Voice Cascade on Social Media

Huaxia Rui

April 26, 2018

Abstract

A voice cascade is a sudden increase in the number of publicly expressed opinions related to a particular event. Over the past decade, social media has given almost everyone unprecedented power to widely distribute their opinions, making voice cascade on social media a phenomenon that is not only scientifically interesting but also practically relevant, especially for potential stakeholders such as firms, governments, and celebrities. In this first paper on voice cascade, I propose Seasonally Adjusted Gamma Aftershocks, or SAGA in short, to model how such a phenomenon unfolds over time so that we can better understand its underlying mechanism. There are two key ingredients of the model. The first one describes the stochastic arrival of voices that would have occurred in the absence of any influence from previous voices. The second component models the temporal *shape of influence*, which specifies how each voice alters the arrival process of future voices over time. Together, they form the core of the SAGA model and offer a flexible framework to understand the evolution of a voice cascade. Using data from a recent voice cascade on Twitter, I conduct several tests to evaluate the validity of the SAGA model, the results of which show great promise of this new modeling framework.

**Keywords:** social media; voice cascade; point process; Twitter
1 Introduction

The rise of social media has empowered people in an unprecedented way. Anyone using a smart device can express an opinion publicly, with the potential to reach millions of others or to trigger an avalanche of related voices. This fundamental shift towards the democratization of voicing power has tremendous implications for business and governments, because, as Margaret Thatcher noted, words become actions. Indeed, in a well-connected society with free flow of information and opinions, any misconduct by any organization would face severe consequences. In this sense, the rise of social media is truly a revolutionary development in the advancement of civil societies. Crucial in its powerful mechanism to change society is the role of social media in enabling voice cascade, which I define as a sudden increase in the number of publicly expressed opinions related to a particular event. The #NeverAgain voice cascade\(^3\) on gun control in the aftermath of the shooting at Stoneman Douglas High School, and the #MeToo voice cascade\(^4\) on sexual assault and harassment, are two latest examples of voice cascade on social media, both with tremendous social impact.

As we move further into this new media age, I believe it is not only scientifically interesting but also practically important for researchers to gain a deeper understanding of how voice cascades emerge on social media, both theoretically and empirically. This is the objective of the current paper, which, to my knowledge, is the first attempt to study voice cascade on social media.

The defining feature of a voice cascade, which is also our key intuition of such a phenomenon, is the idea that past voices don’t just pass; they also influence the occurrence of future voices. As a result, a time series of voices without cascade and a time series of voices during a cascade should be structurally different. Indeed, this insight can be clearly seen from Figure I where the left panel shows how the number of tweets changes from 5 AM to noon on a typical day from a sample of Twitter users in the United States, and the right

\(^3\)https://en.wikipedia.org/wiki/Never_Again_MSD
\(^4\)https://en.wikipedia.org/wiki/Me_Too_movement
panel shows how the number of tweets belonging to a voice cascade changes from 5 AM to noon. The nonlinearity of the time series of the right panel reveals the role of influence, which has long been studied in the innovation diffusion literature. In his seminal work, Bass (1969) conceptualized and modeled two types of innovation adopters: those who adopt the innovation without influence from others (i.e., innovators) and those whose adoption is influenced by previous adopters (i.e., imitators). Bass’s famous model naturally gives rise to the nonlinear growth pattern similar to the one on the right panel of Figure 1. However, Bass’s model is an aggregate one, both conceptually and mathematically. To dive deeper into the mechanisms that drive innovation diffusion and voice cascade, we need to go one step further to conceptualize and model the two types of participants at individual level. Mathematically, we need to capture the raw data as a point process instead of a time series.

Figure 1: Comparing the structure of two time series: number of tweets during each 5-min interval from 5 am to noon by a sample of Twitter users in the United States on 2018/01/10 (left); number of tweets during each 5-min interval from 5 am to noon that belong to a voice cascade in the United States on 2017/04/10.

More specifically, in the current context of voice cascade, we need to first model the stochastic arrival of voices that would occur in the absence of any influence from previous voices, henceforth referred to as *exogeneous voices* or *immigrants*. Conceptually, these are akin to innovators in the Bass model. Then, we need to model the arrival process of voices
that occur because of influence from previous voices, henceforth referred to as endogeneous voices, descendants, aftershocks, or echoes. Conceptually, these are similar to imitators in the Bass model.

The nature of time-varying arrival rates of immigrants is the seasonality of user activities during different times of the day. Hence, I collected, for a two-week period, all the tweets from Twitter users located in the U.S. who share their location information in their tweets. Using this data set, I fit a piecewise linear model as the arrival rate function for the exogeneous voices at different times of the day.

To understand how past voices influence the occurrence of future voices, I develop a simple probabilistic model to illustrate that, under some technical assumptions, the temporal distribution of endogeneous voices, or the shape of influence, is a gamma distribution with a left truncation and a right time shift. This shape of influence is intellectually intriguing because it seems to confirm our intuition that the influence of a voice or an idea often rises gradually over time to reach its peak before slowly fading away. Interestingly, such a shape of influence may also give rise to a ripple structure over time which is common in physics and often used metaphorically to describe influence over time.

Combining the seasonality model for exogeneous voices and the gamma aftershocks model for endogeneous voices, I propose Seasonally Adjusted Gamma Aftershocks, or SAGA in short, to model voice cascade on social media. To evaluate the validity of this model, I conduct three empirical tests using data from a recent voice cascade which was triggered by a senior passenger being dragged off a plane in 2017. Estimation results from the baseline SAGA model show that the SAGA model fits the data much better than several seasonality adjusted non-homogeneous Poisson models. A similar test of an extension of the baseline

---

5 These different names are informative when we draw connection to different literature.

6 Of course, this requires the shape parameter of the gamma distribution be larger than 1, which is true in all empirical estimations of the model in this paper.

7 For example, Mother Teresa once said: “I alone cannot change the world, but I can cast a stone across the waters to create many ripples.”
SAGA model reveals that certain content features can elevate the magnitude of influence of a voice. Another validation test using more detailed social network information serves as a unique “out-of-sample” test because the quantity predicted by the SAGA model is constructed from data not used during the calibration of the SAGA model. The results of this test offer strong support to the SAGA model, and suggest that it is possible to infer microscopic information about a voice cascade from macroscopic data. This is particularly encouraging from the scientific perspective, because even in the age of big data, we can rarely access fine-grained data that can directly reveal the mechanism of voice cascade.

Since its emergence over a decade ago, social media has often been perceived as a blessing for brands and has been intensively studied by researchers for its marketing potentials. But as many recent events have shown, social media is a double-edged sword that can easily amplify the damage to the reputation of a brand. From the perspective of a brand, a negative voice cascade on social media is analogous to an epidemic. In this regard, there is an interesting parallel between the study of voice cascade and the field of mathematical epidemiology, the goal of which is to understand the spread of a communicable disease through mathematical modeling so that decision makers can better control the disease spread. In the same vein, with a deeper understanding of the mechanism of voice cascade, stakeholders can make better decisions to manage the situation and might even design and evaluate intervention strategies. By proposing the first mathematical model of voice cascade on social media, this paper lays the theoretical foundation for understanding voice cascade and paves the way for future studies on the management of voice cascade.

