

mixi における関係形成原理

連結とコミュニティの成長

安田 雪

東京大学大学院経済学研究科

E-mail: yasuda@mmrc.e.u-tokyo.ac.jp

松尾 豊

独立行政法人 産業技術総合研究所

スタンフォード大学

E-mail: y.matsuo@aist.go.jp

キーワード：SNS、コミュニティ、社会ネットワーク分析

1. SNS とは

近年、ソーシャルネットワーキングサービス(SNS)が着目を集めている。SNSとは、WWW上で知人関係を形成し相互交流を行うためのプラットフォームであり、日本では2004年から、GREE、mixi、Orkutなどが、それぞれ独立のサービスとして無料で提供している。中でも、最もユーザ数の多いmixi(ミクシィ)は、ユーザ数が2004年5月に1万人、2004年9月に10万人、2005年8月に100万人に達した。² さらに増加のペースは衰えず、2006年10月にはユーザ数600万人を超えている。

SNSでは、ユーザはシステム上で、自分の知り合いや友人を登録することで関係を構築していく。そして、ユーザが次々に新しいユーザを招待することでユーザ数が増え続けていく。では、このように連鎖的にユーザ数を増やす SNS の内部ではいったい何が起きているのだろうか?

¹ 報告内容の一部は、東京大学ものづくり経営研究センターMMRCディスカッションペーパーNo. 107 松尾豊・安田雪「SNSにおける関係形成原理」http://www.ut-mmrc.jp/dp/PDF/MMRC107_2007.pdfにて公開中である。なお、本報告をふまえて内容を修正した論文を人工知能学会誌へ投稿中である。

² 株式会社ミクシィ 2006年3月2日プレスリリースより。

ユーザはどのような人とどのようなコミュニケーションを行っているのだろうか。SNSでは、いかなるネットワークの形成原理が働いており、結果的にどのような関係構造が生成されているのであろうか。現実の人間関係の全体像は、誰にも俯瞰できぬいわば「見えない連鎖」であるが、SNSの関係構造はその見えない連鎖を限定的ではあるが、具現化する可能性があることが、これらの関心を導いていると考えられる。

SNSが生み出す関係構造については、湯田ら(湯田, 藤原, 2005)がmixiを対象とした研究で、(i) スケールフリー性と高い凝集性、(ii) 抽出したクラスタの大きさにおけるZipf則とギャップの存在を指摘している。本研究でも同様に、mixiを対象とした分析を行いたい。友人関係の相互承認によって形成されたネットワークがどのような特徴を持つか、中心性の高いユーザはどのような関係を構築しているか、また、ユーザの関心の共通性を表すコミュニティがどのようなネットワークを構成し、どのような相互作用を行っているのかなどについて分析する。既存の実運用システムを対象とした研究であるので、データ分析と解釈が主な内容であるが、今後、SNSをはじめとするコミュニティWebプラットフォーム(大向, 松尾, 松村, 武田, 2006)を構築していく上で、ひとつの重要な知見を提供するものである。

また、現実社会における友人・知人のホールネットワークの抽出は、極めて困難であり、都市社会学の分野でエゴセントリックな友人・知人ネットワークのサンプル調査がなされているものの、数十万規模の友人・知人の相互連結ネットワークの形態分析は、商品の普及や、流行や疾病の伝播研究などにも有用な知見を提供しうると考える。

要約するに、SNSは、学術的には、異なる社会圏に属する人々の距離、伝播・普及のメカニズム、巨大ネットワークの連鎖構造、スモールワールド実験などにも関連する「人々の社会的相互作用や関心事項は、どのような構造の知人・友人ネットワーク、関心のコミュニティをつくりだすのか」という関心の対象となる。一方、実務的には、情報の伝播・普及、商品や疫病の普及などを統制するために必要な、人間、物財、情報の連結構造の把握のための基礎データをSNSは提供しうる。これらが我々が日本のSNSの代表的な存在であるmixiを分析する理由である。

2. 分析に用いたデータ

本分析では、mixi上で形成されている関係を対象とする。データは、mixi上での友人関係、コミュニティの参加に関するものであり、個人情報は一切含まれておらず、個別のユーザIDも特定できないように加工されている。³

³ データは2005年3月に提供を受けたものである。株式会社ミクシィ様からのデータ提供に際して

mixi の機能について簡単に説明する。mixi は多くの SNS がそうであるように招待制をとっており、mixi ユーザからの招待があってはじめて参加することができる。ユーザ登録後、知人を mixi に招待したり、mixi 上で友人を見つけることによって、紐帯が増加する。この知り合いは「マイミクシィ」(略してマイミク)と呼ばれる。紐帯の形成は、具体的には、ユーザ A が別のユーザ B に対して、マイミクシィに追加してもらうように依頼のメッセージを出し、ユーザ B が承認すると、ユーザ A、B 双方に相手がマイミクシィとして登録される。この相互承認をもって作られたマイミクシィの関係を、本論文では友人関係と呼ぶことにする。

