

学位申請論文 2016 年度（平成 28 年度）

感情分析を用いたショックの前後における
会話ネットワーク上のコミュニティ変化の分析

指導教員 坂田 一郎 教授

東京大学大学 工学部 システム創成学科 C

03-150998 與島 仙太郎

2017 年 2 月 2 日 提出

学位申請論文 2016 年度（平成 28 年度）

感情分析を用いたショックの前後における 会話ネットワーク上のコミュニティ変化の分析

内容梗概

ソーシャルメディアの普及により、情報の動きは双方向的なものとなった。その結果としてソーシャルメディアは情報拡散や意見形成の面において重要な役割を果たしている。選挙活動でのソーシャルメディアの活用などがその顕著な例である。その中で、ソーシャルメディア内のレジリエンスが損なわれてしまう例としてデマの拡散などの問題が起こっている。実際、ソーシャルメディア上では Echo chamber(意見の偏り) や、Silent Majority(無言の大多数) といったものが確認されている。しかし、そのような意見形成や感情変化の過程に関しては十分な研究がなされていない。本研究では、2016 アメリカ大統領選の前後約一ヶ月のデータを使用することで、大統領選の前後において、特定話題に関する感情度極性で分けたコミュニティの規模や密度がどのように変化するのか、またポジティブネガティブの極性を持つコミュニティ間の対立構造はどのように変化するのかについて明らかにすること目的とする。

提案手法においては、ツイートの感情度の分析と、クエリに基づいたユーザ群におけるユーザの感情属性の付与を行い、属性付与に基づくポジティブコミュニティとネガティブコミュニティについてその規模と密度に関する分析、両コミュニティの間でのコミュニケーションの変化を分析した。

実験においては各分野について提案手法を実施し、それぞれのクエリで抽出された投稿とその投稿者であるユーザ群に対して、ほぼ全ての分野についてポジティブコミュニティとネガティブコミュニティの対立構造が深まるという結果を得た。逆に相互のコミュニケーションが増加している分野については、当事者ではないユーザが多く含まれることが推測されるコミュニティであり、その場合は大統領選の前後でネガティブコミュニティの密度が小さくなりネガティブ内でのコミュニケーションも減っているということが示された。

03-150998 與島 仙太郎

目 次

第 1 章 序論	1
1.1 背景	1
1.1.1 ソーシャルメディアの普及	1
1.1.2 情報伝達媒体としてのソーシャルメディアの応用	2
1.1.3 選挙においてレジリエンスを損なう事象の発生とソーシャルメディア	2
1.1.4 大統領選への適用可能性と意見形成過程の発見	2
1.1.5 意見形成の指標としての感情分析	2
1.2 目的	3
1.3 貢献	4
1.4 本論文の構成	4
第 2 章 関連研究	5
2.1 ソーシャルネットワークの分析	5
2.1.1 ソーシャルメディアとしての Twitter 分析	5
2.1.2 ソーシャルメディアのコミュニティ分析	5
2.2 テキストの感情分析	6
2.2.1 感情表現の分類	6
2.2.2 短文における感情分析	6
2.3 ネットワーク変化の分析	7
2.3.1 ネットワーク対立構造の分析	7
2.3.2 ソーシャルメディアにおける意見の偏り	8
2.4 本研究の位置づけ	8
第 3 章 手法	9
3.1 手法の手順	9
3.2 データの収集と前処理	9
3.2.1 データ収集対象となるソーシャルメディアの決定	9
3.2.2 データ収集のクエリ設定とデータの収集	10
3.2.3 収集したデータの前処理	10
3.2.4 メンション情報のカラムの追加	11
3.3 ツイートに対する感情値の付与	11
3.4 コミュニティの分割	12

3.5 ネットワークの構築	13
3.5.1 ユーザのポジティブネガティブ属性の付与	13
3.5.2 エッジに関する属性付与	14
3.6 コミュニティに関する分析指標	14
3.7 ネットワークに関する分析指標	15
3.7.1 ランダムに張ったエッジとの比	15
第4章 アメリカ大統領選	17
4.1 アメリカの政治	17
4.1.1 アメリカの政治構造	17
4.1.2 オバマ大統領期から見るアメリカ政治の変遷	18
4.1.3 オバマ政権における各種政策について	18
4.1.4 ポストオバマの選挙戦について	20
4.2 2016 アメリカ大統領選挙の二大候補者について	20
4.2.1 ヒラリークリントンとその政治的特徴	20
4.2.2 ドナルドトランプとその政治的特徴	20
4.3 2016 アメリカ大統領選挙の争点	21
4.3.1 経済政策	21
4.3.2 外交政策, 国防政策	21
4.3.3 医療政策	21
4.3.4 右傾化	22
4.3.5 階級闘争	22
4.3.6 雇用創出	22
4.3.7 人種問題	22
4.3.8 移民問題	23
第5章 実験	25
5.1 各ケースのデータ設定	25
5.1.1 分析対象の分野とクエリ	25
5.1.2 分析分野とクエリの指定	25
5.1.3 データセットの設定について	26
5.2 評価方法	27
第6章 結果	29
6.1 大統領と大統領候補者	29

6.1.1 ヒラリー	29
6.1.2 トランプ	33
6.1.3 オバマ	36
6.2 選挙関連	40
6.2.1 民主党	40
6.2.2 共和党	44
6.2.3 選挙	48
6.3 外交	53
6.3.1 移民	53
6.3.2 イギリス	55
6.3.3 ロシア	59
6.4 人種問題	63
6.4.1 白人	63
6.4.2 黒人	67
第7章 考察	73
7.1 考察の概要	73
7.2 コミュニティ規模・密度の変化	73
7.2.1 候補者	73
7.2.2 選挙関連	74
7.2.3 外交	74
7.2.4 人種	74
7.2.5 コミュニティ規模・密度の変化の考察まとめ	74
7.3 ランダム比から見る会話量	75
7.3.1 候補者	75
7.3.2 選挙関連	75
7.3.3 外交	76
7.3.4 人種	76
7.3.5 ランダム比から見る会話量への考察まとめ	76
第8章 結論	77
8.1 今後の展望	78
8.1.1 選挙中のデータ分析	78
8.1.2 時系列の細分化による変化過程の分析	78

8.1.3 コミュニティ内の話題特徴量の導入	79
8.1.4 世論調査とのデータ比較	79
8.1.5 支持率導出手法の提案	79
 謝辞	
	80
 第9章 巻末収録	87
9.1 経済政策	87
9.1.1 貿易	87
9.2 多様性への価値観	90
9.2.1 リベラル	91
9.3 性の問題	94
9.3.1 ゲイ	95
9.4 サブクエリ分析	98
9.4.1 ヒラリー&オバマ	99
9.4.2 トランプ&オバマ	102
9.4.3 ヒラリー&黒人	106
9.4.4 トランプ&黒人	110
9.4.5 ヒラリー&移民	114
9.4.6 トランプ&移民	118

図 目 次

2.1 Open network	7
2.2 Turtles-up network	7
6.1 大統領選以前のコミュニティ内のユーザ感情度分布	31
6.2 大統領選後のコミュニティ内のユーザ感情度分布	31
6.3 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	32
6.4 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	32
6.5 大統領選以前のコミュニティ内のユーザ感情度分布	35
6.6 大統領選後のコミュニティ内のユーザ感情度分布	35
6.7 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	36
6.8 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	36
6.9 大統領選以前のコミュニティ内のユーザ感情度分布	39
6.10 大統領選後のコミュニティ内のユーザ感情度分布	39
6.11 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	40
6.12 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	40
6.13 大統領選以前のコミュニティ内のユーザ感情度分布	43
6.14 大統領選後のコミュニティ内のユーザ感情度分布	43
6.15 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	44
6.16 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	44
6.17 大統領選以前のコミュニティ内のユーザ感情度分布	47
6.18 大統領選後のコミュニティ内のユーザ感情度分布	47
6.19 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	48
6.20 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	48
6.21 大統領選以前のコミュニティ内のユーザ感情度分布	51

6.22 大統領選後のコミュニティ内のユーザ感情度分布	51
6.23 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	52
6.24 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	52
6.25 大統領選以前のコミュニティ内のユーザ感情度分布	54
6.26 大統領選後のコミュニティ内のユーザ感情度分布	54
6.27 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	55
6.28 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	55
6.29 大統領選以前のコミュニティ内のユーザ感情度分布	58
6.30 大統領選後のコミュニティ内のユーザ感情度分布	58
6.31 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	59
6.32 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	59
6.33 大統領選以前のコミュニティ内のユーザ感情度分布	62
6.34 大統領選後のコミュニティ内のユーザ感情度分布	62
6.35 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	63
6.36 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	63
6.37 大統領選以前のコミュニティ内のユーザ感情度分布	66
6.38 大統領選後のコミュニティ内のユーザ感情度分布	66
6.39 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	67
6.40 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	67
6.41 大統領選以前のコミュニティ内のユーザ感情度分布	70
6.42 大統領選後のコミュニティ内のユーザ感情度分布	70
6.43 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	71
6.44 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	71

9.1	大統領選以前のコミュニティ内のユーザ感情度分布	89
9.2	大統領選後のコミュニティ内のユーザ感情度分布	89
9.3	大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	90
9.4	大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	90
9.5	大統領選以前のコミュニティ内のユーザ感情度分布	93
9.6	大統領選後のコミュニティ内のユーザ感情度分布	93
9.7	大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	94
9.8	大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	94
9.9	大統領選以前のコミュニティ内のユーザ感情度分布	97
9.10	大統領選後のコミュニティ内のユーザ感情度分布	97
9.11	大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	98
9.12	大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	98
9.13	大統領選以前のコミュニティ内のユーザ感情度分布	101
9.14	大統領選後のコミュニティ内のユーザ感情度分布	101
9.15	大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	102
9.16	大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	102
9.17	大統領選以前のコミュニティ内のユーザ感情度分布	105
9.18	大統領選後のコミュニティ内のユーザ感情度分布	105
9.19	大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	106
9.20	大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	106
9.21	大統領選以前のコミュニティ内のユーザ感情度分布	109
9.22	大統領選後のコミュニティ内のユーザ感情度分布	109
9.23	大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	110

9.24 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	110
9.25 大統領選以前のコミュニティ内のユーザ感情度分布	113
9.26 大統領選後のコミュニティ内のユーザ感情度分布	113
9.27 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	114
9.28 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	114
9.29 大統領選以前のコミュニティ内のユーザ感情度分布	117
9.30 大統領選後のコミュニティ内のユーザ感情度分布	117
9.31 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	118
9.32 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	118
9.33 大統領選以前のコミュニティ内のユーザ感情度分布	121
9.34 大統領選後のコミュニティ内のユーザ感情度分布	121
9.35 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	122
9.36 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)	122

表 目 次

3.1 収集したツイートの個別情報	10
3.2 使用するデータ量	11
3.3 文章と付与される感情値	12
6.1 コミュニティ情報	30
6.2 ネットワーク情報	30
6.3 ネットワークの密度	30
6.4 エッジの割合	30
6.5 ランダムネットワークとの比	30
6.6 コミュニティ情報	34
6.7 ネットワーク情報	34
6.8 ネットワークの密度	34
6.9 エッジの割合	34
6.10 ランダムネットワークとの比	34
6.11 コミュニティ情報	38
6.12 ネットワーク情報	38
6.13 ネットワークの密度	38
6.14 エッジの割合	38
6.15 ランダムネットワークとの比	38
6.16 コミュニティ情報	42
6.17 ネットワーク情報	42
6.18 ネットワークの密度	42
6.19 エッジの割合	42
6.20 ランダムネットワークとの比	42
6.21 コミュニティ情報	46
6.22 ネットワーク情報	46
6.23 ネットワークの密度	46
6.24 エッジの割合	46
6.25 ランダムネットワークとの比	46
6.26 コミュニティ情報	50
6.27 ネットワーク情報	50
6.28 ネットワークの密度	50

6.29 エッジの割合	50
6.30 ランダムネットワークとの比	50
6.31 コミュニティ情報	53
6.32 ネットワーク情報	53
6.33 ネットワークの密度	53
6.34 エッジの割合	53
6.35 ランダムネットワークとの比	53
6.36 コミュニティ情報	57
6.37 ネットワーク情報	57
6.38 ネットワークの密度	57
6.39 エッジの割合	57
6.40 ランダムネットワークとの比	57
6.41 コミュニティ情報	61
6.42 ネットワーク情報	61
6.43 ネットワークの密度	61
6.44 エッジの割合	61
6.45 ランダムネットワークとの比	61
6.46 コミュニティ情報	65
6.47 ネットワーク情報	65
6.48 ネットワークの密度	65
6.49 エッジの割合	65
6.50 ランダムネットワークとの比	65
6.51 コミュニティ情報	69
6.52 ネットワーク情報	69
6.53 ネットワークの密度	69
6.54 エッジの割合	69
6.55 ランダムネットワークとの比	69
 9.1 コミュニティ情報	88
9.2 ネットワーク情報	88
9.3 ネットワークの密度	88
9.4 エッジの割合	88
9.5 ランダムネットワークとの比	88
9.6 コミュニティ情報	92

9.7 ネットワーク情報	92
9.8 ネットワークの密度	92
9.9 エッジの割合	92
9.10 ランダムネットワークとの比	92
9.11 コミュニティ情報	96
9.12 ネットワーク情報	96
9.13 ネットワークの密度	96
9.14 エッジの割合	96
9.15 ランダムネットワークとの比	96
9.16 コミュニティ情報	100
9.17 ネットワーク情報	100
9.18 ネットワークの密度	100
9.19 エッジの割合	100
9.20 ランダムネットワークとの比	100
9.21 コミュニティ情報	104
9.22 ネットワーク情報	104
9.23 ネットワークの密度	104
9.24 エッジの割合	104
9.25 ランダムネットワークとの比	104
9.26 コミュニティ情報	108
9.27 ネットワーク情報	108
9.28 ネットワークの密度	108
9.29 エッジの割合	108
9.30 ランダムネットワークとの比	108
9.31 コミュニティ情報	112
9.32 ネットワーク情報	112
9.33 ネットワークの密度	112
9.34 エッジの割合	112
9.35 ランダムネットワークとの比	112
9.36 コミュニティ情報	116
9.37 ネットワーク情報	116
9.38 ネットワークの密度	116
9.39 エッジの割合	116

9.40 ランダムネットワークとの比	116
9.41 コミュニティ情報	120
9.42 ネットワーク情報	120
9.43 ネットワークの密度	120
9.44 エッジの割合	120
9.45 ランダムネットワークとの比	120

第1章 序論

本章では、本研究の背景、目的、貢献及び構成について述べる。

1.1 背景

1.1.1 ソーシャルメディアの普及

近年、Twitter や Facebook をはじめとするソーシャルメディアの発達が目覚ましい。ソーシャルメディアとは、利用者である一般大衆が情報を発信することができるサービスである。このソーシャルメディアの特徴は、情報の発信が一方向的ではなく双方向のもので、コミュニケーションが可能である点である。例えば Twitter では、140 文字以内の投稿が可能であるだけでなく、フォローをしているユーザの投稿情報を確認することや、それに対してコメントを送り合うことが可能である。ソーシャルメディアの中でもリアルタイム性が高く、情報の伝播が起きやすいソーシャルメディアである Twitter は、上記のような機能をユーザーに提供することによって、現在世界で 3 億人以上のアクティブユーザーを持つサービスとなっている。

このような、Facebook や Twitter を始めとするソーシャルメディアの普及によって社会の構造に変化が起こっている。特に大きな変化は、情報の発信受信の形態に関する変化である。従来は雑誌や新聞といった出版社及びテレビなどの放送局が一方的に情報の発信を行い、一般大衆はそれをただ受信するという構造であった。この状態においては、大衆は受信者としての役割しか持たず、自らの意見や情報を自由に発信していくことはできない。そのため、情報は一方向性のものであり、一般大衆がどのように考えているのかを知る事は困難であった。しかし、ソーシャルメディアが普及することにより、従来とは違い双方向の情報発信が可能となった。

ソーシャルメディアの普及により、数多くの出版社や放送局などの「マスメディア」によって発信される情報に対してもリアルタイムで意見を述べることが可能となっている。これは従来単なる受信者であった大衆が、発信者としての役割を担い始めたことを意味する。このように、今まででは発信力を持たなかった大衆の個人個人が、小さな発信者として他者に影響を与えるような発信をすることが可能となった。

他方、発信者としての役割のみを持っていた「マスメディア」は大衆の意見を受信しそれに応じた発信をする必要に駆られるようになった。自らが発信する情報に対して大衆がどのような意見を持っているのか、あるいはどのような発信方法をすることによって大衆の心を掴むことが可能になるのかという事は、「マスメディア」にとってとても重要な課題となって

いる。すでにマスメディアもソーシャルメディアを利用した情報発信というものには着手している。具体的には、ニュースを報道しながらリアルタイムでTwitterでのつぶやきを流すものや、特定の期間に多く投稿された内容に対するニュースを報道するという類の発信形態もとっている。

1.1.2 情報伝達媒体としてのソーシャルメディアの応用

ソーシャルメディアは、大衆の利用だけでなく大衆に影響を与えようとする利用者も多い。政治におけるソーシャルメディアの利用も盛んに行われており、アメリカでは2008年大統領選挙時のObama元大統領や2016年大統領選挙において炎上を多発させたTrump現大統領、そして日本においては2016年の東京都知事の選挙で当選した小池百合子都知事も選挙戦にTwitterを有効活用している。政治においてソーシャルメディアを活用する際、投稿によってユーザに想定していた感情を持たせることが重要となってくる。そのため、広報活動の一環として使用しているソーシャルメディアの投稿に対する反応や感情の変化を分析することは、昨今の選挙戦術の中でも重要な課題となっている。

1.1.3 選挙においてレジリエンスを損なう事象の発生とソーシャルメディア

ソーシャルメディアが発達して情報発信が双方向性を持っている一方で、選挙を含む政治においては、レジリエンスを損なう事象が発生する。例えば、デマの発生や過激派の誕生などがレジリエンスを損なう事象の例としてあげられる。ソーシャルメディアが、このような事象のメカニズムを内包しているのではないだろうか。例えば、ソーシャルメディア上では意見の偏りであるEcho chamber[1]が発生する。

1.1.4 大統領選への適用可能性と意見形成過程の発見

ソーシャルメディア上での意見の偏りは上述のように政治の面でも確認されているが、それではその形成過程についてはどうであろうか、そもそも、ソーシャルメディア上の話題は流動的なものである。特に情報拡散性の強いTwitterにおいては、コミュニティの中で話題とされる内容は流行のように変化を続ける。それは政治的な分野においても同様であろう、話題の変化とともに利用者の感情などの要素も時事刻々と移り変わっていく。Twitterはリアルタイムなソーシャルメディアであるので、他のソーシャルメディアなどと比べてもより如実に意見形成を観察できる。

1.1.5 意見形成の指標としての感情分析

意見は感情という極性を伴うものである。この感情の分析について、特にソーシャルメディアの利用者の感情を分析するという手法は近年幅広く研究され始めている分野である。

感情分析を行うことによって、情報に対して利用者がどのような感情を抱いているかを分析すれば、芸能や広告など様々な分野に応用していくことができるためである。特に情報拡散の側面の強いTwitterにおいて、利用者の感情を分析することは炎上の把握や情報拡散のモーデリング、そしてコミュニティのネットワークを分析する上で重要な要素となっている。

1.2 目的

本研究では大規模なショックの前後における、コミュニティ規模・密度の変化と、コミュニティ間の対立構造の深まりについて分析する。具体的には、2016年のアメリカ大統領選挙を取り上げ、ユーザ間のツイートのメンションに基づく会話ネットワークを構築する。ツイートの感情分析からあるトピックに対して切り出した上で、ポジティブなコミュニティとネガティブなコミュニティに分け、以下の分析を行う。

コミュニティ規模・密度の変化

選挙の前後における、ポジティブなユーザ群の割合とネガティブなユーザ群の割合変化から、各トピックに対するコミュニティ規模の変化を分析する。また、コミュニティ内における会話の活発さをコミュニティ密度と定義し、その変化を分析する。

コミュニティ間の対立構造の深まり

コミュニティが異なるユーザ間の選挙前後における会話の増減から、対立構造の深まりについて分析する。会話が減少した場合、コミュニティの対立構造が深まったとみなす。

2016年のアメリカ大統領選挙の最終候補者はヒラリークリントンとドナルドトランプの二者であり、この二者のどちらを支持すべきなのかということやどちらが当選するのかということについては大きな波紋を呼んだ。世論調査によれば終始ヒラリーが優勢を保つ状態であったが、最終結果としてはトランプが勝利するという結果に終わった。この、世論や大勢の予想に反する結果となった大統領選挙について、その前後ではどのようなコミュニティの変化、感情の変化、ネットワークの変化が起こっているのかを分析する。分析に使用するデータは、世論調査より一般大衆の素直な感情が表出すると考えられるTwitterのデータを使用する。本研究では特に、話題にされている分野とその感情のポジティブネガティブの度合いに注目して、大統領選の前後でポジティブなグループとネガティブなグループにおけるコミュニケーションはどのように変化したのか、またポジティブの内部あるいはネガティブの内部似おけるコミュニケーションはどのように変化したのかということについて分析を進める。第2章で概説する先行研究によれば、株式価格にショックが起きた際の投資家の行動は、ネットワーク構造上近い相手との限定的なコミュニケーションに偏り、普段であればコミュニケーションをとっていたネットワークの構造上は距離のある相手とのコミュニケーションが減少する。また、対立構造の変化について研究した論文によれば、対立をしているネットワークにおいては相互の間での接続確率が低いことが示されている。この二つの関連研究から推察される仮説として、大統領選の前後において、コミュニティが分断し、ポ

ジティブとネガティブの間での相互のコミュニケーションが減少するということがあげられる。本研究では大統領選挙において争点となるような話題性のある分野を予め抽出し、そのコミュニティにおける変化について分析する。コミュニティの変化やネットワークの構造の変化について、分野を横断して一般的にポジティブとネガティブのコミュニケーションが減少するとき、大統領選の前後において対立構造は深まったのだと結論づける。

1.3 貢献

本研究の貢献は以下の通りである。

- 大統領選挙における話題性の大きな分野について感情分析を用いてネットワーク構造やコミュニティの変化、特に感情極性で分けたコミュニティの規模変化や密度変化を分析した
- 大統領選の前後では、ほぼ全ての分野においてポジティブの感情をもつグループとネガティブな感情を持つグループの間でのコミュニケーションは減り、対立構造が深まる示した。
- 外交に関する投稿など当事者でない投稿者が多く含まれることが推測される分野や、選挙などといったその言葉自体に対立構造のないものに関しては逆に相互のコミュニケーションが増加しており、共通して持つ特徴が存在することを示した。

1.4 本論文の構成

第2章ではソーシャルメディアや感情分析、ネットワークの対立構造とダイナミクスについての関連研究と本研究の位置付けについて述べる。第3章ではTwitterのデータと感情分析を用いたネットワークやコミュニティの変化分析の手法について述べる。第4章ではアメリカ大統領選挙のトピックとしての特性を把握するためにアメリカ大統領選挙とその候補者について述べる。第5章では提案した分析手法と抽出したコミュニティについて述べる。第6章では実験結果に関する結果を述べる。第7章では実験結果に関する考察を行い、分析の有用性・課題について議論する。第8章では本研究の結論を述べる。

第2章 関連研究

本章では、ソーシャルネットワークの分析と感情分析の利用について関連研究を概説し、本研究の位置付けについて述べる。

2.1 ソーシャルネットワークの分析

本節では、ソーシャルネットワークについてどのような分析がなされているのかについて、特に Twitter を用いた分析やネットワークのコミュニティ分析について概説する。

2.1.1 ソーシャルメディアとしての Twitter 分析

Twitter を用いたソーシャルメディアの研究は、Twitter のサービス開始より盛んに行われている。Java らの研究では、フォローとフォロワーに注目をし、そのフォロワーネットワークの分析を行うことによって、フォローネットワーク上で分離のあるコミュニティにおいては使用意図が異なるということを明らかにした [2]。Twitter におけるネットワークの特徴として、情報拡散の要素を強く持つことを示す研究も広く行われており、H. Kwak らは情報共有のメディアとしての可能性を示した [3]。榎らの研究ではユーザーとツイートの双方をクラスタリングすることによって、情報が伝播においてトピックやユーザによって影響を受けることを明らかにした [5]。Twitter を社会的な事象に反応するソーシャルセンサーと捉える研究を行い、スポーツにおいてはポジティブなユーザの反応が情報の拡散を誘発することも示されている [4]。また、Twitter はリアルタイム性の高いサービスであり、ニュースの情報拡散において Facebook, Twitter, Google Plus といったソーシャルメディアの中で Twitter が最も早く情報を伝えやすいプラットフォームであることが示されており、リアルタイム性という面で Twitter の優位性が明らかとなっている [6]。

2.1.2 ソーシャルメディアのコミュニティ分析

ソーシャルメディアにおけるコミュニティ分析は、ネットワークの構造を用いたものとして、ソーシャルネットワークに注目した丸井らの研究がある [9]。この研究においてはユーザのプロフィール情報やメンション情報からネットワークを構築しコミュニティの抽出をし、それらがプロフィールにおける興味に従うということを示した。Huberman らは相互メンションのネットワークを用いた分析を行い、相互メンションで構築されたネットワークは、

フォローフォロワー関係のネットワークよりもツイート数に強い影響を及ぼすということを示した [10].

