Rumor Cascades and Propagation in Facebook and Twitter

Tara Butler

March 1, 2020

Seminar: How do I lie with statistics

Contents

1 Introduction 3

2 Rumor Cascades 4

2.1 Propagation of information on Facebook 4

2.2 Structure and dynamics of rumor cascades 6

2.3 Summary 9

3 Prominent Features of Rumor Propagation in Online Social Media 10

3.1 Feature Identification 10

3.2 Feature selection and Results 13

4 References 15

1 Introduction

Nowadays, the fastest way for spreading different kinds of information is through the internet. In social networks, it is also possible for individuals to share different information, i.e. spreading and sharing both true and false rumors quickly and widely. The question is, which rumors are spreading faster via social networks? True or false rumors or rumors with a truth value that cannot be clearly assigned? Hence, there are websites like Snopes.com that deal with this question and try to determine the truth of many rumors.

In the following two different papers are presented, which deal with the spreading of rumors in social networks.
The first paper "Rumor Cascades" tries to find the exact way to spread rumors on Facebook, for example by uploading and sharing photos. In addition, the longevity and deletion rate of these are examined in more detail.
The second paper "Prominent Features of Rumor Propagation in Online Social Media" examines the spread of rumors and non-rumors through a commented dataset on Twitter. The three categories temporal, structural and linguistic properties are examined.

2 Rumor Cascades

2.1 Propagation of information on Facebook

The basis for the following investigations of the paper "Rumor Cascades" provided the website Snopes.com, where a corpus with rumors and a sample of reshare cascades belonging to the body are required. The following chapter and its results are based on [1].

First, the three di erent categories true, false and maybe true, i.e. the rumors, for which the truth value could not be assigned exactly, as well as the di erent categories were retrieved. A total number of 4761 rumors were considered. Figure 1 shows the distribution of the rumors and their distribution regarding their truth value. It can be seen that the majority of the rumors from the areas of politics (22%), Fauxtos (12%), i.e. changed images or images with fake background stories and inboxer rebellion (11%), i.e. for example chain letters with ambiguous origin and veracity.

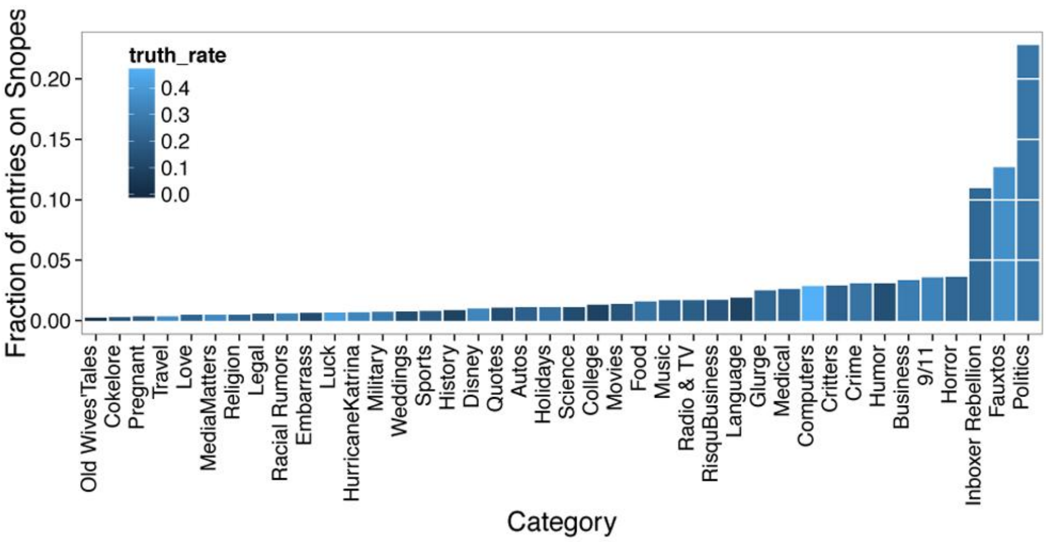


Figure 1: Distribution of rumors [1]

45% of the rumors from the corpus can be assigned to the category false, 26% to the category true and the remaining percent to the category maybe true because either only parts of the rumor are true or the assignment by snopes was not possible.