用いたデータは、ユーザ数が 363,819 人、紐帯総数が 3,813,702 である。ほとんどの紐帯は双方向あるが、108 個の紐帯だけは非対称である。⁴ そのため、本分析では方向ありの紐帯が約 381 万本存在するとして議論を進める。一人当たりの平均友人数は 10.48 人である。

さらに mixi では、ユーザはコミュニティと呼ばれるグループに所属できる。コミュニティは、ユーザが自由に開設でき、設定したテーマに関心のあるユーザが参加して話題を交わすための場となる。コミュニティへの参加は自発的なものであり、誰でも参加可能なもの、管理者の承認が必要なものがある。用いたデータでは、コミュニティの数は 90,795 個であり、コミュニティにひとつでも所属しているユーザ数は 241,423 人(全ユーザの 66.4%)であった。

ここで、重要な点は、互いの友人・知人関係を公開することが、mixi における良質なコミュニケーションの維持に大きく貢献している点である。複数の友人・知人と mixi 内で良好な関係を維持するためには、他者に統一的な自画像を提示しつづけなければならない。相手により態度や言葉遣いを変えたり、異なるパーソナリティを見せることは mixi 上でも可能であるが、マイミク上の一定数の友人と少なくとも良好な関係を維持するためには、そこに提示する自己に一貫性が必要であり、複数の他者との安定的な関係維持が重要になる。参加のための紹介制度、匿名性の原則などを mixi が安心して参加できるサイトである根拠とする論もあるが、むしろ重要なのは関係の開示による信頼性の担保であると、我々は考える。生産現場における「見える化」が「計測可能性」から「管理可能性」へつながるように、SNS においては、他者との「関係の可視化」が、不特定多数と自己の「差別化」(自分は普通の人間であり、このような友人・知人と社会生活を営んでいる)と「信頼性の確保」(これらの人々を傷つけたり、恥をかかせることのないよう誹謗中傷やいやがらせなどはしない人間

は、東京大学大学院経済学研究科新宅純二郎先生のご紹介に預かりました。ここに御礼もうしあげます。

⁴ はっきりした理由は不明である。

であると信頼してほしい) という力になっている。

3. 友人関係の構造

3.1. 次数分布、 C 、 L

友人関係の紐帯の次数分布を示したものが図1である。全体の23.6%にあたる85,851人が友人を1人だけしかもっていない。また友人が2人以下の者は全体の35.7%、3人以下の者は全体の44.3%である。友人数の最大値は1,301人であるが、mixiでは2004年10月以降、マイクシィの数を1,000人に制限している。mixiの友人関係構造がスケールフリー性を持つことは湯田らも確認しているが(湯田, 小野, 藤原, 2006)、この図からも友人数の分布にスケールフリー性を確認できる。

友人関係のネットワークは、全体で1,215の連結コンポネントに分かれており、最大のものは360,801人のユーザから構成される。2番目に大きいものでもノード数は16であり、ほかのコンポネントはそれ以下である(ノード数が2個のコンポネントが895個であった)。

用いたデータでは、各ユーザが登録を行った順序関係はおおよそ保存されている。それを利用して、 C と L の値がユーザ数の増加とともにどのように変化したかを示したのが表1で

