

令和元年6月5日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00418

研究課題名(和文) オンライン写真共有サービスにおけるソーシャルタギングダイナミクスの解明

研究課題名(英文) Study on Social Tagging Dynamics in Online Photo-Sharing Services

研究代表者

橋本 康弘 (Hashimoto, Yasuhiro)

東京大学・大学院総合文化研究科・特任研究員

研究者番号：10376494

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：写真共有サービスなどで一般的となったソーシャルタギングシステムを、タグと人間行動が共進化的に発展していく協同的な現象と捉え、タグの生成と選択のダイナミクス、その背後にあるユーザの選好的な構造に関して、実験と解析の両面から実証的な分析を行った。タグとその組み合わせの選択のメカニズムは、多く使われたものほどさらに使われやすくなるモデルがよい説明を与える一方で、短い時間スケールでは非一様な選択パターンが存在することを明らかにした。また、多数のタグについて、それと隣接する共起構造の変化を可視化した結果、タグとその組み合わせのオープンエンドな進化を示唆する特異的なパターンを見出した。

研究成果の学術的意義や社会的意義

ソーシャルタギングシステムは、新しいタグの発明とタグへの投票行動によって、その意味の構造を発展させながら、それによって次の人々の行動を方向付けていくシステムである。そのメカニズムを理解することは、新しい概念やアイデアを産み出す原動力とその社会的受容、つまりイノベーションを理解する上での示唆を与えると期待できる。本課題ではタグの生成と選択の基本となる数理モデルを発展させ、背後にある人間行動との強い相関の存在を明らかにした。同時に、タグという文化遺伝子のオープンエンドな進化の端緒を可視化することに成功した。

研究成果の概要(英文)：Social tagging dynamics is regarded as a cooperative recursive phenomenon in which the vocabulary system and human behavior coevolutionarily develop in time. The empirical analysis on the dynamics of generation and selection of tags and the structure of preference of users was carried out from both sides of theoretical and numerical analysis. We confirmed that the preferential attachment model provides a good explanation to the selection mechanism of tag and their combination on average, while revealing temporal inhomogeneity such as "bursts" on short time scales. Moreover, as a result of visualizing the change of the adjacent co-occurrence structure for a large number of tags, a specific pattern, which indicates the sign of the open-ended evolution of tags and their combinations, was discovered.

研究分野：計算社会科学

キーワード：ソーシャルタギング ソーシャルネットワーク Yule-Simon過程 優先的選択 進化 新規性 コミュニティ ベキ則

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

ウェブ上で写真や動画を他者と共有するサービスは現代人にとってポピュラーなコミュニケーション手段の一つとなっている。ウェブサービスのアーキテクチャが発達する中で、2000年代以降、“ソーシャルタギングシステム”と呼ばれるしくみがそれらコンテンツ共有サービスを構成する機能の一つとして広く用いられるようになってきた。これはサービスを利用するユーザが、投稿されたコンテンツに対して、“タグ”あるいは“ハッシュタグ”と呼ばれる短い単語やフレーズをアノテーションとして加えることができる機能である。用いられる自由なタグは、サービスが規定する一定の枠の中で、投稿者の選好や、投稿したコンテンツに対する創造的解釈を含むものとなる。それは社会的なトレンドやコンテクストを反映する場合も多い。このように、自由なタグ付けをユーザに許したことで、そこに新たな利用方法が生まれた。ユーザはタグを分類学的な情報ナビゲーションのためのみならず、コンテンツに付随する独創的・創造的な表現として、他者との間接的なコミュニケーションを図る手段としても用いるようになったのである。

ソーシャルタギングとは、日常の様々な事物を短いワードで表現した、いわば民主的な“コード”であり、それは新しい語彙の生成と、既存の語彙への投票行動によって発展していく進化的システムである。これを敷衍すれば、タグの語彙のサイズや出現数、組み合わせの変化は、新奇な概念や解釈が生み出されてから、集団内での選択、淘汰、社会的受容を経て、新しい習慣や文化を形成していく過程とみることができる。同時に、こうして作り上げられた世界観が再帰的に人間自身の行動をも変化させていく。具体例を挙げれば、我々は写真共有ソーシャルネットワークワーキングサービス「RoomClip」において、ある種の写真のテイストを表現する単語が生み出され、その新しい表現が一定の意味を獲得し集団内に定着することによって、そのテイストに追従する写真が投稿される現象を観測している。このような視点でソーシャルタギングという行為を眺める時、そこには現代の情報サービスが媒介する、人と社会の創造的な共進化を見出すことができるのである。