One possibility for the spread of rumors on Facebook is given through the "share button", which means that content can be shared among many people in a very short time. In the further process, only cascades of publicly visible content are included. First, photos are viewed, whereby the rumors can be part of the picture itself or added text captions. Links published on Snopes are used for a comment that belongs to the original photo or the respective reshare. It should be noted that a photo cascade has comments with different links from Snope due to mixing or incorrect allocation of rumors. It should also be noted that the category and correctness of a rumor can have an impact on whether a share has a comment, since rumors that are incorrectly assigned to the category cause more snopes links, which can be seen in Figure 2.

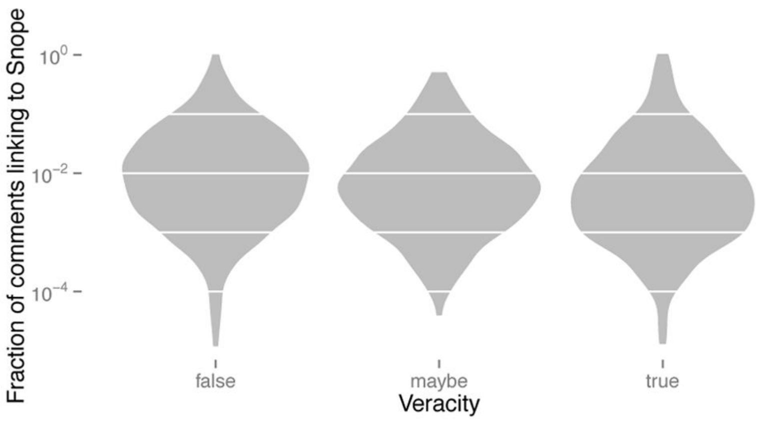


Figure 2: Comments depending on the accuracy of the rumor [1]

The probability that a cascade consisting of n photos and shares is determined can be described by the following equation: $p \quad 1 \quad (1 \quad p_s)^2$ with p_s as the probability of receiving a comment that contains a link to snopes. However, the probability depends on many circumstances, such as whether the rumor is true or false or whether it makes readers curious to know more about the rumor.
For example a photograph of an old 'money bags' text meme states:

This year July has 5 Fridays, 5 Saturdays and 5 Sundays. This happens once every 823 years. This is called money bags. So copy this and money will arrive within 4 days. Based on Chinese Feng Shui, the one who does not copy, will be without money. Figured I'd pass this on!

It was shared 1,259,642 times and received 174,728 comments with 908 linked to Snopes. To estimate the probability of true, false and maybe true content, shares on large cascades (10^4) with a snope comment are considered. The following values therefore result in:

$$p_{true} = 3.0310 \quad ^{3} \qquad p_{false} = 3.4610 \quad ^{3} \qquad p_{maybe} = 3.6810 \quad ^{3}$$

with Figure 3:

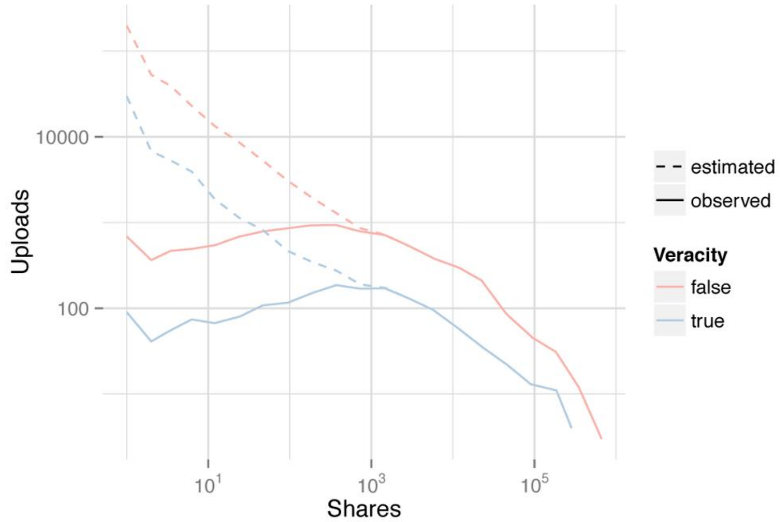


Figure 3: Distribution of the number of shares; estimated and observed [1]