図1 次数分布

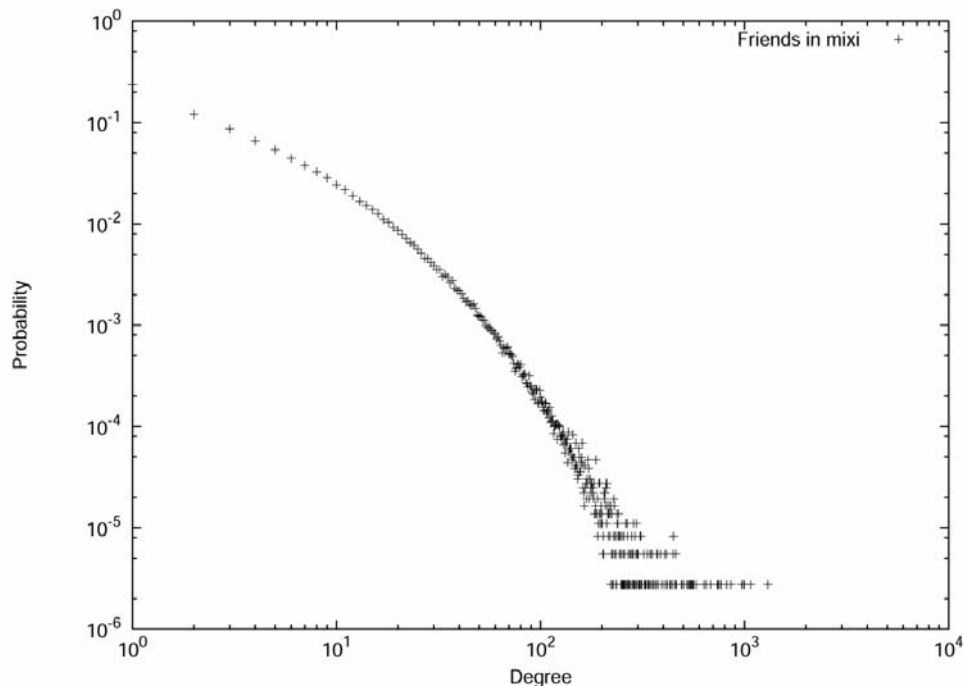


表1 ユーザ数とC、Lの変化

人数	C	L	最大コンポネント
2,500	0.349	3.951	2,466人
5,000	0.382	4.082	4,933人
10,000	0.378	4.279	9,895人
20,000	0.372	4.454	19,798人
30,000	0.366	4.524	29,694人
50,000	0.353	4.634	49,491人
70,000	0.349	4.729	69,319人
100,000	0.344	4.847	99,050人
150,000	0.337	4.998	148,694人
200,000	0.334	5.124	198,269人
300,000	0.333	5.335	297,528人
363,819	0.328	5.528	360,801人

ある。人数が10,000人という行は、初期ユーザ10,000人に限った場合のネットワークの諸量を表している。完全にその時点に遡っているわけではなく、初期ユーザ10,000人がその10,000人内で後から形成した友人関係もあるので、正確な時系列の変化を表しているわけではないが、おおよその概略をつかむことができる。

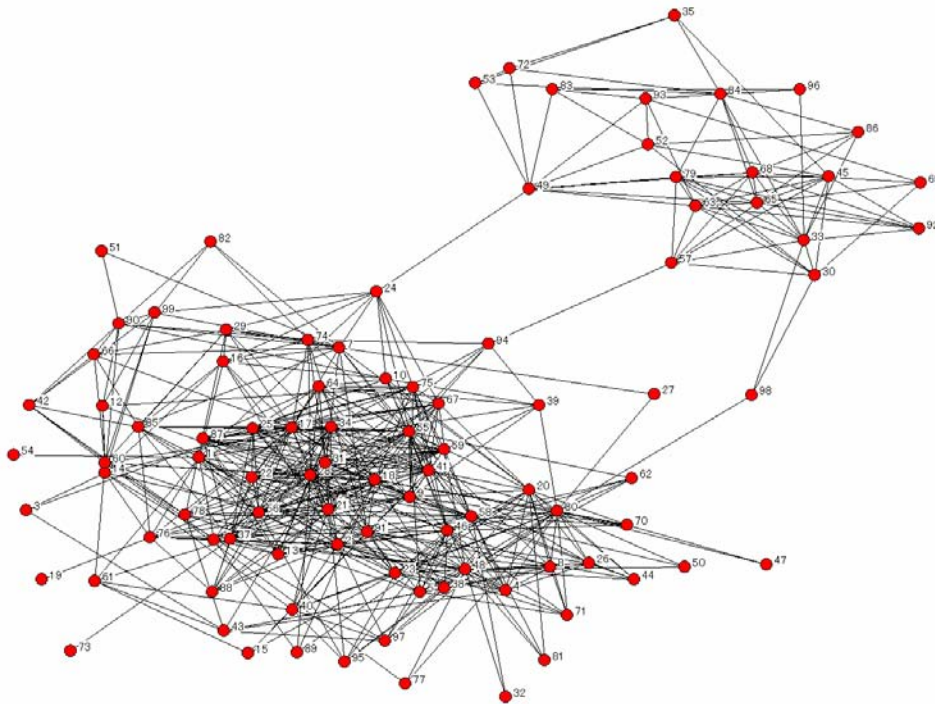
3.2. 中心性上位のネットワーク

次に、mixi内部に多数の友人関係を持つ中心的なユーザのみを抽出し、そのネットワークを調べる。301人以上の友人を持つユーザはわずか99人であり、全体の0.027%ときわめて少ない。この99人は、いわばmixiにおける著名人であって、立ち上げからのメンバーを含んだネットワークのコアであると予想できる。その友人関係を図示したものが図2である。

このネットワークでは、内部に大きく二つのクレークが存在する。中心的なメンバーの間でもかなり密な関係が形成されていることがわかる。

やや範囲を緩めて201人以上の友人をもつユーザに限定すると、mixi内では365人、全体の0.100%である。この365人の相互関係のネットワークでは、Cは0.323でありLは2.877であった。101人以上の友人を持つユーザは2,564人おり、Cは0.233、Lは3.139となる。

図2 300人以上の友人を持つ中心的なユーザ同士の友人関係



理由は定かではないが、mixiには、中心的ユーザが形成する大きな二つのクリーク構造が存在することは興味深い事実である。

3.3. コミュニティの構造

友人関係は承認の連鎖として成長していくのに対して、ユーザの関心により成長するのがコミュニティである。多数のユーザの関心を集めたコミュニティは拡大し、また類似した関心を持つユーザは共通のコミュニティに集まる。コミュニティの共通所属は友人関係の契機になりうると同時に、コミュニティ相互の関連性を形成する。

