

On the Origins of Memes by Means of Fringe Web Communities

Savvas Zannettou, Tristan Caulfield, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Guillermo Suarez-Tangil.

In Proceedings of International Measurement Conference (IMC) '18. ACM, 15 pages.

情報工学科 寺岡研究室
安森 涼 61619027

背景

- Web の台頭
 - 考え, 文化, 画像, 映像などが今までにない速度で拡散
- 政治思想や信条の拡散に **ミーム (meme)** を利用
 - ミーム: インターネットで流行する画像や言い回し
- 攻撃的なミームの特定, 排除の需要増加
 - e.g., 人種差別的, 政治的

ミームの起源と影響力を解釈できるツールの作成が必要



通常



人種差別的



政治的

関連研究

- Twitter のハッシュタグからミームの人気度調査 [69]
- Facebook 間の伝搬の過程で新しいミームが生まれることを証明 [6]
- 4chan, Reddit で口汚い言葉を用いた投稿の検知 [10]

● 検証対象でない SNS データセットは利用不可

- 複数 SNS のミームを意味付けが必要
 - ハッシュ化した画像のクラスタリングを実施

● 複数 SNS 間のミームの伝搬の検知は不可能

- 複数 SNS 間のミームの伝搬を検知が必要
 - 伝搬性をもつ事象の累積発生件数を表現できるモデルを作成

[69] L. Weng, F. Menczer, and Y.-Y. Ahn. Predicting Successful Memes Using Network and Community Structure. In ICWSM, 2014.

[6] L. A. Adamic, T. M. Lento, E. Adar, and P. C. Ng. Information Evolution in Social Networks. In WSDM, 2016.

[10] E. Chandrasekharan, M. Samory, A. Srinivasan, and E. Gilbert. The bag of communities: Identifying abusive behavior online with preexisting Internet data. In CHI, 2017.

■ 前提知識 1 | 利用する SNS データセット

● Mainstream : 悪意のないミームが拡散されやすい SNS

SNS 名	概要
Twitter	140 字以内の短い記事を投稿し合うサイト
Reddit	ニュース記事, 画像やテキストの投稿サイト

● Fringe : 悪意を含むミームが拡散されやすい SNS

SNS 名	概要
The_Donald (T_D)	Reddit のトランプについてのチャンネル
Reddit	匿名の掲示板 4chan の政治チャンネル
Gab	ほぼ規制がない, 言論の自由を尊重するサイト

* 2016/07/01 から 2017/07/31 のデータを利用

■ 前提知識 2 | ミームを保有するデータベース

● Know Your Meme (KYM)

- ミームの辞書をまとめたクラウドソーシングサービス
- ミームに対して役立つメタデータを供給
 - 起源
 - キーワードタグ
 - 説明
 - 例
 - イメージギャラリー

データセットごとの投稿数

	Twitter	Reddit	/pol/	Gab	KYM
投稿数	1.4×10^9	1.0×10^9	4.8×10^7	1.2×10^7	1.5×10^4
画像付投稿数	2.4×10^8	6.2×10^7	1.3×10^7	9.6×10^5	1.5×10^4
画像数	1.1×10^8	4.0×10^7	4.0×10^6	2.4×10^5	7.1×10^5

提案手法 | 複数 SNS のミームを意味付け

1. Fringe データの一部をハッシュ化



2. クラスタリング



3. クラスタごとに Medoid を決定



4. クラスタごとに注釈付け



5. データセットの拡大

1. 64 bit ハッシュ値を利用
 - 似た画像は近い値を保持
2. ハミング距離準拠のアルゴリズムを利用
3. Medoid : クラスタ内の自身以外の画像との距離の総和が最小となる画像
4. ハッシュ化した KYM データを利用
5. ハッシュ化した全てのデータを利用

