

Twitter上における フェイクニュース拡散構造の分析

グループ演習第7班

怒田 晟也

中井 啓太

峰 亮太郎

Wang Zilin

アドバイザー教員 伊藤誠

目次

1. 背景・目的
2. 関連研究紹介
3. データの収集
4. 分析手法・結果
5. 考察・まとめ

背景・目的

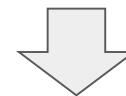
フェイクニュースとは

マスメディアやソーシャルメディア等で

事実とは異なる情報を報道すること



米大統領選挙(2016)で話題



ニュースや新聞でも
取り上げられるように

フェイクニュースの実状

フランス大統領選に関するTweetを

英オックスフォード大学のインターネット研究所が分析



1/4が出所のわからない場所を情報源としている

また、中国やロシアではフェイクニュースの作成・拡散を
請け負うサービスが存在

→フェイクニュースによって金銭的な利益を得ている

過去のフェイクニュース・虚偽報道

朝日新聞の
森友学園報道

自衛隊訓練
拒否報道

オバマ元大統領が
離婚！？

古いものだと . . .

1891年 西郷隆盛生還偽報道事件

1971年 淡路丸沈没虚偽報道事件 など

出典：<https://www.sankei.com/politics/news/180213/plt1802130019-n1.html>
<https://www.j-cast.com/2012/07/23140353.html?p=all>
<https://www.snopes.com/fact-check/obama-divorce-rumors/>

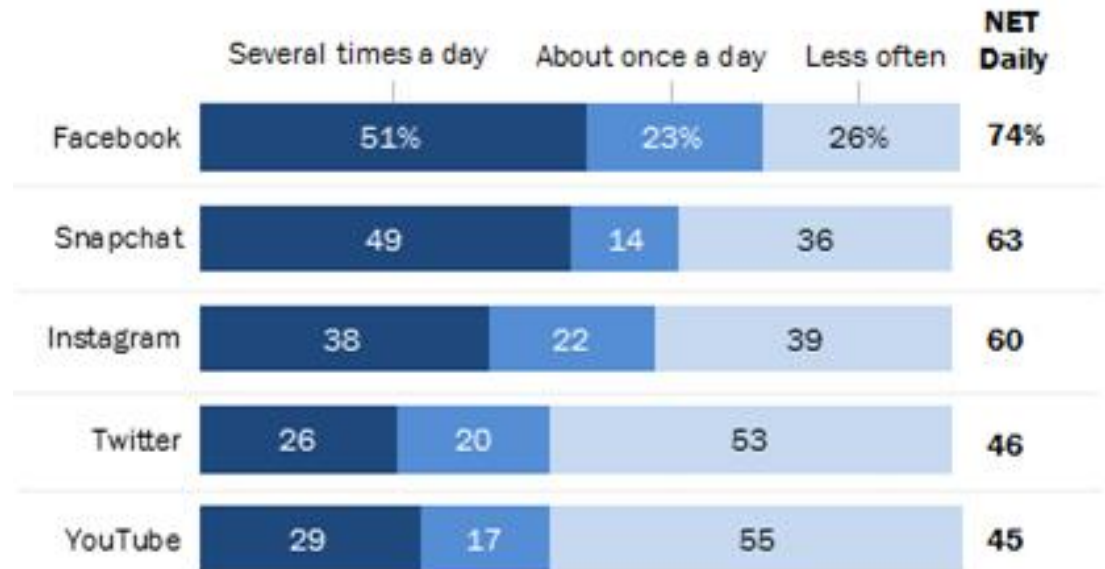
SNS普及の影響

米大統領選挙におけるフェイクニュースには
Facebookが大きな役割を果たしていた

約3/4のユーザが
毎日Facebookに
アクセスしている

A majority of Facebook, Snapchat and Instagram users visit these platforms on a daily basis

