

# 時系列ネットワークの可視化から人間行動のモデリングへ

From visualizing the time-series of social networks to modeling human behaviors

橋本康弘，陳昱，大橋弘忠

Yasuhiro Hashimoto, Yu Chen, Hirotada Ohashi

東京大学大学院工学系研究科システム創成学専攻

Department of Systems Innovation, School of Engineering, The University of Tokyo

**概要:** 近年，電子メールや電子掲示板，ブログ，オンラインアーカイブ，ソーシャルネットワークサービス（SNS）といったオンラインコミュニケーション手段の目覚ましい発展と普及に伴い，人間の社会的活動が次々とデータ化され蓄積されてきている．そしてこれら膨大なデータの山から人間の隠れた行動パターンを発見することが情報社会におけるイノベーション創出，あるいは潜在リスク発見の鍵となると期待されている．我々は電子掲示板や SNS 等におけるコミュニケーションのイベントログからソーシャルネットワークの時系列を構築し，コミュニティの生成や消滅といったソーシャルネットワークが持つダイナミックな構造をコミュニティ系統図として可視化する手法について提案する．そしてコミュニケーションデータの分析から人間行動をモデル化するための新しいアプローチについて議論する．

**Abstract:** In recent decades, a lot of data of various social activities of human beings has been accumulated through the development and spread of on-line communication tools such as e-mail, BBS, blog, on-line archives, and SNS. Consequently, it is expected to reveal the hidden patterns of human behaviors from those massive data for bringing innovation to the information-based society and revealing the potential risk in the complex social systems. We introduce a new method to visualize social community evolution as the community genealogy diagram that enables us to understand the dynamic structure of human networks using common data on social events such as the communication on BBS and SNS. Then, we discuss a new approach to model human behaviors based on the data analysis of human communication.

## 1. 序論

人間の社会的活動を理解する方法論としてネットワーク分析がある[1]．これは個人や組織といった社会的人格とその間の社会的相互作用をネットワークとして抽象化し，ネットワーク構造が持つ統計的性質から人間集団に普遍的に観測される一般法則を導き出そうとするものである．とくに Watts[2] や Barabási[3]らの研究以降，多様なネットワークシステムが示す普遍的性質が注目され，ソーシャルネットワークについても実証的研究が多くなされてきた[4-6]．しかし，平均距離や次数分布といったネットワーク構造全体が持つ統計平均量に着目した研究には限界がある．ソーシャルネットワークの分析にはコミュニケーションの具体性に基づく個別的な分析が不可欠だからである．

コミュニティ分析はこの両極の視点を仲介するも

のである．ネットワーク分析においてコミュニティとは内部では密に接続し，コミュニティ同士では疎な接続を持つサブネットワークを指し[7]，コミュニティ構造を明らかにすることによってネットワークは一段高い意味のレベルに粗視化され，全体構造の俯瞰が容易となる．本論では，ソーシャルコミュニケーションの時系列データからコミュニティの時間発展を可視化<sup>[8]</sup>によって明らかにする手法について紹介し，人間行動をモデル化する前段階としてのシナリオを語る枠組みについて議論する．

## 2. コミュニティ時系列の可視化

コミュニティ時系列の可視化は以下の手順で行う．

1. コミュニケーションの時系列データからネットワークの時系列を生成する
2. 時系列上のすべてのネットワークに対して

コミュニティ検出を行う

3. 連続する時刻においてコミュニティ同士を関連付ける
4. 時間軸方向に関連付けられたコミュニティを可視化する

1~3の基本的手順は[9]に従う。以下では各手順について簡単に解説する。

### 2.1. コミュニティ時系列の可視化

ネットワークを定義するコミュニケーションの例として電話での会話を挙げるならば、それは「いつ」「誰が」「誰に」電話をかけたのかというイベントのログから構築することができる。あるいは会話時間やその内容を用いてイベントを重み付けすることも可能である。したがって、ネットワーク時系列を構築する手順は以下ようになる。

