Chapter 5
Hawkes Point Processes for Social Media Analytics

Amir Hassan Zadeh and Ramesh Sharda

Abstract Online social networks (OSNs) produce a huge volume of content and clickstream data over time as a result of continuous social interactions between users. Because these social interactions are not fully observable, the mining of such social streams is more complex than traditional data streams. Stochastic point processes, as a promising approach, have recently received significant research attention in social network analysis, in attempts to discover latent network structure of online social networks and particularly understand human interactions and behavior within the social networks. The objective of this paper is to provide a tutorial to the point process framework and its implementation in social media analytics. It begins by providing a quick overview of the history of Hawkes point processes as the most widely used classes of point process models. We identify various capabilities and attributes of the Hawkes point processes and build a bridge between the theory and practice of point processes in social network analytics. Then the paper includes a brief description of some current research projects that demonstrate the potential of the proposed framework. We also conclude with a discussion of some research opportunities in online social network and clickstream point process data.

Keywords Point process • Hawkes process • Self-exciting • Mutual-exciting • Online social networks • Twitter • Online content • Stream mining

5.1 Introduction

During the past few years, millions of people and organizations have used online social networks applications (Facebook, Twitter, YouTube, Google+, etc.) as a part of their daily online activities (Guy et al. 2010). As a result of continuous social interactions between participants over these websites, these platforms have generated, and will continue to generate, enormous amount of data over time. Understanding

A. Hassan Zadeh • R. Sharda (✉)
Department of Management Science and Information Systems,
Oklahoma State University, Stillwater, OK, USA
e-mail: ramesh.sharda@okstate.edu; http://spears.okstate.edu/profiles/?id=71

L.S. Iyer, D.J. Power (eds.), *Reshaping Society through Analytics, Collaboration, and Decision Support*, Annals of Information Systems 18,
DOI 10.1007/978-3-319-11575-7_5

rules and structures governing human interactions and collective behavior is a challenging task in the field of social network analysis. This paper is concerned with the latent networks that cannot fully be observed in online social networks, but have to be analyzed for different purposes. For example, videos on YouTube are watched thousands of times; tweets on Twitter are retweeted, replied and marked as favorite many times by followers; Wikipedia pages are edited quite frequently by contributors; online advertisements or brand posts on Facebook are clicked on by users resulting in popping a web page up from which a purchase may be completed. The common thing among these examples is that activities on one piece of information are likely to cause follow-up activities on itself and other related content. How can we elucidate such interactions and similarities, unravel them from massive unstructured data that generate throughout the days and leverage them for businesses purposes? Numerous approaches have been proposed to address this problem in different ways. Point processes are one of these approaches that have recently received substantial attention in social and behavioral sciences due to their ability to resemble the dynamics of social interactions and discover hidden patterns and implicit network structure within the underlying social structures. On the other hand, the availability of time-stamped and geo-located data collected by various data technologies over these platforms have made it possible to treat this data like point process data, work with point process models and study the spatial-temporal properties of the dynamic social networks.

In this paper, we consider a class of point process models capable of resembling the temporal pattern of social interactions and specifying the temporal dependencies of the interaction events with a branching structure within online social networks. Hawke's self-exciting and mutual-exciting point process (Hawkes 1971) and Ogata's ETAS (Epidemic Type After shock Sequences) model (Ogata 1988), an extension of Hawkes process, are very flexible frameworks which best suited for highly dynamic networks. The Hawkes process model was first designed to model the branching structure of earthquakes. It implemented the idea that an earthquake can trigger aftershocks. Later, the Ogata's point process model was developed to implement the idea that aftershocks sequences have an epidemic behavior, i.e. a large earthquake induces more aftershocks than small earthquakes do. An earthquake with a large magnitude triggering aftershocks is analogous to the activity of an influential user on a content causing follow-up activities by other users on his/her network of followers. Therefore, adapting a point process framework similar to the Hawkes' self-exciting and mutual-exciting point process and ETAS model can be beneficial to analysis that we wish to perform within online social networks.

To the best of our knowledge, there are relatively few studies in literature that have used the point process framework to study dynamic online social networks (Crane and Sornette 2008; Lawrence and Michael 2010; Steeg and Galstyan 2012; Hassan Zadeh and Sharda 2014a, b; Yang and Zhao 2013). Building upon these studies, this paper provides a tutorial for a point process framework and its implementation to social media analytics. We first briefly provide a quick overview of the point process, then present self and mutually exciting or Hawkes processes in detail.

This section basically reviews the statistical theory underlying the Hawkes point process approach. In Sect. 5.4, we briefly review some of the applications of the Hawkes process across various areas of research including applications not only in seismology and finance, but also in healthcare and bioinformatics, social networks, social media, sociology, criminology and terrorism. We then explain some current research projects that demonstrate the usefullness of the Hawkes point process and how the point processes framework is mapped out, to some extent, in online social networks. The final section presents a general conclusion of this paper.

5.2 Point Processes

A point process is a type of random process or a stochastic mechanism that can generate times and spaces of events of a point pattern which we may observe. For instance, the occurrence of an earthquake, fire, neuronal spikes, crime or new thread, content and conversation on social media might be regarded as a point process in both time and geographical space (or in even more general dimensions) if every single point is recorded and mapped according to its position either in time or space or both. Point process models have long been used for describing such real-world phenomena occurring at random locations and/or times (Schoenberg et al. 2010). Temporal Point process models are closely rooted in survival analysis which deals with the durations of time between consecutive events. However, point process models focus on times of events that may appear on the timeline (Rathbun et al. 2006).

