稿比率の計 4 つの視点について、指標を作成した。教育現場の協力によって得られたソシオメトリのデータを用いて、分析を行った。親密なペアは、マイリンクの双方向性に高い数値を持つこと、応答率については、ブログのコメントに有意な差が現れること、リアルについては、行為として差異が現れないこと、などが分かった。一般的なネットワーク指標についても、親密なペアとそうでないペアの傾向の違いを調べた。各種機械学習の精度についても評価を行い、デフォルトのパラメータではあるものの、再現率の点で、二者間の相互行為に関する指標がネットワーク指標よりも有効であることなど分かった。

今後は、別の視点での指標についても分析を進めると共に、人間関係推定方式の確立を目指す。

**謝辞** 本研究は、総務省 SCOPE (戦略的情報通信研究開発推進制度) の委託研究に基づく結果である。

## 参考文献

[1] M. Honjo, T. Hasegawa, T. Haseawa, T. Suda, K. Mishima and T. Yoshida, "A Framework to Identify Relationships among Students in School Bullying Using Digital Communication Media," in Proc. of the 3<sup>rd</sup> IEEE International Conference on Social Computing (SocialCom2011) Workshop, SBABC, pp. 1474-1479, Oct. 2011.

[2] N.R. Crick and J.K. Grotpeter, "Relational Aggression, Gender, and Social-Psychological Adjustment," Child Development 66, pp. 710-722, 1995.

[3] 三島浩路, 黒川雅幸, 大西彩子, 本庄勝, 吉武久美, 長谷川輝之, 長谷川亨, 吉田俊和, "ネット上のトラブルや「いじめ」に関する報告," 名古屋大学大学院教育発達科学研究科紀要 (心理発達科学), 57 巻, pp. 61-69, 2010.

[4] 本庄勝, 田上敦士, 三島浩路, 吉田俊和, 長谷川亨, "中高生向けソーシャルメディアにおけるソーシャルグラフ抽出のためのアカウント同定方式に関する一検討," DICOMO 2012, 8H-2, pp. 2271-2278, 2012.

[5] 中村海, 本庄勝, 橋本真幸, 三島浩路, 黒川雅幸, 吉田俊和, 長谷川亨, "人間関係を推定するフレームワークに基づくネットいじめ防止ツールの実装," 情処第 75 回全国大会, 投稿中, 2012.

[6] 榎本淳子, "青年期の友人関係の発達の变化—友人関係における活動・感情・欲求と適応—," 風間書房, 2003.

[7] mixi, <http://mixi.jp/>.

[8] facebook, <http://facebook.com/>.

[9] 前略プロフィール, <http://pr.cgiboy.com/>.

[10] Twitter, <http://twitter.com/>.

[11] 加藤千枝, 堀田香織, "中学校におけるインターネットを介したモニタリング活動の実践—思春期女子のインターネット利用の実態と考察—," 埼玉大学紀要, 61 (1), pp. 107-119, 2012.

[12] M. Fire, L. Tenenboim, O. Lesser, R. Puzis, L. Rokach and Y. Elovici, "Link Prediction in Social Networks using Computationally Efficient Topological Features," in Proc. of the 3<sup>rd</sup> IEEE International Conference on Social Computing (SocialCom2011), pp. 73-80, Oct. 2011.

[13] J. Cheng, D.M. Romero, B. Meeder and J. Kleinberg, "Predicting Reciprocity in Social Networks," in Proc. of the 3<sup>rd</sup> IEEE International Conference on Social Computing (SocialCom2011), pp. 49-56, Oct. 2011.

[14] H.H. Song, T.W. Cho, V. Dave, Y. Zhang and L. Qiu, "Scalable Proximity Estimation and Link Prediction in Online Social Networks," in Proc. of the 9th ACM SIGCOMM conference on Internet measurement conference (IMC '09), pp. 322-335, Nov. 2009.

[15] W.J. Cukierski, B. Hamner and B. Yang, "Graph-based Features for Supervised Link Prediction," International Joint Conference on Neural Networks, 2011.

[16] L. Adamic, "Friends and Neighbors on the Web," Social Networks, 2003.

[17] M.E.J. Newman, "Clustering and Preferential Attachment in Growing Networks," Physical Review E, no.2, pp. 025 102+, Jul. 2001.

[18] A. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert and T. Vicsek, "Evolution of the Social Network of Scientific Collaborations," Physica A: Statistical Mechanics and its Applications, vol. 311, 2002.

[19] L. Katz, "A New Status Index Derived from Sociometric Analysis," Psychometrika, 1953.

[20] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I.H. Witten, "The Weka Data Mining Software: an Update," SIGKDD Explor Newsl., vol. 11, pp. 10-18, Nov. 2009.