Blue Light Camera Impact Analysis Report

Ning Zhang
University of Rochester
nzhang15@ur.rochester.edu

INTRODUCTION

A Blue Light camera is a surveillance camera with a lamp radiating blue light installed on its top. The camera is usually put in an open area such as road intersection. In addition to monitor crime activities within its viewshed, it is also expected to have an effect in reducing crimes with the help of its symbolic blue light, which is recognizable even from a distance of hundreds of meters.

By the end of 2008, Rochester Police Department purchased 50 cameras and strategically installed them in areas with high crime density based on the past crime records by the end of 2008. Ever since then, more Blue Light cameras were brought in and installed around the city at various points of time. After the latest batch of cameras were installed in July 2016, the total number has increased up to 136.

Since the installation of the first Blue Light camera in 2008, the RPD has not yet conducted effectiveness analysis on those cameras. The objective of this project is to help the RPD to quantitatively evaluate how effective these cameras are in reducing crimes in their vicinity.

Under the request of the RPD, this project put its main focus on felony activities instead of all crimes. Therefore all hypothesis tests in the analyss process are implemented on the dataset of felony activities.

DATA PREPARATION

The data used in this project consist of three parts. The **first part** is records of crime activities that were reported during 12 years from 2005 up to 2016, which was provided by Rochester Police Department (RPD). The original dataset contains 12 RDATA files and a total of 363,215 observations or crime reports. Each RDATA file consists of one yearâĂŹs crime reports. The dataset is cleaned and preprocessed through the following main steps.

- Due to the fact that the RPD switched its crime records systems from PACER to LERMS in 2013 and only brought the data in the five years prior to 2013 into the new systems, many discrepancies in terms of recording methodology exist between the data recorded before 2009 and after those ever since. The issue is solved by filtering out redundant or irrelevant columns to this project in both parts of the dataset. After being merged, the dataset has a total of 21 attributes.

- There are 63 crime records whose report years are prior to 2005. All of these records are removed.

- Geospatial information of all crime records from 2005 to 2008, which is logged in a GIS projection format, are converted to longitude-latitude format. In the process a total of 6456 crime records are found logged with incorrect or null geographical coordinates, and they are all removed.

- Values in the attribute âĂIJCrimeTypeâĂİ are further cleaned. Redundant or small subcategories of crimes are merged.

The final all-crime data contains a total of 356,696 crime records covering all crime types. In addition to the full crime data, a subset that only includes felony crime activities is also prepared. To be more specific about felony crime, the subset contains all crime records with values in attribute *CrimeType* being *aggravated assault, arson, burglary, dangerous weapons, homicide, kidnapping, manslaughter, murder, mv theft, rape, robbery, stolen property* and *non-petit larceny*.

The **second part** is geospatial information of 136 Blue Light Cameras, which was downloaded from RPD Open Data platform. The main issue about this part is that the installation time of each Blue Light camera was not well recorded. To our best knowledge, there are two groups of cameras which have relatively accurate installation times on record. The first group is the batch of 56 cameras that were installed in 2008 and early 2009, and the other consists of 19 cameras that were installed in July 2016. Although it was reported that there were 27 cameras installed in 2016, only 19 were identified in the dataset downloaded from the open data platform.

The **last part** of the data is geospatial data of the City of Rochester, which was downloaded from OpenStreetMap under the Open Data Commons Open Database License (ODbL). Geographical coordinates of 4,255 intersections to the south of Holy Sepulchre Cemetery in Rochester are extracted from the data. A further filtering step is implemented on the intersection data and picks out a total of 2,358 intersections which has a distance of more than 300 meters away from any of the 136 Blue Light cameras installed in the city. The resulted dataset is used as control group against the camera group in following analysis processes.

All processes mentioned above are reproducible and well recorded in the processing codes that are included in this project′s attachment. To recap, the processed datasets are listed as follows.

1. all-crime dataset of 356,696 rows.

2. felony dataset of 91,837 rows.