入力が画像なのでどの SNS にも対応可能

評価 | SNS 毎に投稿されるミームの割合

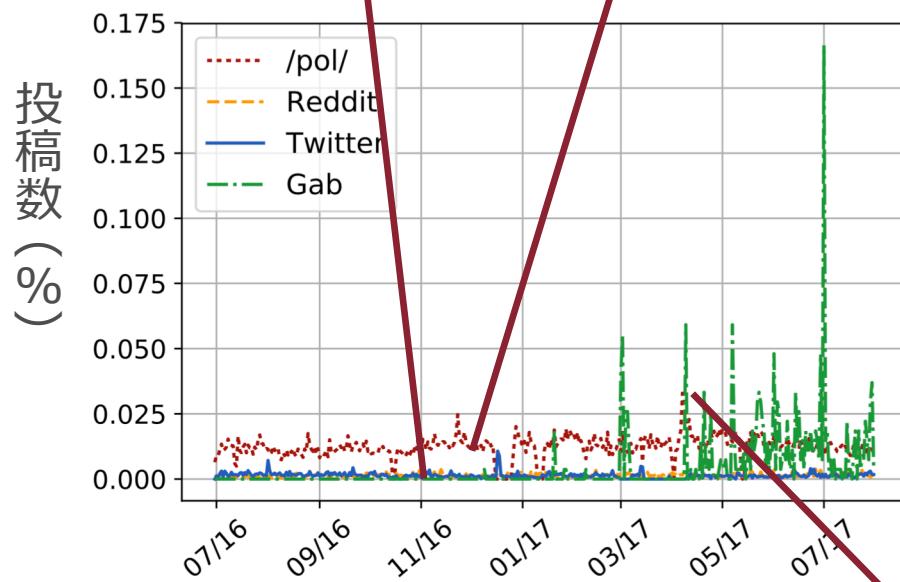
- 人種差別的なミーム (図中赤) は Fringe で上位
 - 'racist', 'antisemitism' などのタグを持つもの
- 害のないミームが Mainstream で上位
- 政治的なミーム (図中黄) 投稿はどこにでも存在
 - 'politics', 'trump', 'clinton' などのタグを持つもの

Fringe		Mainstream	
/pol/	Reddit	Gab	Twitter
Entry	Posts (%)	Entry	Posts (%)
Feels Bad Man/Sad Frog	64,367 (4.9%)	Manning Face	12,540 (2.2%)
Smug Frog	63,290 (4.8%)	That's the Joke	7,626 (1.3%)
Happy Merchant (R)	49,608 (3.8%)	Feels Bad Man/ Sad Frog	7,240 (1.3%)
Apu Apustaja	29,756 (2.2%)	Confession Bear	7,147 (1.3%)
Pepe the Frog	25,197 (1.9%)	This is Fine	5,032 (0.9%)
Make America Great Again (P)	21,229 (1.6%)	Smug Frog	4,642 (0.8%)
Angry Pepe	20,485 (1.5%)	Roll Safe	4,523 (0.8%)
Bait this is Bait	16,686 (1.2%)	Rage Guy	4,491 (0.8%)
I Know that Feel Bro	14,490 (1.1%)	Make America Great Again (P)	4,440 (0.8%)
Cult of Kek	14,428 (1.1%)	Fake CCG Cards	4,438 (0.8%)
Laughing Tom Cruise	14,312 (1.1%)	Confused Nick Young	4,024 (0.7%)
Awoo	13,767 (1.0%)	Daily Struggle	4,015 (0.7%)
Tony Kornheiser's Why	13,577 (1.0%)	Expanding Brain	3,757 (0.7%)
Picardia (P)	13,540 (1.0%)	Demotivational Posters	3,419 (0.6%)
Big Grin / Never Ever	12,893 (1.0%)	Actual Advice Mallard	3,293 (0.6%)
Reaction Images	12,608 (0.9%)	Reaction Images	2,959 (0.5%)
Computer Reaction Faces	12,247 (0.9%)	Handsome Face	2,675 (0.5%)
Wojak / Feels Guy	11,682 (0.9%)	Absolutely Disgusting	2,674 (0.5%)
Absolutely Disgusting	11,436 (0.8%)	Pepe the Frog	2,672 (0.5%)
Spurdo Sparde	9,581 (0.7%)	Pretending to be Retarded	2,462 (0.4%)
Total	445,179 (33.4%)	94,069 (16.7%)	4,808 (17.0%)
			249,047 (26.4%)

