'Fake News', Rumour Propagation and Source Inference on Social Networks

Network Software Modelling Assignment 2 – MSc Business Analytics (Part-Time)

Name	Student ID
Shane McCarthy	14200512
Richard McAleavey	15204643

23rd April 2017

Introduction: Tackling Fake News on Social Media

There has been widespread concern of late with the supposed phenomenon of 'fake news', often based on the observation that online social networks allow for the propagation of information that is misleading or inaccurate. 'Fake news' is considered to have a negative power over political decision-making and accountability, and people's capacity for rational evaluation and deliberation[1]. There are dire warnings about the destructive and manipulative potential of 'fake news'. Such commentary has intensified with the election of Donald Trump to the US presidency, with firms such as Facebook, for instance, has promised to introduce mechanisms to stem the flow of such information.

No rigorous definition of 'fake news' exists, however. In everyday parlance, the term covers anything from a tweet by Donald Trump to sensationalised content produced primarily to generating advertising revenue. Its meaning is contested: Donald Trump himself, considered by many as a major factor in the propagation of 'fake news', uses his own definition of the term.

Fake news as rumour

In our investigation, we shall consider 'fake news' as a form of *rumour*, to understand how it is propagated, and how its source might be identified. Psychological research has described the propagation of rumours as part of the way humans make sense of the world around them (DiFonzo & Bordia, 2007, cited in (Silverman, 2015)). Locating the source of a rumour need not just be a matter of seeking out a culprit for untruthful information. It can be part of a process of establishing the veracity of information. Many news outlets, given commercial imperatives, are compelled to produce news in line with a 24-hour news cycle and in ways that capture audience attention for advertising purposes. At moments of high uncertainty and instability, the allure of rumours propagated through social media is heightened, with news outlets compete intensely for audience and readership share. This environment has led to high-profile instances of established outlets transmitting false information emanating from unverified sources (Silverman, 2015). A tension will always exist in this environment between the production of new and interesting content on the one hand, and, on the other, practices of verifying information and establishing the reliability of sources. Methods that infer the source of a social media rumour could therefore be used to establish the veracity of the rumour's content, and hence help reduce the risk of a news outlet propagating false information.

Experiment Outline

In this experiment, we explore this phenomenon with an algorithm that is a variation on a Susceptible-Infected (SI) epidemic model. It is intended to be open enough to allow us to understand not just how this or that piece of fake news propagates through a network, but rather what it is about the network, the nodes it comprises, and their relations to other nodes, that determines the propagation of an item, whether the content is 'real' or 'fake'. We simulate the propagation of rumours on different types of network: scale-free graphs based on the Barabási–Albert model, and small-world graphs based on the Watts–Strogatz model, as well as a real-world Facebook network from the Stanford Large Network Dataset Collection. We use this to explore, using a source inference algorithm that draws on previous work by Shah and Zaman (Shah & Zaman, 2011), how the veracity of an information item propagated on social media might be established, through the identification of the source of the item.

[1] See (Wall & Cullen, 2017) for an example of how 'fake news' is cited as a factor in reducing the incidence of HPV virus vaccinations

Experiment Hypothesis

We wish to understand how varying attitudes within networks, and differences in the type of network, affect rumour propagation, and how varying attitudes affect the possibility of inferring the source of a rumour across different types of network. In real-world online social networks, there is a vast range of different kinds of information transmitted, with vast variation in the properties of the networks in terms of properties such as size, density, and clustering. We expect that the distribution of attitudes in a network regarding the information transmitted will decisively affect how rumours are propagated, in terms of both speed and coverage. Furthermore, the distribution of attitudes will affect the effectiveness of rumour source inference. In a network where attitudes are broadly similar, inferring the source is likely to prove a greater challenge.

Simulation Rules

Previous work on information flow in social media has considered the propagation of this kind of information in terms of *rumours*. Models used to understand the spread of epidemics have served as a basis for modelling the spread of these rumours. In our work, we adapt this approach and consider ‘fake news’ as a rumour (whilst acknowledging that rumours often prove true), spread on an undirected graph, as in the case of a Facebook network.

Whereas epidemic and rumour source inference models use a small set of states to describe the behaviour of a node (for instance observed - not observed, or susceptible – infected – recovered – dead), we attach importance to the attributes of the node and how these affect the propagation of the rumour. In the real world, the way people encounter information, and what they do with it, likely depend on how their habits and attitudes have been shaped prior to the encounter, and these habits and attitudes will vary from person to person. Our SI model seeks to account for these habits and attitudes.

Attitude and Suggestibility

Initial attitude

We propose two attributes to determine how likely a person is to propagate information. The first concerns their *initial attitude*. In public discourse on political matters, a person’s attitude to the matter under discussion may be modelled as ranging from strong opposition to strong approval. Factors such as confirmation bias mean that someone who strongly approves of the death penalty may be more likely, for example, to share material claiming a causative relationship between the death penalty and low crime rates. Group polarisation across a network is also likely to affect how information is shared. We are particularly keen to explore how the range of attitudes within a given graph can affect the propagation of a rumour. To this end, we use a variety of distributions -beta, exponential, normal and uniform- to simulate the propagation.

Suggestibility

The second concerns the person’s *suggestibility*: that is, how likely is a person to accept the suggestions of others, and act upon them? In the case of social media interactions, this would be a measure of how likely the person is to share the information they encounter.

Acquired attitude

We assume that a person’s attitude does not remain fixed over time, and that it is affected -to a greater or lesser extent- depending on whether they encounter the rumour. In our model, not all nodes that encounter the rumour will transmit it to their neighbours. To pass the rumour on, the node must meet a set acquired attitude threshold. This acquired attitude is a function of the node’s initial attitude, its suggestibility, and its exposure to the rumour.

Influencers

Some individuals, due to their role within the social network, may have a more decisive effect on the propagation of a rumour than others. They may represent an institution with interests that remain fixed over time, or because their importance is a function of their consistency in the kind of messages they communicate. We model these individuals as a special case of node. We treat their attitude as fixed, and their suggestibility as zero. The measure of their influence is their degree centrality. We identify influencers from those nodes in the top 10% of degree centrality for the network.

