Mining 911 Calls in New York City: Temporal Patterns, Detection and Forecasting

Alex Chohlas-Wood, Aliya Merali, Warren Reed, Theodoros Damoulas
Center for Urban Science & Progress (CUSP)
New York University
{alex.cw, ajm777, jwarren.reed, damoulas}@nyu.edu

Abstract

The New York Police Department (NYPD) is tasked with responding to a wide range of incidents that are reported through the city's 911 emergency hotline. Currently, response resources are distributed within police precincts on the basis of high-level summary statistics and expert reasoning. In this paper, we describe our first steps towards a better understanding of 911 call activity: temporal behavioral clustering, predictive models of call activity, and anomalous event detection. In practice, the proposed techniques provide decision makers granular information on resource allocation needs across precincts and are important components of an overall data-driven resource allocation policy.

Introduction

The NYPD is responsible for responding to many different types of frequent incidents, including emergency 911 calls and crimes in progress. Every week, precinct commanders allocate a limited number of patrol cars to respond to these events. Patrol cars are assigned to specific geographical sub-divisions of precincts, referred to as "sectors", of which there are typically ten per precinct. In this process, the primary information considered by commanding officers is the overall recent history of crime events and the "traditional" crime patterns that persist in each geographical area. Precinct commanders make these allocation decisions using high-level aggregate statistics and mental reasoning, a process which may not fully consider the spectrum of call behaviors, correlative conditions, and exceptional anomalies in their precinct.

Much of the recent work in quantifying police-related activity has centered around predictive policing. This term, of rising popularity in the policing community, has been used broadly to describe the quantified prediction of criminal activity in order to intervene before events occur (Perry 2013). Many of these studies are formulations of mathematical models of criminal behavior (Berestycki and Nadal 2010; Mohler et al. 2011; Short et al. 2010; Zipkin, Short, and Bertozzi 2013; Short et al. 2008). These models may be useful for understanding the generating processes behind criminal hotspots, but do not address resource allocation issues, or demand for police resources outside of hotspot locations. A few studies have examined the use of more data-centric police demand forecasting, primarily focusing on spatiotemporal anomaly detection (Neill and Gorr 2007; Woodworth et al. 2013).

In this paper, we outline three distinct techniques that incorporate emergency call data — along with other urban datasets — to provide insight into local demand for police resources. First, we seek to discover temporal behavior patterns in local emergency call behavior, across different time windows and geographical areas. Extracting these patterns will allow precinct commanders to understand different types of typical behavior in their precinct in much greater detail when compared to the overall aggregate statistics that are currently used.

Second, we predict the future demand for police resources in high-resolution geographical areas and time bins. Prediction models can incorporate a wide range of conditions that may correlate with significant changes in demand for police resources. These conditions can be used to predict demand, and therefore automate the mental guesswork that is currently required for resource allocation.

Finally, we seek to detect emergency call behavior that is unusual when compared to an existing "normal" spatiotemporal baseline. This method would allow a precinct commander to quickly detect and respond to exceptional events in their precinct. It would also inform the allocation of resources that are dedicated to respond to extreme or very unusual events.

Each area of focus provides immediate benefit to the weekly decisions of precinct commanders, by focusing on outputs that are easy to interpret and incorporate in the decision-making process. These outputs will also be incorporated in a future optimized allocation system that will make evolving recommendations on the distribution of police resources.

Data

We worked with 2 years of data from the city's 911 dispatch systems: the retired Special Police Radio Intelligence Network (SPRINT) system, and its new replacement, Intergraph Computer Aided Dispatch (ICAD). Much of our analysis focused on ICAD, which describes each call in greater detail than SPRINT. Both datasets record every police response to

Table 1: Datasets

Dataset	Date Range	Size
911: ICAD	6/2013–12/2013	2.7 million
911: SPRINT	2/2012–5/2013	6.3 million
311 Complaints	2/2003–5/2014	7.4 million

Table 2: 311 Complaint Type Groupings

Groupings	Complaint Types
Noise	Noise
Street Condition	Street Condition, Blocked Driveway, Traffic Signal Condition, Illegal Parking
Building	Sewer, Water System, Building/Use, General Construction/Plumbing
Dirty Conditions	Dirty Conditions, Rodent, Sanitation Condition, Missed Collection, Damaged Tree, Derelict Vehicle

an event, called in by the public or self-initiated by officers, but excluding events requiring solely Emergency Medical Services (EMS) and/or firefighters.

