

情報コンテンツの信頼性とその評価技術

Credibility of Information Content: Concepts and Evaluation Technology

加藤 義清*¹, 黒橋 禎夫*²*¹, 江本 浩*¹
Yoshikiyo Kato, Sadao Kurohashi and Hiroshi Emoto

*¹情報通信研究機構

National Institute of Information and Communications Technology

*²京都大学大学院情報学研究科

Graduate School of Informatics, Kyoto University

Abstract: 従来は一部の機関、著者が特権的に有していた情報発信という機能が、ウェブの登場により一般の個人に開放されることとなった。特に、近年では消費者発信メディアの普及に伴いその傾向は顕著となっている。従来では得られなかった情報が得られるようになり、その利用価値は高まっている。その反面、信頼性という観点から情報を評価、選別し利活用することが難しくなっているのも事実である。本稿では、情報の信頼性について、関連する概念について整理し、ウェブ上の情報の信頼性評価を行う上での課題を挙げて要素技術をまとめる。

1 はじめに

90年代半ばにウェブが登場し、誰もが容易に情報発信できるようになった。ウェブの登場以前には、個人で世界中の人に向けて情報発信することはとても考えられなかったことである。更に、近年では、消費者発信メディア (CGM: Consumer Generated Media) として、ブログや SNS (Social Network Service) の普及が注目を集めている。これらのサービスにより、一般の個人が容易に情報発信ができるようになり、商品についての口コミ情報など、CGM が消費者の行動に一定の影響を持つような状況が現れつつある。実際、インターネットの普及に伴う情報収集コストの低下にとともに、商品の購入前にネットで情報を収集する利用者が 62% に上っており、消費購買行動に変化が現れている [22]。

このように、誰もが容易に情報を発信するようになると、今までは得られなかったような情報が得られるなど、ウェブから得られる情報が多様性が増す一方で、その信頼性が問題となってくる。従来のメディアでは、情報が発信される前にその内容について何らかの形でチェックがなされるのが普通である。しかし、ウェブでは何のチェックも無しに情報を発信できるため、ウェブから得られる情報は必然的に玉石混淆となる。現状では、提示された情報そのものの正しさについての判断

は各利用者に任されている。このことにより、利用者は間違った情報を信じることによりリスクを背負ったり、どの情報を信用してよいか分からないことによって、情報を十分に有効活用できなかつたりするなどの問題が生じる。

これらの問題を解決する方法として、2つのアプローチが考えられる。1つ目のアプローチは、情報の信頼性を評価して、信頼性の低い情報をフィルタリングするとともに、利用者がアクセスする情報に評価結果を付与して提示するやり方である。これにより、利用者に質の高い情報を選択的に提供し、情報の利活用に当たってその信頼性を判断するのを支援する。2つ目のアプローチは、ウェブ全体として流通する情報の質を高めるために、質の高い情報の発信を促すような情報発信プロセスを提案、構築していくやり方である。

情報の評価 玉石混淆の関連情報の内、信頼できる情報と出来ない情報を選り分ける。

質の向上 信頼度の高い情報の発信を積極的に促す。

本稿では、この2つのアプローチのうち、1つ目の情報の信頼性の評価について取り扱う*¹。2章で情報の信頼性についての概念を整理する。3章では、情報の信頼性を評価するための4つの視点を提示し、それぞれの視点に関連する技術について述べたのち、4章で本稿をまとめる。

加藤 義清, 情報通信研究機構 知識創成コミュニケーション
研究センター 知識処理グループ

〒 619-0298 京都府相楽郡精華町光台 3-5

Tel: 0774-98-6874, Fax: 0774-98-6960

E-Mail: ykato.at.nict.go.jp

*¹ 2つ目の信頼度の高い情報の発信を促す情報発信プロセスについては、<http://kc.nict.go.jp/c3/infocred.pdf> を参照されたい。

2 情報の信頼性

情報の信頼性について考えるときに、どのようなことが問題になってくるだろうか。まず、「信頼性」と言ったときに何を意味するのか明らかにしなければならない。コンピュータサイエンスの分野において、「信頼性」と言ったときに、おおむね次に述べる3つの概念の何れかを指していると思われる。1つ目は、英語の *reliability* に相当する概念である。これは、計算機のハードウェアやソフトウェアが仕様通りに動作するかという意味における信頼性である。2つ目は、英語の *trustworthiness* に相当する信頼性である。これは、ある主体（人間に限らず、ソフトウェアエージェントも含む）を、ある事柄（「このクライアントはシステムに対する正当なアクセス権を有している」、「このオンラインショッピングサイトは個人情報を不正使用しない」など）に関して信頼できる、という意味における信頼性である。3つ目は、書かれている情報の内容を信頼できるという意味での信頼性で、英語の *credibility* に相当する。この概念では情報の内容の正しさが問題となる。関連する技術として、情報が改ざんされていないことを保証したり、確かに誰かが書いたものであることを保証したりする電子署名技術が挙げられるが、直接情報の内容を対象としたものではない。

本稿で扱う対象となる「情報の信頼性」は、主に上で述べた3つ目の信頼性 (*credibility*) である。ただし、情報の信頼性を考えるときに、その情報の発信者の信頼性が深く関わってくるため、2つ目に挙げた意味での信頼性 (*trustworthiness*) の問題にもふれざるを得ない。以降、それぞれの概念について詳しく見ていくことにする。

