

# 脚注と参考文献を用いた Wikipedia 記事の信頼性評価の一手法

井上 雄介<sup>†</sup> 太田 学<sup>††</sup>

<sup>†</sup> 岡山大学工学部情報工学科 〒700-8530 岡山県岡山市北区津島中 3-1-1

<sup>††</sup> 岡山大学大学院自然科学研究科 〒700-8530 岡山県岡山市北区津島中 3-1-1

E-mail: <sup>†</sup>inouey@de.cs.okayama-u.ac.jp, <sup>††</sup>ohta@de.cs.okayama-u.ac.jp

あらまし 本稿では、脚注と参考文献を用いた Wikipedia 記事の信頼性評価の一手法を提案する。信頼性の高い記事には、多くの脚注と参考文献があり、さらにそれらは信頼性の高い情報源をリンク先とするものが多いと考えられる。そこで本稿では、Wikipedia 記事の一つの特徴である脚注と参考文献の数や種類、リンク先の信頼性を用いて記事の信頼性を評価する。日本語版 Wikipedia で公開されている「秀逸な記事」と「良質な記事」を信頼性の高い記事とし、それらとその他の記事の分類実験を行い、その分類精度を求めた。

キーワード 信頼性, Wikipedia, 脚注, 参考文献

## A Reliability Evaluation Method for Wikipedia Articles Using Footnotes and References

Yusuke INOUE<sup>†</sup> and Manabu OHTA<sup>††</sup>

<sup>†</sup> Department of Information Technology, Faculty of Engineering, Okayama University  
3-1-1 Tsushimanaka, Kita-ku, Okayama-shi, Okayama, 700-8530 Japan

<sup>††</sup> Graduate School of Natural Science and Technology, Okayama University  
3-1-1 Tsushimanaka, Kita-ku, Okayama-shi, Okayama, 700-8530 Japan

E-mail: <sup>†</sup>inouey@de.cs.okayama-u.ac.jp, <sup>††</sup>ohta@de.cs.okayama-u.ac.jp

**Abstract** This paper proposes a reliability evaluation method for Wikipedia articles using their footnotes and references. We assume that reliable articles have and should have sufficient footnotes and references that have links to reliable information sources. Therefore, the proposed method evaluates the reliability of Wikipedia articles by the reliability of linked sources in addition to the numbers and kinds of their footnotes and references. For experiment, we collected “ featured ” and “ good ” articles from the Japanese Wikipedia article set as well as the other articles. We classified the collected articles into three classes according to their reliability and showed the classification accuracy.

**Key words** Reliability, Wikipedia, Footnote, Reference

### 1. はじめに

最近、様々な場面で用いられることが多くなっている Wikipedia は誰でも編集可能なオンライン百科事典である。しかし、誰でも編集可能であるため全ての記事が正確で信頼できる情報であるとは限らない。つまり、ある編集者が誤った情報を載せた場合でも、その情報が削除されるまで誤った情報が多くの閲覧者に示されることになり、Wikipedia の影響力を考えるとこれは無視できない。このような場合、各閲覧者が正しい記述だけを見極めることは困難で、誤った記述を正しい記述であると誤って認識してしまうことも少なくない。現在でも信頼性の低い情報や不適切な情報は、他の多くの編集者による削除や編集によって訂正の努力が続けられている。しかし、現在の Wikipedia のように規

模が大きくなると、全ての記事に対して信頼性を保つことが難しくなると考えられている [10]。そこで、各記事の信頼性を示すことができれば、多くの閲覧者の知識獲得の手助けになると考えた。田中ら [1, 2] は、「ページの品質が高いほどそこで言及されている知識の信頼性も高い」と仮定しており、PageRank に代表されるリンク解析 [3] が、Web ページの信頼性を判断する際にも有用な手段であるとしている。本研究では、信頼性を多くの人からの支持を得ていることと定義する。

本稿の構成は以下の通りである。2 節で関連研究について述べ、3 節で提案する信頼性評価手法について説明し、4 節で評価実験と考察、5 節でまとめについてそれぞれ述べる。

## 2. 関連研究

Wikipedia 記事の信頼性評価に関しては、既にいくつかの方法が提案されている。本節では、信頼性を人手で評価する方法と、編集履歴を用いて自動で評価する方法、リンク解析による自動評価方法を紹介する。