A point process can be viewed in terms of a list of times $t_1, t_2, \ldots, t_n$ at which corresponding events *1, 2,…, n* occur (Daley 2006; Daley and Vere-Jones 2003a, b). Intuitively, a point process is characterized by its conditional intensity $\lambda(t)$ which represents the mean spontaneous rate at which events are expected to occur given the history of the process up to time t (Ogata 1988). In particular, a version of the conditional intensity may be given by the process

$$\lambda(t) = \lim_{\Delta t \to \infty} \frac{E\left(N[t, t+\Delta t] \middle| H_t\right)}{\Delta t}$$

where H_t denotes the history of events prior to time t, and the expectation represents the number of events $N[t, t+\Delta t]$ occurring between time t and $t+\Delta t$. The Poisson process is the prototype of a point process that yields random events in space or time where two consecutive events are independent. In other words, a point process is classified as a Poisson process if events occurring at two different times are statistically independent of one another, meaning that an event at time t_1 neither increases nor decreases the probability of an event occurring at any subsequent time. A Poisson process is governed by a single parameter or Poisson intensity. Although Poisson processes have many nice properties which make them particularly well suited for special purposes, they cannot capture interaction effects between events. In the next section, we turn our attention to a more general point process rather than the stationary Poisson process known as Hawkes process. Very useful sources of theoretical discussions and empirical applications of various types of point processes can be found in the textbooks (Daley 2006; Daley and Vere-Jones 2003a, b).

5.3 Hawkes Point Processes

The origin of Hawkes point process goes back to the seventies when Hawkes was looking for a mathematical model that describes earthquake occurrences. Hawkes (1971) introduced a new class of point process to model earthquake occurrences. What was new in his perspective in contrast to other concurrent approaches was a concrete and mathematically traceable point process model with inclusions of branching, self-exciting and self-similarity behaviors (Liniger 2009). The Hawkes process originally states that when an event occurs, it will increase the chance of occurrence of some future events. Over the past few years, Hawkes process models have received significant attentions from researchers, especially in seismology research in terms of theoretical and empirical implications. Though, there exists several equivalent forms in literature in which Hawkes point process can be defined, standard Hawkes process can be defined as a temporal point process with long memory, branching effect and self-exciting properties. Hawkes process is originally characterized by its associated conditional intensity process which allows us to describe the underlying dynamics of the process in a convenient way. The intensity $\lambda(t)$ at a given time t corresponds to the chance of occurrence of an event between time t and $t+\Delta t$ given the history of the process up to time t. Let X= $\{(t_i, m_i)\}$ to be a marked point process on the timeline, where $t_i \subseteq \mathrm{R}$ is an event of the point process and $m_i \subseteq M$ denotes the corresponding mark. The conditional intensity function of the standard Hawkes process is assumed to be of the form

$$\lambda\left(t\right)=\mu\left(t\right)+\sum_{t_i<t}\alpha\left(\kappa_i\right)\beta\left(t-t_i,m_i\right) \tag{5.1}$$

Where $\mu(t)$ denotes an immigrant intensity function. Function $\alpha\ (\kappa)$ represents total offspring intensity and function $\beta(t,\kappa)$ is a density function on $[0,\infty)$. This function is also called normalized offspring intensity which is allowed to depend on the mark m. They are conditional on past events and marks given by history of the process i.e. $H_t : \left\{\left(t_i,m_i\right)\right\}_{t_i<t}$.

Now, we turn our attention to multivariate marked Hawkes process. In this article, we put the definition of Daley and Vere-Jones (2003) into the prospective as the generalized closed form of the Hawkes point process. Let $N(t)=[N_1(t),\ldots,N_d(t)]$ be a multivariate point process that is the superposition of several univariate point processes of different types $\{N_1(t), \ldots, N_d(t)\}$. $N_j(t): j=1,\ldots,d$ denotes the number of points or events type j in the interval $[0,t)$. By definition, a multivariate marked Hawkes process is a class of d-dimensional point process which has d intensity processes given by:

$$\lambda_j\left(t \mid H_t\right) = \mu_j + \sum_{k=1}^{d} \alpha_{kj} \int_{(-\infty,t)\times R} h_{kj}\left(t-s\right) g_j\left(m\right) dN_j\left(s\right) \qquad j = 1,2,\ldots,d \tag{5.2}$$

where the rate of event type (mark) *j*, $\lambda_j(t)$, is determined by the accumulative self- and mutual-excitement effects of the past occurrence of events of all types. Any one event triggers an increase in the rate of the entire process including its associated intensity process and the other *d-1* marked intensity processes. In other words, every event in one of the components increases the intensity of this and all other components. The functions h_{kj} are called transfer functions (also called response, reaction or decay or triggering function in the literature), which are density functions describing the waiting time (lag) distribution between excited and exciting events. These describe how fast the self- and mutual-excitement effects decay in time. The amount of excitement depends on the magnitude of the mark (type) of the triggering event. The fact that a Hawkes process has an underlying clustering structure appears at parameter α_{kj} which indicates the amount of excitation an event type *k* contributes to the time path of component *j*. These branching coefficients reflect the overall behavior of the point process that Hawkes found in his paper (Hawkes and Oakes 1974) to be in the interval [0,1) as the necessary condition for existence.