1. ネットワーク時系列の時間間隔（時間の解像度）を決める
2. 各時刻に対する全イベントの寄与を重み関数を用いて計算する
3. 寄与が閾値を超えたノード間に対してエッジを作成する

手順2のイベントの寄与を定式化すると以下のようになる。

$$w_{mn}(t) = \sum_i f(t-t_i) \delta(x_i, m) \delta(y_i, n),$$

ここで  $i$  はイベントのインデックス、 $t_i$  はイベント  $i$  の発生時刻、 $t$  はネットワークを構成する時刻、 $x_i$  および  $y_i$  はイベント  $i$  が接続するノード、 $mn$  は任意の2つのノード、 $\delta(a, b)$  はクロネッカーのデルタを表す。代表時刻  $t$  において  $w_{mn}(t)$  が閾値  $w^*$  を超えたノード  $mn$  間に対してエッジを作成する。重み関数  $f(\Delta t)$  ( $\Delta t = t - t_i \geq 0$ ) の形としては、単純な矩形関数や指数関数的に減少する関数などが考えられる。あるいはコミュニケーションの具体性によっては時間反転性  $f(\Delta t) = f(-\Delta t)$  を仮定する場合もある。重み関数の決定には任意性があり、結果に与える効果を確かめながらトップダウン的に決めることになる。

### 2.2. コミュニティ検出

ネットワークからコミュニティ構造を検出する手法については代表的なモジュラリティ最適化法[7]以外にも様々な手法が提案されているが、手法によってコミュニティの定義が異なる。各コミュニティ検出手法にはそれぞれ特徴があり、対象とするネットワークの特徴に応じて適切に選択する必要がある。ここではクリークパーコレーション法（CPM）[9]について簡単に解説する。

CPMでは  $k$  個のノードからなる完全サブグラフを

$k$ -クリークと呼び、 $k-1$  個のノードを共有する  $k$ -クリーク同士は同一コミュニティに所属すると考える。こうして  $k$ -クリーク同士を可能な限り連結したノード集合を  $k$ -クリークコミュニティと呼ぶ。CPMの特徴として

1.  $k$ -クリークを構成しないノードはコミュニティ検出の結果から除外される
2. 複数のコミュニティに所属するノード（オーバーラップノード）が存在する

の2点が挙げられる。

### 2.3. コミュニティの関連付け

コミュニティ検出によって各時刻におけるコミュニティの集合を求めた後、連続する時刻においてコミュニティ同士を関連付ける必要がある。つまり時刻  $t$  の任意のコミュニティが時刻  $t+1$  のどのコミュニティに対応するのか（あるいはどのコミュニティにも対応しないのか）を判定する作業である。この判定はコミュニティに所属するノードの類似度を比較することで行う。類似度は以下で定義される。

$$\sigma(i, j) = |c_i(t) \cap c_j(t+1)| / |c_i(t) \cup c_j(t+1)|$$

$i$  は時刻  $t$  のコミュニティのインデックス、 $j$  は時刻  $t+1$  のコミュニティのインデックス、 $c_i(t)$  および  $c_j(t+1)$  は各時刻のコミュニティのノード集合を表す。 $c(t)$  と  $c(t+1)$  のすべての組み合わせに対して類似度が最も大きいコミュニティのペアから順に決定し、一方が既に決定しているペアはスキップする。この作業をすべての連続する時刻間で行うことで、任意のコミュニティの生成から消滅までを追うことができる。

また、同定されたペア以外で  $\sigma > 0$  のペアは異なるコミュニティ間でメンバーが交換されたとみなすことができる。これらすべての組み合わせを考慮することで、コミュニティ自身をメタノード、コミュニティ間での所属ノードの移動をメタエッジとする時間軸方向への有向グラフとしてコミュニティの系統図を表現することができる。

表1と図1にコミュニティ同定の例を示す。

表1: コミュニティの同定 (図1参照)

