Predicting Power Failures with Reactive Point Processes

Şeyda Ertekin
MIT Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
seyda@mit.edu

Cynthia Rudin
MIT Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
rudin@mit.edu

Tyler H. McCormick
Department of Statistics
University of Washington
Seattle, WA 98195, USA
tylermc@u.washington.edu

Introduction

We present a new statistical model for predicting discrete events continuously in time, called Reactive Point Processes (RPP's). RPP's are a natural fit for many domains where time-series data are available, and their development was motivated by the important problem of predicting serious events (fires, explosions, power failures) in the underground electrical grid of New York City (NYC). RPP's capture several important properties of this domain:

- There is an instantaneous rise in vulnerability to future serious events immediately following an occurrence of a past serious event, and this rise in vulnerability gradually fades back to the baseline level. This is a type of *self-exciting* property.
- There is an instantaneous decrease in vulnerability due to an inspection or repair. The effect of this inspection fades gradually over time. This is a *self-regulating* property.
- The cumulative effect of events or inspections can saturate, ensuring that vulnerability levels never stray too far beyond their baseline level. This captures *diminishing returns* of many events or inspections in a row.

RPP's are related to Hawkes processes (Hawkes 1971a; 1971b), and more specifically, self-exciting point processes (SEPP). Demonstrated to be highly effective for earthquake modeling (Ogata 1988; 1998), SEPP have been used in a vast array of domains, including criminology (Mohler et al. 2011), neuroscience (Krumin, Reutsky, and Shoham 2010) and social networks (Crane and Sornette 2008). A common theme in these applications is that the likelihood of an event at a certain time is probabilistically dependent on earlier events, even after controlling for covariates. RPP's extend these models to a much broader array of settings where external events can cause vulnerability levels to drop or where diminishing returns can occur.

We apply RPP's to the critical problem of predicting power failures in New York City using data from a large-scale ongoing collaboration with Con Edison, New York City's power utility company (Rudin et al. 2010; 2012). Manholes are access points to NYC's underground electrical grid, and we have features of each manhole that can be useful in predicting vulnerability to serious events such as manhole fires and explosions, which can cause power failures. The RPP model takes into account past events that are correlated with increased future vulnerability, past inspections that reduce future risk, and also the manhole's characteristics (e.g., number and age of cables entering the manhole). The dataset is comprised of processed trouble tickets, inspection reports, and cable information for all manholes in Manhattan (total 213,504 records for 53,525 manholes). We have data from the year 1995 up to and including 2010.

Model

Our approach models events as being generated from a non-homogenous Poisson process for entity (manhole) j with intensity $\lambda_j(t)$ where

$$\lambda_j(t) = \lambda_0 \Bigg[1 + g_1\Bigg(\sum_{\forall t_e < t} g_2(t - t_e) \Bigg) - g_3\Bigg(\sum_{\forall \bar{t}_i < t} g_4(t - \bar{t}_i) \Bigg) + c_1 \mathbf{1}_{[N(\tau) \geq 1]} \Bigg]$$

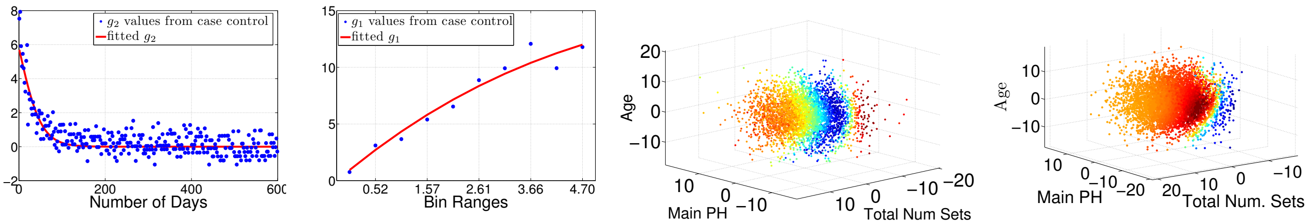
where t_e are event times and $\bar{t}_i$ are inspection times. λ_0 denotes the baseline vulnerability. The vulnerability level permanently goes up by c_1 if the number of events on the interval $[0, \tau]$, $N(\tau)$, is more than zero. The constant c_1 can be fitted from the observed data. Functions g_2 and g_4 are the self-excitation and self-regulation functions, which have initially large amplitudes and fade over time. Note that SEPP's have only g_2, and not the other functions. Functions g_1 and g_3 are the saturation functions, which start out as the identity function and then flatten farther from the origin. If the total sum of the excitation terms is large, g_1 will prevent the vulnerability level from increasing too much. Similarly, g_3 controls the total possible amount of self-regulation.