311 call data was also included for behavioral analysis and as a predictive feature for 911 analysis. This dataset includes not only direct complaints from citizens, but also updates and status reports from other agencies. 311 contains a high number of distinct complaint categories, so we grouped the most common complaint types into four overarching categories, as indicated in Table 2.

Many 311 complaints do not contain location information, and many calls do not fall within sector boundaries. For these reasons, our final sector-labeled 311 dataset contained 6.7 million records. Other open datasets such as weather[1] and Primary Land Use Tax Lot Output[2] (PLUTO) data, were employed as covariates in the forecasting model.

Temporal Call Behavior

Our initial step sought to understand typical temporal behavioral patterns for calls. These distinct behavioral patterns highlight peak times of emergency call demand, as well as highlighting the difference in behavior between different sectors. We chose k-means clustering with an L^1 norm as a way to extract distinct types of behavior. Behavior was represented by hourly time series of total emergency call counts, with separate time series for each sector s, call type m, and day of week d. Given $N = S \times M \times D$ total time series $\mathbf{c}_1, \ldots, \mathbf{c}_n \in \Re^{24}$, the objective function attempts to find clusters $l_1, \ldots, l_j$ to minimize:

$$\underset{\{l_j\}_{j=1}^{k}}{\arg\min} \sum_{j=1}^{k} \sum_{\mathbf{c} \in l_j} |\mathbf{c} - \boldsymbol{\mu}_j|$$

[1]http://www.nrcc.cornell.edu/

[2]http://www.nyc.gov/html/dcp/html/bytes/applbyte.shtml

where $\boldsymbol{\mu}_j$ is the mean of points in l_j

Data were aggregated to create a 24-hour cumulative call profile for each type of call (crime or non-crime), in each geographic region, for each day of the week, and was then z-normalized. The number of clusters (k=3) was selected based on a balance of a high silhouette score while maintaining legibility for precinct commanders. The clusters in Figure 1 demonstrate three distinct call patterns with shifting peak hours. Cluster 0 indicates a peak in non-crime calls during working hours on weekdays, whereas Cluster 1 shows a weekday evening peak in crime calls. Cluster 2 includes the late night weekend activity patterns suggested by anecdotal evidence from NYPD police officers.

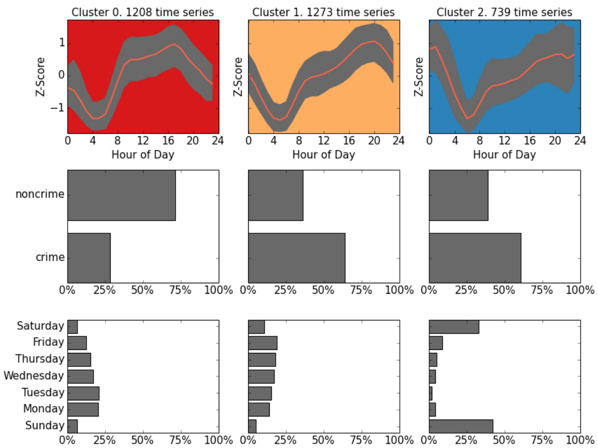

Figure 1: 911 k-means clustering results: cluster 0 peaks during the workday hours with primarily non-crime calls; cluster 1 peaks in the evening with primarily crime calls; cluster 2 peaks in the late night hours with mainly crime calls on the weekend days

Figures 2 and 3 are examples of clustering results by precinct and sector, respectively, during various time periods. These clustering maps provide a deeper level of spatiotemporal detail on historical crime patterns to precinct-level commanding officers. The visual representation of call patterns will allow officers to identify which sectors historically require resources during the various shifts, informing allocation decisions.