2.2 Structure and dynamics of rumor cascades

The second part examines the e□ect of possible reactions on the spread of rumors, such as the removal of rumors that have been passed on.
As one can see in Figure 1, there are more rumors in some Snopes.com categories than in others. It is striking, however, that 45% of the rumors on Snopes belong to the category false, whereas 62% of cascades on Facebook were classified in this category. Furthermore, 26% of rumors were classified as true on Snopes and only 9% of cascades on Facebook. This can be seen in Figure 4:

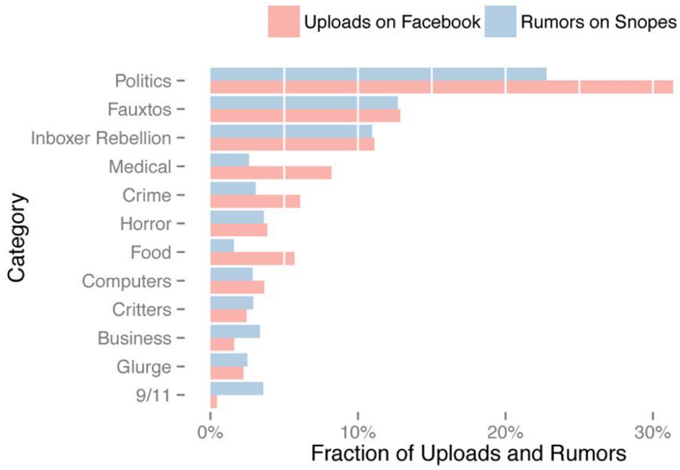


Figure 4: Comparison between rumors on Snopes and cascades on Facebook [1]

It can be seen that false rumors are decisive, but true rumors lead to larger cascades - on average, true rumors have 163 shares per upload and false rumors only 108 shares per upload. However, if the rumors are more popular regarding the number of di□erent cascades, it does not mean that they will also lead to larger cascades. This can be shown with an example: The fauxton category is the second most popular with fake images, but was only shared about 60 times per upload, whereas the less popular category Inboxer Rebellion was shared more than 250 times per upload. This can be seen in Figure 5:

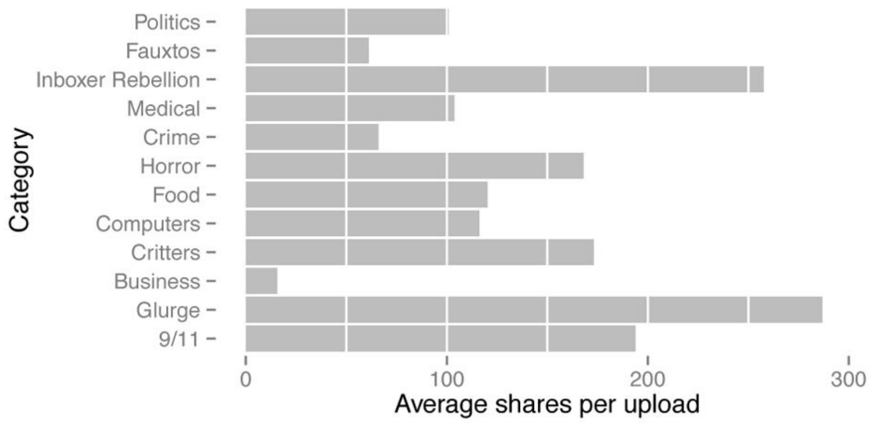


Figure 5: Average of shares per upload [1]

Photos can be uploaded either through users or through pages on Facebook, whereby the possibility of spreading between users and pages di�ers, since users maintain friendships with other users, but can decide whether they grant access to just friends or other users. The posts on pages are public. The investigation below only includes publicly released photos.

Individuals who have spread a false rumor have the option of distancing themselves from it, such as when others comment on the reshare regarding external sources by discussing the truth value of the rumor e.g. when they snope the reshare. However, there is an assumption that in the case of false rumors it leads to a higher deletion rate of reshares to avoid being associated with a false rumor.

As can be seen in Figure 6, the deletion rate for reshares about false rumors when they are snoped, is noticeably higher.