表2は、メンバー数の上位20位までのコミュニティ名及び所属人数を示している。

図3は、ユーザがどのくらいの数のコミュニティに所属しているかの分布を示している。各ユーザの所属コミュニティ数の平均は24.97個であり、所属コミュニティの数が増えるにしたがって人数は少なくなる。これは所属コミュニティ数が50、100、200などのプロットである。特に、所属コミュニティ数が1,000近辺で大きく上に上がっているが、これは2004年10月から所属コミュニティの上限を1,000としているため、所属コミュニティの数が1,000（もしくは999、998）のユーザ数が多くなっているためである。

表2 コミュニティのメンバー数上位

	コミュニティ名	人数	中心性*
1	資料になりそうなウェブサイト	10,238 人	1
2	Mac ユーザー	8,826 人	3
3	まったくわけがわかりません！	7,787 人	6
4	Photoshop	7,407 人	2
5	クリエイター・デザイナー	7,163 人	5
6	iPod User's	7,061 人	11
7	面白ネタで笑おう！	7,022 人	4
8	美術館・博物館 展示情報	6,824 人	13
9	空を見る人	6,581 人	7
10	料理作るのが好き	6,467 人	15
11	笑える画像	5,828 人	12
12	フォント	5,516 人	8
13	にゃんこ組	5,342 人	38
14	O 型	5,337 人	27
15	Illustrator	5,309 人	9
16	ヴィレッジヴァンガード	5,107 人	18
17	名前覚えられません	4,892 人	16
18	めんどくさい	4,850 人	10
19	カレー大好き	4,849 人	22
20	水曜どうでしょう	4,831 人	118

*) 固有ベクトル中心性による順位

コミュニティのメンバー数の分布を示したものが図4である。各コミュニティのメンバー数の平均は66.40人であり、所属人数が最大のコミュニティは10,238人、所属者1人だけのコミュニティは5,395個であった。一部のコミュニティで非常にメンバー数が多く、メンバー数が1,000を超えるものは全体の7.9%だけである。

分析に用いたデータではコミュニティの時系列の順序関係が保存されており、これを利用してコミュニティの新しさと人数の関係を示したものが図5である。

次にユーザがどのようなコミュニティに所属しているかを使って、コミュニティ間の関連を見てみよう。対象としたのはメンバー数の上位200位までのコミュニティで、メンバー数は最大10,238人から最小で2,016人、mixi全体のコミュニティの延べ人数の10.6%を構成する。ユーザとコミュニティの所属関係は二分グラフを形成するが、これをコミュニティ間の関係性に縮退させ、ネットワークを抽出する。

このネットワークを本論文ではコミュニティ連関ネットワークと呼ぶ。コミュニティ間の

図3 ユーザの所属コミュニティ数の分布

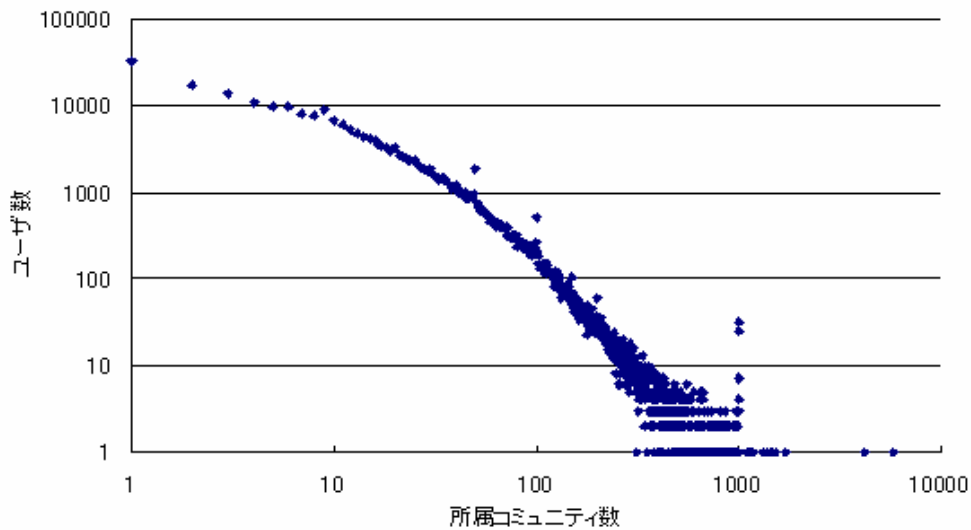
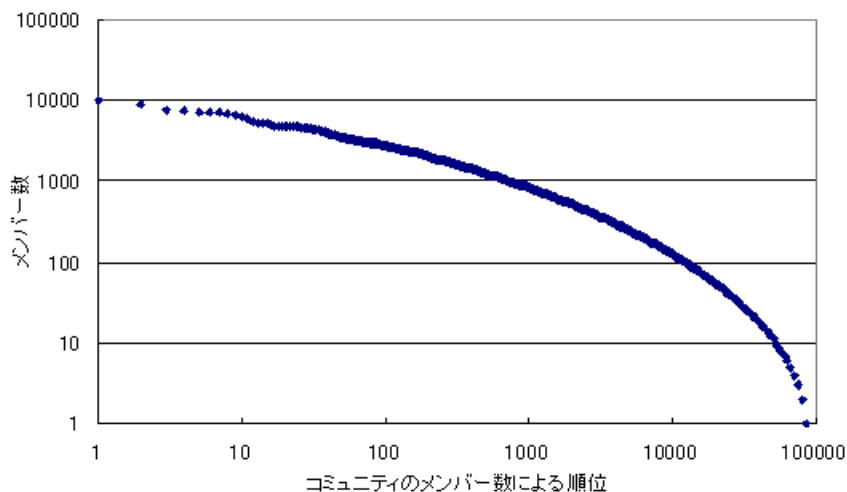