Predictive Model

A reliable predictive model of call demand in each sector will be a central component to optimize precinct-level resource allocation. The output of such a model has clear immediate value, along with providing material for subsequent optimization and detection measures. In the short term, the output of a predictive model can be used by precinct commanders to predict demand and inform where they allocate resources. In the long term, the outputs of such a model will be used as a reliable baseline for automated allocation techniques and event detection.

Given that patrol units are assigned in eight-hour shifts to sectors, and that crime-related calls are the highest pri-

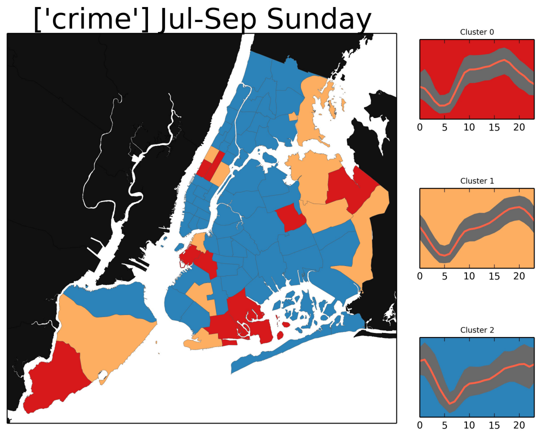

Figure 2: Summer Crime Clusters by Precinct: the "Late Night" cluster dominates, as expected, while regions peaking in crime calls during the day include the Upper East and Upper West sides, which are typically quiet in the overnight hours

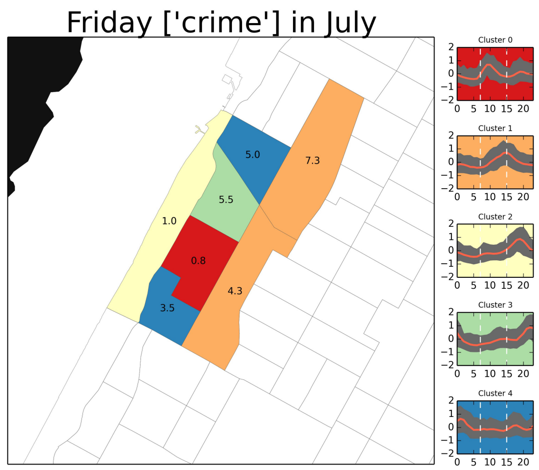

Figure 3: July Crime Clusters in the 26th Precinct on Fridays; labels on each sector indicate the average number of crime calls in that sector for the given temporal specification

ority for response, we built a supervised learning model to forecast the number of crime calls in each sector for each shift. We chose to forecast counts specifically because they are easily understood by precinct commanders and have a natural connection to both manual and automatic resource allocation.

Feature Extraction

Most of the extracted features were two statistics – historical means and medians – for ICAD 911 calls and 311 complaint calls. These features were calculated over retrospective windows of differing sizes: from as little as one hour to as much as eight weeks before each prediction. Additionally, we included more of the same statistics, but restricted them to retrospective windows over the same day of week and/or time

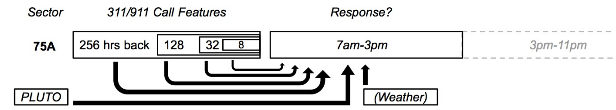

Figure 4: An illustration of the feature extraction process for one (of many) prediction targets.

of day. These statistics were calculated not only for the predicted target, but also for all calls, for every call type, and for every call outcome. The same statistics were also aggregated for 311 calls over identical historical windows. The statistics were calculated for all 311 calls, for the complaint type groupings, for the 20 most common complaint types, and the 20 most common call descriptors.

In addition to the autoregressive-type features built from past 911 and 311 data, we also incorporated weather data, PLUTO building features, and contextual information. Weather metrics included temperature, humidity, solar radiation, and precipitation. Features extracted from PLUTO provided information on the characteristics of the sector, including the total area and count of each building type (i.e. two-family residential vs. commercial high-rise) and owner type, along with land value estimates and building construction dates. Finally, a number of contextual categorical variables were included using one-in-K encoding: day of week, location of prediction bin in day, and the current sector's 311 and 911 cluster classifications, as produced by k-means clustering discussed above.

