Modeling Social Media Memes as a Contagious Process

S.Towers[1,*], A.Person[2], C.Castillo-Chavez[1]

1 Arizona State University, Tempe, AZ, USA

2 Some University, Nowhereville, NE, USA

∗ **E-mail: smtowers@asu.edu**

Abstract

In recent years, memes have become increasingly prevalent on the Internet and social media platforms such as Facebook and Twitter. Understanding the dynamics behind the spread of memes has important implications for understanding the spread of general information on social media, and how the spread of such information might be promoted or suppressed.

In this analysis, we use a mathematical model to simulate the spread of memes as a contagious process among a heterogenous population in which one group is quick to be "infected" with the meme (ie; they spread the meme on social media), and quick to "recover" (ie; they quickly cease spreading the meme). A second group in the population is slower to begin spreading the meme, and slower to stop spreading it. We will refer to the former group as "Hip", and the latter group as "Grannies". We show that the contagion model has good explanatory capabilities for describing the spread of popular memes on the Internet.

1 Introduction

Here is where you motivate the analysis: why do we care about memes? While memes might seem like a frivolous topic, understanding the dynamics behind their spread has important implications for understanding the spread of other information, including rumours, on the Internet and social media. There are many people and/or companies that would like to exploit these dynamics to maximize the spread of their message, such as advertising as product, or maximizing their influence on the way a population thinks. In the motivation, we would include references to relevant literature. For instance, Reference [1] used a network model to simulate the spread of information on networks; they developed criteria for identifying influential spreaders of information, and point out that their work has implications for both sociological studies and the design of efficient commercial viral processes. Reference [3] also examines the issue of treating memes as a contagious process.

Here is where you give the objective of the analysis.

Here we will use a mathematical compartmental contagion model of a heterogenous population to simulate the spread of memes in a population that is composed of "Hip" users who are quick to adopt popular practices, but also quick to abandon them in favour of a newly emerging popular practice. The population is also composed of "Granny" users who are much slower to adopt popular practices, but also much slower to realize that the practice is no longer popular among the "Hip", and thus are slower to cease the practice.

We examine various memes that have been popular in recent years, and show that the mathematical model provides a good fit to the data. We estimate that $YYY\%$ of the population are Hip, and the remainder are Grannies. The reproduction number of the memes ranges from YY to ZZ.

In the following sections we describe the sources of data used in this study, and the mathematical model employed, followed by Results and Discussion.

2 Methods

2.1 Data

In this analysis we examine the following memes:

- Good Guy Greg,
- Bad Luck Brian,
- Scumbag Steve,
- Seems Legit.

Data used in this analysis consist of the time series of Google search trend data in the US between 2010 to present for each of the memes, as shown in Figure 1.

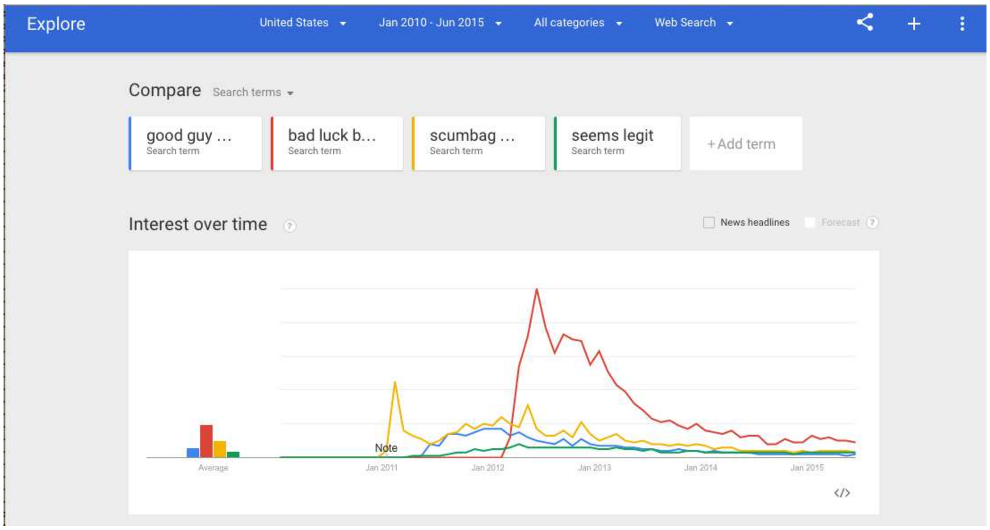

Figure 1: Time series of monthly Google search trend data related to the memes examined in this analysis.