評価 | ミームの投稿数の推移

政治的ミームの投稿数推移は
実世界の出来事と密接に関係

Mainstream で
ほとんど投稿無し



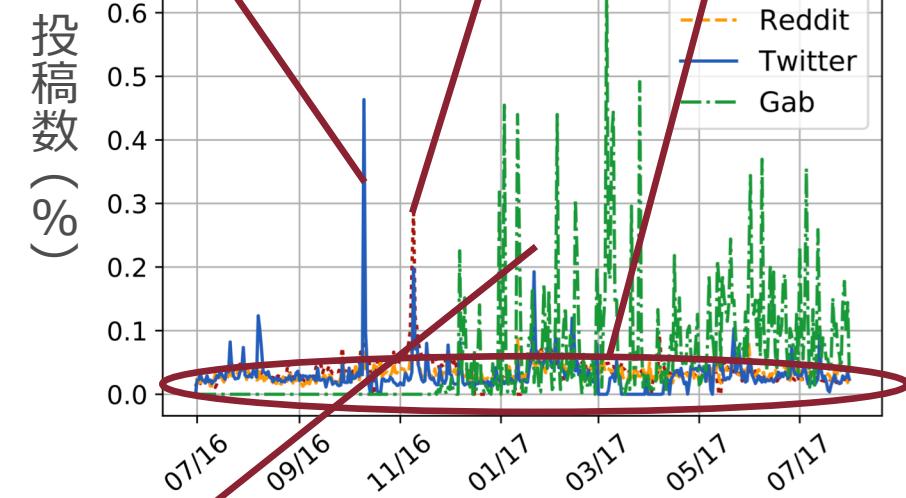
継続的な
/pol/ への投稿

第 2 回 US
大統領選挙討論会

2016 US
大統領選挙

継続的な投稿

2017 年
Gab の投稿増加



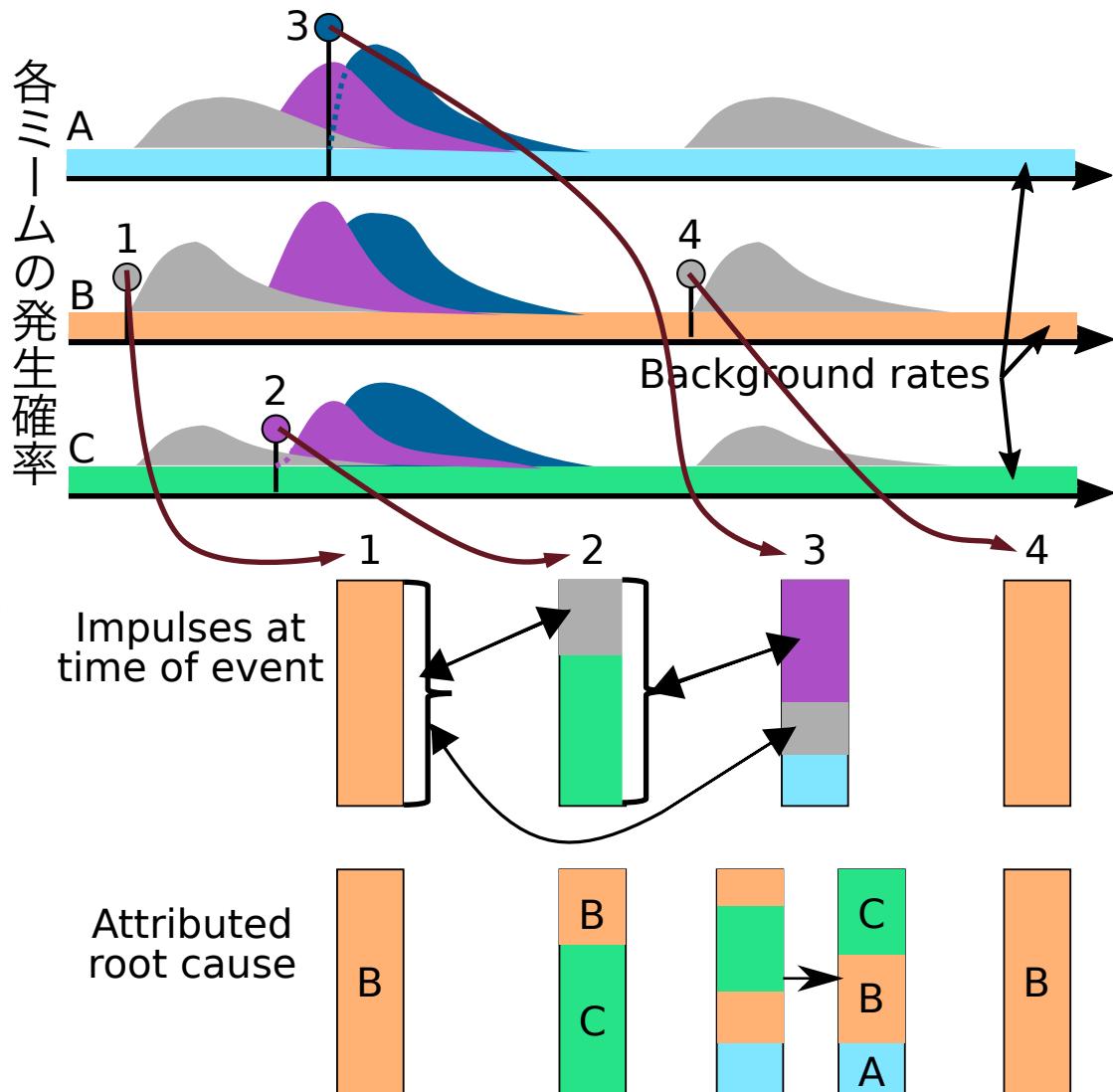
人種差別的ミーム