$\sigma(i, j)$	$t$	$t+1$	
0.75	<b>B</b>	<b>D</b>	同一コミュニティ
0.5	<b>A</b>	<b>C</b>	同一コミュニティ
0.4	<b>B</b>	<b>E</b>	メンバー交換
0.17	<b>B</b>	<b>C</b>	メンバー交換
0	<b>A</b>	<b>D</b>	関連なし
0	<b>A</b>	<b>E</b>	関連なし

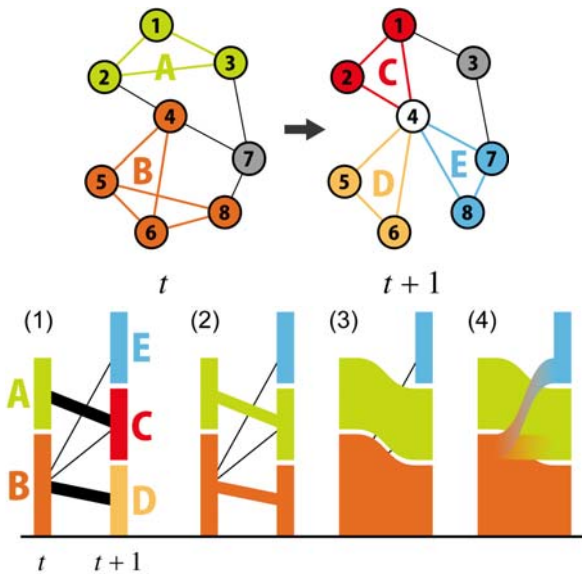


図1：コミュニティの同定と可視化

## 2.4. コミュニティ系統図の描画

コミュニティ系統図の描画は以下の手順で行う。

1. 横軸を時刻とし、過去から順に未来方向へ描画する
2. 各時刻のコミュニティを矩形で描画する。矩形の高さはコミュニティに所属するノード数に比例し、原則ノード数の多いコミュニティから順に積み上げるように描く。
3. 時刻  $t$  を描画後、 $t+1$  を描画する場合には  $t$  に存在するコミュニティを優先的に下方に描く

4. 時刻の前後で同一であるとみなしたコミュニティは同色とし、両者をスプライン曲線で接続する
  5. 前後の時刻に同一コミュニティが存在しない場合、曲線は閉じる
  6. 他のコミュニティからの合流あるいは分岐を同様に描く。合流・分岐する曲線の幅はコミュニティ間を移動するノード数に比例する
- 以上の手順によってコミュニティの変化の視覚的連続性を保ちながら、その遍歴を表示する。

## 3. 可視化結果

図2に電子掲示板のコメントング関係を可視化した例を示す。図中、下図がコミュニティの系統図、左図が系統図中央の時刻における構造のスナップショットを示す。時刻を切り替えることでスナップショットは逐次更新される。これら2つのビューにより、時刻をインタラクティブに前後しながら系統図により注目すべき時刻を特定し、詳細な構造から注目すべきコミュニティを特定し、さらに注目すべき時刻を新たに発見するという再帰的なシナリオ分析が容易に可能となる。

また、系統図中で暗色で示した部分は、ユーザーがスナップショットビューで選択したノードがコミュニティに関わった期間を示している。これによって個人がネットワークのライフタイムを通してコミュニティ空間でどのようなパスを描いたかが一目瞭然であり、より詳細に検討すれば、そこからコミュ