2.1 Trust としての信頼

Luhmann[5] は、信頼は社会の複雑性を縮減するためのメカニズムだとしている。高度に複雑化した社会において、ある行動を起こそうとする際に、関係する情報を集めて、自ら合理的に判断しようとした場合、人間の能力に限界があるために、取りうる行動が極度に制約されてしまうことになってしまう。他人、ならびに社会システムを信頼することにより、リスクを負うことにはなっても、行動の可能性は信頼を用いない場合に比べてはるかに広がるのである。これを情報の信頼性について考えてみると、情報発信者に対する信頼をもって情報の信頼性を判断すると言うことは我々がごく普通に行っていることである。例えば、メディアで報道される内容について、視聴者や読者の各個人が逐一その正しさについて検証している訳ではない。

全く鵜呑みにするわけでもないが、ある程度信頼性があるものと受け止めるのが普通であろう。

Solomon と Flores[12] は、信頼は無条件なものではなく、信用する対象や文脈に応じてその範囲が限定されるとしている。例えば、郵便配達人を、郵便を配達してくれることについては信頼するが、突然家に上がり込んだりすれば、疑念を抱く。このように、信頼というのは、「誰を」（信頼の対象）「何について」（信頼の範囲）信頼するかという形で限定される。

山岸 [20] は、信頼の概念について社会心理学の立場から整理を行っている。その中で、信頼と信頼性の違いについて重要な区別を行っている。信頼は「信頼する側の特性」であり、信頼性は「信頼される側の人格特性」であるとしている。更に、人を信頼するときの相手の意図に対する期待を中心に据えて、人間関係や人格特性に基づく信頼を詳しく分類している。意図に対する期待としての信頼と区別されるべき信頼として、能力に対する期待や、相手の損得勘定に対する評価に基づく安心 (*assurance*) など挙げられている。これらの区別は、Solomon と Flores のいう信頼の範囲を決めるための条件として考慮すべき項目に挙げられるだろう。

Falcone と Castelfranchi[1] は認知社会学的な分析に基づく信頼のモデルを提案している。このモデルでは、あるエージェントがある目標を達成しようとしているときに、その目標を実現するために実行されなければならないアクションを他のエージェントに委任するかどうかを決定するという状況が前提としておかれている。すなわち、他のエージェントにあるタスクを委任をするかどうかの意思決定をするときの、エージェントの信念状態として信頼をモデル化している。このモデルでは、信頼という行為は3つの段階を経て行われる。信頼をする側のエージェントには、まず最初に「信頼の意向」 (TD: Trust Disposition) があり、次に「信頼することの決定」 (DtT: Decision to Trust) をして、最後に信頼の実行 (AoT: Act of Trusting) に至る。信頼の意向は、ゴール g を達成しようとする信頼する側のエージェント x が g を達成するのに関係するタスク α を信頼される側のエージェント y が実行可能であること（山岸の整理における「能力に対する期待」に相当）、更に y がそのタスクを行う意図を持っていること（山岸の「意図に対する期待」に相当）についての信念を x が有している状態としてモデル化されている。信頼することの決定は、TD の条件に加えて、ゴール g を達成するのに、タスク α を委任しなければならない、あるいはした方が良いという判断が加わったものである。信頼の実行は DtT を受けて実際

にタスクを他のエージェントに委任する行為である。
このように、様々な分野で扱われている信頼について見たときに、以下のことが言える。

- 信頼は、意図を持った主体を対象としたものである。
- 信頼は、信頼する相手の能力と相手の意図に対する期待により構成される。
- 信頼は、信頼する側の目標など、状況によってその範囲が限定される。

以上の議論より、情報発信者に対する信頼を考えるときに、発信者の能力、発信者の意図、さらに利用者の目的と発信者の能力や意図、その他の状況の関係を考慮する必要があることが明らかであろう。

2.2 Credibility としての信頼性

情報の信頼性 (credibility) とは何であろうか。情報の内容にもよるが、情報が事実や知識に関するものである場合には、その内容の真偽や正確さが問題となるであろう。一方、内容が意見である場合には、情報を受け取る人の立場や状況によって当然受け止め方が変わるので、真偽そのものは一意に決定できない。

情報の信頼性というのは、情報の真偽や正確さと等価ではない。もし、情報の真偽や正確さが分かっているのであれば、その情報の信頼性を問題にする必要はそもそもないであろう。むしろ、情報の信頼性というのは情報の真偽や正確さそのものは分からないが、それを推定するために利用するものであると言える。つまり、情報の信頼性とは、ある主体にとって、ある情報に含まれている内容が真実であるか、どれくらい正確であるという信念を形成するのに利用される様々な特性のことを指すと考えられる。

Fogg と Tseng[2] は計算機の出力について利用者がどのように受け止めるかという意味での情報の信頼性について論じている。この場合、情報の発信者は計算機で、計算機の出力が信頼性を判断する対象となる情報にあたる。Fogg らは計算機についての信頼性について、利用者がその出力をどのように受け止めるかという観点から、次の4つの型に分類している。

仮定された信頼性 (presumed credibility) 情報を受け取り手が、信頼する対象についての「一般的な想定 (general assumptions)」に基づいて決定される信頼性。「デフォルト値としての『一般的な信頼 (general trust)』 ([20], p.42)」に基づいた信頼性と言える。

