

## Blog スレッドの解析に基づく重要な blogger の発見 -Blog スレッド形成における Trackback の影響に関する考察-

中 島 伸 介<sup>†1</sup> 館 村 純 一<sup>†2</sup>  
原 良 憲<sup>†3</sup> 田 中 克 己<sup>†4</sup>

近年、取得可能な情報量の増大に伴い、信頼できる情報を効率的に取得することが困難になっている。我々は Web 上での有識者を発見し、この有識者が発信するコンテンツに基づいた情報フィルタリングを行うことで、信頼できる情報の取得が可能ではないかと考え、重要な blogger の発見手法の確立を目指した研究を行ってきた。しかしながら、blog スレッドの検出や、重要な blogger の発見に関して、これまでは Trackback リンクを考慮できておらず、blog エントリ同士の関係を適切に抽出することが十分にできているとはいえない。したがって、Trackback の利用状況の調査と、この結果に基づく Trackback の利用方法に関する検討を行った。

## Discovering Important Bloggers based on Analyzing Blog Threads - Influence of Trackback Links on Blog Thread Formation-

SHINSUKE NAKAJIMA,<sup>†1</sup> JUNICHI TATEMURA,<sup>†2</sup> YOSHINORI HARA<sup>†3</sup>  
and KATSUMI TANAKA<sup>†4</sup>

With rapid spread of amount of information, it become difficult for us to discover trusty information efficiently. Consequently, we try to extract blog threads and to discover important bloggers, because we consider that it is possible to provide trusty information by using information filtering based on important bloggers as learned people. Now, we have not used trackback data of blog entries, so that we could not extract blog threads and could not discover important bloggers. Thus, we investigate influence of trackback links on blog thread formation and use of trackback link for discovering important bloggers.

### 1. はじめに

ユビキタス・ブロードバンド基盤は人々が常にオンラインであるという環境をもたらしつつある。このような中で、Web を介したユーザ間の即時的情報流通が広まりつつある。ただし、取得可能な情報量の増大に伴い、信頼できる情報を効率的に取得することが困

難になっている。ここで人間同士での情報のやり取りにおいて信頼性を評価する際には、人はその情報の出所を重要視する傾向がある。すなわち、出所不明の情報よりも、権威者や有識者からの情報の方が人にとっての信頼性は高くなるといえる。

我々は Web 情報の信頼性に関しても同様であると考え、Web 上での有識者を発見し、この有識者が発信するコンテンツに基づいた情報フィルタリングを行うことで、信頼できる情報の取得が可能ではないかと考えた。そこで著者らはこれまでの研究において、blog における議論の一つの形態とも考えられる blog スレッドを提案し、この blog スレッドにおいて重要な役割を担う blog の書き手（以下、blogger）の特徴について検討した。また、これら重要な blogger の判別手法に関して検討を行うとともに、妥当性について議論している。

これまでの我々の研究においては、blog スレッドの検出や、重要な blogger の発見に関しては Track-

<sup>†1</sup> 奈良先端科学技術大学院大学  
Graduate School of Information Science, Nara Institute of Science and Technology  
<sup>†2</sup> NEC Laboratories America, Inc.  
<sup>†3</sup> 日本電気株式会社 インターネットシステム研究所  
NEC Corporation  
<sup>†4</sup> 京都大学大学院情報学研究科社会情報学専攻  
Department of Social Informatics, Graduate School of Informatics, Kyoto University  
および  
独立行政法人 情報通信研究機構  
National Institute of Information and Communications Technology

back リンクを考慮していない。Trackback とは、多くの blog サイトが有する機能であり、他のエントリーに対してリンクを貼った際に、そのリンク参照先へリンクを貼った事を知らせる機能である。したがって、Trackback を伴うリンクと Trackback を伴わないリンクでは、リンクを付けた際の blogger の意識が異なると考えられ、これら Trackback を考慮した方が、blogger 同士の関係をより的確に認識することができ、blog スレッドの抽出や重要な blogger の判別をより適切に行うことができる可能性がある。ただし、最近では Trackback スпамといった Trackback 機能を悪用するケースや、悪用ではないにせよ通常のリンクを貼ることなしに Trackback ping のみを送信するケースもあり、Trackback を利用する場合には Trackback の利用状況の調査が不可欠であるといえる。