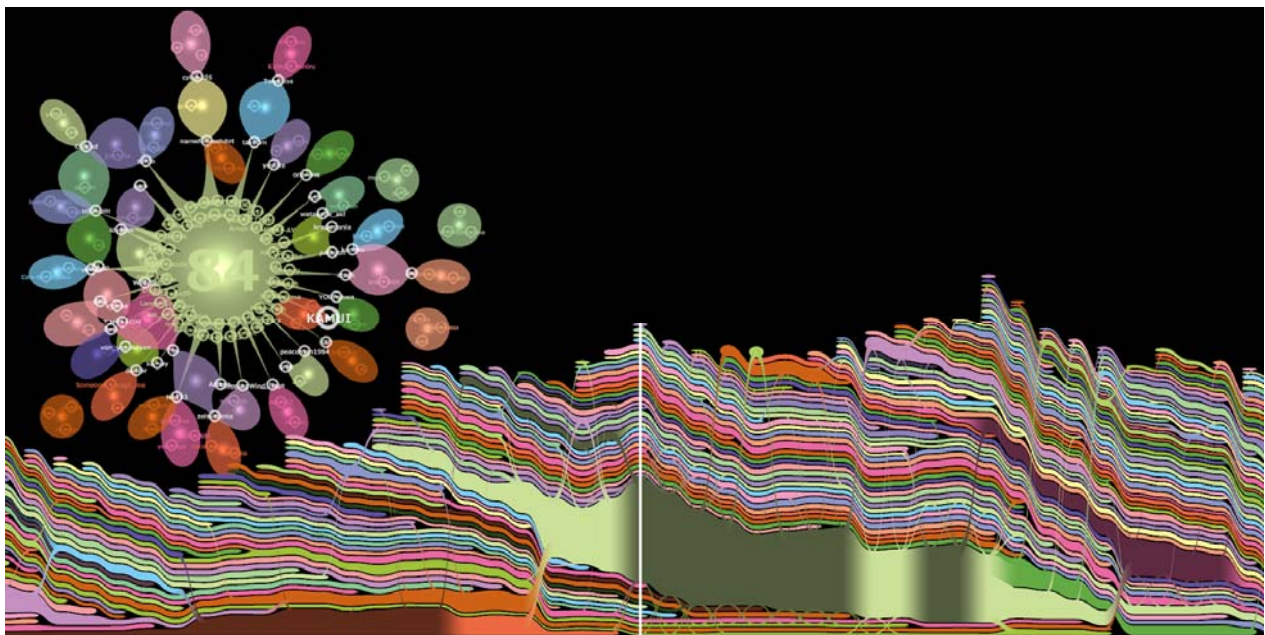


図2：電子掲示板のコメントング関係を可視化した例

ニティへのコミットという観点から個人の行動パターンを抽象化できる可能性も見える。具体的に言えば、例えばコミュニティが生まれた初期に頻繁に出現する傾向を持つノード（コミュニティを生み出す性格を持った人）、あるいはコミュニティが消滅あるいは分裂する時期に頻繁に出現するノード（コミュニティを解体する性格を持った人）といった分類ができるかもしれない。個人の属性を分類できれば、次はコミュニティを構成するメンバーの属性分布を調べることで、そのコミュニティの将来の発展を予測できる可能性も出てくるだろう。

#### 4. 結論

コミュニケーションのイベントログからソーシャルネットワークの時系列を構築し、そこからコミュニティの変化をコミュニティ系統図として可視化する手法を提案した。

#### 謝辞

本研究は科研費 19700085 の助成、および東京大学人工物工学研究センター価値創成イニシアティブ（住友商事）寄付研究部門の支援を受けたものである。

#### 参考文献

- [1] M. Granovetter, The strength of weak ties: a networks theory revisited, *Sociological Theory*, 1, 201-233(1983).
- [2] D. J. Watts, et al., Collective dynamics of 'small-world' networks, *Nature*, 393, 440-442 (1998).
- [3] A.-L. Barabási, et al., Emergence of scaling in random networks, *Science*, 286, 509-512 (1999).
- [4] A.-L. Barabasi, et al., Evolution of the social network of scientific collaborations, *Physica A*, 311, 590-614 (2002).
- [5] M. E. J. Newman, et al., Why social networks are different from other types of networks, *Phys. Rev. E*, 68, 036122 (2003).
- [6] G. Kossinets, et al., Empirical analysis of an evolving social network, *Science*, 311, 88-90 (2006).
- [7] M. E. J. Newman, Fast algorithm for detecting community structure in networks, *Phys. Rev. E*, 69, 066133 (2004).
- [8] L. C. Freeman, Visualizing Social Networks, *J. Soc. Struct.*, 1, 1 (2000).
- [9] G. Palla, et al., Quantifying social group evolution, *Nature*, 446, 664--667 (2007).