

物理の現象論に学ぶ

——通信ネットワークに現れるべき乗則を利用した 社会ネットワーク構造の解明——

Structures of Social Networks and User-dynamics Revealed by Power Laws : Inspired from Phenomenology

会田雅樹

Abstract

情報通信ネットワークの利用履歴等の各種データには、背後にある社会ネットワークの特性が何らかの形で反映しているはずである。本稿では、物理学で用いられる現象論的な考え方により、情報通信ネットワークの様々なデータに現れるべき乗則を解析し、その解析結果から社会ネットワークの構造やユーザの行動規則で見られるグラフの構造を読み解く試みを紹介する。

キーワード：トラヒック、べき乗則、グラフ、社会ネットワーク、スケールフリー性

1. はじめに

物理学の「ものの見方」には、一見複雑に感じる雑多な現象に対して、それらの背後にある比較的単純な基本法則で統一的理解ができることが望ましい、という美意識なり世界観がある。このような「ものの見方」は、目の前で起きている現象をそのまま記述するものではないため、工学的に直接役立つことは少ないかもしれない。しかし、背後にある基本法則を理解することは、いまだ見ぬ状況変化などによって引き起こされる未知の現象を予測するための知見として、工学の立場でも有用であると考えられる。このような「ものの見方」を具体化する分析方法の一つが「現象論」と呼ばれるアプローチである。これは、目の前で起きている現象や実験結果を正しい仮説として受け入れ、そこから論理的に導かれる結果として背後の基本法則の構造を探る方法である。

情報通信ネットワークの利用履歴等の各種データは、背後にある社会ネットワークの特性を何らかの形で反映しているはずである。しかし、通信履歴自体の中身を詳細に分析するだけでは、例えば組織内の人間関係や法人

間の取引といった「ネットワークの中で起こっている個別の状況」の理解には有効であっても、背後にある社会ネットワークの全体像を理解することは難しい。本稿では、通信ネットワークに現れる各種のべき乗則を基にして、それらの出現理由を背後にある社会ネットワークの特性に求めることで、社会ネットワークの全体像の理解を目指す試みを解説する。ここで扱うべき乗則は、細部の構造が分からない粗視化されたデータであるため、個々のユーザの詳細な振舞いを見ることはできないが、細部の構造によらないより普遍的で一般的な構造が潜んでいることが期待できる。このような分析は、直接的な応用を目指した工学的な必要性というより、マクロレベルの実験事実から背後の基本法則を探る物理学の現象論的な立場に近い。得られた構造は、直接的には「個別の状況」を記述するものではないが、ミクロレベルの個別状況によらない普遍的な性質を把握することは、いまだ見ぬ新サービスの展開やそれに伴う新たな技術課題にアプローチする際の指針を与える可能性がある。

ここで、本研究においてデータのべき乗則に着目する意図を補足しておく。べき乗則には、指数分布や正規分布のような「分布の出現要因が何らかのランダム性に起因するもの」とは異なり、確定的な出現要因の存在が期待できる。このため、べき乗則の出現要因を探索することで、探索の過程がランダムな要因によって乱されることなく、何らかの別の性質に対応することが期待できる。

会田雅樹 正員 首都大学東京大学院システムデザイン研究科経営システムデザイン専修

E-mail maida@sd.tmu.ac.jp

Masaki AIDA, Member (Graduate School of System Design, Tokyo Metropolitan University, Hino-shi, 191-0065 Japan).

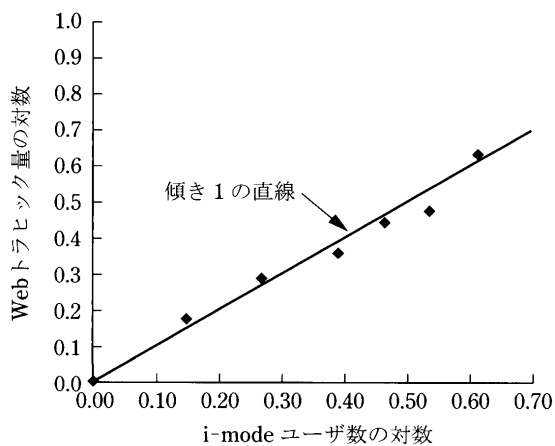
電子情報通信学会誌 Vol.91 No.10 pp.891-896 2008年10月

