Deep Stochastic Block Models for Measuring Political Engagement

by

Cameron Raymond

A thesis submitted to the
School of Computing
in conformity with the requirements for
the course of CISC 500 - Undergraduate Thesis

Queen's University
Kingston, Ontario, Canada
March 2020

Supervisor: Dr. Robin Dawes

Abstract

This is my abstract.

Acknowledgments

Blah blah blah.

Contents

Chapter 1

Introduction

1.1 Background

The advent of social media has enabled political parties to engage with the broader populous in new and unforeseen ways. The ability to bypass the traditional mediating forces of mass media allows for an unfiltered promotion of policy, ideology and party stances. This is specifically interesting in Canada's political system which has historically been defined by large brokerage parties. In order to win a diverse range of electoral districts across Canada, these "big tent" parties try to appeal to various political persuasions. Political adverts, and policy have traditionally been the conduits through which brokerage parties attempt to accommodate different ideologies, but social media allows for a direct, granular approach to political messaging which is completely novel. Social media, a form of new media, facilitate the creation of social networks, DEFINE. Social networks such as Twitter, are inherently relational and thus lend themselves well to being represented as graphs. Therefore, as political strategy becomes increasingly digital, the use of graph theory can potentially illustrate how large brokerage parties organize and along what axes Canadians engage with political parties.

1.2 Motivation

The way information is distributed and received has changed significantly over the past decade. Cogburn and Espinoza-Vasquez argue that Barrack Obama's 2008 presidential campaign was a watershed moment in social media campaigning - and in the subsequent decade, from Macron to Brexit to the Five Star Movement, social media has played an increasing role in how politics is conducted (CITE). Over the past XAMOUNT of years, money spent on digital marketing has increased by XPERCENT - making the study of new media critical from a social science perspective. Additionally, rises in political polarization, populism and a decline in trust in political institutions in the 21^{st} century has been a topic of popular debate. Ezra Klein argues in his 2019 book, *Why We're Polarized*, that this due to a shift in preferences for parties over policies. If this is the case, then this preference to engage along party lines rather than along the axis of issues should pattern engagement. An empirical analysis of how users behave and engage with political parties online should privilege the relational aspect of social media. Social network analysis helps avoid the pitfalls of survey data, famously described by Allen Barton as "a sociological meat grinder, tearing the individual from [their] social context"[1].

Graph theory's use in social network analysis, also called network science, has already been applied to explore problems in marketing, sociology and epidemiology - but there is a gap in analysis of political engagement online. Therefore, the contribution of this thesis is a novel, robust mathematical process for analyzing different axes of political engagement online in a purely relational manner. A secondary outcome of making relationships a "first-class citizen" in this work will be an organic analysis of what issues produce the most engagement. This deviates from traditional survey data that ask test subjects which issues concern them - and instead uses observations of past behaviour to model variations in local connectivity to answer the question: "what do people actually care about?" All of this will be done in the context of the 2019 Canadian federal election and the tweets of Canada's five major, english speaking party leaders: Andrew Scheer, Elizabeth May, Jagmeet Singh, Justin Trudeau, and Maxime Bernier.

1.3 Research Question

The primary question concerning this project is: in the lead up to the 2019 Canadian federal election, did politically active users on Twitter engage with political elites along the axis of issues (also referred to as topics or policies) or parties? If Ezra Klein is correct, then *who* produces the message will pattern engagement more than *what* the message is; this will act as the initial null hypothesis.

The secondary question to be explored is: during this period, what topics produced the highest level of engagement? And, what topics were important to the broader populous versus a party leader's base?

Chapter 2

Modelling Unstructured Text Data

2.1 Background

Blah blah blah

2.2 Latent Dirichlet Allocation

Blah blah blah

Chapter 3

Graph Theory and Computational Social Science

3.1 Graph Theory

Where the graphs at...

3.1.1 Background

Woohoo euler

3.1.2 Spectral Graph Theory

An excerpt from an Alloy

3.1.3 Network Laplacian Spectral Descriptor

Hellll ya

3.2 Random Graphs

Where the graphs at...

3.2.1 Background

Hellll Ya

3.2.2 Stochastic Block Models

Hellll Ya

3.3 Measures of Centrality

Where the graphs at...

3.3.1 Background

Hellll Ya

3.3.2 Eigenvector Centrality

Hellll Ya

Chapter 4

Social Media and Politics

Listing 4.1 Alloy specification of a singly-linked list using only binary relations

```
module List

sig Node {
    next : option Node
}

sig List {
    first : Node
}

fact NodeInOneList {
    all n : Node | one l : List | n in (l.first).*next
}

fact NoCycle {
    all n : Node | n ! in n.^next
}

fun Show() {}

run Show for 4
```

4.0.1 Political Communication in the Digital Era

Lorem Ipsum

test test test

4.0.2 Technologies Implications for Democratization

4.0.3 Canadian Brokerage Politics

ahh yeah boiii. Explain what brokerage parties are

Rationale

Why canada is a good case for this kind of question

Chapter 5

Thesis Contribution

Here we go This is the big one

5.1 Topic Modelling

Model topics?

5.1.1 Hyper-Parameter Tuning

5.1.2 Results

5.2 Topic Centrality

5.2.1 Total Network Topic Centrality

Define what it is, show the full graph again.

5.2.2 Party Leader Topic Centrality

Now, consider a fact stating that relation `r` is total, i.e.,

5.2.3 Results

...upper bound on Embee's performance:

$$\text{upper bound is} \begin{cases} O(bN^2) & \text{if } \mathit{scope} \leq 16 \\ O(bF) & \text{if } \mathit{scope} > 16 \end{cases}$$

5.3 Deep Stochastic Block Models

5.3.1 Artificial Neural Network Adaptation

5.3.2 NetLSD for Describing Political Engagment

5.3.3 Results

Chapter 6

Summary and Conclusions

6.1 Summary

6.2 Future Work

6.3 Conclusion

Bibliography

[1] Linton Freeman. The development of social network analysis. *A Study in the Sociology of Science*, 1:687, 2004.