

論文の内容の要旨

論文題目 **Friendster** におけるユーザーネットワーク崩壊のメカニズム

氏 名 関 和 紀

Friendster は、SNS（ソーシャル・ネットワーキング・サービス）の黎明期を代表する Web サービスの一つである。2003 年にサービスが開始されると、開始から 7 か月で登録ユーザー数は 200 万人に達した。後発の **MySpace** や **Facebook** 等の台頭の中でも成長を続け、最終的に登録ユーザー数は一億人を超えた。しかし、2009 年頃にアクティブユーザー数が減少に転じ、2011 年にサービスが終了した。

近年は、多種多様な SNS が人々の生活に浸透し、社会に大きな影響を与えている。それらの SNS が **Friendster** と同じ失敗を繰り返し、人々や社会に悪影響を及ぼすことを防ぐためにも、**Friendster** の衰退過程に関する理解を深めることは大きな意義を持つと考えられる。

Friendster の衰退過程に関して、**Friendster** 上で形成されたユーザーネットワークの構造に着目して分析した先行研究として、Garcia (2013) と Yu (2016) がある。Garcia は、「各ユーザーは **Friendster** 上に留まり続けるコストと、**Friendster** 上で得られる利益との大小関係によって、**Friendster** に留まり続けるか離脱するかを決定する」と仮定した。利益はその時点でアクティブである隣接ユーザー数に比例する値（比例定数 b ）、コストは全ユーザーに共通の定数 c であると仮定した。このとき、ネットワークにおける中心性を示す指

標であるコア値 k_c がある閾値 (b/c) 以上であるユーザーのみが **Friendster** に留まり続ける。**Garcia** は、この閾値が **Friendster** の衰退開始以降一定のペースで増加したと仮定することで、衰退過程におけるアクティブユーザー数の時間変化をよく説明できると主張した。**Yu** は、**Garcia** が置いた「閾値が一定のペースで増加した」という仮定の根拠が乏しいことを問題視し、「ユーザーの意思決定はアクティブである隣接ユーザーの数と割合に影響される」と仮定することで、閾値増加の仮定を加えることなくアクティブユーザー数の時間変化を説明できると主張した。

アクティブユーザー数のみに着目すれば、**Garcia**, **Yu** のモデルはいずれも **Friendster** の衰退過程を良く説明することができる。しかし、**Friendster** の衰退過程におけるアクティブネットワーク（アクティブユーザーからなる部分ネットワーク）の時間変化は、十分に検証されていない。

そこで、本論文の第4章では、**Friendster** の衰退過程においてアクティブネットワークがどのように縮小したのかを分析した。

分析対象としたデータセットは **Internet Archive** が 2011 年 6 月に収集したものであり、**Friendster** のサービス終了直前のある一時刻における、ユーザーネットワーク全体の構造を示している。このデータセットでは、あるユーザーA が他のユーザーB を自らのフレンドリストに登録しているとき、これをユーザーA からユーザーB へ向かう有向辺で表現している。データセットには、各ユーザーが会員登録した日時や、辺が生成された日時に関する情報は含まれていない。そのため、このデータセットからネットワーク構造の時間変化を正確に再現することは不可能である。しかし、各ユーザーはデータセット中で連続的な整数値である ID によって表されているため、ここから各ユーザーが会員登録した順序を知ることが可能である。

そこで、ユーザーの ID を基に、各時刻におけるアクティブネットワークを推定する手法を考案した。各ユーザーの ID と時刻とを同一視し、時刻 T におけるアクティブユーザーとは、1) ID が $T + \delta$ 以下であり、かつ 2) 「ID が T 以上 $T + \delta$ 未満である一人以上のユーザー」への辺を持つユーザー、と定義した (δ はパラメータ)。また、アクティブユーザー同士を結ぶ辺をアクティブな辺と定義し、ユーザーネットワーク全体のうちアクティブユーザーとアクティブな辺のみから構成される部分ネットワークを、時刻 T におけるアクティブネットワークと定義した。以上の手法によって推定した各時刻のアクティブネットワークのコア構造を、**k-core decomposition** 法を有向ネットワーク用に定義し直した有向 **k-core decomposition** 法で分析した。

以上の分析から得られた各時刻のアクティブネットワークのコア構造を見比べることで、コア構造がどのように時間変化したのかを確かめた。その結果、**Friendster** のユーザーネットワークにおいては「アクティブユーザー数が減少に転じる前に、アクティブネットワークにおいて高コア値を持つユーザーの数が減少に転じた」という傾向がみられること

が明らかとなった。いくつかの指標による定量的な分析からも同様の傾向が確かめられた。さらに、ID を時刻とみなしたことによる影響が小さいこと、幅広いパラメータ領域で同様の傾向がみられること、アクティブユーザーやアクティブネットワークやコア構造の定義にいくつかの変更を加えた場合でも同様の傾向がみられることが確認された。以上から、Friendster においては「ユーザーネットワークがコア構造の中心部から空洞化するように縮小した」ことが示唆された。

