The Degree-Dependent Threshold Model: Towards a Better Understanding of Opinion Dynamics on Online Social Networks

Ece Çiğdem Mutlu[1,2], Ivan Garibay[1,2,3*]
{ece.mutlu@ucf.edu,igaribay@ucf.edu}
[1]Complex Adaptive Systems Laboratory, [2]Department of Industrial Engineering, [3]Department of Computer Science
University of Central Florida, FL 32816 USA

Abstract

With the rapid growth of online social media, people are increasingly overwhelmed by the quantity and variety of information available online. Researchers from various disciplines focus on understanding the mechanism of information contagion because people are influenced by the opinions they see on social media, and in turn influence others via their own opinions. The threshold model is currently one of the most common methods to capture the effect of people on others' opinions. The heterogeneity in thresholds of individuals is oftentimes poorly defined, which leads to the rather simplistic use of uniform and binary functions that are far from representing reality. In this study, we use 30,704,025 tweets from Twitter to mimic the adoption of a new opinion. Our results show that the threshold is not only correlated with the out-degree of nodes, but also the nodes' in-degree, which contradicts other studies. Therefore, we simulated two separate cases in which thresholds are out-degree and in-degree dependent. We concluded that the system is more likely to reach a consensus when the thresholds are in-degree dependent; however, the time elapsed until all nodes fix their opinions is significantly higher. Additionally, we did not observe a notable effect of the mean-degree on either the average opinion or the fixation time of opinions for both cases, and we noted that increasing the seed size had a negative effect on reaching a consensus. Although threshold heterogeneity has a slight influence on the average opinion, the positive effect of heterogeneity on reaching a consensus is more pronounced when thresholds are in-degree dependent.

CCS Concepts • **Methods** → *Social Network Analysis*; • **Theoretical Foundations** → *Thresholds, Social contagion.*

*Corresponding author: Ivan Garibay[1], Ivan.Garibay@ucf.edu

CSS'10, 2019, Santa Fe, NM

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Keywords consensus, in-degree, out-degree, polarization, social contagion, twitter

ACM Reference Format:
Ece Çiğdem Mutlu[1,2], Ivan Garibay[1,2,3]. 2019. The Degree-Dependent Threshold Model: Towards a Better Understanding of Opinion Dynamics on Online Social Networks. *Submitted to The Computational Social Science Society of the Americans (CSS'10).* CSS, Santa Fe, NM, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

While network study is not new, its focus has shifted from physical proximity and socio-economic networks to social media based networks. This change is arguably the product of the fast-paced information flow that is engendered by the technological advances of the 21st century, and the resulting impact on the needs and lifestyles of people. The need to address the newly emerged phenomenon of the creation and dissemination of information on social media has amplified interest in the field of network science. Indeed, network science applications have extended to the field of marketing [2, 7], sociology [6, 10], political science [17], physics [16], economics [13], machine learning [14] and biology [11], all attempting to reveal the interdependency between units of interest. Social media has allowed people's opinions to be voiced freely, with far reaching consequences; this phenomena has affected many things including the shift to online marketing and social media based political campaigning. The interplay between information from average users and bigger entities is much larger than in the past due to social media. This reciprocity in information flow, the increase in the volume of information received and sent, and the ease of relaying information has made it imperative that researchers understand the dynamics of information and opinion formation, propagation, and exchange [1, 20, 21, 23].

In the mid-20th century, sociology pioneered the development of information and opinion diffusion as a subject of study. One of the early studies is the Markovian linear threshold model introduced by Granovetter [5]. According to the threshold model, individuals adopt a new opinion only if a critical fraction of their neighbors have already adopted the new opinion. Granovetter suggests that the threshold