2. 研究の目的

本研究が解明しようとする現象は、協同的再帰現象としての、語彙システムと人間行動の共進化的な時間発展である。そのために、ここではとくにオンライン写真共有サービスの一つである Flickr と RoomClip にフォーカスし、ソーシャルタギングシステムにおけるタグの生成と選択のダイナミクス、その背後にあるユーザの選好的な構造を実証的に明らかにしていくことを目的とする。

ソーシャルタギングシステムを扱った研究は 2000 年代半ば以降盛んに行われてきたが、その動機は、サービス運営の効率化やユーザ体験の新規開拓を目的とした、タグ推薦アルゴリズムの開発、タグのセマンティックな構造の再構築、可視化やユーザインターフェイスの改善、あるいは純粋に言語学的な見地から語彙の時間発展が示す統計則を調べたものなど、多種多様である。これに対して本研究では、タグの進化的なダイナミクスを、その背後にあるユーザ行動と関連付けながら分析していく。

ソーシャルタギングシステムは、新しいタグの発明と、ある種の投票行動としての既存のタグの使用によって、セマンティックな構造を発展させ、それによって次の行動を方向付けていくシステムである。この再帰的な枠組みの中では、人々が日常を生きる世界観が外部化されながら、行動と行動を導く環境そのものが更新されていく。この進化的ダイナミクスを明らかにすることは、新しい概念やアイデアを生み出す原動力とその社会的受容、すなわちイノベーションの創発といった生成的価値をもたらす原理について、大きな示唆を与えると我々は期待する。

3. 研究の方法

- ① 本課題に先行する研究の中で、我々は個々のタグの出現数が従来の典型的な確率過程が予測する成長から大きく逸脱することを確かめており、ソーシャルタギングシステムにはユーザ選好に基づく強い非線形効果が存在することを予見していた。したがって、まずタグの生成と選択のプロセスをシンプルな確率過程としてモデリングし、これを基礎に各種統計則の解析的・実験的導出を行うところから研究を開始した。そして、その振る舞いを実データと照らし合わせ、妥当性と限界を検証した。
- ② 基礎となる時間発展モデルでは新規な語彙の生成レートを一様としたが、新規性の生成はユーザの属性やコミュニティの構造と関連している可能性がある。ここでは個々の投稿者が用いた語彙の確率分布によって投稿者同士の類似度を定義し、投稿者をノードとする類似ネットワークを構築した。そこからクラスタ分析によって、語彙プロファイルの近い集団と新しいタグの生成レートの関係について分析を行った。

- ③ 一般に一つの写真には複数のタグがセットで与えられ、そこではタグの共起が重要な意味を持つ。タグを生物遺伝子になぞらえタグのセットを一つの表現型と捉えたとき、組み合わせとして意味を持つタグのセットが存在し、その組み合わせの変化にも特徴的なパターンが存在すると考えられる。このタグの組み合わせについても、タグ単体と同様、生成のパターン、および選択のパターンに関する分析を行った。
- ④ タグの共起グラフ構造の時間発展に着目し、タグのセマンティクスの変化について分析を行った。ここでのタグのセマンティクスとは、着目するタグに隣接する他のタグの共起サブグラフを指し、この隣接グラフ構造がタグの意味を規定していると考えられる。この隣接グラフ構造の変化を定量的に追うことで、タグの「意味」が変化していく過程を可視化し、語彙のオープンエンドな進化の可能性について検証した。

4. 研究成果

- ① タグの選択確率に見られる多様性が、優先的選択性を持つ確率過程の一つである Yule-Simon 過程からの逸脱として定量化できることを明らかにした。例えば、逸脱を定量化するスケーリング係数の確率密度分布の裾野が実データでは広がる要因について、優先的選択とは異なる固定レートの成長メカニズム—つまり、競合となる他のタグに依存しない、時間的に集中したタグの選択行為がタグの成長の初期段階で存在することを明らかにした。また、その固定レートの成長は長時間スケールでは消滅し、最終的には累積出現数が大きなタグは優先的選択の平均場解の周辺に集まることを確かめた。これは、最終的に大きな累積出現数を得るタグは成長が速いタグであるという直感に反する結果である。しかし、ソーシャルタギングにおける出現数の確率分布が Zipf 則（つまり優先的選択の帰結）に従うというマクロな結果との整合性は取れており、ソーシャルタギングを特徴付ける、ミクロな構造の一つであると結論付けられる。（図 1）