2.2 テキストの感情分析

テキストの感情を分析することは、意見を分析することにおいて重要であり、感情分析に関する研究は広く行われている。本節では、テキスト情報の感情分析の手法について関連研究を概説する。

2.2.1 感情表現の分類

テキストに含まれる感情を推定する研究は広く行われているが、その中でも英語を用いたものは広く行われており、その感情表現の分類にも多くの方法がある。興奮－恐怖と快－不快に緊張－弛緩で感情を表現する Wundt の 3 軸モデルや、憧憬－憎悪、歓喜－悲嘆、警戒－驚嘆、激怒－恐怖の 4 つの軸を用いて分類を行う Plutchik のモデルの軸モデルが存在する [7]。その他にも、6 つの第一感情と 26 の第二感情、137 の第三感情に分類を行った Parrott のモデル [8] など、数多くの感情表現が存在する。

2.2.2 短文における感情分析

ソーシャルメディアが普及することによって短文によるコミュニケーションが増加しており、短文テキストの分析は重要度を増している。特に、テキストに含有された感情を推定することは炎上予測において重要な特徴量となる [11]。上述の通り感情分析に使用する感情表現の分類は多岐にわたるが、短文の場合感情を正確に推定する手法は長文とは異なったアプローチが必要である。ソーシャルメディアの中でも文字数が 140 字までと制限のある Twitter は短文の感情分析に数多く使用されており、2010 年にはツイートデータから感情コーパスを自動で作成して感情分析する手法が提案されている [12]。また、学習データの必要なないものとして、2014 年には Hutto らがソーシャルメディアのテキストに適する感情分析の手法を提案している。[13]

この分析手法においては、顔文字やインターネットスラングを含む 900 程度の語彙を収集し、各語の positive/negative の度合いをレーティングする。レーティングは大学レベル以上の学力を持つ人の手によって行われ、感情のレーティングの後に、感情の強さについての分析を行った。感嘆符や大文字などからなる 5 つの特徴量が感情の強度に影響を与えることを発見し、強度を応用することによってそれぞれのツイートの感情を強さまで含めて予測することを実現している。

2.3 ネットワーク変化の分析

ソーシャルネットワークの分析については様々な媒体を用いた数多くの研究がなされている。特に近年注目され始めたのが、ソーシャルネットワーク上のダイナミクスの分析である。この分析は、イベントやショックの前後においてどのようにネットワークの構造が変化するのかを見るものであり、主に社会学に応用されている。ダイナミクスの分析は、株価変動のショック状況下においてネットワーク構造は turtles-up するという Daniel らの研究がある [14]。turtles-up とは、普段のネットワーク構造において距離は近いがタイの弱い部分において、ショック状況下においては強まって閉鎖的なコミュニケーションを行い、距離は遠いが普段強いタイはむしろ弱まりコミュニケーションがなくなる、という構造上の変化である。Daniel らは投資家たちのプライベートメッセージのネットワークを分析対象にし、株価変動のストレス状況下においては人は平常時と異なる行動をとること、ネットワークの構造的に近しいものとより密に連携をとるようになることが示されている。

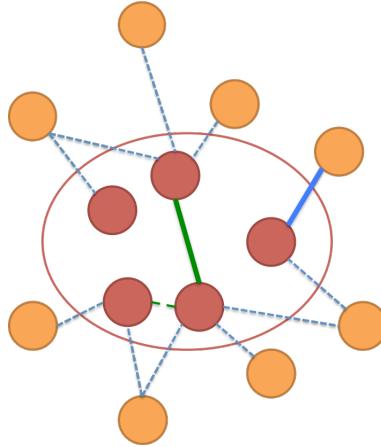


図 2.1: Open network

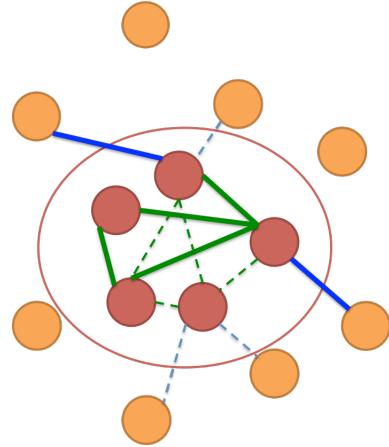


図 2.2: Turtles-up network

オープンネットワークとは、図 2.1 のような状態で閉じていないもので、逆に turtles-up network とは図 2.2 であり、ネットワーク上の関係が近いもので閉じてしまっている状態である。

2.3.1 ネットワーク対立構造の分析

社会には同性結婚や中絶など社会的な敵対を誘発する様々な意見の対立構造が存在する。ネットワークの構造については、そのネットワークのモジュラリティを用いるのでは不十分であり、所属するグループ内部への接続の度合いと、他のグループへの接続の度合いを比較することが敵対構造の把握に有用であるという研究がペドロ（英語にする）らによって行われている。

2.3.2 ソーシャルメディアにおける意見の偏り

ソーシャルメディアの使用方法は人によって異なる。ソーシャルメディア上には Echo chamber[1] と呼ばれる意見の偏りや, Silent Majority[15] の存在が確認されている。しかし、発言をしないサイレントマジョリティの動向の観測や、意見の偏りの形成過程についての研究は十分にはなされていない。

2.4 本研究の位置づけ

既存研究と照らし合わせた本研究の位置づけ・有用性は以下の通りである。

- 大統領選のデータに感情の極性を付与して大統領選の前後のデータ比較をすることによって、対立構造の変化が観察できた
- 対立構造が深まらない一部のデータは、題材に対して当事者でないユーザが含まれるためである可能性があることを示唆した

第3章 手法

本章では、感情分析を用いてショック事象の前後における会話ネットワークとコミュニティの構造変化を分析するための手法を提案する。まず初めに、分析手法の全体のフレームワークについて概説し、その後個別の分析について述べる。

3.1 手法の手順

本研究で使用する手法の手順について以下にリストアップする。

- データの収集と前処理の実施
- ツイートに対する感情値の付与
- コミュニティの分割
- ネットワークの構築
- コミュニティに関する分析指標
- ネットワークに関する分析指標

3.2 データの収集と前処理

3.2.1 データ収集対象となるソーシャルメディアの決定

本研究では、大統領選の前後においてのネットワーク構造の変化について感情分析を用いて分析する。大統領選のデータの分析に際して、データの収集にはtwitterのデータを使用した。Twitterのデータを採用するのには以下のようない由がある。

- ニュース性の高い分析対象である選挙について、Twitterというソーシャルメディアはもっともリアルタイムに反応をするメディアである
- 情報拡散の側面の強いTwiiterは話題に対して敏感に反応し、その話題の大きさによってコミュニティの大きさも変化する
- 匿名性があり、より各利用者の素直な感情が表出する
- RTやメンションによるネットワークが存在し、その変化を観測することができる

上記のようなTwitterの特性を利用して、目的となるコミュニティやネットワーク構造の変化及びコミュニティ内のユーザの感情の変化について分析を進める。

3.2.2 データ収集のクエリ設定とデータの収集

データ収集のクエリと、収集したデータのフォーマットについて述べるデータ収集に使用した検索クエリは次のものである。“trump”, “hillary”, “election 2016”, “us president”, “republican”, “democratic”, “tax”, “job”, “immigration”, “foreign policy”, “trade policy”, “refugee”, “climate change”, “gun”, “abortion”, “childcare”これらのうちどれか一つ以上を含むものを収集した。また、収集の際にTwitterのデータから取得したものは

表 3.1: 収集したツイートの個別情報

収集データ	データ内容
id	ツイートの id
created at	作成日時
id str	ツイートの id(String 型)
username	投稿者のユーザー名
text	本文内容
source	ツイートの url

3.2.3 収集したデータの前処理

ノイズ成分を多く含むクエリの除外

上記のクエリでツイートを収集したが、その中には選挙に関連のないツイートがほとんどであるものが複数存在したため、以下のクエリを除外する“tax”, “job”, “refugee”, “climate change”, “gun”, “abortion”, “childcare”

データ重複の削除

収集したデータは、TwitterのAPI取得制限のためもあり、重複するデータが存在する。重複するデータに関しては以下のように処理をした。

- 収集したツイートの個別情報が全て同じものは削除する
- 収集期間に username が変わっているものがごく少数あるが、どちらが主に使用されていた username でも適切にエッジを張れるようその重複については削除せずに保持する
- ツイートが他者と繋がりのないもの(@を持たないもの)を削除する

表 3.2: 使用するデータ量

期間	ツイート数
before	869435
after	1626309

重複データの削除後に残ったデータは 3.2 の通りである。このデータから、本研究では 10/27-11/6 の大統領選前 (before) と 11/10-11/15 の大統領選後 (after) のデータを使用する（選挙當日前後 1 日については、直前直後ということから除外をした）前処理を終えた状態での使用するデータは以下の通りである。

3.2.4 メンション情報のカラムの追加

収集したデータから、各ツイートがどのユーザに対してメンションを飛ばしているのかについての情報を追加する。

- @を含む文字列を抽出
- それぞれへのメンションとして値を保持する

3.3 ツイートに対する感情値の付与

収集したツイートの全てに対して、positive/negative/neutral が全て 0-1 の値で割り当てられ、その合計が 1 となるようにツイートに対して感情値を割り当てる。感情値の付与については、00 (vader の関連論文) で作成された vader という手法を使用する。vader を作成した論文では、以下のステップを踏んで感情がポジティブネガティブであるかの判定と、その感情の強度の判定を行なっている。

- 顔文字やネットスラングを含む様々な言語資源から収録語彙を収集
- 人手を使用してスコアリング
- 感情度の強度を評価する

ここで、二点目のスコアリングに際してその精度を上げるために、

- 大学レベル以上の英語読解能力の人のみ使用
- 事前に sentiment rating のトレーニング
- 前もって分布を計算した特定の語 (golden items) から大きく離れたレートをつける人のデータは外れ値として、使わない
- 大多数の人が選んだレートと一致している場合には報酬を与える

というような手順を踏んでいる。また、感情の強度の分析においては以下の点を考慮してレーティングをしている。

- 感嘆符 (!) の数で評価
- 大文字である語 (good を GOOD として表記するなど)
- 程度を表す語 (extreamy など) の有無
- ”but” の後の文においては極性をシフト

例えば、以下のような文章にはそれぞれ表のように感情値が付与される。

表 3.3: 文章と付与される感情値

text	pos	neu	neg
Yay. Another good phone interview	0.68	0.32	0
Yay! Another good phone interview!	0.70	0.30	0
Yay! Another extremely good phone interview!	0.64	0.36	0
Nooo! I am so sad x(.	0	0.48	0.52
Nooo! I am so sad.	0	0.41	0.59

この感情値の付与については、ツイート本文がポジティブネガティブニュートラルの要素をどれくらい持つのかということについての値の付与であり、特定のトピックに対する賛同と反対などの指標を付与している訳ではないことをここで注記する。

3.4 コミュニティの分割

コミュニティの抽出の方法には、ネットワーク構造を用いたクラスタリングと投稿の内容によって抽出するものの大きく二つの手法が存在する。

ネットワークの構造を用いてクラスタリングするモデルに関してはモジュラリティ最大化 [16] が頻繁に使用される。モジュラリティとは、ネットワーク上で所与の分割に対して、グループ内のノード同士が繋がるリンクの割合からリンクがランダムに配置された場合の期待値を引いた値として定義される。

$$Q = \sum_i (e_{ii} - a_i^2) \tag{3.1}$$

ここで Q はモジュラリティ, e_{ii} はグループ内のノードが繋がるリンクの割合, a_i^2 はリンクがランダムに配置され場合の期待値である。モジュラリティの最大化は計算時間がとても大きいため、その計算時間を短縮する手法も提案されている [17]。ここで、ネットワーク構造を用いてモジュラリティ最大化によってコミュニティの抽出をするとき、その利点はネット

ワーク構造上近しいものでクラスタリングを行うことが可能であるという点である。対して欠点は、Twitterのような情報拡散の側面が強いソーシャルメディアにおいて、特定期間のメンション情報のみを使用してネットワークを構築する場合、うまくクラスタリングをすることができず、ほとんどのクラスタが所属数1となる点である。本研究のデータに対しても同様にクラスタリングを行なった場合、多くのクラスタが自らのみのクラスタとなった。これは取得データが一ヶ月に満たないデータを使用しておりネットワーク構造としては疎であるためである。

コミュニティの抽出方法としてもう一つの大きな方法が、内容を元にコミュニティの抽出を行う方法である。内容を元にしたコミュニティの抽出には、純粋なキーワードによる抽出の他にもLDAという手法などが広く用いられる。LDAはLatent Dirichlet Allocation[18]の略であり、文章中のトピックを分析する手法である。LDAの利点としては、特定の言葉に言及していないなくても、話題が同じものを同一のコミュニティとして抽出することが可能であること、想定されていない話題のグループなどを発見することができるなどがあげられる。対してLDAを使用することではコントロールできない部分があるという欠点もある。具体的には、分割の数を先に指定するため、類似の話題が別のコミュニティへと別れることがあり適切な分割数が所与でないこと、分割される話題を指定することが不可能であるため、欲しい話題を得られない可能性があることなどが欠点として存在する。

以上の特性から、ネットワークとして適切に分割できない点や、予め設定した話題に対する変化を分析するという目的から本研究ではコミュニティへの分割の際に、非常に簡単な方法として大統領選に関係する言葉を選択し、その言葉に直接言及しているツイートの集合をコミュニティとして抽出する。

3.5 ネットワークの構築

各分野の検索クエリに言及している投稿をコミュニティとして抽出してきた後に、メンションによるネットワークの構築を行う。このメンションネットワークでは、その際に、投稿の感情度を元にユーザの感情度やエッジの特性によって属性を付与した

3.5.1 ユーザのポジティブネガティブ属性の付与

コミュニティ内に存在するユーザはそれぞれ一つ以上の投稿をしており、その投稿にはすべてpos/neg/neuの感情値が割り当てられている。ここで、ユーザ*i*の投稿数が N_i であるとする。ユーザの投稿に対して感情値 $s_{pos}, s_{neg}, s_{neu}$ とするとすべての感情値の和は1であるので

$$s_{pos} + s_{neg} + s_{neu} = 1 \quad (3.2)$$

である。ここで、投稿の感情値 s_p を0から1の値で表すために

$$s_p = 1 * s_{pos} + 0.5 * s_{neu} + 0 * s_{neg} \quad (3.3)$$

と計算する。このとき、ユーザの感情度 s_u は各投稿の感情度の平均として考えられるので、

$$s_u = \sum_{j=1}^{N_i} s_{p_j} \quad (3.4)$$

$s_u >= 0.5$ であるものをポジティブユーザ、 $s_u < 0.5$ であるものをネガティブユーザと定義する。

3.5.2 エッジに関する属性付与

ネットワークに貼られるエッジにはポジティブユーザ内のエッジ、ネガティブユーザ内のエッジ、そしてポジティブユーザとネガティブユーザの間のエッジの三種類が存在する。ポジティブユーザ内のエッジを pos-pos のエッジ、ネガティブユーザ内のエッジを neg-neg のエッジ、そしてポジティブユーザとネガティブユーザの間に張られたエッジを interactive なエッジと呼ぶ。

3.6 コミュニティに関する分析指標

コミュニティに関する分析をするための指標は以下のものを使用する。

- 投稿数
- ポジティブユーザとネガティブユーザの数
- ネットワークの密度
- コミュニティ内のユーザ感情度の分布
- 投稿数の推移
- コミュニティの投稿の感情度平均の推移

ここで、ネットワークの密度は、それぞれのポジティブなユーザの数とネガティブなユーザの数をそれぞれ N_p, N_n , pos-pos で貼られたエッジの数と neg-neg で貼られたエッジの数をそれぞれ e_p, e_n とする時に、密度 d_p, d_n は、

$$d_p = \frac{e_p}{N_p}, d_n = \frac{e_n}{N_n} \quad (3.5)$$

で表される。

3.7 ネットワークに関する分析指標

ネットワークの構造の変化について分析するための指標は以下のものを使用する.

- ネットワークのエッジ数とその属性内訳
- エッジの割合
- ランダムにエッジを張った場合との比

ここで、ランダムにエッジを張った場合の導入とその算出方法を述べる.

3.7.1 ランダムに張ったエッジとの比

本研究では、大統領選の前後のデータについてそのネットワーク構造などの変化を分析する。しかし、ネットワークは固定されたものではなく、それぞれの時期に条件を満たす投稿をしたユーザとメンション関係のエッジを使用している。そのため、大統領選の前後のネットワーク構造を比較するために、ランダムに張ったエッジとの比を導入する。

ポジティブなユーザの数を N_p , ネガティブなユーザの数を N_n , 張られているエッジのうち $pos - pos$ のものを e_p , $neg - neg$ のものを e_n , *interactive* なものを e_i とする。このとき、エッジの総数 e_{total} は

$$e_{total} = e_p + e_n + e_i \quad (3.6)$$

であり、ユーザ総数 N_{total} は

$$N_{total} = N_p + N_n \quad (3.7)$$

で表せる。ここでランダムにエッジを張った時の $pos - pos$ のエッジの期待値を E_p , $neg - neg$ のエッジの期待値を E_n , *interactive* なエッジの期待値を E_i とすると、それぞれの期待値は

$$E_p = \frac{e_{total} * N_p * (N_p - 1)}{N_{total} * (N_{total} - 1)} \quad (3.8)$$

$$E_n = \frac{e_{total} * N_n * (N_n - 1)}{N_{total} * (N_{total} - 1)} \quad (3.9)$$

$$E_i = \frac{2 * e_{total} * N_p * N_n}{N_{total} * (N_{total} - 1)} \quad (3.10)$$

で表される。よってランダムにエッジを張った場合との比は、 $\frac{e_p}{E_p}, \frac{e_n}{E_n}, \frac{e_i}{E_i}$ で表すことができる。この値が大統領選の前後によってどのように変化するかを考えることによって、どれだけネットワークの内部構造が変化したのかを観察することができる。

第4章 アメリカ大統領選

本章では、今回の分析の対象としたアメリカ大統領選挙についての背景知識の説明と、今回のアメリカ大統領選挙において注目すべき争点とそれぞれの候補者の立場等について記述する。

4.1 アメリカの政治

本研究では大統領選挙の前後におけるネットワークの構造やコミュニティの変化を分析することを目的とするが、ネットワークの変化やコミュニティの変化は、トピックであるアメリカ大統領選挙の特性に少なからず影響されるものである。そのため本節では、トピックとして取り扱うアメリカ大統領選挙について記述する。

4.1.1 アメリカの政治構造

この小節では、大統領選挙に関連する背景としてアメリカの政治構造について概説する。

アメリカの二大政党

アメリカ大統領選挙を分析するにあたっては、ヒラリートランプの両候補者の主義主張を知ると同等に、その所属政党がどのような政策方針であるかを知ることが必要である。民主党と共和党の政治的方針について簡単に述べる。

共和党とその政策方針

共和党は民主党と並んでアメリカの二大政党を形成する党であり、伝統的な保守層とキリスト教主義の層が支持層となる、政策方針も保守的な政党である。シンボルは象であり、保守派の主な政策方針は、人工中絶反対、銃規制への反対、不法移民への反対など、保守的であり、キリスト教に従うものが多い。

民主党とその政策方針

民主党もアメリカ二大政党の一つで、こちらは逆にマイノリティからの支持が厚く、主にリベラルな政策方針をもつ。そのシンボルはロバであり、宗教多様化への容認や人工中絶の許容、不法移民の容認などが主な政策方針である。

4.1.2 オバマ大統領期から見るアメリカ政治の変遷

この小節では、2016 大統領選挙での争点についてより明確に把握するため、オバマ政権から 2016 アメリカ大統領選挙に至るまでの政策等の変遷を概説する。オバマ政権の以前より共和党支持者と民主党の支持者層は分裂をしており、「2つのアメリカ」と呼ばれるほどであったが、8 年のオバマ政権を経てそれがより深刻化することとなり、結果として 2016 年の大統領選挙とトランプ勝利という結果の足がかりとなっている。

(1) 選挙戦におけるメディアの活用

2008 年の大統領選挙においては、ブロードバンドや携帯電話の急速な発達を背景としてユーチューブやフェイスブックを始めとするソーシャルメディアが取り入れられた。従来の選挙戦術と比べて、効果的に使用することができれば広範に影響を与えることが可能であるのがインターネットメディアの活用である。ソーシャルメディアの特徴は、一方向性のものではなく双方向的であるところにある。ソーシャルメディアを使用することで一般大衆にアピールをすることが可能なだけでなく、大衆の意見等を吸い上げて反映していくことや、支持率をあげる戦略を講じていくことが可能となる。事実、オバマ大統領はソーシャルメディアの駆使によって支持者層を拡大していき、200 ドル以下の個人献金は個人献金総額 7 億 4,600 万円のうち 24 % を占めている。オバマ陣営は選挙の資金を史上最多の額集めることに成功しているが、これもソーシャルメディアの活用が成功した影響を受けている [19]。

4.1.3 オバマ政権における各種政策について

(1) 人種政策

オバマは選挙中より、「脱人種」路線をとっており、人種を特定しての政策を避けていた。この方針はオバマの政権第一期から第二期の途中まで一貫しており、このオバマの脱人種路線において支持率の低下を招くことはなかった。事実、オバマの支持率は黒人からは概ね 80-90 % 台の支持率を集めており、当選当時 58 % の支持率だったものが 30 % までに落ち込んだ白人によるオバマの支持率とは大きく差が開いている。オバマは二期目においては黒人を明確に対象とする政策を複数実施するようになる。

オバマが第二期で人種政策を講じるようになったきっかけとされているのは 2013 年の夏で、2012 年に起きた黒人少年が射殺された事件が不起訴処分となつたことに起因する。不

起訴処分を受けてオバマは会見において自らが受けてきた人種差別に対してもコメントを述べ、アメリカ社会に根付く人種差別に関して言及することとなった。第二期オバマ政権では黒人に関連する諸政策が取られており、その中でも代表的なものは、米国雇用法・刑事司法制度の改革・兄弟の万人イニシアチブ・住宅の公正な供給、の四点である。それぞれの政策がどのようなものかを概説する。米国雇用法はリーマンショック後の不況を極めて高まる失業率に対する政策であり、黒人のためだけの政策ではない。しかしながらその法案の中では黒人を含む特定のコミュニティに対してどのような利益をもたらすのかを具体的な数値とともに示しており、アメリカ全体の失業率7.8%に対して12.7%の失業率となっていた黒人コミュニティに受け入れられた。

刑事司法制度改革においては、コカインに対する刑罰を見直し、人種的不均衡の少ない法律の整備に努めた。第三の政策として取り上げられる「兄弟の番人」イニシアチブは、有色人種の男性及び青年に対して焦点を当てたものである。有色人種は環境要因によって十分な教育が受けられなかつたり犯罪に走りやすい傾向があつたが、官民連携してその環境を変え、機会の平等を保障しようとするものである。

オバマの人種政策として取り上げられる四つ目のものは住宅の公正な供給である。アメリカの社会においては、インナーシティと呼ばれる都市部に低所得層の黒人が、郊外に中流以上の白人がすむという状況が存在していた。低所得層のすむ環境においては教育環境等が十分でなく、治安も不安定であり格差が生じていた。このような制度的差別に対して、オバマは3年から5年おきに居住パターンを人種的分布、貧困率、学校の質などに基づいて報告するよう取り決め、制度的差別の是正に取り組んだ。

オバマ政権における人種への取り組みは細かく分かれているが、人種政策全体としては国益を最優先に考えられており、その上で特定の人種を差別せず不遇にあつているものについては改善を試みるという方針が持たれていた。人種の壁のない社会を作ることにオバマは取り組み、多様性の受け入れられやすい社会構造への変革に大きく貢献したと言える。

(2) 医療政策

オバマはアメリカ史上初の医療政策を行なった。オバマケアと呼ばれるそれは、それまで保険への加入が自由であったアメリカにおいて全市民に対して医療保険への加入を義務付け、また50人以上を雇用する雇用主に対してはその保険提供を義務付けるものであった。オバマケアは民間保険が一般に普及しており、所得などの格差によって受けられる医療に差が出るアメリカの社会に一石を投じる政策であったが、保険加入の義務付けはなされたもののその質についての決まりなどについては未完成な部分が多く、低保険者の問題と、医療費や保険料の高騰の問題が残されている。

(3) 外交政策

外交政策におけるオバマの立場は、理想や合意の重要性を強調しつつも現実問題としての軍事力の必要性は否定しないというものであった。オバマは外交政策において「対話」を強

調する姿勢をとったが、それは単に理想を掲げ現実主義から外れているわけではない。オバマの立場は、互いの利益相反を認識した上で共通の目的に向かっての対話を心がけるもので、具体的な政策としては既存の大国間政治にそぐわない辺境からの脅威に対しては、アメリカへの不信を払拭し信頼を獲得する方針、そして国家間の安全保障問題については、従来通り軍事力による抑止に比重をおいた政策をとる、というものである [ポストオバマ]

4.1.4 ポストオバマの選挙戦について

(1) 選挙戦術におけるメディアの重要性

オバマが選挙戦を勝ち取った2008年以降も、インターネットなどのITサービスは急速な勢いで普及した。そのため、より一層ソーシャルメディアを使用した選挙戦術は重要度を増している。ソーシャルメディアを使用した選挙戦術の重要性を把握しているためか、2016年の大統領選挙においてはトランプがTwitterを用いて炎上を数多く起こしている。ヒラリーもそれに対してTwitter上で対抗するなど、両者ともにTwitterの使用を重要視した選挙戦術を取っている。