### 2.1 人手による信頼性評価

Kramer[4]らは利用者の投票を利用して、Wikipedia 記事の信頼性を算出している。彼らは、MediaWiki に利用者による記事への投票システムを付加している。この方法は利用者が記事の信頼性を判断するので、その信頼性は、我々の定義する信頼性と同じであると考えられる。さらに全て人手で評価するため十分な数の利用者が投票すれば高精度の信頼性評価が期待できる。

しかし、Wikipedia が大規模になると、全ての記事の信頼性を評価するには非常に多くの利用者の判断が必要になる。また全ての利用者が正確に各記事の信頼性を判断することも困難である。一方提案手法ではこのような利用者の判断を必要としない。

### 2.2 編集履歴を用いた信頼性評価

Adler ら [5, 6, 7] や Hu ら [8] は Wikipedia の編集履歴を用いることで記事の信頼性を評価している。記事の残存率から著者の信頼性を求め、記事の中で信頼性の高い著者の書いた部分ほど信頼性が高いとしている。Adler らはその手法を実装し、WikiTrustModule として公開している。さらに、鈴木ら [10] は評価の対象とする著者を編集量の多い著者だけに限定した手法を提案しており、計算量の削減と精度の向上を実現している。

これらの手法は Wikipedia が大規模になり記事の量が増えると、より正確に著者の評価ができるようになり、その結果、記事についてもより精度の高い信頼性評価が可能になると考えられる。しかし、著者同士の編集合戦になった場合は両者の記事の残存率が下がり、著者の信頼性が低下してしまう。我々の提案手法では、信頼性評価に編集履歴を用いないので、このような問題は発生しない。

### 2.3 リンク解析による信頼性評価

McGuinness ら [9] は、Wikipedia のリンク構造を利用した信頼性評価を行っている。彼らの研究では他の Wikipedia 記事からより多くのリンクを獲得している記事ほど信頼性が高いとしている。しかし、Wikipedia におけるリンクは通常の Web ページとは異なり、記事中の単語に関する記事があれば積極的にリンクを生成する構造なので、リンクの数がそのまま記事の信頼性につながるとは限らない。

## 3. 提案手法

本研究では、Wikipedia の「脚注」と「参考文献」を用いて記事の信頼度を算出する手法を提案する。参考文献には編集者が著述する際に参考にした図書や文献の書誌事項が示されている。脚注には記事中の該当部分の情報源や補足説明が書かれている。そこで、参考文献や脚注が多く含まれている記事は、内容の根拠となる情報が示されているので、信頼性が高いと考える。さらに、脚注の種類や脚注の情報源の信頼性を判断し、それによって提案する記事の信頼度への反映方法を変える。

以下で本研究で利用する参考文献と脚注について述べ、それらの信頼度への反映方法について説明する。

### 3.1 参考文献

Wikipedia 記事の参考文献には、通常書籍が挙げられている。本研究では、各参考文献の信頼性は同一であるとみなして、参考文献については数のみを信頼度に反映させる。つまり、参考文献の数が多ければ記事の信頼度を高くする。

### 3.2 脚注

脚注には書籍名、URL、ニュース記事のタイトル、新聞記事のタイトル、補足説明など様々な情報が書き込まれている。そこでまずどの種類の情報であるかを判断し、その種類や情報源の信頼性によって信頼度への反映方法を変える。以下で本研究における脚注の分類について説明する。

#### (1) 書籍名

書籍名が挙げられている場合は参考文献と同様に記事の信頼度に反映する。本研究では、ISBN やページが示されている脚注を書籍名と判断する。

#### (2) URL

URL の場合は信頼するドメインとその他に分けて、信頼度への反映方法を変える。信頼するドメインの種類と例を以下に示す

- 政府系機関

例) go.jp, gov

- 教育機関

例) ac.jp, edu

- 新聞社

例) asahi.com mainichi.jp

#### (3) 新聞記事

URL が付されていなくても、信頼する新聞社の新聞記事の場合は、全て信頼性が高い情報源であると考えられる。

#### (4) ニュース記事

ニュース記事は、リンクがあれば (1) に示した URL の分類に従う。リンクがなければ、ニュース記事の信頼性は不明とする。

#### (5) その他

その他の場合は、記事の信頼性判断の根拠としない。

### 3.3 記事の信頼度

3.1 節と 3.2 節で説明した、記事の信頼性を求めるための手かかりを 3 種類に分類し、それぞれの数を  $N_1, N_2, N_3$  とする。そして、記事の信頼度  $R$  の算出法として式 (1)~(8) の 8 種類を提案する。