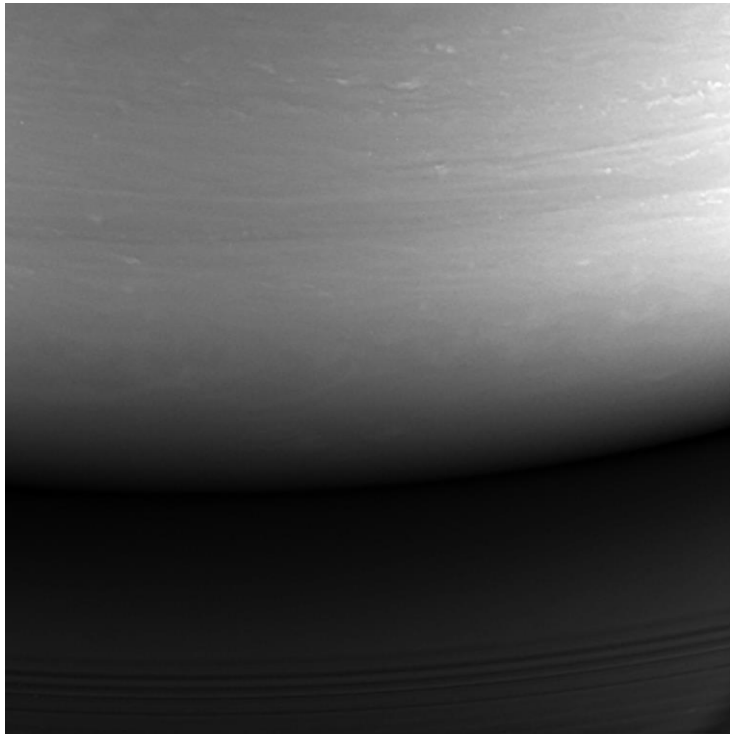
Among U.S. adults who say they use ____, the % who use each site ...



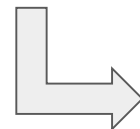
出典 : Survey conducted Jan. 3-10, 2018
PEW RESEARCH CENTER

フェイクニュースの発信が容易に

実際の画像



Impact Site: Cassini's Final Image
Image Credit: NASA/JPL-Caltech/Space Science Institute