評判に基づく信頼性 (reputed credibility) 第三者の報告に基づいて受け止められる信頼性。例えば、あるコンピュータ雑誌の、XYZ社のソフトウェアが非常に優れているというテスト結果についての記事を読むことによって、XYZ社のコンピュータ関連製品の品質を信頼するような場合にあたる。

表面的な信頼性 (surface credibility) 単純な検査 (simple inspection) に基づいて判断される信頼性。例えば、表紙で本についての評価を下したり、ウェブページの見た目のデザインでその信頼性を判断したりするようなこと。

経験に基づく信頼性 (experience credibility) 直接の経験に基づいて判断される信頼性。

この分類は、人間が信頼性を判断する様式を分類したものと見える。

一方、武田 [17] はページの信頼性について、信頼性を推定するのに利用する情報という観点で分類を行っている。まず、情報は5W(who, what, where, when, why)で規定されるとし、それらに加えてウェブの特徴として情報についての情報(それを武田は「評判 (reputation)」と呼んでいる)も、情報を規定するものに含めている。このうち、「何を(内容)」,「誰が(情報提供者)」,「コンテキスト (where, when, whyで規定される情報)」,「評判」の4つに大きく分類して、ページの信頼性を整理している。

ページの内容に基づく信頼性 内容として記述されている事実の真偽に基づく信頼性。

ページのコンテキストに基づく信頼性 ページの記述・公開のコンテキストに基づいて判断される信頼性。公開されているサイト、公開日時、電子署名などの有無などがコンテキストの例として挙げられる

ページの評判に基づく信頼性 ページを利用している他のユーザの情報に基づく信頼性。PageRankなど。

エージェントの信頼に基づく信頼性 著者に対する信頼 (trust) に基づいて計られる信頼性。

3 情報の信頼性評価の視点

前章までに、「信頼性」という言葉に関連する概念を整理し、「信頼 (trust)」および「信頼性 (credibility)」について様々な側面があることを紹介した。ここでは、ここまでに述べた側面を踏まえつつ、情報の流通モデ

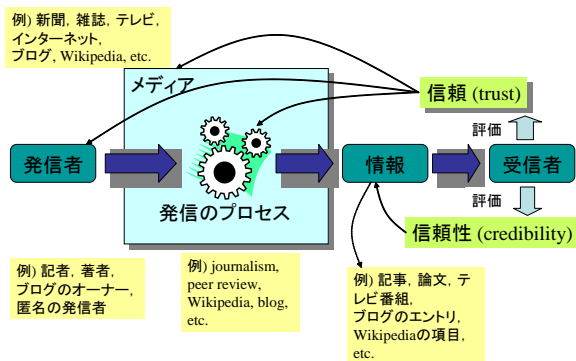


図 1: 情報の流れと信頼性

ルを提示し、更に情報の信頼性を評価するための4つの視点を導入する。

図1は情報の発信者から受信者までの情報の流れと情報の信頼性の関係を示したものである。発信者から送られた情報は、何らかのメディアを経由して、受信者の元に届く。このメディアは、新聞であったり、雑誌であったり、テレビだったり、あるいはインターネットのサイトであったり、ブログであったり、Wikipediaであったりする。それぞれのメディアでは何らかの発信プロセスが働いている。例えば、新聞や雑誌では、記者が書いた記事がそのまま掲載される訳ではなく、編集局の方針に従って編集され、加工された情報が受信者の手元に届くことになる。学術雑誌であれば、ピアレビュー(査読)を通して、その道の専門家が学術的に出版に値すると判断された論文のみが出版されることになる。インターネットの場合、ページを公開した時点で世界中からアクセス可能となる。しかし、その情報の存在を知ってもらわなければ誰もアクセスしないので、注意を集めるために様々な手段を使うことになる。例えば、検索エンジンで特定のキーワードで上位にランクされるように工夫をしたり(SEO: Search Engine Optimization)、Trackbackを使って他人のサイトに自分のサイトへのリンクを張ったり、RSSにより更新情報を配信したりすることがなされる。

この情報の流通モデルと情報の信頼性の関連であるが、ここでは受信者が受け取った情報そのものを見て判断する信頼性(credibility)と、受け取った情報の発信者、メディア、発信のプロセスといったいわゆる「情報の出所」に対する信頼(trust)に基づいて判断される信頼性に分けられる。これはちょうど、2章で見たtrustとしての信頼とcredibilityとしての信頼性の区別に対応している。

次に、図2に情報の信頼性評価に関わる要素と課題を示す。この図の中で、周辺にある4つの要素が、信

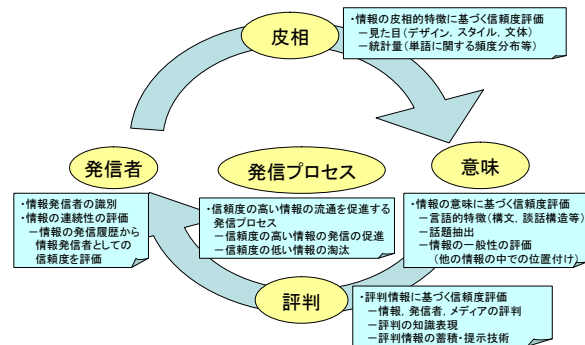


図 2: 情報の信頼性評価に関わる要素と課題

頼性を評価するための4つの視点に対応する。これらの4つの視点は前章で示した信頼性に関する様々な側面に基づいて、ウェブ上の情報の信頼性を評価する上で必要な要素としてまとめたものである。図の中央にある「発信プロセス」は、信頼性評価の問題と情報発信プロセスの関係を課題に挙げたもので、これについては、次章で詳しく述べる。以下、それぞれの視点について述べていく。