$$R = N_1 \quad (1)$$

$$R = N_2 \quad (2)$$

$$R = N_3 \quad (3)$$

$$R = N_1 + N_2 \quad (4)$$

$$R = 2 * N_1 + N_2 \quad (5)$$

$$R = N_1 + N_2 + N_3 \quad (6)$$

$$R = 2 * N_1 + N_2 + N_3 \quad (7)$$

表1 評価実験用データ

記事の分類 (信頼性)	本数
秀逸な記事 ( )	87
良質な記事 ( )	200
その他の記事 ( )	300

$$R = 4 * N_1 + 2 * N_2 + N_3 \quad (8)$$

上記の各式において、

$N_1$ : 参考文献, 書籍名の脚注の数

$N_2$ : 信頼する URL(ニュース記事を含む), 新聞記事の脚注の数

$N_3$ : 信頼性の不明な URL, リンクのないニュース記事の脚注の数

である。式 (1), (2), (3) はそれぞれ  $N_1, N_2, N_3$  だけを信頼度の算出に利用した場合の算出法で、式 (4)~(8) は  $N_1 \sim N_3$  を組み合わせた算出法である。

- 式 (4), (5) は  $N_3$  に含まれる脚注を、信頼度の算出に利用しない。
- 式 (4), (6) は、脚注と参考文献の数のみを考慮して信頼度を算出する。
- 式 (5), (7), (8) は、種類によってこれらの手がかりの信頼度が異なると考え、それを重みとして記事の信頼度計算に反映させる。具体的には、参考文献等に挙げられる書籍の重みを大きくする。

なお記事の信頼度を下げるような手がかりについては考慮していない。

## 4. 評価実験と考察

本節では、定義した信頼度を利用して、Wikipedia 記事の分類実験を行う。実験では、記事をその信頼性に応じて、 $\alpha, \beta, \gamma$  の3段階に分類する。以下ではまず評価実験用データについて説明し、続いて本稿で定義した信頼度の閾値の設定、さらにそれを利用した記事の分類実験について述べる。

### 4.1 評価実験用データ

本研究では、信頼性が高いということ、多くの人から支持されていることと同義と考え、実験では日本語版 Wikipedia で公開されている「秀逸な記事」と「良質な記事」を信頼性の高い記事と仮定し、「削除依頼」と「修正依頼」が出されている記事を信頼性の低い記事と仮定した。しかし、「削除依頼」や「修正依頼」が出されている記事の数は、全記事数の約 0.1% しか存在しないため、以下の評価実験では対象としない。

また評価実験用データとしては、表 1 に示す記事を Wikipedia データベースから収集した。表 1 において、「秀逸な記事」を最も信頼性の高い記事の集合とし、「良質な記事」をその次に信頼性の高い記事の集合とする。「その他の記事」はそのどちらにも含まれない記事の集合とする。

### 4.2 信頼度の閾値

各式ごとの信頼度の閾値を「秀逸な記事」、「良質な記事」、「その他の記事」の信頼度の分布から求める。Wikipedia の記事集合における 3 種類の記事の割合は同一ではないので、まず全ての Wikipedia 記事の中に含まれるそれぞれの割合を求める。

Wikipedia で公開されている「秀逸な記事」と「良質な記事」の数は分かっているが、そう明示されていない記事の中にも、それらに相当する記事が含まれていると考えられる。そこで、表 1 の評価実験用データとして収集した記事とは別に Wikipedia 記事の中から、無作為に抽出した 300 件の記事を「秀逸な記事」と「良質な記事」の定義に照らして人手で分類した。その結果、6% が「秀逸な記事」、14% が「良質な記事」だった。

そこで、表 1 の評価実験用データの中から「秀逸な記事」を 12 件、「良質な記事」を 28 件、「その他の記事」を 160 件選び、その計 200 件の記事の信頼度の分布について調べた。図 1 に式 (1) により算出した各記事の信頼度の分布を、信頼度のカテゴリ別に正規分布で近似したグラフを示す。図 2 には図 1 を拡大したグラフを示す。また、式 (1)~(8) の信頼度の分布から決定したその閾値を表 2 に示す。