How the model works

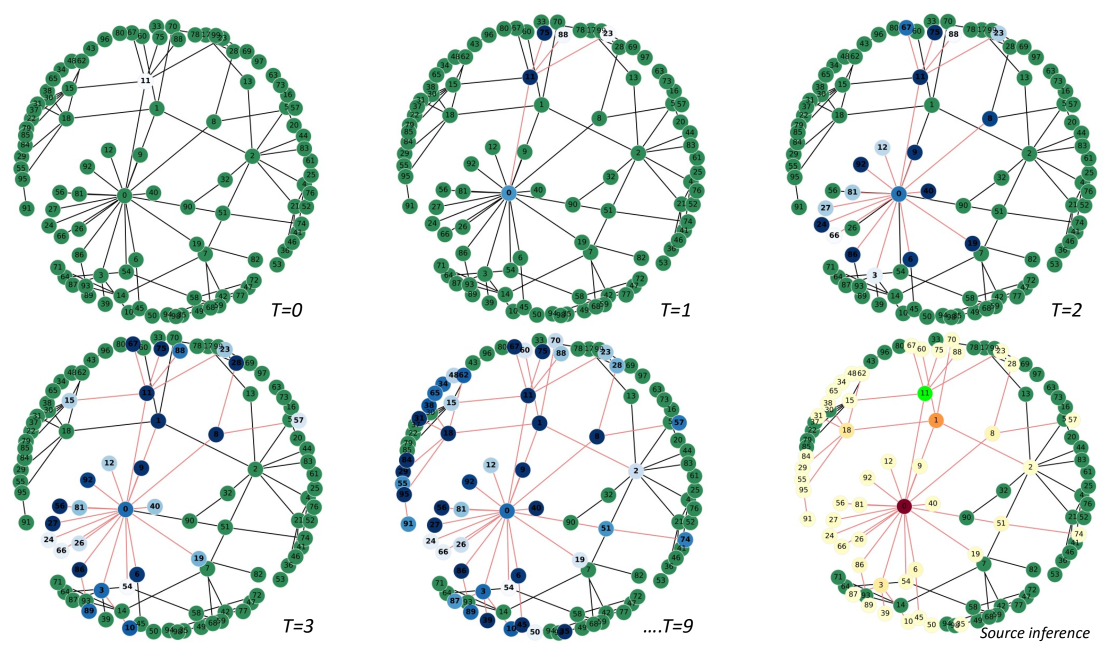


Figure 1 End-to-end rumour propagation and source interference simulation

The model we use is a variation of a standard SI epidemic model adapted to model the propagation of a rumour through a network. An end-to-end rumour propagation and source inference simulation for a Barabási-Albert random graph with 100 nodes is illustrated above in Figure 1. At time step = 0, all nodes are assigned an initial attitude. This is a real value between -1 and 1. They are assigned a *suggestibility* factor between 0 and 1, and a source node is selected at random from the graph (node 11). This is assigned the rumour value, which can be a real value between -1 and 1.

Before the time step increments, the algorithm evaluates how many neighbours of node 11 will be transmitted the rumour. As node 11 has 6 neighbours, and as the probability of those neighbours observing the rumour is set to 0.7 (for each time step), 4 nodes {88,0,75,23} will observe the rumour. The high probability of observing the rumour per hour we assign here is based on empirical findings that Facebook users check their timeline 13.8 times a day on average (Levitas, 2013). Given that the average human is awake for 15 hours a day, 0.7 is a conservative estimate. Once these nodes have observed the rumour, their acquired attitudes are updated using the equation below.

$$Acquired\ attitude = \ initial\ attitude + suggestibility \times rumour$$

$$where\ Acquired\ attitude\ \{\mathbb{R}| -1 \geq x \leq 1\}$$

Only nodes with an acquired attitude that exceeds the transmission threshold can propagate the rumour to their neighbours. In Figure 1, light blue nodes indicate an acquired attitude close to -1 and dark blue nodes indicate an acquired attitude closer to 1. For this set of simulations, we maintain the transmission threshold of 0.25. Hence once the acquired attitude is updated for the 4 nodes {88,0,75,23}, only two of them {0,75} can transmit the rumour to their neighbours. This reflects how we tend to share and retweet information that resonates with our own beliefs and values and pass over information that does not.

This process repeats for each time step until the rumour stops propagating, that is, when all nodes in the graph have observed the rumour (i.e. convergence) or no node that has observed the rumour has an acquired attitude above the transmission threshold. The latter stopping reason can be observed at time step = 9 in Figure 1.

Once the rumour has stopped propagating, we seek to determine the source of the rumour without using any knowledge of the randomly selected source or the time step sequencing of the rumour. To do this we implement Shah and Zaman's method of rumour centrality (Shah & Zaman, 2011) which casts determining the source of a rumour as a maximum likelihood estimation problem.

The key concept behind the rumour centrality algorithm is that whilst we do not know the spreading order (that is, we opt to ignore it for our simulations) the spreading constraints are available to us (i.e. if node A is not connected to C, then A could not have transmitted the rumour to C). For general networks, a breadth-first search heuristic is used to determine the path/tree that is intuitively is most likely to have spread the rumour. All permitted permutations of this tree are then calculated, as there is a unique sequence of nodes for the rumour to spread to each node in the original graph. Rumour centrality can be expressed as:

$$\mathcal{R}(v, G) = \frac{N!}{\prod_{k \in G} T_k^v}$$

where T_k^v is the number of nodes in the subtree rooted at node k, with node v as the source, and N is the number of nodes in the network.

Results

The rumour source inference algorithm was tested on three different graph types: Barabási-Albert random graphs and Watts-Strogatz small-world graphs, with graphs of each type varying in size from 100 to 2000 nodes; and a real-world Facebook network comprising 500 nodes. In total, we ran 3600 simulations. The results are displayed in Table 1 below.

Observations on results

Of the distributions used to assign attitudes to graphs, the exponential distribution corresponds most closely to a network with a high uniformity of opinion. The results show that it is on networks with this distribution that rumours are propagated for longer and spread further, and where the greatest difficulty for detecting the source of rumour exists. This has significant consequences for our understanding of 'fake news' phenomena: it suggests that on networks where people think alike regarding certain topics, the spread of false or misleading information is harder to contain. This corresponds to our intuition regarding confirmation bias. Also of note is the nature of the rumour propagation for exponentially distributed attitudes on Barabási-Albert and Watts-Strogatz graphs. On the former, which are characteristically scale-free, the rumour spreading for exponentially distributed attitudes is prolonged and wide. By contrast, the rumour propagation on the latter, though far more prolonged and wide compared to other distributions, is less extensive, although it propagates for longer.

