

## 平成25年度 情報工学コース卒業研究報告要旨

石川 研究室	氏 名	近 藤 弘 隆
卒業研究題目	Wikipediaのノートページにおける 編集者の重要度算出手法	
<p>Webの発達により、一般の利用者によって作成されるコンテンツが増加した。それらのコンテンツはUGC( User Generated Content) と呼ばれ、その情報の豊富さ、幅広さなどの理由により重要視されている。</p> <p>巨大で有用なUGCの一つとしてWikipediaが存在する。Wikipediaは誰でも編集が可能なWeb上の百科辞典である。Wikipediaには記事の質や信頼性、内容の方向性を決定する機関は存在しない。編集者同士が編集、議論、合意を繰り返して中立的な観点で書かれた記事を作成していくことになっている。誰でも編集が可能であるため、異なる意見を持った複数の編集者が存在する。異なる意見を持った編集者が記事を編集するため、意見が対立することがある。意見が対立した場合、Wikipediaでは編集者同士が対話を行い、合意を得て編集を行う必要がある。中立的な観点から書かれた質の高い記事を作成していくためには、うまく合意を形成することが求められる。そのためには、対立の原因となっている話題、支持されている意見を知る必要があると考えられる。また、支持されている編集者が分かれば、その編集者の意見から支持されている意見、そして対立の原因となっている話題を把握することができると考えられる。</p> <p>そこで本研究では、Wikipediaのノートページにおける編集者の支持されている度合いを重要度とし、その重要度を算出する手法を提案する。ノートページとは、Wikipediaの記事に自動的に付随する、議論や会話のために用いられるページである。提案手法では、ある記事における編集者の重要度を以下の手法で算出する。ノートページにおいて意見が支持された編集者は支持された内容を本文に反映させると考えられる。よって、発言内容を本文に反映させる編集を取得する。また、発言内容が重要であるとき、つまり他の編集者に支持された発言内容であるとき、発言内容が本文に反映されており、その反映された本文は他の編集者から削除されないと考えられる。よって、編集内容が残っていれば、その編集のもとになった発言は支持されているといえる。そのため、発言を反映させる編集が残っているかどうかを求める。また、発言を反映させる編集において、発言を反映させるために重要な単語について考える。ある記事の編集において定常的に使われる単語が発言を反映させる編集に用いられていた場合、その単語は発言を反映させるために重要であるとは考えにくい。逆に、ある発言を反映させる編集においてのみ使用される単語はその発言を反映させるために必要な単語であると考えられる。よって、発言を反映させるために重要な単語とは、その発言を反映させる編集に多く使われ、その他の編集に定常的には使用されない単語であると考えられる。発言を反映させる編集における重要な単語が残されたか、元に戻されたかという情報は、発言が支持されているかどうか大きく影響すると考えられる。よって、発言に対する編集によって追加、削除された単語が取り消されていないか、その単語は特徴的かという情報を用いて、発言の重みを算出する。そして、その発言の重みを用いて編集者の重要度を算出する。</p> <p>また、評価実験では提案手法によって求めた重要度が実際の編集者の支持、不支持を表すかどうかを確かめた。具体的には、記事ごとに、算出された重要度と人手で求めた実際の支持、不支持に対しピアマンの順位相関係数を求め、相関関係があるのかどうかを調べることで確かめた。結果としては、重要度と人手で求めた実際の支持、不支持の間に高い相関を示す記事が存在した。また、この研究を行ったことによって、重要度算出における課題についても明らかにすることができた。</p>		