(2) 各種政策について

オバマ政権は上述のように数多くの政策を実行しているが、それに対して時期大統領候補がどのような政策案を打ち立てるのかという点もまた、選挙戦の注目点となる。

4.2 2016アメリカ大統領選挙の二大候補者について

4.2.1 ヒラリークリントンとその政治的特徴

ヒラリークリントンは弁護士の出身で、オバマの前の大統領であるクリントンのファーストレディでもある。ヒラリーの政治的方針は、党の方針がリベラルである民主党の中でも特にリベラルな立場をとっており、多様性を認める政治方針である。特に宗教的に議論の分かれる人工中絶に対して支持的である。ヒラリーの支持層は特に女性層や都市部の非白人層が多い。外交政策の立場としてはリベラルホークと呼ばれる、軍事力の使用を厭わずに秩序の安定を図るという考え方を持っている。

4.2.2 ドナルドトランプとその政治的特徴

ドナルドトランプはコメディアンであり不動産王とも呼ばれる、芸能とビジネスにおいての大成功者であり、共和党の最終候補者として2016アメリカ大統領選挙を戦い勝利した。彼の政治的特徴は共和党の特徴に従い右翼で、アメリカの国益、中でもトランプの支持層の

中流階級労働者層の利益を追求するものである。不法移民に反対しており、メキシコとの間に壁を作ると発言するほどの過激な発言をも厭わない。

4.3 2016 アメリカ大統領選挙の争点

各争点となる部分について、ヒラリーとトランプの政策方針とその相違点について述べる。

4.3.1 経済政策

ヒラリーは経済政策においては、累進課税の促進や最低賃金の引き上げを掲げており、国内全体の経済力の底上げを図り全体の成長速度を上げることを説いている。TPPにも反対して貿易に関して国内の利益を優先している。対するトランプは経済政策的にも保護貿易的で、TPPにはヒラリー同様反対している。国内の製造業を保護することを特に第一と考えており、そのためには極端な手段を取ることも厭わない姿勢である。

4.3.2 外交政策、国防政策

次に、外交政策についてのトランプとヒラリーの政策案について述べる。トランプは、軍事費を増強して軍隊再建を図り、サイバーの強化も図ると主張する。安全保障や難民政策がアメリカ人の安全を第一とすることを説き、ISISについては強固な姿勢をとって対処することを述べている。対してヒラリーは、特に国防の面において ISISへの具体的な対応策を論じている。パリを始めとするテロの攻撃の恐ろしさを深刻に捉え、テロリスト及びジハズムを打ち負かすための計画を策定していると述べる。具体的な施策は、一つ目が現地での対策として、アラブやクルドに対する支援を強化することである。そして二つ目は、世界のネットワークを強固なものとしてテロに対応すること、金銭や武器などの面においてジハード主義のテロリストに供与するテロネットワークを解体することなどの外部からの対策である。そして三つ目が、国内のテロ対策を強化して、テロ攻撃が発生する前に未然にそれを防ぐことなどを説いている。

4.3.3 医療政策

オバマののこした皆保険制度、通称オバマケアに対してその制度をどのように引き継ぐのか、または撤廃するのかという点についても両候補者においては主張の分かれる部分となっている。ヒラリーはオバマケアを改善するとは行かずとも維持するという方針でいるが、トランプはオバマケアの撤廃を主張している。

4.3.4 右傾化

トランプの政策は何よりもアメリカを第一に考えるものである。そのことは彼のスローガンである、”Make America Great Again”からも見て取れる。アメリカを第一に考えるトランプの政策方針の中でも特に保守的であるのが、人種問題、生の価値観、貿易問題、宗教問題である。トランプはいずれに対しても保守的排他的な姿勢をとっている。対してヒラリーはそのような問題に対してリベラルな立場をとっており、人種の多様性や不法移民についても寛容な立場である。

4.3.5 階級闘争

アメリカが多様性のある社会であるのは人種だけではない。階級という面においてもアメリカは幅広い階級の人々がおり、貧困層と富裕層の格差は問題となっている。トランプはアメリカの産業を支えている製造業を始めとする中流労働者層の活力増強を説いていた。ヒラリーがエリート層やリベラル層の支持を集めてトランプ支持者層を嘆かわしい人々の集まりと称したのに対して、トランプは反エスタブリッシュメントを主張してアメリカの中流労働者層の支持を集めていった。

4.3.6 雇用創出

大統領選挙の争点の一つには雇用の創出という問題もある。トランプはアメリカの中間層の底上げを図ることを政策としており、10年間で2500万人の雇用を創出するという主張をしている。トランプが雇用の創出のために重要視しているのが、外国への雇用の流出の防止である。特にトランプはアメリカの中間層を支えている製造業の外国への雇用流出を懸念しており、国内の法人税率を35%から15%へ引き下げることで対処するとの主張を行なっている。インフラ事業への投資や貿易赤字の削減、規制緩和などを駆使して国内の雇用を確保するというのが雇用に関するトランプの政策方針である。対してヒラリーは中小企業への投資を行うことによって雇用の創出を図ると主張しており、雇用の創出のために投資する分野も先端技術やそれらを用いた製造業、再生可能エネルギーなどをあげている。ヒラリーは富裕層からの税収を財源として用いて、雇用の創出のための職業訓練の機会も創出するとして、推計では1000万人分の新規雇用を生むと主張する。

4.3.7 人種問題

人種の問題に対するヒラリーとトランプの立場を概説する。ヒラリーは、人種の多様性を容認する立場である。頻発する白人による黒人の射殺問題を受けての銃規制に対しても乗り気であり、人種によるプロファイリングを終わらせるような法案を通し、幼少期の教育機会からの改善を主張する。対するトランプはマイノリティが大好きであるとしつつも人種差

別的で、メキシコ人は麻薬密輸犯であり強姦魔であると言った過激な発言もしている。イスラム教徒をアメリカから追い出すとの主張をするなど、人種差別的な姿勢が見て取れる。

4.3.8 移民問題

移民に対する認識という点についても両候補者の主張は異なっている。ヒラリーは、移民に対して寛容な態度をとり、アメリカは移民による国家であるとして壊れた移民制度を修復して行くことを述べる。具体的には、医療へのアクセスの拡大や健康保険の購入を可能にすること、そして帰化費用の軽減による帰化の促進などを掲げている。対してトランプは移民に対して強硬措置を取ることも辞さない立場で、新しい移民規制を確立することを述べる。アメリカ人労働者の雇用を第一に考え、管理されていない外国人労働者の入国は抑制し、合法移民の経済的福利を守ることも同時に目指す方針だ。アメリカを第一に考えるため、不法移民などを防ぐためにメキシコとの国境にメキシコに費用を負担させて壁を作る、という過激な方針を主張している。また、オバマ政権時代に築かれた執行猶予を終了して全ての移民法を適用するとも語っている。ヒラリーが移民に寛容でアメリカの多様性を認める方向であるのに対して、トランプはアメリカの中流労働者を第一優先として移民に対しては厳しい態度で臨むと語っている。

第5章 実験

本章では提案手法による各実験の設定と、その結果について述べる。実験においてデータセットから用意した各ケース、分析・評価方法について説明し、最後に個別の分析結果と結果のまとめについて述べる。

5.1 各ケースのデータ設定

本研究では第3章にて前処理までを済ませたデータの中から分析対象の分野を定め、対応するクエリを含むデータを各ケースのデータセットとして用意した。全体のデータは（検索の）具体的な分野とクエリについては以下に説明する。

5.1.1 分析対象の分野とクエリ

今回の分析対象であるアメリカ大統領選挙においては争点となる分野が複数存在する。これに対して本研究では、あらかじめ争点となるような各分野についてクエリを設定してコミュニティを抽出して分析をする。それににより、分野の中におけるユーザの感情の変化を知ることやその分野におけるネットワークの構造の変化を分析することが可能となる。

例えば、大統領選挙においてどちらの候補が獲得するかわからないスウィング州とされていたオハイオ州について分析をすることを考える。この場合、収集した大統領選挙データ内において”ohio”というクエリを設定することで、オハイオについて言及しているデータセットが得られる。このデータセットについて、大統領選挙の前後においてネットワークにはどのような変化があるのか、言及している呟きの感情値はどのように変化したのかを分析することによって、ネットワーク構造やその密度、感情度の分布がどのように変化したのかという分析を行うことができる。

5.1.2 分析分野とクエリの指定

分析分野は第4章での大統領選挙の争点などを元に策定した。策定した分野について、それぞれ話題性のあるクエリを設定して、その言葉に言及しているか否かによってデータを抽出した。

分析分野とその中のトピック及び検索クエリを示す。

タイトル	クエリ
ヒラリー	”hillary” or ”Hillary”
トランプ	”trump” or ”Trump”
オバマ	”obama” or ”Obama”
民主党	”democratic”
共和党	”republican”
選挙	”election”
移民	”immigration” or ”immigrants” or ”immigrate”
イギリス	”UK” or ”uk”
ロシア	”russia” or ”Russia”
白人	”white”
黒人	”black”
貿易	”trade”
リベラル	”lberal”
ゲイ	”gay”

5.1.3 データセットの設定について

ソーシャルメディアのようなデータからコミュニティを抽出する方法は大きく分けて二つ存在する。一つが、メンションやフォローフォロワーといったネットワークを基準として分ける方法（論文を引用する），そしてもう一つが投稿の内容によってコミュニティを抽出する方法である。コミュニティのデータセット設定において今回はクエリを使用して，該当クエリをツイートの本文内に含むものを抽出しコミュニティとした。これには主に以下の二つの理由がある。一つ目の理由は，話題を同じくするコミュニティにおける大統領選前後の変化を観測するためである。つまり，コミュニティの抽出は内容によるものを使用するため，である。例えば大統領選の前後でオバマに言及しているコミュニティを分析することを考える。この場合，求めているのはオバマに関する投稿をしているコミュニティであるので，クエリにオバマを設定してコミュニティを抽出する。コミュニティに含まれる全てのデータにはオバマという単語が含まれているので，本文の話題の中にオバマが登場していることは確実である。本研究では，上記の例のように特定の話題に言及しているコミュニティを抽出してそのコミュニティについて，大統領選の前後においてネットワークの構造はどのように変化するのか，投稿者数やその感情度はどのように変化するのかということを分析する。本研究においては，内容ベースによってクラスタリングをすることを目的として，またクエリに直接言及しているユーザの感情値に関して

二つ目の理由は，Twitterの特性に起因する。Twitterは情報伝播の特性の強いソーシャルメディアであるので，本研究のように短い期間における会話情報を元にネットワークを構築するとき，そのネットワークは疎なネットワークとなる。そのため，クラスタリングをかける際に，ほとんどのクラスタが自身のみのクラスタとなってしまうという問題が生じる。

上記二つの理由から，本研究では，クエリを事前に設定しそれに言及しているユーザ群

を、感情の極性によりポジティブなコミュニティ、ネガティブなコミュニティと捉え、そのコミュニティ構造やネットワーク構造の変化を分析する。

5.2 評価方法

各コミュニティの大統領選前後での変化については以下の点を分析の評価として使用した

- 投稿数によるコミュニティの大きさ
- 投稿数の時系列変化
- 感情値がポジティブなユーザの数とネガティブなユーザの数及びその割合
- 投稿の感情値の時系列変化
- メンションネットワークの性質
- ポジティブコミュニティとネガティブコミュニティでのネットワークの密度
- ランダムネットワークとの比較

それぞれの指標の計算方法は、3.6にて詳述されている。

第6章 結果

6.1 大統領と大統領候補者

6.1.1 ヒラリー

表6.6では、ヒラリーに言及したコミュニティの情報を示している。コミュニティの大きさは話題自体の大きさを表すが、大統領選の前後でヒラリーへの言及は210135から153269へと変化した。また、コミュニティのユーザの感情度については、大統領選の前はポジティブの割合が0.60と過半数を占めるのに対し、大統領選の後ではポジティブの割合が0.49と割合が逆転している。表??ではネットワークの情報を示している。大統領選の前後ではエッジ数が46397から8798と大きく減少しているが、その大きな要因となっているのはどちらも1/7程度に減少しているneg-negのエッジとinteractiveなエッジである。ネットワークの密度としては、大統領選の前では0.49であったネガティブ内の密度が、大統領選後のデータでは0.05と大きく小さい値へと変化した。エッジの割合については大きな変化は観測されないが、エッジ数ポジティブユーザ数ネガティブユーザ数が同一であるものに対してランダムにネットワークを張った場合との比の値では、ポジティブの割合が0.53から1.52へと三倍ほどまで増加したのに対して、ネガティブの割合が2.99から1.56へとほぼ半減しており、ポジティブとネガティブの間のインタラクティブなエッジについても減少している。

グラフからはユーザの感情などの分布とコミュニティ内の投稿の感情などの日毎の時系列変化、そして投稿数の推移が得られる。ユーザの感情値の分布としては、0.5から0.7のものの割合が増えており、時系列変化としては投稿数が選挙の当日地点に近いほど多くなっている。

表 6.1: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	210135	68598	45560	0.60	0.40
after	153269	65450	68923	0.49	0.51

表 6.2: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	46397	8974	22097	15326
after	8798	3178	3616	2004

表 6.3: ネットワークの密度

時期	pos	neg
before	0.13	0.49
after	0.05	0.05

表 6.4: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.19	0.48	0.33
after	0.36	0.41	0.23

表 6.5: ランダムネットワークとの比

時期	pos	neg	interactive
before	0.53	2.99	0.69
after	1.52	1.56	0.46

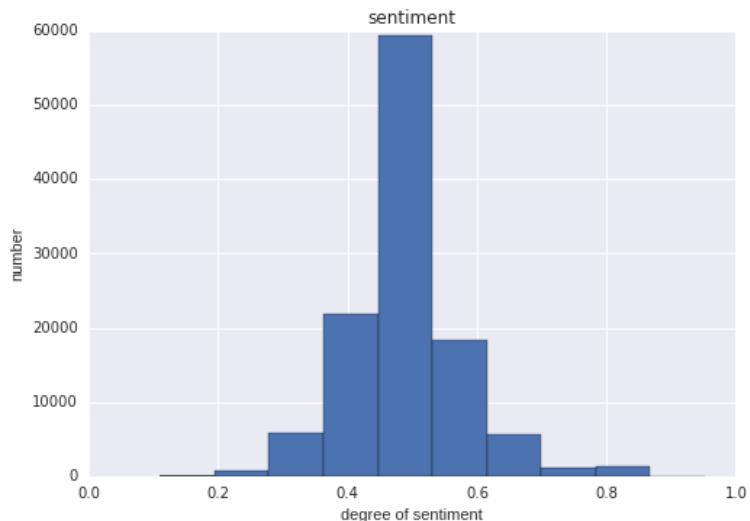


図 6.1: 大統領選以前のコミュニティ内のユーザ感情度分布

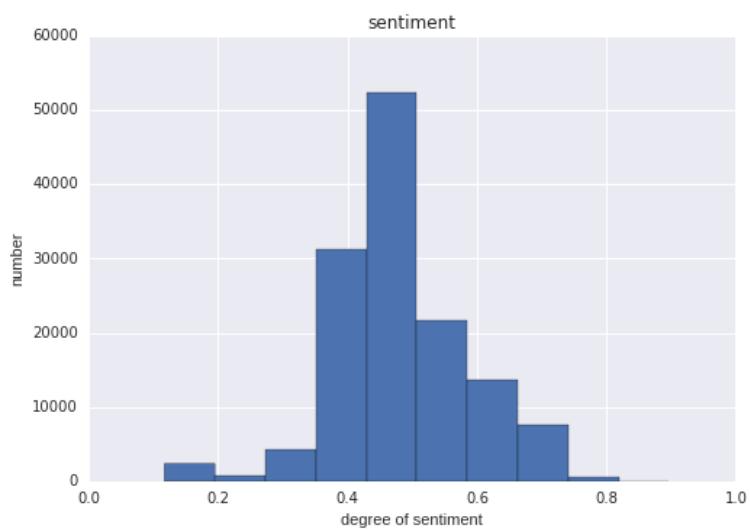


図 6.2: 大統領選後のコミュニティ内のユーザ感情度分布

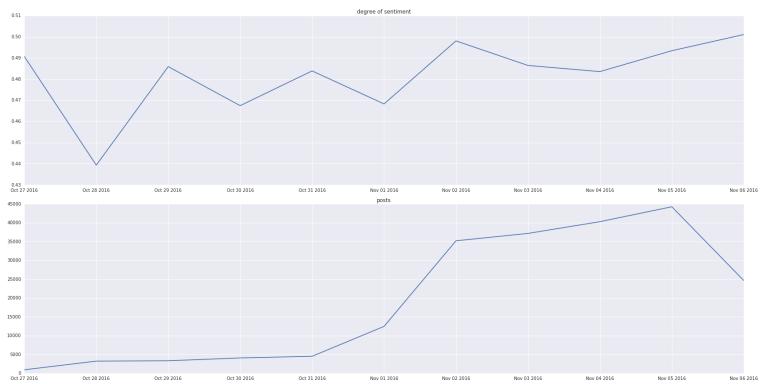


図 6.3: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

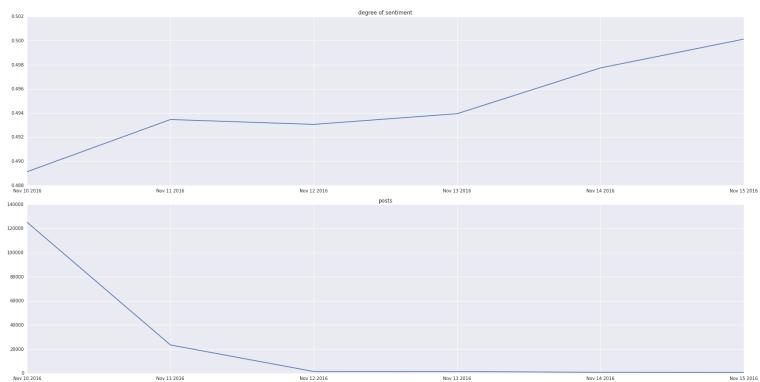


図 6.4: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.1.2 トランプ

次に、トランプをツイートに含むコミュニティに関する結果を示す。トランプのコミュニティ抽出のクエリは'trump' or 'Trump' である。トランプにおいては、話題の大きさを表す投稿数が大統領選の前では 535032 であったのに対して大統領選の後では 1253356 と倍以上に大きくなっている。そのコミュニティの中でのポジティブユーザとネガティブユーザの割合は、大統領選の前後で多少拮抗はしたものの中のポジティブユーザが過半数を占めている。ネットワークの密度に注目すると、ポジティブの密度は半減しており、さらにネガティブの密度については 0.36 から 0.06 へとほぼ $1/6$ にまで小さくなっている。ノードに対してランダムにネットワークを張った場合と比較した値においては、大統領選の前においてはネガティブな割合が 2.57 と大きくあったのが大統領選の後には 1.31 にまで半減しており、逆にポジティブな割合は大統領選の前からほぼ倍増している。ポジティブとネガティブの間のコミュニケーションを表すインタラクティブなエッジの比については、0.45 から 0.36 と減少している。

グラフ 1 からはコミュニティ内の感情度の分布とその推移が得られる。ユーザの感情度合いの分布を見ると、トランプに関しては大統領選挙前の方が 0.5 を超えているもの分布数が多くあるのが判明する。

表 6.6: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	535032	190307	96770	0.66	0.34
after	1253356	498296	366468	0.58	0.42

表 6.7: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	119553	37057	34903	47593
after	88427	39503	20834	28090

表 6.8: ネットワークの密度

時期	pos	neg
before	0.19	0.36
after	0.08	0.06

表 6.9: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.31	0.29	0.40
after	0.45	0.24	0.32

表 6.10: ランダムネットワークとの比

時期	pos	neg	interactive
before	0.70	2.57	0.89
after	1.35	1.31	0.65

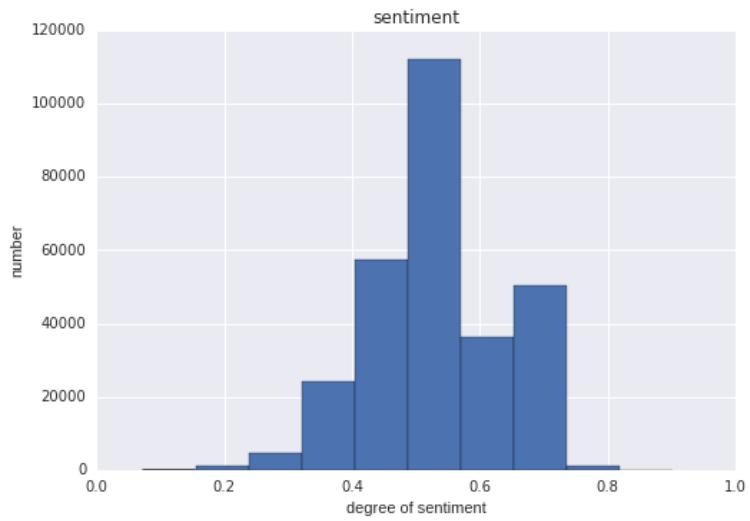


図 6.5: 大統領選以前のコミュニティ内のユーザ感情度分布

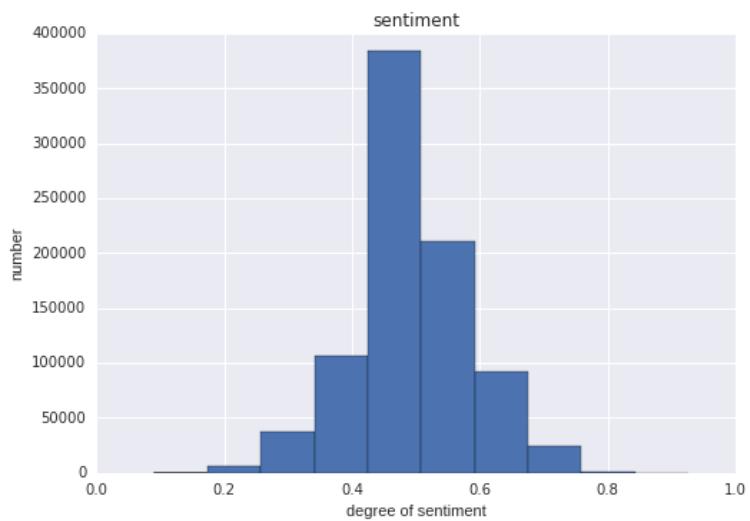


図 6.6: 大統領選後のコミュニティ内のユーザ感情度分布

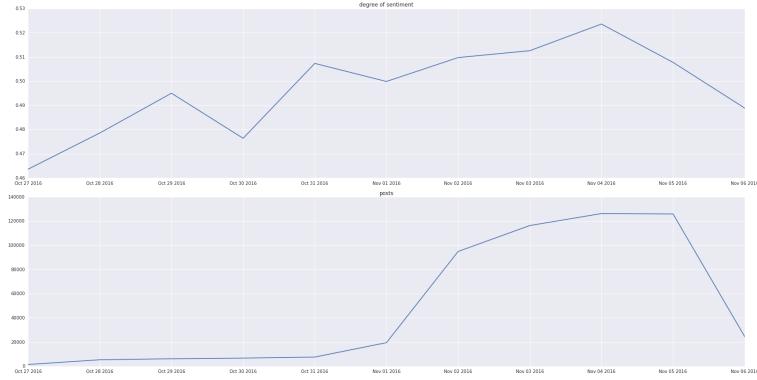


図 6.7: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

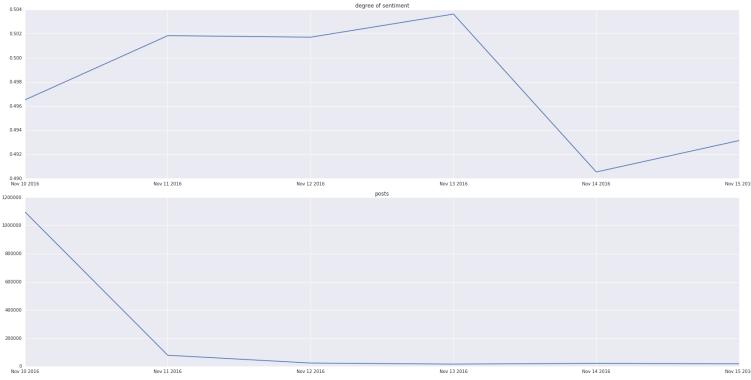


図 6.8: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.1.3 オバマ

全アメリカ大統領であるオバマについて言及しているコミュニティを分析する。オバマに関するコミュニティは”obama”もしくは”Obama”を含むツイートを抽出した。どれだけ話題に出ているかについての指標であるコミュニティの大きさの大統領選前後についての変化を見ると、大統領選の前の投稿数が 19278 であるのに対して大統領選後の投稿数は 77796 と四倍近くまで増加している。コミュニティ内のポジティブなユーザとネガティブなユーザの数については、どちらも大きく増加しているが特にネガティブなユーザは 6711 から 42834 と 6 倍以上の大きさへと増加している。また、それに伴いコミュニティ内のポジティブユー

ザとネガティブユーザの比率は、大統領選以前はポジティブユーザが多かったのに対して大統領選後はネガティブユーザが過半数となり逆転をしている。ネットワークの構造についての分析結果を見ると、ネガティブであるもののネットワーク密度が 0.10 から 0.01 へと $1/10$ までに小さくなっている。そして、同一のポジティブユーザとネガティブユーザに対して同数のエッジをランダムに張った場合との比を見るとポジティブは 1.44 から 4.02 まで増加し、ネガティブは 1.67 から 0.44 まで約 $1/4$ にまで減少している。そしてポジティブとネガティブの間でのエッジに関しては、0.45 から 0.36 と減少している。

オバマのデータについて感情度の分布を見てみると、大統領選以前にはポジティブなユーザ及びポジティブに近いネガティブユーザの数が多くあったが、大統領選後には、大きくネガティブであるユーザの多い分布であることがわかる。

表 6.11: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	19278	9647	6711	0.59	0.41
after	77796	29662	42834	0.41	0.59

表 6.12: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	2342	1171	657	514
after	3756	2529	576	651