Table 1: Simulation Results for Rumour Propagation and Source Inference

Simulation Settings			Mean Results		
Graph	**Initial Attitude Distribution**	**Number of Simulations**	**Mean number of time steps until rumour stopped propagating[2]**	**Mean proportion of graph to observe rumour**	**Mean error between actual source and inferred source (number of hops)**
Facebook (500 Nodes)	Beta	100	7.07	59.03%	0.97
	Exponential	100	8.56	98.50%	12.18
	Normal	100	7.80	75.05%	1.89
	Uniform	100	8.13	72.52%	1.16
Barabási-Albert (100-2000 Nodes)	Beta	400	5.23	9.01%	1.25
	Exponential	400	18.49	96.51%	3.53
	Normal	400	8.31	20.91%	1.84
	Uniform	400	5.55	10.34%	1.35
Watts-Strogatz (100-2000 Nodes)	Beta	400	3.68	1.50%	0.97
	Exponential	400	43.73	42.87%	12.18
	Normal	400	6.59	2.66%	1.89
	Uniform	400	4.30	1.59%	1.16

This suggests that the high degree of clustering that is characteristic of small-world graphs can produce environments where rumours persist and propagate extensively within these clusters. Again, this corresponds to our intuition regarding online 'echo-chambers'. We need not conclude from this, however, that such phenomena are a natural feature of online social networks. One area for further inquiry might be to understand what features of the network cause such clustering in the first instance.

Conclusion

It is beyond the scope of this investigation to offer a comprehensive account of attitudes on social networks, as such attitudes depend on the nature of information under consideration. However, we have demonstrated how the attitude of individuals in a network, towards a given item of information, can significantly affect the way information flows through the network. We have also shown how similarity in attitudes throughout a network is a considerable factor in the extent and persistence of rumour propagation, and how it makes the task of identifying the source of a rumour more difficult. Nonetheless, if a rumour can be identified as emanating from a cluster where a high similarity of attitude can be discerned, and this is one potential use of a rumour source inference algorithm, this would prove useful for institutions that seek to preserve their credibility, as a form of insurance against the risks of 'fake news'.

---

[2] Our model allows rumour propagation to cease for two reasons: (1) all nodes in the graph have observed the rumour (i.e. convergence); or (2) no node that has observed the rumour has an "acquired attitude" above the threshold allowed to transmit the rumour to its neighbours. The threshold for this set of simulations was set to 0.25. As that the spectrum of attitude is between -1 and 1, this threshold is relatively low. However, in applying such a threshold to a real-world graph, some additional work would be required to judge the correct setting.

References

Bakshy, E., Messing, S. & Adamic, L. A., 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 05 June, pp. 1130-1132.

DiFonzo, N. & Bordia, P., 2007. *Rumor psychology: Social and organizational approaches.* s.l.:American Psychological Association..

Shah, D. & Zaman, T., 2011. Rumors in a Network: Who's the Culprit?. *IEEE Transactions on Information Theory,* 57(8).

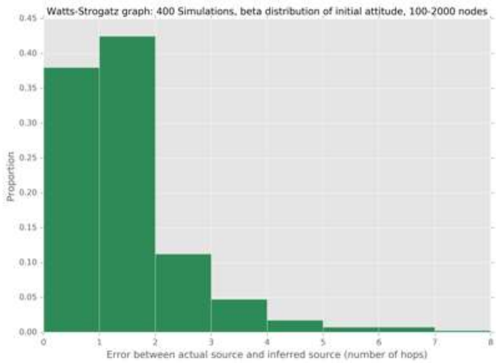
Silverman, C., 2015. *Lies, Damn Lies, and Viral Content: How news websites spread (and debunk) online rumors, unverified claims, and misinformation,* s.l.: Tow Center for Digital Journalism.

Wall, M. & Cullen, P., 2017. 'Uninformed nonsense' about HPV vaccine is endangering lives. *The Irish Times*, 23 April.

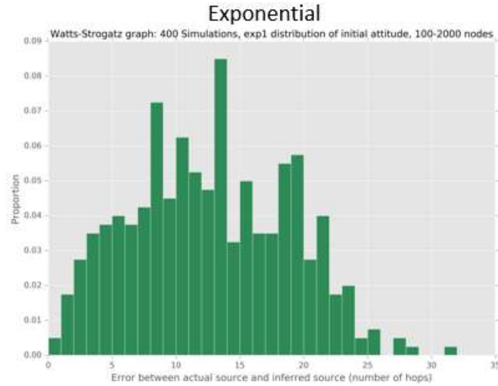
Appendices

Appendix 1: Rumour Source Inference for initial attitudes by distribution on Barabási–Albert graphs