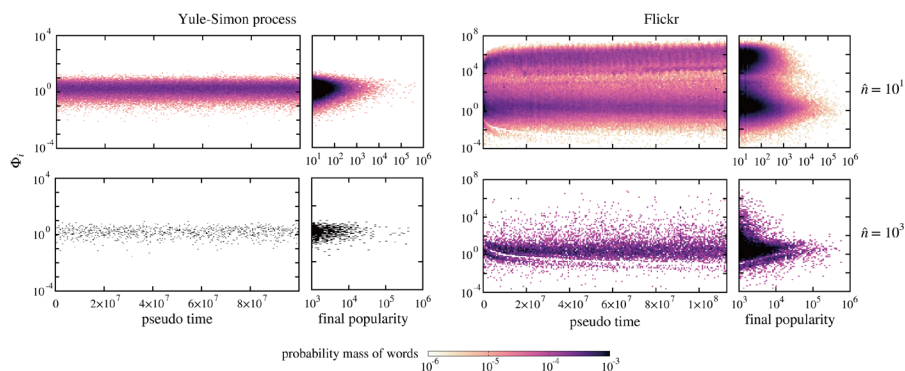


図 1 タグの確率質量分布。横軸は時刻（左パネル）と最終的な累積出現数（右パネル）。縦軸は成長速度の平均場解からのズレ。実データ（左）には Yule-Simon 過程には見られない、成長の早いタグの集団が存在する。短い時間スケール（上）では顕著だが、長い時間スケール（下）ではその傾向は消えていく。

- ② 投稿者がタグに用いた語彙の確率分布の類似度を Jensen-Shannon Divergence を用いて比較し、得られた類似度グラフについてクラスタリングを行った結果、語彙の類似性の高い投稿者の集団ほど高いモジュール性（何らかのトピック構造）を持ち、グラフ全体は、モジュール性を持ったコア構造と、緩い接続関係を持ったペリフェリ構造として表現できることが明らかになった。このとき、新しいタグを生成する投稿者は、コア部分に比べてペリフェリ部分に多く存在することが確かめられた。（図 2）

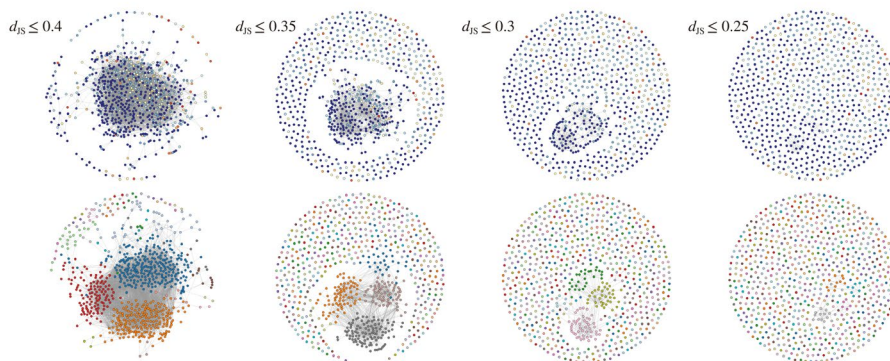


図 2 投稿者のタグのプロファイルによる類似度グラフと新規タグ生成レート。ノードは個々のユーザ、ノードの色は新規タグの生成レート（上）とクラスタリング結果（下）。右へ行くほどリンクの閾値を高くしたもので、つまり類似度の強いコアの部分抽出している。グラフモジュール性も高い。

③ Yule-Simon 過程の拡張として、タグのペアの選択をモデル化した結果、実データにおいても全体的には優先的選択性が働いていることが示された。この数値計算のために、タグのペアに関する“アタッチメントカーネル”の計算アルゴリズムを工夫し、高い精度でこれを確かめた(図3)。一方で、組み合わせに用いられるタグは一樣な確率で選択されるのではなく、個々のタグが作られた時刻によるバイアスが働いていることを明らかにした。これは①で明らかにしたタグのバースト的な使用とも関連があり、さらなる検証が望まれる。