図4 コミュニティのメンバー数の分布



関連度は Jaccard 係数を用いて計算する。

図6に、コミュニティの連関ネットワークを示す。Jaccard 係数が0.2以上の紐帯のみを表示している。

このネットワークでの中心性(固有ベクトル中心性)を表2に示している。メンバー数の多いコミュニティは中心性も高い傾向があるが、例えば「にゃんこ組」(中心性38位)や「水曜どうでしょう」(中心性118位)など、中心性が低くてもメンバー数が上位20位までに入っているものもある。

図5 コミュニティの新しさとメンバー数

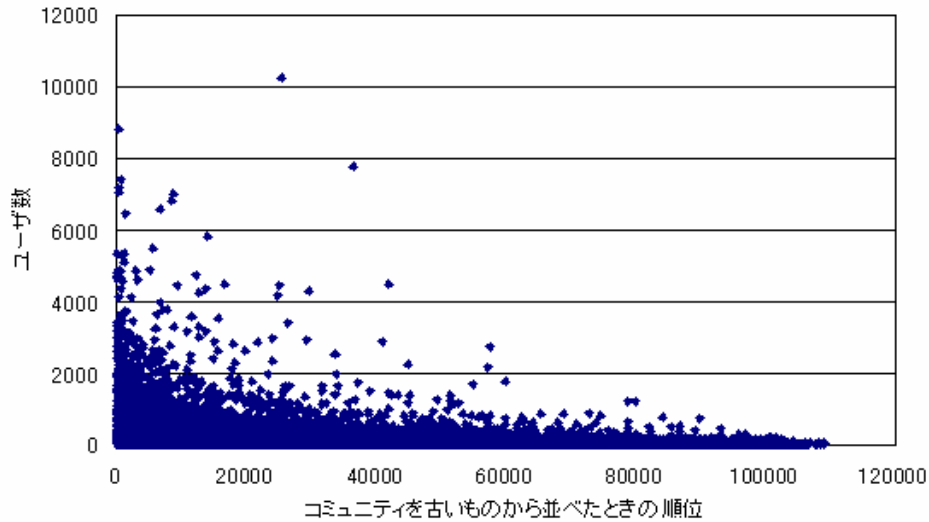
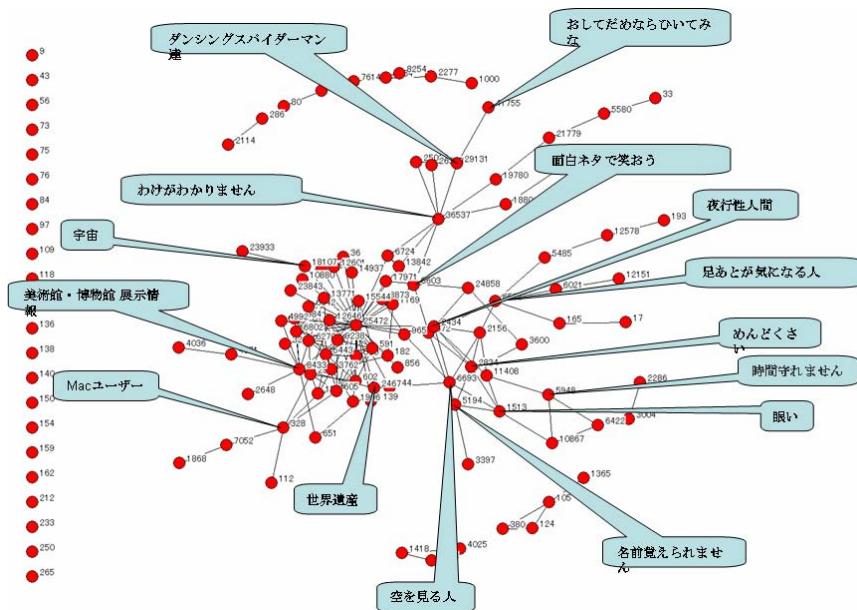


図6 コミュニティの連関構造



さらにこれらのコミュニティのテーマをブロックモデルという手法 (安田, 2001) によりクラスタリングし、ラベル付けしておおまかに分類した結果が図7である。8分割の時点で、共通要素系、共通嗜好系、芸能人系、面白ネタ系、趣味系、TV系、Mac系、デザイン系の