表 6.13: ネットワークの密度

時期	pos	neg
before	0.12	0.10
after	0.09	0.01

表 6.14: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.50	0.28	0.22
after	0.67	0.15	0.17

表 6.15: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.44	1.67	0.45
after	4.02	0.44	0.36

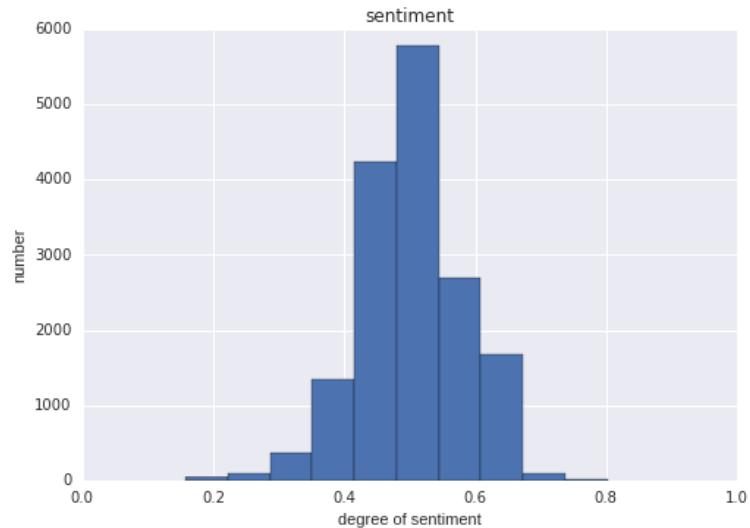


図 6.9: 大統領選以前のコミュニティ内のユーザ感情度分布

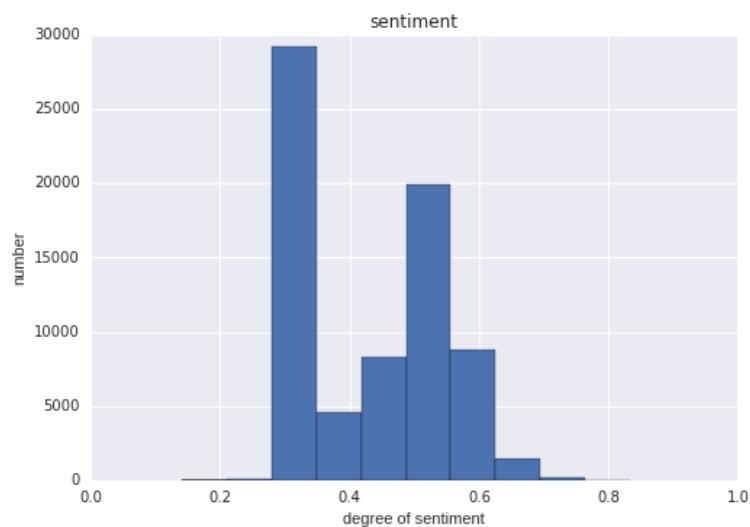


図 6.10: 大統領選後のコミュニティ内のユーザ感情度分布

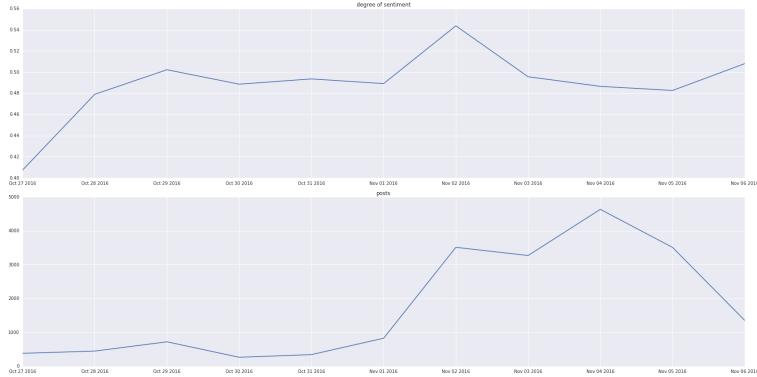


図 6.11: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

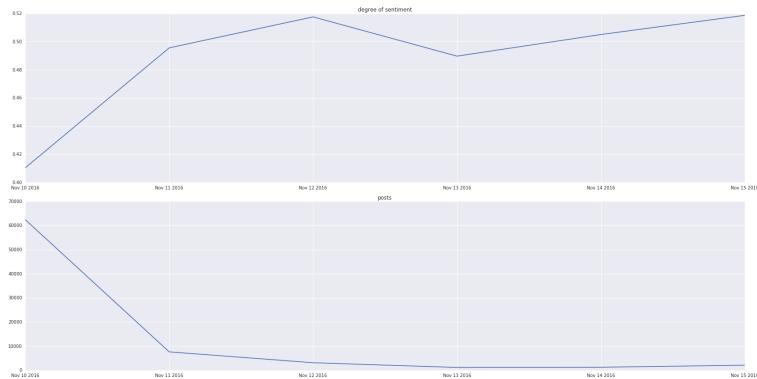


図 6.12: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.2 選挙関連

6.2.1 民主党

民主党に関して言及しているものについてのコミュニティを考える。大統領選の前後で、コミュニティの大きさは 111793 から 181867 へと大きくなっている。コミュニティ内のポジティブなユーザーとネガティブなユーザーの割合についてはほとんど変化をしておらず、過半数はポジティブなユーザーで占められている。ネットワークの構造としては、大統領選後のポジティブなユーザーからポジティブなユーザーへのエッジが大統領選前の倍以上に増加しており、

ポジティブなものの密度も大きくなっている。逆にネガティブ内の密度はほぼ半分に減少している。エッジをランダムに張った場合との比較を見ると、ネガティブなものの割合が半減しており、ポジティブとネガティブの間でのツジの割合は 0.56 から 0.59 へとわずかであるが増加している。

グラフからはユーザの感情度の分布がどのように変化したのかを見ることができるが、民主党に関するコミュニティに所属するユーザの感情度の分布の傾向はほとんど変化していない。投稿の感情値については、選挙前になるに従い感情値がネガティブな方向へ下落していくことが見て取れる。

表 6.16: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	111793	50749	32458	0.61	0.39
after	181867	76903	57332	0.57	0.43

表 6.17: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	13715	5759	4326	3630
after	22045	11455	4275	6315

表 6.18: ネットワークの密度

時期	pos	neg
before	0.11	0.13
after	0.15	0.07

表 6.19: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.42	0.32	0.26
after	0.52	0.19	0.29

表 6.20: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.13	2.07	0.56
after	1.58	1.06	0.59

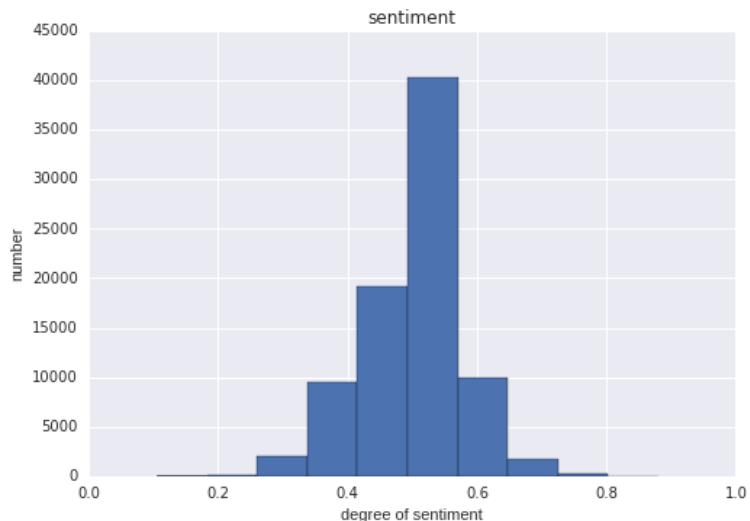


図 6.13: 大統領選以前のコミュニティ内のユーザ感情度分布

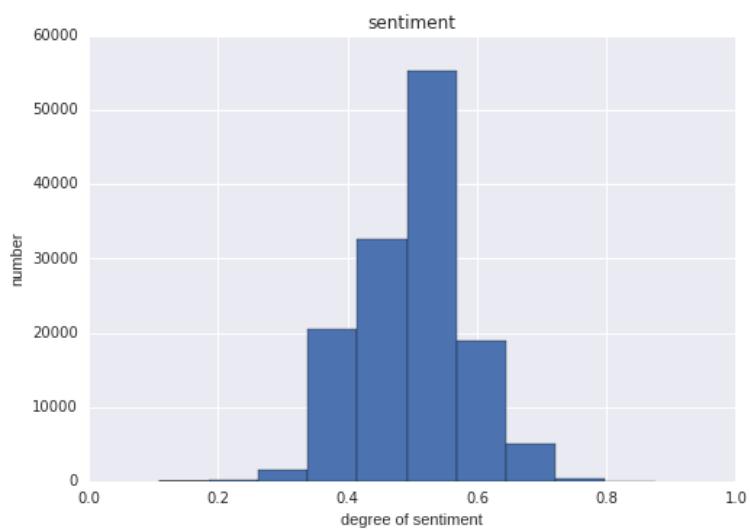


図 6.14: 大統領選後のコミュニティ内のユーザ感情度分布

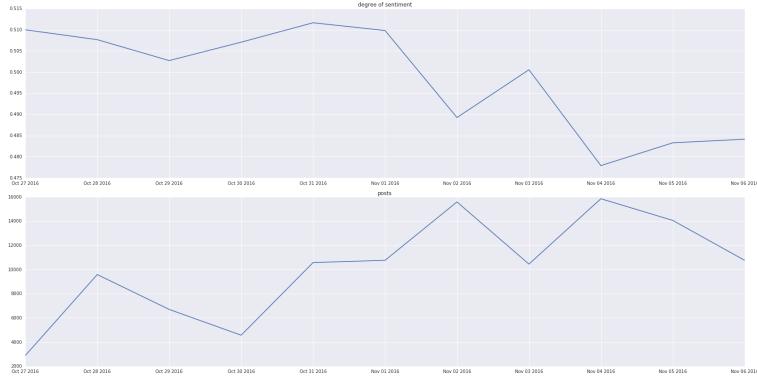


図 6.15: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

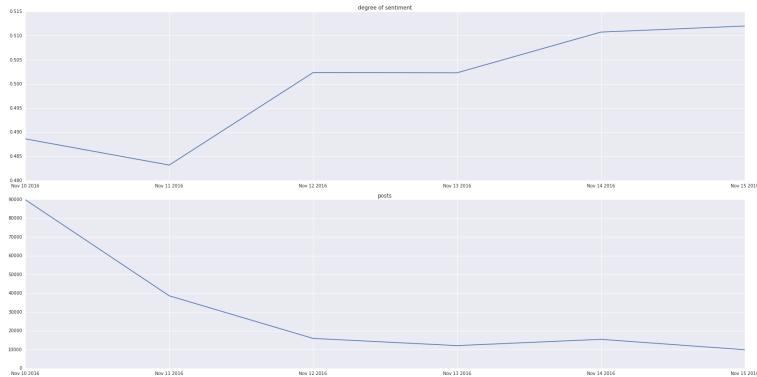


図 6.16: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.2.2 共和党

共和党に関して言及しているコミュニティを分析する。まず、大統領選の前後でのコミュニティの投稿数に関しては 54933 から 76649 と変化しており、コミュニティ内のポジティブなユーザは 266134 から 48014 へとほぼ倍増している。コミュニティ内でのポジティブユーザとネガティブユーザの割合には大きな変化がなく、大統領選の前後どちらにおいても 6 割以上がポジティブなユーザであった。コミュニティの密度に注目すると、ポジティブなもの の密度は 0.16 から 0.08 に、ネガティブなもの密度は 0.19 から 0.06 にと両方疎に変化している。エッジをランダムに張った場合との比の変化についてみてみると、ポジティブなもの

は比率1少々を保持しておりネガティブは前後共に比率は2台でありほとんど大きな変化はないがポジティブとネガティブの間に貼られているインタラクティブなもエッジの比は0.62から0.46へと減少した。

グラフから共和党に関するコミュニティ内のユーザの感情度の分布と投稿数及び投稿の感情度平均の時系列推移を見てみる。大統領選の前後で、ユーザの感情値の分布はポジティブなものの割合が増えていることが見て取れる。

表 6.21: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	54993	26134	15162	0.63	0.37
after	76649	48014	20123	0.70	0.30

表 6.22: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	9835	4130	2852	2853
after	6116	3700	1246	1170

表 6.23: ネットワークの密度

時期	pos	neg
before	0.16	0.19
after	0.08	0.06

表 6.24: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.42	0.29	0.29
after	0.60	0.20	0.19

表 6.25: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.05	2.15	0.62
after	1.22	2.34	0.46

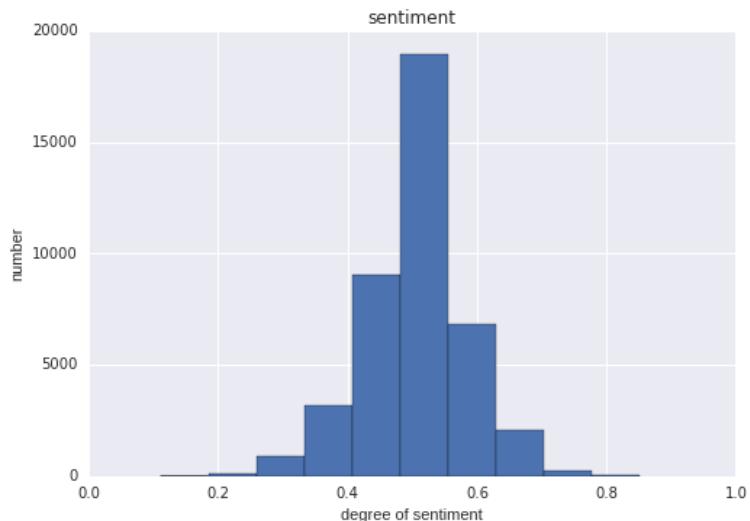


図 6.17: 大統領選以前のコミュニティ内のユーザ感情度分布

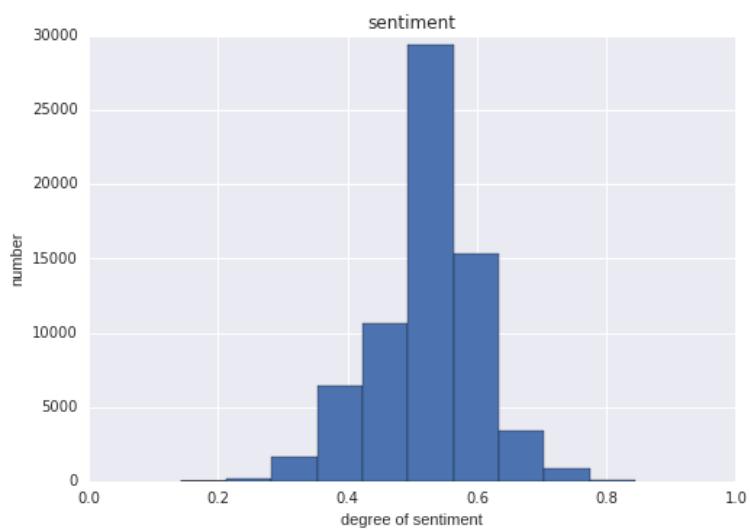


図 6.18: 大統領選後のコミュニティ内のユーザ感情度分布

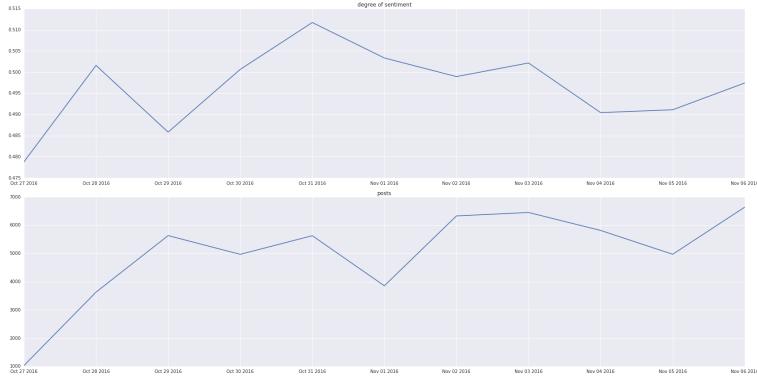


図 6.19: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

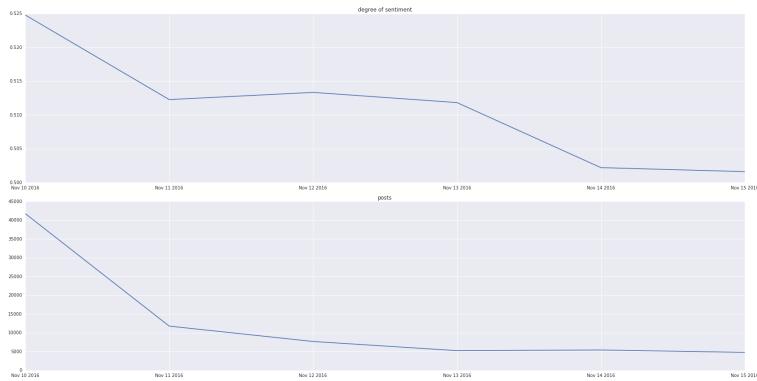


図 6.20: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.2.3 選挙

大統領選について言及しているコミュニティについて分析をする。コミュニティ内の投稿数は大統領選の前後で 25921 から 124594 へと 5 倍程度に増加している。コミュニティ内のポジティブユーザとネガティブユーザの比率に大きな変化はなく、大統領選の前後どちらにおいてもポジティブなユーザが過半数を占めている。ネットワークに貼られたエッジについてはポジティブなユーザからネガティブなユーザへのインタラクティブなエッジの数が、279 から 1036 へと大幅に増加している。コミュニティ内のポジティブなものとの密度とネガティブものの密度に関しては大きな変化があり、ポジティブなものは 0.81 から 0.03 へと、

ネガティブなものは 0.16 から 0.02 へと、どちらも大きく疎に変化している。ランダムにエッジを張った場合との比を見てみると、ネガティブなものの比が 4.40 から 1.48 へと激減しており、さらにポジティブとネガティブなものの間に貼られるインタラクティブなもののひが 0.24 から 0.53 へと増加している。

グラフによってコミュニティ内のポジティブユーザとネガティブユーザの分布変化を見てみると、大統領選の前と比べて大統領選の後のデータにおいては感情度が 0.3-0.5 のネガティブユーザの割合が圧倒的に増えていることが見て取れる。また、投稿の感情値の推移は日毎に大きくばらつきのある推移をしている。

表 6.26: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	25921	15700	7130	0.69	0.31
after	124594	59121	41890	0.59	0.41

表 6.27: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	2707	1266	1162	279
after	4011	1952	1023	1036

表 6.28: ネットワークの密度

時期	pos	neg
before	0.81	0.16
after	0.03	0.02

表 6.29: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.47	0.43	0.10
after	0.49	0.26	0.26

表 6.30: ランダムネットワークとの比

時期	pos	neg	interactive
before	0.99	4.40	0.24
after	1.42	1.48	0.53

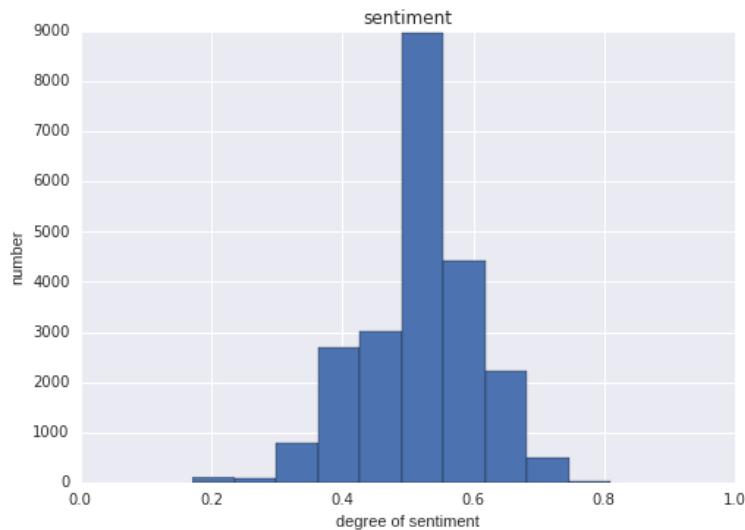


図 6.21: 大統領選以前のコミュニティ内のユーザ感情度分布

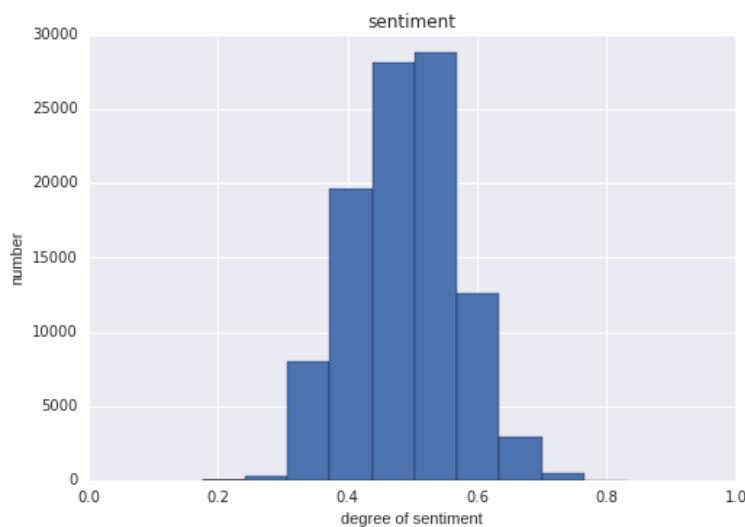


図 6.22: 大統領選後のコミュニティ内のユーザ感情度分布

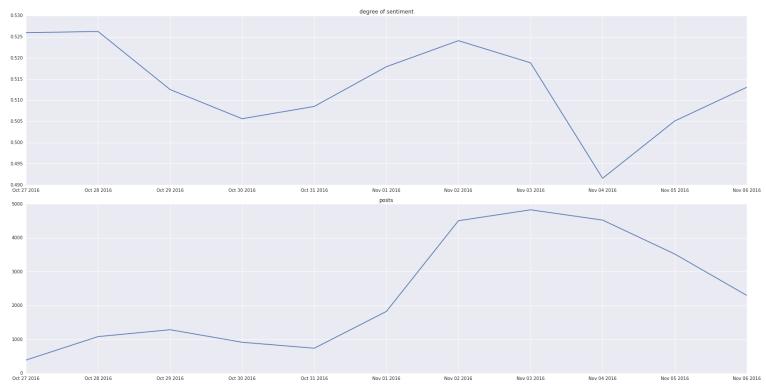


図 6.23: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

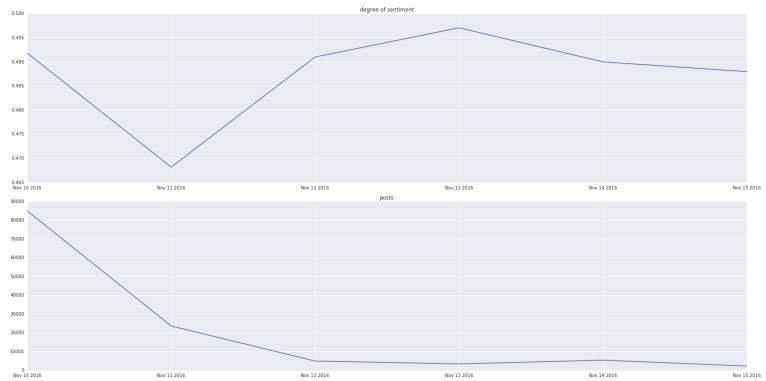


図 6.24: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.3 外交

次に、外交についてのクエリを使用してその構造の変化を分析した結果を述べる。外交関係のデータの特徴は、この分野でのみ、移民や英国のデータにおいて、ポジティブユーザとネガティブユーザの間のコミュニケーション比率が上がったことが特徴である。

6.3.1 移民

次に、移民について言及しているコミュニティを分析する。コミュニティ内の投稿数は 136184 から 194498 へと増加しており、大統領選の前後においてポジティブユーザの割合は 0.47 から 0.55 へと変化しており、ポジティブユーザとネガティブユーザの割合は逆転している。ネットワークのエッジについては、ネガティブなもののエッジ数が 10086 から 5755 へとほぼ半減しており、コミュニティの密度に関してはネガティブなものが 0.23 から 0.09 へと大きく減少している。エッジをランダムに張った場合との比を見てみると、ネガティブなものの比は 1.47 から 1.34 にかすかに減少しており、ポジティブとネガティブの間のインタラクティブなエッジの比は 0.57 から 0.72 へと増加している。

グラフからコミュニティ内のユーザの感情度の分布を見てみると、ユーザの感情度の分布がポジティブとネガティブの境目へと近づいていることが見て取れる。投稿における感情度の推移は、感情度の推移が選挙直前に揺れている状態が見受けられる。

表 6.31: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	136184	37788	43307	0.47	0.53
after	194498	76607	62044	0.55	0.45

表 6.32: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	24011	7147	10086	6778
after	21488	8014	5755	7719

表 6.33: ネットワークの密度

時期	pos	neg
before	0.19	0.23
after	0.10	0.09

表 6.34: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.30	0.42	0.28
after	0.37	0.27	0.36

表 6.35: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.37	1.47	0.57
after	1.22	1.34	0.72