本稿では、NTT DoCoMo の i-mode サービス⁽¹⁾と、(株)ミクシィにより提供されている日本最大手のソーシャルネットワーキングサービスである mixi⁽²⁾に現れるべき乗則に注目し、背後にある社会ネットワーク構造を探る。

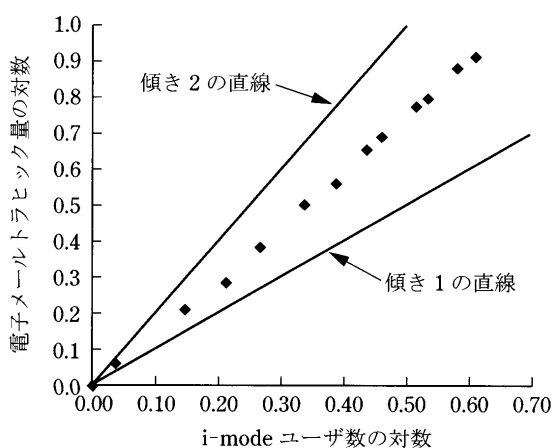
2. 通信ネットワークに現れる各種のべき乗則

本章では i-mode と mixi に現れるべき乗則を紹介し、それらのデータの重要性について述べる。

まず、i-mode のサービス開始後の初期普及期間における、ユーザ数とトラフィック量の関係を表すデータを図 1 に示す。このデータは 1999 年 8 月～2000 年 1 月末までの半年間の初期普及期間のもので、この間に i-mode ユーザ数は約 1,290,000 から 3,740,000 までほぼ 3 倍に伸びている。この期間で i-mode ユーザ数と Web アクセスのトラフィック量（アクセス回数）の関係を調べると、i-mode ユーザ数を n としたとき、



(a) Webトラフィック



(b) 電子メールトラフィック

図 1 両対数表示による i-mode ユーザ数とトラフィック量の関係 (a)は Web トラフィック、(b)は電子メールトラフィックの結果。縦軸及び横軸は、1999 年 8 月 1 日のトラフィック量と i-mode ユーザ数をそれぞれ 1 とし、2000 年 1 月 31 日までの 6 か月間について i-mode ユーザ数とトラフィック量の関係を対数（底は 10）で表示。比較のため傾き 1、2 の直線を併せて表示。

$$(\text{Web トラフィック量}) \propto n \quad (1)$$

となる(図 1 (a))。ユーザー人当りの平均 Web アクセス回数が一定であれば、この結果は明らかである。逆にいえば、この期間でユーザの平均的な i-mode 使用特性は変化していないことを示唆していて、初期に加入するユーザがヘビーユーザであるといったような偏りは見られない。これを前提として、同期間の電子メールトラフィック量（メッセージ数）を調べると、

$$(\text{電子メールトラフィック量}) \propto n^{1.6} \quad (2)$$

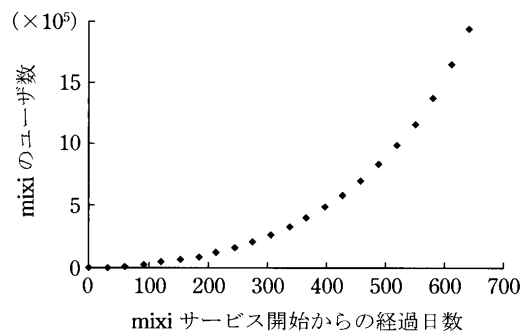
のようなべき乗則となる(図 1 (b))。

次に、mixi がサービスを開始した後のユーザ数の経時変化を図 2 (a)に示す。横軸はサービス開始からの経過日数、縦軸は mixi ユーザ数である。これを両対数表示したものが図 2 (b)である。mixi ユーザ数が時間に対してべきで増加することは既に報告されている⁽³⁾。初期条件に左右されるサービス開始直後を別にとすると、サービス開始からの経過時刻 t に対する mixi ユーザ数 $m(t)$ は