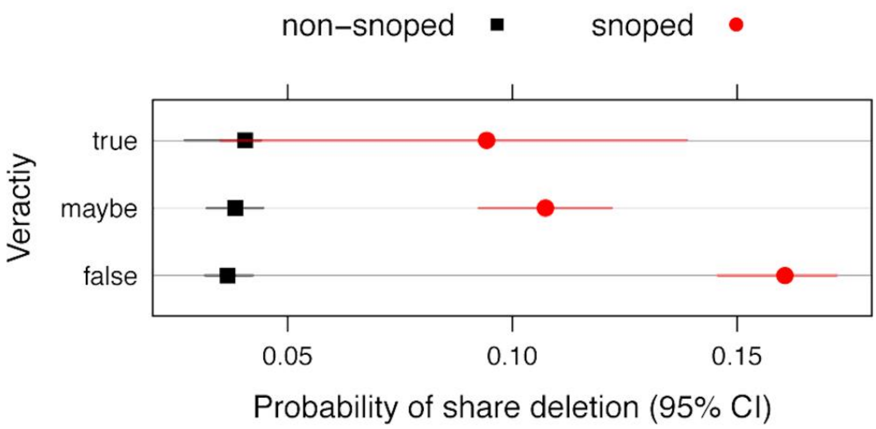


Figure 6: Probability of deletion rate [1]

In addition, because of their higher deletion rate - regardless of whether they are snoped or not - false rumors also have a higher relative risk of deletion than true or maybe true rumors. Furthermore, in figure 7 can be seen that the likelihood of deleting a reshare that was snoped shortly after posting is very high:

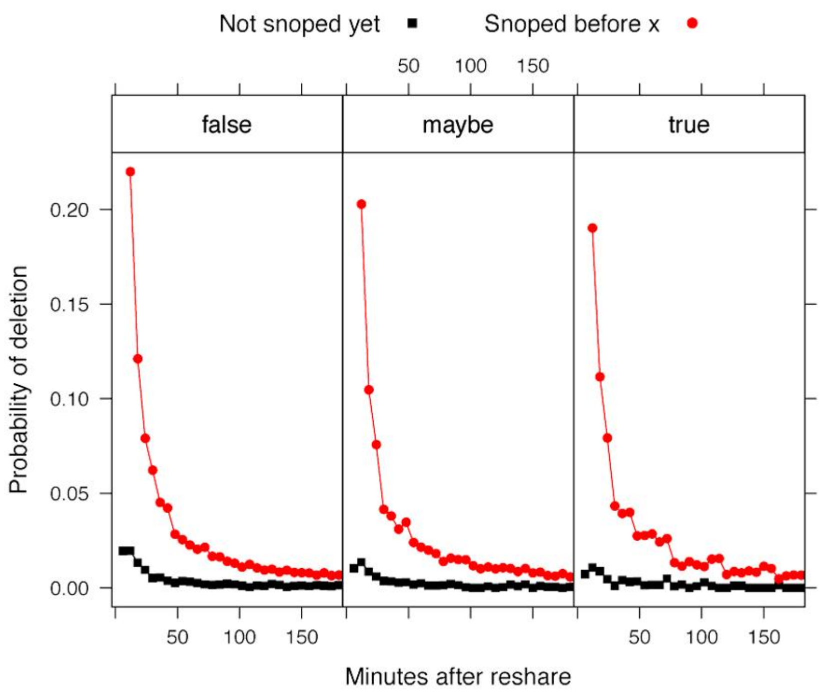


Figure 7: Probability of deletion rate [1]

2.3 Summary

It was shown how quickly content such as photos, regardless of their correctness, can spread on social networks like Facebook. In addition, it was found that while false rumors are uploaded and snoped more often, true rumors showed the greatest cascades. Furthermore it was found out, that the deletion rate for reshares about false rumors when they are snoped, is higher than when not-snoped, but the cascades continue to spread due to the fact that there are many more non-snoped than snoped resahres.

3 Prominent Features of Rumor Propagation in Online Social Media

The basis for the following investigations of the paper "Prominent Features of Rumor Propagation in Online Social Media" provided the website Snopes.com, urbanlegends.about.com, pcmang.com and times.com from which rumors and non-rumors were examined, which were spread during the period of the Twitter dataset from [3]. The following chapter and its results are based on [2].