3. batch 2008: the batch of cameras that were installed by the end of 2008 or early 2009
4. batch 2016: the batch of cameras that were installed around July 2016.
5. control group: 2,358 road intersections that are at least 300 meters far away from any of installed Blue Light cameras.

EXPLORATORY ANALYSIS

The goal of this part is to explore interactions between crime distribution and Blue Light cameras as well as general patterns hidden in the crime records.

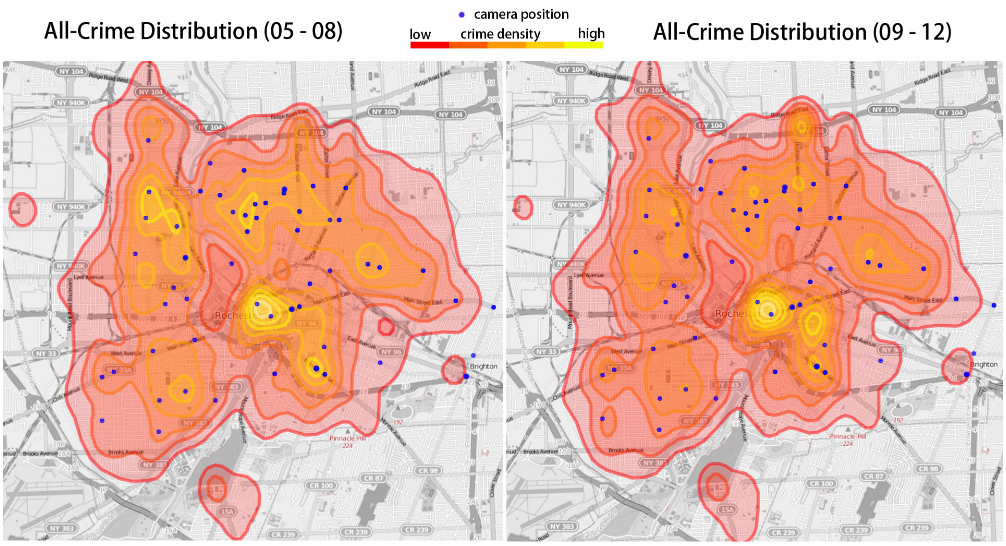

Figure 1: Comparison Between All-Crime Distributions

In order to have an idea of crime distributions in the city, several contour maps are drawn as shown in **Figure 1** and **2**. In **Figure 2**, the left contour plot demonstrates the density distribution of all-crime activities from 2005 to 2008, which in this project is treated as pre-blue-light period, while the right plot shows the distribution in the period from 2009 to 2012, during which the batch 2008 cameras had taken effect for a while. The comparison of the two plots reveals information. On one hand, the center of city (yellow area in the center of the plots) consistently remains a hot bed of crime activities throughout the 8 years. On the other hand, some highlighted area in the left plot is no longer so in the right plot. Taking the city section of Clinton for example, which is located to the north of the city center, it was clearly one of the areas with highest crime density in the left plot, but its crime density reduced in the right plot. Meanwhile, a new hightlighted area shows up in the north of Clinton, which is the location of a Walmart supermarket frequented by crime of larceny.

Since felony activities are the main focus of this project, we also draw similar contour plots demonstrating felony distributions in the two four-year periods as shown in **Figure 2**. By observation, however, **Figure 2** does not show any significant density shifting in the case of felony actitivities. Areas with high crime density remains, if not deteriorates to, the crime-highlighted areas during 09-12 period.