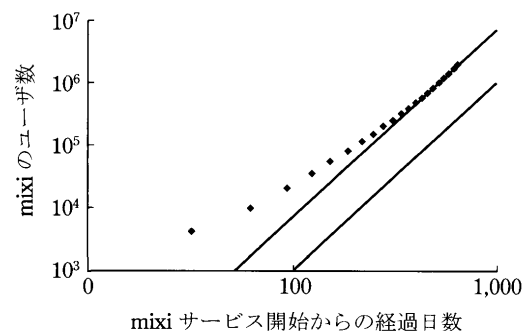
$$m(t) \propto t^3 \quad (3)$$

が成り立つことが確認できる。

本稿でのデータ分析の目的は、i-mode や mixi といっ



(a) 線形スケールでの表示



(b) 両対数表示

図 2 mixi ユーザ数の経時変化 (a)は線形スケールでの表示、(b)は両対数表示で比較のため傾き 3 の直線を併せて表示。

た特定のサービスの利用状況を調べるのではなく、これらのサービスの背後にある人間（日本人）の社会ネットワーク構造を調べることにある。つまり、個別の通信サービスによらない「社会ネットワーク」という実体が存在するとして、具体的な通信サービスによってデータの中に顕在化する「社会ネットワーク」の構造の断片から、「社会ネットワーク」構造の全体像を抽出しようとしている。上記べき乗則のデータは、社会ネットワーク構造把握の目的を遂行するのに適した以下の特徴を持っている。

- i-mode と mixi は（固定電話等に比べ）短期間に爆発的に普及したため、ライフスタイルの変化等の外部要因がデータに与える影響は小さい。
- ほとんどの携帯電話端末はユーザ個人が独占的に使用するため、i-mode 端末間のトラヒックは端末ユーザ間の情報交換が支配的である。
- ほとんどの i-mode メールは一対一の通信であるので、電子メールトラヒック量と情報交換を行う関係にある i-mode ユーザペアの数には密接な関連がある。
- 新たに mixi ユーザになるには既存ユーザからの紹介が必要であるため、mixi の普及過程は社会ネットワークのリンクと密接に関係している。

以降では、i-mode と mixi に現れるべき乗則を基にして、これらのサービスの背後にある「共通の」社会ネットワーク構造を調べる^{(4), (5)}。

3. i-mode の分析

3.1 記号の定義と問題設定

i-mode ユーザ以外の人間を含む一般の人間（通信サービスの潜在的ユーザ）の集合を V 、 V に含まれる人間の数を $|V|=N$ 、何らかの形で情報交換がある人間のペア $(x, y) \in V \times V$ （リンク）の集合を E とし、人間の情報交換関係（社会ネットワーク）を表すグラフを $G(V, E)$ とする。 $G(V, E)$ は時間とともに変化がないと仮定する。

次に、i-mode ユーザ（ノード）の集合を $V_i(n)$ ($V_i(n) \subset V$)、i-mode のユーザ数を $|V_i(n)|=n$ 、i-mode ユーザの集合 $V_i(n)$ によって $G(V, E)$ から誘導される誘導部分グラフを $G_i(V_i(n), E_i(n))$ とし、 $E_i(n)$ は $G_i(V_i(n), E_i(n))$ のリンクの集合とする。

Web トラヒックの振舞い（式（1））によると、i-mode ユーザ数が増加してもユーザー一人当り Web トラヒック量は変わらないことになる。一方で、式（2）のように電子メールトラヒック量が $n^{1+\alpha}$ ($\alpha \approx 0.6$) となるのは、ユーザが増えることでユーザの中の知人数が増え、電子メールの送信相手が増加する効果によると考えられる。つまり、 α の大きさは社会ネットワークの接続の強さを反映した量であると考えられる。電子メールトラヒックについて、リンク当りの平均トラヒック量は n によらず一定であると仮定する。この仮定の妥当性については、携帯電話の音声通信トラヒックのデータを用いて間接的に確認済みである⁽⁵⁾。このとき、式（2）の条件は

$$|E_i(n)| = O(n^{1+\alpha}) \quad (4)$$

となる。

繰返しになるが、ここでの課題は、i-mode ユーザ間の社会ネットワーク $G_i(V_i(n), E_i(n))$ を調べるのではなく、i-mode のトラヒックデータを利用して i-mode ユーザ以外の人間を含む一般の社会ネットワーク $G(V, E)$ を調べることである。図 3 は $G(V, E)$ と $G_i(V_i(n), E_i(n))$ の関係を示したものである。一番上のグラフは一般の社会ネットワーク $G(V, E)$ を表し、そこから i-mode ユーザのみを抜き出して作った i-mode ユーザ間の社会ネットワーク $G_i(V_i(n), E_i(n))$ が一番下のグラフである。i-mode のユーザ数と電子メールトラヒック量から得られる情報は、i-mode ユーザ間の社会ネットワーク $G_i(V_i(n), E_i(n))$ におけるノード数とリンク数の関係（式（4））である。この関係を実現するための条件として、一般の社会ネットワーク $G(V, E)$ の構造や新規 i-mode ユーザの加入規則を考察する。

