

複数のモダリティを考慮した

災害関連ツイートのノイズ除去の可能性

関西大学大学院 総合情報学研究科 総合情報学専攻

森野 穰

要 旨

本研究の目的は、災害関連ツイートから被害や救援要請とは直接関わらないノイズとしての情報を削除することである。災害発生時には、ソーシャルメディアには様々な災害情報が投稿される。特に Twitter 上ではその速報性と拡散力の強さから、自身の被害状況や救助・支援を求める投稿が見られる。一方で、災害時の Twitter にはテレビ報道の拡散や個人の感想といった救援要請に直接関わらないツイートが多く含まれているため、被害情報や救援情報に関わるツイートを的確に抽出することが難しい。また、Twitter に投稿された災害時の被害情報・救援要請情報を活用するためには、投稿情報の信頼性を担保する必要がある。最終的には人手による投稿の確認が必須となる。災害時という人的リソースが欠如した状態で Twitter に投稿された災害情報を効率的に確認し活用するためには、膨大な SNS 情報の中から災害被害とは無関係な情報をノイズとして扱うことで取り除き、人が確認すべき情報の絶対量を減らすことが望ましい。そういった社会的課題に対し、ツイートのテキスト情報に加え、画像の特徴を考慮しノイズとなる情報を検出することで、被害情報や救援情報の抽出精度向上を試みる。本稿では、令和 2 年 7 月豪雨時のツイートを対象とし、テキストと画像の用いられ方に関する調査を行う。ツイート画像に付随したテキストの特徴は他の類似ツイート画像のテキスト特徴と類似するという仮定のもと、人手で分類したツイート中の画像とテキストの傾向を分析し、ノイズとして扱うべきデータを明らかにする。これにより、Twitter に投稿された災害情報を効率的に取得するためのデータクレンジングの可能性について考察する。

キーワード：Twitter, 情報信頼性, 災害情報処理, 情報トリアージ