The most commonly used form of the response function is an exponential decay distribution as follows:

$$h_{kj}\left(t-s\right) = \beta_{kj} e^{-\beta_{kj}\left(t-s\right)} \qquad \left(\beta_{kj} > 0\right) \tag{5.3}$$

Based on the Hawkes process, intensity function specified in Eq. 5.2 and the density function specified in Eq. 5.3, the conditional intensity for the type-*j* point process can be written as:

$$\lambda_j\left(t \mid H_t\right) = \mu_j + \sum_{k=1}^{d} \rho_{kj} \sum_{\substack{n=1 \\ \{t_i < t\}}}^{N_k(t)} \beta_{kj} e^{-\beta_{kj}\left(t-t_n^k\right)} \qquad j = 1,2,\ldots,d \tag{5.4}$$

It indicates that if an event has occurred at time *s*, the intensity is increased at time *t* by amount of $h(t-s)$. The functions $g_j(m)$ are so-called boost functions which are a distinct feature of marked Hawkes point process describing the strength of the event. In other words, if an event type *j* with mark *x* occurs at time *t*, the effect of this event on the time path of the component *j* is proportional to $g_j(m)$. While the transfer and boost functions deal with the relative effects of an event, the branching coefficients imply the absolute influence of the event to the timeline (Steeg and Galstyan 2012).

As pointed out earlier, one of the nice properties of Hawkes point process is the ability to handle a branching structure which facilitates incorporating self-excitement, self-excitation and self-similarity behaviors without even taking into consideration the time and the location of the event. This is a different way of relating events to each other in the way as ancestors and offspring are linked together and allows us to bring the theory of branching processes to the context of point process.

Hawkes point processes are commonly fitted with both parametric and non-parametric estimation techniques. Based on the Hawkes process, intensity function written in Eq. 5.1, the likelihood function for any of the individual point process embedded in the entire process can be written as (Daley and Vere-Jones 2003)

$$L = \prod_{k=1}^{d}\prod_{j=1}^{d}\left[\left[\prod_{v=1}^{N_k(t)}\lambda_j\left(t_n^k \mid H_{t_v^k}\right)\right]e^{-\int_0^T \lambda_j(t \mid H_t dt)}\right] \tag{5.5}$$

Numerical maximization algorithms such as the quasi-Newton method, the conjugate gradient method, the simplex algorithm of Nelder and Mead and the simulated annealing procedure are often implemented to compute maximum log-likelihood estimation of Hawkes process models (Daley and Vere-Jones 2003) as there are no analytically close-form solutions available. Veen and Schoenberg (2008) observed that the log-likelihood functions of branching processes of Hawkes type are complex and extremely flat and numerically unstable due to the multidimensionality, incomplete data and hidden network of branching structure of the Hawkes process. They implemented the idea that Hawkes point process data can be viewed as an incomplete data problem in which the unobservable or latent variables ascertain whether an event belongs to a background event or whether it is a foreground event and was triggered by a preceding occurrence. Therefore, they investigated the expectation-maximization (EM) algorithm as an alternative parameter estimation to estimate Hawkes process parameters and found that it is very efficient compared to traditional methods.

The Bayesian nonparametric inference can also be built as an alternative parameter estimation approach for Hawkes process. Rasmussen (2013) implemented an MCMC (Markov Chain Monte Carlo) algorithm i.e. Metropolis-within-Gibbs algorithm, to perform posterior approximations. Usually, a nonparametric approach leads to a more accurate and robust estimation of parameters.

To assess the goodness-of-fit of the fitted conditional intensity, Q-Q plots of the residual process and the durations (the time intervals between the events of residual process) are drawn. The so-called compensator process is used to perform Kolmogorov- Smirnov (K-S) test to assess the reliability of each model as to the extent to which the model fits the data. This criterion provides useful information of the absolute goodness-of-fit of candidate models. Furthermore, the relative ability of each model to describe the data is measured by computing the Akaike information criteria (AIC) (Akaike 1992). The Akaike statistic provides germane numerical comparisons of the global fit of competing models.

When it comes to the simulation of the point process, the thinning method of Ogata (2006) is often used for the simulation of a point process with the estimated intensity function. This method calculates an upper process for the intensity function which is used to simulate a frequency rate on each mark for the time of the next possible event. Then, the ratio of this rate to the upper bound is compared with a uniform distribution to decide whether the simulated occurrence time is accepted or not (Møller and Rasmussen 2005; Harte 2010). This method is also used for the purpose of prediction. The probability distribution of the time to the next possible event is obtained empirically by simulation outcomes. Based on the in-sample and out-sample performance measures, such as mean absolute error (MAE) (Hyndman and Koehler 2006), the predictive performance of the model can be assessed.

5.4 Hawkes Process Modeling Applications

As mentioned earlier, Hawkes process models have long been used in seismology to recognize similar clustering patterns in earthquake occurrences and to predict subsequent earthquakes, or aftershocks (Adamopoulos 1976; Ogata and Vere-Jones 1984; Ogata 1988; Ogata 1999; Veen and Schoenberg 2008; Wang et al. 2012). In the past few years, point process models have attracted the attentions of researchers from various areas ranging from finance, healthcare and bioinformatics, social networks, to even sociology, criminology and terrorism.