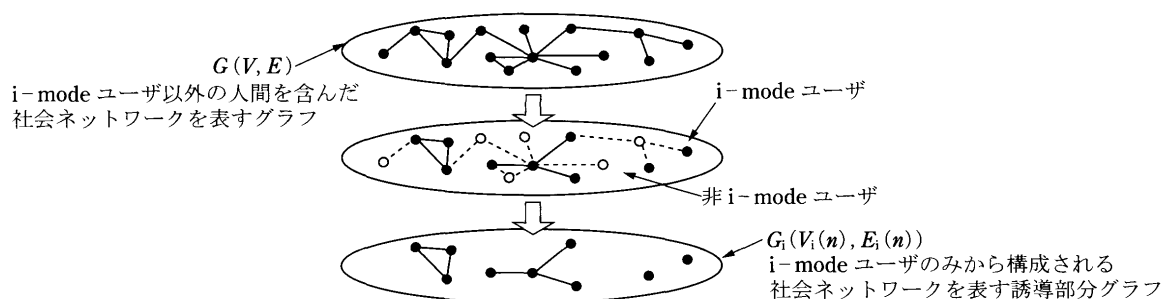


図 3 社会ネットワークの構造を表すグラフ $G(V, E)$ と i-mode ユーザによって誘導される部分グラフ $G_i(V_i(n), E_i(n))$ の例

3.2 i-mode 新規ユーザの加入規則と社会ネットワークの構造

V の要素を $G(V, E)$ に関して次数 (各ノードの持つリンク数) の大きい順に並べ, j 番目の要素の次数を $D(j)$ とする.

更に, V の要素を i-mode 加入順に並べたときの, h 番目の要素の $G(V, E)$ に関する次数を $D_i(h)$ とし, $V_i(n)$ の要素を i-mode 加入順に並べたときの, h 番目の要素の $G_i(V_i(n), E_i(n))$ に関する次数を $d_i(n, h)$ とする.

i-mode ユーザの $G_i(V_i(n), E_i(n))$ に関する次数と, $G(V, E)$ に関する次数とを関係付けるため, $\sum_{h=1}^n d_i(n, h)$ と $\sum_{h=1}^n D_i(h)$ の比 $c_i(n)$ を導入して

$$\sum_{h=1}^n d_i(n, h) = c_i(n) \sum_{h=1}^n D_i(h) \quad (5)$$

とする. $c_i(n)$ は i-mode ユーザの知人が i-mode ユーザになっている割合を示し, $c_i(1)=0$, $c_i(N)=1$ となる単調増加関数である.

ここで, 関数 $c_i(n)$ の性質として, 以下のようなべきの振舞いを仮定する.

$$c_i(n) \propto n^{1-\delta} \quad (6)$$

ここで δ は定数である. もし $\delta > 0$ であれば, ユーザの知人のうち i-mode ユーザである割合は, 初期においてその増加率が高く, 加入順序が早いユーザ同士は互いに知人同士である確率が高いという, クラス的な構造があることを意味する. $\delta=0$ であれば, そのようなクラス的な構造は存在しない. 関数 $c_i(n)$ が式 (6) の構造を持つという仮定の妥当性と, 定数 δ の値は, 4. の mixi の分析で明らかになる.

式 (4), (5) 及び $|E_i(n)| = \frac{1}{2} \sum_{h=1}^n d_i(n, h)$ から,

$$\sum_{h=1}^n D_i(h) = O(n^{\alpha+\delta}) \quad (7)$$

となり, これがすべての n で成立すれば

$$D_i(h) = O(h^{\alpha+\delta-1}) \quad (8)$$

となる.

現時点では δ の値が決まらないが, 仮に $\alpha+\delta < 1$ であるとして, もう少し考察を進めてみる. この場合 $D_i(h)$ は h の増加に対して減少する傾向があるので, i-mode 加入順に並べた次数 $D_i(h)$ は, オーダの意味で次数の大きい順

$$D_i(h) \simeq D(h) \quad (\text{オーダの意味で}) \quad (9)$$

であり, h の増加に対して $D(h)$ はべきで減少することが分かる. したがって, $\alpha+\delta < 1$ としたとき以下の結論を得る.