The rest of the paper is organized as follows: In Section 2, I connect voice cascade and the SAGA model with several streams of literature from different disciplines. I then develop the SAGA model in Section 3. After describing the data in Section 4, I conduct three sets of empirical tests in Section 5 to provide supporting evidences for the SAGA model. Finally, in Section 6, I summarize the paper and its contributions to the literature, before discussing its limitations and suggesting some future directions to advance this new stream of research.
2 Literature Review

The current paper relates to several streams of earlier works in multiple disciplines, yet is the first of its kind in many aspects. Cascade phenomenon on social media has been studied in the context of content diffusion in online social network, but there are important differences between the diffusion of content and the cascade of voices. The conceptual similarity between voice cascade and innovation diffusion relates the current paper to the famous work of Bass (1969), but the granularity of social media data allows us to model the branching structure at the individual rather than at the aggregate level, thereby giving us deeper understanding of the underlying mechanism. There is also an interesting parallel between this nascent research area of voice cascade on social media and the century-old field of mathematical epidemiology. The research context and potential applications of the current paper naturally connect it to the vast literature on social media marketing. However, the idea and the mathematical formulation of voice cascade has hitherto been overlooked by this stream of works. Finally, this paper is also distantly connected to the nascent literature studying the role of social influence on the generation of online product review which can be viewed as a type of voice on social media.

To broadly position the current paper among these different streams of works, I organize the rest of the literature review into five parts, each corresponding to one stream.

2.1 Content Diffusion on Social Media

The literature on social media content diffusion typically concerns how a piece of content disseminates in a given social network and has been a major area of research in the past decade (Sundararajan et al., 2013). Conceptually, sharing a piece of content is different from voicing an opinion. Whereas the former is all about the diffusion of a particular piece of content, the latter is centered around some triggering event that often leads to a public display of voices. There is certainly an overlap between the two, especially on social media.
For example, retweet cascade can be viewed both as content diffusion and as a very special type of voice cascade where all voices are exactly the same as the first one. Clearly, on social media, both the cost and the reward of voicing is very different from those of sharing a piece of content posted by another individual. Hence, both the temporal pattern and the type of individuals who participate a voice cascade are likely different from those of a content diffusion.

Most of the works in this stream either try to understand how social network characteristics affect content diffusion or to predict the final popularity of a given content. A an example of the first type of work is Shi et al. (2014), where they studied the relation between a Twitter user’s retweeting decision and his or her social network information. Drawing upon the strength of weak ties theory and the social capital theory, they hypothesized and verified that users with weak ties to the author of a tweet are more likely to retweet than users with strong ties to the author. More recently, Peng et al. (2018) proposed a novel hazard model to flexibly capture the impact of three different measures of network overlap (i.e., common followees, common followers, and common mutual followers) on content sharing. Model validation using data from Twitter and Digg shows great promise and suggest the proposed dyadic hazard model applies to the sharing of content generated by both firms and users.

Works of the second type are more abundant where the focus is often to predict the final popularity of a given content. Naturally, most works (Cheng et al., 2014) explored features related to the original poster, the resharers, and the content of the original post. These works typically differentiate by their particular methods and their choice of features. For example, Suh et al. (2010) used a generalized linear model to identify key factors that affect retweet rate. Zaman et al. (2014) combined directed graph and Bayesian approach to predict the total number of retweets a tweet will receive using early retweet times, the retweets of other tweets and summaries of the follower graphs. Zhang et al. (2015) proposed a non-parametric statistical model to combine structural, textual, and temporal information for the prediction of retweet behavior. Zhao et al. (2015) proposed a self-exciting point process
model to predict the final number of retweets of a given tweet. While most researchers seem to focus on the understanding of the popularity of a given tweet, some studies also investigate other types of content. For example, using data from Youtube, Susarla et al. (2012) showed that the diffusion rate of videos relies heavily on the temporal network structure of social interactions, such as in-degree centrality, friend network, etc. Rizoiu et al. (2017) studied the relation between external promotions of a Youtube video and its final popularity.

Besides the key difference that the current paper studies the phenomenon of voice cascade rather than content diffusion, the research objective here is also different from those in the literature of content diffusion. Instead of predicting the final number of voices related to the triggering event, which is an important task and a potential future application of the current study, the primary goal of the current paper is to reveal the dynamics of a voice cascade, both theoretically and empirically. As the first paper to study voice cascade, a deeper understanding of how it works could be the foundation of various future studies and potential applications. From the perspective of a stakeholder of a voice cascade, it is certainly more valuable to know how the phenomenon unfolds than to be told of a total number of voices based on some model prediction.

2.2 Innovation Diffusion

The literature on innovation diffusion concerns with how innovations diffuse over time. Of particular interest is the diffusion of new products, probably because of its direct marketing applications.

The most well-known model of innovation diffusion is that of Bass (1969). Conceptualizing a population as consisting of both innovators (those with a constant propensity, \( p \), to adopt a new product or technology) and imitators (those whose propensity to adopt is influenced by previous adopters), Bass assumes the adoption density function \( f(t) \) satisfies the differential equation \( f(t) = (p + qF(t))(1 - F(t)) \) where \( F(t) \) is the cumulative distribution function of adoption and \( q \) measures the degree to which adopters can influence imitators to
adopt. Starting from this differential equation, Bass derived the formula of the cumulative density function $F(t)$ which closely resembles the often observed S-curve of new product adoption. The main advantage of the Bass model is its simple but elegant mathematical formulation that explains the empirical data well.

Following Bass’s seminal work, various extensions have been proposed to add flexibility to the model. For example, Bewley and Fiebig (1988) extended the basic Bass model using four parameters in order to accommodate the flexible locations of inflection points. These diffusion models have also been generalized to incorporate more variables designed for particular applications. For instance, the pioneering work of Robinson and Lakhani (1975) estimated the product demand using a modified Bass model by including marketing variables in the parameterization. The main models of innovation diffusion and their modifications have been applied to forecast market adoptions of consumer durables (Grewal et al., 2004), telecommunications (Gruber and Verboven, 2004), generations of technologies (Norton and Bass, 1987), etc.

There are several important differences between this family of innovation diffusion models and the voice cascade model proposed in the current paper. First, unlike a Bass-type model which is a deterministic and aggregate-level model offering a macroscopic view, a voice cascade model is a stochastic and individual-level model offering a microscopic view. Consequently, compared with a Bass-type model, a voice cascade model requires and takes advantage of much fine-grained data. For example, the Bass model is estimated only using a time series consisting of the number of new adopters during each time period (e.g., a quarter or a year). In contrast, a voice cascade model is estimated using the path of a point process consisting of the exact time of each point where a point is a voice in the current context and an adoption in the context of innovation diffusion. Hence, the temporal information

---

8The macroscopic nature of a Bass-type model is independent of how time unit is defined. As an imperfect analogy, consider the standard Brownian motion, which provides us a macroscopic framework to understand particle movement. Even though a Brownian motion can be conceptually motivated by random walks, there is no actual “jump” in the path of a Brownian motion. In fact, the path of a Brownian motion is almost surely continuous.
that feeds into the Bass model is a strict subset of the temporal information that feeds into a voice cascade model. The less demanding data requirement is an advantage when data is scarce and computation is slow. However, in the current age of big data and abundant computing resources, models that exploit more fine-grained data can potentially be more powerful and informative.

Second, models of innovation diffusion implicitly assume that individuals’ interactions with each other occur randomly, and heterogeneity is only allowed at an aggregate level. In reality, people are only influenced by their social network neighbors, which differ from individual to individual. Unlike innovation diffusion models, the conceptual and mathematical framework for a voice cascade model is conceived with each individual as its building block. Therefore, the flexibility to incorporate heterogeneity at the individual level is an innate advantage of a voice cascade model. For example, a measure of the direct size of influence of each point, an important piece of information not present in the Bass model, enters directly into a voice cascade model. It is also natural to include additional features associated with each point in such models, as I will later illustrate in an extension of the SAGA model in Section 5.3.