For the electrical grid data, we used a family of functions of the form

$$g_1(x) = a_{11} \times \left(1 - \frac{\log(1 + e^{-a_{11}x})}{\log(2)}\right), \quad g_2(x) = \frac{1}{1 + e^{\beta x}}$$

$$g_3(x) = -a_{31} \times \left(1 - \frac{\log(1 + e^{-a_{32}x})}{\log(2)}\right), \quad g_4(x) = \frac{-1}{1 + e^{-\gamma x}},$$

where $a_{11}, a_{12}, a_{31}, a_{32}, \beta$, and γ are parameters that can be either constant across all manholes, or can be adapted based



(a) Fitted g_2 (b) Fitted g_1 (c) KL Divergence (d) Log-Likelihood

Figure 1: Results for the Manhattan Dataset. Figures 1(a) and 1(b) show the fitted functions for the Manhattan dataset from Case Control. Figures 1(c) and 1(d) show the KL Divergence from ABC and the Log-Likelihood, respectively.

on the manhole's features. For instance, our experimental analysis covers both constant β across all manholes as well as an adaptive β, modeled for manhole k as:

$$\beta_k = \log\left(1 + e^{-\sum_{i=1}^{n} \theta_i f_{ik}}\right)$$

where θ_i is the coefficient of feature i and f_{ik} is the value of feature i for manhole k. For our dataset, the features include the number of main phase cables, age of the oldest cable, and the total number of cable sets inside the manhole.

Model Fitting and Experiments

We will introduce several methods for fitting RPP's below. To verify that the methods are working properly, we used simulated data where ground truth was available (not shown here) to demonstrate that the underlying function could be recovered.

Case Control

The Case Control method allows us to nonparametrically trace out the shapes of g_1, g_2, g_3, and g_4, using only relevant manholes to estimate specific points of each function. Figures 1(a) and 1(b) show g_2 and g_1 from Manhattan manholes, traced out (in blue) using the Case Control method, and parameterized (in red) afterwards. For the g_1 curve, the range of observed $\sum g_2$ values in the data was discretized into equal sized bins, and the likelihood of an event if the sum value falls within the range of a bin was used to fit g_1.

Bayesian and Maximum Likelihood Approaches

The Case Control method is useful for finding parameterizations of the functions, but there are not enough data to permit its use when we would like to customize g_1, g_2, g_3, and g_4 to each individual manhole. Based on insights gleaned from the Case Control method we explore both Bayesian and Maximum Likelihood approaches to model fitting.

For our Bayesian approach, we use the model for the β_k's with covariates described above, where each coefficient has a Gaussian prior distribution. We fit the model using Approximate Bayesian Computation (ABC). To do this, we evaluated the suitability of simulated data generated using a given candidate parameter via two statistics, one consisting of the number of events and another related to the distribution of time between events. Since RPP's model association between event times, the second statistic is important. We use KL Divergence to measure differences in the distribution of time between events for the observed data and the distribution of time between events generated using our candidate parameter values. In Figure 1(c) the blue boundary corresponds to the region that yields the minimum divergence, and the optimum parameter values can be derived by fitting a manifold on that region.

We also experimented with model fitting using Maximum Likelihood techniques. Figure 1(d) demonstrates the region of maximum likelihood (shown in dark red), where the simulated vulnerabilities are in closest proximity of actual vulnerabilities.

Prediction Performance Evaluation

We assessed the predictive performance of our RPP model by comparing it against Cox Proportional Hazards Model (Cox 1972), which is a commonly used technique in survival analysis. We trained both an RPP model and a Cox model (available in the survival package in R) on the historical event records up to the year 2010 and used the models to predict vulnerabilities of each manhole on each day of 2010. We then generated a ranked list of manholes for each day with the most vulnerable manhole ranked at the top. For each manhole that had an event during a day, we compared the rank of that manhole in the two ranked lists. Aggregated over days in 2010 where an event was recorded, the RPP model ranked the manholes with events significantly higher than the Cox model in its lists, indicating that the RPP model more accurately measures short term vulnerabilities.

Understanding the Impact of Inspection Policies

Beyond identifying the vulnerable manholes to be targeted for repair and inspection, we used the RPP model to simulate the effect of different broad inspection policies. Our simulations show that there is a natural inverse relationship between the number of events and the number of inspections: more inspections lead to fewer events. Using our simulations, we can quantify an optimal trade off between the anticipated number of events and the cost of inspections. This can assist Con Edison to make decisions between candidate inspection policies.

References

Cox, D. R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 187–220.

Crane, R., and 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.

Mohler, G.; Short, M.; Brantingham, P.; Schoenberg, F.; and Tita, G. 2011. Self-exciting point process modeling of crime. *Journal of the American Statistical Association* 106(493):100–108.

Krumin, M.; Reutsky, I.; and Shoham, S. 2010. Correlation-based analysis and generation of multiple spike trains using Hawkes models with an exogenous input. *Frontiers in computational neuroscience* 4.

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. 1998. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* 50(2):379–402.

Rudin, C.; Passonneau, R.; Radeva, A.; Dutta, H.; Ierome, S.; and Isaac, D. 2010. A process for predicting manhole events in Manhattan. *Machine Learning* 80:1–31.

Rudin, C.; Waltz, D.; Anderson, R. N.; Boulanger, A.; Salleb-Aouissi, A.; Chow, M.; Dutta, H.; Gross, P.; Huang, B.; Ierome, S.; Isaac, D.; Kressner, A.; Passonneau, R. J.; Radeva, A.; and Wu, L. 2012. Machine learning for the New York City power grid. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34(2):328–345.

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

Hawkes, A. 1971. Point Spectra of Some Mutually Exciting Point Processes. *Journal of the Royal Statistical Society* 33(3):438–443.