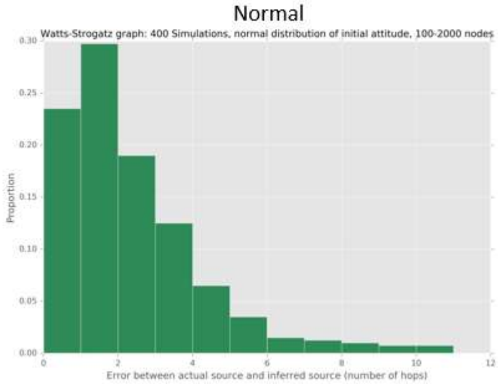
Beta



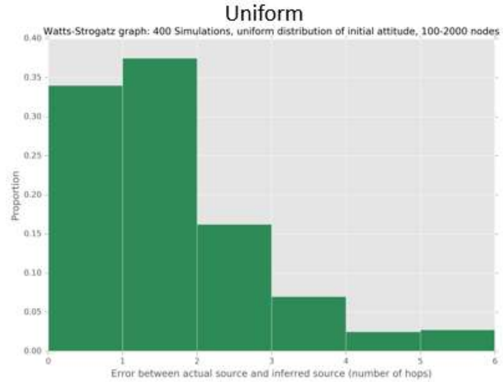
Exponential



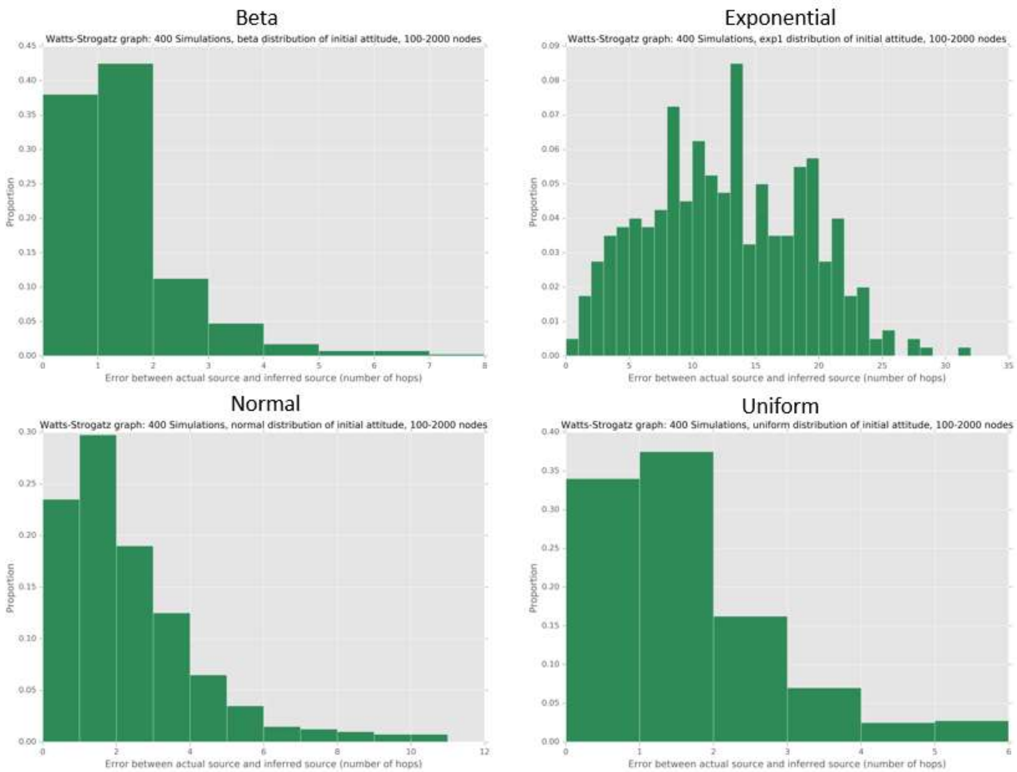
Normal



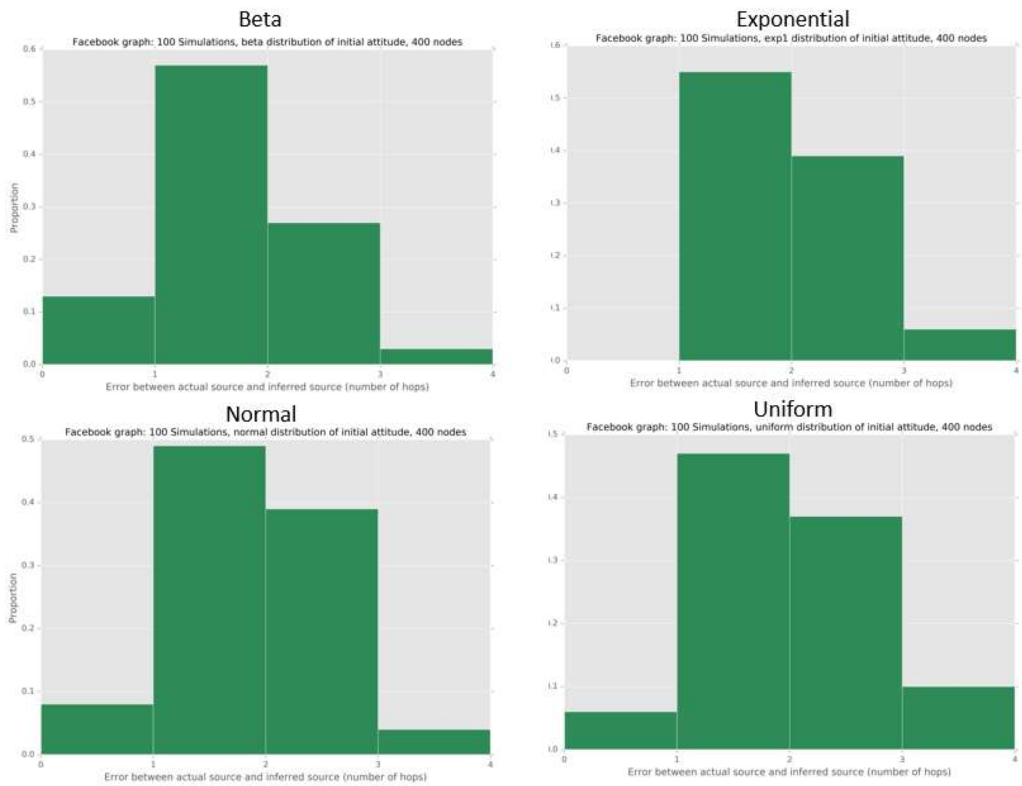
Uniform



Appendix 2: Rumour Source Inference for initial attitudes by distribution on Watts and Strogatz graphs

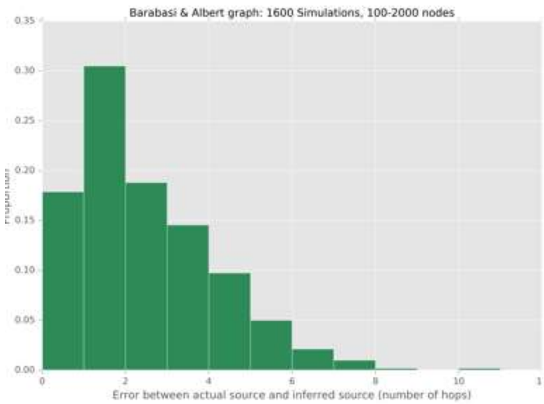


Appendix 3: Rumour Source Inference for initial attitudes by distribution on real world Facebook graph of 500 nodes.