Finally, there is a conceptual difference between the adoption of an innovation and the decision to voice. For example, most people adopt an innovation at most once while they can certainly voice multiple times on social media. Therefore, the idea of saturation, which is a key ingredient in innovation diffusion models, might play a less important role in a voice cascade. On the other hand, the role of seasonality, which is critical in many voice cascades, seems not significant in this stream of literature.

2.3 Mathematical Epidemiology

A voice cascade on social media often follows after a brand makes some mistake and ultimately culminates in a public relation (PR) disaster. For the troubled brand, the cascade of voices on social media is analogous to the spread of a communicable disease. From this
perspective, there is an interesting parallel between the research area of voice cascade and
the field of mathematical epidemiology. An epidemic is a sudden outbreak of a disease that
infects a substantial portion of the population in a region before it disappears. The objec-
tive of mathematical epidemiology is to capture the dynamics of disease transmission using
mathematical models so that policy makers can be more informed and make better decisions
to combat the disease. Similarly, the practical goal of studying voice cascade is to better
understand the stochastic process governing the cascade phenomenon so that managers can
better evaluate and manage the situation. Because of the conceptual similarity between
mathematical epidemiology and the study of voice cascade, I briefly review some notable
mathematical models of disease transmission.

Most epidemic models start by dividing the population into groups of homogeneous
hosts, known as compartments. In the classical SIR model studied by Kermack and McK-
endrick (1932), there are three compartments: Susceptible—those at risk of being infected;
Infectious—infected hosts who can transmit the infection to susceptible hosts; and Removed—
those immune to the infection. The SIR model specifies the differential equations for the
evolution of the number of hosts in each compartment. There are many variations of the
basic SIR model. For example, the SIS model assumes that hosts have no immunity against
re-infection and return to the susceptible class after recovery; the SEIS model assumes an
exposed period between being infected and becoming infective. To account for population
heterogeneity, Arino et al. (2007) first divide the population into several subpopulations, then
apply a basic compartment model for each subpopulation and specify how infection spreads
across groups. This class of epidemic models is deterministic, and like those innovation diffu-
sion models, uses time series data for model estimation. A voice cascade model differs from
these models in similar way as how it differs from the innovation diffusion models.

The second class of epidemic models uses stochastic process to describe disease trans-
mission. Such a stochastic approach is suitable when the population size is small, especially
during the early stage of an outbreak. A simple example of stochastic epidemic models is
the Reed-Frost model ([Daley and Gani, 2001]) where at each period, each infected individual independently infects each susceptible individual with some probability $p$, and an infected individual at $t$ is removed at $t + 1$. Therefore, the independence assumptions give rise to a binomial distribution of the number of infected at $t + 1$ which allows further analysis of the progression of the epidemic. Another type of stochastic epidemic models is rooted in the mathematical theory of branching process. In the simple example of a Galton-Watson process ([Brauer, 2017]), at each period, each infected individual has a probability $p_k$ of infecting $k$ susceptibles in the next period before being removed, independent of any other infections or any history. Such a model can be analyzed to offer insights such as whether the infection will die out or persist. By explicitly modeling how each individual voice affects voices occurrence in future periods, the voice cascade model of the current paper (i.e., the SAGA model) is in the same spirit of the stochastic epidemic models described above. However, there are also important differences. First, the modeling approaches are fundamentally different. In the SAGA model, the focus is on the probability distribution of descendants in all future periods, rather than the number of descendants at the next period as in most stochastic epidemic models. The total number of descendants in the SAGA model is partly determined by the data (e.g., the number of followers) and partly inferred from the model (e.g., through the quality score). In fact, the partial availability of social network information is the key motivation of the particular modeling approach adopted in the SAGA model. Second, because of the different modeling approach, I use a Hawkes process to specify the SAGA model, rather using a branching process. However, it should be pointed out that there is a deeper connection between Hawkes processes and branching processes.

2.4 Social Media Marketing

The IS and marketing literature on social media marketing is vast, almost all of which focus on the offensive marketing applications of social media. Most of these papers either explore the link between firms’ social media activities and consumers’ social media engagements
(e.g., Lee et al. (2018a)), or studying the link between consumers’ engagements and their brand choices (e.g., Goh et al. (2013); Rishika et al. (2013)). More recently, researchers have also started investigating the defensive marketing applications of social media (e.g., He et al. (2018); Gunarathne et al. (2018)). Because most voice cascades on social media are about some negative PR events, direct applications of the present paper are likely along this new line of research regarding defensive marketing on social media.

### 2.5 Online Product Review and Social Influence

Online product reviews can be considered as a type of voice on social media. In this regard, there is also a connection between the literature on the generation of online product review and voice cascade. In particular, some recent papers have examined the role of social influence on product review generation. For example, Wang et al. (2018) investigated online friends’ social influence in online book ratings. They found that with social-networking functions, online rating contributors are socially nudged when giving their ratings. More recently, Lee et al. (2018b) studied how social imitation and learning affect user rating generation in the context of movie ratings. A major difference between this nascent stream of literature and the current study is that the dynamics of event-driven voice cascade is fundamentally different from product-driven review cascade, both in terms of intensity and mechanism, as well as potential applications. However, the important role of social influence in the generation of voices is shared by the current study and this stream of literature.

### 3 Model

The mathematical foundation of the SAGA model is a generalization of the Poisson process called Hawkes process (Hawkes, 1971), where the independent arrival assumption of Poisson process is relaxed. Generally, a Hawkes process is a simple point process $\{N_t\}_{t \geq 0}$
characterized by a conditional intensity function

\[ \lambda(t | \mathcal{F}_t) \equiv \lim_{h \to 0^+} \mathbb{E} \left[ \frac{N(t + h) - N(t)}{h} \bigg| \mathcal{F}_t \right] = \lim_{h \to 0^+} \frac{1}{h} \mathbb{P}[N(t + h) - N(t) > 0 | \mathcal{F}_t], \]

where \( \mathcal{F}_t \) is the natural filtration associated with \( N(t) \). A Poisson process is a degenerated Hawkes process with \( \mathcal{F}_t = \mathcal{F}_0 \). Because Hawkes processes can generate clustering patterns resembling certain real-world point processes, it has been used in disciplines such as seismology (Ogata, 1998) and more recently in IS and marketing (Xu et al., 2014; Zadeh and Sharda, 2013) as well.

The conditional intensity function can be decomposed into an exogeneous component that does not depend on the arrival history (i.e., \( \mathcal{F}_0 \)-measurable), and an endogeneous component that depends on information associated with previous arrivals (i.e., \( \mathcal{F}_t \)-measurable but not \( \mathcal{F}_0 \)-measurable). Hence, I will model these two parts separately. The first component, which will be discussed first, is motivated by the seasonality of the voicing pattern on social media. The second component, which I call gamma aftershocks, is motivated by a probabilistic voicing model. Together, these two components form the core of the SAGA model.

### 3.1 Seasonality

One of the key challenges in understanding a voice cascade process is to separate it from seasonality, because an increase in the arrival rate of voices might indicate a true cascade event but also may simply reflect people’s different tendency to voice during different times of the day. However, a point process due to a cascade might be structurally different from a point process purely driven by seasonality, which is the key to identify a voice cascade using point process data. Therefore, it is crucial that we model the seasonality during the sample period.

