

コンテンツ間の双方向リンクを用いたネットワークモデルに関する考察

古賀隆浩[†] 伊藤正詩[‡] 大園忠親[†] 新谷虎松[†][†]名古屋工業大学 大学院工学研究科 情報工学専攻 [‡]名古屋工業大学 知能情報システム学科

e-mail: {koga, ozono, tora}@ics.nitech.ac.jp

1 はじめに

本論文では、Web におけるコンテンツが、自律的かつ双方向にリンクするネットワークモデルを提案し、提案するモデルについてのシミュレーションを行う。筆者らは、ある Web コンテンツからハイパーリンクが作成されたときに、自動的にリンク元へのリンクが作成される環境である、双方向リンクを提案している [2]。双方向リンクの特長は、ハイパーリンクのようなページを単位とするリンクではなく、ページ内の任意の部分をコンテンツとして指定し、リンクすることを可能とすることである。本論文では、双方向リンク環境において、i) 通常の Web 検索では発見が困難なコンテンツとリンク可能、ii) 新しく作成されたコンテンツが多くのコンテンツとリンク可能となるモデルの構築を目的とする。アプローチとして、コンテンツの情報を保持するエージェントが、コンテンツの最新性、ネットワーク上におけるコンテンツ間の距離などを考慮し、リンクするコンテンツを決定する。

2 ネットワークモデル

本モデルでは、双方向リンクの対象となるコンテンツを決定することを、マッチングとよぶ。コンテンツが作成されると、コンテンツの情報を保持するエージェントが生成される。生成されたエージェントは、共有空間に集まり、互いにコンテンツの情報を評価し、マッチングする候補を選択する。エージェント間に合意が成立した場合、マッチングが成功する。

本モデルでは、Web サイトをネットワークのノードで表す。また、Web サイト間のハイパーリンクを、あるノードから他のノードへの有向のリンクで表す。コンテンツは、各ノードから複数生成されるものとする。マッチングが成功すると、コンテンツの生成元であるノード間を相互に結ぶハイパーリンクが作成される。一方、マッチングに失敗したエージェントは、より遠くのコンテンツとマッチングするため、他のノードへ移動する。各エージェントは、生成された時に生成元のノード ID を内部状態として保持する。移動の際には、現在のノードからリンクを辿って到達可能なノ

ードを選択し、内部状態を変化させる。ノード間の距離計測、および閉路での循環を避けるため、エージェントは通過したノードの履歴を保持し、一度通ったノードには移動しないものとする。移動可能なノードが存在しない場合は、1 ステップ前のノードに戻る。マッチングに用いるパラメータを以下に示す。

R:コンテンツの最新性

コンテンツの最新性は、シミュレーションのステップ数と、コンテンツが生成されたステップ数から、次の式で表される。

$$R = \frac{\text{生成されたステップ数}}{\text{現在のステップ数}}$$

D:ノード間の距離

WWW のような大規模なネットワークでは、ネットワークの全体像を把握するのが困難であるため、任意の二つのノード間の距離を正確に求めるのは困難である。そこで本モデルでは、個々のエージェントが持つ情報に基づいて二つのコンテンツ間の距離を決定する。コンテンツ間の距離は、エージェントの経路について、共通部分までの間隔の和とする。例えば、エージェント A の経路が [a,b,c,d]、エージェント B の経路が [e,f,b,g] である場合を考える。共通部分 b までの間隔は、エージェント A については [a,b] 間で間隔 1、エージェント B については [e,f,b] 間で間隔 2 であることから、距離を 3 とする。この値を、到達可能な最も遠いノードとの距離を用いて正規化する。共通部分が無い場合は、後に示す評価値の式において、距離の値を最大とする。

Rest:残りリンク数

本モデルにおけるコンテンツは、複数のコンテンツとリンクすることが可能である。コンテンツにはリンク数の上限 ($ML: MaxLinks$) が設定されており、上限の値と現在のリンク数の差が Rest となる。ここで、Rest の値が大きい、つまり現在のリンク数が少ないコンテンツは、より多くのリンクを獲得することを目標とする。一方、Rest が小さいコンテンツは、可能な限り評価の高いコンテンツとリンクすることを目標とする。以上の点を考慮し、Rest の値が小さいコンテンツほど評価が高くなるように評価関数を設定する。

以上の値を用いて、ノード i に対するノード j の評価関数を以下に示す。

$$V_{ij} = \frac{wR_j + (1-w)D_{ij}}{\sqrt{Rest_i Rest_j}}$$

Growing Network Simulation using Bi-directional Link
Takahiro KOGA, Noriharu TASHIRO, Tadachika OZONO, Takayuki ITO, Toramatsu SHINTANI

Dept. of Intelligence and Computer Science, Nagoya Institute of Technology. Gokiso, Showa-ku, Nagoya 466-8555 JAPAN