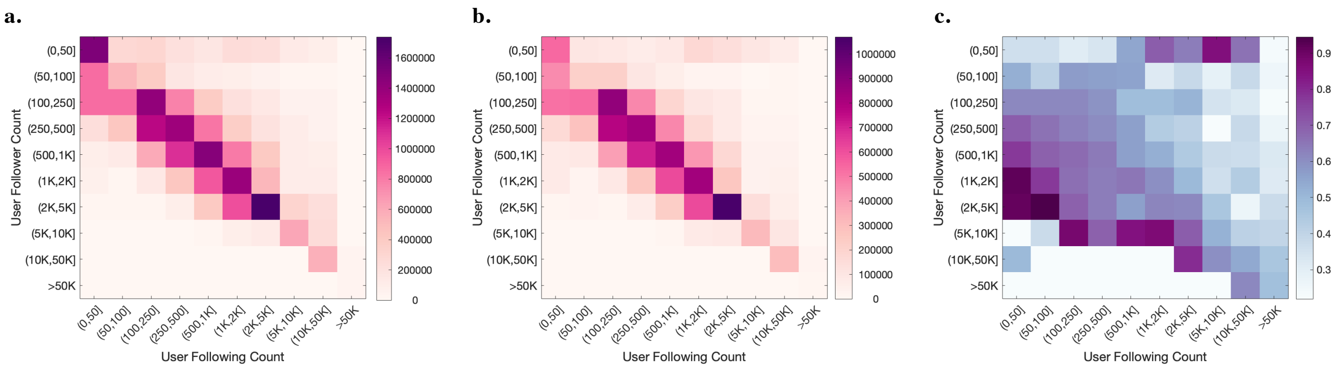


Figure 1. a. Number of users b. Number of retweeters c. The retweeting probability of users in each cluster -the element-wise ratio of number of retweeters in b to the number of users in a.

of individuals can be different, and are influenced by demographic and psychographic factors. However, this heterogeneity among researchers is poorly-defined, which leads to an extensive use of uniform [12, 18, 19] and binary [22] thresholds in many studies. Arguably, this assumption of homogeneous or binary thresholds is an oversimplification of reality and may produce misleading results. To remedy this oversimplification and thereby provide a more holistic and accurate model, more complex threshold models such as the tent-like function [24], the truncated normal distribution function [8] or the sigmoid function [3] are also used in the literature. Our Twitter data mining results show that the threshold of an individual for adopting a new opinion (retweeting a tweet) is affected either by their out-degree (number of followers) or in-degree (number followed). Some studies have already employed degree-dependent threshold models in explaining the dynamics of information diffusion [4, 9], however the degree dependency of an individual's threshold is associated only with their out-degree. Additionally, these studies have implemented threshold heterogeneity by using custom threshold functions, which renders the results less robust and less reliable. Therefore, we seek to analyze the sensitivity of information diffusion dynamics to in-degree and out-degree dependencies of thresholds. Another purpose of this study is to understand how threshold heterogeneity and network properties (seed size, mean-degree) affect information diffusion dynamics when thresholds are in-degree or out-degree dependent.

2 Methods

2.1 Data set and Twitter Analysis Results

The Twitter data set used for this study contains 30,704,025 tweets from the cybersecurity-related events from March 2016 to August 2017, of which 16,884,353 are retweets. We first collected follower and following counts of each user to relate the retweeting probability of users with the two aforementioned counts. We then generated a matrix of rows representing follower count clusters, and columns representing following count clusters of all users in our data set (Figure 1.a). We also filtered users who have retweets only and generated the same matrix (Figure 1.b). Preliminary results showed that a majority of users are clustered around the areas where follower and following counts are not extreme, and the matrix of retweeted users show a similar pattern. Since retweeting probabilities of users in each cluster are not clear from these matrices only, we calculated the element-wise division of these two matrices to figure out the ratio of number of retweeters to the number of all users in each cluster. Results show that the retweeting probability of users who have a relatively low following count is higher, i.e., the threshold of a node seems to be positively correlated with its out-degree. On the other hand, the effect of varying follower count on the retweeting probability is unclear since the left-bottom of the matrix is empty (Figure 1.c). To remedy this, we extracted the 3 most retweeted tweets (RT1, RT2, RT3) of retweet sizes 138,969, 58,546, and 57,280, respectively. We divided users into 8 clusters with respect to their follower (Figure 2.a-2.c) and following counts (Figure 2.d-2.f) independently rather than jointly. For each cluster, we calculated the ratio of the number of users who retweeted RT1, RT2, or RT3 to the number of all users, respectively, as in Figure 2.c. The only difference is that instead of all retweeters, we focused on retweeters of RT1, RT2 and RT3. Thus, we could prevent the masking effect of non-active users in the whole data set. The results show that both follower and following counts have a negative effect on the retweeting probability of users. Furthermore, we applied a one-sided Chi-square test ($\alpha = 0.05$) to determine whether this decreasing pattern is statistically significant. We included relative χ^2 values if the retweeting ratio in the cluster is significantly higher than that of next cluster (p-value is lower than 0.005). We observed that the retweeting probability decreases when follower count increases and this decreasing pattern is significant for almost all consecutive clusters. Nevertheless, the decrease between the consecutive