Placed in the classical framework of causal inference, we need to find a counterfactual point process in the absence of any voice cascade event. To construct this “control group,” I
first collected, from January 10, 2018 to January 23, 2018, all the tweets whose longitude and latitude in the metadata are located within the continental United States, henceforth referred to as geo tweets. The rationale to choose tweets originating from the U.S. is because the event that triggered the voice cascade studied in the empirical analysis is largely a U.S. event. Only a small percentage of Twitter users reveal their geo locations in their tweets. Hence, these users and their tweets are unlikely to be representative among all Twitter users located in the United States. However, their seasonality patterns are likely to be representative. In total, there are 13,839,940 geo tweets during the two-week period. Figure 2 shows the time series of these tweets, averaged over the 14 days, with each point corresponding to a 5-min interval from midnight to midnight (left panel) and from 5 AM to 1 PM (right panel).

![Voice seasonality from midnight to midnight (left) and from 5 AM to 1 PM (right), Eastern Standard Time, January 10, 2018 to January 23, 2018.](image)

Although not plotted in the paper, the seasonality patterns of different days during the two-week period are very similar. Because the sample period of the empirical analysis in Section 5 is on Monday, April 10, 2017, to be conservative, I will only use Monday, January 15 and Monday, January 22 to construct the seasonality model.

From Figure 2, it’s clear that the seasonality pattern between 5 AM and 1 PM is the relatively simple during the day. The seasonality during this 8-hour period also seems to be

---

9The seasonality from midnight to 5 AM is also simple. However, in the main dataset, there are very few observations related to the event that triggered the voice cascade.
more stable than other periods of the day across different days. Thus, to reduce statistical errors and for better identification of the proposed voice cascade model, I will focus on the hours between 5 AM and 1 PM in the empirical analysis of Section 5.

To model the seasonality from this 8-hour period, I use piecewise linear functions. Mathematically, for \( k \in \mathbb{Z}^+ \), let the arrival process of exogeneous voices on social media be a non-homogeneous Poisson process where the arrival rate is a piecewise linear function of time \( t \) with \( k \) segments. Note that when \( k = 1 \), the arrival rate is simply a linear function of time.

Assuming that the arrival rate is constant during each of the 5-minute interval between 5 AM and 1 PM, I estimate the arrival rate for each 5-minute interval using the number of geo tweets on Monday, January 15 and Monday, January 22. Then, I approximated the evolution of arrival rate using a piecewise linear functions of \( k \) segments where \( k = 1, 2, 3, 4 \). Figure 3 illustrates two fitted piecewise linear functions with \( k = 3 \) (left panel) and with \( k = 4 \) (right panel).

![Figure 3: Modeling seasonality using piecewise linear functions.](image)

To distinguish different choices of the seasonality component in a point process model, I append the number \( k \) to the name of the model to indicate the number of segments used...
in the piecewise linear function. For example, the model labeled as Poisson4 is a non-homogeneous Poisson process model that uses a 4-segment piecewise linear approximation of the seasonality pattern, and the model labeled as SAGA3 is a SAGA model that uses a 3-segment piecewise linear approximation of the seasonality pattern. For convenience, \( k = 0 \) is used to indicate that the model does not have a seasonality component, that is, with a constant arrival rate of exogeneous voices.

### 3.2 Gamma Aftershocks

To understand the shape of influence, we need to understand the statistical property of time-to-echo, \( T \), which is the elapsed time between a voice—immigrant or descendant—and its descendant which is also referred to as an echo in this section. Note that although we can observe the time stamp of each voice in a voice cascade, we cannot observe whether the voice is an immigrant or a descendant, not to mention the realization of \( T \) associated with the voice if it is a descendant.

Intuitively, once a voice appears, echoes by different users among its audience should be largely unrelated to each other. Hence, for simplicity, we think of the audience as consisting of multiple cohorts where there is at most one echo in each cohort and the time-to-echo in different cohorts are independent of each other. Our next task is to focus on a representative cohort and study the distribution of \( T \) in this cohort. Unless otherwise stated, all voices in this subsection belong to this representative cohort.

To derive the probability distribution function of \( T \), we make two modeling assumptions. First, we assume that upon the occurrence of a voice, the process of the voice being actually consumed by social media users (e.g., followers) is a Poisson process with a time shift to the right. In other words, the Poisson process only starts after the elapse of a strictly positive amount of time. Intuitively, such a positive shift reflects the minimal amount of time for a social media user to digest the information in the voice. The Poisson process assumption is reasonable because each voice consumption is by a different user and should be stochastically
independent to other voice consumptions.

Second, suppose the user echoing the voice is the $K$-th user who consumed the original voice. Here, $K$ is a discrete random variable with a finite support $\{1, 2, \cdots, N + 1\}$. Intuitively, $K$ should follow a unimodal probability distribution because the probability of $K = 1$ and $K = N + 1$ should both be small compared with intermediary values of $K$. Two natural candidates are the binomial distribution and the truncated Poisson distribution. Because a Binomial distribution converges to a Poisson distribution as the number of trials becomes very large and the probability of success becomes very small, a scenario that fits well our context, the truncated Poisson distribution seems a natural choice for the distribution of $K$. However, for technical reason, instead of directly using a truncated Poisson distribution, I will use its mirror. Formally, $K$, the mirror of a truncated Poisson random variable with rate parameter $\nu > 0$, has the following probability density function:

$$P(K = k) = (1 - \epsilon)^{-1} \frac{\nu^{N-k+1} e^{-\nu}}{(N - k + 1)!}, \forall k \in \{1, 2, \cdots, N + 1\},$$

where $\epsilon = \sum_{k=N+1}^{\infty} \frac{\nu^k e^{-\nu}}{k!}$.

Under these assumptions, we have the following important theoretical result which is the motivation of the SAGA model.

**Theorem 1.** Time-to-echo follows an Erlang distribution with a left truncation and a right time shift.

The proof is by direct construction of the density function of $T$. Let $\Gamma(k, x) = \int_x^\infty t^{k-1} e^{-t} dt$ be the unnormalized upper incomplete gamma function.

**Proof.** Let $\lambda$ and $\tau_1$ be the rate and shift of the right-shifted Poisson process of voice consumption, and let $\tau_2$ be the time it takes for the $K$-th user to echo. Hence, $T = \tau_2$ is the time when the $K$-th user completed consuming the voice. Let $T' = T - \tau_1 - \tau_2$. For any
given \( k \in \{1, 2, \cdots, N + 1\} \) and \( t \geq 0 \), by the second assumption, we have

\[
F_k(T') \leq t) = 1 - F_k(T' > t) = 1 - e^{-\lambda t} \sum_{n=1}^{k-1} \frac{(\lambda t)^n e^{-\lambda t}}{n!},
\]

so,

\[
\frac{dF_k}{dt} = \lambda e^{-\lambda t} - \sum_{n=1}^{k-1} \frac{1}{n!} \left( (\lambda t)^n (-\lambda e^{-\lambda t}) + e^{-\lambda t} n \lambda (\lambda t)^{n-1} \right)
\]

\[
= \lambda e^{-\lambda t} + \lambda e^{-\lambda t} \sum_{n=1}^{k-1} \left[ \frac{(\lambda t)^n}{n!} - \frac{(\lambda t)^{n-1}}{(n-1)!} \right]
\]

\[
= \lambda^k e^{-\lambda t} \frac{t^{k-1}}{(k-1)!}.
\]