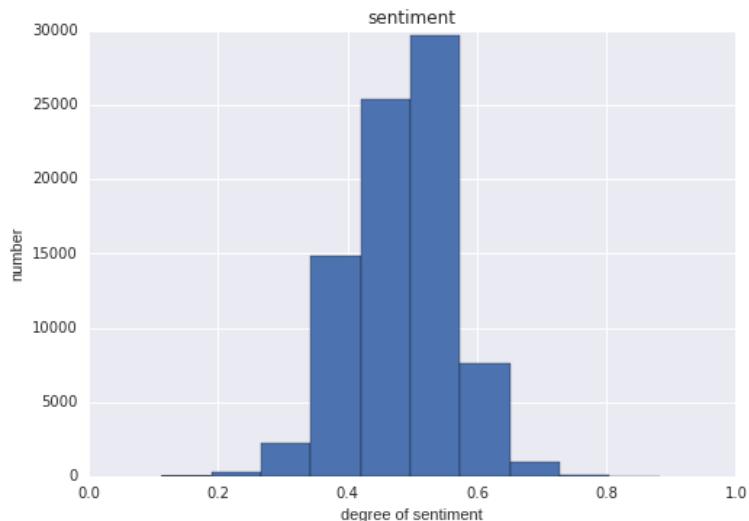


図 6.25: 大統領選以前のコミュニティ内のユーザ感情度分布

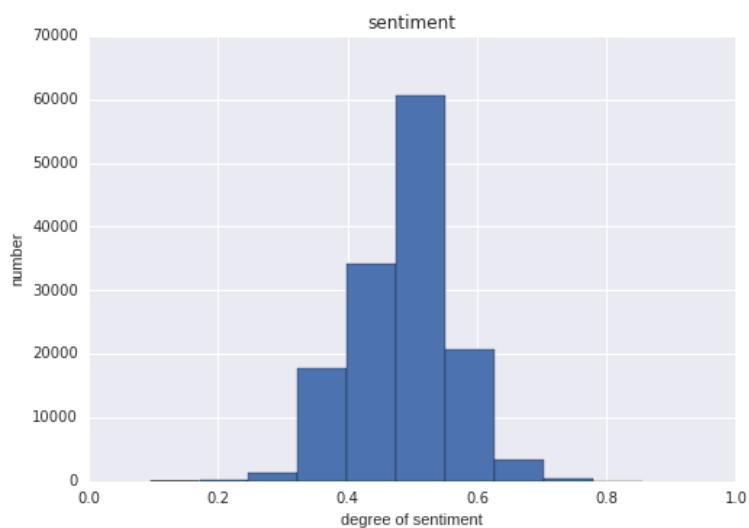


図 6.26: 大統領選後のコミュニティ内のユーザ感情度分布

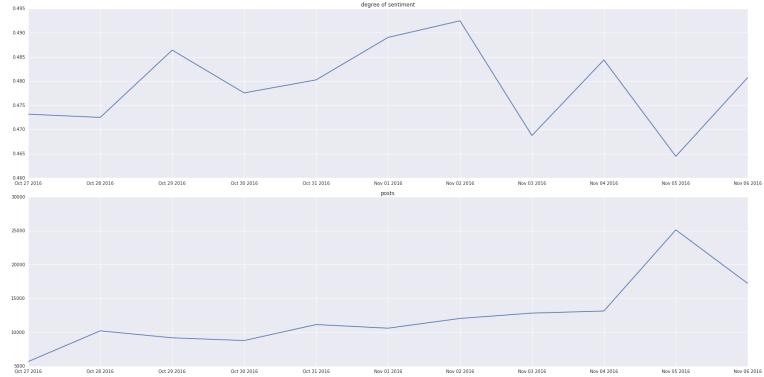


図 6.27: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

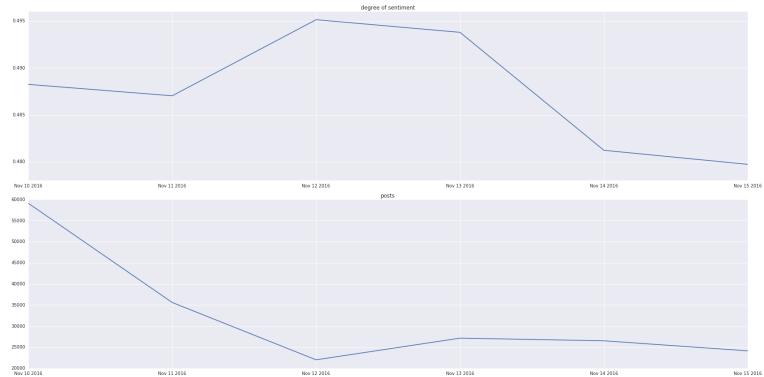


図 6.28: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.3.2 イギリス

次にイギリスについて言及しているものを取り上げる。まず、大統領選の前後でのコミュニティ内の投稿数の変化は 17862 から 13834 であり、やや投稿数が減っている。コミュニティ内のポジティブユーザとネガティブユーザの割合はほとんど変わらず 6 割弱がポジティブユーザである。ネットワークのエッジについてみると、ポジティブなもの同士のエッジは 2884 から 674 に、ネガティブなもの同士のエッジは 1331 から 333 へとどちらも $1/4$ 程度にまで激減している。コミュニティ内の密度については、ポジティブなものが 0.34 から 0.10 へ、ネガティブなものが 0.22 から 0.06 へとどちらも大きく減少している。ランダムに張つ

たエッジとの比については、大統領選の前後において大きな差はないが、ポジティブとネガティブの間のインタラクティブなエッジの比が 0.27 から 0.47 へと増加している。

ユーザの感情値の分布については大きな変化はほとんどなく全体としての形もほぼ同じ状態である。大統領選挙の前に一度小さくイギリスに関する投稿数が増えたときに、感情値が大きく下がっている状態も見受けられる。

表 6.36: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	17862	8361	6177	0.58	0.42
after	13834	7026	5229	0.57	0.43

表 6.37: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	4866	2884	1331	651
after	1309	674	333	302

表 6.38: ネットワークの密度

時期	pos	neg
before	0.34	0.22
after	0.10	0.06

表 6.39: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.59	0.27	0.13
after	0.51	0.25	0.23

表 6.40: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.79	1.51	0.27
after	1.57	1.40	0.47

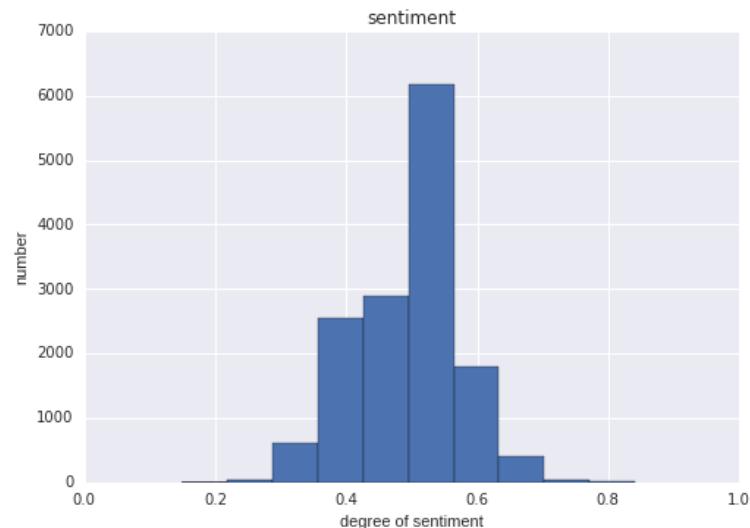


図 6.29: 大統領選以前のコミュニティ内のユーザ感情度分布

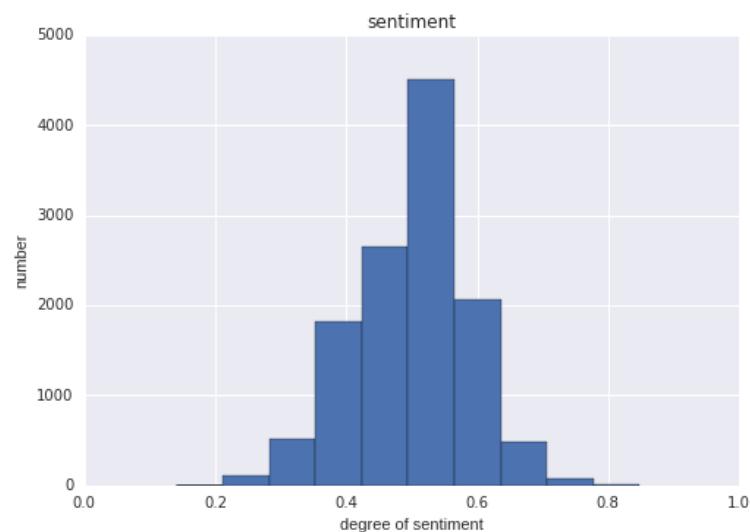


図 6.30: 大統領選後のコミュニティ内のユーザ感情度分布

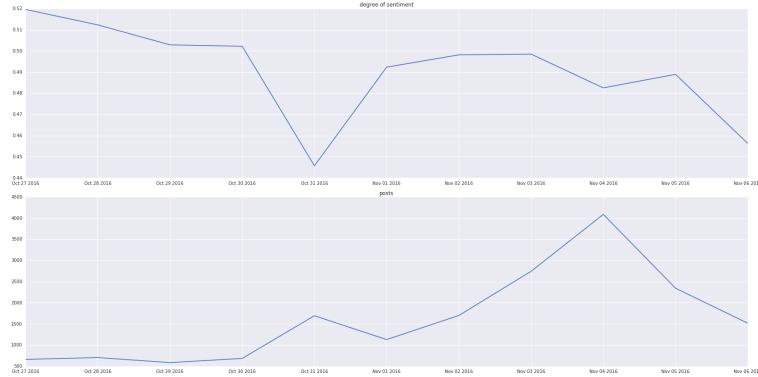


図 6.31: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

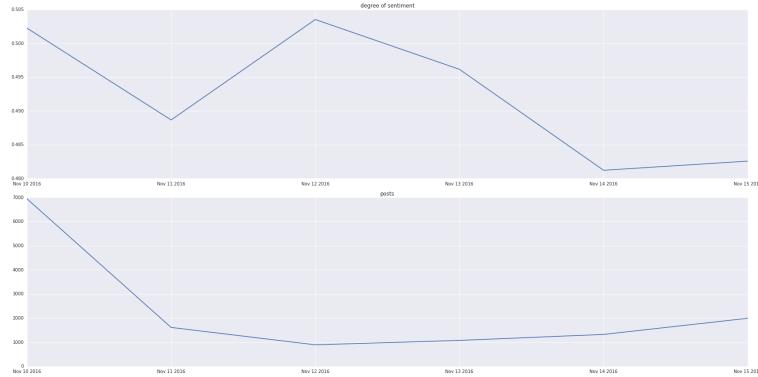


図 6.32: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.3.3 ロシア

外交の最後としてロシアについて言及しているコミュニティを取り上げる。まず、コミュニティの規模、活発さを表す投稿数については、大統領選挙前の 13793 から 5138 へと半分以下に減っている。次にコミュニティ内のポジティブなユーザとネガティブなユーザの割合は、大統領選の前後でどちらもポジティブユーザが過半数であり、さらに大統領選後においては 2 割ほどの割合が増加した 7 割程度がポジティブユーザである状態となっている。ネットワークの情報としてエッジ数に着目すると、ポジティブとネガティブの間で生じるインタラクティブなエッジの数が 472 から 40 と大幅に減っている。ネットワーク内での密度はどう

ちらも大統領選の前に比べて後の方が疎な状態となっている。ポジティブユーザとネガティブユーザに対してランダムに同数のエッジを張った場合との比を分析すると、ネガティブ内のエッジの数がほぼ倍増している。対してポジティブユーザとネガティブユーザの間のインタラクティブなエッジの比率は 0.49 から 0.22 へと半分以下に小さくなっている。

ロシアについてのコミュニティ内のユーザの感情度分布には大きな違いはなく、投稿数も選挙に合わせて増えては選挙後に収束するという推移を見せている。

表 6.41: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	13793	5870	5332	0.52	0.48
after	5138	3420	1379	0.71	0.29

表 6.42: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	1933	790	671	472
after	449	290	119	40

表 6.43: ネットワークの密度

時期	pos	neg
before	0.13	0.13
after	0.08	0.09

表 6.44: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.41	0.35	0.24
after	0.65	0.27	0.09

表 6.45: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.49	1.53	0.49
after	1.27	3.21	0.22

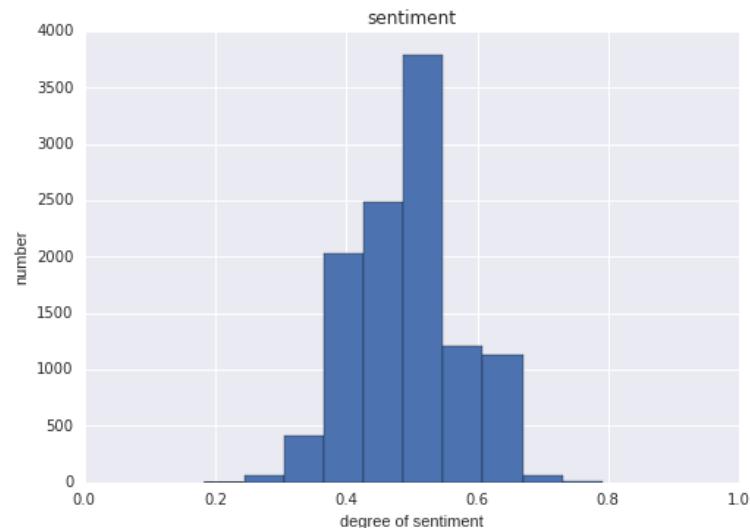


図 6.33: 大統領選以前のコミュニティ内のユーザ感情度分布

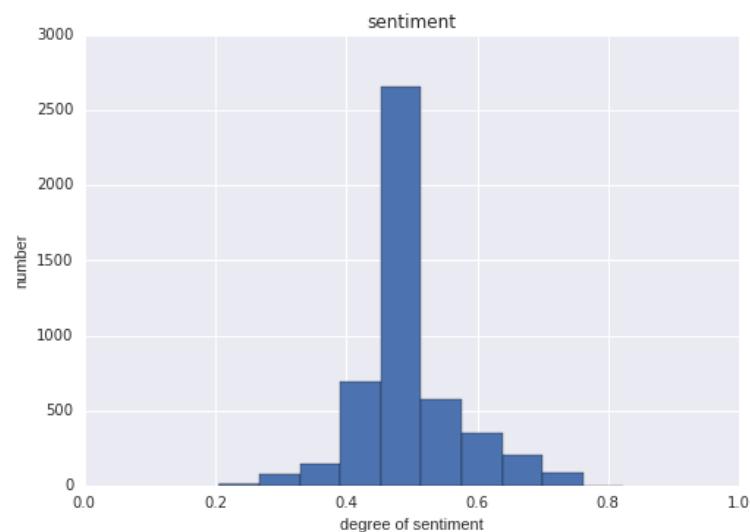


図 6.34: 大統領選後のコミュニティ内のユーザ感情度分布

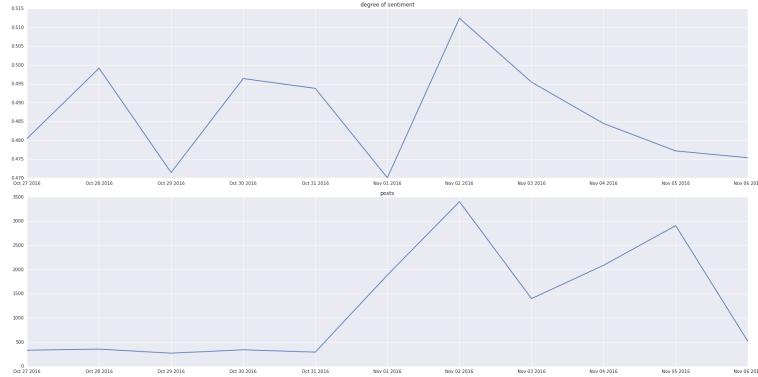


図 6.35: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

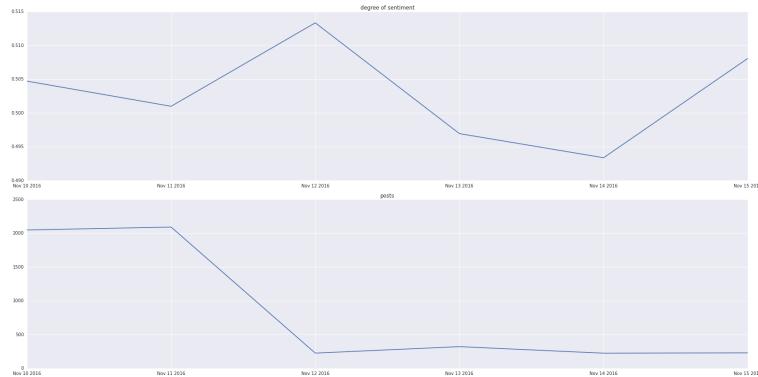


図 6.36: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.4 人種問題

6.4.1 白人

次に人種関係の分野で、白人について言及しているコミュニティを分析する。コミュニティの投稿数は大統領選の前が 12133、大統領選後が 59553 であり、投稿数は 5 倍弱に増加している。コミュニティ内のポジティブユーザとネガティブユーザに関しては大統領選の前後で大きな変化はなく、6 割以上がポジティブユーザで占められている。ネットワークのエッジについては、ポジティブなユーザ内のエッジが 387 から 1502 と大きく四倍ほどに増加

している。コミュニティの密度についてはポジティブなものは大きな変化はないがネガティブなものについては、0.09から0.03と三分の一程度まで減少している。エッジをランダムに張った場合との比についてはポジティブなものの比が0.94から1.70とほぼ倍増しているのに対して、ネガティブなものは5.33から1.68へと大きく四分の一程度まで減少している。ポジティブとネガティブの間のインタラクティブなエッジの比は0.35から0.22へと減少している。

グラフをみて白人について言及しているコミュニティ内のユーザの感情度の分布や投稿の時系列変化等について分析した結果をみると、大統領選以後のデータにおいては、ネガティブユーザの分布が多くなっているものの、もっとも大きなユーザの分布の値に関しては、選挙前の方がネガティブに近く、大統領選挙後の分布では最頻値は0.5から0.6の値へと変化している。

表 6.46: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	12133	8388	2848	0.75	0.25
after	59553	34312	21520	0.61	0.39

表 6.47: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	738	387	253	98
after	2338	1502	582	254

表 6.48: ネットワークの密度

時期	pos	neg
before	0.05	0.09
after	0.04	0.03

表 6.49: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.52	0.34	0.13
after	0.64	0.25	0.10

表 6.50: ランダムネットワークとの比

時期	pos	neg	interactive
before	0.52	0.34	0.13
after	0.64	0.25	0.10

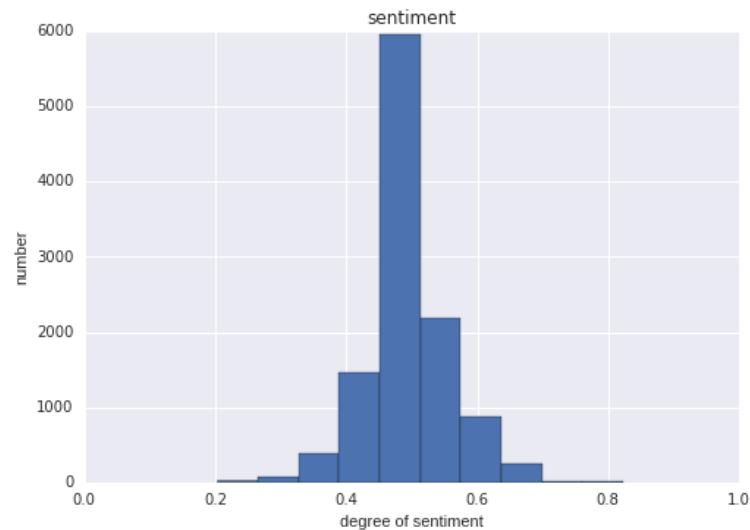


図 6.37: 大統領選以前のコミュニティ内のユーザ感情度分布

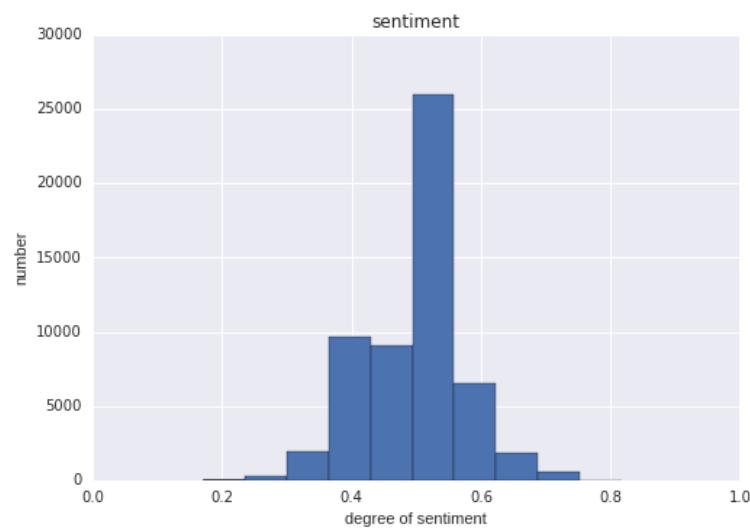


図 6.38: 大統領選後のコミュニティ内のユーザ感情度分布

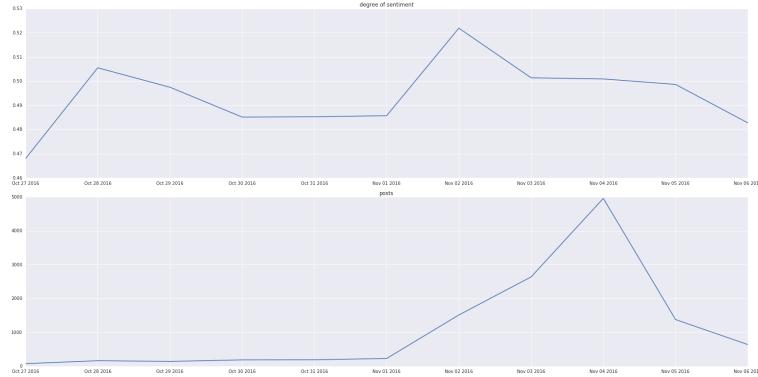


図 6.39: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

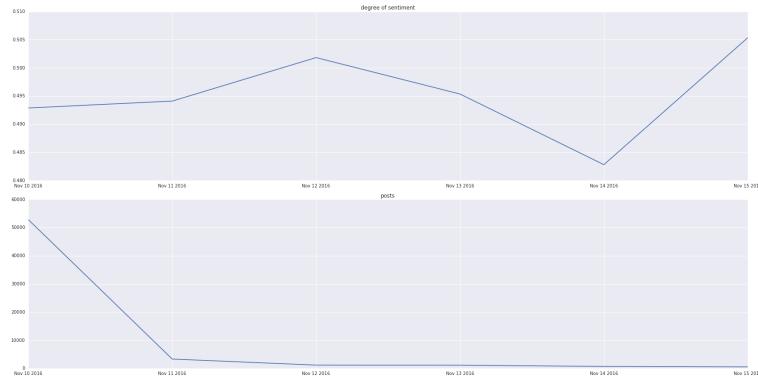


図 6.40: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

6.4.2 黒人

次に人種に関して黒人に言及しているコミュニティを取り上げる。コミュニティ内での投稿数は 11330 から 23882 へと 2 倍程度へ増加しており、コミュニティ内のポジティブユーザとネガティブユーザの比率は大統領選の前後で縮まりはしたもののポジティブユーザが過半数を占めている。ネットワークのエッジについてはネガティブ内のエッジが 280 から 1146 へ増加しているのに対してポジティブとネガティブの間のインタラクティブなエッジは 162 から 65 と大きく減少している。コミュニティ内の密度については、ポジティブ内もネガティブ内も密になっており、これは他の分野には見られない変化である。ランダムにエッジを張つ

た場合との比については、ポジティブとネガティブの間のインタラクティブなエッジの比が 0.43 から 0.09 へと大きく減少することとなった。

感情度の分布についてのグラフを見ると、大統領選挙の前後によって、0.5-0.6 程度の感情度を持つものの割合が増えていることがわかる。投稿数と感情度の推移には大きな相関はない。

表 6.51: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	11330	6898	3220	0.68	0.32
after	23882	12164	10633	0.53	0.47

表 6.52: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	876	434	280	162
after	1534	323	1146	65

表 6.53: ネットワークの密度

時期	pos	neg
before	0.06	0.09
after	0.11	0.11

表 6.54: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.50	0.32	0.18
after	0.21	0.75	0.04

表 6.55: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.07	3.16	0.43
after	0.74	3.43	0.09

