

# 個人の嗜好と情報カテゴリを考慮した 情報発信モデルによるウェブリンク構造生成

Generation of web link structure based on posting information model  
including personal preference and content categories

熊澤 健人<sup>†</sup>  
Kento Kumazawa

中平 勝子<sup>†</sup>  
Katsuko T. Nakahira

北島 宗雄<sup>†</sup>  
Muneo Kitajima

## 1 はじめに

昨今の情報化社会において、ICTの発達とともに個人によるインターネットを介した情報発信活動は増加の一途をたどっている。情報化社会におけるインターネット上の様々な現象の理解には、現象をマクロに見るだけでなく、ミクロに、すなわち、個の情報送受信活動にも着目することが必要となる。インターネットにおける情報発信活動は、古典的なWebページ作成のほか、SNSをはじめとする娯楽的な情報授受、eコマースサイトにおける推薦など様々な手法がある。情報発信活動は個によって様々な制約や条件の下で行われる。本稿では、それら制約や条件をeネットワーク[1]に基づいて設計されたウェブ空間成長過程モデル[4]に基づいて情報発信活動を考える。eネットワークにおける個人特性としての嗜好や情報カテゴリの選択性を考慮し、ウェブ空間におけるウェブリンク構造を時間発展させることにより、人やウェブコンテンツの持つ特性がウェブリンク構造に与える影響を分析する。

## 2 情報発信における個人特性の役割

本稿では、情報発信活動をウェブリンク構造で捉え、シミュレーションによりその再現を行う。そのため、個人によるリンク生成という行動の再現が核となる。情報発信者が自分のコンテンツから他のコンテンツに対してリンクを張るといって情報発信活動をするとき、被リンクコンテンツが情報発信者にとって発信したい情報であったと考えられる。この情報発信者にとって、閲覧した情報が発信したいと思う情報であるか発信したいと思わない情報であるかは、情報発信者の嗜好に依存する。情報が情報発信者の嗜好や意図に合致した時に、リンクが張られる。そのため、リンク生成という形で営まれる情報発信活動を考えるには、情報発信者の嗜好を考慮する必要がある。

ウェブコンテンツの多くは人によって保守運営されている。人がコンテンツの保守運営を行うにあたり、すべてのコンテンツを平等にメンテナンスできるとは限らない。コンテンツの更新頻度には、コンテンツ自身の活性度や重要度、話題性などが関係している。同時期に管理することができるコンテンツ数は、管理者の能力や性格に依存して上限値が決定される。

## 3 実験におけるパラメータ

本稿では、情報発信者とウェブコンテンツによるマルチエージェントシミュレーションを用いてウェブ空間成長の分析を行

う。本実験では、ウェブ空間成長過程モデルに基づいて「情報発信者個々の嗜好」「コンテンツカテゴリ」「コンテンツの活性」「コンテンツのハブ性」を考慮した情報発信活動の再現を行う。【情報発信者個々の嗜好数・ $N_{p,i}$ 】シミュレーション開始時に全エージェントに対して複数の嗜好パラメータを設定する。個人によって生成されるコンテンツのカテゴリは、この嗜好パラメータの中から確率的に選択される。

【コンテンツカテゴリ・ $N_c$ 】 $N_c = 1, 2, \dots, 20$ とし、コンテンツ $j$ はカテゴリ $N_{c,j}$ に分類される。コンテンツ $C_j - - C_k$ 間のリンク生成は、 $N_{c,j} = N_{c,k}$ であればリンクが張られやすく、 $N_{c,j} \neq N_{c,k}$ であればリンクは張られにくくなる。

【コンテンツの活性度・ $D_u$ 】コンテンツがアップデートされる状態にあるか否かを表すものである。情報発信者によって管理されるコンテンツは、アクティブなものとそうでないものがあると考えられる。アクティブなコンテンツは、内容の編集などが行われる。それに対して、非アクティブなコンテンツは編集などが行われぬ。個人のICTリテラシーや管理能力などによって、同時期に何個のコンテンツをアクティブにできるかは異なる。本稿では簡単のために、個人が同時期に管理できるアクティブなコンテンツを10個とした。