Denote \( \theta = \lambda^{-1} \), the density function of \( T' \) can be written as

\[
f(t) = \sum_{k=1}^{N+1} P(K = k) \frac{t^{k-1}}{(k-1)!} \theta^k e^{-\frac{t}{\theta}} = (1 - \epsilon)^{-1} \sum_{k=1}^{N+1} \frac{\nu^{N-k+1} e^{-\nu}}{(N-k+1)!} \cdot \frac{t^{k-1}}{(k-1)! \theta^k} e^{-\frac{t}{\theta}}.
\]

By change of index and the following binomial expansion,

\[
\left( \nu + \frac{t}{\theta} \right)^N = \sum_{k=0}^{N} \frac{N!}{(N-k)!k!} \left( \frac{t}{\theta} \right)^k \nu^{N-k},
\]

we have

\[
f(t) = (1 - \epsilon)^{-1} e^{-(\nu + \frac{t}{\theta})} \sum_{k=1}^{N+1} \frac{\nu^{N-k+1}}{(N-k+1)!} \cdot \frac{t^{k-1}}{(k-1)! \theta^k}
\]

\[
= (1 - \epsilon)^{-1} e^{-(\nu + \frac{t}{\theta})} \theta^{-1} \sum_{k=0}^{N} \frac{\nu^{N-k}}{(N-k)!} \cdot \frac{1}{\theta^k} \left( \frac{t}{\theta} \right)^k
\]

\[
= (1 - \epsilon)^{-1} e^{-(\nu + \frac{t}{\theta})} \theta^{-1} \frac{(\nu + \frac{t}{\theta})^N}{N!},
\]

which is the density function of a (left) truncated Erlang distribution with shape parameter is \( N + 1 \), scale parameter \( \theta \), and truncation point \( \nu \theta \). To verify that \( f(t) \) is indeed a density
function, note

\[
\int_0^\infty f(t)dt = (1 - \epsilon)^{-1} \int_\nu^\infty x^N e^{-x} dx = \frac{\Gamma(N + 1, \nu)}{N!(1 - \epsilon)} = 1
\]

where we used the identity

\[
\Gamma(N + 1, \nu) = N! \sum_{k=0}^{N} \frac{\nu^k e^{-\nu}}{k!}.
\]

Finally, because \( T = T' + \tau \) where \( \tau \equiv \tau_1 + \tau_2 \), the distribution of \( T \) is simply the distribution of \( T' \) with a time shift \( \tau \) to the right, and the conclusion follows.

### 3.3 SAGA Model

Inspired by Theorem 1, I propose the following mathematical model for the voice cascade phenomenon.

**SAGA:**

\[
\begin{align*}
E_i &= \Psi(M_i, q_i) \\
\lambda(t) &= \mu(t) + \frac{1}{\Gamma(k, \nu)^\theta} \sum_{i: t_i + \tau < t} E_i \cdot \left( \frac{t - (t_i + \tau)}{\theta} \right)^{k-1} e^{-\frac{t - (t_i + \tau)}{\theta} \nu \theta}
\end{align*}
\]

The SAGA model consists of two components. The first component specifies the energy level of a voice, \( E_i \), which measures the total number of echoes it ultimately triggers. The energy level of a voice is determined by its mass — \( M_i \) and its quality score — \( q_i \), through the energy function \( \Psi \). The mass \( M_i \) captures the scope of influence of a voice. On Twitter, a natural measure of mass is the number of followers of the voicer. The quality score \( q_i \) is introduced to reflect the fact that not all receivers of a voice will echo. In the simplest case of a linear energy function, \( \Psi(m, q) = mq \), we can interpret the quality score as the propensity to echo by the audience. However, one might also introduce concavity or convexity with respect to \( m \) if one believes that there is a decreasing or increasing marginal effectiveness (with respect to the mass, \( m \)) of a voice in triggering aftershocks.

In the baseline SAGA model, the same quality score is shared by all voices. In a het-
erogeneous extension of the SAGA model, different voices can have different quality scores. For example, we can use a quality function, $\Phi(X_i \beta)$, to map a voice with a feature vector $X_i$ to its quality score in a traditional regression framework. If one wishes to restrict the value of quality score to be between 0 and 1, we can let the quality function be the cumulative distribution function of some probability distribution. For example, we can let $\Phi$ be the logistic function or the CDF of the standard normal function. Alternatively, we can more generally allow the value of quality score to be any non-negative real number. There are at least two advantages of this relaxed configuration. First, while the followers of a Twitter user automatically receive a copy of his/her tweets, one does not have to be a follower in order to consume these tweets as long as such tweets are not protected. For example, some tweets may reach more people after being picked up by bloggers, news websites, or users on other social media platforms. So there is a conceptual advantage of not restricting the quality score to be less than one and let the data reveal its most likely value. Second, computationally, it might be easier to estimate the variance of $\beta$ if the optimal parameters lie in a relatively flat region of the parameter space, which could occur in the current context due to the challenge of modeling content quality.

The second component of SAGA specifies the temporal pattern of echoes triggered by a voice given its energy level, $E_i$. This part follows directly from the prescription of Theorem 1 if we notice that an Erlang distribution is a special type of Gamma distribution where the shape parameter takes an integer value. Because optimizing over a continuous space is easier than optimizing over a discrete space, I expand the family of distributions in the SAGA model to all truncated Gamma distributions with a time shift for ease of empirical estimation.

It should be noted that the current approach of modeling voice heterogeneity through a quality score function is a convenient but not necessarily the best approach. For example, it’s possible to introduce voice heterogeneity through content-specific parameters in the Gamma distribution. However, for simplicity and ease of estimation, I will focus on the quality score
approach in this paper.

To estimate the SAGA model, we need to efficiently compute the compensator of the point process, which is defined as the integrated intensity function  \( \Lambda(0, T) \equiv \int_0^T \lambda(s)ds \) for any  \( T > 0 \). Luckily,  \( \Lambda(0, T) \) can be easily computed as the following:

\[
\Lambda(0, T) = \int_0^T \mu(t)dt + \frac{1}{\Gamma(k, \nu)\theta^k} \sum_{i: t_i + \tau < T} E_i \cdot \int_{t_i + \tau}^T (t - (t_i + \tau) + \nu \theta)^{k-1} e^{-\frac{t - (t_i + \tau) + \nu \theta}{\theta}} dt
\]

\[
= \int_0^T \mu(t)dt + \frac{1}{\Gamma(k, \nu)} \sum_{i: t_i + \tau < T} E_i \cdot \left( \Gamma(k, \nu) - \Gamma \left( k, \frac{T - (t_i + \tau) + \nu \theta}{\theta} \right) \right)
\]

\[
= \int_0^T \mu(t)dt + \sum_{i: t_i + \tau < T} E_i \cdot \left( 1 - \frac{\Gamma \left( k, \frac{T - (t_i + \tau) + \nu \theta}{\theta} \right)}{\Gamma(k, \nu)} \right).
\]

Given the conditional intensity and the compensator of a point process, the log-likelihood for the observed data  \( \{t_i | t_i \in [0, T]\} \) is given by the following expression:

\[
\ln L(N(t)_{t \in [0, T]}) = \int_0^T \ln \lambda(s)dN(s) - \int_0^T \lambda(s)ds = \sum_{i: t_i \in [0, T]} \ln \lambda(t_i) - \Lambda(0, T).
\]

4 Data

I estimate the SAGA model using data from a recent voice cascade event on social media. Below, I first describe the event and then discuss the data. To avoid confusion, all times in this paper have been converted to Eastern Standard Time.