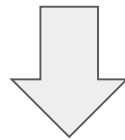


1000件以上リツイート

研究の目的

安易に誤った情報を拡散しないためにも、
フェイクニュースの特徴について知る必要がある

Twitterにおけるフェイクニュースの拡散に注目



拡散しているユーザやツイートの内容から
フェイクニュースが拡散していく構造の分析を行う

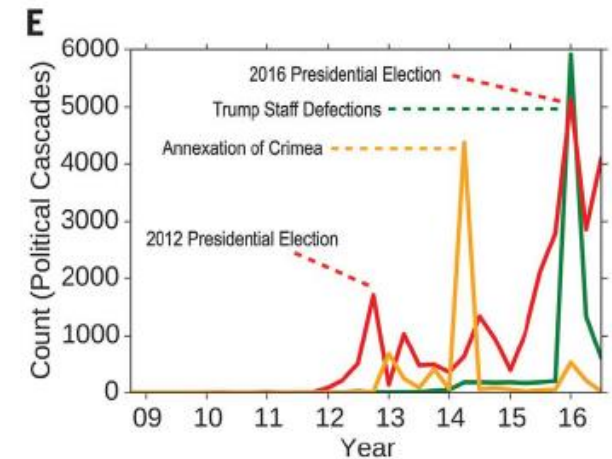
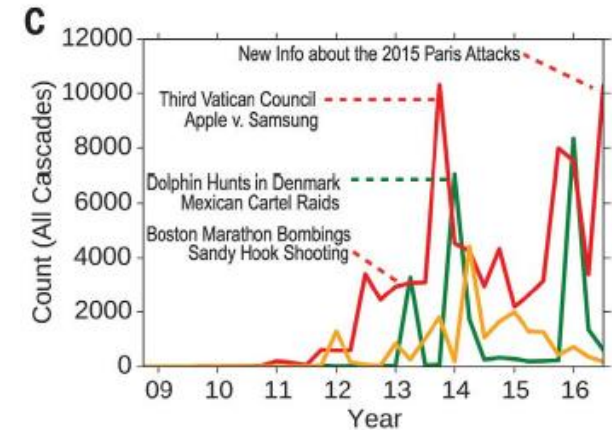
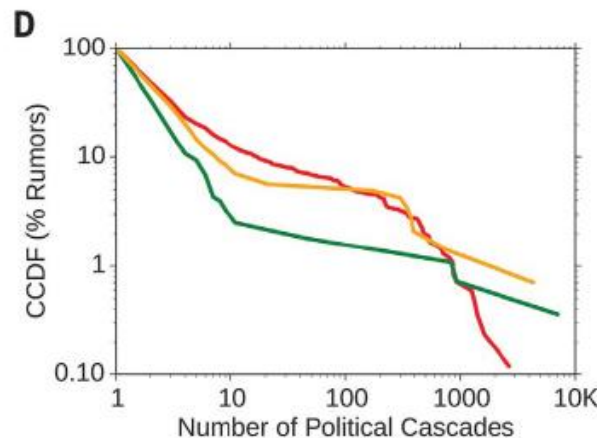