2.2 Mathematical Contagion Model

In our analysis, we model the spread of memes on the Internet using a compartmental Susceptible, Infected, Recovered (SIR) mathematical model with two groups that can be infected. On group (the "Hip") is quick to adopt a meme, but quick to cease spreading it. Another group (the "Grannies") is slower to adopt, and slower to cease spreading it. The model equations are:

$$
\begin{aligned}
S'_{\text{hip}} &= -\beta_{hh} S_{\text{hip}} I_{\text{hip}} / N_{\text{hip}} \\
S'_{\text{granny}} &= -\beta_{hg} S_{\text{granny}} I_{\text{hip}} / N_{\text{hip}} - \beta_{gg} S_{\text{granny}} I_{\text{granny}} / N_{\text{granny}} \\
I'_{\text{hip}} &= +\beta S_{\text{hip}} I_{\text{hip}} / N_{\text{hip}} - \gamma_h I_{\text{hip}} \\
I'_{\text{granny}} &= +\beta_{hg} S_{\text{granny}} I_{\text{hip}} / N_{\text{hip}} + \beta_{gg} S_{\text{granny}} I_{\text{granny}} / N_{\text{granny}} - \gamma_g I_{\text{granny}} \\
R_{\text{hip}} &= \gamma_h I_{\text{hip}} \\
R_{\text{granny}} &= \gamma_g I_{\text{granny}}
\end{aligned}
\tag{1}
$$

In our model, we assume that Hip people can never be infected by a Granny, but Grannies can be infected by the Hip. We also ignore vital dynamics in the population, under the assumption that the duration of the meme "epidemic" is much shorter than the average human lifetime, thus the population size is constant, and equal to $N = N_{\text{hip}} + N_{\text{granny}}$, with $N_{\text{hip}} = S_{\text{hip}} + I_{\text{hip}} + R_{\text{hip}}$ and $N_{\text{granny}} = S_{\text{granny}} + I_{\text{granny}} + R_{\text{granny}}$.

Table 1: Compartments and parameters of the mathematical contagion model of Equations 1. In the analysis, we fit the parameters of the model by comparing the model prediction of the total prevalence, $I_{\text{hip}} + I_{\text{granny}}$, to the observed total prevalence as determined from the Google Trends data.

Compartment	Description
S_{hip}	Number susceptible Hip people
I_{hip}	Number infected Hip people
R_{hip}	Number recovered Hip people
S_{granny}	Number susceptible Granny people
I_{granny}	Number infected Granny people
R_{granny}	Number recovered Granny people
Parameter	
β_{hh}	Transmission rate between Hip people
β_{hg}	Transmission rate from Hip people to Grannies
β_{gg}	Transmission rate between Grannies
γ_h	Recovery rate of Hip people
γ_g	Recovery rate of Grannies people

The parameter β_{hh} is the contact rate, sufficient to transmit infection, among the Hip population. The parameter β_{hg} is the transmission rate between Hip and Grannies, and β_{gg} is the transmission rate among Grannies. The parameters γ_h and γ_g are the recovery rates of the Hip and Grannies, respectively. We assume that $\gamma_h > \gamma_g$.

A summary of the model parameters is shown in Table 1.

Here is where I would also put in a figure showing the compartmental diagram.

Here is also where I would discuss the mathematical analysis of the model, and the expression for the reproduction number, $\mathcal{R}_0$.

2.3 Statistical Methods

To estimate the model parameters that optimally describe each data sample, we use the Monte Carlo method for solution of inverse problems to randomly sample the parameters of the model from broad uniform distributions, and calculate the Pearson χ^2 goodness-of-fit statistic comparing the model to the data sample [2, 5]. The uniform distribution sampling range is chosen to be large enough to ultimately include the parameter optimal value and at least a ± 5 standard deviation range about that value. In order to determine the parameter hypotheses that minimize the Pearson χ^2 goodness-of-fit statistic, this procedure is repeated at least one million times To determine the parameter 95% confidence intervals, the Pearson χ^2 statistic is corrected for over-dispersion using the ansatz of McCullagh and Nelder (1989) [4].

3 Results and Discussion

4 Summary

References

[1] Javier Borge-Holthoefer and Yamir Moreno, *Absence of influential spreaders in rumor dynamics*, Physical Review E **85** (2012), no. 2, 026116.

[2] Glen Cowan, *Statistical data analysis*, Oxford University Press, 1998.

[3] Patrick Grim, Daniel J Singer, Christopher Reade, and Steven Fisher, *Germs, genes, and memes: Function and fitness dynamics on information networks*, Philosophy of Science **82** (2015), no. 2, 219–243.

[4] Peter McCullagh and John A Nelder, *Generalized linear models.*, London England Chapman and Hall 1983., 1989.

[5] Klaus Mosegaard and Albert Tarantola, *Monte carlo sampling of solutions to inverse problems*, Journal of Geophysical Research: Solid Earth (1978–2012) **100** (1995), no. B7, 12431–12447.