5.4.1 Finance

Starting with papers (Bowsher 2003; Engle and Lunde 2003; Bowsher 2007; Carlsson et al. 2007), Hawkes point process showed the potential to be applicable to a wide variety of problems in economics, finance and insurance. Bowsher (2003, 2007), Engle and Lunde (2003) and later, Bauwens and Hautsch (2009) showed that Hawkes process is able to capture some of the typical characteristics of financial time series. They used Hawkes process models to study the high-frequency price dynamics of financial assets. They proposed mutually exciting or bivariate Hawkes processes as models for the arrival times of trades and succeeding quotes in stock markets and observed that changes in price of a given asset may lead to subsequent quote revisions. Their model helps sellers and buyers determine their pricing strategies by taking into consideration the past prices and trades to decide what price and quote to post. In another study, Hawkes processes have also been proposed as models for the arrival process of buy and sell orders (Carlsson et al. 2007). Bacry et al. (2012a, b) developed multivariate Hawkes process models associated with positive and negative jumps of the asset prices. Their model captures upward and downward changes of prices of assets. Zheng et al. (2014) also extended previous Hawkes process models and proposed a multivariate Hawkes process to describe the dynamics of the bid and ask price of a financial asset.

Another area of finance where Hawkes processes have received significant attentions is risk management and portfolio credit risk analytics. Giesecke and Tomecek (2005), Giesecke et al. (2011) and Errais et al. (2010) revealed that Hawkes processes can also model the credit risk process. They observed that Hawkes processes with exponential transfer function (markov-type Hawkes process) is consistent with the theory of affine jump-diffusion processes in portfolio credit risk and can analyze price processes for certain credit derivatives. Later, Dassios and Zhao (2012) presented a new point process as a generalization of the cox process with short noise intensity and Hawkes process with exponential decay which combines both self-excited (endogenous) and externally excited (exogenous) factors of the underlying system. They used it to model the credit risk process with the arrival of claims and assumed that bankruptcy is caused by primarily a number of bad events such as credit rating downgrades by rating agencies (endogenous factors) and also other bad news on the company front such as bad corporate financial reports (exogenous factors). Their model is capable of capturing additional aspects of the risk, particularly during the economic downturn which involves plethora of bad economic events. In the same line of research, Chavez-Demoulin and McGill (2012) used Hawkes process models featured by a Pareto distribution for the marks to estimate intraday value-at-risk as one of the important metrics used by market participants engaged in high-frequency trading. In another study, Herrera (2013) applied a marked self-exciting point process model to arrival times of extreme events to estimate value-at-risk in oil markets. These are some among many applications of Hawkes process in finance and other related areas which demonstrate that Hawkes processes have some of the typical characteristics of financial time series.

5.4.2 Healthcare and Bioinformatics

Point process models have also been successfully applied in the analysis of a variety of problems in bioinformatics and healthcare domain. In neurosciences, the spikes are the major components that elicit from real-time information processing in the brain. Brillinger (1975, 1988) and Brillinger et al. (1976) was the first one who proposed the Hawkes process in the field of neurophysiology as a model for neural activity in networks of neurons for understanding the mechanisms of what causes a neuron to spike. They used Hawkes point processes as a tool for identifying the relation between connectivity and spike train correlations in small neural networks. Dahlhaus et al. (1997) used Hawkes processes as a tool for identifying direct and indirect synaptic connections in relatively large neural networks. However, in the following years (till 2010), the literature disregarded the linear Hawkes models perhaps due to nonlinear aspects in spike trains, focusing instead on non-linear point process-type models such as the generalized linear models (GLMs), and related multiplicative models (Cardanobile and Rotter 2010) while not offering the same mathematical exposure and simplicity as Hawkes process does. Krumin (2010), Pernice et al. (2011) and Reynaud-Bouret et al. (2013) used Hawkes process models as a tool for spike train analysis for relating neural spiking activity to spiking history, neural ensemble and exogenous effects. They analyzed effects of different connectivity patterns on correlation structure of neuronal activities and observed that the Hawkes process framework is capable of capturing the dynamics of the spike trains in a linear manner as Hawkes process counterparts.

Recently, Hawkes process has been has been used as an analytical model in human physical activity and health. Paraschiv-Ionescu (2013) proposed Hawkes process as a model for understanding human physical activity patterns in health and disease, particularly physical behavior in chronic pain. The central question in their research was how chronic pain affects individuals' daily life and movement. They studied the interactions between chronic pain and regular physical activity and observed that Hawkes process is able to capture the temporal dynamics of human activity patterns between periods and events. They concluded that Hawkes process can improve the clinical understanding of chronic pain behaviors by quantifying the complex dynamics of various human activities.

In the area of disease epidemiology, Meyer (2009) and Kim (2011) used Hawkes' self-exciting point process and Ogata's ETAS (Epidemic-Type After shock Sequences) models to study the spread of infectious disease like flu virus during an epidemic or pandemic. They demonstrated that Hawkes process type models can incorporate spatial and temporal dependencies of outbreaks by specifying a branching structure among the outbreaks in order to predict future occurrences of infectious disease and epidemics.