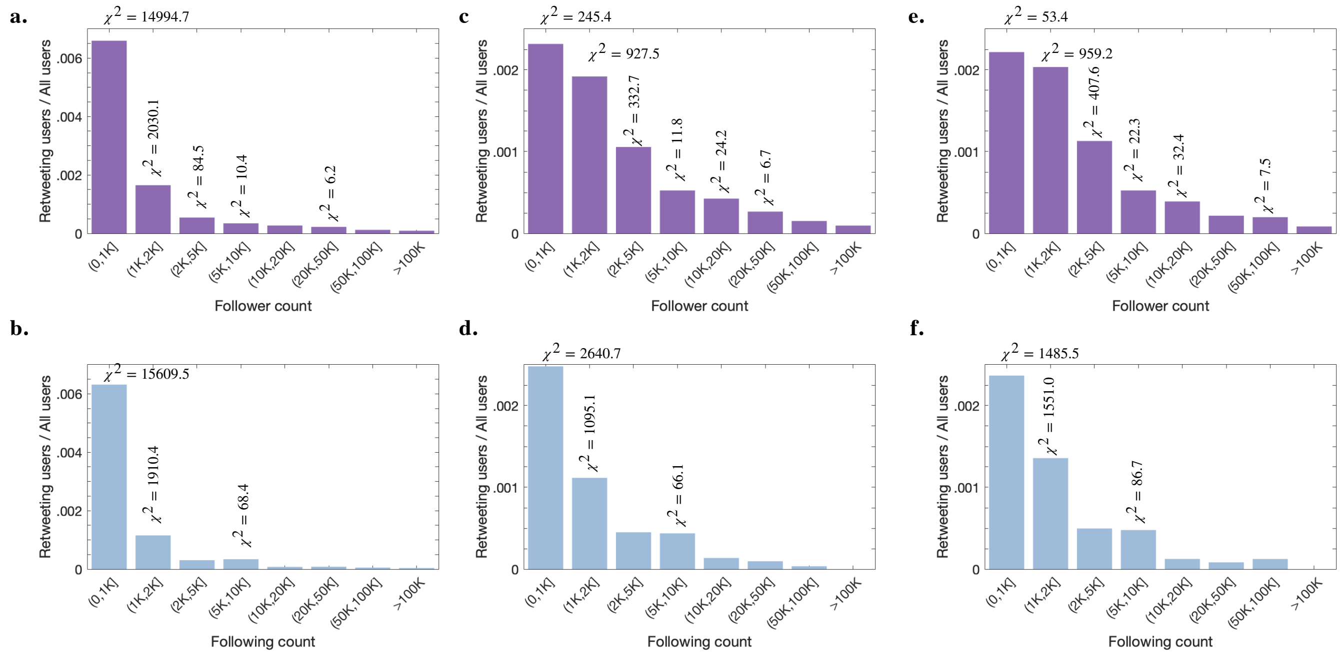


Figure 2. The ratio of retweeters of RT1, RT2 and RT3 to all users when we cluster users according to the their follower counts and following counts independently. Bar plots in a and b shows the results for RT1 users, c and d for RT2 users and e and f for RT3 users.

clusters defined by following counts were significant only when following counts are not high. This is probably because clustering users according to their follower and following counts with the same limits has a notable effect on the test statistics. Distributions of follower and following counts of users are not similar in the data set, i.e, the .8 and .9 quantiles and the maximum of the user following counts are 1916, 3860, and 3,136,215, while it is 2332, 5639, and 94,833,565 for user follower counts. When we decrease the number of clusters from 8 to 3 ($(0, 1K], (1K, 10K], (> 10K]$), we observed that the retweeting probability decreases when following count increases, and this decreasing pattern is significant for all consecutive clusters ($\chi^2 = \{2602.8, 18087, 272.3\}$ for RT1, $\{2762.3, 1807.2, 52.0\}$ for RT2, and $\{1299.3, 987.1, 87.1\}$ for RT3). Thus, our data analysis shows that thresholds of individuals to accept a new opinion is positively correlated with their in-degree and out-degree.

2.2 Generating Networks

The main aim of this paper is to investigate the effect of threshold heterogeneity on opinion spreading dynamics when thresholds are correlated with the degree-distribution of the nodes. For this purpose, we generated power-law distributed random numbers (x_i) to further assign them to the desired degree-distribution of the network. To understand the effect of out-degree dependent threshold and in-degree dependent threshold on the dynamics of opinion spreading separately, we created two independent networks:

1. **CASE I:** Out-degrees of the nodes (k_{out}) are power-law distributed and has the form $\sqrt{N}x^{\gamma}$ and in-degrees are kept constant (M_{in}).
2. **CASE II:** In-degrees of the nodes (k_{in}) are power-law distributed and has the form $\sqrt{N}x^{\gamma}$ and out-degrees are kept constant (M_{out}) .