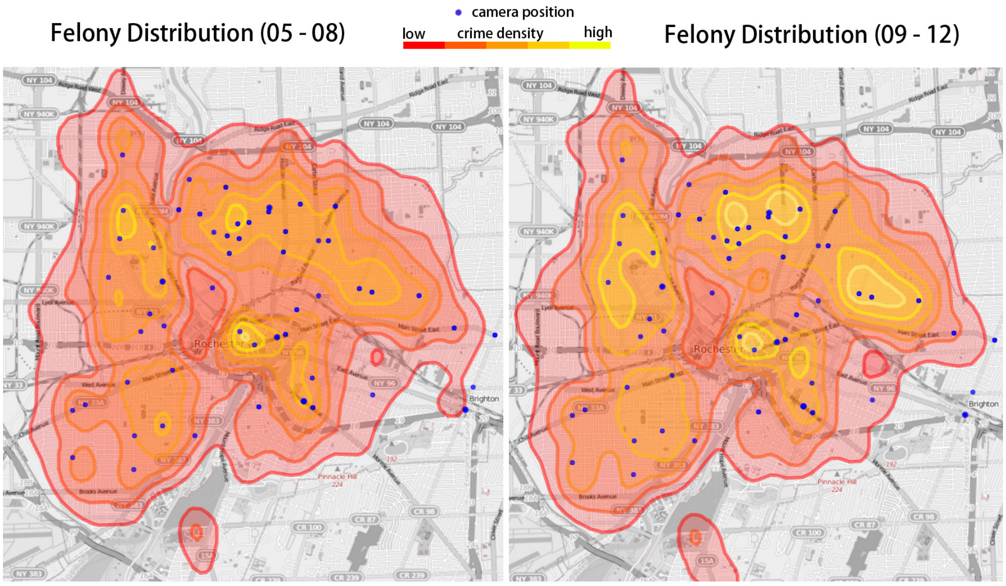

Figure 2: Comparison Between All-Crime Distributions

While contour maps perform well in revealing where hot areas of crime are located, they do not tell much about quantitative variation of crime counts in the dimension of time. Therefore we also made time-series plots, as shown in **Figure 3**, using all-crime and felony datasets respectively.

When crime records are summarized on monthly base, a strong patten of seasonality is revealed. The frequency of crime activities, no matter it is for all-crime or only felony, tends to peak in summer time, and then dips down to the bottom during December and January. More interesting than the pattern of seasonality is a decreasing trend in both all-crime and felony activities. The patterns can be observed in **Figure 3**.

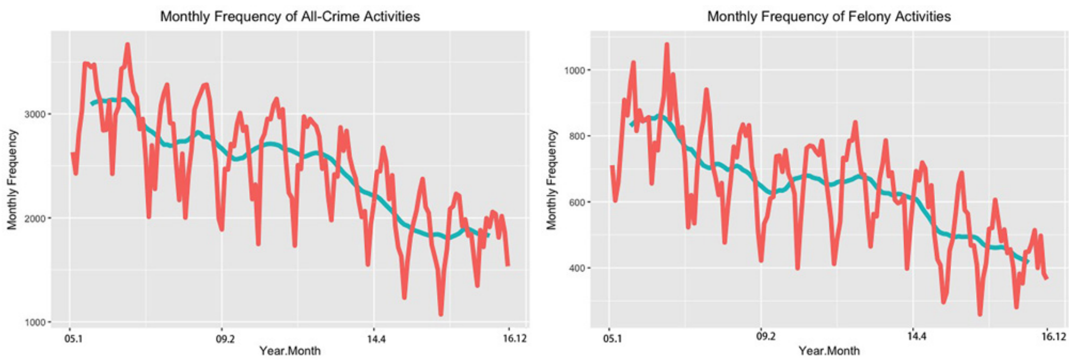

Figure 3: Monthly Crim Count Time-Series Plot; **Left**: All-Crime; **Right**: Felony.

Methods

Crime Count. In order to quantify a camera′s impact on its nearby crime activities we first developed a measuring metric *crime count*. *Crime count* is the sum of reports of crime activities that happened within the distance of r away from a location l during a period of time t, where r is the radius of a circle centered at l, which stands for the geographical coordinates of a camera or control intersection, and t is a time interval, which could be one month, three months, etc. In this project r is fixed at 50 meters, while t is various across different hypothesis tests.