そこで最終的には blog スレッドの検出や重要な blogger の発見において Trackback リンクを考慮することを目指す。そこでその前段階として、本論文では、Trackback リンクの検出方法の検討を行うとともに、スパム等を含めた Trackback の利用状況に関する調査および考察を行うことで blog スレッド形成における Trackback の影響に関して考察する。

以下、本論文の構成を示す。2 節では blog および Trackback の概要と関連研究について述べる。3 節では重要な blogger 発見のための blog スレッド解析方針について述べる。4 節では blog スレッド形成における Trackback の影響について述べる。5 節ではまとめと今後の方向性について述べる。

## 2. blog および Trackback の概要と関連研究

### 2.1 blog および Trackback の概要

blog は、アメリカにおいては 1999 年以降、急速に発達し標準化が進みつつある Web コンテンツであり、最新ニュースをいち早く取りあげ、独自の視点で解説をする blog や、ある特定の分野に対してコラムを展開する blog、写真を掲載しそれらにコメントを付ける blog などその内容は様々である。日本では元々「Web 日記」と呼ばれるサイトが数多く存在しており、広義での blog と定義できる。現在では「MovableType<sup>1)</sup>」などの blog サイト構築ツールなどが公開されており、また、ホスティングサービスを行っているサイトも増えており、誰でも簡単に blog サイトを立ち上げるための環境が整っている。2005 年 5 月の総務省の発表<sup>2)</sup>によれば、2005 年 5 月現在で小規模な事業者を含めて 115 社の事業者が一般ユーザ向けに blog サービスを提供している。また、2005 年 3 月末時点で延べ約

335 万人の国内 blog 利用者は、2007 年 3 月末には延べ約 782 万人に上ると予測している。

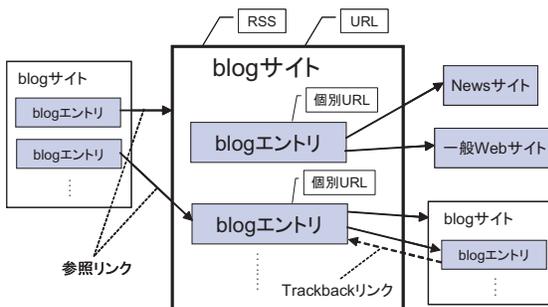


図 1 典型的な blog サイトの例

図 1 に典型的な blog サイトの例を示す。blog サイトは、そのトップページに最近（例えば一週間）書かれた「エントリー」と呼ばれる個別書き込み記事を複数表示している。通常は blog サイトの管理者のみがエントリーを追加することができ、この点が Web 掲示板とは異なる。新しいエントリーが追加されれば、古いエントリーはトップページからは削除されるが、各エントリーが保持している個別 URL を辿れば、トップページから削除された後でも閲覧することが可能である。

また、blog サイトトップページについては、RSS (Rich Site Summary もしくは RDF Site Summary) と呼ばれる XML で記述されたサイトの要約を公開していることが多く、RSS のみを巡回することで blog サイトの更新情報等を取得することが可能となっている。

他人の blog エントリーに対して意見を延べる際には、自分の blog エントリー内において相手のエントリーをリンク参照するということがしばしば行われるが、この際に Trackback という機能を利用することができる。Trackback の手順としては、まず他人の blog エントリーをリンク参照した際に、リンク参照したことを知らせる Trackback ping の送信を行う。すると Trackback ping を受け取った blog エントリーにて、自分のエントリーを参照している blog エントリーへのリンクが自動で貼られることになる。通常のリンクでは、参照元はリンクを貼られたことを知ることはできないが、Trackback により参照された blog エントリーの投稿者がリンクを貼られたことを知ることができる。

blog サイトであることと条件は明確なものはないが、本研究では RSS を保持するものを blog と扱うことにしている。ただし、ニュースサイトの中には RSS を公開しているものもある。したがって、RSS が存

