

記事コメント生成によるフェイクニュースの早期検出

柳 裕太^{1,a)} 折原 良平^{1,b)} 清 雄一^{1,c)} 田原 康之^{1,d)} 大須賀 昭彦^{1,e)}

概要：SNS 上でフェイクニュースが拡散されて事実と異なる風評が広がりやすくなった。誤った風評に騙された人々が社会的損害を与えるためこの問題は深刻である。ファクトチェックがフェイクニュース対策として行われているが、属人的な作業である上に時間がかかるため、フェイクニュースと比べ拡散されにくい課題がある。自動でフェイクニュースを検出することが広く研究されており、記事に加えてリツイートやリプライといったソーシャルコンテキストが検出性能を改善することが確認されている。しかしながら、ソーシャルコンテキストは SNS ユーザの拡散によって生まれる情報であるため、同じく検出に時間がかかる。我々はフェイクニュースの早期検出に向けて、ソーシャルコンテキスト情報として記事へのコメントを生成することで検出を補助するフェイクニュース自動検出モデルを提案する。コメント生成モデルと真偽分類モデルは記事とコメントを併せ持つデータセットから学習される。検証時は実在コメント件数を制限した状況から新たにコメントを生成した上で真偽分類を補助させる。実際に生成コメントを付加して分類した場合と、付加せず分類した場合を比較した結果、生成コメントを付加した方がより多くのフェイクニュースを検出した。これは、我々の提案したモデルが早期検出に向くことを示唆している。

¹ 電気通信大学
UEC, Chofu, Tokyo 182-8585, Japan
^{a)} yanagi.yuta@ohsuga.lab.uec.ac.jp
^{b)} orihara@acm.org
^{c)} seiuny@uec.ac.jp
^{d)} ohsuga@uec.ac.jp
^{e)} tahara@uec.ac.jp