3.1 Feature Identification

With regard to the spread of rumors, the three categories temporal, structural and linguistic properties are examined.

1. Temporal properties:
Figure 8 shows examples of time series of rumors and non-rumors, whereby it can be clearly seen that rumors have more - and also periodic peaks - peaks than the non-rumors, which are usually characterized by a high peak.

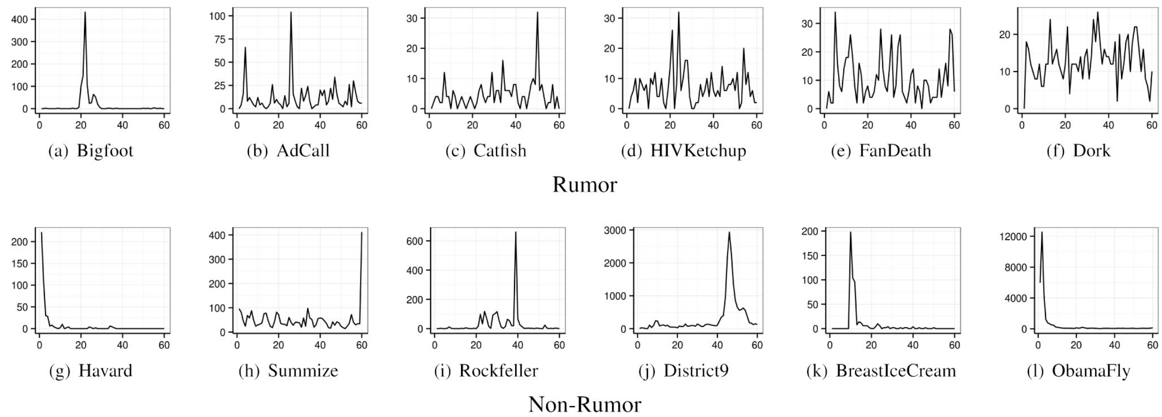


Figure 8: Time series of rumors and non-rumors [2]

The so-called SpikeM model, which is defined in equation (3.1), is used to cover periodic spiky behavior:

$$\Delta B(n+1) = p(n+1) \quad [\frac{\beta}{N} \quad U(n) \quad \sum_{t=n_b}^{n} (\Delta B(t) + S(t)) \quad (n+1 \quad t) \quad ^{1.5} + \epsilon] \tag{3.1}$$

with

$$p(n) = 1 \quad \frac{1}{2} p_a \left[1 + \sin \frac{2\pi}{p_p} (n + p_s) \right]$$

and

$$S(t) = S_b \text{ when } t = n_b, \quad \text{otherwise } 0 \ .$$

U(n) is the number of uninfected nodes at time n and B(n) is the number of newly infected nodes. However, from the parameters of the SpikeM model, no investigation can be made about the peaks of the rumors and non-rumors. For this reason, the SpikeM model is expanded below and defines the so-called Periodic External Shocks (PES) model - a periodic time series model that includes daily and external shock cycles:

$$\Delta B(n+1) = p(n+1) \quad [\frac{\beta}{N} \quad U(n) \quad \sum_{t=n_b}^{n} (\Delta B(t) + \bar{S}(t)) \quad (n+1 \quad t) \quad ^{1.5} + \epsilon] \tag{3.2}$$

with

$$S(t) = S(t) + q(t)$$

and

$$q(t) = q_a \left[1 + \left(\sin \frac{2\pi}{q_p} (t + q_s) \right) \right] \ .$$

q(t) represents the periodic external shock function. Figure 9 shows the different parameters and their meaning.