On Sunday, April 9, 2017, at Chicago’s O’Hare International Airport, after passengers were seated in the United Express flight 3411 aircraft, the gate agent announced that they need four passengers to get off the plane to accommodate four staff members. Because there is no volunteer, a manager informed the flight that four passengers were chosen by a computer to leave the plane and the 69-year-old Dr. David Dao was one of them. When Dr. Dao refused to leave the plane, the airline called security personnel who injured and forcibly dragged Dr. Dao along the aircraft aisle out of the plane. At 7:22 PM, a tweet
with the content “@united just had the police forcefully remove an older passenger, medical Doctor, who had to work tomorrow.” appeared on Twitter. In the next few hours, more people expressed their dismay on social media. The next morning, a voice cascade occurred, as is shown in the bottom panels of Figure 4.

I obtained all tweets directed towards the official Twitter account of United Airlines (i.e., @united) from April 9 to April 10 that contain at least one of the following terms: drag, knock, beat, overbook, doctor, physician, forced off, violent, banish, removed, forcibly, assault, appall, terrify, horrific, 3411. These terms are selected to best capture tweets related to the incident. In total, I obtained 19,084 relevant ones. To check whether these tweets are truly related to the incident, I randomly selected 100 of these and manually evaluated their relevance. Out of these 100 tweets, 98 are truly relevant. Hence, I conclude that the sample is a good representation of the underlying voice cascade.

One concern regarding this sample is that some relevant voices are not included. For example, some tweets may not mention the official Twitter handle of United Airlines. To evaluate this possibility, I utilize Twitter’s spritzer data, which contains a small percentage of public tweets randomly selected by Twitter. The spritzer data for April 10 of 2017 contains 1,293,706 tweets. From this random sample of tweets, I extracted those satisfying the following two criteria. First, it contains the term doctor and united. Second, it contains at least one of the following terms: drag, knock, beat, forced off, assault, violent. A total of 42 tweets satisfy both criteria, out of which 39 of them, or 93%, mentioned the official Twitter account of United Airlines. It’s also possible that some relevant tweets are more implicit in writing without any mention of the terms I selected. I consider such tweets as “weak” voices because their aftershocks are likely not as strong as those more explicit voices. Because we are mostly interested in voice aftershocks, I think focusing on strong voices with some weak voices potentially missing in the data will not severely impair the research objective of the

---

10 These tweets are available from the author upon request.

Figure 4: The top left panel shows Dr. David Dao after being forcibly removed from United Express Flight 3411 on April 9, 2017. According to the tweets I collected, the incident occurred between 6:40 PM and 7:22 PM. The top right panel shows the number of incident-related tweets for each 5-min interval from 5 AM to 1 PM on April 10, 2017. The bottom left panel shows in a Archimedean spiral incident-related tweets as a point process from 7 PM April 9, 2017 to 4:59 AM, April 10, 2017. Every point in this spiral plot represents one voice. The graph begins at the center of a spiral which represents the first point in the sample, and then progresses outwards as time flows. Similarly, the bottom right panel shows in a Archimedean spiral incident-related tweets as a point process from 5 AM to 8:59 AM, April 10, 2017.
current paper. That being said, I acknowledge that the lack of a complete set of all voice
data is a limitation in the empirical validation of the current paper.

As is shown in the bottom left panel of Figure 4, there are few event-related tweets before
5 AM. This is not very surprising given the daily seasonality pattern. Hence, I use 5 AM as
the start of the sample period. For the end of the sample period, I use four different hours:
10 AM, 11 AM, 12 PM, and 1 PM. This is because the seasonality pattern is simpler and
more stable before 1 PM.

The most important variable in a point process is the time stamp of each point. Consistent
with the notations in the SAGA model, I use $t_i$ to denote the time of the $i$-th voice, in seconds
since 5 AM, April 10, 2017. Equally important in the SAGA model is the mass of each point,
measured by $M_i$, the number of followers of the author of the $i$-th voice at the time. In the
homogeneous SAGA model, the complete data is the ordered set $\{(t_i, M_i)|i \in I\}$ where $I$
the index set of all points during the sample period. For the heterogeneous SAGA model,
additional features about each voice is used to infer heterogeneous quality scores. In Section
5.3, I will use three features of a voice: $m_i$, $h_i$, and $u_i$. The variable $m_i$ is the number
of Twitter accounts mentioned in the $i$-th voice, other than the official account of United
Airlines. The variable $h_i$ is the number of hashtags embedded in the $i$-th voice. The variable
$u_i$ is a dummy variable that is 1 if the $i$-th voice contains a URL link and 0 otherwise. Table
1 reports the summary statistics of the five variables used in the empirical estimation for
four different sample periods.

5 Empirical Results

In this section, I report three sets of estimation results. In Section 5.1, I focus on the
estimation of the baseline SAGA model, that is, the homogeneous SAGA model where the
same quality score is shared by all voices. Hence, in the quality function $ \Phi(X_i^{\prime}\beta)$, $X$
is a constant and its coefficient is the average quality score. This baseline SAGA model is the
Table 1: Summary statistics of variables for different sample periods.

<table>
<thead>
<tr>
<th></th>
<th>Percentile</th>
<th>Mean</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0th 25th 50th 75th 100th</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5:00 - 9:59</td>
<td>t</td>
<td>27 13516 15655 16952 17999</td>
<td>14533</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0 72 290 1016 506977</td>
<td>2618</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>0 0 0 1 7</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>h</td>
<td>0 0 0 0 5</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>u</td>
<td>0 0 0 1 1</td>
<td>0.39</td>
</tr>
<tr>
<td>5:00 - 10:59</td>
<td>t</td>
<td>27 15974 18380 20160 21599</td>
<td>17494</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0 83 311 1073 1203591</td>
<td>3547</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>0 0 0 1 7</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>h</td>
<td>0 0 0 0 5</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>u</td>
<td>0 0 0 1 1</td>
<td>0.38</td>
</tr>
<tr>
<td>5:00 - 11:59</td>
<td>t</td>
<td>27 17750 20708 23024 25199</td>
<td>19882</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0 78 297 1061 6445108</td>
<td>4988</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>0 0 0 1 7</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>h</td>
<td>0 0 0 0 6</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>u</td>
<td>0 0 0 1 1</td>
<td>0.37</td>
</tr>
<tr>
<td>5:00 - 12:59</td>
<td>t</td>
<td>27 19359 23081 26275 28799</td>
<td>22317</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0 71 272 990 6445108</td>
<td>4609</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>0 0 0 1 7</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>h</td>
<td>0 0 0 0 10</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>u</td>
<td>0 0 0 1 1</td>
<td>0.36</td>
</tr>
</tbody>
</table>

foundation of all variations and extensions of the SAGA model. Its estimation and validation are the most important empirical part of this paper. The homogeneous specification is also a "fair" benchmark when we compare the SAGA model with a non-homogeneous Poisson process for which there is no direct way of incorporating voice heterogeneity. In Section 5.2, I provide a validation test of the baseline SAGA model in the same vein as the out-of-sample test in machine learning. In Section 5.3, I consider a natural extension of the baseline SAGA model with heterogeneous quality score in which voices with different content features have different quality scores.

Throughout this section, the energy function in the SAGA model is specified as a linear function, i.e., $\Psi(m, q) = mq$. 