図7 ブロックモデルによるコミュニティの分類

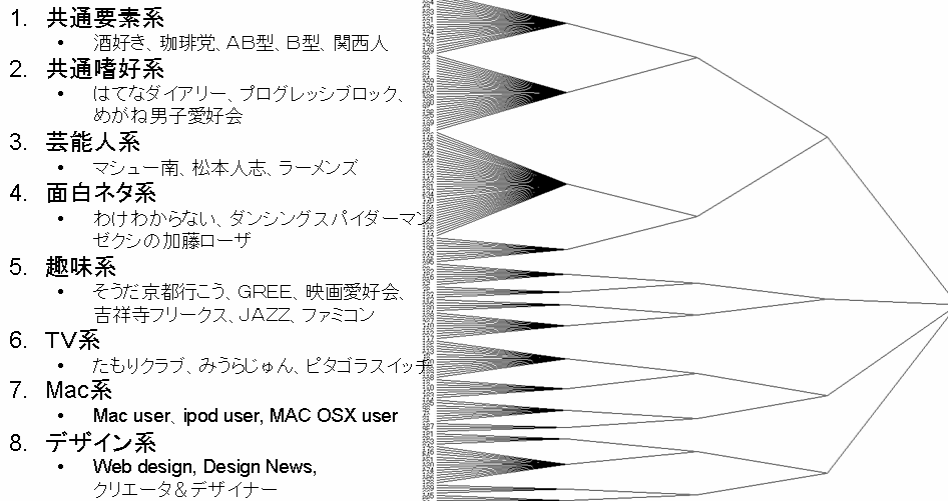
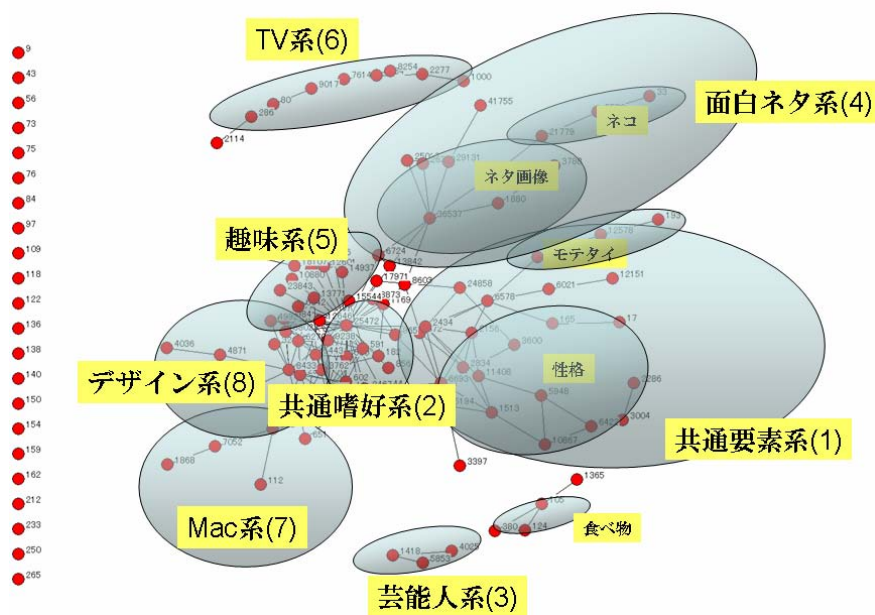


図8 コミュニティの領域図



八つに分けられる。それぞれがネットワーク図のどのあたりの領域であるかを図8に示す。2005年3月の時点でのコミュニティは、このような領域をカバーしながら相互に関連の構造を形作っていたわけである。

4. 友人関係とコミュニティ

mixi では、友人登録とコミュニティへの参加が相互発展的に起こるので、友人関係と共通のコミュニティがあるかどうかはある程度相関している。友人関係のある・なしと、共通のコミュニティのある・なしの分割表を作成したものを表3に示す。あまりに全体の数が多いので、友人関係がない 共通のコミュニティがないという負の相関は読み取りにくいですが、友人関係がある 共通のコミュニティがあるという正の相関ははっきりと読み取れる。

では、コミュニティによっては友人関係が多いものとそうでないものがあるのだろうか？
 n 人から成るコミュニティであれば、最大で $n(n-1)/2$ 組の友人関係が存在し得る。このうち、実際に友人関係が存在する割合を、本論文ではコミュニティの結合性 (CC: community connectedness) と呼ぶことにする。結合性が1であれば、コミュニティ内の全てのメンバーは友人関係であり、0であればどの2人をとっても友人関係でない。

図9は、コミュニティのサイズと結合性を調べたものである。コミュニティのサイズが大

表3 全ユーザ間の関係における分割表

	共通コミュあり	なし	計
友人関係あり	2.04×10^7	1.77×10^7	3.81×10^7
なし	2.80×10^{10}	1.30×10^{12}	1.32×10^{12}
計	2.80×10^{10}	1.30×10^{12}	1.32×10^{12}