- 社会ネットワーク $G(V, E)$ の次数分布は以下のべき乗則 (Zipf 則)

$$D(h) = O(h^{-(1-\alpha-\delta)}) \quad (10)$$

に従う.

- i-mode への加入は $G(V, E)$ に関して次数の高い人から順番に行われる傾向がある. つまり知人の多い人から順に加入する傾向がある.

社会ネットワークを表すグラフのノード次数が Zipf 則を満たす場合, 社会ネットワークがスケールフリーネットワークになることがいえる. スケールフリーネットワークとは, ノードの次数分布が以下のべき乗則を満たすグラフである^{(6),(7)}.

$$n_d(k) \propto k^{-\gamma} \quad (11)$$

ここで, k はノードの次数, $n_d(k)$ は次数 k を持つノードの数, $\gamma > 0$ は定数である. 社会ネットワークのグラフ構造はべき指数 γ が

$$\gamma = \frac{1}{1-\alpha-\delta} + 1 \quad (12)$$

で与えられるスケールフリーグラフとなる.

以上の議論からは δ の具体的な値については分からないが, 他の $c_i(n)$ に関する仮定と併せて 4. の mixi の分析で明らかになる.

4. mixi の分析

ここでは, i-mode の分析によって得られた結果に mixi のデータを使った別の観点からの分析結果を加えて考察することで, 社会ネットワーク構造のモデルを詳細化する.

mixi ユーザになるためには, 既に mixi のユーザである人から招待を受けなければならない. このため, mixi の普及過程は社会ネットワークのリンク構造と密接に関係していると考えることができる.

時刻 t での mixi ユーザ数 $m(t)$ に関するべき乗則 (式 (3)) から, mixi ユーザの増加率 dm/dt は

$$\frac{dm}{dt} \propto t^2 \quad (13)$$

となり, mixi ユーザの増加率を t を陽に含まないように $m(t)$ の関数として書けば

$$\frac{dm}{dt} \propto m^{2/3} \quad (14)$$

となる.

次に, mixi ユーザの次数を考える. i-mode でも mixi でも潜在的ユーザの社会ネットワーク $G(V, E)$ は同じであるとする. V の要素を mixi 加入順に並べ, h 番目の要素の $G(V, E)$ に関する次数を $D_x(h)$ とし, mixi ユーザのみで作られるグラフに関する次数を $d_x(m, h)$ とする.

i-mode の場合と同様に, $D_x(h)$ と $d_x(m, h)$ を関係付けるため, 関数 $c_x(m)$ を用いて

$$\sum_{h=1}^m d_x(m, h) = c_x(m) \sum_{h=1}^m D_x(h) \quad (15)$$

とする. $c_x(m)$ は, mixi ユーザの知人のうちの mixi ユーザの割合を示す. $c_x(m)$ は $c_x(1)=0$, $c_x(N)=1$ となる単調増加関数である.

mixi ユーザが持つ $G(V, E)$ に関するリンクのうち, mixi ユーザ同士をつないでいないリンクを外線と呼ぶことにする(図4). mixi ユーザ数が m のときの外線数は

$$\sum_{h=1}^m D_x(h) - \sum_{h=1}^m d_x(m, h) = (1 - c_x(m)) \sum_{h=1}^m D_x(h)$$

となる. mixi の普及はユーザによる招待に限られることを考えると, mixi ユーザ数の増加率は外線数に比例すると考えるのが自然である. つまり,

$$\frac{dm}{dt} \propto (1 - c_x(m)) \sum_{h=1}^m D_x(h) \quad (16)$$

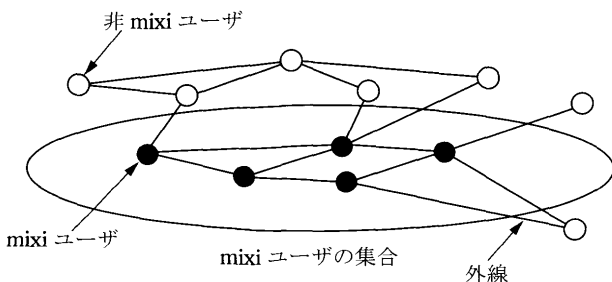


図4 mixi ユーザと非 mixi ユーザの社会ネットワークの例 外線は mixi ユーザの集合から外に向かう線で, この例では6本ある.