25
5.1 Baseline SAGA

Table 2 reports the estimation results of the SAGA3 model using four different sample periods. I chose SAGA3 because the seasonality model with three segments seems to perform better than the seasonality model with four segments, as is indicated by the comparison of the log-likelihood values of different non-homogeneous Poisson models. Although not reported in this paper, estimation results using SAGA4 model suggests little difference in performance between SAGA3 and SAGA4.

Except the truncation parameter, \( \nu \), all SAGA parameters are statistically significant, which is the first piece of evidence supporting our theoretically motivated model. The fact that we cannot statistically distinguish the amount of left truncation, \( \nu \), from zero is not surprising given the large estimated values of the shape parameter \( k \). Recall that the density function of a gamma distribution with shape parameter \( k \) is 
\[
    f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}},
\]

hence, 
\[
    \lim_{x \to 0} f(x) = f(0) = 0 \text{ if } k > 1, \text{ and for any } n \leq k-1, \text{ we have } \lim_{x \to 0} f^{(n)}(x) = f^{(n)}(0) = 0.
\]

Therefore, the left truncation of the gamma distribution did not change much the density near 0 which is likely the main information source for identifying \( \nu \).

The average quality score, \( q \), is between 2% and 5%, with an overall decreasing trend over time. For example, between 5 AM and noon, on average, roughly 3% of the audience (i.e., followers) of a voice would eventually echo. The overall decreasing trend indicates that voices that occurred earlier are more likely to generate aftershocks, probably due to information novelty at the earlier hours or because those who voiced earlier are more influential.

I estimated five Poisson process models using different seasonality modules. By comparing the log likelihood values of these Poisson process models and that of SAGA3, it’s clear that the SAGA model explains the point process data much better than any of the Poisson models, thereby providing the second piece of evidence supporting the SAGA model.

The estimated shape parameter \( k \) suggests a unimodal and slightly right-skewed temporal pattern of voice aftershocks. So after a voice is generated, the intensity of aftershocks
gradually rises to the peak and then slowly decreases over time. Using the estimated parameters, we can calculate the average duration between a voice and all the echoes triggered by a voice. I will derive this quantity and use it as the model prediction for the validation test next.

### 5.2 A Validation Test

In an ideal research context, we can evaluate the validity of the SAGA model by directly estimating the distribution of all the aftershocks triggered by a voice. However, this would

<table>
<thead>
<tr>
<th>SAGA3</th>
<th>5:00-9:59</th>
<th>5:00-10:59</th>
<th>5:00-11:59</th>
<th>5:00-12:59</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_0$</td>
<td>0.0205***</td>
<td>0.0207***</td>
<td>0.0143***</td>
<td>0.0059***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$k$</td>
<td>79.73***</td>
<td>60.82***</td>
<td>75.63***</td>
<td>127.23***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>209.58***</td>
<td>283.65***</td>
<td>222.70***</td>
<td>151.74***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.0736</td>
<td>0.0481</td>
<td>0.0033</td>
<td>0.0667</td>
</tr>
<tr>
<td></td>
<td>(0.972)</td>
<td>(0.968)</td>
<td>(0.980)</td>
<td>(0.961)</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>7.31e-05***</td>
<td>7.13e-05***</td>
<td>1.123e-04***</td>
<td>2.025e-04***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$q$</td>
<td>0.0392***</td>
<td>0.0480***</td>
<td>0.0298***</td>
<td>0.0256***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>LL</td>
<td>-6816.7</td>
<td>-10342.6</td>
<td>-13878.4</td>
<td>-16846.7</td>
</tr>
<tr>
<td>LL.Poisson0</td>
<td>-9084.8</td>
<td>-15475.0</td>
<td>-21324.8</td>
<td>-26896.8</td>
</tr>
<tr>
<td>LL.Poisson1</td>
<td>-7639.2</td>
<td>-12124.5</td>
<td>-16003.9</td>
<td>-19206.8</td>
</tr>
<tr>
<td>LL.Poisson2</td>
<td>-7638.8</td>
<td>-12123.8</td>
<td>-16068.1</td>
<td>-19494.8</td>
</tr>
<tr>
<td>LL.Poisson3</td>
<td>-7214.5</td>
<td>-11229.3</td>
<td>-14882.2</td>
<td>-17972.2</td>
</tr>
<tr>
<td>LL.Poisson4</td>
<td>-7228.3</td>
<td>-11264.3</td>
<td>-14897.0</td>
<td>-18008.2</td>
</tr>
<tr>
<td>$N$</td>
<td>3,414</td>
<td>7,548</td>
<td>12,606</td>
<td>19,027</td>
</tr>
</tbody>
</table>

Table 2: Estimation results of the baseline SAGA model. p-value in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
require us know, for each voice, whether it is an endogeneous voice or an exogeneous voice, and in the case of an endogeneous voice, which previous voices are the triggers of this voice. Unfortunately, both pieces of information are latent and it’s unclear how they can even be collected.

On the other hand, if we can observe the social network structure of the author of each voice, we may get a glimpse of the set of its potential triggers. Formally, given a voice $i$, define the ancestors of voice $i$, denoted by $A_i$, as the set of users who voiced before $i$ and are also followed by the author of voice $i$. Intuitively, voice $i$ is likely the aftershock of voices in $A_i$. Define

$$d_i \equiv \frac{1}{|A_i|} \sum_{j \in A_i} (t_i - t_j),$$

and

$$d = \frac{1}{N} \sum_i d_i,$$

where $|A_i|$ is the cardinality of the set $A_i$. We can interpret $d$ as the average duration between a voice and its echo, which, according to the SAGA model, can be derived and expressed as

$$d_{SAGA} = \tau + \frac{1}{\Gamma(k, \nu)\theta} \int_0^\infty t \left( \frac{t + \nu\theta}{\theta} \right)^{k-1} e^{-\frac{t + \nu\theta}{\theta}} dt = \tau - \nu\theta + \frac{\Gamma(k + 1, \nu)}{\Gamma(k, \nu)} \theta.$$

Therefore, comparing the values of $d$ and $d_{SAGA}$ offers a nice “out-of-sample” validation test. Such a test is an “out-of-sample” test because the data used to construct $d$ is not used when the parameters of the SAGA model are optimized.

To conduct this validation test, I collected the social network information of all Twitter users in my sample to construct $d$. One caveat of this data is that its collection time is about six months after the voice cascade. Hence, some measurement error exists for some $d_i$. However, if the number of voices is large enough, by the Law of Large Number, I expect the measurement error of $d$ to be small.

Table 3 reports the values of $d$ and $d_{SAGA}$ for each of the four sample periods. From the percentage difference between $d$ and $d_{SAGA}$, we see that the SAGA3 model predicted very well the mean of the actual temporal distribution of aftershocks. In particular, the percentage difference for the longest sample period, between 5 AM and 1 PM, is less than
1%. In general, this validation test should be more accurate with longer sample period because more observations implies more accurate estimation of \( d \), thanks to the Law of Large Number. Given that this is an out-of-sample test, I consider this as the strongest evidence that the SAGA model captures well the underlying stochastic process that generated the voice cascade data.

<table>
<thead>
<tr>
<th></th>
<th>5:00-9:59</th>
<th>5:00-10:59</th>
<th>5:00-11:59</th>
<th>5:00-12:59</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d_{SAGA} )</td>
<td>16813.73</td>
<td>17358.11</td>
<td>16962.86</td>
<td>19416.6</td>
</tr>
<tr>
<td>( d )</td>
<td>15881.58</td>
<td>16113.17</td>
<td>16383.18</td>
<td>19575.85</td>
</tr>
<tr>
<td>Percentage difference</td>
<td>5.87%</td>
<td>7.73%</td>
<td>3.53%</td>
<td>-0.81%</td>
</tr>
<tr>
<td>( N )</td>
<td>3,414</td>
<td>7,548</td>
<td>12,606</td>
<td>19,027</td>
</tr>
</tbody>
</table>

Table 3: Comparing average duration between a voice and its echo predicted by SAGA and approximated from data.