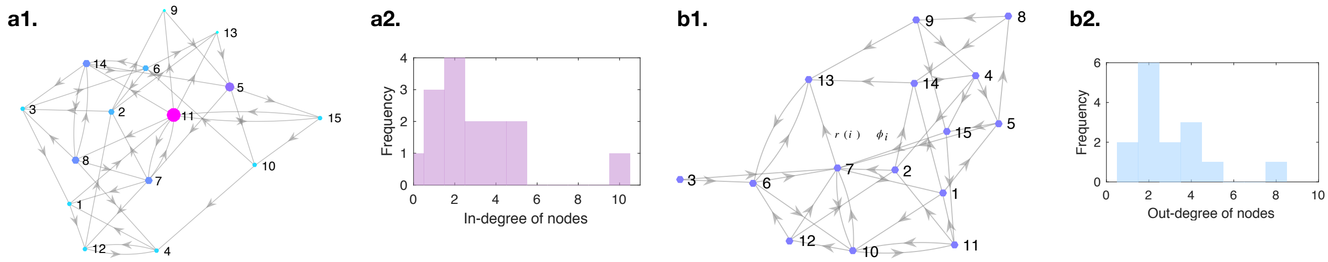
Here, N denotes number of nodes in the network (seed size) and $\gamma = 3$ for both cases for a fair comparison. Then, we added directed links between randomly selected node pairs (i, j) by employing configuration model [15] if $i \neq j$ and $k_{out} < x_i$ for Case 1, $k_{in} < x_i$ for Case 2. This wiring process continued until all possible links are formed. In this network structure, self-edges are not allowed while multiple edges between same node pairs are possible. Since total in-degree in the network should be equal to the total out-degree in the network and total in-degree is equal to total out-degree, one can easily realize that the mean-degree of the network is equal to:

1. **CASE I:** Fixed in-degree (M_{in}).
2. **CASE II:** Fixed out-degree (M_{out}).

2.3 Assigning Thresholds

After generating networks, we employed the degree-dependent threshold model by assigning the threshold of node i to accept a new opinion (ϕ_i) as correlated with:

1. **CASE I:** its out-degree .
2. **CASE II:** its in-degree.



a3.

	out-degree	in-degree	$r(i)$	ϕ_i
1	2	3	8	0.75
2	3	3	6	0.75
3	2	3	9	0.75
4	2	3	10	0.75
5	5	3	2	1
6	3	3	7	0.75
7	4	3	3	1
8	4	3	4	1
9	1	3	14	0.5
10	2	3	11	0.5
11	8	3	1	1
12	2	3	12	0.5
13	1	3	15	0.5
14	4	3	5	1
15	2	3	13	0.5

b3.

	out-degree	in-degree	$r(i)$	ϕ_i
1	3	2	11	0.5
2	3	3	6	0.75
3	3	0	15	0.5
4	3	2	10	0.75
5	3	2	9	0.75
6	3	5	3	1
7	3	10	1	1
8	3	1	13	0.5
9	3	1	14	0.5
10	3	4	4	1
11	3	2	8	0.75
12	3	1	12	0.5
13	3	5	2	1
14	3	4	5	1
15	3	3	7	0.75

Figure 3. The representation of network when $N = 15$ and a1. out-degrees are power-law distributed and in-degrees are kept constant as $M_{in} = 3$, b1. in-degrees are power-law distributed and out-degrees are kept constant as $M_{out} = 3$. Histogram plots of a2. out-degrees in a1, b2. in-degrees in b1. In addition to out-degree and in-degree of nodes, their ranks $r(i)$ and thresholds ϕ_i are also given in the table for a3. the network in a1, b3. the network in b1.

Since threshold heterogeneity is one of our main concerns in this study, we divided nodes into N_{th} groups by their ranks which can be obtained by sorting their

1. **CASE I:** out-degrees.
2. **CASE II:** in-degrees.

Then, we assigned thresholds as evenly spaced N_{th} points between 0.5 and 1 to prevent the confounding effect of the mean-threshold, i.e. the average threshold is always constant as 0.75. Thus, increasing N_{th} yields more heterogeneity among thresholds of individuals.

$$\phi_i = \begin{cases} 0.5 & \text{if } r(i) \leq \frac{N}{N_{th}} \\ 0.5 + \frac{0.5}{N_{th}-1} & \text{if } \frac{N}{N_{th}} < r(i) \leq \frac{2N}{N_{th}} \\ \dots & \dots \\ 0.5 + \frac{0.5(N_{th}-2)}{N_{th}-1} & \text{if } \frac{(N_{th}-2)N}{N_{th}} < r(i) \leq \frac{(N_{th}-1)N}{N_{th}} \\ 1 & \text{if } \frac{(N_{th}-1)N}{N_{th}} < r(i) \leq N \end{cases}$$

where $r(i)$ represents the rank of the node when they are sorted according to their out-degree in Case 1 and their in-degree in Case 2.

An example of network generation for two cases, out-degrees and in-degrees and relative threshold values of the nodes are shown in Figure 3.