在しても、明らかにニュースサイトであり blog とは考えにくいと認められる場合には、これを除外して考える。

## 2.2 関連研究および技術

### ● blog による情報の広がり

Kumar ら<sup>3)</sup> および, Gruhl ら<sup>4)</sup> は, blog 空間の進化や広がりに関する調査研究を行っている。Kumar らは, 25,000 の blog サイトとその中の 750,000 本のリンクについて解析している。また, blogspace と名づけたハイパーリンクによる blog 群のつながりに注目し, この Blogspace における blog コミュニティの抽出とこの blog コミュニティの進化に関する調査研究を行っている。

Gruhl らは, 11,000 以上の blog サイトにおける 400,000 以上の blog エントリについて解析している。この中で, blogspace におけるマクロな視点によるトピックの伝播の特徴付けと, ミクロな視点による個々の blog 同士のトピックの伝播の特徴付けを試みている。この中で, blogspace において内部的に発生する議論である Chatter と, 外的要因により発生する Spikes という尺度を用いて, トピック伝播のモデル化を行っている。ただし, これらの研究では Trackback による blog エントリ同士の関係についてはほとんど言及されていない。

### ● リンク構造の時間特性に着目した blog 時系列解析

著者らは, Web コンテンツの信頼性評価を目的とした blog 解析手法に関して提案している<sup>5)6)</sup>。この中で, blog エントリが形成する blog スレッドを定義し, この blog スレッド内における blog サイトの役割の判別方法に関して議論している。blog サイトの役割としては, “Topicfinder”, “Agitator”, “Summarizer”などを定義し, blog スレッドのリンク構造の解析および時系列解析によってこれらの判別方法を提案している。

ただし, これらの研究では, エントリ同士の関係を調べる際には Trackback リンクによるつながりは考慮できていない。

### ● blog 情報に基づく信頼値の算出方式

竹原らは, blog サイトが, 参照している Web コンテンツに対して何らかの評価を示しているケースに着目し, blog 情報に基づく Web コンテンツの信頼値の算出方式を提案している<sup>7)</sup>。この中で blog サイトの熟知度と, blog エントリ内での評価度という指標を提案し, これに基づい

て Web コンテンツの信頼値を定義している。この研究は, blog エントリ内の Web コンテンツに対する評価を利用した, 検索エンジン結果の修正方法を提案しているものであり, 重要な blog データそのものを取得および検索しようとするものではない。

### ● Trackback 機能を応用したシステム開発

Trackback に着目した研究は多くはないが, 渡邊らは, 公共性の高い情報や評判情報を蓄積して提供するコミュニケーションに関する研究を行っている<sup>8)</sup>。ただし, これは Trackback 機能を応用した新たなシステムの提案であり, Trackback の解析に基づいたエントリ同士の関連について調べたものではない。

## 3. 重要な blogger 発見のための blog スレッド解析方針

本節では, これまでに行ってきた重要な blogger 発見のための blog スレッド抽出とその解析方法について説明する。

### 3.1 blog データモデル

ここでは, 図2および図3に blog サイトおよび blog エントリの例を示すと共に, blog サイトおよび blog エントリのデータモデルを定義する。

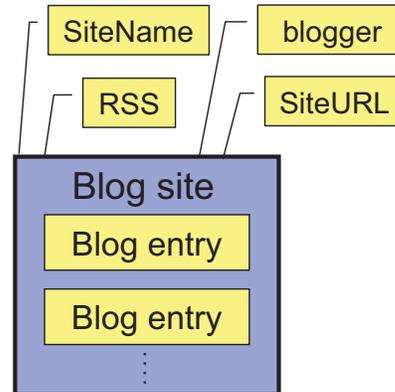


図2 blog サイトの例

$$site = (siteURL, RSS, blogger^+, siteName, entry^+)$$

$$entry = (permaLink, blogger, time, title^?, description, comment^*)$$

$$comment = (blogger, permaLink, content, time)$$

なお, *siteURL* は, blog サイトのトップページの URL であり, *blogger* は, 各 blog サイトを管理し エントリを書き込むユーザである。多くの場合は,

