

# Twitterを用いた交通事故を含む語と実際の 事故率との関係性の考察

D-4

Consideration of Relationship Between the Tweets including  
Traffic Accidents and the Real Percents of Accidents

橋本 大樹

伊與田 光宏

Taiki HASHIMOTO

Mituhiro IYODA

千葉工業大学情報工学科

Department of Computer Science, Chiba Institute of Technology

## 1.はじめに

Twitterは、10代から60代まで、幅広い層から注目されているソーシャルメディア(SNS)である。日本でのアクティブユーザーは4500万人を超えており、特に若年層が多く割合を占めている。気軽に情報の発信や収集ができる。

事故率とは、自動車の保有数や使用数と年間の事故件数の割合によって表されるものである。本研究では事故率を表す上で重要な年間の事故件数を扱う。

## 2.目的

本研究では、Twitterを使用し交通事故に関する言葉を含んだツイートの収集したものと、実際の交通事故件数との関係性を調べることを目的とする。

## 3.システム構成

関係性を調べる為の流れを下記に示す。

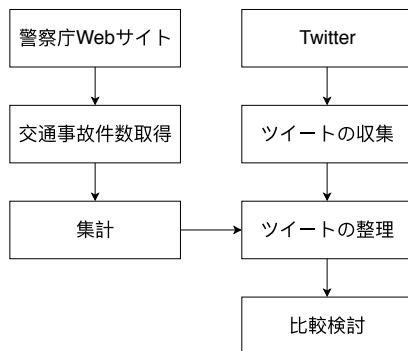


図1.フローチャート

### 3.1 GetOldTweet-Python

本研究では、GetOldTweet-Pythonというものを用いてツイートの収集を行う。現在、公式のAPIでは過去ツイートが約1週間までのツイートしか取得できない。GetOldTweet-Pythonでは詳細な検索をすることができず、ツイートの取得制限がない。多くのツイートを取得可能なため、これを使用する。

## 4.比較方法

GetOldTweet-Pythonで集めたツイートを分類したものと交通事故件数の相関係数を求めることで関係性を示す。

### 4.1 比較対象

本研究で比較するものは、警察庁が公開している2016年から2019年までの月報交通事故統計とGetOldTweet-Pythonで収集した2016年から2019年の交通事故を含んだツイートを整理したものを比較対象とする。

### 4.2 ツイートの整理

ツイートの整理として、収集したツイート2,557,394件からURLを含んだスパムツイートや交通事故とは関係のない#(ハッシュタグ)を含んだツイートなど、合計989,478件を比較対象から外した。

残った1,567,916件のツイートを本研究の比較対象とし実験を行った。

## 5.結果

実験の結果を以下の表1に示す。

表1 各年代における交通事故件数とツイート数

年代	交通事故件数	ツイート数	相関係数
2016年	499,232	351,576	0.13
2017年	472,165	358,599	0.53
2018年	430,601	373,584	-0.12
2019年	381,002	484,157	-0.13
合計	1,783,000	1,567,916	

各年代ごとに分けて相関係数を求めた結果、2017年代が正の相関があり、他の年代にはほとんど相関がないという結果になった。

## 6.おわりに

本研究では、実際の交通事故件数と整理したツイートを相関係数を用いて相関性を求めた。