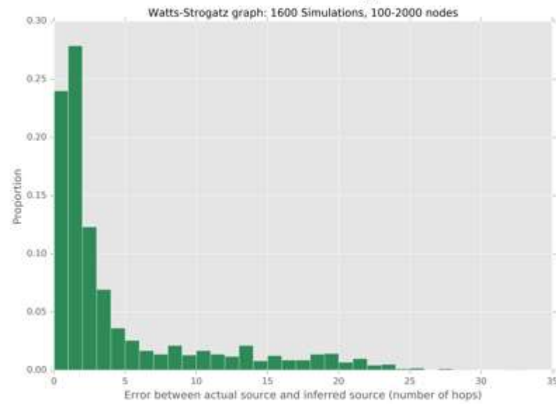


Appendix 4: Comparison of Rumour Source Inference on Barabási–Albert, Watts and Strogatz, and real world Facebook graph

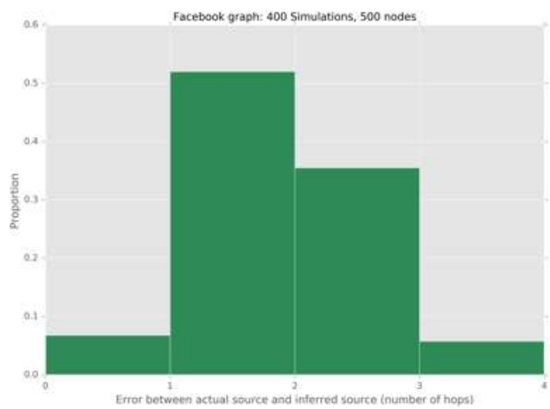
Barabási-Albert



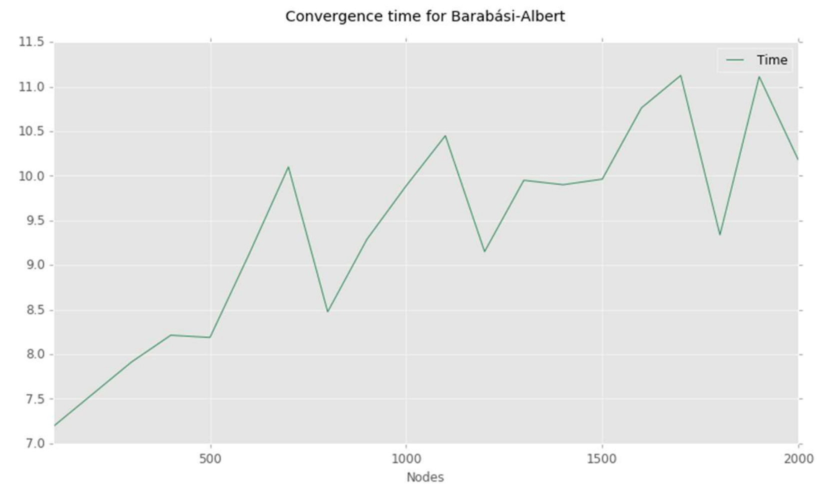
Watts-Strogatz



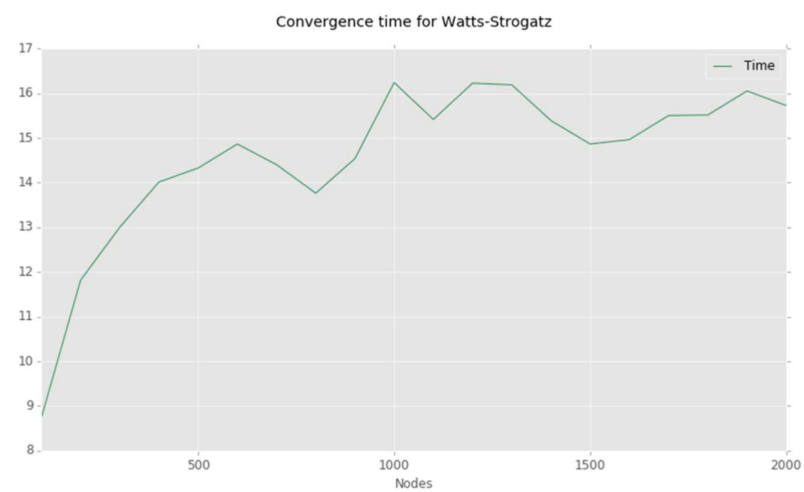
Facebook



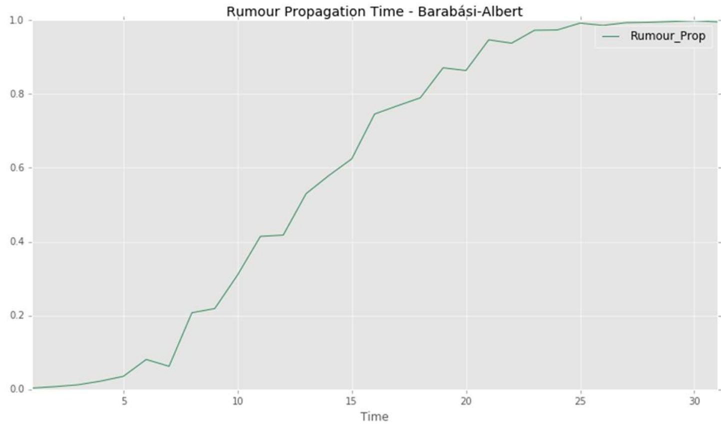
Appendix 5: Convergence time for rumour source inference algorithm for Barabási-Albert graphs

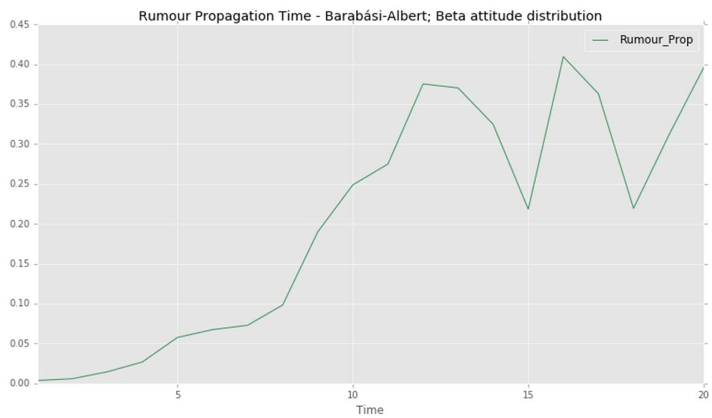


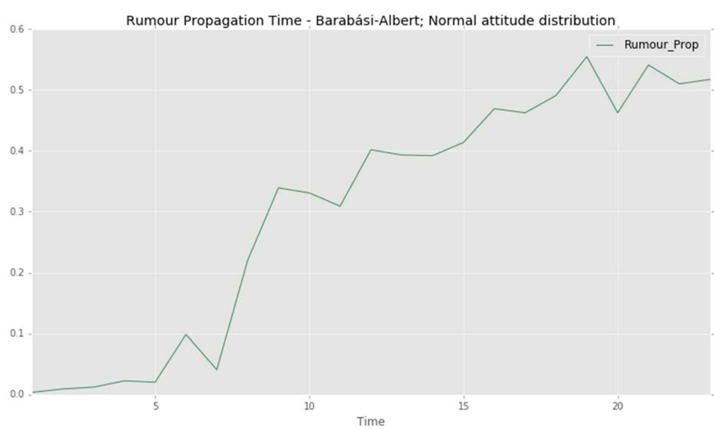
Appendix 6: Convergence time for rumour source inference algorithm for Watts-Strogatz graphs