Control Group. As it can be seen in figures, crime activities were decreasing across the city. Empirically speaking, many factors such as economy, demography, government policies and even some unconceivable elements can be attributed to the decreasing trend. With respect to this project, we must take into account the confounding effects of these factors on crime. To overcome this challenge the control group which consists of 2,358 intersections are brought in and the crime activities near control group are used as a benchmark in comparison with those near cameras. As mentioned earlier, a 300-meter threshold is set in the process of choosing control intersection. It is so done in order to reduce the probability that members of the control group are covered in the buffer zone of any cameraâĂŹs impact area.

Hypothesis tests. If a Blue Light camera has solid impact on nearby crimes, then a decreasing pattern of crime is supposed to be observable along a certain period of time. Otherwise, the camera′s impact on crime should be considered as obscure. Based on this assumption we designed a one-sided hypothesis in the following expression, where $prior\ crime\ count$ stands for a sequence of counts of crimes recorded prior to the stallation of a batch of cameras, and $post\ crime\ count$ for the post-installtion data.

$$H_o : pre-installation\ crime\ count <= post-installation\ crime\ count$$

$$H_a : pre-installation\ crime\ count > post-installation\ crime\ count$$

Results

In this section we conducted a series of one-sided paired t-tests and Wilcoxon rank sum tests to assess the null hypothesis. In t-tests, for each camera in each batch we computed its felony crime count within time interval t starting at $T_{install}$, the time point of camera installation, and paired it with the count whithin the same period one year earlier. Here t was set at three-month, six-month, and one-year respectively, and $T_{install}$ was set at a time point upon which most of cameras in the batch had been installed. Taking the 2008 batch for example, since most of the cameras had been installed by the end of 2008, $T_{install}$ was set at January 1st, 2009. The same crime count computation was also conducted on each intersetion in the control group. Then we implemented one-sided t-tests on the paired felony crime counts.

Table 1 and **2** enumerate the test results for batch 2008 and 2016 respectively. The paired felony crime counts of batch 2008 cameras, no matter in what time interval they were computed, fail to reject the null hypothesis that crime count in time interval t is less than or equal to the value one year earlier. In another word, batch 2008 cameras do not see statistically significant decrease in felony activities during a time interval t following their installation, where t is three-month, six-month, or one-year. We ended up with the same result in tests on batch 2016 cameras as shown in **Table 2**.

Table 1: One-Sided Paired Difference T-Test for Batch 2008 (**df = 55**)

Camera Batch	Time Interval	P-Value	95% C.I.
batch 2008	three-month	0.383	$[-0.246, \infty)$
control group	three-month	0.057	$[-0.0009, \infty]$
batch 2008	six-month	0.288	$[-0.354, \infty)$
control group	six-month	0.001	$[0.031, \infty)$
batch 2008	one-year	0.208	$[-0.286, 1.286]$
control group	one-year	0.560	$[-0.051, 0.094]$

Table 2: One-Sided Paired Difference T-Test for Batch 2016 (**df = 18**)

Camera Batch	Time Interval	P-Value	95% C.I.
batch 2016	three-month	0.250	$[-1.276, \infty)$
control group	three-month	3.176e-05	$[0.033, \infty)$
batch 2016	six-month	0.167	$[-1.298, \infty)$
control group	six-month	3.427e-08	$[0.071, \infty)$

In Wilcoxon tests, we narrowed the time interval t down to one month. We first computed monthly felony crime counts from 2005 to 2016 for each camera in batch 2008. Then we conducted a sequence of 8 Wilcoxon tests for each camera. More specifically, when selecting $pre-installation\ crime\ count$ for each test in a sequence we always included the whole dataset of pre-installation felony crime counts recorded from 2005 to 2008; On the other hand, we accumulatively added one year's post-installation felony crime counts into $post-installation\ crime\ count$. For example, in the first Wilcoxon test in a sequence, its $post-installation\ crime\ count$ only included monhtly crime counts in 2009, the second test contained the data in 2009 and 2010, and each following test had one more year's post-installation crime counts than the previous one.