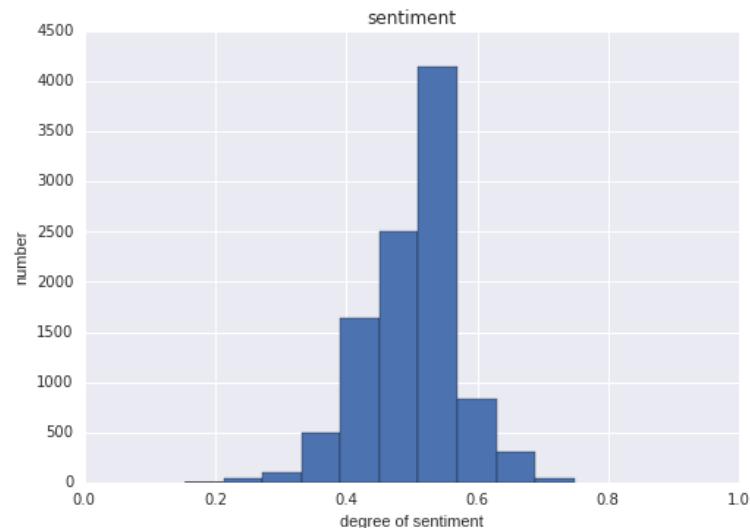


図 6.41: 大統領選以前のコミュニティ内のユーザ感情度分布

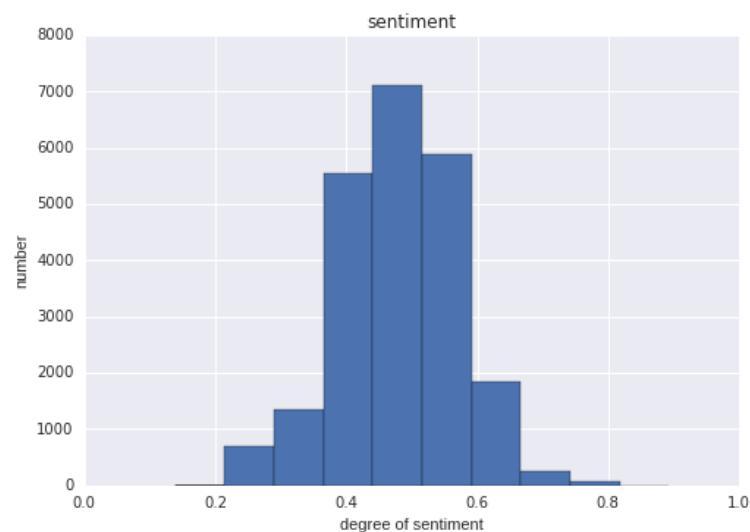


図 6.42: 大統領選後のコミュニティ内のユーザ感情度分布

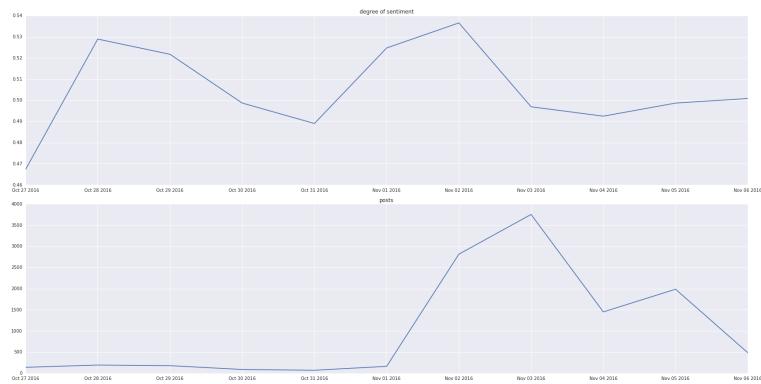


図 6.43: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

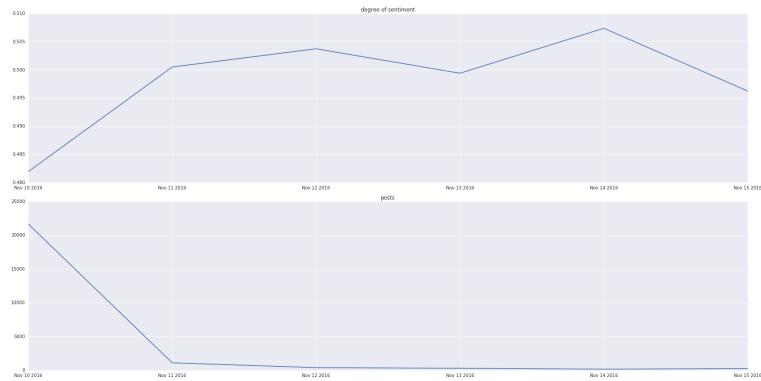


図 6.44: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

第7章 考察

7.1 考察の概要

この章では、コミュニティの規模・密度についての分析とランダムな場合との比を用いたコミュニケーション状態の変化に関する分析について述べる。まず初めに考察で得られた知見をまとめ、次に具体的に分析の考察について詳しく記述する。

考察で得られた大きな知見は以下の三つである。

- コミュニティ規模が大きく変化する時は、結果の意外さにショックを受ける場合の他に、結果が出ることによって自分の立場を発言しやすくなるような場合であることがわかった
- 大統領選の前後において各ネットワーク内のインタラクティブな会話は減少したが、外交分野などの当事者以外を含むトピックに関しては、増加が確認された。
- インタラクティブなコミュニケーションが増えたコミュニティでは同時にネガティブの密度減少とネガティブ内ランダム比がどちらも減少している

7.2 コミュニティ規模・密度の変化

この節では、コミュニティの規模と密度変化についての考察を行う。まず、全体としてコミュニティの規模変化や密度の変化がどのように捉えられるかを概説する。

ポジティブネガティブのコミュニティ規模変化は、大統領選の前後でのそのトピックに関する話題の大きさとして捉えられる。そして、密度の変化については、会話の活発さがどれだけ変化しているのかがわかる。例えば規模として投稿数やユーザ割合が増えていたとしても密度が疎になっていれば、その中の会話や議論が活発に繰り広げられている訳ではなく、より単発の会話などが生じているということを示唆していると言える。

7.2.1 候補者

候補者の分野について考察を行う、ヒラリーとオバマに対してはポジティブなコミュニティとネガティブなコミュニティの割合が逆転して、大統領選後には過半数がネガティブなコミュニティとなっている。対するトランプは大統領選の前後共にポジティブなコミュニティが過半数を占めている状態となっている。密度に関しては、三者のデータとも、ポジティブネガティブ共に減少しており、特にヒラリーとトランプに関しては、両者の値はとても小さ

くなっている、内部での会話の活発さは下がっているということがわかる。これは、トランプとヒラリーについて呴く場合、直接選挙の結果に同時に言及するが多く、そのため大統領選挙後は、選挙結果に対して、議論よりも情報拡散や単発の会話が多くなった、ということが考えられる。

7.2.2 選挙関連

選挙に関するクエリについては、全て共通してポジティブとネガティブの割合は大きく変化することなく、ポジティブなものが過半数を占めていた。これは、選挙関連というクエリ自体が他のワードに比べると抽象的かつメタなものであるため、結果によって大きく構造が変わる訳ではないから、という理由が考えられる。密度に関しては、民主党のポジティブなコミュニティの密度以外は全て選挙後に小さくなっている。

7.2.3 外交

外交に関しては、ロシアとイギリスにおいては大統領選の前後共にポジティブコミュニティが過半数となっているが、移民については大統領選後に逆転して、ポジティブコミュニティが過半数となっている。また、移民とイギリスのデータに関してはポジティブネガティブ合わせた投稿数やユーザ数といったものの値が大きくなっているが、ロシアに関しては小さくなっている。これは、移民とイギリスについては、イギリス人や移民関連で英語話者であるものなどが結果に対して英語で会話をしており、ロシアに関してはその会話をロシア語で行うため収集されなかったと考えられる。密度については三つ全てにおいて減少している状態となっている。

7.2.4 人種

人種については白人と黒人に関して分析をした。こちらは、コミュニティ規模に関しては、どちらも過半数がポジティブコミュニティであったが、両方について大統領選前後でポジティブコミュニティの割合は減少していた。そしてネットワークの密度については、黒人の方ではポジティブネガティブどちらも大きくなっているのに対して、白人の方ではどちらも小さくなっているという変化をえた。これは、人種差別的であるトランプが当選したことによって、将来的に差別を受けるかもしれない黒人に関して感情値としてのポジティブネガティブはあっても議論が生まれていることを示唆する。

7.2.5 コミュニティ規模・密度の変化の考察まとめ

コミュニティの割合については、大きく逆転するものから割合を変えないものまで幅広く存在した。大きく割合を変えるものについてはどのようなものが考えられるかというと、結

果の意外さにショックを受ける分野の他に、結果が出ることによって自分の立場を発言しやすくなるような話題の場合が考えられる。

具体的には、トランプに関するポジティブコミュニティの数の増加や、白人に関するポジティブユーザ及び投稿数の急激な増加が挙げられる。トランプは人種差別的で白人中流労働階級層からの支持が厚かったが、その点に関する投稿が増えている。この点は、ランダムとの比においてトランプについても白人についても割合がほぼ倍近くまで大きくなっていることからも見て取れる。トランプと白人に関しては投稿数がもともと大きいためサイレントマジョリティとまでは言えないが、一定数の反応を示していなかったトランプ支持者が選挙後に喜んでいるという状態の可能性が示唆される。また、コミュニティ内の密度に関しては基本的には小さくなることが多い、それは選挙に関連する話題それ自体が、結果に関する簡単な会話や情報拡散を多く含むようになっているからだということが考えられる。

7.3 ランダム比から見る会話量

この節では、ポジティブユーザとネガティブユーザに対して、ランダムに同数のエッジが貼られた場合との比が、どのように大統領選の前後で変化したかの値について考察する。ランダム比から見る会話量は、特にインラクティブなものとの増減に注目する。ポジティブコミュニティとネガティブコミュニティの間でのコミュニケーションが増加するか減少するかによって、議論や融和が生まれているのか、または対立や極性の近いもの同士でのみの対話が起きているのかということが分析できる。

7.3.1 候補者

ランダム比は、ヒラリートランプオバマのいずれについてもポジティブなものが上がり、ネガティブなものが下がり、インラクティブなものが下がるという結果となった。ポジティブなコミュニティとネガティブなコミュニティの間でのコミュニケーションは小さくなっている。それだけでなく、ネガティブなコミュニティの中における比も下がっており、ポジティブコミュニティとネガティブコミュニティ間のコミュニケーションが減って対立構造が深まっただけでなく、ネガティブな感情を持つものは他とあまり活発な議論をしないで悲嘆しているのではないかということも考えられる。

7.3.2 選挙関連

選挙に関連するものに関して分析したところ、ポジティブ内の比は全てにおいて上がったが、ネガティブ内の比は民主党と選挙に関して下がり共和党に関してのみ上がった、そして、ポジティブコミュニティとネガティブなコミュニティの間のインラクティブなものは選挙と民主党において上がった。民主党のデータはごくわずかの変化であるが、選挙のデータは0.24から0.53に変わる大きな変化である。これについては、選挙という全体への言及で

抽出したものを見ると、選挙後の両コミュニティ間での議論が生まれているということが示唆される。これは、ヒラリーなどといった抽出とは異なり、選挙という言葉自体には支持不支持などの二分化される対立構造がないためと考えられる。

7.3.3 外交

外交において分析した結果、移民、ロシア、イギリス全てについてポジティブ内のコミュニケーションの比は小さくなつた。また、イギリスと移民においてはネガティブ内は小さくなりインラクティブなものは増加したが、ロシアにおいてはネガティブ内のものは大きくなつたがインラクティブなものは小さくなっている。インラクティブなものが外交の分野において増加している原因として考えられるのは、アメリカ大統領選挙の当事者、つまりアメリカ国民でないものの投稿が多く含まれるということである。移民やイギリスについては本国での投稿も英語でされうるので、当事者ではない情報が多く含まれるような場合、当事者のような強い対立構造が存在せずインラクティブな対話が起こりやすいのだと考えられる。対してロシアのデータは、ロシアの国民が投稿する場合ロシア語を使用するため、当事者以外のデータが多く混入することなく、他の分野と同じような結果を得られるのだと判断した。

7.3.4 人種

人種について、白人と黒人の変化で共通しているのはポジティブコミュニティとネガティブコミュニティとの間でのインラクティブなコミュニケーションは大統領選の前後で小さくなっているということである。これは、それぞれの極性同士の対立構造が深まり、議論や融和をしていないものだと考えられる。

7.3.5 ランダム比から見る会話量への考察まとめ

各分野について見た場合、外交と選挙というメタな抽出以外については全て、ランダムなものとの比においてインラクティブなコミュニケーションは小さくなるという結果であった。これは争点となるような各話題においてはその対立構造は選挙後に深まるということを示唆している。また、比が小さくならないものに関しては共通点が存在して、それはネガティブ内の密度もランダム比で見た時のネガティブ内コミュニケーションも減るという点である。これに関しては、ネガティブなコミュニティに置いて、内部での活発な会話をせずに外部とのみコミュニケーションを行っているという状態であることが読み取れ、これは対立構造においては融和という状態というよりは対立構造との議論が発生しているのではないかと考えられる。

第8章 結論

本研究では大規模なショックの前後における、コミュニティ規模・密度の変化と、コミュニティ間の対立構造の深まりについて分析を行なった。具体的には、2016年のアメリカ大統領選挙を取り上げ、ユーザ間のツイートのメンションに基づく会話ネットワークを構築した。ツイートの感情分析からあるトピックに対しポジティブなコミュニティとネガティブなコミュニティに分け、以下の分析を行った。

コミュニティ規模・密度の変化

選挙の前後における、ポジティブなユーザ群の割合とネガティブなユーザ群の割合変化から、各トピックに対するコミュニティ規模の変化を分析した。また、コミュニティ内における会話の活発さをコミュニティ密度と定義し、その変化を分析した。

コミュニティ間の対立構造の深まり

コミュニティが異なるユーザ間の選挙前後における会話の増減から、対立構造の深まりについて分析した。会話が減少した場合、コミュニティの対立構造が深まったとみなした。

コミュニティ規模の変化は、大統領選の前後でのそのトピックに関する話題の変化として捉えられる。そして、密度の変化については、会話の活発さがどれだけ変化したかわかる。例えば規模として投稿数やユーザ割合が増えているとしても密度が疎になっていれば、その中の会話や議論が活発に繰り広げられている訳ではなく、より多くのユーザによる単発の会話などが生じているということを示唆していると考えられる。

- コミュニティ規模が大きく変化する時は、結果の意外さにショックを受ける場合の他に、結果が出ることによって自分の立場を発言しやすくなるような場合であることがわかった
- 大統領選の前後において各ネットワーク内のインタラクティブな会話は減少したが、外交分野などの当事者以外を含むトピックに関しては、増加が確認された。
- インタラクティブなコミュニケーションが増えたコミュニティでは同時にネガティブの密度減少とネガティブ内ランダム比がどちらも減少している

コミュニティ規模については、具体例としてトランプに関するポジティブコミュニティの数の増加や、白人に関するポジティブユーザ及び投稿数の急激な増加が挙げられる。トランプは人種差別的で白人中流労働階級層からの支持が厚かつたが、その点に関する投稿が増えている。この点は、ランダムとの比においてトランプについても白人についても割合がほぼ

倍近くまで大きくなっていることからも見て取れる。トランプと白人に関しては投稿数がもともと大きいためサイレントマジョリティとまでは言えないが、一定数の反応を示していなかったトランプ支持者が選挙後に喜んでいるという状態の可能性が示唆される。また、コミュニティ内の密度に関しては基本的には小さくなることが多く、それは選挙に関連する話題それ自体が、結果に関する簡単な会話や情報拡散を多く含むようになっているからだということが考えられる。

対立構造が深まっているものに関しては、一般的には極性の違いによりユーザ間の距離が離れているためランダムな場合よりもコミュニケーションが起きにくくなっていると考えられる。対して、そのようにコミュニケーションが減っていないものに関しては、具体的には外交の例が挙げられるが、移民やイギリス人などの当事者でない者の議論が活発化していると考えられる。

以上のように、本研究では大規模なショックとしてアメリカ大統領選挙を取り上げ、その前後におけるコミュニティ規模・密度の変化と、コミュニティ間の対立構造の深まりについて分析を行なった。

8.1 今後の展望

本研究では、大統領選挙の前後におけるネットワークの変化がどのように生じているかを観測し、感情という極性で見た場合、その対立構造の間でのコミュニケーションは減少するという結果が得られた。本研究において使用したデータセットについては、より一層の研究の方向性があるため、今後の展望について記述する。

8.1.1 選挙中のデータ分析

本研究では選挙の前と後のデータのみを使用したが、同様の方法で選挙日のデータについても収集してあるため、選挙当日と開示の最中のデータについて同様の手法でそのネットワーク構造の変化過程を研究することが可能である。

8.1.2 時系列の細分化による変化過程の分析

本研究においては、分割の対象を大統領選の前後という二値に分けたが、そのコミュニケーションやネットワーク構造の変化過程を追い、どのように人々のコミュニケーションや集団の意見が形成されていくかという点についての細かい観測もまた、集団の意見形成とサイレントマジョリティーの出現過程、Eco chamber の形成過程として有意義であろう。

8.1.3 コミュニティ内の話題特徴量の導入

本研究の主題は、大統領選のようなショックの大きい議論の分かれるトピックについて、感情という極性を導入した際にその対立構造はどのように変化するのか、対立する極性同士の間でのコミュニケーションはどのように行われ、変化するのかということであった。そのため、トピックそれ自体の内容という特徴量に主軸において導入することはなかつたが、変化についてより訂正的な意味づけを考える際には、トピック分析の手法である LDA や DTM の導入などをコミュニティ内について使用して分析を行うことが可能である。この時、より細かい時系列コミュニティ変化と DTM などを合わせることによって、集団としての話題がどのように変遷して、コミュニケーションの対象はどの話題に占有されているのかということをモデリングすることが可能となるであろう。

8.1.4 世論調査とのデータ比較

本研究では感情の極性を分析のための主な指標として導入したが、これは、人間や集団の感情は話題に対してどのように変化し、どのように関わりあうのかを分析するためでもある。2016 年のアメリカ大統領選挙においてはメディアによる世論調査においては終始ヒラリー優勢出会ったのに対して、結果としてはトランプの勝利に終わった。この原因として考えられるのは二点ほど存在する。一つが、世論調査のサンプリング対象の偏りである。ヒラリー支持層は特に都市部の高学歴層などに多いが、トランプの支持層はより内陸の中流労働者層に多い。この二つの層の居住区などの差異はそのまま世論調査のサンプリングに影響を与え、偏りを生じさせる。二つ目の原因として考えられるのは、サイレントマジョリティーが世論調査に対しても影響して、世論調査のレベルでは支持を表明するのも躊躇われたのではないかということである。この二つの考えられる原因について、Twitter のデータを使用して世論調査と選挙結果が相反する形になったことの原因分析をすることも可能であろう。

8.1.5 支持率導出手法の提案

支持率の世論調査には問題点があるという状態であるため、近年より普及しているソーシャルメディアから支持率を計測する仕組みについて検討することができるのではないだろうか。世論調査その他において問題となるのは、サイレントマジョリティーの存在である。そのサイレントマジョリティーの特性や出現について Twitter で観測をすることが可能であれば、それを考慮した新しい支持率導出の手法を考えることができるであろう。これについても本研究の発展課題として有意義である。

謝辞

本研究を進めるにあたり、坂田・森研究室の多くの皆様にご指導いただきましたこと、この場を御借りして深く感謝し申し上げます。特に指導教員の坂田一郎教授、並びに共同研究室の森純一郎准教授には論文の題目決定から、設計・執筆に至るまで、非常に多くのアドバイスを頂きました。株式会社ホットリンク兼本学研究員の榎剛史氏、同じく本学研究員の浅谷公威氏、博士課程5年の大知正直氏には研究アイデアのブラッシュアップや基礎知識の取得、論文校正等で大変お世話になりました。坂田森研究室のメンバーの諸先輩方には、週次の研究会にて丁寧に私の研究へのアドバイスをしていただきました。特に磯沼さんには論文の再三の論文チェックを引き受けていただき大変感謝しております。そして、同期の小林くんと田爪くんとは、日々技術的な質問をしあったり、研究の息抜きをしたりと日常から研究に至るまで様々な面でお世話になりました。半年の期間でしたが坂田・森研究室に所属し、皆様の多大なご協力の下、研究テーマの設定から論文執筆に至る全ての過程を踏み、一つの研究をやり遂げることができたこと、またそのような機会を頂けたことに、大変感謝とともに、この経験を今後に生かせるよう精一杯努力して参りたいと思っております。

皆様への感謝の念を綴ると共に、今後の研究室の更なる発展を願って、謝辞とさせて頂きます。どうもありがとうございました。

参考文献

参考文献

- [1] Goldie, David, et al. "Using bibliometric and social media analyses to explore the "echo chamber" hypothesis." *Educational Policy* 28.2 (2014): 281-305.
- [2] Java, Akshay, et al. "Why we twitter: understanding microblogging usage and communities." *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. ACM, 2007.
- [3] Kwak, Haewoon, et al. "What is Twitter, a social network or a news media?." *Proceedings of the 19th international conference on World wide web*. ACM, 2010.
- [4] Takeichi, Yuki, et al. "Twitter as social sensor: Dynamics and structure in major sporting events." *Artificial Life* 14 (2014): 778-784.
- [5] 鳥海不二夫, 榊剛史, 岡崎直觀. 「人工知能」の表紙に関するツイートの分析・続報. 第4回 Web インテリジェンスとインタラクション研究会, 2014.
- [6] Osborne, Miles, and Mark Dredze. "Facebook, Twitter and Google Plus for breaking news: Is there a winner?." ICWSM. 2014.
- [7] Plutchik, Robert. "A general psychoevolutionary theory of emotion." *Theories of emotion* 1.3-31 (1980): 4.
- [8] Parrott, W. Gerrod. *Emotions in social psychology: Essential readings*. Psychology Press, 2001.
- [9] Marui, Junki, et al. "Empirical Study of Conversational Community Using Linguistic Expression and Profile Information." *International Conference on Active Media Technology*. Springer International Publishing, 2014.
- [10] Huberman, Bernardo A., Daniel M. Romero, and Fang Wu. "Social networks that matter: Twitter under the microscope." (2008).
- [11] 上子優香, et al. "テキスト分析に基づくソーシャルメディア上でニュースの影響度予測に関する研究." *人工知能学会全国大会論文集* 29 (2015): 1-4.
- [12] Pak, Alexander, and Patrick Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." *LREc*. Vol. 10. No. 2010. 2010.

- [13] Gilbert, CJ Hutto Eric. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." Eighth International Conference on Weblogs and Social Media (ICWSM-14). Available at (20/04/16) http://comp.social.gatech.edu/papers/icwsm14_vader_hutto.pdf. 2014.
- [14] Romero, Daniel M., Brian Uzzi, and Jon Kleinberg. "Social Networks Under Stress." Proceedings of the 25th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2016.
- [15] Mustafaraj, Eni, et al. "Vocal minority versus silent majority: Discovering the opinions of the long tail." Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on. IEEE, 2011.
- [16] Newman, Mark EJ, and Michelle Girvan. "Finding and evaluating community structure in networks." Physical review E 69.2 (2004): 026113.
- [17] Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." Journal of statistical mechanics: theory and experiment 2008.10 (2008): P10008.
- [18] Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." Journal of machine Learning research 3.Jan (2003): 993-1022.
- [19] 吉野孝 前島和弘 "オバマ政権と過渡期のアメリカ社会 選挙, 政党, 制度, メディア, 対外援助" 2012: 98-99
- [20] ヒラリークリントン公式 HP <https://www.hillaryclinton.com>
- [21] ドナルドトランプ公式 HP <https://www.donaldjtrump.com>

第9章 卷末収録

卷末に、そのほかの分析における結果を記す。

9.1 経済政策

次に、経済政策の分野を分析する。経済政策に関する分野は、貿易にコミュニティについて分析を行なった。

9.1.1 貿易

貿易についてのコミュニティを分析する。コミュニティの活性度や規模に関わる、投稿数の変化は 6470 から 3895 であり、コミュニティ内に存在するポジティブユーザとネガティブユーザの割合は大統領選以前はネガティブユーザが微かに過半数であったが、大統領選後のデータではポジティブなユーザが 6 割以上を占める。ネットワーク内の密度はポジティブネガティブどちらにおいても増加している。ユーザに対して、ランダムに同数のエッジを張った場合との比を考えると、ポジティブ内のものは 2.75 から 1.71 に減少、ネガティブなものは 0.65 から 1.68 に倍増、そしてポジティブなものとネガティブものの間のインタラクティブなものは 0.31 から 0.21 へと減少している。

コミュニティ内の感情度の分布のグラフに関しては大統領選の前後によって大きな変化等はほとんどなく、投稿数とその時の感情度の推移についての大きな相関は見受けられない。

表 9.1: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	6470	2951	2960	0.498	0.501
after	3895	2257	1392	0.62	0.38

表 9.2: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	483	330	79	74
after	404	265	99	40

表 9.3: ネットワークの密度

時期	pos	neg
before	0.11	0.03
after	0.12	0.07

表 9.4: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.68	0.16	0.15
after	0.66	0.25	0.10