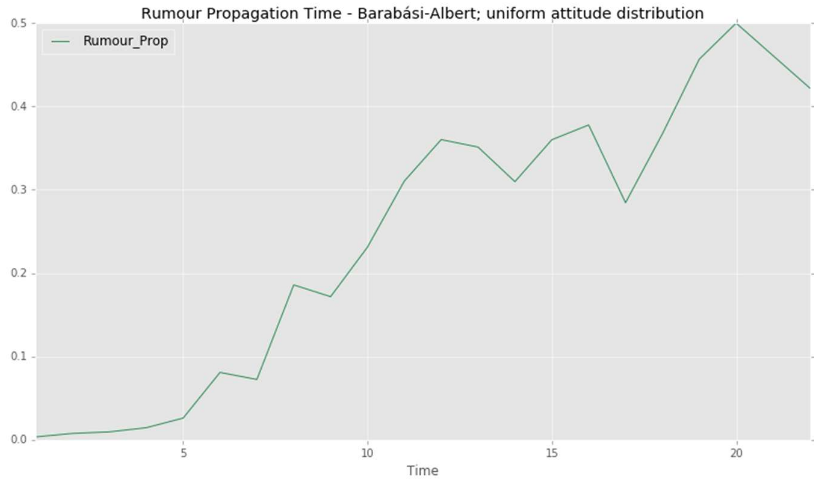


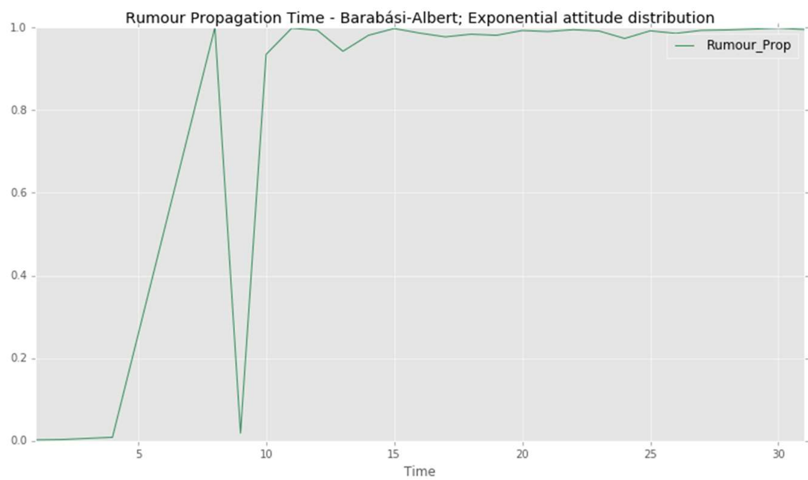
Appendix 7: Rumour Propagation time-steps for Barabási-Albert graphs