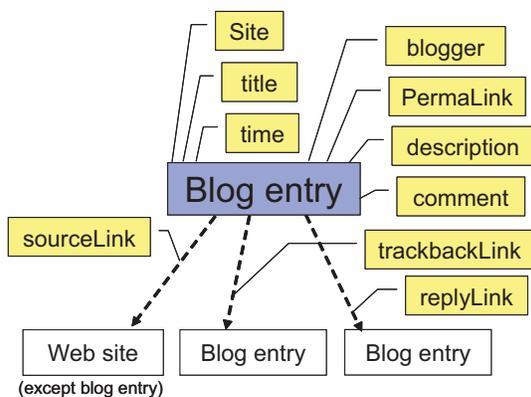


図 3 blog エントリの例

1つのblogサイトには1人のbloggerが存在する。permaLinkとは、blogサイトを経由せず、エントリに直接アクセスするための個別URLである。timeは、エントリを投稿した際のタイムスタンプであり、descriptionはユーザにより実際の書き込み内容であり、blogエントリページ内のフォーマットに含まれるリンクやcomment等は含まない。

なお、リンクに関しては、replyLink, sourceLinkおよびTrackbackLinkが存在する。replyLinkおよびsourceLinkは、共にdescription内からのリンクであるが、replyLinkは他のblogエントリへのリンクであり、sourceLinkはエントリ以外のWebページへのリンクである。TrackbackLinkは、他のblogエントリから送信されたTrackback pingに基づいて自動生成されるリンクである。

図4に、blogスレッドの例を示す。なお、本研究ではblogスレッドを以下のように定義する。

$$\begin{aligned}
 \text{thread} &:= (V, E) \\
 V &= V_W \cup V_E, \quad E = L_s \cup L_r \\
 L_s &\subseteq \{(e, e') | e \in V_E, e' \in V_W\} \\
 L_r &\subseteq \{(e, e') | e \in V_E, e' \in V_E\}
 \end{aligned}$$

すなわち、blogスレッドは、blogエントリをノード、replyLinkおよびsourceLinkをエッジとした、有向連結グラフとして定義できる。ただし、 $V_W$ はWebページ集合であり、 $V_E$ はblogエントリ集合である。また、 $L_s$ は、sourceLinkの集合を示し、 $L_r$ は、replyLinkの集合を示す。

これまではblogスレッド形成におけるTrackbackリンクの存在意義については考慮していないが、通常のreplyLinkを貼る際に、Trackback pingを送信しているかどうかということは、エントリ同士の関係

を把握する上で重要であると考えられる。また、最近では通常のreplyLinkを貼ることなしにTrackback pingを送信することもある。したがって、blogスレッドの抽出においては、Trackbackの有無を考慮する必要があると考えている。

### 3.2 blogデータの収集およびスレッドの抽出

以下にblogデータの収集およびスレッドの抽出方法を示す。

- (1) システムは、“ping.bloggers.jp”<sup>9)</sup>等で公開しているopmlファイルを巡回し、クローリング時に利用するRSSリストの更新を行う。
- (2) 更新されたRSSリストに基づいて各RSSを巡回し、未登録のblogエントリが存在すればその個別URLをblogエントリリストに登録する。
- (3) blogエントリリストに基づいて、未収集のblogエントリを巡回してhtmlファイルを保存する。保存したhtmlファイルのタグ解析により、書き込み部分であるdescriptionを特定し、replyLinkおよびsourceLinkを抽出して、リンクリストに保存する。
- (4) 保存されたリンクリストの中から、リンク先がエントリであるreplyLinkをピックアップして、このreplyLinkにてつながっているblogエントリ集合をblogスレッドとする。
- (5) (1)へ戻る。