In summary, each sample n contains the count of crime calls $c_{s,t}$ in a given sector s over an eight-hour period Δt starting at t, as well as a feature set $\mathbf{x}_{s,t}$:

$$n = \{\mathbf{x}_{s,t}, c_{s,t}\}_{s=1:770, t=0:\Delta t:T}$$

There are $N = S \times (T/\Delta t)$ samples overall available for both training and testing.

Each $\mathbf{x}_{s,t}$ contains a number of different feature sets, derived from 911, 311, PLUTO, and Weather data:

$$\mathbf{x}_{s,t} = \{\mathbf{C}^{911}_{s,t-\boldsymbol{\tau}}, \mathbf{C}^{311}_{s,t-\boldsymbol{\tau}}, \mathbf{x}^{\text{PLUTO}}_{s}, \mathbf{x}^{\text{Weather}}_{t}, \mathbf{x}^{\text{Contextual}}_{t}\}$$

Where $\boldsymbol{\tau}$ is the set of retrospective windows, from 1 to τ_{max}, in hourly units:

$$\boldsymbol{\tau} = \{1, 2, 4, \ldots, \tau_{\text{max}}\}$$

with τ_{max} chosen as 1,344 hours (8 weeks) back. Over these retrospective windows, we derive the first moment $\bar{c}$, second moment $\text{var}(c)$, and median count $\tilde{c}$ of various types of calls.

Initial runs ran on a set of about 2,600 features. In subsequent steps, many of the historical counts of 911 and 311 call records that had extremely low variance and/or many zero-counts were discarded. Features were also pruned based on feature importances from early runs with random forest regression. Final runs had comparative predictive power with just over 300 features. The large majority of this reduced set of features were the historical statistics on 911 calls.

Model

To mirror how this may be implemented at NYPD, the model was trained and tested using a "rolling forecast" method, as

Algorithm 1 Rolling Forecast Prediction Model

for $s = 1$ **to** S **do**
  for $t = 0$ **to** T **by** Δt **do**
    for τ in $\boldsymbol{\tau}$ **do**
      $\mathbf{C}^{911}_{t-\tau} = \{\bar{c}^{911}_{t-\tau}, \mathrm{var}(c^{911}_{t-\tau}), \tilde{c}^{911}_{t-\tau}\}$
      $\mathbf{C}^{311}_{t-\tau} = \{\bar{c}^{311}_{t-\tau}, \mathrm{var}(c^{311}_{t-\tau}), \tilde{c}^{311}_{t-\tau}\}$
    end for
  end for
  $\mathbf{x}_{s,t} = \mathbf{C}^{\ 911}_{\ s,t}, \mathbf{C}^{\ 311}_{\ s,t}, C^{\ \text{PLUTO}}_{\ s}, C^{\ \text{Weather}}_{\ t}, C^{\ \text{Contextual}}_{\ t}$
end for
Input: $\{\mathbf{x}_{s,t}, c_{s,t}\}^N_{n=1}$
for $n = n_{\text{boundary}}$ **to** N **do**
  fit $\{\hat{c}_{s,t}\}^{n-1}_{n=1}$ = f($\{\mathbf{x}_{s,t}\}^{n-1}_{n=1}$, w)
  predict $c^{n*}_{s,t}$ = f($\mathbf{x}^n_{s,t}$, ŵ)
end for
Output: $\{c_{s,t}\}^N_{n=n_{\text{boundary}}}$

demonstrated in Algorithm 1. With this technique, each target is predicted using a model trained on all available historical data. Modeling begins with a sufficient amount of training data, noted as n_{boundary} and chosen as 90 eight-hour periods. As soon as prediction for one date is complete – say, October 5th, 2013 – the model incorporates October 5th into the training set, retrains the model, and issues predictions for October 6th. Although we trained one ensemble model for the entire city, ongoing work explores building more localized models for smaller geographic areas.