【コンテンツのハブ性・ $D_h$ 】ウェブリンク構造を分析すると、コンテンツにはハブになるものとそうでないものがある。ハブになるコンテンツは、生まれた時からハブコンテンツとして生成される[2]。そこで本稿では、生成されたハブコンテンツを常にアクティブなコンテンツとして個人に管理させることでハブ性の再現を行なった。

## 4 シミュレーション結果

図1に $N_c, N_{p,i}, D_u, D_h$ の組み合わせを変えたシミュレーション結果を示す。ユーザが持つ嗜好カテゴリの数を $N_{p,i}$ 、コンテンツ活性度を $D_u$ 、コンテンツハブ度を $D_h$ 、情報発信者数を $H$ 、シミュレーションのステップ数を $T$ として表現する。コンテンツ活性度とコンテンツハブ度はそれぞれ、考慮している場合は"1"、考慮していない場合は"0"で表記する。複数パラメータが設定される $N_{p,i}$ の数について、今回の実験では試験的に $N(3, 5)$ の正規分布に基づいて決定している。

【コンテンツカテゴリ分類の関係】コンテンツをカテゴリ分けし情報発信者に個々の嗜好パラメータを設定した結果リンク生成の頻度が変わることがわかった。コンテンツをカテゴリ分けした結果、情報発信者がウェブ空間から無作為に選択したコンテンツと自らのコンテンツをリンクで接続する際にコンテンツ内

<sup>†</sup>長岡技術科学大学

容が一致しているか否かでリンク生成の結果に変化が出た。カテゴリ分けを行わない場合は、抽出したコンテンツが接続済みである場合の重複リンクや、コンテンツ自身を参照するリンクでなかった場合接続される。これは、情報の内容を考えずにリンクを張っている状況、もしくは情報発信者が必ず求めた情報を自らのコンテンツと接続することができる状態である。このような場合は、(a)(b)(c) と (d)(e)(f) を比較すると多くのリンクが生成され outdegree 図のプロットが密になる。それに対して、カテゴリ分けを行う場合、ウェブ空間から無作為に選択したコンテンツと情報発信者の嗜好カテゴリが異なった場合接続に失敗する。こうして情報発信者が情報の選り好みをする事で、リンクが間引きされ outdegree 図のプロット点が疎になる。コンテンツをカテゴリ分けの有無によって、outdegree 図には疎密の変化が現れるが、どちらの場合においてもべき的減衰構造は確認することができた。

【コンテンツ活性・ハブ性の関係】本実験では、情報発信者の嗜好に加えてコンテンツの特性について考慮した実験も行なっている。コンテンツ活性を考慮することによって (a)(d) と (b)(e) を比較すると僅かに右に遷移するが概形の大きな変化を確認することはできなかった。しかし、コンテンツのハブ性を考慮した (c) の概形には五月雨的構造 [3] を確認できた。全ての要素を考慮した (f) の場合では  $T = 50$  から  $T = 199$  においてそれぞれハブコンテンツとそれ以外のコンテンツが混成している状態を確認することができる。ハブコンテンツとそれ以外混成しているウェブ空間では、べき的減衰構造にノイズが生じ、直線的ではなく曲線的な概形になった。ハブコンテンツを再現することによって、人の営みから現れるべき的減衰構造以外の構造を再現できた。

### 5 まとめ・今後の課題

本稿では、ウェブリンク生成に着目した情報発信活動を再現し、ウェブリンク構造を時間発展させ分析を行なった。より現実的なウェブリンク生成の再現を行うために、情報発信者個人の嗜好を生成されるコンテンツとウェブリンク生成に反映させた。ウェブリンク生成において、情報発信者の嗜好を反映させると、リンク生成が嗜好とコンテンツカテゴリの不一致などで間引きされウェブリンク構造は緩やかに発展する。outdegree 図にはプロットされる点が疎になるという構造的特徴が現れることがわかった。加えて、コンテンツの特性として活性度とハブ性を導入した。活性度の導入によるウェブリンク構造の大きな変化は現れなかった。しかし、常に更新などが繰り返されるハブ性のあるコンテンツを導入することで五月雨的構造を再現することができた。人の嗜好などのパラメータはべき的減衰構造に影響を与え、コンテンツのハブ性など人以外の特性がべき的減衰構造から逸脱した構造的特徴を表していた。