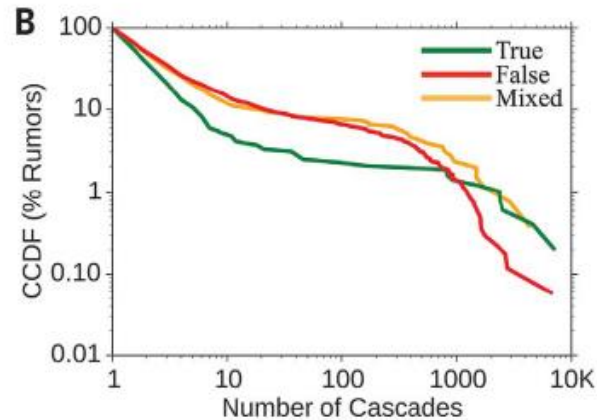
関連研究

関連研究①：「滝」グラフ

■The spread of true and false news online (Soroush Vosoughi, et al. 2018)

フェイクニュースを
広めやすいユーザー
の特徴:

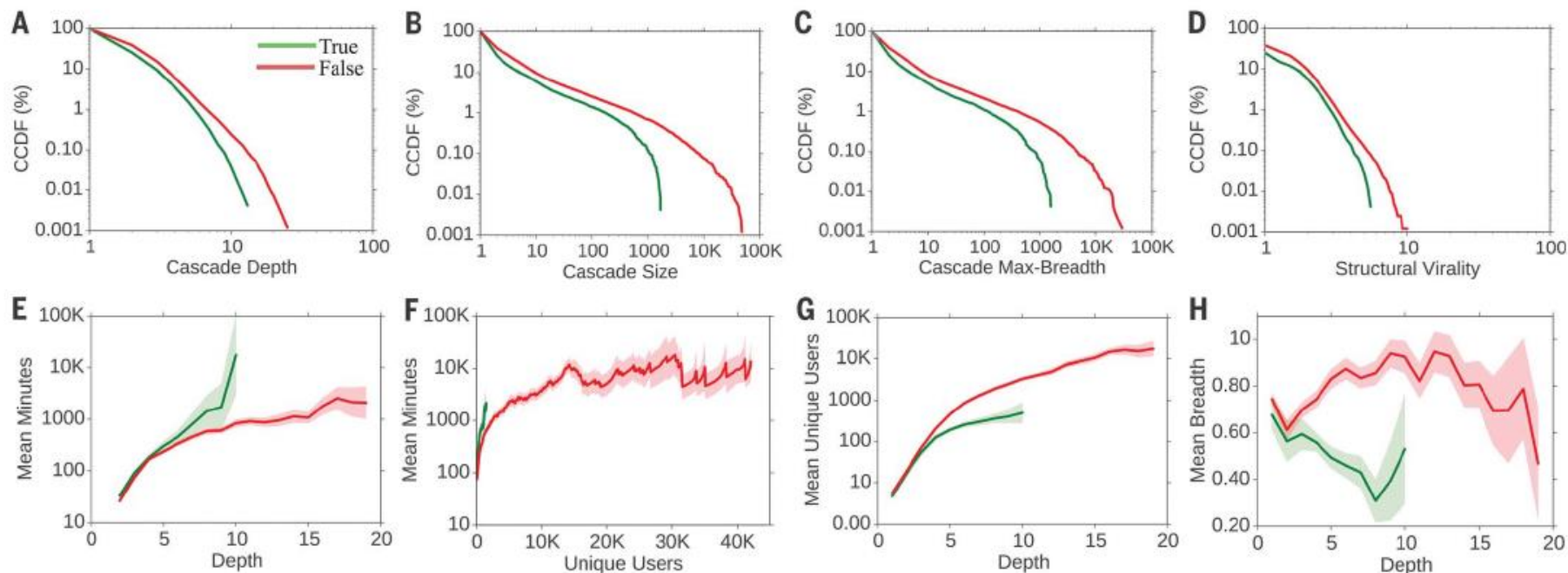
- フォロワー数 **少**
- フォロー数 **少**
- ツイートの頻度 **低**
- 認証ユーザーの割合 **少**
- アカウント保有期間 **短**