3.1 発信者に基づく情報の信頼性評価

情報の信頼性を評価する上で、情報の発信者の性質は重要な要因の一つである。発信者と情報の信頼性は、発信者への信頼から情報の信頼性が導出されるという関係にある。発信者は言うまでもなく情報を発信する主体であるが、2.1節で述べたように、主体への信頼は、能力に対する期待と意図に対する期待で構成される。

発信者の識別

情報の信頼性を考える上で、誰がその情報を発信したのかは重要な判断材料の一つである。例えば、病気のことについて、専門家である医師の発言であれば重みを持って受け止められるが、正しい医学的知識を有しているかどうか分からない一般の人の意見だと、情報の正しさについてある程度割り引いて考える必要がある。現実の世界であれば、誰が発言しているのかを確かめる術はあるが、ネット上では、それが困難になってくる。特に、CGMでは、どこの誰だか分からない人が発言をしていることが普通であり、情報発信者の信頼性を評価することが難しい。情報発信者の信頼性を評価するには、まずその前提として、誰が情報を発信しているのかを識別し、いわばネット上の人格を与えることが必要になってくる。ここで言う「識別」とは、ネット上の情報について現実世界の誰が発信したのかその身元を突き止めることとは異なる。重要なのは現実世界の身元ではなく、ネット上に複数の発言が

あった時に、どの情報が同じ発信者から発せられたものなのかを区別することである。こうした区別をして、発信者としての人格を与えることによって初めて、発信者の能力や意図といったものが扱えるようになる。

それでは、現状で発信者を識別する技術にはどのようなものがあるだろうか。ブログを例にして見てみることにする。同一のブログに掲載されている記事は同一の発信者（個人とは限らない）により発信されていると見なすことができる。これは、ブログで記事を掲載するには、ブログシステムの認証を経る必要があることによっている。ただし、あるブログの記事に対して、そのブログにアカウントを有していない第三者がつけるコメントについては発信者を識別するのにブログシステムの認証機構は利用できない。

ブログの所有者以外についても、何らかの認証を行った上でコメントを許す方法としては、2通りのアプローチが考えられる。1つは1か所で認証を管理する集中管理型のアプローチである。集中管理型の認証サービスとして代表的なのはシングル・サインオン (SSO: Single Sign-On) サービスで、分散されたネットワークサービスに対して、一つの認証サーバで一旦認証を行えば、サービス毎に認証をせずにサービスの提供を受けられるものである。Microsoft Passport Network^{*2} が挙げられる。しかし、情報の発信者の識別という目的で言えば、必ずしも認証 (authentication) が必要な訳ではなく、識別 (identification) だけで十分である^{*3}。集中管理型の識別サービスの例として、TypeKey^{*4}や、SXIP^{*5}が挙げられる。集中管理型の識別サービスを利用して、ブログにコメントを残そうとする場合、利用者はブログから一旦識別サービスのサイトへリダイレクトされ、識別サービスでの認証を経た後に、再び元のブログサイトへとリダイレクトされる。その時に一緒に識別名 (ユーザ名) や認証結果がブログサイトに送られるので、ブログ側ではその情報を利用して、ユーザを識別することになる。

サイトにアカウントを持たない情報発信者を識別するもう1つのアプローチは、分散型のアプローチである。OpenID^{*6}、mIDm^{*7}、SAML 2.0^{*8}などが挙げら

れる。ここでは、OpenID を例に分散型識別サービスの仕組みを述べる。集中管理型識別サービスの場合、識別サービスを提供する特定のサーバにリダイレクトすることにより、識別を行う。一方、OpenID では、ユーザの識別にユーザが管理権限を有しているページの URL を用いる。この URL でユーザが独自に識別サービスを運用することも可能であるし、代理サーバを指定することにより、第三者の提供する識別サービスを利用することも可能である。ここで注意する必要があるのは、あくまでも識別は URL により行われるので、識別サービスを変更しても発信者としてのアイデンティティは継続されるということである。このようにして、分散型のアプローチでは、特定の識別サービスに依存することなく、発信者としてのアイデンティティを確立することが可能になる。

ブログ以外でも発信者の識別は問題になる。会員制のサイトであれば、ユーザ名により識別が可能である。特に SNS サイトなどでは、他人の日記へのコメントなどもどのユーザが書いたのかを識別できる。非会員制の誰もが書き込める掲示板サイトなどの場合、IP アドレスを表示したり、IP アドレスに基づいて計算される固有の ID を付して、匿名性は確保しつつも、発言者の識別はできるように工夫されているところもある。これは、一人の人間が複数の人間を装う自作自演などを防ぐ目的で導入される。

このように、情報発信者を識別するのに様々な手段があるが、複数のサイトに渡って比較的汎用的に識別が可能である分散型の識別サービスの普及によって、現在では容易でないネット上での情報発信者の識別について、新たな展開が期待される。