5.4.3 Sociology, Criminology and Terrorism

Hawkes process has been used in many other areas even in sociology, criminology and terrorism. Very recently, several works addressed the potential of the Hawkes-type models to understand and predict future patterns of violent events and security threats. The fact that some crimes, such as burglary and gang violence tend to happen close to each other in both time and space and spread through local environments contagiously, Mohler et al. (2011), Alexey et al. (2011), Hegemann et al. (2013) and Mohler (2013) took advantage of multidimensionality of Hawkes process across time and space as it was implemented in seismology research, studied the behaviors and rivalries of street gangs and observed that Hawkes process and territorial street gangs exhibit similar behavioral characteristics. They used this model to determine the future urban crime hotspots. Porter and White (2012) and Mohler (2013) used Hawkes-type process models to detect terrorist activities and determine the probability of a terrorist attack occurring in a day, location and the severity of the attack. In similar works, the temporal patterns of violent civilization deaths from the Iraq and Afghan conflicts were explored using self-exciting point processes (Erik et al. 2010; Lewis et al. 2012; Zammit-Mangion et al. 2012).

5.4.4 Social Network Analysis

Recently, there has been a growing interest to use Hawkes point process models for social network analysis. Crane and Sornette (2008) and Lawrence and Michael (2010) developed a family of self-exciting point processes to explore the dynamics of viewing behavior on YouTube. They demonstrated that a Hawkes process with a power law response function exhibits similar characteristics of the viral process of a video on YouTube. These characteristics were classified by a combination of motivational factors (endogenous/exogenous) of user interactions and the ability of viewers to influence others to respond across the network (critical/subcritical). In another study, Lawrence and Michael (2010) used mutually exciting Hawkes process models to understand rules governing collective behaviors and interactions between contributors over Wikipedia. Blundell (2012), Halpin and Boeck (2013) and Masuda (2013) used Hawkes process models to model dyadic and reciprocal interaction within e-mail and conversation networks. Golosovsky (2012) studied the growth patterns of the citation networks and observed that the citation process cannot be a memoryless Markov chain; instead it is consistent with self-exciting point process, since there is an extensive correlation and temporal dependency between the present, recent and past citation rates of a paper. Very recently, Xu et al. (2014) used mutual-exciting Hawkes process to study the dynamic interactions and effect of various types of online advertisements clicks (display, search, purchase) on purchase conversion. Also, Hassan Zadeh and Sharda (2014a, b) and Yang and Zha (2013) built different Hawkes' self- and mutual exciting point process models to investigate the effect of viral diffusion processes on popularity of contents on online social networks.

In summary, Hawkes process has been successfully used in many areas ranging from seismology, finance, medicine, social networks to even criminology and terrorism. Hawkes process models have still this potential to apply to wide variety of other problems to study events or behaviors of interest. Next section outlines two of our previous research projects to demonstrate the capability of Hawkes point process and how point processes models are actually formulated, mapped out and operationalized to some extent in Twitter.

5.5 Hawkes Process Applications in Social Media

Over the past few years, big brands have started taking social media seriously, and social media marketing has been an inevitable part of their marketing plan. As more and more major brands have established their communities within online social networks (OSNs), understanding the behavior of the fans on these platforms became important to the marketers and online content providers in order to enable better organization of online activities, and effective execution of successful marketing campaigns. The central question in our previous works has been to determine the popularity growth patterns of a brand's tweet by analyzing the time-series path of its subsequent activities (i.e. retweets, replies and marks as favorite). Understanding this type of information spreading in social media platforms would potentially allow marketers to predict which trends or ideas will become popular, how fast they will become popular, how much impression a tweet will receive, how long it will be popular, and how often they should tweet. Drawing inspiration from Hawkes process models, we built a self-exciting process model, Ogata's ETAS model and Hawkes mutual-exciting model respectively in order to implement our ideas.

An activity on tweet causing follow-up activities by other users on their network of followers is analogous to an earthquake triggering aftershocks. Second, an earthquake with a large magnitude triggering more aftershocks is analogous to the activity of an influential user on a tweet, inducing more follow-up activities by other users on his/her network of followers; Third, excitation and interaction effects among different types of users' activities (retweets, replies, favorites) is something that needs to be more thoroughly investigated.

The data we crawled from Twitter contained a corpus of an individual brand post tweet, its subsequent activities (retweets, replies, and marks as favorite), along with their timestamps, user IDs and number of followers of the user who contributes to the tweet stream. We took into consideration the timestamp of events, the number of followers and the index or mark attached to it specifying event type "retweet", "reply", and "mark as favorite".

To apply a self-exciting Hawkes process model, we aggregated all users' activities into one single stream of information irrespective to the types of events. As mentioned earlier, self-exciting point process is a simple case of multivariate marked Hawkes point process which is actually a univariate point process, and therefore there is only one intensity process. It indeed ignores the exciting effects among different types of users' activities on a brand's tweet.

To apply an ETAS model, we assumed that the content popularity can be a joint probability function of time and the number of followers. We focused more on incorporating the number of followers as an influential metric into the predictive model of the content popularity, explicitly looking at the impact of influential users on their followers to persuade them to contribute to brand post popularity.