ここで、blogエントリのhtmlファイルのタグ構造は、blogサイトにより異なるため、blogサイト毎に解析ルールを設定する必要がある。本研究では25の主要なblogホスティングサイトに対するタグ解析用ルールを作成し、これらのblogサイトのhtmlタグ解析を可能にしている。なお、2005年6月1日までに収集したデータは、RSSが約100万、blogエントリが約1,500万、リンク数が約300万となっている。

### 3.3 重要なbloggerの判別

重要なbloggerとは、あるトピックに関する一時的なコミュニティとも言える、blogスレッド内において、重要な役割を担うbloggerのことを言う。本研究にて提案する重要なbloggerを以下に示す。

#### • Topicfinder

Topicfinderとは、議論が盛んに行われたWeblogスレッドにおいて、スレッドが立ち上がった初期段階において、エントリを提供することが多いWeblog投稿者である。つまり、Topicfinderは、成長前の段階からWeblogスレッドにて議論するための良いトピックを見つけることが多いWeblog

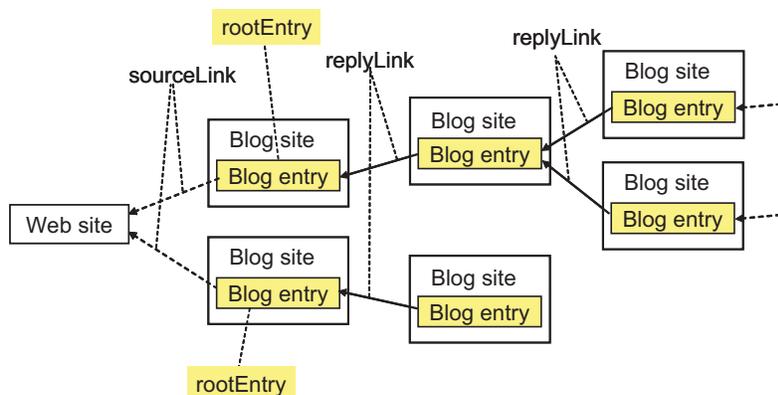


図 4 blog スレッドの例

投稿者であるといえる。

- **Agitator**

Agitator とは、議論が盛んに行われた Weblog スレッドにおいて、スレッドでの議論が盛んになる直前にエントリを提供することが多い Weblog 投稿者である。Agitator のエントリを監視することで、Weblog スレッドが成長する時期を予測するための判断材料にすることができる。

- **Summarizer**

Summarizer とは、あるトピックに関するスレッド内において、他の多くの Weblog エントリを参照することが多い Weblog 投稿者である。Summarizer のエントリを監視することで、あるトピックに関する Weblog スレッドをまとめたような書き込みを効率よく取得できる可能性がある。

これら重要な blogger を判別する際には、上記の特徴に基づいて判別ルールを設定することで、重要な blogger の発見を試みるものである。この際に、どのようなデータに基づいて blog スレッドを抽出するか、どのようなデータに基づいて重要な blogger の判別を行うのかということは、非常に重要な問題である。これまでは Trackback リンクを用いた解析は行っていないが、本研究にて Trackback の blog スレッド形成への影響を把握することができれば、より適切に重要な blogger の発見が行えるものと考えている。

#### 4. blog スレッド形成における Trackback の影響

第 1 節でも述べたが、blog スレッドの検出や、重要な blogger の発見において、Trackback を考慮した方が blogger 同士の関係をよりの確に認識することができ、blog スレッドの抽出や重要な blogger の判別