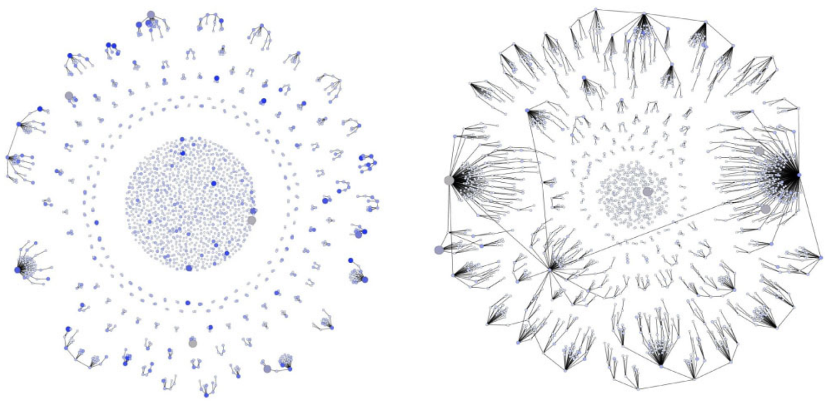
Symbols	Definition
N	Total population of available users
β	Probability of infection
n_b	Starting time of breaking news
S_c	Strength of external shock at birth (time n_b)
ϵ	Background noise
p_a	Strength of interaction periodicity
p_s	Interaction periodicity offset
q_a	Strength of external shock
q_p	Periodicity of external shock
q_s	External shock periodicity offset

Figure 9: Temporal features [2]

2. Structural properties:

In this section two terms are needed: the friendship network and the diffusion network. A friendship network is a subgraph created by users who have posted at least one matching tweet and are connected by a link. Diffusion between two users 1 and 2 means the transfer of content due to the following on Twitter of 2 regarding 1 and the posting of content of 2 with the use of keywords only after 1. The diffusion network is then composed of these diffusions.

In Figure 10 the diffusion networks are depicted once by a rumor (Bigfoot) and once by a non-rumor (Summize) as an example. It can be seen that the Rumor "Bigfoot" has significantly more singletons than the non-rumor "Summize", whereby the edges stand for the occurrences regarding the spread and the nodes for the causes of the spread.



(a) Bigfoot (rumor) (b) Summize (non-rumor)

Figure 10: Diffusion network [2]

3. Linguistic properties:
The Linguistic Inquiry and Word Count (LIWC) - a text analysis program - can assign words based on their psychological category (social, a ective, cognitive, perceptual and biological processes). More information can be found on this website:
http://www.liwc.net/descriptiontable1.php. This program was used for the investigation of di erences between rumors and non-rumors and the respective predominant categories.

3.2 Feature selection and Results

Figure 11 shows a summary of the most important and examined parameters of the three categories and their evaluation:

Symbol	Definition	Type	RF	LR
Temporal features				
q_p	Periodicity of external shock	N	✓	✓
q_s	External shock periodicity offset	N	✓	✓
p_s	Interaction periodicity offset	N	✓	
Structural features				
C_g	Clustering of the friendship network	R		
D_l	Density of the LCC	R		
C_l	Clustering of the LCC	R		
S_d	Fraction of isloated nodes	R		✓
F_d	Fraction of low-to-high diffusion	R	✓	
Linguistic features				
posemo	love, nice, sweet	N	✓	✓
negate	no, not never	R	✓	✓
social	mate, talk, they, child	R	✓	
cogmech	cause, know, ought	N	✓	✓
excl	but, without, exclude	R	✓	✓
insight	think, know, consider	R		
tentat	may be, perhaps, guess	R	✓	✓
see	view, saw, seen	N	✓	✓
hear	listen, hearing	R		

Figure 11: Evaluation of temporal, structural and linguistic features [2]

The abbreviations N and R stand for rumor and non-rumor, whereby the letter that achieved a higher value in the respective categories was chosen. RF stands for Random Forest and LF stands for Logistic Regression.
Regarding the temporal features area, it was found out, that the periodicity of external shocks has given the best prognosis and that rumors are mainly caused by root nodes influenced by external shocks.

In the area of structural properties, it was noticed that the effectiveness of the information flow with regard to the nodes was greatest when there was a diffusion from less influential to more influential people.
In the last category, the linguistic characteristics, it was noticed that the probability of finding positive words in rumors was significantly lower than in non-rumors.

4 References

References

[1] Adamic L.A, Cheng J., Eckles D., Friggeri A., *Rumor Cascades*, 2014

[2] Cha M., Chen W., Jung K., Kwon S., Wang Y., textitProminent Features of Rumor Propagation in Online Social Media, 2013

[3] Benevenuto F., Cha M., Gummadi K., Haddadi H., *Measuring User Influence in Twitter: The Million FollowerFallacy*. In ICWSM, 2010