Rumor spread in social network

Description of the Problem

Rumors refer to statements or claims of questionable accuracy from no known reliable sources. With increasing number of users on social media networks and its emergence as a platform for content sharing of various types including texts, images and other medias; it accounts for more than 90% overall rumor spreads. Fast and wide spread of news and messages is usually considered as an advantage of online social networks, however rumours and misinformation can also be propagated fast and widely in these networks. Rumour spread leads to decreasing trust to social media content and can also have a considerable impact on society resulting in significant damage and/or monetary cost. Minimizing the spread of rumors is therfore a hot topic in social network analysis.

Explaination why the problem is important for economists

Real share prices and financial performance of firms are also dependent on expectations formed by the investors. In the real world, most investor decisions are heavily based on nudges and their perception of the company based on news and other information obtained through various other mediums. Therefore, as we are quickly moving towards a world where technologies like deepfake videos about politicians or tweets regarding a CEO can have a significant impact on bottom lines of public and private corporations respectively. Therefore, economists need to find a way to effectively quanitfy the impact of these rumours and attempt to measure the elasticity of change in the financial condition of the company based on the intensity of the rumours. We can also check if positive and negative rumours have proportionate or disproportionate impact .

Research Hypothesis

Our initial focus is on the study of rumor spread process in a social network. Our goal would be to identify the underlying paramerers that governs the process of rumor spread using simulation and modelling on collected data. We will further build a classification model to categorize social media content as rumor vs non-rumor.

Research methods (including used data)

The overall methodology is described as follows:

- Construction of graph of connections between twitter users
- Build temporal paths on how rumor spread in this network
- Build a rumor spreading model, considering rumor credibility, correlation between rumors and people's lives and crowd classification based on personality.

We use Pheme dataset to verify the rumor spreading model. First, we obtain categories that correspond to the classes in our model by classifying the tweets in the Pheme dataset according to their properties. Second, we adjust the parameters in our model to make the curve obtained by the model simulation best fit the curve generated by the data from the Pheme dataset. Last, we analyze the reasons for the slight difference between the two curves.We also study about existing rumor classification models.

Table of contents of future thesis:

As a next step, the study focuses on minimizing rumor spreads in social networks. After identifying rumor using our model; we further explore both source-based and source-ignorant approach to minimize its spread. The source based approach focusses on identifying rumor sources and minimizing their spreading ability in their network. Source ignorant approach focuses on decreasing the information flow in the network ignoring the source.