表 9.5: ランダムネットワークとの比

時期	pos	neg	interactive
before	2.75	0.65	0.31
after	1.71	1.68	0.21

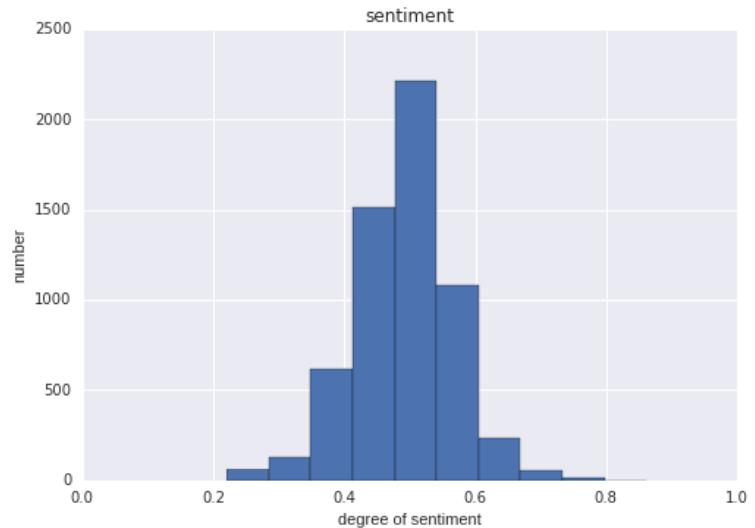


図 9.1: 大統領選以前のコミュニティ内のユーザ感情度分布

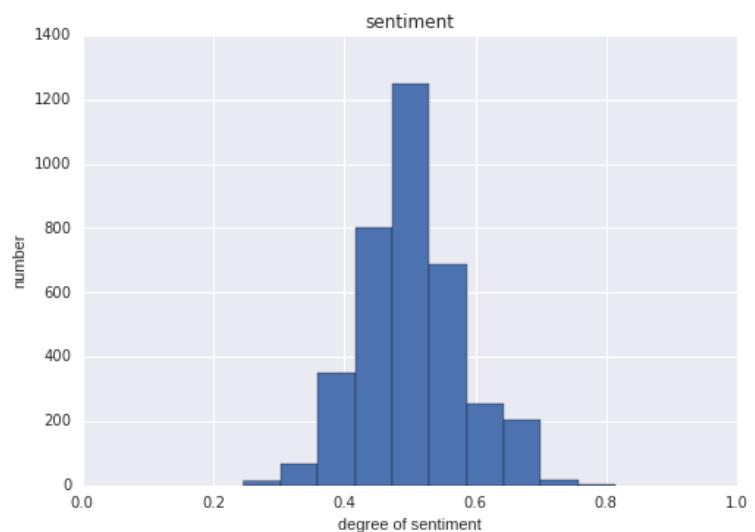


図 9.2: 大統領選後のコミュニティ内のユーザ感情度分布

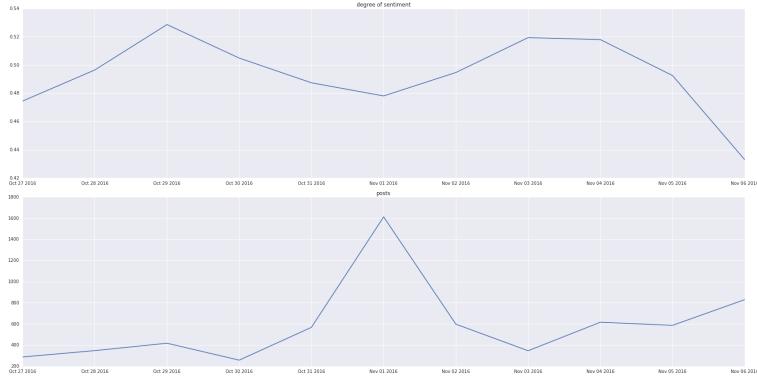


図 9.3: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

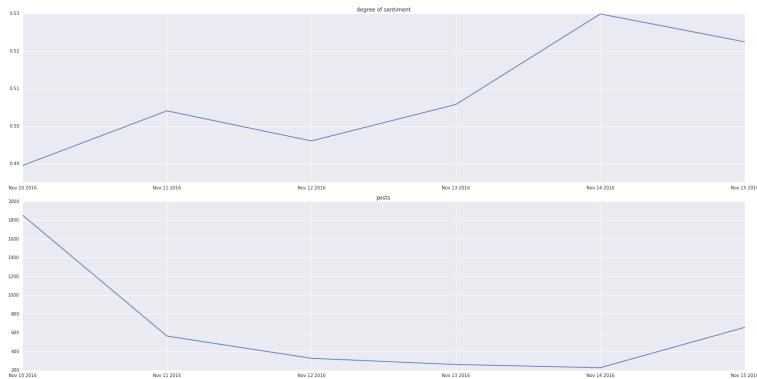


図 9.4: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.2 多様性への価値観

次に、多様性への価値観に関する分野について分析する。この分野についてはリベラルをクエリに用いて抽出をした。結果としては、コミュニティの活発さや規模に関しては大きく増加しており、ポジティブユーザとネガティブユーザの割合に関してはポジティブなものが過半数を占めており大きな変化はなかった。ポジティブなものとネガティブものの間のコミュニケーションは減少していた。

9.2.1 リベラル

多様性への価値観については、リベラルについて言及しているものを抽出した。その分析結果について記す。まず、コミュニティの活発さや規模の指標となる投稿数については、3624 から 11546 と大きく増加した。コミュニティ内部のポジティブなユーザとネガティブなユーザの割合については、大統領選の前後どちらにおいてもポジティブユーザが過半数と大きな変化はなかった。ネットワークの情報については、全体のエッジ数が 358 から 1323 へと増加しており、ネットワーク内での密度に関してはネガティブなものの中でも密度は 0.12 から 0.20 へと増加している。ポジティブユーザとネガティブユーザに対してランダムに同数のエッジを張った場合の比を考えると、ネガティブ内のコミュニケーションは 2.94 から 3.86 へと増加しているのに対して、ポジティブ内のコミュニケーションは 1.23 から 0.79 へ減少、ポジティブユーザとネガティブユーザ間のインタラクティブなコミュニケーションについても 0.16 から 0.09 へと減少している。

ユーザの感情度の分布については大きな割合の変化は観測されず、投稿数と感情度の相関も見受けられない。投稿数については選挙の当日に近いものほど投稿数が多いということがわかる。

表 9.6: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	3624	2042	1406	0.59	0.41
after	11546	6237	4616	0.57	0.43

表 9.7: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	358	155	175	28
after	1323	343	924	56

表 9.8: ネットワークの密度

時期	pos	neg
before	0.08	0.12
after	0.05	0.20

表 9.9: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.43	0.49	0.08
after	0.26	0.70	0.04

表 9.10: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.23	2.94	0.16
after	0.79	3.86	0.09

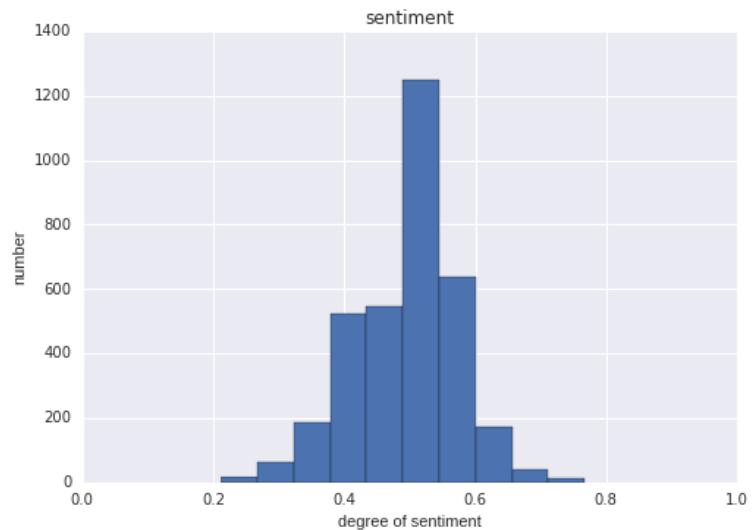


図 9.5: 大統領選以前のコミュニティ内のユーザ感情度分布

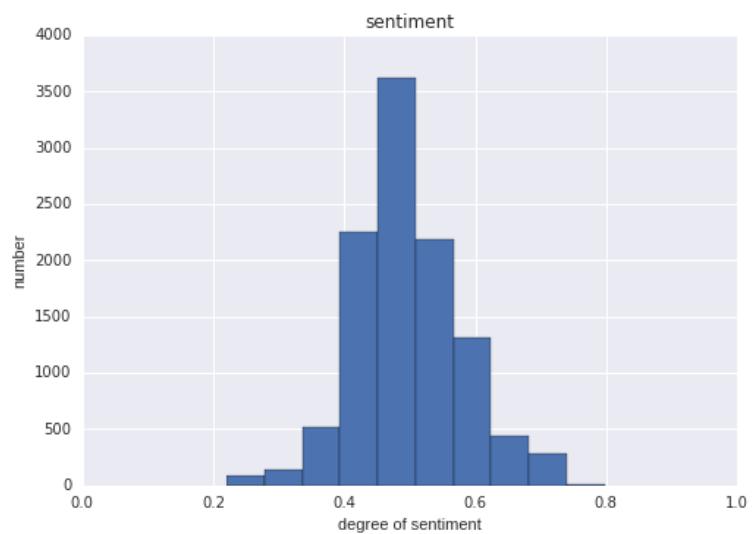


図 9.6: 大統領選後のコミュニティ内のユーザ感情度分布

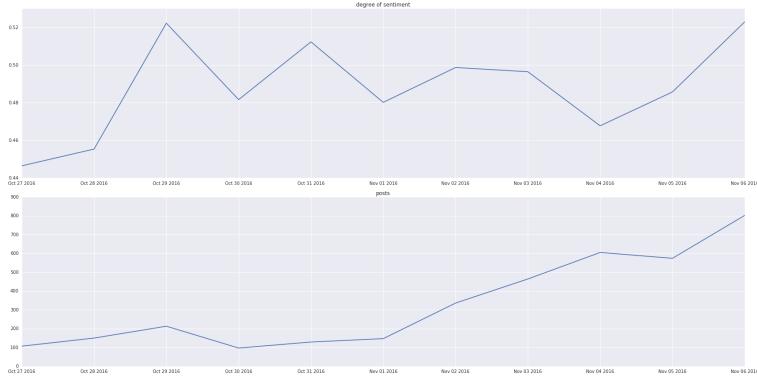


図 9.7: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

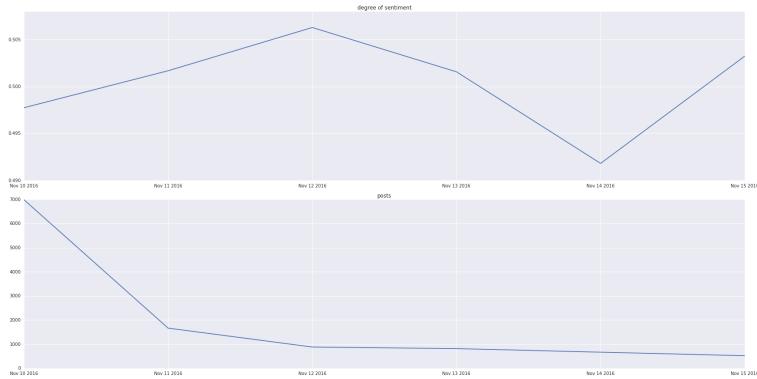


図 9.8: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.3 性の問題

次に、性の問題についての分析をする。抽出したコミュニティはゲイについてである。結果のまとめとしては、コミュニティの活発度や規模は大統領選後に大きく増加しており、その中のポジティブなユーザとネガティブなユーザの比率は大きくは変化しなかった。そしてポジティブなものとネガティブものの間のインタラクティブなコミュニケーションは減少し、ネガティブなものの中でのコミュニケーションが大きく増加するというものになった。

9.3.1 ゲイ

ゲイについて抽出したコミュニティについて分析結果を記す。コミュニティの規模や活発さを表す投稿数については、大統領選の前後で 3172 から 13612 と大きく増加している。その中でポジティブなユーザとネガティブなユーザの割合に関しては、ポジティブなものが一貫して過半数と大きな変化はなかった。ネットワーク内の情報についてはネガティブからネガティブへのエッジ数が 40 から 117 と大きく増加しており、ランダムに張ったエッジとの比で見た場合においても、ポジティブ内のコミュニケーションは 1.93 から 1.27 に減少、ポジティブなものとネガティブものの間のインタラクティブなものは 0.35 から 0.22 と減少する中でネガティブな中のコミュニケーションについては 1.11 から 2.90 と 2 倍以上に増加している。

大統領選前後の感情度の分布を見て見ると、大統領選前については 0.4 付近と 0.5 付近に大きな山をもつ分布であるが、大統領選後には 0.5 と 0.6 付近に大きな山を持つ分布になっている。時系列データに関しては感情度と投稿数については相関は見られないが、大統領選直後に反応が大きくなっていることが受けられる。

表 9.11: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	3172	1714	1327	0.56	0.44
after	13612	8419	4770	0.64	0.36

表 9.12: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	189	116	40	33
after	308	160	117	31

表 9.13: ネットワークの密度

時期	pos	neg
before	0.07	0.03
after	0.02	0.02

表 9.14: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.61	0.21	0.17
after	0.52	0.38	0.10

表 9.15: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.93	1.11	0.35
after	1.27	2.90	0.22

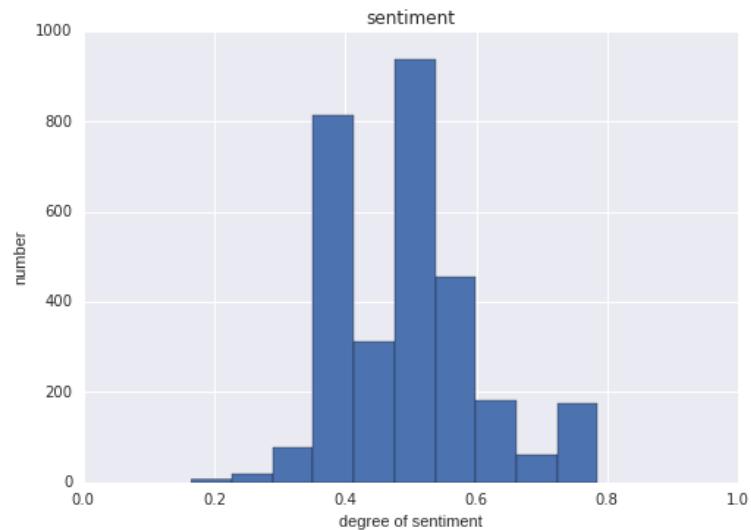


図 9.9: 大統領選以前のコミュニティ内のユーザ感情度分布

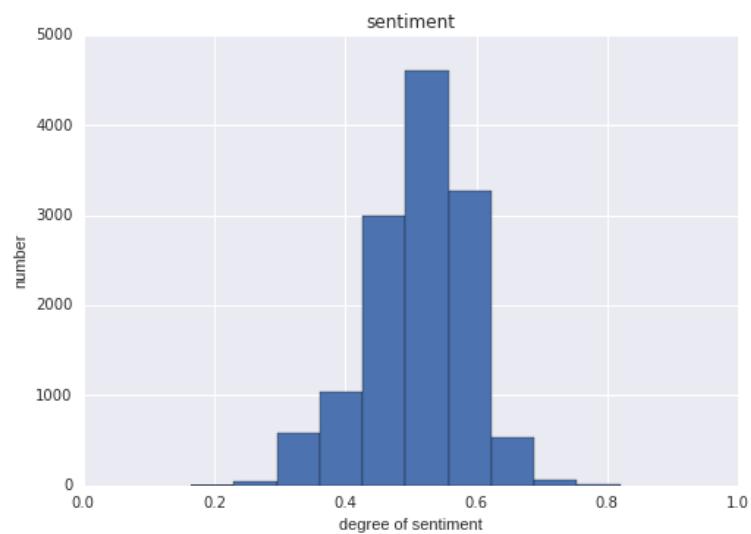


図 9.10: 大統領選後のコミュニティ内のユーザ感情度分布

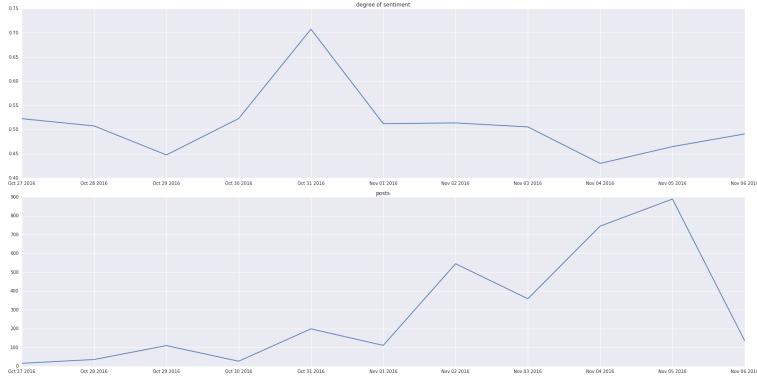


図 9.11: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

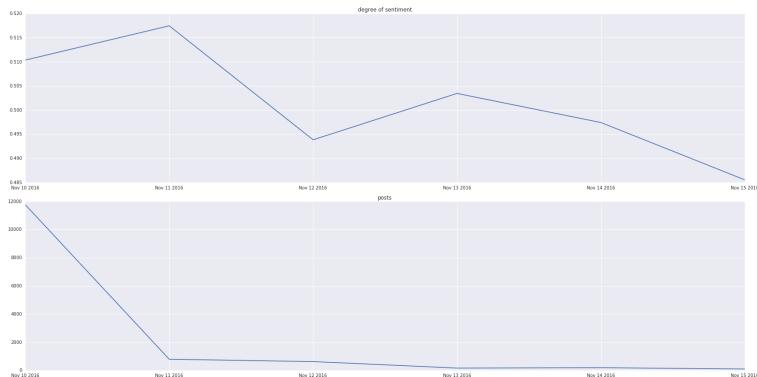


図 9.12: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.4 サブクエリ分析

この節では、ヒラリーとトランプのコミュニティから、さらにクエリを指定して、二つのクエリ両方に言及しているコミュニティを抽出して分析を行う。サブクエリに設定するものは、オバマ、黒人、移民の三つとする。

9.4.1 ヒラリー＆オバマ

ヒラリーとトランプの両方に言及しているコミュニティを抽出して分析をする。まず、コミュニティ活性度と規模の指標となる投稿数は 8568 から 10767 へと増加している。コミュニティ内のポジティブユーザとネガティブユーザの割合については大統領選の前後一貫してポジティブユーザが過半数を占めているが、選挙後には特にポジティブユーザの割合が 0.84 と多くを占める。ネットワーク内での密度は大統領選の後にはポジティブ内のものもネガティブ内のものも減少している。ポジティブユーザとネガティブユーザに対してランダムに同数のエッジを張った場合との比を考えると、ネガティブ内のコミュニケーションは 1.46 から 7.25 と大幅に増加しているのに対して、ポジティブ内は 1.92 から 0.97 へと半減しており、ポジティブユーザとネガティブユーザ間のインタラクティブなコミュニケーションは 0.27 から 0.46 へと増加している。

ユーザの感情度の分布を見てみると、大統領選挙の後においては 0.5 を少々超えるユーザの割合がとても多く存在していることが判明する。

表 9.16: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	8568	4032	3515	0.53	0.47
after	10767	8727	1708	0.84	0.16

表 9.17: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	892	488	283	121
after	505	344	98	63

表 9.18: ネットワークの密度

時期	pos	neg
before	0.12	0.08
after	0.04	0.06

表 9.19: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.55	0.32	0.14
after	0.68	0.19	0.12

表 9.20: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.92	1.46	0.27
after	0.97	7.25	0.46

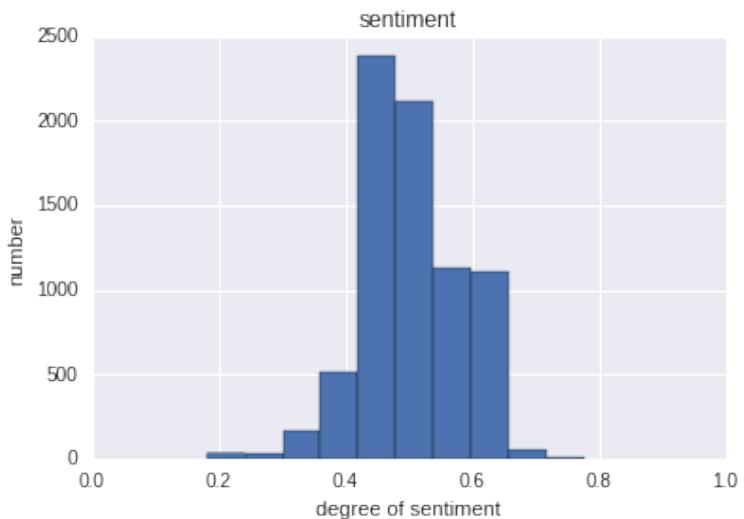


図 9.13: 大統領選以前のコミュニティ内のユーザ感情度分布

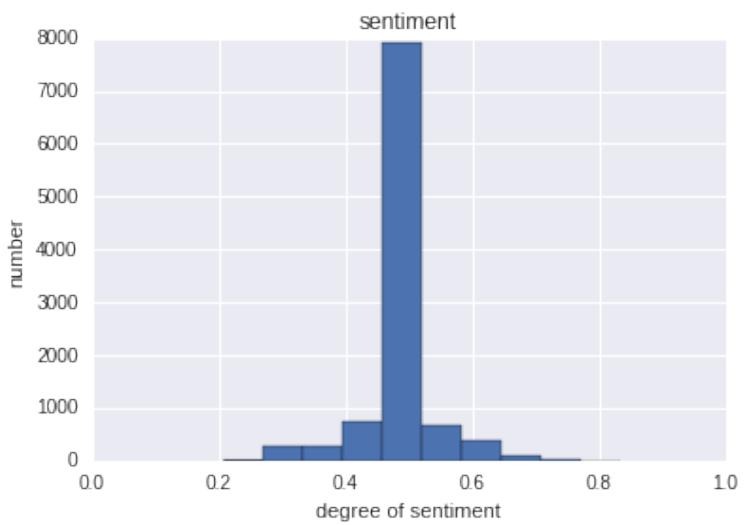


図 9.14: 大統領選後のコミュニティ内のユーザ感情度分布

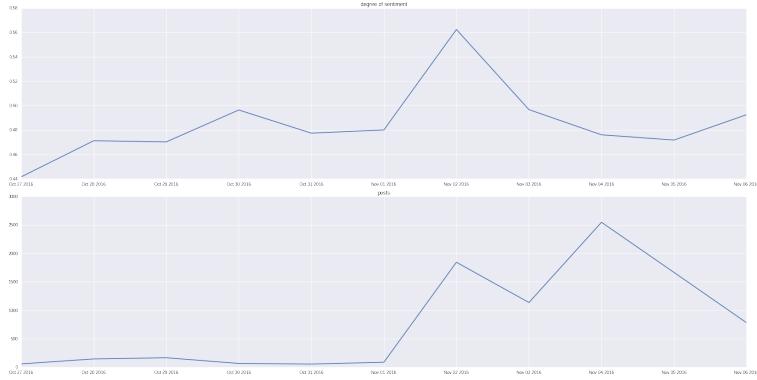


図 9.15: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

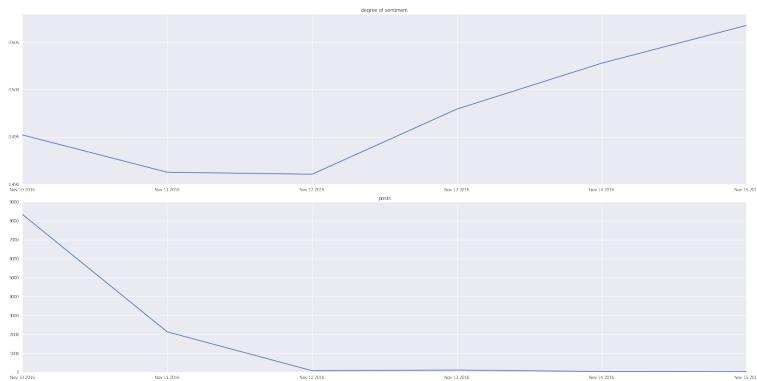


図 9.16: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.4.2 トランプ&オバマ

トランプとオバマの両方について言及しているコミュニティを抽出して分析した結果を見てみる。まず、規模や活発性を表す投稿数については、10761 から 59921 と 6 倍近くまで大きくなっている。そのコミュニティ内のユーザ感情度の割合については、大統領選前はポジティブユーザが 7 割を占めたが、大統領選挙後においては、ネガティブユーザが 6 割以上を占めており、逆転していることが伺える。

コミュニティ内の密度はポジティブ内ネガティブ内ともに減少している。そして、ポジティブユーザとネガティブユーザに対してランダムに同数のエッジを張った場合との比をみ

ると、ポジティブ内のコミュニケーションは5倍以上に増加しており、対してネガティブ内のコミュニケーションは1/5へと減少している。ポジティブユーザとネガティブユーザ間のインタラクティブなコミュニケーションについては0.48から0.32へと減少している。

コミュニティ内のユーザの感情度の分布に関して見てみる。大統領選以前については0.5程度の部分に大きな山がある状態であったが、大統領選後のデータにおいては0.3付近に大きな山があるという状態へと変化したということがわかる。投稿の時系列と感情度は主に呼応しており、特に選挙直前と直後においては数多くの投稿があったことがわかる。

表 9.21: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	10761	6580	2813	0.70	0.30
after	59921	19541	37629	0.34	0.66

表 9.22: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	1281	741	280	260
after	2630	1862	387	381

表 9.23: ネットワークの密度

時期	pos	neg
before	0.11	0.10
after	0.10	0.01

表 9.24: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.58	0.22	0.20
after	0.71	0.15	0.15

表 9.25: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.18	2.44	0.48
after	6.06	0.34	0.32