発信者の意図に対する期待としての信頼性評価

2.1 節で述べたように、主体に対する信頼は意図に対する期待と能力に対する期待から構成される。本節では、主に前者について扱う。意図に対する期待としての信頼とは、信頼の対象となっている主体 (信頼対象) について、基本的に悪意を持っておらず、(例え可能であっても) 自分を搾取しないだろうという、人間性や行動傾向に対する期待である。

意図に対する期待としての信頼を取り扱った例として、社会ネットワークに基づいて信頼を計算する手法について紹介する。

Richardson ら [11] は、利用者の信頼関係のネットワーク (Web of Trust) を利用して、情報の信頼性を計算する手法を提案している。この研究では、個々のユーザについて、ある情報についてどれだけ信用しているかという信用度 (Belief) と、他のユーザをどれだけ信頼しているかという信頼度 (Trust) を用いて Web

*2 <https://accountservices.passport.net/ppnetworkhome.srf?lc=1041>

*3 認証と識別の違いについてはここで深く論じないが、認証は識別に加えて、識別されたユーザがどのサービスを受けられるか、そのアクセス権を管理することも含む。

*4 <http://www.sixapart.jp/typekey/>

*5 <http://www.sxip.com/>

*6 <http://openid.net/>

*7 <http://www.downes.ca/idme.htm>

*8 Security Assertion Markup Language. OASIS による認証プロトコルに関する標準。 <http://www.oasis-open.org/specs/index.php#samlv2.0>

of Trust をモデル化している。ある情報についての信用度を計算するとき、直接知っている（信頼度の値を与えている）ユーザだけでなく、直接は知らないが、信頼しているユーザが信頼しているユーザといった形で、信頼関係ネットワークの上でパスが存在するユーザ全ての信用度を考慮して計算する（経路代数アプローチ）。

Golbeck ら [4] は、人についての情報を表現するためのメタデータスキーマ FOAF (Friend-of-a-friend)^{*9} を拡張して、人に対する信頼度を 9 段階で表現している。そうして表現されたユーザ間の信頼ネットワーク上でユーザ間のパスを考慮して、最大・最小容量法、最長・最短法、加重平均法などによって、任意のユーザの信頼度を計算する手法を提案している。

これまでに述べた研究では、ユーザの他のユーザに対する信頼度というのはユーザが与えるものとして扱われてきた。しかし、近年データマイニング技術をウェブに適用して、ウェブ情報から現実の社会ネットワークを抽出しようという試みが始まっている。森らは [21] は、そのようにして抽出された社会ネットワークに対して、ネットワーク分析手法を適用することにより、ネットワーク上の 2 者間の信頼関係を推定する手法を提案している。

以上、主に社会ネットワークから主体の信頼度、更には情報の信頼度を計算する手法に関する研究について紹介した。この種の研究では、信頼を 1 変数として扱う場合が多く、意図に対する期待としての信頼を考えた場合、十分でない可能性がある。特に情報発信者としての信頼性を考える場合、情報発信者がどのような意図を持っているのか（商品を売ろうとしているのか、アフィリエイトで収入を得るために広告を出しているのか、政治的なプロパガンダが目的なのか）等々単純な 0-1 では信頼を測れない。このような意図をどのように表現し、計算に取り込むかは今後取り組むべき課題である。

発信者の能力に対する期待としての信頼性評価

発信者の信頼性を構成するもう一つの要素は、発信者の能力に対する期待としての信頼である。発信者が医者であれば、医学的知識に関しての情報は信頼できる、といったように、個々の発信者の知識、専門性、経験といったものに対する評価に基づいて、発信者にある程度の能力があることを期待し、発信される情報が信頼される。以下、発信者の情報発信履歴に基づいて発信者の信頼性を評価する研究について紹介する。

Mohan は、情報の発信履歴に基づき発信者の信頼性を評価する手法を提案している [7]。PageRank[9] な

どリンク構造に基づいてページの重要度を計る手法では、新しいけれども重要な情報については、情報が出て間もなく、リンクがまだ少ない間には評価されない。Mohan の提案する手法では、後に重要だと知られる情報を、その重要性が広く認知される前に取り上げたサイトを養育者 (nurturer) と定義し、過去の養育者としての実績を測定することで、養育者として評価の高いサイトが取り上げる情報に高い重要度を与える。この手法では、新しい情報の重要度を評価できる先見性という能力についての期待を評価していると言える。

山本らは、地域のロコミ情報サイトにおける投稿者の信頼性について、投稿者が過去にどの場所についての情報を投稿したかを指標に信頼性を評価する手法を提案している [19]。過去に情報を投稿したことのある場所の分布に基づいて、いわば投稿者の「地元」がどこであるかを推定し、新規に投稿された情報の場所と投稿者の「地元」の距離を元に、情報の信頼度を算出する。これは、「地域の情報については『地元』の人がよく知っている」という仮定に基づき、「よく知っている」という能力に対する期待としての信頼を評価したものだと言える。

3.2 情報の皮相的特徴に基づく信頼性評価

人間は情報の信頼性について、その中身を詳しく吟味せずとも、ある程度表面的な特徴から判断することが可能である。例えば、過度に派手なデザインであったり、どこから読んでいいのかわからないようなレイアウトであったり、誤字脱字が多かったり、文体や言い回しなどが不適切だったり、様々な皮相的要因からその情報には信頼性が無いと判断できる。