The main difference between these two models is that ETAS model is a Bayesian version of self-exciting process model with a dependent mark that treats the number of followers as a mark. Both formulations essentially lead to the same model. However, ETAS model leads us to a more accurate estimation of parameters due to the underlying Bayesian inference.

In Hassan Zadeh and Sharda (2014a), we observed that incorporating the number of followers into the predictive model of popularity of content presumably provides better results. In behavioral terms, it confirmed our hypothesis that the greater the number of followers per event, the greater the influence.

Also, in the follow-up paper (Hassan Zadeh and Sharda 2014b), we implemented the idea of excitation effects between different types of activities. To apply the mutual-exciting Hawkes model, we separated the dynamics of different types of activities that a given tweet receives over its lifetime in order to measure the popularity of a given online content. This type of Hawkes model includes the exciting effects among different types of users' activities on a brand's tweet into the predictive model. There were three point processes associated with each individual event type category (i.e. retweets, replies and favorites). It allowed us to capture the interacting effects between a stream of events from one to another. Our findings determined that incorporating the type of events into the predictive model of the brand post popularity provides a better understanding of such phenomena.

Several interesting managerial implications were derived from the mathematical models of Hawkes point process presented in previous research regarding the effects of different types (retweet, reply and favorite) of user activities on the popularity of the online content. For example, our model revealed that there are significant exciting effects between the same type of user activities as well as exciting effects between different types of user activities. Our model parameters indicated that retweeting is more likely to excite the other two types of events. This is consistent with the observed data as retweet action is more powerful. Furthermore, we concluded that it is more likely for users to behave like their friends and create the same type of event. For instance, once a brand's post receives "replies" multiple times, it can indicate origination of a conversation thread, and followers reply to the post rather than retweeting. Also, the mathematical model confirmed that users may use "mark as favorite" button to bookmark a tweet that contains subject matter they found interesting but they do not feel like broadcasting to the universe. By marking a tweet as favorite, users just take an action, have it saved and later refer back to it as needed.

Also, our model's parameters suggested that past retweets and replies visible in the timeline are more likely to excite a user to mark the tweet as favorite than being excited by itself. In behavioral terms, it is seldom that users check their friends' favorite tweets and decide to retweet, reply or favorite them.

As seen above, Hawkes process offers a powerful modeling method in social network analysis and has flexibility that can be applied to various kinds of time-dependent data to study events or behaviors of interest occurring in social media streams. There are several research opportunities in line with this research. Hawkes process-based analysis can be done in the context of Facebook, LinkedIn and other social media to study how the formation of relationships and interactions on Twitter is different than other social media platforms. Understanding which types of contents are appealing to audience and how users respond to various stimuli like videos, contests, applications or posts are something that can be more thoroughly investigated with the help of sentiment analysis tools. Also, many applications on social media involve more than a single type of event. It may be useful to treat repeated events of a single type (univariate) on multiple contents with multiple types as forming a multivariate point pattern.

In summary, recently there has been a growing interest to use Hawkes point process models for social network analysis. One of the convincing reasons for growing this interest is that Hawke process models offer a natural and traceable way of modeling time dependencies between events that are arisen as a result of branching, self-exciting and self-similarity behaviors in social networks. The underlying self- and mutual exciting mechanism in Hawkes process models is consistent with the structure observed in social networks. It leads to a nice representation that combines both branching process and conditional intensity representations in one solid model.

The reader should note that in this paper the terms "activity", "event" and "point" are used interchangeably. The term "activity" is used frequently in the context of online social network analysis; however the terms "event" and "point" are often used in the context of stochastic point processes.

5.6 Conclusion

This paper provides an introduction to the Hawkes point process as a very powerful and versatile tool for modeling and understanding social media traffic. These models are used to structure spatial and temporal dependencies between events that are arisen as a result of branching, self-exciting and self-similarity behaviors. It allows us to unravel implicit network structure that cannot fully be observed in online social networks, but have to be analyzed for different purposes in a natural, concrete and traceable computational way. Hawkes process models can be potentially applied to any kind of intensive time-stamped data to study events or behaviors of interest. Two of our previous research papers demonstrate the potential of the Hawkes processes in order to understand the dynamics of human interactions and collective behaviors on social media. Such analysis is fundamental to be able to predict how organization's social media campaign will evolve and grow to achieve the objectives of the campaign.

Biography

Amir Hassan Zadeh is pursuing a Ph.D. in Management Science and Information Systems within the Spears School of Business at Oklahoma State University. He received his master's in Industrial and Systems Engineering from Amirkabir University of Technology, and his bachelor's from Department of Mathematics and Computer Science, Shahed University, Tehran, Iran. His research has been published in Decision Support Systems, Production Planning & Control, Advances in Intelligent and Soft Computing, African Journal of Business Management, and also conference proceedings of DSI, INFORMS and IEEE. His research addresses questions at the interface of operations management and information systems. His current research interest is also focused on the use of big data analytics for social media and recommender systems.