We focused on two classes of models: random forest regression (Breiman 2001) and Poisson regression. We began with a random forest model due to its ability to handle high-dimensional data and scale up to millions of samples while capturing nonlinear relationships. Individual trees used mean squared error to measure the quality of each split, considering $\sqrt{\|\mathbf{x}\|}$ features at each split, trained on a bootstrapped sample drawn from all N available samples. The results reported from the random forest model used 100 trees; higher numbers of trees did not result in noticeably improved performance.

Our Poisson regression modeled the expected call-crime count of each sector as a log-linear model of the extracted autoregressive features and the spatiotemporal descriptors of weather and PLUTO data, as follows:

$$\mathrm{E}(Y|\mathbf{X}) = \exp(\mathbf{w}^{\mathrm{T}}\mathbf{X}),$$

where $\mathbf{X} = \{\mathbf{x}_{s,t}\}^N_{n=1}$.

Performance

We began by predicting call counts for entire days using random forest regression. Results for one day on this model are shown in Figure 5. As performance metrics, we used the RMSE, coefficient of determination R^2, and correlation coefficient. The relatively high R^2 of $\sim 0.7(\rho = 0.83)$ indicates the presence of a detectable signal.

After predicting on the day-by-day level, we predicted 911 call counts for each 8-hour officer shift. These shifts take place from 7am to 3pm, 3pm to 11pm, and 11pm to 7am. Results from this level of prediction are shown in Figure 6.

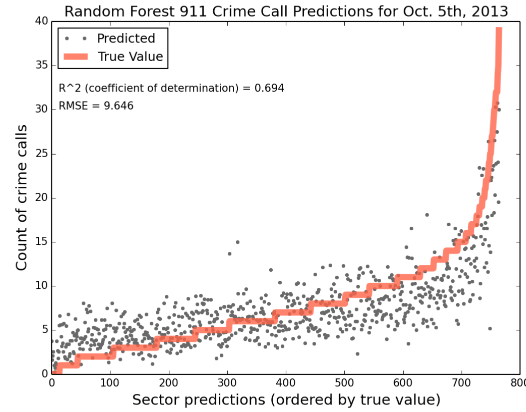

Figure 5: Random forest model predictions for the 24-hour period of October 5th, 2013. Sectors are ordered by actual target count along the x-axis. The red line indicates the target count of the sector on this date; each dot represents the corresponding model prediction for that sector.

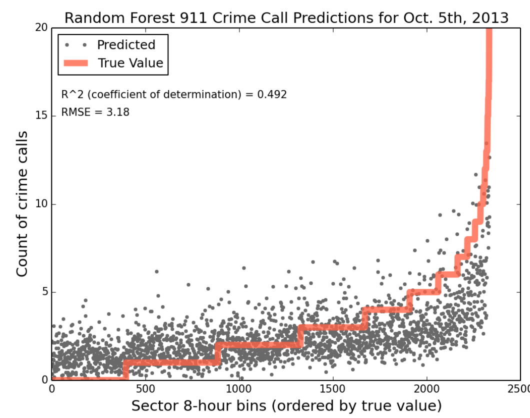

Figure 6: Random forest model predictions for the 8-hour officer shifts on October 5th, 2013. Similar to Figure 5, but predicting on a finer timescale.

The model at this level has R^2 of $\sim 0.5(\rho = 0.7)$, significantly lower than the previous example and in response to the more difficult task of predicting at a higher resolution.

Notably, the model tends to over-predict low values, and under-predict higher values. It is possible that the model cannot respond to quicker variations in local phenomena given the absence of features that respond in real-time. Many of the short-term retrospective features from 311 and 911 are sparse and likely do not provide enough resolution on the immediate history preceding a prediction. However, deeper retrospective features are stable and predictive, indicating that these descriptors may be better for indicating long-term ground level conditions.