With data prepared, we conducted the sequence of Wilcoxon rank sum tests for each camera in order to test the null hypothesis that pre-2009 felony crime counts is greater than the post-2009 data. The same process of compution was also implemented on the control group. We chose Wilcoxon tests over t-tests in this process because Wilcoxon tests loose the constaint of normality assumption on the monthly crime counts, which are apparantly time-series.

Wilcoxon Rank Sum Tests		
Group	**Accumulative Years**	**Percent Of Members**
batch 2008	one	3.6%
control group	one	0.5%
batch 2008	two	7.1%
control group	two	2.0%
batch 2008	three	16.1%
control group	three	3.9%
batch 2008	four	19.6%
control group	four	4.5%
batch 2008	five	19.6%
control group	five	4.5%
batch 2008	six	21.4%
control group	six	7.7%
batch 2008	seven	26.8%
control group	seven	8.5%
batch 2008	eight	28.6%
control group	eight	10.6%

As a result, the sequence of tests show a trend that more members in batch 2008 rejected the null hypothesis tests as more years of post-installation monthly felony crime counts were added into $post\ crime\ count$. While control group showed a similar pattern in the tests, batch 2008 had faster growth in percentage (see **Figure 4**).

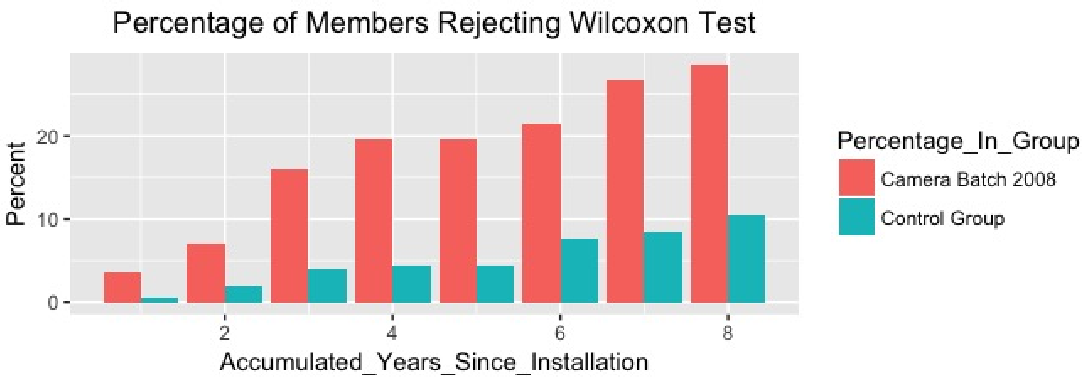

Figure 4: Results of 8-Wilcoxon-Test Sequence for Each Group

Conclusion and future work

From 2005 to 2016, the City of Rochester saw a steady decreasing trend in crime. For all-crime activities, a peak-shifting pattern in all-crime density distribution was observed since the installation of the first batch of Blue Light cameras, while similar pattern did not show up in the case of felony activities. By conducting a series of hypothesis tests on felony records, we found that the crime count in a Blue Light camera's viewshed within the time interval following its installation was statistically significant different from the value in the same period one year earlier. In constrast, when it came to long term patterns, more Blue Light cameras had seen statistically significant deceasing trend in felony activities within their viewshed. In another word, it takes time for a Blue Light camera to have impact on felony activities.

For future work, it may be interesting to drill down into the level of individual cameras to explore why some cameras have impact when others do not. Another question that has not been answered in this project is searching for appropriate locations for five next new Blue Light cameras. One proposal for this question is taking advantage of the data of city intersections and implementing Markov Chain Monte Carlo methods to simulate top 5 candidate locations.

References

References

[1] Rochester Police Department Open Data Portal, http://www.data-rpdny.opendata.arcgis.com

[2] Introduction to Blue Light Camera System, http://www.cityofrochester.gov/article.aspx?id=8589936528

[3] Open Street Map, https://www.openstreetmap.org