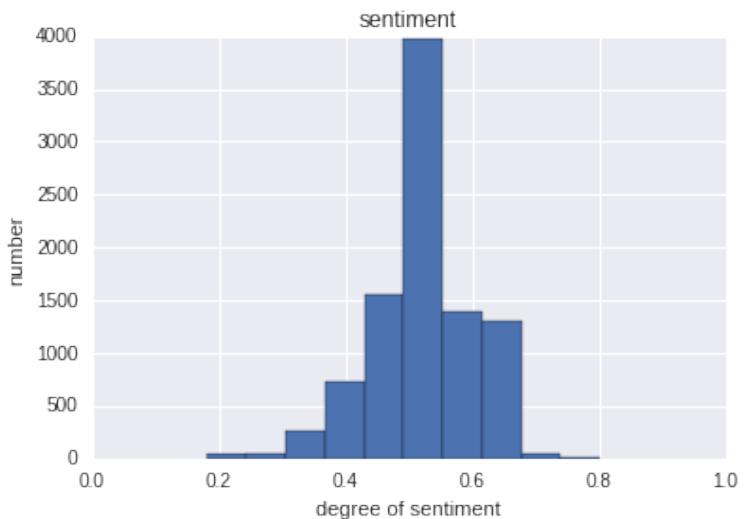


図 9.17: 大統領選以前のコミュニティ内のユーザ感情度分布

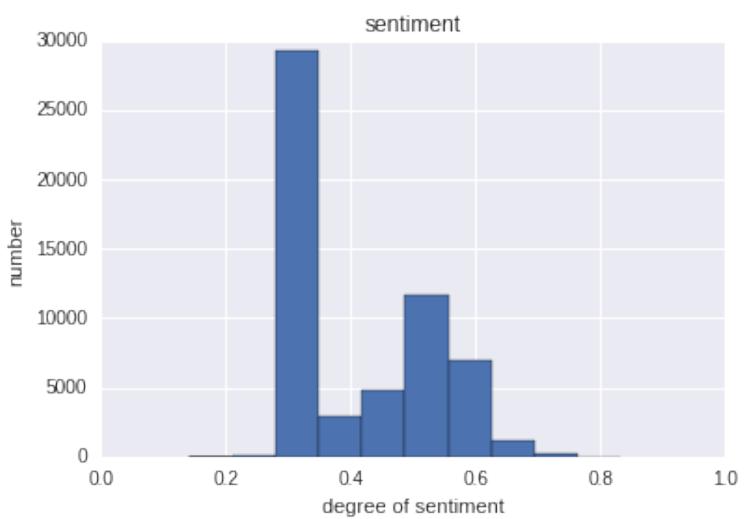


図 9.18: 大統領選後のコミュニティ内のユーザ感情度分布

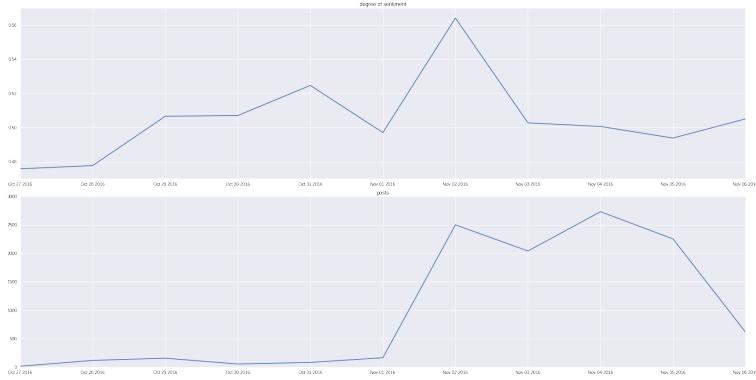


図 9.19: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

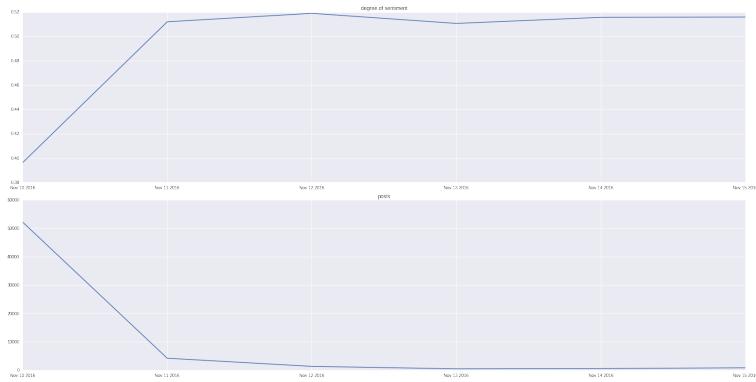


図 9.20: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.4.3 ヒラリー&黒人

ヒラリーと黒人の両方について言及しているコミュニティを分析する。コミュニティの活性化や規模を表す投稿数は 4482 から 3734 と少々減少している。大統領選の前後でのコミュニティ内のユーザの感情度については、大統領選の前ではポジティブユーザが 0.77 と過半数であるのに対して、大統領選の後においてはネガティブユーザが 0.78 と逆転している状態となっている。コミュニティ内の密度については、ネガティブなものは 0.06 から 0.35 へと増加している。ランダムにエッジを張った場合との比較においては、ポジティブ内のコミュニケーションは半減、ネガティブ内のコミュニケーションも 1/3 に減少しており、さらにポ

ジティブユーザとネガティブユーザ間のコミュニケーションも 0.20 から 0.02 へと 1/10 までに減少している。

ユーザの感情度の分布について見てみると、このコミュニティがとても偏りの大きい分布をしていることがわかる。大統領選の前については、0.6 付近に感情値を持つユーザが多く存在するが、大統領選の後においては圧倒的に 0.4 付近のユーザが多くなっている。

表 9.26: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	4482	3099	942	0.77	0.23
after	3734	797	2902	0.22	0.78

表 9.27: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	183	114	56	13
after	1041	26	1009	6

表 9.28: ネットワークの密度

時期	pos	neg
before	0.04	0.06
after	0.03	0.35

表 9.29: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.62	0.31	0.07
after	0.02	0.97	0.01

表 9.30: ランダムネットワークとの比

時期	pos	neg	interactive
before	1.06	5.64	0.20
after	0.54	1.57	0.02

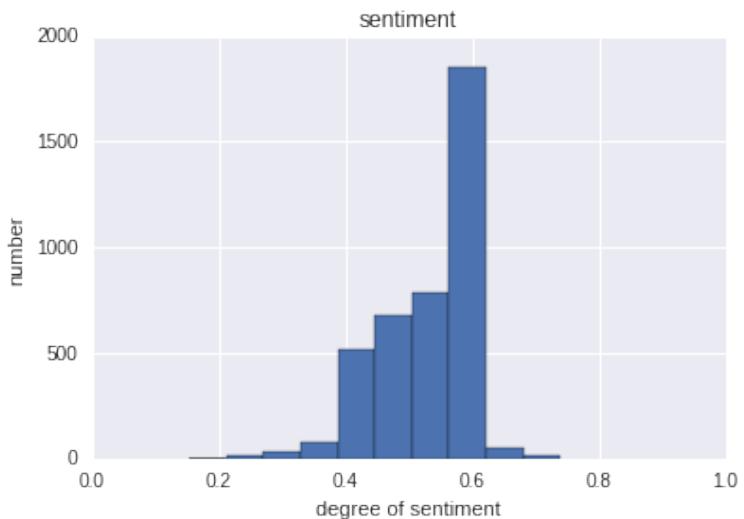


図 9.21: 大統領選以前のコミュニティ内のユーザ感情度分布

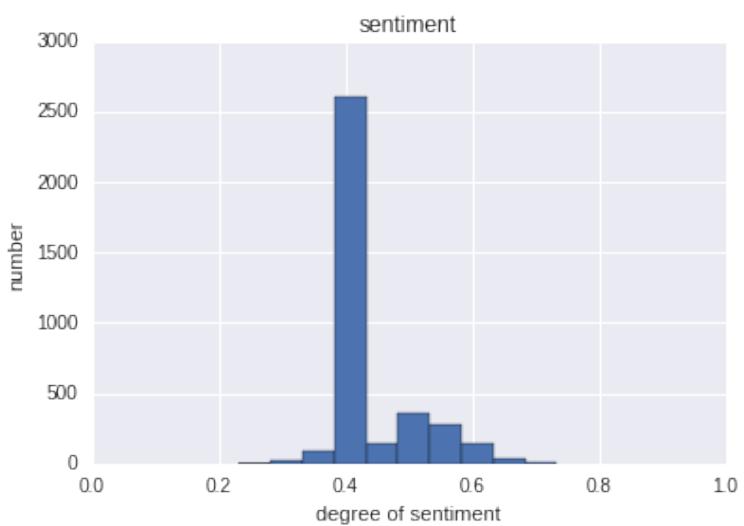


図 9.22: 大統領選後のコミュニティ内のユーザ感情度分布

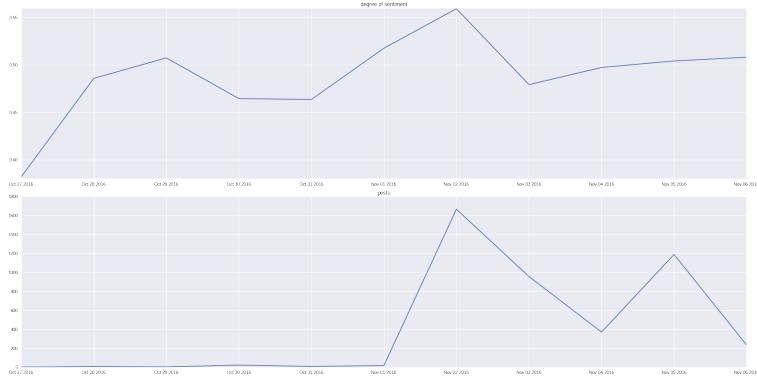


図 9.23: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

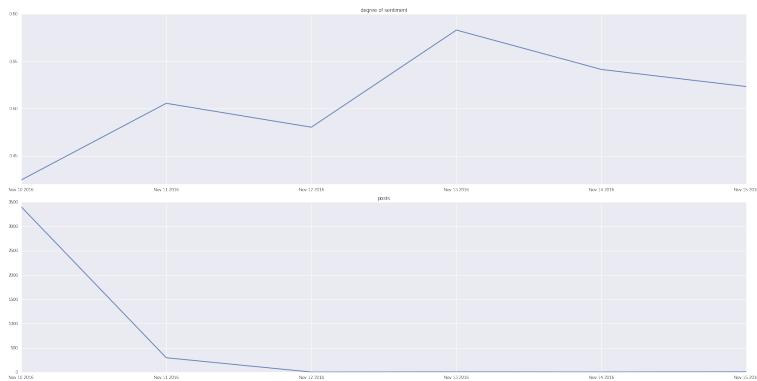


図 9.24: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.4.4 トランプ&黒人

トランプと黒人の両方に言及しているコミュニティについて見てみる。コミュニティの活性度と規模の指標となる投稿数は 8942 から 19613 と大幅に増加している。ポジティブユーザーとネガティブユーザーの割合については、大統領選の前後共にポジティブユーザーが過半数だが、差は縮まっている。コミュニティの密度についてはポジティブ内ネガティブ内共に減少しており、ネットワークにランダムにエッジを張った場合との比較に関しては、ポジティブ内のコミュニケーションは倍増、ネガティブ内でのコミュニケーションは 1/3 程度に減少、ポジティブユーザーとネガティブなユーザーとのコミュニケーションは 0.47 から 0.20 へ半減し

ている。

ユーザの感情度の分布については、前後共にあまり大きな変化はない分布をしている。

表 9.31: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	8942	5774	2312	0.71	0.29
after	19613	10295	8531	0.55	0.45

表 9.32: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	624	291	214	119
after	333	198	102	33

表 9.33: ネットワークの密度

時期	pos	neg
before	0.05	0.09
after	0.02	0.01

表 9.34: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.47	0.34	0.19
after	0.59	0.31	0.10

表 9.35: ランダムネットワークとの比

時期	pos	neg	interactive
before	0.91	4.20	0.47
after	1.99	1.49	0.20

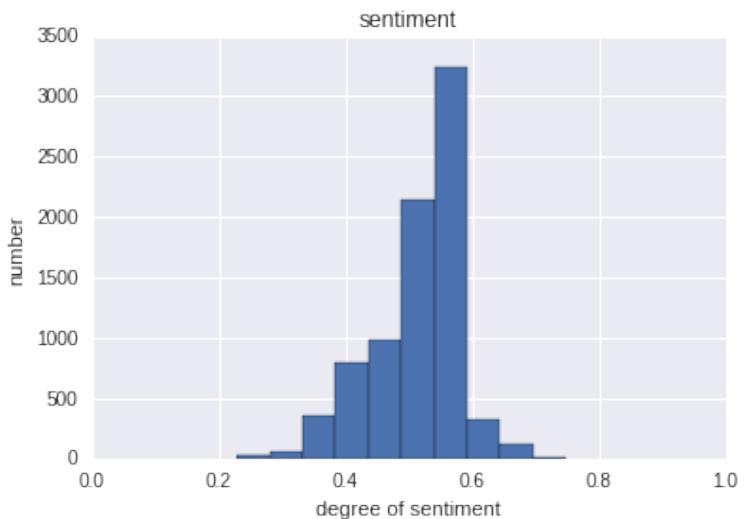


図 9.25: 大統領選以前のコミュニティ内のユーザ感情度分布

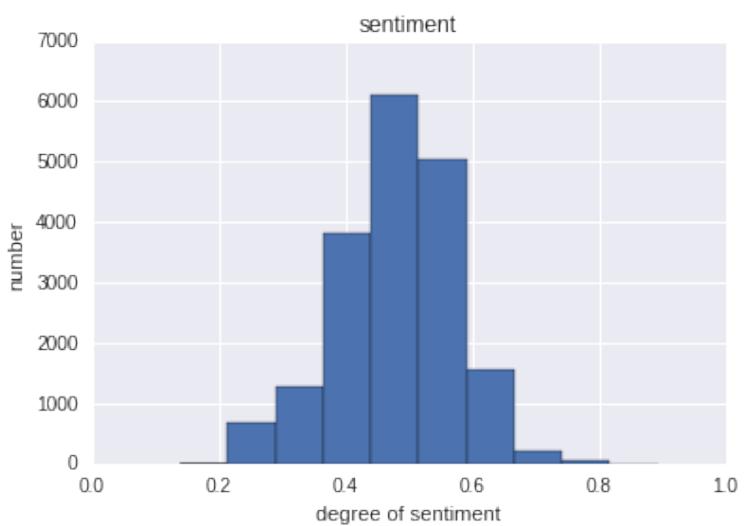


図 9.26: 大統領選後のコミュニティ内のユーザ感情度分布



図 9.27: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

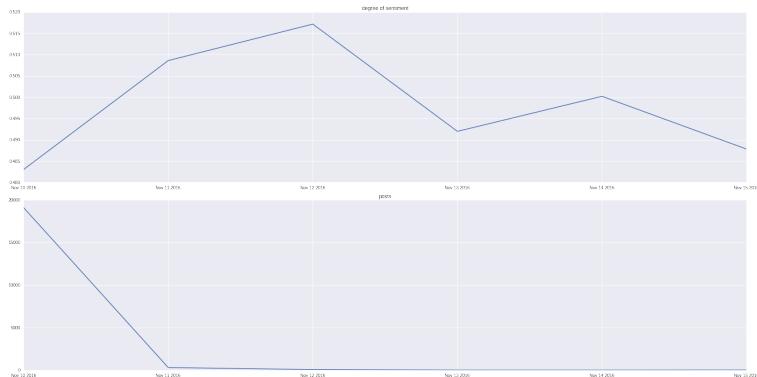


図 9.28: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.4.5 ヒラリー&移民

ヒラリーと移民の両方について言及しているコミュニティについて分析する。まず、コミュニティの活性度と規模の指標となる投稿数は 9391 から 3273 と 1/3 程度に減少している。その中のポジティブユーザーとネガティブユーザーの割合は、大統領選の前後一貫してネガティブなユーザーが過半数を占めているが、特に大統領選の後においては 0.87 がネガティブユーザーとなっている。ネットワーク情報としてはエッジ数が 2943 から 77 へと大きく減少している。また、ネガティブ内の密度に関しても 0.51 から 0.02 へと大きく減少をしている。そして、ランダムにエッジを張った場合との比をみると、ポジティブなユーザーに関してはコ

ミュニケーションは 0.64 から 20.1 へと大幅な増加，ネガティブ内では 2.24 から 0.76 への減少，そしてインラタクイブなものに関しては，0.11 から 0.33 へと増加している。

ユーザの感情度の分布に関しては大統領選の前は 0.4-0.7 の間に一定量のユーザが存在しているが，大統領選の後には 0.4 付近にほとんどのユーザが固まる状態となっている。

表 9.36: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	9391	3410	4827	0.41	0.59
after	3273	424	2784	0.13	0.87

表 9.37: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	2943	322	2463	158
after	77	27	44	6

表 9.38: ネットワークの密度

時期	pos	neg
before	0.09	0.51
after	0.06	0.02

表 9.39: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.11	0.84	0.05
after	0.35	0.57	0.08

表 9.40: ランダムネットワークとの比

時期	pos	neg	interactive
before	0.64	2.44	0.11
after	20.1	0.76	0.33

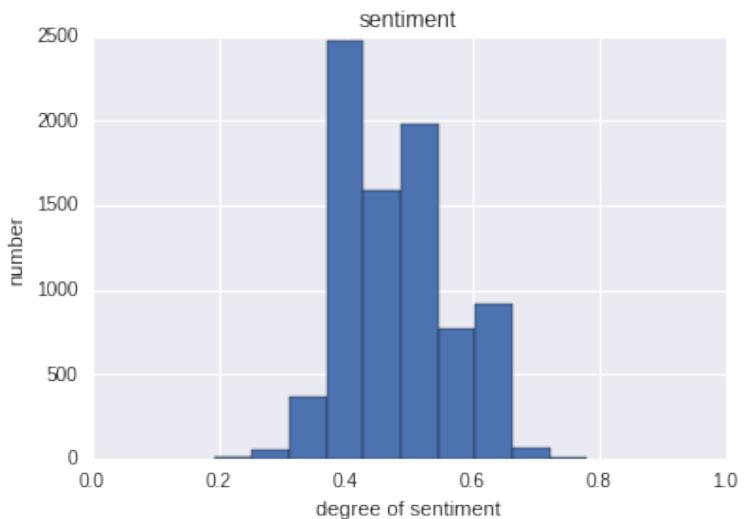


図 9.29: 大統領選以前のコミュニティ内のユーザ感情度分布

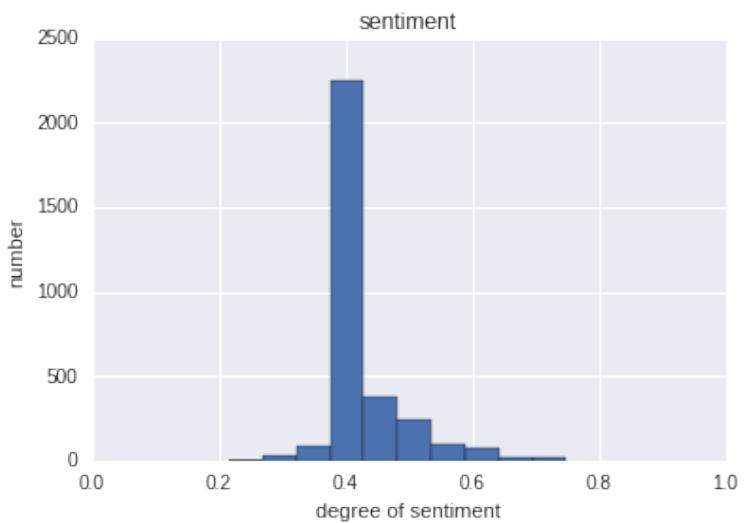


図 9.30: 大統領選後のコミュニティ内のユーザ感情度分布



図 9.31: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)



図 9.32: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

9.4.6 トランプ&移民

トランプと移民の両方に言及しているコミュニティに関して分析をする。コミュニティの活性度と規模の指標となる投稿数は 35933 から 72317 へと倍増している。また、コミュニティ内のポジティブユーザとネガティブユーザの感情度の割合に関しては、大統領選の前はネガティブユーザが過半数を占めていたが、大統領選の後についてはポジティブユーザが過半数となっており、逆転している状態となっている。コミュニティ内の密度に関しては、ポジティブ内ネガティブ内ともに大統領選を経て減少している。また、ポジティブユーザとネガティブユーザの間のコミュニケーションについて、ランダムにエッジを張った場合との比

を見てみると、大統領選のまえが 0.41 であったのに対して大統領選の後においては 0.38 と減少していることがわかる。

ユーザの感情値についての分布を大統領選の前後で比べてみると、大統領選の前に関してはネガティブな 0.4 付近にユーザの山が固まっていたが、大統領選後においては 0.6 付近にも山ができており、中間部のユーザが減っているため、ユーザの感情値が二分されて対立していることが見て取れる。コミュニティ内の日毎の感情値平均と投稿数の推移については、投稿数が選挙当日に近いほど上がっていること、そして選挙終了後も大きく減りはするものの一定の投稿数が保たれているということがわかる。

表 9.41: コミュニティ情報

時期	投稿数	ポジティブユーザ	ネガティブユーザ	ポジティブ割合	ネガティブ割合
before	35933	10333	16857	0.38	0.62
after	72317	32365	27758	0.54	0.46

表 9.42: ネットワーク情報

時期	エッジ数	pos-pos	neg-neg	interactive
before	6319	2104	2980	1235
after	7082	3916	1836	1330

表 9.43: ネットワークの密度

時期	pos	neg
before	0.20	0.18
after	0.12	0.07

表 9.44: エッジの割合

時期	pos-pos	neg-neg	interactive
before	0.33	0.47	0.20
after	0.55	0.26	0.19

表 9.45: ランダムネットワークとの比

時期	pos	neg	interactive
before	2.31	1.23	0.41
after	1.91	1.21	0.38

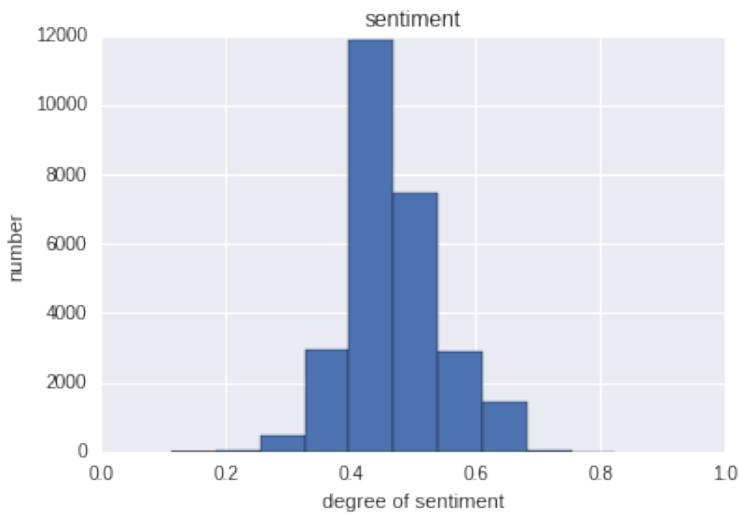


図 9.33: 大統領選以前のコミュニティ内のユーザ感情度分布

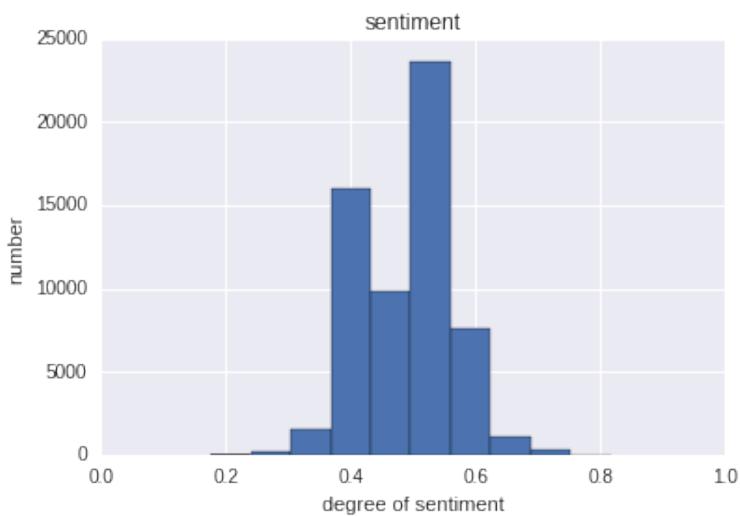


図 9.34: 大統領選後のコミュニティ内のユーザ感情度分布

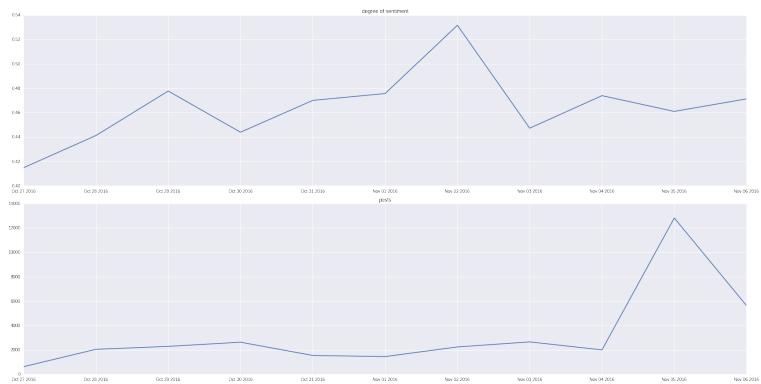


図 9.35: 大統領選以前のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)

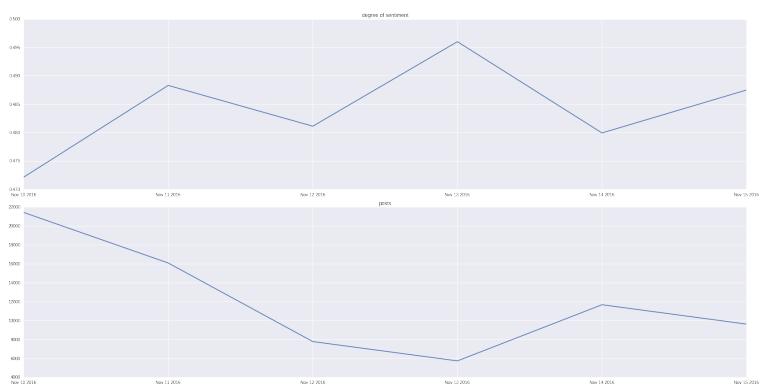


図 9.36: 大統領選後のコミュニティ内の投稿数時系列変化(上)と投稿の平均感情度の時系列変化(下)