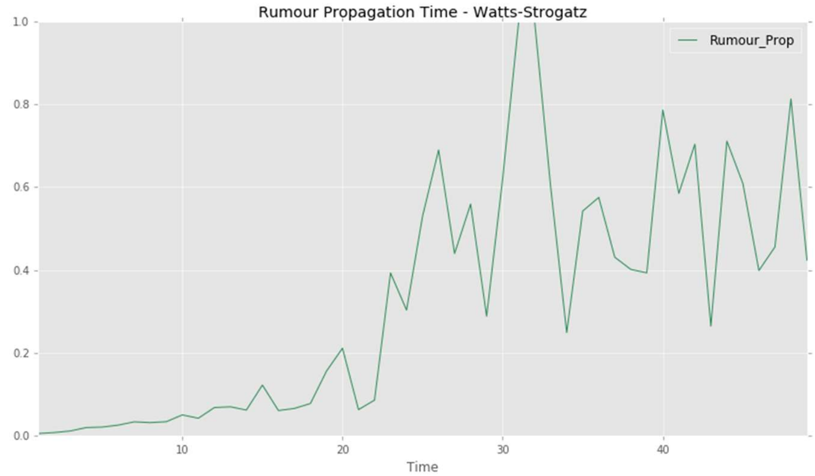


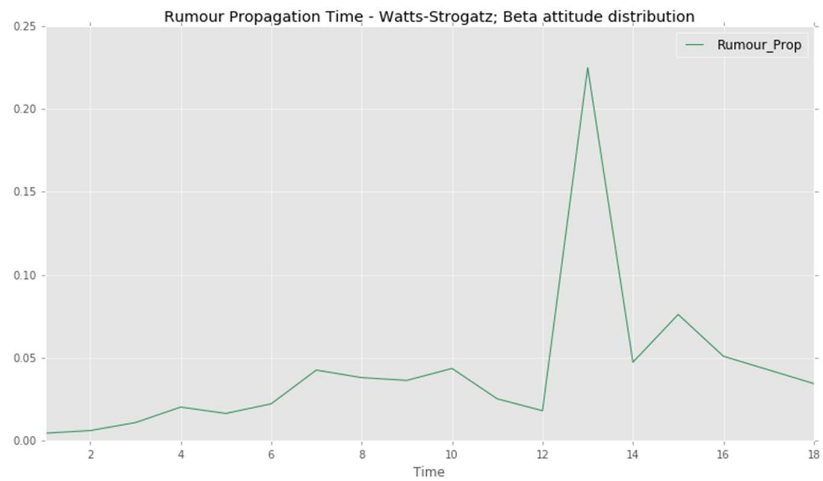


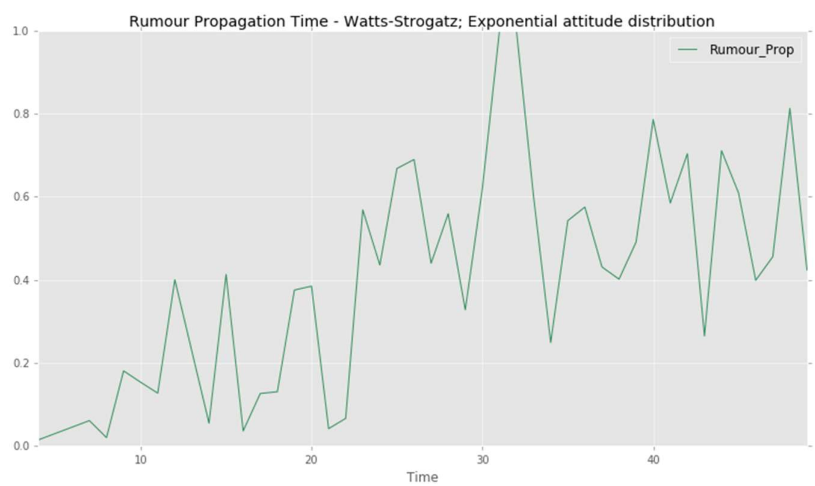

Rumour Propagation Time - Barabási-Albert; uniform attitude distribution
Rumour_Prop
Time

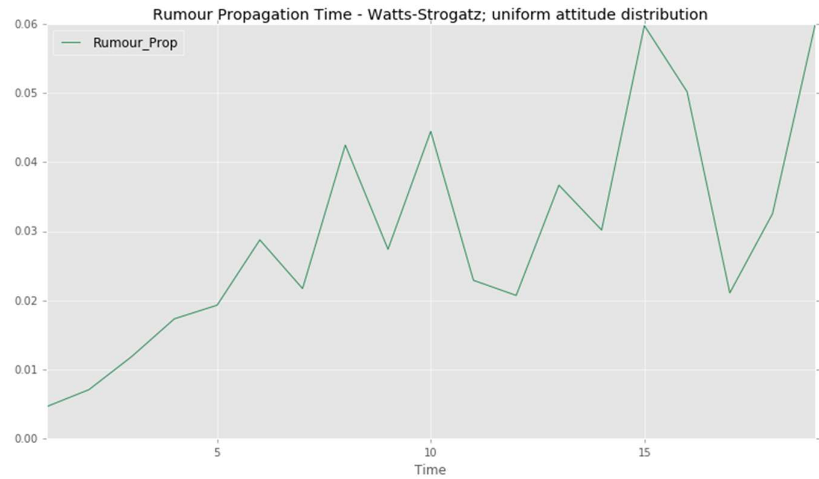

Rumour Propagation Time - Barabási-Albert; Exponential attitude distribution
Rumour_Prop
Time

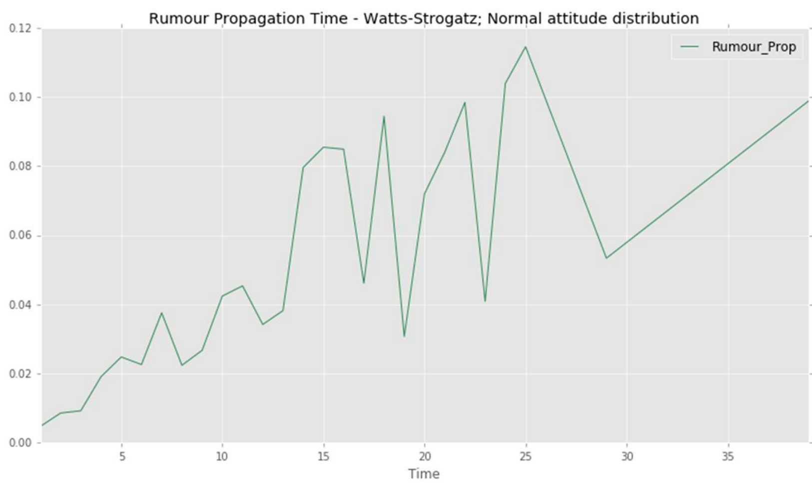
Appendix 8: Rumour Propagation time-steps for Watts-Strogatz graphs