をより適切に行うことができるようになってきている。ただし、Trackback 機能を悪用するユーザも存在するため、Trackback を blog スレッド抽出や重要な blogger 発見に利用する場合には Trackback の利用状況の調査が不可欠である。本節では、Trackback 利用の現状に関する調査を行い、blog スレッド形成への影響について考察する。

なお、通常、tarackback リンクは *description* とは別の場所に掲載されるものであり、これをシステムが認識するために、blog サイト毎に Trackback 抽出のためのタグ解析ルールを設定する必要がある。本論文では、URL の一部が以下のものと合致する blog ホスティングサイトに対して、Trackback 抽出用タグ解析ルールを設定した。

```
nifty.com/          typepad.com/      mo-blog.jp/
typepad.jp/        weblogs.jp/       no-blog.jp/
blog.ocn.ne.jp/   blogzine.jp/      de-blog.jp/
```

以下、これらのサイトを Trackback 解析対象 blog サイトと呼ぶ。

##### 4.1 Trackback の形式的状況

本来 Trackback とは、ある blog エントリから他の blog エントリにリンクを貼ったことを相手の blog エントリに通知 (Trackback ping の送信) するものであり、通知を受けた blog エントリは、どの blog エントリからリンクを貼られているのか (Trackback リンク) を表示することになる。しかしながら、実際にはその blog エントリへのリンクを貼ってなくても、Trackback ping を送信することは可能であり、対応するリンクが存在しなくても相手の blog エントリから Trackback リンクを貼ってもらうことが可能である。本論文ではこれを“空 Trackback ”と呼ぶ。最近はこのことを悪用して、自分の商業用サイトへのリンクを増やしたりするトラックバックスパムが問題となって

表 1 Trackback の形式的状況

	TB(j-i)		TB(j-i) ×	
	link(j-i)	link(j-i) ×	link(j-i)	link(j-i) ×
link(i-j)	2.3%	1.7%	0.2%	9.8%
link(i-j) ×	1.7%	5.8%	10.3%	68.2%

表 2 Trackback の内容的状況

	TB(j-i)		TB(j-i) ×	
	言及 (j-i)	言及 (j-i) ×	言及 (j-i)	言及 (j-i) ×
関連	2%	34%	8%	56%
関連 ×	0%	0%	0%	0%

いる。

そこで、まずは、実際の Trackback の中で、空 Trackback がどの程度存在するのかを調査した。2004 年 11 月 1 日から 2005 年 2 月 28 日までの Trackback 解析対象 blog サイトにて調査したところ、Trackback リンクが 15,035 存在し、その中で空 Trackback は 12,722 存在した。すなわち、調査対象の Trackback うち、84.6% が空 Trackback であり、実際のリンク参照無しに Trackback ping を送信しているケースがかなり多いことが分かった。

実際に幾つかの空 Trackback のエントリを閲覧してみたところ、Trackback に対応する通常にリンクは存在しないものの、Trackback を返しているケースが幾つか認められた。すなわち、相互 Trackback である。これは一般的に認識されている Trackback の利用方法とは異なると考えられるので、より詳細に Trackback の形式的状況について調査する。

blog エントリ (i) から、エントリ (j) への Trackback リンク (TB(i-j)) が存在する時の、この 2 つのエントリ間の Trackback と replyLink の状況について調査した。TB(i-j) は、エントリ (i) に存在する Trackback リンクであり、エントリ (j) から送信した Trackback ping に基づいて生成されるものを指す。なお、対象データは、2004 年 12 月 1 日から 31 日に投稿された blog エントリのうち、Trackback 元も含めて、本論文の Trackback 解析対象サイトのエントリ (総数 1874) とした。表 1 に、Trackback の形式的状況を示す。

なお、TB(j-i) とはエントリ (j) からエントリ (i) への Trackback リンクであり、および × にてその存在を表現している。link(i-j) は、(i) から (j) への通常のリンクを示す。また、表中の全てのケースにおいて、TB(i-j) は存在している。