注) 全ユーザ数 \times (全ユーザ数 - 1) 1.32×10^{12} である

図9 コミュニティのサイズと内部での結合性

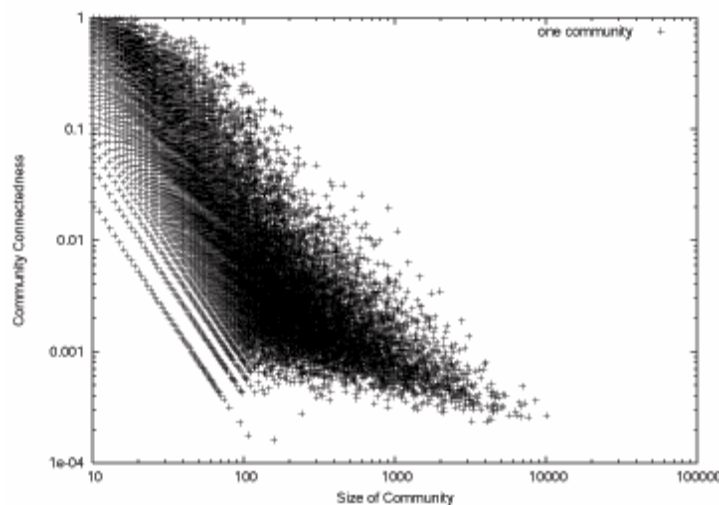


表4 結合性の強いコミュニティ

	コミュニティの概要	人数	結合性
1	Keio SFC のダンス系のコミュニティ	97	0.377
2	鍋と焼肉の集まり	97	0.334
3	割烹のお店の集まり	56	0.539
4	ホームパーティのコミュニティ	84	0.356
5	ある人の家でときどきご飯を食べる会	45	0.656

表5 結合性の弱いコミュニティ

	コミュニティの概要	人数	結合性
1	常盤貴子	107	0.000176
2	バイオハザード4(ゲーム)	94	0.000229
3	カレーうどん	80	0.000316
4	マカロニほうれん荘(漫画)	158	0.000161
5	鷺沢萌(作家)	74	0.000370

きくなると一般的に結合性は弱くなる。しかし、同じサイズのコミュニティでも、結合性が大きなものもあれば小さなものもある。図中では、右上に近いものがサイズの割に結合性が高い。

そこで、結合性 CC を n で割った $f=CC/n$ を指標とし、その値の上位5位までのコミュニティを示したものが表4である。⁵ つまり結合度が高いのは、リアルな空間での関係とmixi上でのコミュニケーションが融合し、うまくサイクルとなって回っているコミュニティであった。対照的に、表5は指標が最も小さいコミュニティである。純粋にユーザの趣味や興味の対象であり、友人関係の契機となるのが難しいという印象をもつコミュニティである。

以上見てきたように、コミュニティは、リアルな友人関係を支える結合性の高いコミュニティから、純粋に興味だけでつながっている結合性の弱いコミュニティまで幅広く存在する。もちろん、結合性が高ければよいというものではなく、情報収集という観点からは結合性が低くても十分にその機能を果たしている場合もあるだろうし、弱い紐帯の強さという観点で言えば、結合性が低いコミュニティにこそ貴重な友人関係の可能性があると言えるかもしれない。いずれにしても、mixiのコミュニティ機能は、ユーザの多様なニーズに応じてうまく機能しており、結果的に独特の生態系を形作っていると言えるだろう。

⁵ コミュニティが比較的小規模であるため、コミュニティ名は示していない。

5. 議論

SNS では、友人関係を形成することによるインタラクションの負荷が小さい。社内での職務遂行上の関係や地域での友人関係は、関係構築にも負荷がかかり、関係維持にはさらに継続的な負荷がかかる。それに対して mixi では、わずか 2 クリックで関係が形成できるし、ログインしなければ関係維持の負荷も 0 である。また、これまで社会ネットワーク分析で扱われてきた企業の取引関係や産業連関ネットワークと比較して、SNS 上での関係は、経済合理性の影響をほとんど受けていない。だからこそ、形成されたネットワークには何らかの意味があり、合理性の観点から分析の対象となり得る。

しかし、SNS を日常的に使っているユーザにとって、関係維持の負荷が皆無というわけではない。友人の日記にコメントしたり、友人に自分の所属するコミュニティを見せる（見られる）ことは、多少なりとも社会的な圧力となる。ユーザの所属する集団によっては、友人のメッセージに答えないというのは不評を買うかもしれないし、友人の日記を読んでいないとリアルな世界での友人関係に支障があるかもしれない。したがって、SNS では SNS の世界の中での規範が存在する。この SNS が内部にもつ社会性こそが、SNS のネットワークを他の社会ネットワークと差異づけ、独特の形成原理の要因となるものであろう。