関連研究①：「滝」グラフ

■ The spread of true and false news online (Soroush Vosoughi, et al. 2018)

フェイクニュースの特徴：より多くの人に、より速く広まり、何層にも渡ってリツイートされている。どのカテゴリーにおいても、嘘は真実よりも遠く、早く、深く、広範囲に広がっていった。

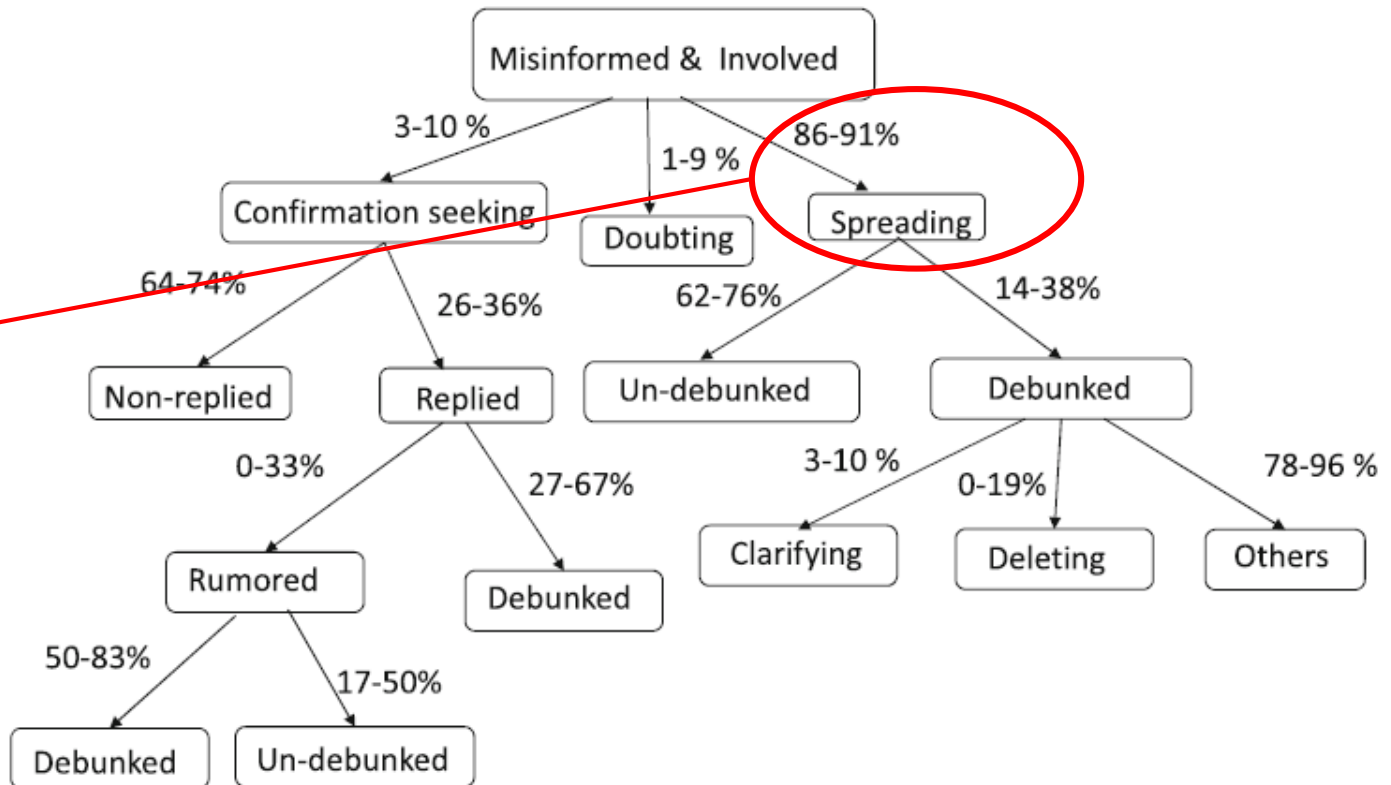


関連研究②：ハリケーン・サンディとボストンマラソン爆破テロ事件に関連するツイッター分析

■Rumor response, debunking response, and decision makings of misinformed Twitter users during disasters (Wangら, 2018)

結論：ツイッターユーザーはうわさの検知能力が低く、すぐに拡散してしまう。

86～91%のツイッターユーザーは考えずに拡散する



データの収集

データの収集

対象とする「フェイクニュース」と「本当のニュース」を以下の様に定義する

フェイクニュース

- 普通のニュースサイトの様に見えるWebページへのリンクを含むツイート
- 何らかの公式アカウントのように見えるユーザのツイート

本当のニュース

- BBCやNew York Timesといった著名な報道機関が総じて報じている内容

収集したデータ

- 対象ツイートの全リツイートの投稿時間
 - ただし非公開ユーザは除く
- 対象ツイートをリツイートしたユーザの情報
 - プロフィール, ツイート数, フォロワー数など
- 対象ツイートをリツイートしたユーザ同士のフォロー関係