Ramesh Sharda is the Vice Dean of the Watson Graduate School of Management, Watson/ConocoPhillips Chair and a Regents Professor of Management Science and Information Systems in the Spears School of Business at Oklahoma State University. He also serves as the Executive Director of the Ph.D. in Business for Executives Program. He has coauthored two textbooks (Business Intelligence and Analytics: Systems for Decision Support, 10th edition, Prentice Hall and Business Intelligence: A Managerial Perspective on Analytics, 3rd Edition, Prentice Hall). His research has been published in major journals in management science and information systems including Management Science, Operations Research, Information Systems Research, Decision Support Systems, Interfaces, INFORMS Journal on Computing, and many others. He is a member of the editorial boards of journals such as the Decision Support Systems and Information Systems Frontiers. He is currently serving as the Executive Director of Teradata University Network and received the 2013 INFORMS HG Computing Society Lifetime Service Award.

References

Adamopoulos, L. (1976). Cluster models for earthquakes: Regional comparisons. *Journal of the International Association for Mathematical Geology, 8*(4), 463–475.

Akaike, H. (1992). Information theory and an extension of the maximum likelihood principle. In *Proceedings of the second international symposium on information theory* (Vol. 1, pp. 610–624). New York: Springer.

Alexey, S., Martin, B. S., et al. (2011). Reconstruction of missing data in social networks based on temporal patterns of interactions. *Inverse Problems, 27*(11), 115013.

Bacry, E., Dayri, K., et al. (2012a). Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *The European Physical Journal B, 85*(5), 1–12.

Bacry, E., Delattre, S., et al. (2012b). Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance, 13*(1), 65–77.

Bauwens, L., & Hautsch, N. (2009). Modelling financial high frequency data using point processes. In T. Mikosch, J.-P. Kreiß, R. A. Davis, & T. G. Andersen (Eds.), *Handbook of financial time series* (pp. 953–979). Berlin/Heidelberg: Springer.

Blundell, C., Heller, K. A., Beck, J. M., & NIPS. (2012). Modelling reciprocating relationships with Hawkes processes. *Advances in Neural Information Processing Systems, 4*, 2600–2608.

Bowsher, C. (2003). *Modelling security market events in continuous time: Intensity based, multivariate point process models*. Oxford: Nuffield College.

Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics, 141*(2), 876–912.

Brillinger, D. R. (1975). The identification of point process systems. *The Annals of Probability, 3*(6), 909–924.

Brillinger, D. R. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cybernetics, 59*(3), 189–200.

Brillinger, D., Bryant, H., Jr., et al. (1976). Identification of synaptic interactions. *Biological Cybernetics, 22*(4), 213–228.

Cardanobile, S., & Rotter, S. (2010). Multiplicatively interacting point processes and applications to neural modeling. *Journal of Computational Neuroscience, 28*(2), 267–284.

Carlsson, J., Foo, M. C., Lee, H. H., & Shek, H. (2007). High frequency trade prediction with bivariate hawkes process. http://users.iems.northwestern.edu/~armbruster/2007msande444/report1b.pdf.

Chavez-Demoulin, V., & McGill, J. A. (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking & Finance, 36*(12), 3415–3426.

Crane, R., & Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences, 105*(41), 15649–15653.

Dahlhaus, R., Eichler, M., et al. (1997). Identification of synaptic connections in neural ensembles by graphical models. *Journal of Neuroscience Methods, 77*(1), 93–107.

Daley, D. J. (2006). *An introduction to the theory of point processes elementary theory and methods*. Retrieved from http://public.eblib.com/EBLPublic/PublicView.do?ptiID=264777

Daley, D. J., & Vere-Jones, D. (2003). Conditional intensities and likelihoods. In *An introduction to the theory of point processes* (pp. 211–287). New York: Springer.

Daley, D. J., & Vere-Jones, D. (2003b). *An introduction to the theory of point processes* (Vol. 1).

Dassios, A., & Zhao, H. (2012). Ruin by dynamic contagion claims. *Insurance Mathematics and Economics, 51*(1), 93–106.

Engle, R. F., & Lunde, A. (2003). Trades and quotes: A bivariate point process. *Journal of Financial Econometrics, 1*(2), 159–188.

Lewis, E., Mohler, G., Brantingham, P. J., & Bertozzi, A. L. (2012). Self-exciting point process models of civilian deaths in Iraq. *Security Journal, 25*(3), 244–264.

Errais, E., Giesecke, K., et al. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics, 1*(1), 642–665.

Giesecke, Kay., & Pascal, T. (2005). Dependent events and changes of time. Cornell University

Giesecke, K., Goldberg, L. R., et al. (2011). A top-down approach to multiname credit. *Operations Research, 59*(2), 283–300.

Golosovsky, M., & Solomon, S. (2012). Stochastic dynamical model of a growing citation network based on a self-exciting point process. *Physical Review Letters, 109*(9).

Guy, I., Jacovi, M., et al. (2010). Same places, same things, same people?: mining user similarity on social media. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 41–50). Savannah: ACM.

Halpin, P., & Boeck, P. (2013). Modelling dyadic interaction with Hawkes processes. *Psychometrika, 78*(4), 793–814.

Harte, D. (2010). PtProcess: An R package for modelling marked point processes indexed by time. *Journal of Statistical Software, 35*(8), 1–32.

Hassan Zadeh, A., & Sharda, R. (2014a). Modeling brand post popularity dynamics in online social networks. *Decision Support Systems, 65*(0), 59–68.

Hassan Zadeh, A., & Sharda, R. (2014b). *A point process framework for predicting popularity of online content in online social networks*. Available at SSRN 2331565.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika, 58*(1), 83–90.

Hawkes, A. G., & Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability, 11*(3), 493–503.