人間が見れば分かるような皮相的特徴に基づく信頼性を、機械が自動的に情報の信頼性を評価する意味はあるだろうか？この場合、一つ一つのページについて、ユーザに信頼性の有無を提示する意義はあまりない。むしろ、大量の情報の中から信頼性の高い情報を探そうとするような場合にこのような技術は重要になってくる。つまり、利用者が全ての情報を一々見るのが不可能な時に、一次的なフィルタリングに皮相的特徴による信頼評価を用いるのである。皮相的特徴による信頼性評価は、フィルタリングのための特徴量の一つとして、他の評価手法とは異なる観点を与えるという意味で、有効であると期待される。

Fogg ら [3] は人がどのようなウェブサイト信頼性があると判断するのかについて、大規模な調査を行っている。Fogg らは信頼性 (Credibility) を情報そのものに備わる性質ではなく、情報を受け取る利用者が「知覚する品質 (perceived quality)」であると定義している。その上で、ウェブサイトの様々な要素がウェブサ

*9 <http://www.foaf-project.org/>

イトの利用者が知覚する信頼性にどのように影響を与えるのかを調査している。この調査では、300の調査項目から初めて、複数回の予備調査を経て信頼性に関して重要なものとして51項目に絞り込んでいる。調査の後、各項目の相関分析に基づき以下の7つの尺度に分類している。

1. 現実世界感 (Real-World Feel Scale)
2. 使いやすさ (Ease of Use Scale)
3. 専門度 (Expertise Scale)
4. 信頼度 (Trustworthiness Scale)
5. オーダーメイド度 (Tailoring Scale)
6. 商業的度合い (Commercial Implications)
7. アマチュア度 (The Ammateurism Scale)

各尺度に含まれる項目としていくつか例をあげると、現実世界感には「問い合わせにはすぐに回答がくる」や「組織の所在地を明記している」などの項目が、信頼度には「信用できるサイトからリンクされている」や「情報の内容について方針を明記している」などの項目が含まれる。この調査の結果は、ウェブサイトのデザイナーに対して信頼されるサイトを作成するためのガイドライン [13] としてまとめられている。

続いて、皮相的な特徴量を利用してウェブページの信頼性を評価する手法についての研究を紹介する。Velayathan[16]らは、情報の信頼度を評価する上で、ページのトポロジー、ドメイン名、ページのデザインや表現、アクセス時間、サーバの情報などを利用することを提案している。福島ら [23] は、Fogg らの調査結果に基づいて、各調査項目に対応する自動処理方法を実現、組み合わせることによりページの信頼性を自動判定する手法を提案している。

本節の冒頭で述べたように、皮相的特徴に基づく信頼性評価は大量の情報を選別するための一次フィルターとして有効であると期待されるが、この種の研究はまだ十分に手がつけられておらず、今後の展開が期待される。

3.3 情報の意味内容に基づく信頼性評価

情報の意味内容に基づく信頼性評価は、本稿で挙げた4つの信頼性評価の要素のうち、従来あまり研究がなされてこなかった分野である。その原因としては、そもそも情報の意味内容を扱うのが困難であることが挙げられる。しかし、近年自然言語処理技術の発展と共に、徐々に情報の意味内容にまで踏み込んだ研究が

増えつつある。例えば、blogを対象として商品に関する肯定意見や否定意見といった評判情報を抽出する研究が盛んになりつつある [18]。

意味内容に基づいて信頼性を評価するという研究はあまりないので、ここでは、いくつかのアイデアを示すにとどめておく。

意味の整合性 情報の意味内容を抽出することができれば、その整合性を検証することにより、信頼性を評価することが可能である。例えば、同じ文章の中で矛盾した主張をしていたり、信頼のおける第三者の情報と矛盾していたりすれば、その情報の信頼性が疑われる。

話題抽出 与えられた情報がどのような話題についての情報かを識別できれば、様々な形で信頼性を評価することが可能になる。例えば、多くの情報をクラスタリングすることにより、個々の情報の全体の中で位置づけることにより、利用者の判断を助けることができる。また、同じ話題について述べている情報でも、より多くの論点を含む情報の方が、複数の視点を持って書かれているという点で信頼できる可能性がある。更には、発信者との関連において発信者の専門と情報が扱っている話題の比較により、その信頼性を評価することもできる。

談話解析 文章の談話構造を解析することにより、主張が説得力のある形で記述されているかどうかを判断する。

3.4 情報の評判に基づく信頼性評価

評判とは、あることがらについて人々が持っている意見のことを指す。つまり、ここでいう「情報の評判に基づく信頼性評価」とは、ある情報について人々が持っている意見を利用して、情報の信頼性を評価するための方法ということになる。

ウェブ上の情報の特徴は、ハイパーリンクにより情報（ページ）がつながって、ネットワーク構造をなしていることにある。このハイパーリンクを一種のページについての評判だと捉え、ページの重要度を計算するのに利用したのが Google*¹⁰の PageRank[9] アルゴリズムである。PageRank アルゴリズムでは、重要度の高いページからリンクされているページの重要度は高いというアイデアに基づいて、ページの重要度を算出している。Google では PageRank の結果に基づき、