政治的ミーム

提案手法 | 複数 SNS 間のミームの伝搬を検知

- ミームの投稿 (事象) がどの SNS に起因するかを調査

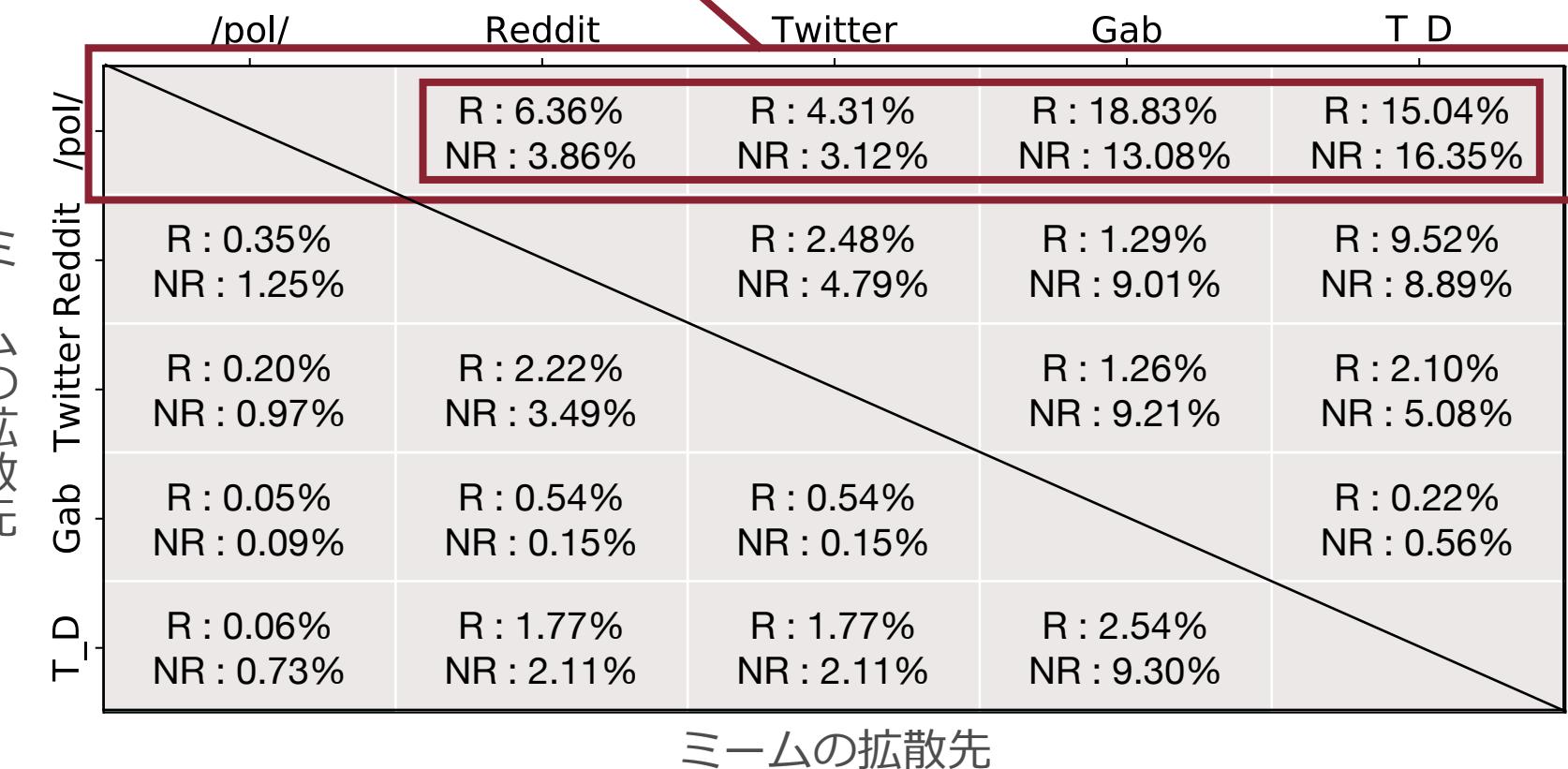
- I. 各事象発生時にポールを立て、
後ろに波を描画
- II. 各ポール時点での発生確率を
「事象の発生要因」として抽出
- III. 「事象の発生要因」内で事前に
発生した事象を再帰的に置換
- IV. 「投稿に起因する SNS」を特定