### 4.3 評価実験

表 1 に示した評価実験用データの中から、閾値の決定に利用した記事以外の記事を 150 件選んで分類対象とするデータセットを作成した。これは、4.2 節で閾値の決定に利用したデータセットと同様に、6% の「秀逸な記事」と 14% の「良質な記事」、80% の「その他の記事」で構成している。これらの記事を 4.2 節で決定した閾値を用いて分類し、その分類精度を計算した結果を表 3 に示す。

表 3 から、式 (7) が最も良い結果を示していることが分かる。式 (1)~(3) と式 (4)~(8) の比較から、参考文献と脚注を組み合わせると信頼度を算出することは精度の向上につながる。また、式 (4) と式 (5)、式 (6) と式 (7) を比較すると、参考文献や書籍名を表す脚注の重みを、他の脚注の重みよりも大きくする方が良い結果となった。さらに、式 (4) と式 (6)、式 (5) と式 (7) を比較すると、信頼性が不明とした URL やリンクのないニュース記事の脚注でも、信頼度の算出に考慮した方が良い結果となった。

式 (7) の実験結果で分類に失敗した記事 6 件を分析した。Wikipedia の「秀逸な記事」の脚注の中で、書籍のタイトルだけを示した脚注を、書籍名として数えていなかった。そのため、このような記事の信頼度が下がり、誤って「良質な記事」のカテゴリである  $\beta$  に分類されたものが 2 件あった。さらに「良質な記事」の中に、脚注や参考文献が少ないために、式 (7) の信頼度が低くなり  $\beta$  に分類された記事が 2 件あった。逆に脚注や参考文献が多いために、信頼度が高くなり  $\alpha$  に分類された記事が 1 件あった。そして、「その他の記事」の中に、脚注や参考文献が多いために、式 (7) の信頼度が高くなり、誤って  $\alpha$  に分類された記事が 1 件あった。これらの記事の分類方法については今後検討したい。

## 5. まとめ

本稿では脚注と参考文献を用いた簡便な Wikipedia 記事の信頼性評価手法を提案した。提案手法は脚注と参考文献の数、リンク先の信頼性などを記事の信頼度算出に利用する。評価実験用データとして、日本語版 Wikipedia で公開されている「秀逸な

表2 各式の閾値

	式(1)	式(2)	式(3)	式(4)	式(5)	式(6)	式(7)	式(8)
との閾値	37	15	27	50	84	65	86	175
との閾値	9	2	5	10	19	11	20	39

表3 実験結果

	式(1)	式(2)	式(3)	式(4)	式(5)	式(6)	式(7)	式(8)
分類精度	0.927	0.887	0.833	0.947	0.953	0.953	0.960	0.947

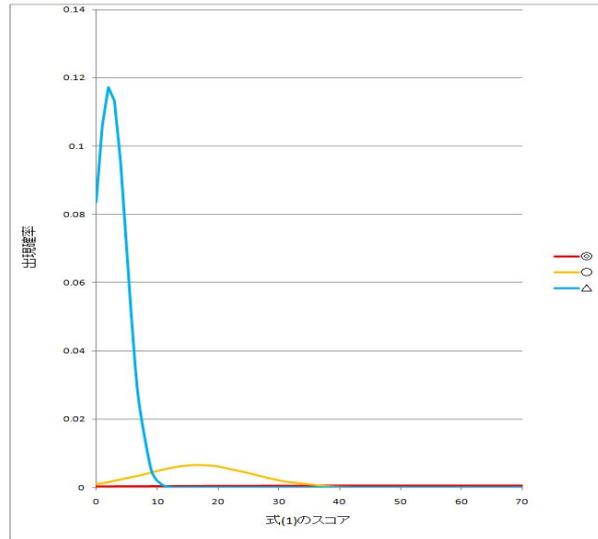


図1 式(1)の信頼度の分布

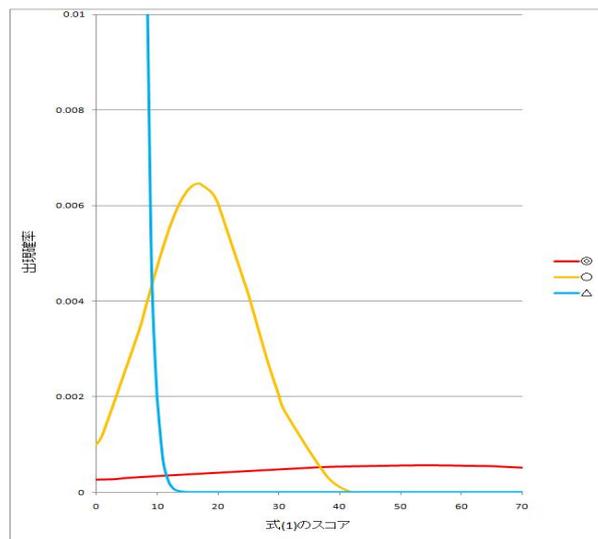


図2 式(1)の信頼度の分布(拡大)