ここで空 Trackback とは、“link(j-i)”が存在しない状態であり、合計 85.5% となる。すなわち、かなりの確率で空 Trackback が存在していることがわかる。また、ここで注目したいのは、“TB(j-i)”が存在する、

相互 Trackback の状態である。これは、合計 11.5% となる。

これまで我々は空 Trackback は、基本的には悪意があるものと考えていた。つまり、リンクを相手に対して貼らないのに、Trackback ping を送信することで自分宛のリンクを生成して、自分のサイトへのアクセス量の増大や、PageRank 値の向上を狙うためである。しかしながら、相互に Trackback をかけるということは、お互いに相手の存在を認めていると考えられ、“空 Trackback は悪意を伴う”という認識が当てはまらないことになる。

そこで、次節にて、これら Trackback を伴う 2 つの blog エントリ同士は、内容的にどのような関係になっているのかということについて調べる。

#### 4.2 Trackback の内容的状況

ここでは特に、空 Trackback を伴う 2 つの blog エントリ同士の内容的関係について調べる。対象データは、2004 年 11 月から 2005 年 2 月までのエントリで空 Trackback を伴うものから無作為に選択した 50 ケースに関して、人手により内容的状況の確認を行った。表 2 に、Trackback の内容的状況を示す。なお、表中の全てのケースにおいて、TB(i-j) は存在している。

ここで、“関連”とは、2 つのエントリ同士の内容的関連の有無であり、例えば言及している内容が似通っているかどうかということを示している。“言及”とは、Trackback ping を送信したエントリ (j) 内において、エントリ (i) に対する言及の有無を表している。

調査実験を行った blog データ量が十分とはいえないが、関連のないエントリに対する Trackback は存在しなかった。内容的に明らかに関連のない迷惑な Trackback は、Trackback を受けた blogger が手動で削除した可能性があるため、元々存在しないのか、存在していたが削除されたのか、については不明である。

また、Trackback ping を送信していながら、送信先のエントリに関する言及がないものが 90% を占めるといった結果となった。そのうち、相互 Trackback と

なっているケースが、34%も存在した。相互 Trackback という、お互いの存在を認知していると考えられるが、それにもかかわらず相手への直接のリンクも貼らず、そのエントリの内容に関する言及もないという Trackback の利用方法はかなり独特なものであり、これまでは一般的にはあまり認識されていなかったものである。

これらのエントリを実際に閲覧してみたが、内容的には非常に関連の深いものでありながら、他のエントリの内容に関しては全く触れていない。さらにこのような関係は1つのエントリに対してだけでなく、複数のエントリと同時に相互 Trackback の関係を保っているケースも少なくない。そうすることで、これらのエントリははコミュニティのようなものを形成しているようである。このように明示的かつ強固なつながりではなく、緩やかな関係の中でつながりを保つというのは日本独特の Web 文化のようにも考えられる。今回の実験データは全て JP ドメインのサイトであるので、今後、JP ドメイン以外のサイトとの比較を行うことで、国毎の Trackback の利用形態の違いがあるのかどうかということについても明確にすることも必要であると考えられる。

#### 4.3 考 察

本論文における Trackback に関する調査実験では、空 Trackback や、対象エントリへの言及のない Trackback がどの程度存在するのかということに関して調査すると共に、相互 Trackback や相互リンクのパターンに関して調査した。この中で、空 Trackback は必ずしもスパムではなく、あまり一般的には認識されていないような緩やかなコミュニティを形成しているケースも数多く存在することを明らかにした。

今回の実験結果に基づいて、即座に Trackback データを blog スレッド抽出や、重要な blogger の判別に利用することはできないが、通常のリンクのつながりとは違う、blog 同士のつながりの存在を明らかにしたことで、より詳細な blog 解析が行える可能性が出てきたということが言える。