Finally, we compared these results over time with other versions of appropriate models. First, we compare the rolling forecast model to a nearly identical "extended" forecast, where the model is only trained once on a month of data. Second, we compare the random forest (RF) rolling forecast model to the poisson rolling forecast model. R^2 over time for all three models are displayed in Figure 7. As expected, the rolling forecast RF performance behaves as an upper bound to the extended RF model. The Poisson model underperforms due to significant overdispersion that is expected at this estimation level. Ongoing work is looking at localized Poisson and Negative Binomial models that can avoid or handle overdispersion.

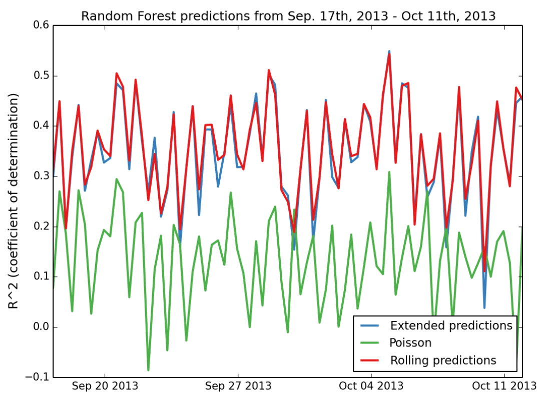

Figure 7: r^2 metrics for three different model types (random forest rolling forecast, random forest extended forecast, and poisson regression) over more than three weeks.

The top 10 predictive features for the random forest regression model were almost entirely long-term time-of-day historical aggregates on the target variable, as shown in Table 3. The model's preference for long-term features is in accordance with the autoregressive nature of the task that subsumes the majority of the signal into past counts from similar hourly period, day and season. The importance of the time-of-day specific counts highlights the natural variability of the number of 911 crime-calls across different hours of the day and their consistency when examining the same time-of-day.

Event Detection

"Anomalous" behavior is present throughout the 911 call data and is due to various events, some well-known such as natural disasters and social events, and others less obvious like socio-economic processes. Apart from their effect on any forecasting model they are also crucial for early detection and appropriate resource allocation.

One of the critical aspects of effectively allocating resources is knowing how to pinpoint regions of interest that may require a modification of police resources. Event detection, or spatial cluster detection, is a technique that identifies

Table 3: Random forest model: top features

Feature	Importance
$\bar{c}^{\,crime\ calls}_{\tau=42\ days}$, same time of day	0.063
$\bar{c}^{\,crime\ calls}_{\tau=35\ days}$, same time of day	0.056
$\bar{c}^{\,crime\ calls}_{\tau=21\ days}$, same time of day	0.046
$\bar{c}^{\,crime\ calls}_{\tau=49\ days}$, same time of day	0.044
$\bar{c}^{\,crime\ calls}_{\tau=56\ days}$, same time of day	0.041
$\tilde{c}^{\,crime\ calls}_{\tau=42\ days}$, same time of day	0.040
$\tilde{c}^{\,crime\ calls}_{\tau=56\ days}$, same time of day	0.038
$\tilde{c}^{\,crime\ calls}_{\tau=35\ days}$, same time of day	0.036
$\bar{c}^{\,crime\ calls\ in\ progress}_{\tau=56\ days}$, same time of day	0.035
$\bar{c}^{\,crime\ calls}_{\tau=28\ days}$, same time of day	0.035

spatial regions where the value of some quantity, e.g. 911 crime-call count, is significantly higher than expected, given some underlying baseline quantity. These quantities vary by application domain, but for police resource allocation count c_i represents the number of 911 calls in a given area s_i, while the baseline b_i is the historical number of expected calls for that area s_i.