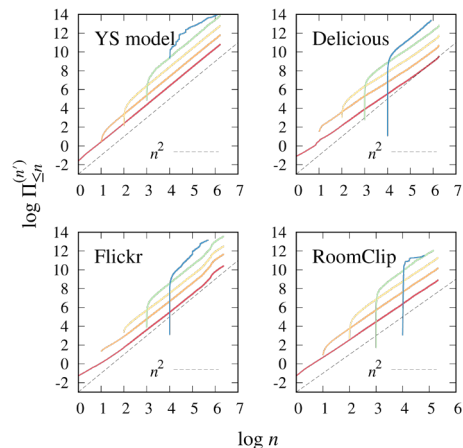


図3 タグのペアに関するアタッチメントカーネル。赤から順に出現数 $n=1, 10, 10^2, 10^3, 10^4$ と n =横軸のタグペアの選択確率。

④ タグの隣接グラフ構造をグラフ要素の重みに関する確率分布とみなし、その時刻間でのグラフ構造の距離を Jensen-Shannon Divergence で定量化することで、タグの隣接グラフ構造の時間変化を可視化することに成功した(図4)。その結果、多くのタグはサービス開始後、グラフ構造の変化が緩やかになる=意味の収束を見せるものの、一部のタグでは構造が複数の時間的クラスタとして現れる「セマンティクスの変遷」を示した。これはタグの内容によっては季節性を反映した周期構造と捉えることもできるが、それでは説明がつかないドリフト的な変遷を示すタグも観察され、ソーシャルタギング固有のオープンエンドな意味の進化を示唆する結果を得た。

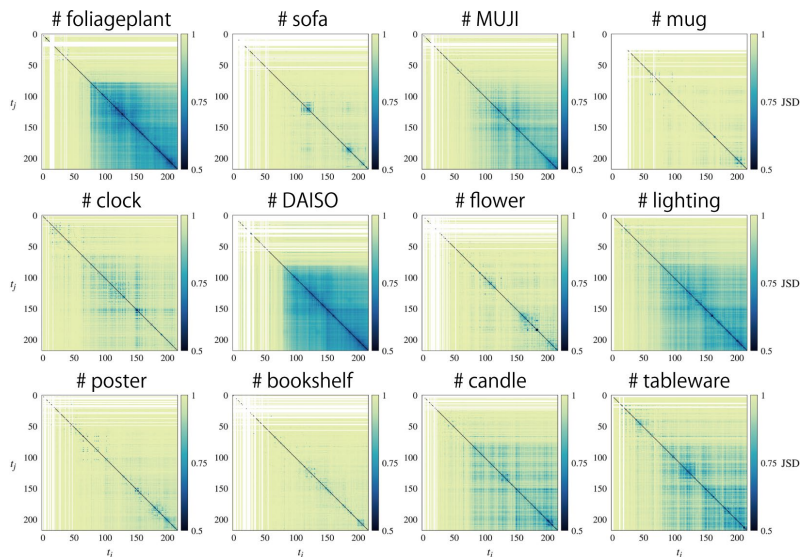


図4 タグのセマンティクスの変化 (RoomClip)。縦横軸はサービス開始 (左上) から4年経過後 (右下) までの週のインデックス、セルの色は2つの週のグラフ構造の類似度 (薄いほど小)。濃いクラスタは時間的な意味の安定を示している。

5. 主な発表論文等

[雑誌論文] (計4件)

- 1) Takashi Ikegami, Yasuhiro Hashimoto, Mizuki Oka: Open-Ended Evolution and a Mechanism of Novelties in Web Services, *Artificial Life*, (2019) 印刷中
- 2) 西川仁将, 岡瑞起, 橋本康弘, 池上高志: Yule-Simon 過程に基づくソーシャルタギングの分析とモデリング手法, *人工知能学会論文誌*, 34 (4) (2019) 印刷中
- 3) Takashi Ikegami, Yoichi Mototake, Shintaro Kobori, Mizuki Oka, Yasuhiro Hashimoto: Life as an Emergent Phenomenon: Studies from a Large-Scale Boid Simulation and Web Data, *Philosophical Transactions A*, 375, 20160351 (2017)
- 4) Yasuhiro Hashimoto: Growth Fluctuation in Preferential Attachment Dynamics, *Physical*

[学会発表] (計 20 件)

(招待講演)

- 1) 橋本康弘: 新規性の発生則と成長則: Yule-Simon 過程を基礎にして, 電気情報通信学会 複雑コミュニケーションサイエンス研究会, 信学技報, 117, 520, CCS2017-41, pp. 49-54, 東京 (2018/3/27)
- 2) 橋本康弘: 新規性の発生則と成長則: Yule-Simon 過程を用いた分析, 計測自動制御学会 第 23 回創発シンポジウム「創発夏の学校」(チュートリアル講演・ワークショップ), 茅野 (2017/9/9)

(査読付き国際会議)