第 5 章では、Friendster のユーザーネットワークがコア構造の中心部から縮小したメカニズムを説明するためのモデルとして、irSIR モデルを基にした影響伝播モデルを使うことが可能であるかを検証した。

irSIR モデルは、1) 集団内の人々は S (まだその SNS に会員登録していない)、I (現在アクティブに活動している)、R (すでにその SNS を離脱した) のいずれかの状態をとり、2) 状態 S の人は状態 I の人との接触頻度に比例する確率 (比例係数 β) で状態 I に、状態 I の人は状態 R の人との接触頻度に比例する確率 (比例係数 γ) で状態 R に遷移する、と仮定する。irSIR モデルは、MySpace という SNS の成長と衰退の過程を説明する上で効果的であることが知られている (Cannarella, 2014)。

ただし、irSIR モデルはネットワーク構造を考慮せず、全ての人が他の全ての人と等しい頻度で接触するような状況を仮定している。一方、本論文ではネットワーク構造に着目しているため、辺で隣接している 2 頂点間でのみ影響伝播が発生しうると仮定するネットワーク irSIR モデルへと拡張した。

本章では第一に、Friendster におけるユーザーネットワークの時間変化がネットワーク irSIR モデルの遷移確率と同様の式を満たすかを検証した。具体的には、各ユーザーの ID を用いて Friendster における各ユーザーの状態 S、I、R を定義した上で、「状態 S の各ユーザーが時刻 A において所持していた状態 I (R) の隣接ユーザー数」と、「次の時刻 B までの間にそのユーザー自身が状態 S から I (状態 I から R) に遷移する確率」との関係調べた。その結果、状態遷移確率は幅広い期間 (時刻 3×10^7 から 8×10^7) においてネットワーク irSIR モデルと同様の式を満たすことが確かめられた。

次に、ネットワーク irSIR モデルにしたがって時間変化するネットワークでは、コア構造の中心部が全体よりも先に縮小するのかを検証した。具体的には、実ネットワークと近い性質をもつネットワークを生成し、その上でネットワーク irSIR モデルのシミュレーションを行い、アクティブネットワーク (状態 I の頂点からなる部分ネットワーク) のコア構造の時間変化を確かめた。その結果、幅広いパラメータ領域において、コア構造中心部の縮小はアクティブユーザー数の減少に先立って発生することが確かめられた。

以上から、Friendster のネットワーク構造の時間変化について分析するモデルとして、ネットワーク irSIR モデルを使える可能性があることが示唆された。

第 6 章では、ネットワーク **irSIR** モデルの遷移規則に変更を加えたときに、アクティブネットワークの縮小のしかたがどのように変化するかを、シミュレーションによって分析した。

状態 **S** から状態 **I** への遷移確率に変更を加えたシミュレーションからは、ここで加えた変更が前章の結論に影響を及ぼさないことが確かめられた。

状態 **I** から状態 **R** への遷移確率に次数に反比例する項を加えたシミュレーションからは、ネットワーク **irSIR** モデルに「次数の小さな頂点は、短期間のうちに状態 **I** から状態 **R** へ遷移する確率が高い」という規則を加えることで、コア中心部の急激な縮小が抑えられることが確かめられた。

影響伝播による状態 **R** から状態 **I** への遷移規則を加えた「ネットワーク **irSIRI** モデル」のシミュレーションからは、状態 **R** から状態 **I** への遷移確率を決定するパラメータ γ の値を十分に大きくすれば、状態 **I** の頂点数は正のある値に収束することが確かめられた。

ネットワーク **irSIRI** モデルの遷移規則にさらに「状態 **R** の頂点を直接選択し、状態 **I** に遷移させる」という規則を加えた「直接遷移ネットワーク **irSIRI** モデル」のシミュレーションからは、次のことが明らかになった。1) 単位時間あたりに状態 **R** から状態 **I** に遷移させる頂点数を多くするほど、平衡状態における状態 **I** の頂点数は多くなる。2) 状態 **R** から状態 **I** に遷移させる頂点をランダムに選んでも、次数に比例する確率で選んでも、平衡状態における状態 **I** の頂点数はほとんど変わらない。

状態 **R** から状態 **I** への遷移が連鎖反応を起こすと仮定する「復帰の連鎖反応モデル」のシミュレーション結果からは、次のことが明らかになった。1) 状態 **R** から状態 **I** への遷移が連鎖反応を起こすと仮定するとき、状態 **I** へ遷移させる頂点を次数に比例する確率で選択するほうが、ランダムに選択する場合よりも、平衡状態における状態 **I** の頂点数は大きくなる。2) 上のような現象がみられるのは、連鎖反応の起こりやすさを司るパラメータ b_3 の大きさが $10^{-2} \leq b_3 \leq 10^{-1.5}$ であるときに限られる。

以上のモデルのシミュレーション結果を、**SNS** の用語で解釈すると以下ようになる。

- 1) 離脱したユーザーの復帰がある閾値以上の確率で発生するとき、アクティブユーザー数は正のある値に収束する。
- 2) 復帰の連鎖反応がなければ、**SNS** の管理者などが直接復帰を働きかけるとき、働きかけるユーザーをランダムに選択する場合と、次数の高いユーザーを優先的に選択する場合との間で、平衡状態におけるアクティブユーザー数には差がない。一方、あるユーザーが **SNS** に復帰することによって復帰の連鎖反応が引き起こされると仮定すると、次数の高いユーザーに優先的に働きかけることは効果的である。

ただし、以上はその **SNS** が本章のモデルに沿っている場合にのみ言えることであることには留意する必要がある。