記事」と「良質な記事」を信頼性の高い記事として、「その他の記事」とともに収集した。実験では、評価実験用データを用いて決定した信頼度の閾値に基づいて、記事を信頼性に応じて3段階に分類し、分類精度を評価した。その結果、参考文献と脚注を組み合わせると信頼性評価を行うことの一定の有効性を確認できた。また実験では、必ずしも信頼性の明らかでない脚注でも、ないよりはましという結果が得られた。また、参考文献や書籍名を表す脚注は、他の種類の脚注よりも、その記事の信頼性を高めているという予想に矛盾しない実験結果となった。

今後は、評価実験の結果を基に提案手法の改良を行い精度向上について検討する。提案手法では、脚注のリンク先の信頼性をドメイン名のみで判断し、実際にそのリンク先が存在しているかどうかを確認していない。また、参考文献や書籍名の脚注に関しても、実際にその書籍が存在しているかの確認をしていないため、タイトルやISBNから確認できないか検討したい。さらに、脚注や参考文献以外の記事の信頼性評価につながる手がかりについて検討したい。これらを提案手法と組み合わせることで精度を向上させることができると考えている。また関連研究で挙げた編集履歴を用いる手法との比較など、客観的な評価も行いたい。

- [10] 鈴木 優, 吉川 正俊, “ Wikipedia におけるキーパーソン抽出による信頼度算出精度および速度の改善 ”, 人工知能学会 第 21 回セマンティックウェブとオントロジー研究会 SIG-SWO-A901-1, 2009.
- [11] 鈴木 優, 金本 径卓, 川越 恭二, “ Wikipedia の編集履歴に基づく記事の信頼性導出 ”, 人工知能学会 第 20 回セマンティックウェブとオントロジー研究会 SIG-SWO-A803-09, 2008.

## 文 献

- [1] 山本 祐輔, 手塚 太郎, アダム ヤフト, 田中 克己, “ ページ特性を考慮した Web 検索結果の集約とページ生成時間分析による知識の信頼性判断支援 ”, 電子情報通信学会論文誌 Vol.89, No.1, pp.1-9, 2006.
- [2] 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 ”, DEWS2004 I-2-05, 2004.
- [3] L. Page, S. Brin, R. Motwani and T. Winograd, “ The PageRank Citation Ranking: Bringing Order to the Web ”, Stanford Digital Libraries Working Paper, 1998.
- [4] Mark Kramer, Andy Gregorowics, and Bala Iyer, “ Wiki trust metrics based on phrasal analysis ”, WikiSym '08: International Symposium on Wikis, ACM, 2008.
- [5] B.Thomas Adler and Luca de Alfaro, “ A contentdriven reputation system for the Wikipedia ”, WWW '07: Proceedings of the 16th international conference on WorldWideWeb, pp.261-270, 2007
- [6] B. Thomas Adler, Krishnendu Chatterjee, Luca de Alfaro, Marco Faella, Ian Pye, and Vishwanath Raman, “ Assigning trust to Wikipedia content ”, WikiSym '08: International Symposium on Wikis, 2008.
- [7] B. Thomas Adler, Ian Pye, and Vishwanath Raman. “ Measuring author contributions to the Wikipedia ”, WikiSym '08: International Symposium on Wikis, 2008.
- [8] Meiqun Hu, Ee-Peng Lim, Aixin Sun, Hady Wirawan Lauw, and Ba-Quy Vung, “ Measuring article quality in Wikipedia: models and evaluation ”, CIKM, pp.243-252, 2007.
- [9] Deborah McGuinness, Honglei Zeng, Paulo Pinheiro da Silva, Li Ding, Dhyanesh Narayanan, and Mayukh Bhaowal. “ Investigation into trust for collaborative information repositories: A Wikipedia case study ”, The Workshop on the Models of Trust for the Web(MTW '06), 2006.