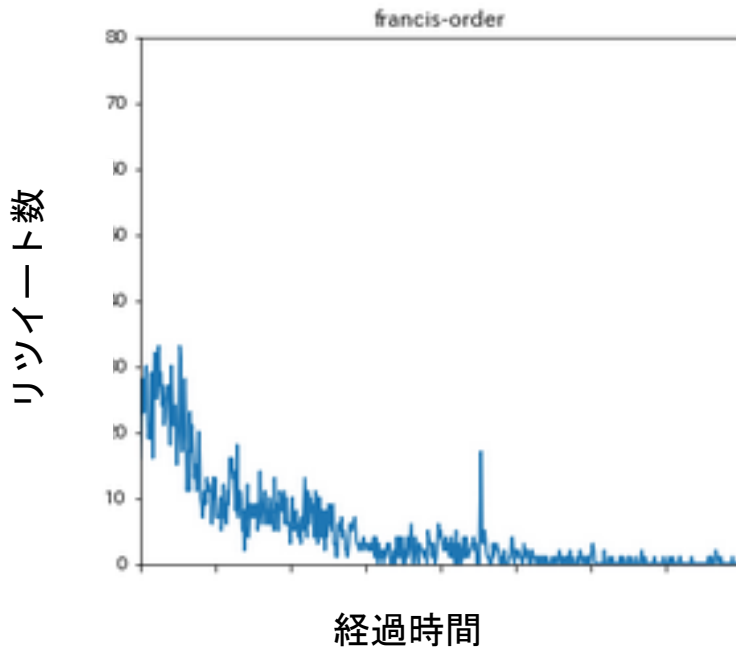
収集したデータのサイズ

- フェイクニュース9件
- 本当のツイート7件
- リツイート総数19,770件
- ユーザ総数18,336人
- フォロー関係総数306,483件

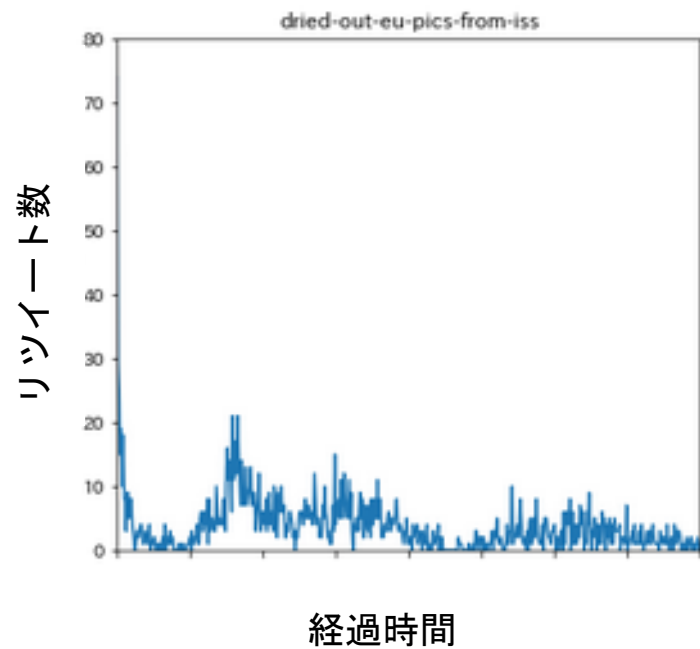
分析手法・結果

リツイートの時系列データ

フェイクニュース



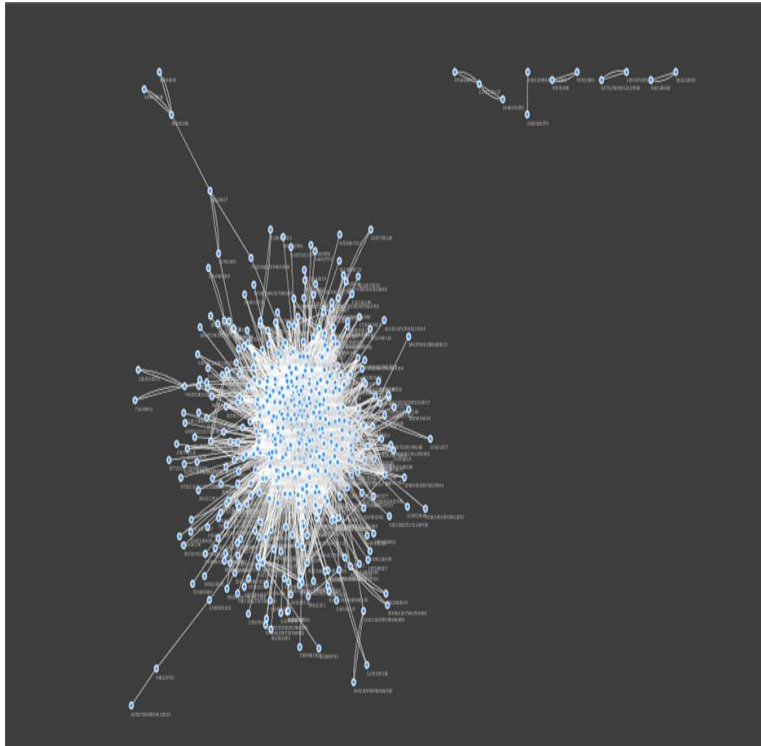
本当のニュース



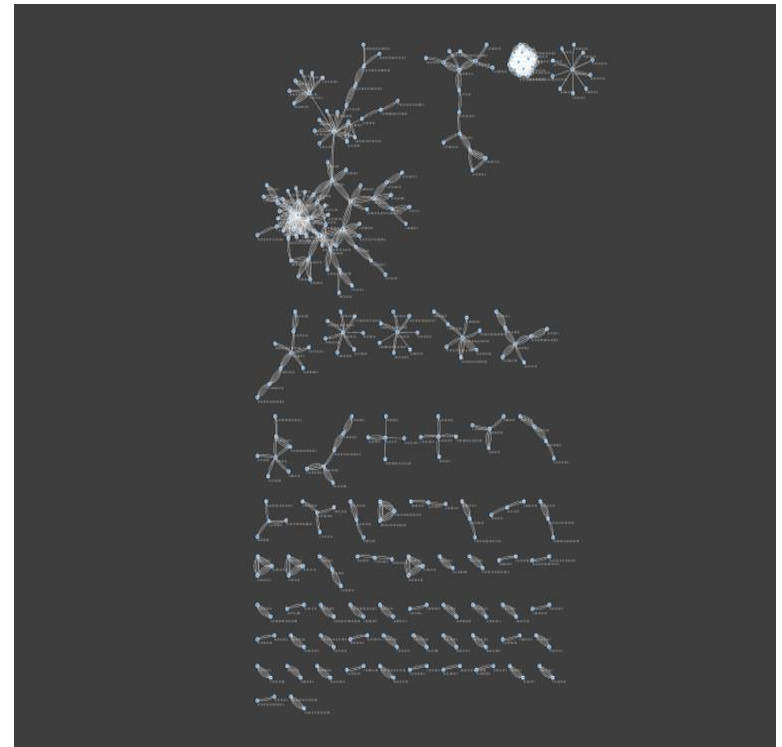
フェイクニュースの方が本当のニュースより短期間で拡散される？

リツイートしたユーザーのフォロー関係

フェイクニュース



本当のニュース



フェイクニュースをリツイートしたユーザーは相互フォローの関係にある？

仮説

仮説① フェイクニュースが本当のニュースより短期間で拡散される？

⇒ 1時間以内におけるリツイートの割合に有意差があるか検証

仮説② フェイクニュースをリツイートしたユーザーは相互フォローの関係にある？

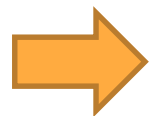
⇒ ユーザー1人当たりの相互フォローの人数に有意差があるか検証

仮説①フェイクニュースの方が本当のニュースより短期間で拡散される?

ツイートがされてから1時間以内のリツイートの割合

| | フェイクニュース | 本当のニュース |
|----|----------|---------|
| 平均 | 16.57% | 25.30% |
| 分散 | 1.63 | 3.43 |