今後は、人が持つ嗜好以外のパラメータがウェブリンク構造に与える影響と、自動推薦リンクや同一 URL 内の内部リンク構造などを再現によって、人以外が持つ特性がウェブリンク構造に与える影響についてそれぞれ分析を行う。

### 参考文献

[1] 岩田 翔太, 中平 勝子, 北島 宗雄: 通信基盤と人との相互作用による情報生成・流通プロセスの記述とバイラルメ

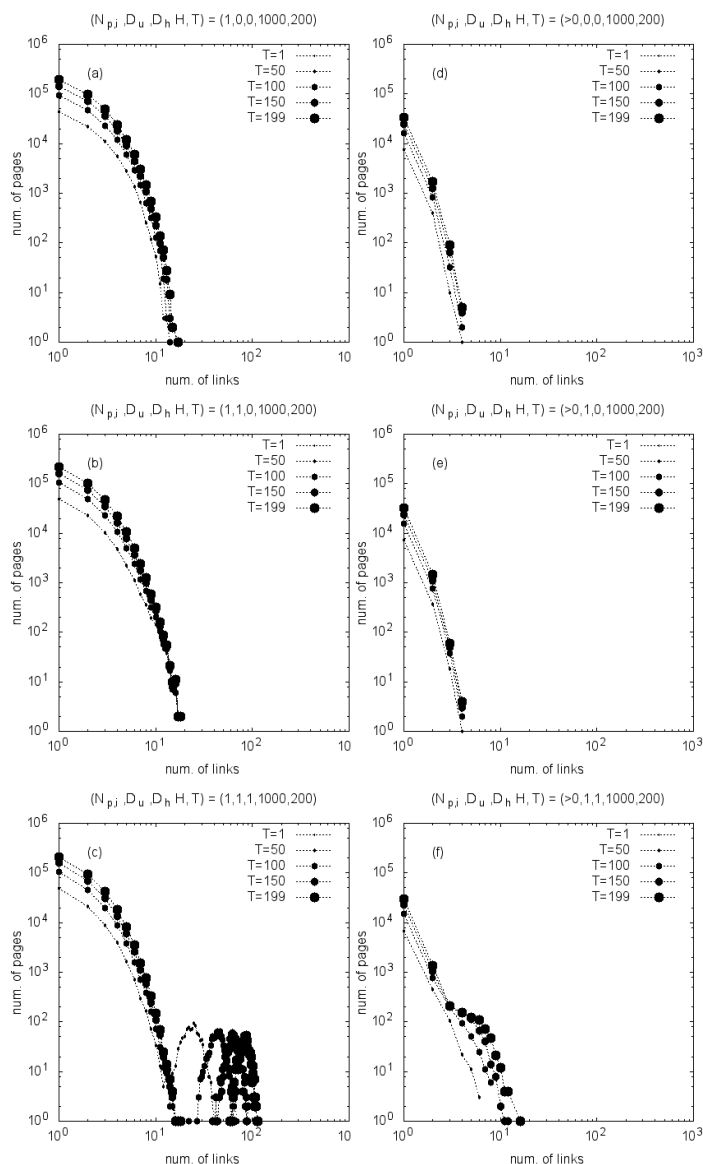


図1 シミュレーション結果

ディアへの適応, FIT2014(第14回情報科学技術フォーラム)pp.337-338, 2014.

[2] A.-L. Barabási, R. Albert, H. Jeong, G. Bianconi : Power-Law Distribution of the World Wide Web, SCIENCE VOL 287 p481 p484 , 2115a 2000 .  
 [3] 熊澤 健人, 中平 勝子, 北島 宗雄: outdegree 図に着目した、ウェブリンク構造変容分析によるデバイド状態の評価, FIT2016(第15回情報科学技術フォーラム) 講演論文集, 第1分冊, pp.161-162.  
 [4] 熊澤 健人, 中平 勝子, 北島 宗雄: ウェブを介した情報発信における提供者と提供物の特性を考慮したページ・リンク生成過程の記述, FIT2017(第16回情報科学技術フォーラム) 講演論文集, 第1分冊, pp.105-108.