Hegemann, R., Lewis, E., et al. (2013). An "Estimate & Score Algorithm" for simultaneous parameter estimation and reconstruction of incomplete data on social networks. *Security Informatics, 2*(1), 1–13.

Herrera, R. (2013). Energy risk management through self-exciting marked point process. *Energy Economics, 38*, 64–76.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688.

Kim, H. (2011). *Spatio-temporal point process models for the spread of avian in anfuenza virus (H5N1)*. Retrieved from http://digitalassets.lib.berkeley.edu/etd/ucb/text/Kim_berkeley_0028E_11401.pdf

Krumin, M., Reutsky, I., & Shoham, S. (2010). Correlation-based analysis and generation of multiple spike trains using Hawkes models with an exogenous input. *Frontiers in Computational Neuroscience, 4*, 147.

Lawrence, M., & Michael, E. C. (2010). Hawkes process as a model of social interactions: a view on video dynamics. *Journal of Physics A: Mathematical and Theoretical, 43*(4), 045101.

Lewis, E., Bertozzi, A. L., Mohler, G., & Brantingham, P. J. (2012). Self-exciting point process models of civilian deaths in Iraq. *Security Journal, 25*(3), 244–264.

Liniger, T. J. (2009). *Multivariate Hawkes processes*. Doctoral dissertation, Eidgenössische Technische Hochschule ETH Zürich, Nr. 18403.

Masuda, N., Takaguchi, T., Sato, N., & Yano, K. (2013). Self-exciting point process modeling of conversation event sequences. In *Temporal Networks* (pp. 245–264). Springer Berlin Heidelberg.

Meyer, S. (2009). *Spatio-temporal infectious disease epidemiology based on point processes*. Retrieved from http://epub.ub.uni-muenchen.de/11703/1/MA_Meyer.pdf

Mohler, G. (2013). Modeling and estimation of multi-source clustering in crime and security data. *The Annals of Applied Statistics, 7*(3), 1525–1539.

Mohler, G. O., Short, M. B., et al. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association, 106*(493), 100–108.

Møller, J., & Rasmussen, J. G. (2005). Perfect simulation of Hawkes processes. *Advances in Applied Probability, 37*(3), 629–646.

Ogata, Y. (1988). Statistical models for Earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association, 83*(401), 9–27.

Ogata, Y. (1999). Seismicity analysis through point-process modeling: A review. In M. Wyss, K. Shimazaki, & A. Ito (Eds.), *Seismicity patterns, their statistical significance and physical meaning* (pp. 471–507). Basel: Birkhäuser.

Ogata, Y. (2006). On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory, 27*(1), 23–31.

Ogata, Y., & Vere-Jones, D. (1984). Inference for earthquake models: A self-correcting model. *Stochastic Processes and their Applications, 17*(2), 337–347.

Paraschiv-Ionescu, A., Buchser, E., & Aminian, K. (2013). Unraveling dynamics of human physical activity patterns in chronic pain conditions. *Scientific Reports, 3*, 2019. doi:10.1038/srep02019.

Pernice, V., Staude, B., et al. (2011). How structure determines correlations in neuronal networks. *PLoS Computational Biology, 7*(5), e1002059.

Porter, M. D., & White, G. (2012). Self-exciting hurdle models for terrorist activity. *Annals of Applied Statistics, 6*(1), 106–124.

Rasmussen, J. (2013). Bayesian inference for Hawkes processes. *Methodology and Computing in Applied Probability, 15*(3), 623–642.

Rathbun, S. L., Shiffman, S., & Gwaltney, C. J. (2006). Point process models for event history data: applications in behavioral science. *Models for intensive longitudinal data, 219*.

Reynaud-Bouret, P., Rivoirard, V. et al. (2013). *Inference of functional connectivity in neurosciences via Hawkes processes*. Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE, Piscataway.

Schoenberg, F. P., Cochran, J. J., et al. (2010). Introduction to point processes. In *Wiley encyclopedia of operations research and management science*. John Wiley & Sons, Inc.

Steeg, G. V., & Galstyan, A. (2012). Information transfer in social media. In *Proceedings of the 21st international conference on World Wide Web* (pp. 509–518). Lyon: ACM.

Veen, A., & Schoenberg, F. P. (2008). Estimation of space–Time branching process models in seismology using an EM–Type algorithm. *Journal of the American Statistical Association, 103*(482), 614–624.

Wang, T., Bebbington, M., et al. (2012). Markov-modulated Hawkes process with stepwise decay. *Annals of the Institute of Statistical Mathematics, 64*(3), 521–544.

Xu, L., Duan, J. A., Whinston, A. B. (2014). *Path to purchase: A mutually exciting point process model for online advertising and conversion. Management Science, 60*(6), 1392–1412.

Yang, S. H., & Zha, H. (2013). Mixture of mutually exciting processes for viral diffusion. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)* (pp. 1–9).

Zammit-Mangion, A., Dewar, M., et al. (2012). Point process modelling of the Afghan War Diary. *Proceedings of the National Academy of Sciences, 109*(31), 12414–12419.

Zheng, B., Roueff, F., et al. (2014). Modelling bid and ask prices using constrained Hawkes processes: Ergodicity and scaling limit. *SIAM Journal on Financial Mathematics, 5*(1), 99–136.