ここで、 w は最新性 R の重みである。本モデルでは、最新性の高いコンテンツは、移動したステップ数が少ないため距離の値が小さくなる。一方、距離が大きいコンテンツは、最新性が低くなる。エージェントが最新性、または距離のどちらを優先するかを w の値によって調整する。

3 シミュレーション

3.1 シミュレーションの設定

本シミュレーションでは、WWW を模した有向グラフで表されるネットワークにおいて、ステップごとに新たなコンテンツが作成される状況を想定する。作成されたコンテンツについて、マッチングが成立すれば、ノード間に新たなリンクが作成される。重み w 、およびコンテンツごとのリンク数の上限 ML を変化させることにより、リンク生成状況の変化を観測する。WWW を表す単純なグラフとして、文献 [1] に示される、インターネット大陸とよばれる構造を利用する。各ノードを、リンクの状況により In, Center, Out に分類する。In ノードとは、他のノードへのリンクが多く、自身へのリンクが少ないノードである。Center ノードとは、他のノードへのリンク、および自身へのリンクが多く、リンク先は Center ノードが多い。Out ノードとは、他のノードへのリンクが少なく、自身のノードへのリンクが少ないノードである。リンクが有向であるため、In ノードから Center ノード、または Center ノードから Out ノードへの移動は可能であるが、逆方向に移動することはできない。

ノード数は、In, Center, Out について各 10 ずつ、計 30 とした。本モデルでは、エージェントの経路に共通部分がない場合は、距離が最大となる。コンテンツに対してノード数が多い場合、エージェントの経路に共通部分が存在せず、多くのコンテンツ間において距離が最大になってしまう。ここでは、経路に基づく距離を利用するため、少数のノードにおいてシミュレーションを行った。初期状態のリンクとして、各 In ノードから Center へ 2 本ずつ、各 Center ノードから他の Center ノードへ 2 本ずつ、そして各 Center ノードから Out ノードへ 2 本ずつリンクを作成した。コンテンツの追加方法は、ステップ毎に In, Center, Out に属するノード 1 ずつについて、それぞれ 1 つ新たなコンテンツを作成した。 ML の値は、全てのコンテンツについて共通の値を用いた。また、 w の値は、全てのエージェントについて共通の値を用いた。

3.2 考察

実験結果から、ノード数が同一の環境においては、 $ML = 1, 2$ のとき、 w の値を変化させても、シミュ

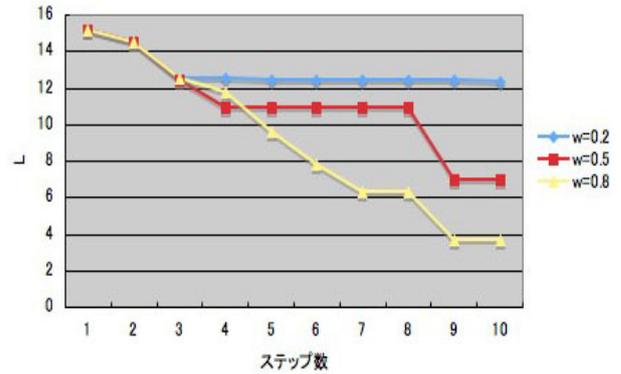


図 1: $ML=5$ のときの L の変化

レーションの過程に目立った変化は見られなかった。 ML が 3 以上のとき、 ML の値が大きいほど w の値による差が顕著に現れている。図 1 は、 $ML = 5$ のときのシミュレーションにおける、平均経路長 L の値である。図 1 より、 w の値が大きいほど、 L の変化が大きく減少していることがわかる。ステップ 10 までに成功したマッチング回数は、 $w = 0.2$ のとき 23、 $w = 0.5$ のとき 33、 $w = 0.8$ のとき 49 であった。したがって、図 1 にみられるような L の値の変化は、マッチングの成功回数の差異によるものであると考えられる。 w の値が小さいとき、つまりノード間の距離を優先した場合、マッチングできずにネットワーク上を移動し続けるエージェントが多く存在するといえる。本モデルでは、エージェントが情報を保持するコンテンツについて、エージェント同士が共通のノードを通っていない場合は、距離 D の値を最大とした。そのため、比較的新しく、移動距離の短いエージェント同士がマッチングしやすくなる。一方、何度かマッチングに失敗したエージェントは、新たなコンテンツが生成され続ける限り、マッチングできずに移動し続けると考えられる。

4 おわりに

本論文では、コンテンツ間のマッチングによるネットワークモデルについて、シミュレーションを行った。シミュレーションの結果、最新性 / コンテンツ間の距離のどちらを優先するかによって、リンクの生成過程が異なることが示された。今後の課題として、初期ネットワークの構成方法、ノード数、コンテンツの追加方法を変化させ、検証することがあげられる。

参考文献

- [1] Barabási, A, “新ネットワーク思考 -世界のしくみを読み解く-”, 2000.
- [2] 古賀 隆浩, 大園 忠親, 新谷虎松 “P2P に基づく双方向 HTML リンクの実現” 合同エージェントワーク ショップ & シンポジウム 2006(JAWS 2006), 2006.