である. したがって, 式(14)により

$$(1 - c_x(m)) \sum_{h=1}^m D_x(h) \propto m^{2/3} \quad (17)$$

となる. $c_x(m)$ については m が小さい領域で

$$(1 - c_x(m)) = \text{constant} \quad (18)$$

としてよいので,

$$\sum_{h=1}^m D_x(h) = O(m^{2/3}) \quad (19)$$

となる. これが任意の m に対して成立すれば,

$$D_x(h) = O(h^{-1/3}) \quad (20)$$

を得る. したがって, mixi でもオーダの意味で加入順序は次数の大きい順

$$D_x(h) \simeq D(h) \quad (\text{オーダの意味で}) \quad (21)$$

であり, h の増加に対して $D(h)$ はべきで減少することが分かる. 更に式(19)と式(7)を比較することにより

$$\alpha + \delta = \frac{2}{3} \quad (22)$$

となる. また, i-mode の分析だけでは決まらなかった δ の値が

$$\delta = 0.066 \cdots \simeq 0 \quad (23)$$

と決定される.

ちなみに, ここまでの分析結果により, $c_i(n)$ と $c_x(m)$ はオーダの意味で同じ量を表していることが分かるので, i-mode の分析において使用した仮定,

- 関数 $c_i(n)$ が式(6)の構造を持つ ($0 < \delta < 1$).
- δ は $\alpha + \delta < 1$ を満たす.

が検証された.

i-mode と mixi の分析を組み合わせると, 両者のデータを共に満足するような社会ネットワークの構造は以下のような性質を持つことが分かる.

- 一般の社会ネットワークは, 次数分布が

$$n_d(k) \propto k^{-4} \quad (24)$$

となるスケールフリーグラフである.

- 人間同士は次数の大小による偏りがほとんどない知人関係を持つ ($c_i(n) \propto n$).

なお, 次数分布 (式(24)) の妥当性については, au by KDDI の提供する携帯電話サービスの音声通信トラヒックデータを用いた検証がなされている⁽⁸⁾. 個々の通信サービスのデータには, 社会ネットワーク構造の部分的な情報しか含まれていないかもしれないが, このような独立な通信サービスを用いた分析と検証によって, 社会ネットワーク構造のより詳細な「立体像」を浮かび上がらせることができるかもしれない.

5. お わ り に

本稿では, 二つの異なる通信ネットワークサービスに関するべき乗則を利用して, それらの通信サービスの背後にある社会ネットワーク構造を分析する方法を解説した. この分析は, 目の前で起きている現象をそのまま分析するのではなく, その背後に潜む構造を探るものである. このような考え方は, データが直接的に表している個々のサービスの特性を超えて, より普遍的で本質的な構造に迫ることができるかもしれない.

文 献

- (1) DoCoMo Net, i モードご利用状況 .
www.nttdocomo.co.jp/
- (2) mixi, Inc.
<http://mixi.co.jp/>
- (3) 湯田聡夫, 小野直亮, 藤原義久, “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,” 情報学論, vol. 47, no.3, pp.865-874, 2006.
- (4) M. Aida, K. Ishibashi, C. Takano, H. Miwa, K. Muranaka, and A. Miura, “Cluster structures in topology of large-scale social networks revealed by traffic data,” IEEE GLOBECOM 2005, St. Louis, 2005.
- (5) 会田雅樹, 佐々木 純, “通信サービスの普及プロセスを利用した社会ネットワークの構造分析,” 信学技報, IN2006-41, pp.37-42, July 2006.
- (6) A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” Science, vol.286, pp.509-512, 1999.
- (7) R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” Rev. Mod. Phys., vol.74, pp.47-97, 2002.
- (8) 会田雅樹, 小頭秀行, 中村 元, 高野知佐, “情報ネットワークサービスのデータを利用した社会ネットワークの次数分布の分析と検証,” 信学技報, Sept. 2008.

(平成 20 年 4 月 10 日受付 平成 20 年 7 月 8 日最終受付)



あいだ まさき
会田 雅樹 (正員)

昭 62 立教大・理・物理卒. 平元同大学院博士前期課程了. 同年 NTT 入社. 平 17 首都大東京・システムデザイン・准教授. 平 19 同大学院・教授. 博士 (工学).