Anomalous Spatial Cluster Detection

We employ Neil's Expectation-based Poisson scan statistic model (Neill 2009) whch was derived from Kulldorff's population-based Poisson scan statistic (Kulldorff and Nagarwalla 1995) and is widely used in the epidemiological community for finding significant spatial regions of disease outbreaks. The technique provides a statistically robust way of detecting spatial overdensities, or spatial regions that are likely to be significant. Neill's statistic assumes that the regions s_i are aggregated to a uniform two-dimensional, N×N grid G, and we search the set of rectangular regions S $\subseteq$ G. The counts $c_{s,t}$ are Poisson distributed with $c_{s,t} \sim \text{Poisson}(q_{s,t} b_{s,t})$, where $q_{s,t}$ is the (unknown) underlying call rate and $b_{s,t}$ represents the known historical number of expected calls for area region s_i. The method searches over a set of spatial regions, identifying regions which maximize a Poisson likelihood ratio statistic, given by

$$F(S) = \frac{P(D|H_1(S))}{P(D|H_0)},$$

where $P(D|H)$ is the Poisson likelihood given hypothesis H and the null hypothesis H_0 assumes $q_{s,t}=1$ everywhere (counts = expected). Under H_1(S), we assume that $q_{s,t}=q_{in,t}$ in S and $q_{s,t} = 1$ outside, for some $q_{in,t} > 1$ (counts > expected in S). Once we found the region $S^* = \arg\max_S F(S)$ of grid G, and its score F^* = F(S^*), we must determine the statistical significance of this region by randomization testing. This is done by randomly creating a large number R of replica grids by sampling under the null hypothesis. The highest scoring region and its score is tracked for each replica grid, and a p-value of S^* is calculated given by

$$\rho_{S^*} = \frac{R_{\text{beat}} + 1}{R + 1},$$

Algorithm 2 Expectation Based Poisson Scan Statistic

Input: $C^{911}_{s,t}$, G
for $s = 0$ **to** S **do**
  Likelihood Ratio Statistic: With grid G, calculate F(S) based on H_1(S) and H_0 for each region. **s**
  Max Region: S^* = $\arg\max_{S} F(S)$ of grid G
  Max Score: F^* = $F(S^*)$
end for
Input: S^*, F^*
for $r = 0$ **to** R **do**
  for $s = 0$ **to** S **do**
    calculate: Likelihood Ratio Statistic, Max Region and Max Score for replica region. **s**
    if: Max Score $> F^*$, R_{beat} += 1
  end for
end for
calculate: ρ_{S^*}
Output: ρ_{S^*}, S^*

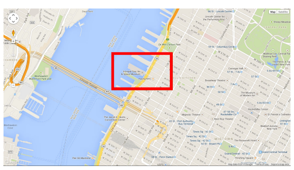

Figure 8: Most significant spatial region cluster identified on July 4th, 2012 in Manhattan

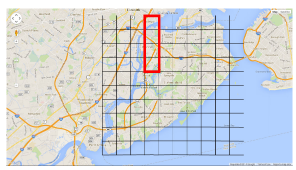

Figure 9: Most significant spatial region cluster identified (in red) on July 6th, 2012 in Staten Island

where R_{beat} is the number of replica grids with F^* higher than the original grid G.

The main disadvantage of the proposed approach is that it is computationaly expensive. It requires exhaustively searching over all rectangular regions, both for the original and the replica grids. The computational complexity for an N×N grid G with R replica grids is O(RN^4). In order to tackle this we implemented an overlap kd-tree data structure (Neill and Moore 2004) which improves efficiency with computational complexity of O($N \log N$).

Event detection of 911 calls

For validation, we first looked at historical days that were likely to be "eventful" such as major holidays or weather events. In the case of July 4th, 2012, we successfully detected the most significant cluster region in Figure 8 as adjacent to the fireworks barges along the Hudson River.

Finally, we focused our search area on Staten Island from July 4-10, 2012, and found significant clusters on the western coast of Staten Island as depicted in Fig. 9. Upon examination we discovered that this area is heavily industrial with many manufacturing buildings and the reason for this significant clustering of 911 calls is due to unexpected commercial alarms that led to a series of "unfounded" 911 calls.