$$P(T \leq t) = 0.177 > 0.05$$



有意差はない

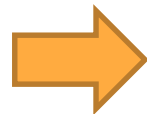
- フェイクニュースは本当のニュースと比較し、短期間で拡散されるとは言えない

仮説② フェイクニュースをリツイートしたユーザーは相互フォローの関係にある？

ユーザー一人当たりの相互フォローの割合

| | フェイクニュース | 本当のニュース |
|----|----------|---------|
| 平均 | 5.70 | 0.72 |
| 分散 | 80.80 | 1.11 |

$$P(T \leq t) = 0.0787 > 0.05$$



有意差はない

- フェイクニュースと本当のニュースのリツイートしたユーザーの相互フォロー数に差があるとは言えない
- しかしフェイクニュースの相互フォロー数は多い傾向にある

まとめ

考察・まとめ

- 今回2つの仮説を立て分析を行ったが、残念ながらこれらを実証することはできなかった。
 - データの取得時間などの問題によりデータが十分でなかったこと
- しかし有意差は見られなかったが、フェイクニュースをリツイートしたユーザの相互フォロワー数が本当のニュースより多い傾向にあった。

今後の課題

- 今回用いたツイートのデータは、フェイクニュース9件
本当のツイート7件とデータ数が十分でなかった
→ 今後はより多くのツイートのデータを用い分析する
必要がある。
- フェイクニュースの拡散過程を分析するには、今回収集
したデータだけでは不十分かもしれない
→ リツイートした人物のツイート歴やプロフィールな
どの点からも分析を行う必要がある。

ご清聴ありがとうございました