2.4 Running Simulations

We initialized the opinions of individuals as a Bernoulli distributed random variable with an initial probability (p), i.e. the opinion of the node i (s_i) might equal to 1 with a probability p and equal to 0 with a probability $1-p$. We assumed that the opinion change process is reversible; thus, individuals may change their opinions continuously rather than only once.

After generating the network, assigning thresholds, and initializing the opinions, we ran the opinion change simulations. The process of updating their opinions is as follows:

1. Picking a node i randomly.
2. Calculating the weighted average of the opinions of its in-neighbors ($\bar{o}_i$). Here, weights are the multiple edges formed between node i and its neighbors.
3. Updating the opinion of node i (s_i) according to the criteria as follows:

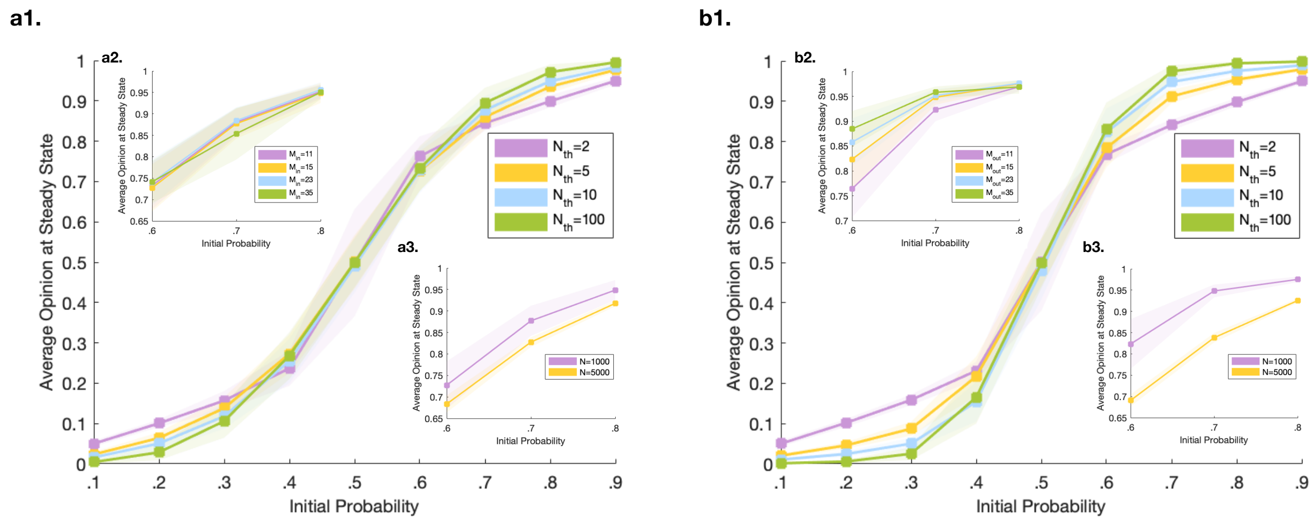


Figure 4. Simulation result of average opinion at steady state as a function of initial probability (p) a1. with varying threshold heterogeneity (N_{th}) when $N = 1000$ and $M_{in} = 15$, a2. with varying in-degree (M_{in}) when $N = 1000$ and $N_{th} = 10$ and a3. with varying seed size (N) when $M_{in} = 15$ and $N_{th} = 10$ if thresholds are out-degree dependent, and out-degrees are power-law distributed. Additionally, the simulation result of the average opinion at steady state as a function of p b1. with varying N_{th} when $N = 1000$ and $M_{in} = 15$, b2. with varying out-degree (M_{out}) when $N = 1000$ and $N_{th} = 10$ and b3. as a function p when $M_{out} = 15$ and $N_{th} = 10$ if thresholds are in-degree dependent, and in-degrees are power-law distributed.

a. **if** $s_i = 0$ and $\bar{o}_i - s_i > \phi_i$,
then $s_i = 1$ in the next step.
b. **if** $s_i = 1$ and $\bar{o}_i - s_i < -\phi_i$,
then $s_i = 0$ in the next step.

This Markovian chain is repeated until all possible opinion changes are made and individuals fix their opinion. We carried out all the simulations on MATLAB and repeated these simulations 50,000 times.