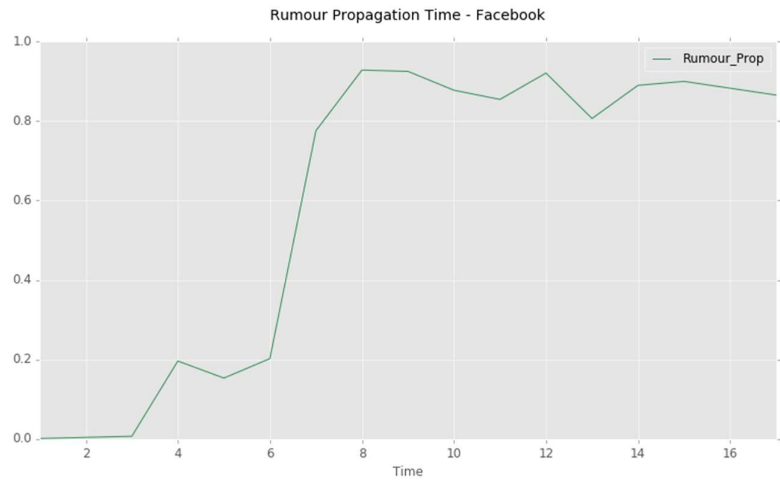


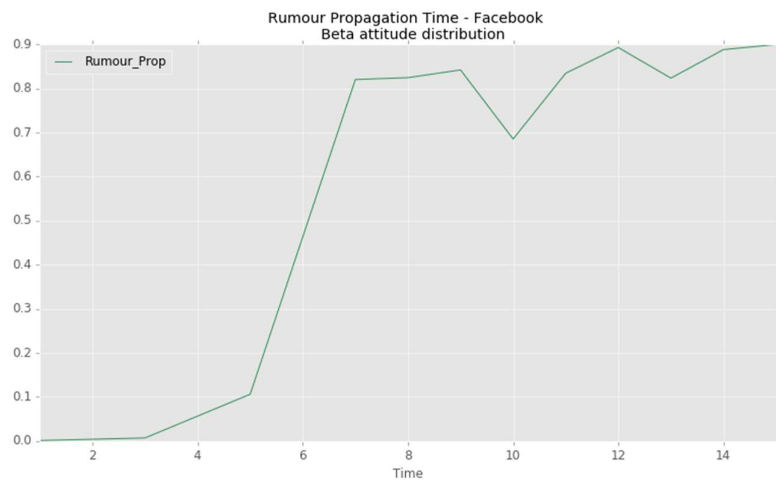


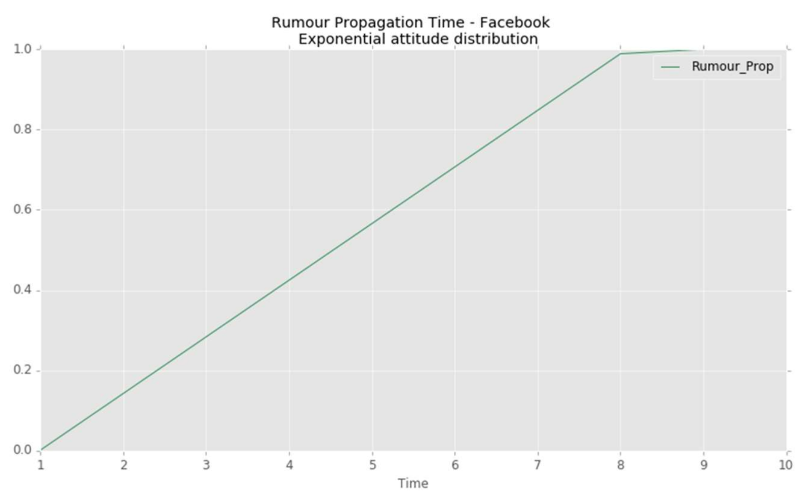

Rumour Propagation Time - Watts-Strogatz; uniform attitude distribution
Rumour_Prop
0.06
0.05
0.04
0.03
0.02
0.01
0.00
5
10
15
Time

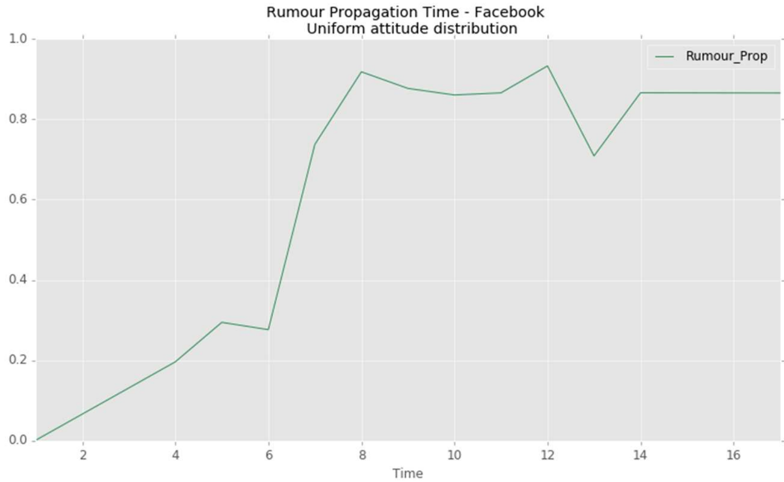

Rumour Propagation Time - Watts-Strogatz; Normal attitude distribution
Rumour_Prop
0.12
0.10
0.08
0.06
0.04
0.02
0.00
5
10
15
20
25
30
35
Time

Appendix 9: Rumour Propagation time-steps for Facebook graph

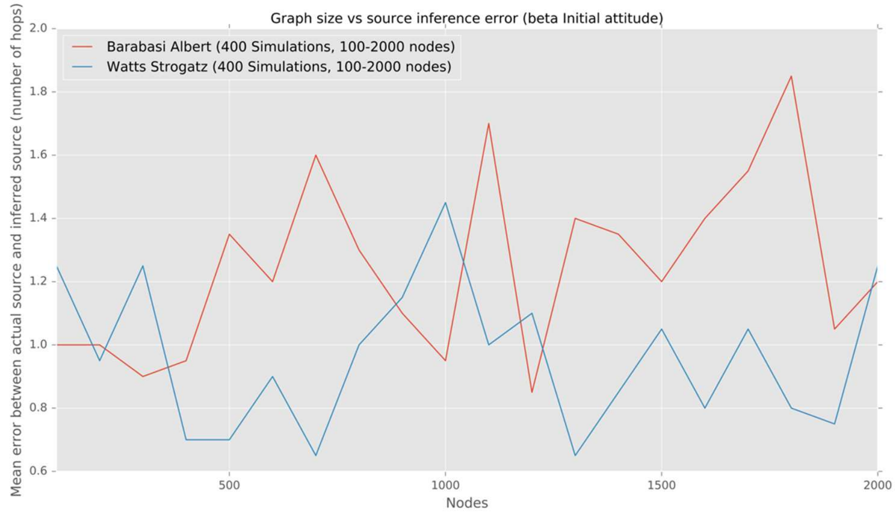


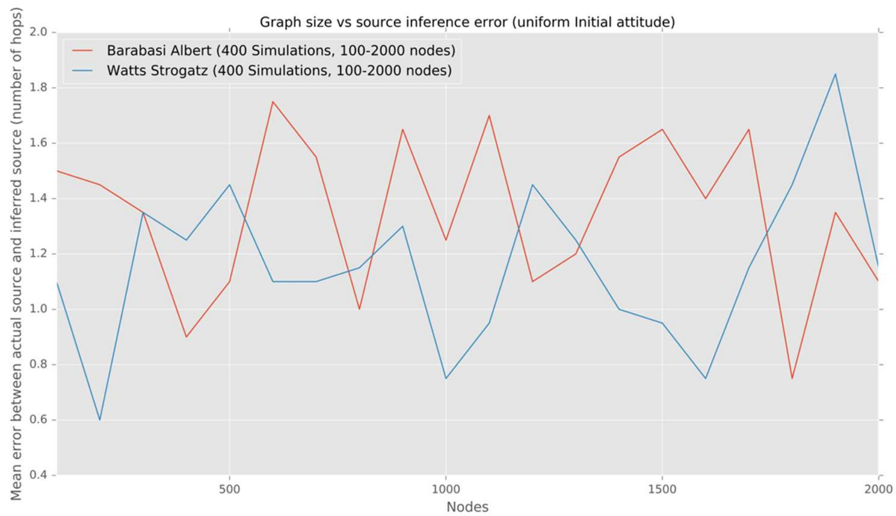





Rumour Propagation Time - Facebook
Uniform attitude distribution
Rumour_Prop
Time
1.0
0.8
0.6
0.4
0.2
0.0
2
4
6
8
10
12
14
16

Appendix 10: Graph Size vs Source Inference Error





Appendix 11: Proportion of Graph Observing Rumour vs Source Inference Error