*10 <http://www.google.com/>

検索結果のランキングを行っている。PageRank で計算されるページの重要度は、評判に基づいた信頼性評価の一種だと言える。

他に情報の評判を利用する方法として、直接ユーザから情報についての評判を集めて利用することが挙げられる。Slashdot^{*11}では、モデレーションシステム (moderation system) によりユーザからの評判を利用してコメントの評価およびフィルタリングを行っている。Slashdot は様々な (コンピュータ技術関係に偏っているがそれに限らない) ニュースについて、掲示板で自由に議論ができるサイトである。モデレーションシステムでは、モデレータ (moderator) に選ばれたユーザが、個々のコメントについて「参考になる (informative)」「すばらしい洞察 (insightful)」「おもしろおかしい (funny)」「フレームの餌 (flamebait)」などの肯定的/否定的評価を行い、それに応じてコメントにポイントが加算/減算される仕組みになっている。一般のユーザは、指定した閾値以上のポイントのコメントのみを表示することができるので、モデレータの評価に基づいてコメントをフィルタリングすることが可能である。ユーザから情報についての評判を集めてランキング等に利用するサイトとしては、他に digg^{*12}、reddit^{*13} などがある。また、投票によるブログのランキングを行うサイト^{*14}もある。

次に、評判に基づいて信頼性を評価するときを考慮すべき問題について検討する。情報についての信頼性ではないが、ネット上で一般的になっている評判を利用した仕組みとして、オークションサイトにおける出品者/落札者の評価システムが挙げられる。例えば、Yahoo!オークション^{*15}では、取引終了後に、出品者と落札者がお互いに相手を「非常に良い」「良い」「どちらでもない」「悪い」「非常に悪い」の5段階で評価を行う。Resnickら [10] このようなオークションサイトにおける評価システムの問題として、ユーザが提供する評判について、肯定的な評価に大きく偏っていることを指摘している。その理由の一つとして、オークションサイトにおける評価システムが相互的であり、仕返し (悪い評価を与えると、その仕返しとして悪い評価を与えられる) を恐れて、よほどのことがない限り肯定的な評価を残すためであると説明している。Traupmanら [15] は、この評価の偏りのために「普通の参加者」と「極めて素晴らしい参加者」の区別がで

きないという問題を指摘し、EM法により参加者の信頼性を推定する手法を提案している。今後、情報について評判を情報の信頼性を評価するための基盤として利用することを考えるときに、オークションサイトにおけるのと同様なユーザ同士の社会的な要因の評判へ影響を考慮する必要が出てくるであろう。

評判に基づいて情報の信頼性を評価するときの問題として、他に評判情報の表現力が挙げられる。これまで見てきた評判情報はおおそ肯定か否定という一つの対立軸で表現されている。Slashdot の場合は多少形容詞の多様性はあるものの、最終的に集計するときには肯定的な形容詞であれば+1点、否定的な形容詞であれば-1点として計算している。しかし、情報の信頼性を考慮する際、本稿でも4つの要素を挙げているように、本来、人であれば様々な要素を勘案して、最終的な判断をしているはずである。情報に対する評判を表現する段階で肯定-否定の軸に位置付けることは、途中の過程での様々な判断に関する情報が失われてしまっていることを意味する。結果、集約された評判も、評価対象の情報について正確に反映しなくなる可能性がある。事実、先に述べたオークションサイトにおける評価システムではそのようなことが起きている。肯定-否定にとどまらず、より豊かな表現力をもった評判情報により、より精密に情報の信頼性を評価することが期待される。ただし、表現力が豊かになるということは、表現が複雑になることを意味しておりユーザの負担が増えると共に、集めた評判情報を集約してユーザに提示することも難しくなるという問題も抱えていることに留意せねばならない。

近年、「群衆の叡智 (Wisdom of Crowds)」 [14] という考え方が注目を集めている。これは、特定の条件を整えば、あることがらについて多数の「普通」の人の判断を集約することにより、少数の優秀な人に勝る意思決定ができるというものである。一方、ウェブの世界では *Web 2.0* [8] という言葉で代表される、従来のウェブのように一方的な情報発信にはとどまらない、ウェブを基盤としたユーザ同士の連携、サービス間の連携が盛んになりつつある。ユーザ同士の連携の例としては、ブックマーク共有サイト del.icio.us^{*16}、写真共有サイト Flickr^{*17}、各種ソーシャルネットワークサービス^{*18}が広く普及するに至っている。このような状況は、ウェブ上で群衆の叡智を発揮する仕組みを作るための素地は十分に整っていることを物語っている。

*11 <http://slashdot.org/> (日本語版, <http://slashdot.jp/>)

*12 <http://digg.com/>

*13 <http://reddit.com/>

(日本語版, <http://ja.reddit.com/>)

*14 Blog Ranking (<http://blog.with2.net/>) など

*15 <http://auctions.yahoo.co.jp/>

*16 <http://del.icio.us/>

*17 <http://www.flickr.com/>

*18 海外では Friendster (<http://www.friendster.com/>), Orkut.com (<http://www.orkut.com/>) など。国内では mixi (<http://mixi.jp/>) など

例えば, Mika[6] は, del.icio.us においてユーザがブックマークにつけるタグの使用状況を分析することにより, いわばコミュニティのオントロジーと呼べるものが抽出できる可能性を示している. このように, ユーザから表現力の豊かな評判情報を収集し, 情報の信頼性評価を行うことは十分に可能であると考えられる.