今後は、Trackback 解析対象 blog サイトの範囲を拡大すると共に、海外の blog サイトの Trackback 解析も可能にすることで、より大きなデータセットの獲得が必要である。さらに、Trackback リンクにより形成される blogger 同士のつながりに関する解析を行うことで、Trackback 解析の意義を明確にし、最終的には blog スレッド抽出および重要な blogger の発見に利用したいと考えている。

## 5. おわりに

本論文では、Trackback リンクの検出方法の検討を行うとともに、Trackback の利用状況に関する調査を行うことで blog スレッド形成における Trackback の影響に関して考察した。以下に、本論文のまとめを示す。

- blog サイト、blog エントリ、blog スレッドのデータモデルを定義した。
- blog エントリからの Trackback リンクの抽出のためのルールベースを作成し、JP ドメインの blog サイトにおける Trackback リンクに関する調査実験を行った。
- Trackback の形式的状況に関する調査実験では、Trackback のほとんどが空 Trackback (対応する通常リンクを伴わない Trackback) であることや、相互 Trackback の存在について明らかにした。
- Trackback の内容的状況に関する調査実験では、Trackback ping を送信しながらも、相手の blog エントリに関する言及がない blog エントリの存在や、相互の空 Trackback に基づく緩やかなコミュニティ形成が行われていることを明らかにした。

今後は、Trackback 解析対象 blog サイトの範囲拡大と、重要な blogger 発見における Trackback 解析の意義に関して検討するつもりである。

## 謝 辞

本研究の一部は、平成 17 年度科学研究費補助金 若手研究 (B) 「重要な Weblog 投稿者の発見と信頼性の高い情報の検索・呈示に関する研究」(課題番号: 17700132) による。ここに記して謝意を表します。

## 参 考 文 献

- 1) movabletype.org, <http://www.movabletype.org/>
- 2) ブログ・SNS (ソーシャルネットワーキングサイト) の現状分析及び将来予測, 総務省報道資料 (平成 17 年 5 月 17 日), [http://www.soumu.go.jp/s-news/2005/050517\\_3.html](http://www.soumu.go.jp/s-news/2005/050517_3.html)
- 3) Ravi Kumar, et al: "On the Bursty Evolution of Blogspace", *The Twelfth International World Wide Web Conference (2003)*. <http://www2003.org/cdrom/papers/refereed/p477/p477-kumar/p477-kumar.htm>
- 4) D. Gruhl, et al: "Information Diffusion Through Blogspace", *The Thirteenth International World Wide Web Conference (2004)*.

- <http://www2004.org/proceedings/docs/1p491.pdf>
- 5) 中島伸介, 館村純一, 日野洋一郎, 原 良憲, 田中克己: リンク構造の時間特性に着目した Weblog 解析に基づくコンテンツの信頼性評価の検討, DBSJ Letters, Vol.3, No.1, pp.109-112, 2004 年 6 月.
  - 6) Shinsuke Nakajima, Junichi Tatemura, Yoichiro Hino, Yoshinori Hara, and Katsumi Tanaka: "Discovering Important Bloggers Based on Analyzing Blog Threads", WWW2005 Workshop on the Weblogging Ecosystem, May 2005.  
<http://www-idl.hpl.hp.com/blogworkshop2005/nakajima.pdf>
  - 7) 竹原幹人, 中島伸介, 角谷和俊, 田中 克己: Web 情報検索のための Blog 情報に基づくトラスト値の算出方式, DBSJ Letters, Vol.3, No.1, pp.101-104, 2004 年 6 月.
  - 8) 渡邊正弘, 大倉典子: blog の機能を利用した位置情報コミュニケーションシステム, データベースと Web 情報システムに関するシンポジウム (DB-Web2004) 情報処理学会シンポジウムシリーズ Vol.2004, No.14, pp.73-80, 2004 年 11 月.
  - 9) PING.BLOGGERS.JP, <http://ping.bloggers.jp/>