- 3) Koya Sato, Mizuki Oka, Yasuhiro Hashimoto, Takashi Ikegami, Kazuhiko Kato: How the Nature of Web Services Drives Vocabulary Creation in Social Tagging, *The 2nd International Conference on Information Science and System (ICISS 2019)*, Tokyo, Japan, Mar 16-19 (2019)
- 4) Yasuhiro Hashimoto, Takashi Ikegami: Novelty Production in Tagging Crowds, *SWARM2017*, pp. 311-312, Kyoto, Japan, Oct 29-Nov 1 (2017)
- 5) Yasuhiro Hashimoto, Mizuki Oka, Takashi Ikegami: Large Fluctuation in Social Tagging Ecosystems, *International Conference on Computational Social Science (IC2S2) 2017*, no-proceedings, Cologne, Germany, Jul 10-13 (2017)

(その他国際会議)

- 6) Yasuhiro Hashimoto, Takashi Ikegami: Growth Dynamics of Pairwise Novelty in Social Annotation, *Workshop on Social Simulations and Supercomputers, The 21th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA2018)*, Tokyo, Oct 29-Nov 2 (2018)
- 7) Norihiro Maruyama, Daichi Saito, Yasuhiro Hashimoto, Takashi Ikegami: Dynamic Organization of Flocking Behaviors in a Large Scale Boids Model, *Workshop on Social Simulations and Supercomputers, The 21th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA2018)*, Tokyo, Oct 29-Nov 2, (2018)
- 8) Yasuhiro Hashimoto, Mizuki Oka, Takashi Ikegami: Measuring Bursty Growth of Words in Social Tagging Ecosystems, *10th International ACM Web Science Conference 2018 (WebSci 2018)*, Amsterdam, The Netherlands, May 27-30 (2018)
- 9) Yasuhiro Hashimoto, Mizuki Oka, Takashi Ikegami: Anomalous Popularity Growth in Social Tagging Dynamics, *CompleNet 2018*, Boston, USA, Mar 5-8 (2018)
- 10) Takashi Ikegami, Mizuki Oka and Yasuhiro Hashimoto: How Novelty is Created in a Web Service, *OEE3: The Third Workshop on Open-Ended Evolution, The 2018 Conference on Artificial Life (ALIFE2018)*, Tokyo, Jul 23-27 (2018)
- 11) Mizuki Oka, Koya Sato, Yasuhiro Hashimoto, Takashi Ikegami: Emergence of Individuality on Social Tagging Dynamics, *Complex Networks: From Theory to Interdisciplinary Applications*, no-proceedings, Marseille, France, Jul 11-13 (2016)

(その他国内会議)

- 12) 橋本康弘, 岡瑞起, 池上 高志: ソーシャルタギングシステムに見られる特異的なタグの成長, 人工知能学会第 32 回全国大会, 鹿児島 (2018/6)
- 13) 西川仁將, 岡瑞起, 橋本康弘, 池上高志: ソーシャルタギングにおけるタグ出現数のゆらぎ, 人工知能学会第 32 回全国大会, 鹿児島 (2018/6)
- 14) 橋本康弘, 佐藤晃矢, 岡瑞起, 池上高志: Yule-Simon モデルをベースにしたソーシャルタギングの進化計算, 日本物理学会第 72 回年次大会, 大阪 (2017/3)
- 15) 岡瑞起, 西川仁將, 佐藤晃矢, 橋本康弘, 池上高志: SNS にみる新規語彙の生成・選択・揺らぎの現象, 情報処理学会ヒューマンコンピュータインタラクション, 沖縄, (2017/1)
- 16) 橋本康弘, 佐藤晃矢, 岡瑞起, 池上高志: ソーシャルタギングにおける語彙出現のアンマリー, 日本物理学会秋季大会, 金沢 (2016/9)
- 17) 佐藤晃矢, 岡瑞起, 橋本康弘, 池上高志, 加藤和彦: ソーシャルネットワークのダイナミクスにみる, ゆらぎ, 応答, 進化, 日本物理学会秋季大会, 金沢 (2016/9)
- 18) 橋本康弘, 佐藤晃矢, 岡瑞起, 池上高志: ソーシャルタギングにおけるタグ出現数のゆらぎ, 人工知能学会第 30 回全国大会, 北九州 (2016/6)
- 19) 西川仁將, 岡瑞起, 橋本康弘, 池上高志: ソーシャルタギングの形質表現と進化メカニズム, 人工知能学会第 30 回全国大会, 北九州 (2016/6)
- 20) 佐藤晃矢, 岡瑞起, 橋本康弘, 池上高志, 加藤和彦: 新しいタグの出現とソーシャルタギングシステム, 人工知能学会第 30 回全国大会, 北九州 (2016/6)

〔図書〕 (計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

なし

6. 研究組織

(1) 研究分担者 なし

(2) 研究協力者 なし

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。