3 Simulation Results

In the current study, we first aimed to analyze the effect of in-degree and out-degree dependence of thresholds on the average opinion at steady state ($\bar{s}$). Therefore, after all the individuals fix their opinions in a network, we averaged their opinions by using the equation below:

$$\bar{s} = \frac{1}{N} \sum_{i}^{N} s_i \tag{1}$$

where s_i is the opinion of node i at steady state. We conducted our simulations to measure $\bar{s}$ as a function of the initial probability (p) with a varying mean-degree (M_{in}/M_{out}), seed size (N), and threshold heterogeneity N_{th}. Line plots in Figure 4 represent the expected value of 50,000 Monte Carlo simulations, and shaded areas with the same colors denote the relative one standard deviation from the expected value of these simulations. Here, Figure 4.a1-4.a3 shows the simulation results when thresholds are out-degree dependent and out-degrees are power-law distributed, while in-degrees are kept constant. Figure 4.b1-4.b3, on the other hand, indicates the simulation results when thresholds are in-degree dependent and power-law distributed, while out-degrees are kept constant.

Figure 5 shows the time elapsed until all individuals fix their opinion (t_f) as a function of p with varying M_{in}/M_{out} and varying N_{th}. We did not simulate the effect of varying N on t_f since it is already clear that increasing the seed size would cause more deviation in the opinions and increases t_f.

Figure 4.a1 and 4.b1 show $\bar{s}$ as a function of p at various N_{th} values. Since the standard deviation of the simulations are highest in the range $0.35 \lesssim p \lesssim 0.65$, we especially focus on the results when $p \lesssim 0.35$ and $p \gtrsim 0.65$. In general, the system is more likely to reach a consensus when thresholds are in-degree dependent, and there is a clear asymmetry before and after $p = 0.5$ in both cases. Therefore, we just focused on the region $0.6 \leq p \leq 0.8$ for further analysis. Although threshold heterogeneity of nodes in the system has a slight effect in the resulting average opinion when thresholds are out-degree dependent, we can conclude that the probability that the system reaches a consensus increases as the threshold heterogeneity increases; this increase is more pronounced when thresholds are in-degree dependent, e.g. $\bar{s} = 0.8412$ when $N_{th} = 2$, while $\bar{s} = 0.9742$ when $N_{th} = 100$ at $p = 0.7$ (Figure 4.b1). When it comes to the effect of heterogeneity on the opinion fixation time, t_f increases with