### 5.3 An Extension of SAGA with Heterogeneous Quality

The heterogeneous SAGA model extends the baseline SAGA model by allowing different voices to have different quality scores. So different voices could have different energy levels even if their masses (e.g., number of Twitter followers) are the same. By doing so, we basically conduct a joint estimation of the SAGA model and a voice heterogeneity model. Conceptually, the problem of inferring voice heterogeneity is somewhat orthogonal to the problem of inferring the SAGA model because the latter mostly concerns the temporal shape of influence while the former concerns potential determinants of influence other than the mass.

A natural cause of differential quality scores is content heterogeneity: certain content might be more effective in stimulating others to voice. For simplicity, I measure content heterogeneity in three dimensions. First, because tweets explicitly mentioning other users are more targeted, they are more likely to trigger echoes from those users mentioned. On the flip side, it’s also possible that these voices are more likely to receive less attention from others who are not mentioned. Second, a tweet containing hashtags is more likely to stimulate its
reader to echo. On Twitter, there is a hyperlink embedded in each hashtag, and content consumers can easily click on those hyperlinks to bring out other tweets containing the same hashtag which are often related content. With more exposure, this can potentially increase the content consumer’s propensity to echo. Third, tweets including a URL often contain more information which might increase the probability of echoes. Based on these arguments, I measure content heterogeneity with three variables, $m_i$, $h_i$, and $u_i$, corresponding to the number of other users mentioned, the number hashtags, and whether a URL is included.

Estimating the heterogeneous SAGA model is computationally challenging. Hence, I choose a relatively short sample period to facilitate statistical inference. The sample period between 5 AM and 11 AM seems to give a good balance between estimation time and length of sample period. Hence, I focus on this sample period and report the estimation results in Table 4.

In columns (1) through (3) of Table 4, I only included one of the three variables for content heterogeneity. In column (4), I included all three variables. The coefficients of $m_i$, $h_i$, and $u_i$ in column (4) are all positive and statistically significant, suggesting that a voice that mentions other users, uses hashtags, or includes URL in its content is associated with a higher energy level in generating aftershocks. Comparing the log likelihood value of the baseline SAGA model and those of the four heterogeneous SAGA models, we see there is little increase. This suggests that, in the current empirical analysis, incorporating content heterogeneity does not add much explanatory power over the baseline SAGA model. There could be many reasons for this. For example, maybe there is just very little variation in the latent quality scores of different voices and one’s number of followers already captures one’s influence very well. Alternatively, it’s possible that the current approach of measuring quality score with $m_i$, $h_i$, and $u_i$ is too crude to accurately capture the true heterogeneity of quality scores. Regardless, I think it is an important future research direction to extend the baseline (homogeneous) SAGA model to a heterogeneous SAGA model with more sophisticated method of measuring voice quality.
<table>
<thead>
<tr>
<th>SAGA3</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_0 )</td>
<td>0.021***</td>
<td>0.021***</td>
<td>0.021***</td>
<td>0.020***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>( k )</td>
<td>59.76***</td>
<td>60.03***</td>
<td>59.75***</td>
<td>67.74***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>( \theta )</td>
<td>289.81***</td>
<td>287.9***</td>
<td>289.4***</td>
<td>202.96***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>( \nu )</td>
<td>0.046</td>
<td>0.065</td>
<td>0.046</td>
<td>1.676e-05</td>
</tr>
<tr>
<td></td>
<td>(0.967)</td>
<td>(0.962)</td>
<td>(0.970)</td>
<td>(0.996)</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>7.084e-05***</td>
<td>7.093e-05***</td>
<td>7.084e-05***</td>
<td>7.395e-05***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>mention</td>
<td>0.0323</td>
<td></td>
<td></td>
<td>2.225***</td>
</tr>
<tr>
<td></td>
<td>(0.789)</td>
<td></td>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>hashtag</td>
<td>-0.024</td>
<td></td>
<td></td>
<td>2.492***</td>
</tr>
<tr>
<td></td>
<td>(0.860)</td>
<td></td>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>URL</td>
<td></td>
<td>0.032</td>
<td></td>
<td>3.640***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.823)</td>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td>constant</td>
<td>-3.058***</td>
<td>-3.024***</td>
<td>-3.047***</td>
<td>-13.94***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>LL</td>
<td>-10342.6</td>
<td>-10342.6</td>
<td>-10342.6</td>
<td>-10338.7</td>
</tr>
<tr>
<td>LL of the baseline SAGA</td>
<td>-10342.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Estimation results of extended an SAGA model with heterogeneous quality scores. p-value in parentheses. * \( p < 0.1 \), ** \( p < 0.05 \), *** \( p < 0.01 \)
6 Conclusions

This paper initiated the study of voice cascade on social media, a phenomenon that is both theoretically interesting and practically important, but to my knowledge, has not yet been studied in the IS or marketing literature. As a first attempt to understand voice cascade, I proposed the SAGA model, where the seasonality adjustment is calibrated by real-world data and the gamma aftershock is motivated by a probability model. Using the point process data from a recent voice cascade, I found that the baseline SAGA model fits the data much better than Poisson family models. The heterogeneous SAGA model also reveals some interesting patterns regarding how certain features of a voice are associated with its aftershocks. More importantly, I conducted a validation test, the results of which show promising signs that the SAGA model might have captured some key aspects of voice cascade on social media.

The paper makes two important contributions to the literature. First, as the first study of voice cascade, this paper can potentially stimulate more researchers to explore the mechanism of voice cascade, both theoretically and empirically. Such an endeavor not only is intellectually intriguing, but can also inform managers and help them make better decisions during volatile situations on social media, which is becoming increasingly important as social media further integrates with various aspects of modern societies. In this regard, I believe the potential to start a new stream of scientifically interesting and practically relevant research is the most important contribution of this work. Second, this paper proposed and empirically tested the SAGA model which offers a conceptually broad and mathematically rigorous framework to study voice cascade. This modeling framework is flexible enough to allow extensions in many different directions, but also specific enough to serve as the benchmark model for potential applications.

The paper has several limitations, which also present us with important future research opportunities. First, as mentioned earlier, some voices during the cascade are missing from my data. Also, the social network data I used for a validation test in Section 5 is not the real-time social network information during the voice cascade. Therefore, a valuable and
important extension of this paper is to use more complete and more real-time data for the validation of the SAGA model and for further studies of voice cascade. Second, my current exploration of the heterogeneous SAGA model only use three most salient features of a voice. Admittedly, this is a very crude approach to measure voice quality. An important future direction is to explore more informative features to measure the quality score. Relating to this, it will also be interesting to conduct more research on alternative and potentially more efficient approaches to introduce heterogeneity into the baseline SAGA model. Finally, I think the SAGA model and its current theoretical justification is probably just the tip of the iceberg. More fundamental research on the generalization and variation of the SAGA model is needed to further our understanding of voice cascade and to establish a more elegant mathematical foundation for this phenomenon.

References


Lee, Y., Y. Tan, K Hosanagar. 2018b. Do i follow my friends or the crowd? information cascades in online movie rating.