評価 | 拡散された人種差別的ミームの割合

/pol/ が拡散元のとき
他の SNS への影響力が最大

R : 人種差別的
NR : 非人種差別的



評価 | 人種差別的ミームが拡散される確率

T_D にミームを投稿すると
他の SNS に広がる可能性が高い

R : 人種差別的
NR : 非人種差別的

		/pol/	Reddit	Twitter	Gab	T_D	Total Ext
		/pol/	R : 0.4% NR : 1.5%	R : 0.3% NR : 1.8%	R : 0.2% NR : 0.4%	R : 0.2% NR : 0.9%	R : 1.1% NR : 4.5%
		Reddit	R : 5.1% NR : 3.3%	R : 2.9% NR : 7.1%	R : 0.2% NR : 0.7%	R : 1.4% NR : 1.3%	R : 9.5% NR : 12.4%
		Twitter	R : 2.4% NR : 1.7%	R : 1.9% NR : 2.3%	R : 0.1% NR : 0.5%	R : 0.3% NR : 0.5%	R : 4.7% NR : 5.0%
		Gab	R : 5.3% NR : 3.0%	R : 4.0% NR : 1.9%	R : 0.5% NR : 3.1%	R : 0.2% NR : 1.0%	R : 10.0% NR : 9.1%
		T_D	R : 6.3% NR : 13.6%	R : 12.2% NR : 15.0%	R : 2.5% NR : 12.6%	R : 2.3% NR : 5.1%	R : 23.3% NR : 46.2%

まとめ

- 政治思想や信条の拡散にミームを利用
 - 攻撃的なミームの排除, 特定の需要増加
 - ミームの起源と影響力を解釈できるツールは少ない
- 複数 SNS のミームを意味付け
 - 人種差別的ミームが Fringe に多く投稿
 - 政治的ミームは普遍的に存在
 - 実世界の出来事は政治的ミームの投稿数に影響
- 複数 SNS 間のミームの伝搬を検知
 - /pol/ の他の SNS への影響力が最大
 - T_D にミームを投稿すると他の SNS に広がる可能性が高い