本報告では、mixi 上の「マイミクシィ」を友人関係と呼び替えて議論を進めてきた。しかし、上述の意味では、友人関係とマイミクシィは全く異なるものである。マイミクシィに登録するのは、自己言及的になるが、マイミクシィに登録してもいいと思うからであって、それは mixi 上の自分のさまざまな行動と関係している。

本論文では、友人関係とコミュニティに関してさまざまな角度から分析結果を報告した。こういったコミュニティは常に多くの人の参加を呼び、時間とともに大きくなる。時間とともにユーザが増えて規模が大きくなったコミュニティは、さらに特化したコミュニティを生み出す（図 5）。一方で、友人関係（特にリアルの世界での友人関係）をきっかけにしたコミュニティもある。

このようなコミュニティの像が正しいかどうかは、今後の mixi やその他の SNS の分析を待たなければならない。しかし、本論文の分析から、SNS におけるユーザやコミュニティの生態系のある側面が垣間見えるのではないだろうか。

6. まとめ

本報告で明らかにしたように、mixi のネットワークは、ユーザ数が 36 万人の時点でクラ

スタ係数 C が 0.328、平均パス長 L が 5.53 であった。 C は高く L は小さく、この値は SNS としての優秀さ、すなわち友人の局所的連結機能と、効率の高い情報伝播機能を示す値ではないかと思う。また、コミュニティと友人関係は相互の関連を持っており、その両者を考慮することで、ユーザが自分の友人から友人関係を広げていること、結合性の高いコミュニティではオフ会が重要な役割を果たしていることなどを示した。

mixi では、コミュニティの検索や人の検索は、(技術的には) 比較的単純な機能に限られている。もちろん、プライバシーを守るという実際的な問題のためであると思うが、この検索機能が限られていることが、逆に、友人関係のつながりによる友人の探索を促し、またコミュニティの連関を生み出し、興味深いネットワークを作り出しているとも見える。

このように、システムのデザインがユーザの挙動に与える影響は少なくないはずであり、こうした影響についてより一般的な知見を得ることは、SNS をはじめとするコミュニティ Web プラットフォームで今後ますます重要になる課題であろう。

参考文献

- Adamic, L., Buyukkokten, O., & Adar, E. (2003). A social network caught in the web. *First Monday*, 8(6).
http://firstmonday.org/issues/issue8_6/adamic/index.html
- Backstrom, L., Huttenlocher, D., Lan, X., & Kleinberg, J. (2006). Group formation in large social networks: Membership, growth, and evolution. *Proceedings of the Twelfth Annual SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- Kautz, H., Selman, B., & Shah, M. (1997). Referral web: Combining social networks and collaborative filtering. *Communications of the ACM*, 40(3), 63-65.
- Matsuo, Y., Ohsawa, Y., & Ishizuka, M. (2003). Average-clicks: A new measure of distance on the world wide web. *Journal of Intelligent Information Systems*, 20(1), 51-62.
- 松尾 豊 (2006). 「Web2.0 時代の個人とコラボレーション」『情報処理』47(11), 1229-1236.
- Mika, P. (2005). Flink: Semantic web technology for the extraction and analysis of social networks. *Journal of Web Semantics*, 3(2-3), 211-223.
- Milgram, S. (1967). The small-world problem. *Psychology Today*, 2, 60-67.
- Mori, J., Ishizuka, M., Sugiyama, T., & Matsuo, Y. (2005). Real-world oriented information sharing using social networks. *Proceedings of ACM Group 2005 Conference*.
- 沼晃介, 大向一輝, 濱崎雅弘, 武田英明 (2004)「Weblog におけるエゴセントリック検索の提案と実装」第 6 回人工知能学会セマンティックウェブとオントロジー研究会.
- 大向一輝, 松尾豊, 松村真宏, 武田英明 (2006)「Community Web プラットフォーム」『人工知能学会論

文誌』3(21), 251-256.

Spertus, E., Sahami, M., & Buyukkokten, O. (2005). Evaluating similarity measures: A large-scale study in the orkut social network. *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2005*.

安田雪 (2001) 『実践ネットワーク分析』新曜社.

湯田聡夫, 藤原義久 (2005) 「SNS における人のネットワーク構造: その地平線の超え方」Web が生み出す関係構造と社会ネットワーク分析ワークショップ.

湯田聡夫, 小野直亮, 藤原義久 (2006) 「ソーシャル・ネットワーキング・サービスにおける人的ネットワークの構造」『情報処理学会論文誌』47(3) 865-874.

赤門マネジメント・レビュー編集委員会

編集長 新宅 純二郎

編集委員 阿部 誠 粕谷 誠 高橋 伸夫 藤本 隆宏

編集担当 西田 麻希

赤門マネジメント・レビュー 6巻4号 2007年4月25日発行

編集 東京大学大学院経済学研究科 ABAS/AMR 編集委員会

発行 特定非営利活動法人グローバルビジネスリサーチセンター

理事長 高橋 伸夫

東京都千代田区丸の内

<http://www.gbrc.jp>