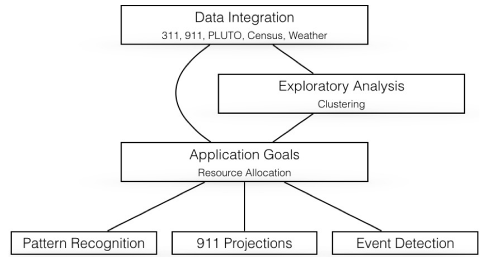

Figure 10: Overview of resource allocation framework, including components described in this paper

Discussion and Conclusion

Additional steps are underway to improve the performance of the forecasting model. Apart from handling overdispersion, ongoing work is focusing on i) extracting better predictive features from 311, PLUTO, and weather data sets, ii) fitting localized models for each sector that are later combined (Fink, Damoulas, and Dave 2013), iii) the inclusion of longer-term historical features and shorter-term "dynamic" features, and iv) more training data to capture seasonal effects. As part of this effort, we are compiling other data sources, including social network data, that could explain a larger percentage of the variation — especially in extreme cases where behavior is unusual compared to the past.

Event detection will also be improved in two ways. First, implementing a Bayesian scan statistic (Neill, Moore, and Cooper 2006) rather than derivations of the Kuldorff statistic, would allow us to monitor several event types simultaneously in a probabilistic framework and gain computational efficiency by avoiding randomization testing. Moreover, this would enable us to apply the statistic at a higher resolution, providing more information on anomalies in fine-grained spatial regions. Second, the predictive model's forecast of counts will also be integrated into the scan statistic as a more developed baseline for the likelihood ratio statistic.

This work represents the first steps in building a comprehensive system to effectively allocate police resources, using all three techniques discussed in this paper (Figure 10). In this framework, there are both immediate practical uses for the outputs of our model, as well as additional long-term goals that will be aided by these techniques. The outputs from this work can be used immediately by precinct commanders when allocating resources. However, they will also feed into an overarching resource allocation scheme as critical inputs for understanding and forecasting demand. Given the size of New York City's police force — the biggest in the United States — the opportunities for improving resource allocation are significant and potentially very impactful to human quality of life in cities.

References

Berestycki, H., and Nadal, J.-P. 2010. Self-organised critical hot spots of criminal activity. *European Journal of Applied Mathematics* 21(4-5):371–399.

Breiman, L. 2001. Random forests. *Machine learning* 45(1):5–32.

Fink, D.; Damoulas, T.; and Dave, J. 2013. Adaptive spatio-temporal exploratory models: Hemisphere-wide species distributions from massively crowdsourced ebird data. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, July 14-18, 2013, Bellevue, Washington, USA.*

Kulldorff, M., and Nagarwalla, N. 1995. Spatial disease clusters: detection and inference. *Statistics in Medicine* 14:799–810.

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

Neill, D. B., and Gorr, W. L. 2007. Detecting and preventing emerging epidemics of crime. *Advances in Disease Surveillance* 4(13).

Neill, D. B., and Moore, A. W. 2004. Rapid detection of significant spatial clusters. In *Proceedings of the 10th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 256–265.

Neill, D. B.; Moore, A. W.; and Cooper, G. F. 2006. A bayesian spatial scan statistic. *Advances in Neural Information Processing Systems* 18:1003–1010.

Neill, D. B. 2009. Expectation-based scan statistics for monitoring spatial time series data. *International Journal of Forecasting* 25:498–517.

Perry, W. L. 2013. *Predictive policing: The role of crime forecasting in law enforcement operations*. Rand Corporation.

Short, M. B.; D'Orsogna, M. R.; Pasour, V. B.; Tita, G. E.; Brantingham, P. J.; Bertozzi, A. L.; and Chayes, L. B. 2008. A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences* 18(supp01):1249–1267.

Short, M. B.; Brantingham, P. J.; Bertozzi, A. L.; and Tita, G. E. 2010. Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences* 107(9):3961–3965.

Woodworth, J.; Mohler, G.; Bertozzi, A.; and Brantingham, P. 2013. Nonlocal crime density estimation incorporating housing information. Manuscript submitted to journal.

Zipkin, J. R.; Short, M. B.; and Bertozzi, A. L. 2013. Cops on the dots in a mathematical model of urban crime and police response. Manuscript submitted to AIMS? Journals.