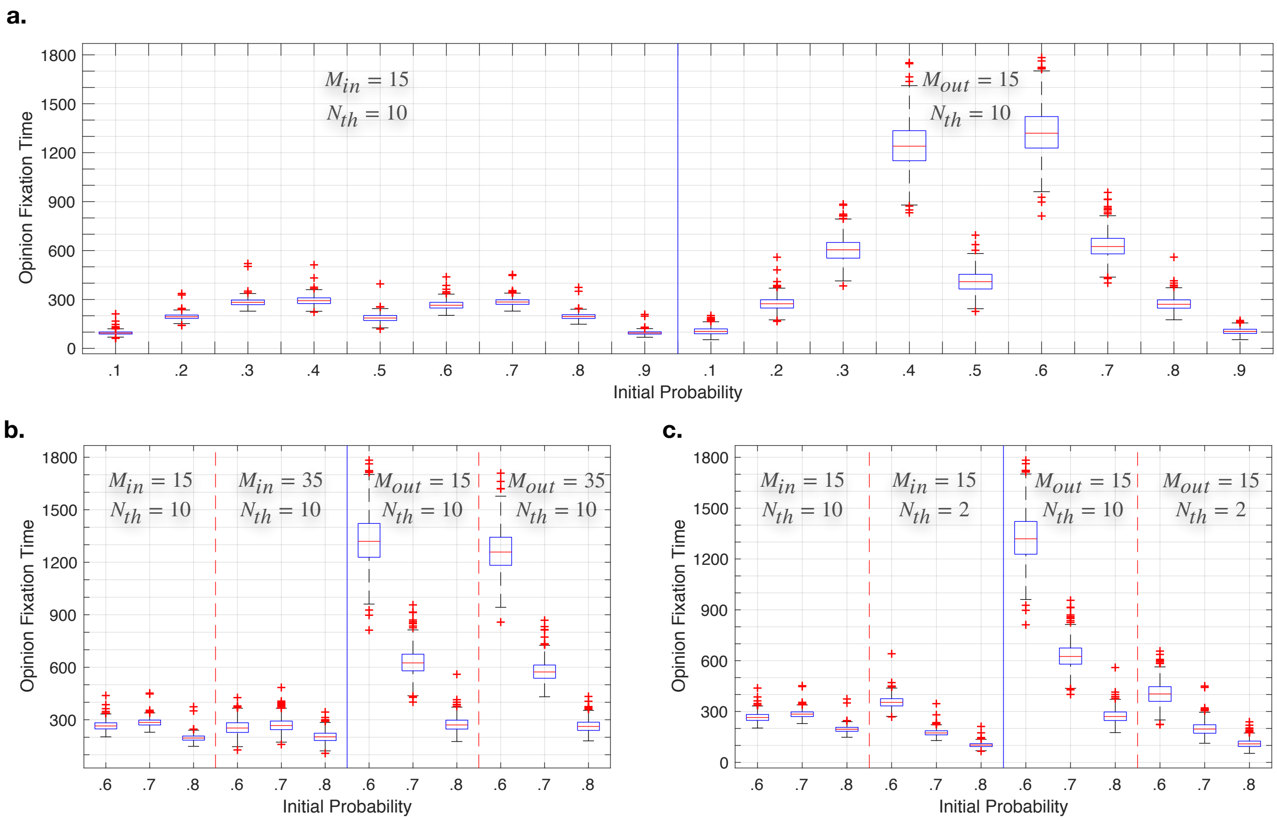


Figure 5. The comparison of fixation time of individuals as a function of the initial probability a. when thresholds are out-degree dependent (left) and in-degree dependent (right) b. with varying mean-degree (M_{in}/M_{out}) when thresholds are out-degree dependent (left) and in-degree dependent (right) c. with varying threshold heterogeneity (N_{th}) when thresholds are out-degree dependent (left) and in-degree dependent (right).

increasing N_{th} when thresholds are in-degree dependent. When thresholds are out-degree dependent, the effect of N_{th} on f_t is very minimal and the relation between N_{th} and t_f depends on p, thus, increasing N_{th} causes the people to fix their opinions later when $p \gtrsim 0.7$, while the effect is opposite when $0.6 \gtrsim p \gtrsim 0.7$.

Figure 4.a2 and 4.b2 show $\bar{s}$ as a function of p at various M_{in} and M_{out} values. Results show that the change in the mean-degree has no prominent effect on the average opinion at the steady state when thresholds are out-degree dependent; however, increasing the mean-degree seems to facilitate reaching a consensus when thresholds are in-degree dependent if $p \lesssim 0.7$. In the same situation, if $p \gtrsim 0.7$, $\bar{s}$ values are very close to each other. Since the standard deviations of the results are high, we can conclude that the mean degree does not affect average opinion at a steady state when thresholds are in-degree dependent or out-degree dependent. This is not surprising when we redefine the threshold model. The threshold model takes the ratio of the node's threshold to the average opinion of his neighbors and changes the node's opinion if the ratio is higher than 1. The ratio does not change with changing mean-degree when the initial, $\bar{s}$, is not affected from M_{in} and M_{out}. We would expect t_f to increase because when the number of links between nodes increased and caused more changes in ideas, but the results show that the change in mean-degree has no effect on t_f when thresholds are in-degree or out-degree dependent.

Increasing node size in the network decreases $\bar{s}$ significantly when thresholds are in-degree dependent, whereas it has very little effect when thresholds are out-degree dependent. The effect of seed size on $\bar{s}$, when $0.6 \leq p \leq 0.8$, shows that there is more diversity in the opinions when the seed size is higher, e.g. $\bar{s} = 0.9480$ when $N = 1000$, while $\bar{s} = 0.7155$ when $N = 5000$ at $p = 0.7$. Low standard deviation in the Monte Carlo simulations also demonstrates the consistency of simulation results in every trial.

4 Conclusion

In modern times, face-to-face communications have decreased in importance, and now, social media is a large part of humanity's social communication. Therefore, social network analysis has become very important to understand the dynamics of opinion formation, change, and propagation. One of the most common methods used to understand these dynamics is the threshold model. The first studies of social contagion used homogeneous binary threshold model due to its simplicity; however, people show more heterogeneous attitude to adopt a new opinion, which renders the use of heterogeneous thresholds a must. Even though more complex thresholds are now used in social contagion analysis, none of them validates their model with real data analysis. Our study is thus novel because the degree-dependency of thresholds is inferred by using real world Twitter data. Social data analysis shows that the threshold of a node does not only depend on its out-degree but also depend on its in-degree. Although the examples of out-degree dependent threshold models can be found in some studies, we also examined the results of opinion change simulations of the in-degree dependent threshold model. This study also investigates the effect of heterogeneity in thresholds reaching a consensus for the first time. Our simulations demonstrated that the system is more likely to reach a consensus when thresholds are in-degree dependent, rather than out-degree dependent. However, people change their opinion more and fix their opinion later in this case. Thresholds with higher heterogeneity are more likely to come to a consensus, but reaching a consensus takes more time than thresholds with lower heterogeneity, and this change is more significant when threshold are correlated with in-degree of nodes. Additionally, increasing seed size in the network makes the formation of a consensus more difficult regardless of the dependence of threshold to the in-degree or out-degree. We also note that as mean degree increases, diversity in opinions of individuals decrease when thresholds are in-degree dependent, while it has no effect when thresholds are out-degree dependent.

Acknowledgments

This work was partially supported by grant FA8650-18-C-7823 from the Defense Advanced Research Projects Agency (DARPA).