4 おわりに

本稿では, 主にウェブから得られる情報コンテンツの信頼性評価技術について議論をした. まず, 信頼性の概念について整理をおこない, 信頼性を評価する4つの視点を提案し, 関連技術を紹介した.

ウェブが普及して以来, 大量の情報の中から必要な情報をいかに取り出すかという問題は常に語られてきたが, 情報の信頼性について注目されるようになったのは比較的最近である. 今後, 様々な技術が開発され, 信頼性の高い情報の発信を促す情報発信プロセスへの実現に結びつくことを期待するものである.

参考文献

- [1] Rino Falcone and Cristiano Castelfranchi. A belief-based model of trust. In *Trust in Knowledge Management and Systems in Organizations*, chapter XI, pp. 306–343. Idea Group Publishing, 2004.
- [2] B. J. Fogg and Hsiang Tseng. The elements of computer credibility. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 80–87. ACM Press, 1999.
- [3] BJ Fogg, Jonathan Marshall, Othman Laraki, Alex Osipovich, Chris Varma, Nicholas Fang, Jyoti Paul, Akshay Rangnekar, John Shon, Preeti Swani, and Marissa Treinen. What makes web sites credible?: A report on a large quantitative study. In *SIGCHI'01*, pp. 61–68, 2001.
- [4] Jennifer Golbeck, Bijan Parsia, and James Hendler. Trust networks on the semantic web. In *Proceedings of Cooperative Intelligent Agents 2003*, 2003.
- [5] Niklas Luhmann. 信頼—社会的な複雑性の縮減メカニズム. 大庭健, 正村俊之 (訳) 勁草書房, 1990. (Ferdinand Enke Verlag, 2nd edition, 1973).
- [6] Peter Mika. Ontologies are us: A unified model of social networks and semantics. In Yolanda Gil, Enrico Motta, V. Richard Benjamins, and Mark A. Musen, editors, *The Semantic Web – ISWC 2005: 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10, 2005. Proceedings*, Vol. 3729 of *Lecture Notes in Computer Science*, pp. 522–536, 2005.
- [7] Bharath Kumar Mohan. Searching association networks for nurturers. *IEEE Computer*, Vol. 38, No. 10, pp. 54–60, 2005.
- [8] Tim O'reilly. What is web 2.0: Design patterns and business models for the next generation of software. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>, September 2005.
- [9] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical report, Stanford University, 1999. [Available at <http://dbpubs.stanford.edu/pub/1999-66>].
- [10] Paul Resnick and Richard Zeckhauser. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. In Michael R. Baye, editor, *The Economics of the Internet and E-Commerce*, Vol. 11 of *Advances in Applied Microeconomics*. Elsevier Science, Amsterdam, 2002.
- [11] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. Trust management for the semantic web. In D. Fensel et al., editor, *ISWC 2003*, Vol. 2870 of *LNCS*, pp. 351–368, Berlin Heidelberg, 2003. Springer-Verlag.
- [12] Robert C. Solomon and Fernando Flores. 「信頼」の研究—全てのビジネスは信頼から. 上野正安 (訳) シュプリンガー・フェアラーク東京, 2004. (*Building Trust*. Oxford University Press, 2001).
- [13] Stanford Persuasive Technology Lab. Stanford guidelines for web credibility. <http://www.webcredibility.org/guidelines/>, 2002.
- [14] James Surowiecki. *The Wisdom of Crowds: Why the Many Are Smarter than the Few*

and How Collective Wisdom Shapes Business, Economies, Societies, and Nations. Doubleday, 2004.

- [15] Jonathan Traupman and Robert Wilensky. Robust reputations for peer-to-peer marketplaces. In Ketil Stølen, William H. Winsborough, Fabio Martinelli, and Fabio Massacci, editors, *Trust Management: 4th International Conference, iTrust 2006, Pisa, Italy, May 2006. Proceedings*, Vol. 3986 of *Lecture Notes in Computer Science*, pp. 382–396, 2006.
- [16] Ganethan Velayathan, 山田誠二. Web ページの相対信頼度. 2004 年度人工知能学会全国大会 (第 18 回) 論文集, pp. 3F1–05, 2004.
- [17] 武田英明. 信頼の Web. コンピュータソフトウェア, Vol. 22, No. 4, pp. 19–25, 2005.
- [18] 奥村学. blog マイニング: インターネット上のトレンド, 意見分析を目指して. 人工知能学会誌, Vol. 21, No. 4, pp. 424–429, 2006.
- [19] 山本浩司, 片上大輔, 新田克己, 相場亮, 桑田仁. 地図上の投稿情報の信頼度. 2006 年度人工知能学会全国大会 (第 20 回) 論文集, pp. 3G1–3, 2006.
- [20] 山岸俊男. 信頼の構造-こころと社会の進化ゲーム. 東京大学出版会, 1998.
- [21] 森純一郎, 武田英明, 石塚満. 信頼の構造: 社会ネットワークの構造に基づく Trust モデル. 2006 年度人工知能学会全国大会 (第 20 回) 論文集, pp. 3D4–5, 2006.
- [22] 総務省. 平成 18 年度版情報通信に関する現状報告 (情報通信白書), 2006.
- [23] 福島隆寛, 内海彰. Web ページの信頼性の自動推定. 2005 年度人工知能学会全国大会 (第 19 回) 論文集, pp. 3C2–03, 2005.