References

[1] Alexandre Bovet and Hernán A Makse. 2019. Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications* 10, 1 (2019), 7.

[2] Aron Culotta and Jennifer Cutler. 2016. Mining brand perceptions from twitter social networks. *Marketing science* 35, 3 (2016), 343–362.

[3] Clay Fink, Aurora C Schmidt, Vladimir Barash, John Kelly, Christopher Cameron, and Michael Macy. 2016. Investigating the observability of complex contagion in empirical social networks. In *Tenth International AAAI Conference on Web and Social Media.*

[4] James P Gleeson. 2013. Binary-state dynamics on complex networks: Pair approximation and beyond. *Physical Review X* 3, 2 (2013), 021004.

[5] Mark Granovetter. 1978. Threshold models of collective behavior. *American journal of sociology* 83, 6 (1978), 1420–1443.

[6] Peter Hedström. 1994. Contagious collectivities: On the spatial diffusion of Swedish trade unions, 1890-1940. *Amer. J. Sociology* 99, 5 (1994), 1157–1179.

[7] Raghuram Iyengar, Christophe Van den Bulte, and Thomas W Valente. 2011. Opinion leadership and social contagion in new product diffusion. *Marketing Science* 30, 2 (2011), 195–212.

[8] Panagiotis D Karampourniotis, Sameet Sreenivasan, Boleslaw K Szymanski, and Gyorgy Korniss. 2015. The impact of heterogeneous thresholds on social contagion with multiple initiators. *PloS one* 10, 11 (2015), e0143020.

[9] Eun Lee and Petter Holme. 2017. Social contagion with degree-dependent thresholds. *Physical Review E* 96, 1 (2017), 012315.

[10] Kevin Lewis and Jason Kaufman. 2018. The conversion of cultural tastes into social network ties. *American journal of sociology* 123, 6 (2018), 1684–1742.

[11] Taibo Li, Rasmus Wernersson, Rasmus B Hansen, Heiko Horn, Johnathan Mercer, Greg Slodkowicz, Christopher T Workman, Olga Rigina, Kristoffer Rapacki, Hans H Stærfeldt, et al. 2017. A scored human protein–protein interaction network to catalyze genomic interpretation. *Nature methods* 14, 1 (2017), 61.

[12] Quan-Hui Liu, Feng-Mao Lü, Qian Zhang, Ming Tang, and Tao Zhou. 2018. Impacts of opinion leaders on social contagions. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28, 5 (2018), 053103.

[13] Shraddha Mishra, Surya Prakash Singh, John Johansen, Yang Cheng, and Sami Farooq. 2019. Evaluating indicators for international manufacturing network under circular economy. *Management Decision* 57, 4 (2019), 811–839.

[14] Ece C Mutlu and Toktam A Oghaz. 2019. Review on Graph Feature Learning and Feature Extraction Techniques for Link Prediction. *arXiv preprint arXiv:1901.03425* (2019).

[15] Mark Newman. 2010. *Networks: an introduction.* Oxford university press.

[16] Zhongyuan Ruan, Gerardo Iniguez, Márton Karsai, and János Kertész. 2015. Kinetics of social contagion. *Physical review letters* 115, 21 (2015), 218702.

[17] Feng Shi, Misha Teplitskiy, Eamon Duede, and James A Evans. 2019. The wisdom of polarized crowds. *Nature Human Behaviour* (2019), 1.

[18] Pramesh Singh, Sameet Sreenivasan, Boleslaw K Szymanski, and Gyorgy Korniss. 2013. Threshold-limited spreading in social networks with multiple initiators. *Scientific reports* 3 (2013), 2330.

[19] Daniel A Sprague and Thomas House. 2017. Evidence for complex contagion models of social contagion from observational data. *PloS one* 12, 7 (2017), e0180802.

[20] Johan Ugander, Lars Backstrom, Cameron Marlow, and Jon Kleinberg. 2012. Structural diversity in social contagion. *Proceedings of the National Academy of Sciences* 109, 16 (2012), 5962–5966.

[21] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.

[22] Wei Wang, Ming Tang, Panpan Shu, and Zhen Wang. 2016. Dynamics of social contagions with heterogeneous adoption thresholds: crossover phenomena in phase transition. *New Journal of Physics* 18, 1 (2016), 013029.

[23] Duncan J Watts and Peter Sheridan Dodds. 2007. Influentials, networks, and public opinion formation. *Journal of consumer research* 34, 4 (2007), 441–458.

[24] Xuzhen Zhu, Wei Wang, Shimin Cai, and H Eugene Stanley. 2018. Dynamics of social contagions with local trend imitation. *Scientific reports* 8, 1 (2018), 7335.