Research Science Press

International Journal of Data Analysis and Information Systems
VOLUME 8 • NUMBER 1 • JUNE 2016

journal homepage : serialsjournals.com

ISSN : 2229-5887

Crime Analytics Across Neighborhoods of Chicago Using Spatiotemporal Correlation

Prithvi Yarlagadda*, Madhumita Duvvuri*, Natasha Kapoor*, Gowtham Srungarapu*, Hakan Gogtas, Edward Moskal* and Sylvain Jaume*****

**Data Science Graduate Program, Saint Peter's University, Jersey City, NJ 07306*
***American Express, New York, NY 10285*

Authors emails
sjaume@saintpeters.edu

ABSTRACT

The availability of detailed crime records over long period of time offers new applications for data mining and the emergence of crime analytics. Are crimes happening in one neighborhood related to crimes happening later in another neighborhood? By applying time series and correlation analyses on crime records collected from 2001 to 2015, we investigate the existence of this relation across the City of Chicago and at different time scales. Our results indicate that the most significant relation exists at the quarterly and bi-monthly scale.

Keywords: *Crime analytics, Spatiotemporal modeling, Correlation, Time series, Geospatial analysis, Big Data, Data visualization, Smart cities, Computational social science*

INTRODUCTION

Mining crime data is a new field that is of high value for law-enforcement and intelligence-gathering organizations. In an effort to accurately and efficiently analyze the growing volumes of crime data, Chen and coworkers (Chen, Chung, Qin, Chau, Xu, Wang, Zheng, & Atabakhsh, 2003) reviewed crime data mining techniques and presented four case studies done with Phoenix police departments. Nath (Nath, 2006) focused on k-means clustering with some enhancements to identify crime patterns. Estivill-Castro and Lee (Estivill-Castro, & Lee, 2001) incorporated clustering and association-rule mining into an exploratory tool for the discovery of spatiotemporal patterns. They have also developed an application to visualize cluster boundaries and mining association rules. Keyvanpour *et al.* (Keyvanpour, Javideh, & Ebrahimi, 2011) proposed to extract important entities from police narrative reports written in plain text and entered into a database in law enforcement agencies. They applied self-organizing maps clustering and plan to use the clustering results to perform a crime matching process. Yu and coworkers (Yu, Ward, Morabito, & Ding, 2011) built a model that takes advantage of implicit and explicit spatial and temporal data to make reliable crime predictions.

Crime analytics helps police departments in finding patterns and relationships between the crimes. It provides us with who, what, when and where about crimes. With this information we can build better strategies to prevent or reduce crime from happening. Crime analytics' main focus is to find useful patterns and relationships from previously committed crimes and data sets. This in turn helps law enforcement to establish crime prevention strategies, leading to the improvement of overall safety and quality of life in a state or country. Advances in research and development in this area are producing new software that can help in understanding, reporting and analyzing crime patterns. Many discoveries in relationships between locations and crimes are emerging from these technologies using spatial location data.

Crime data can be classified into point data and areal data based on the spatial locations. Point data contains information about where the crime happened with locations at maximum accuracy, usually in latitudes and longitudes. Areal data contains information about crime locations within predefined boundaries. Point data can be used as areal data in some situations, as location points can easily be identified within area to which they belong (e.g. state, county or district). For visualizing different crime patterns there are numerous tools that allow us to plot time-series graphs, bar graphs, choropleth maps or scatter plots.

Crime analytics is emerging worldwide in law enforcement and intelligence agencies. There is great potential for mining these data sets for predicting crime trends and patterns and it may be possible to transfer the lessons learned in one area to crime prevention to another. On the other hand, many crime patterns are likely to be local in nature. It is important that crime analysts adopt a set of definitions and terms related to crime analytics that constitute the best practice in the profession and permit global data sharing. Among the many existing definitions given by practitioners in the field, the most fitting way to describe crime analytics is in the definition given by the U.S. Department of Justice i.e. "the qualitative and quantitative study of crime and law enforcement information in combination with socio-demographic and spatial factors to apprehend suspects, prevent crime, reduce disorder, and evaluate organizational procedures."

The analysis can be made at different levels, including tactical, operational, and strategic. Analysts observe crime reports, arrests information as well as detective reports to identify various patterns, series or trends as rapidly as possible. They analyze all significant aspects, occasionally make predictions or forecast future occurrence rates, and issue reports, bulletins and real time alerts to law enforcement agencies.

Jayaweera and coworkers (Jayaweera, Sajeewa, Liyanage, Wijewardane, Perera, & Wijayasiri, 2015) analyzed the crimes described in newspaper articles and observed that geographical diversity and complexity of crime patterns have made the recording and analyzing of crime data more difficult. They proposed a web-based system comprised of crime analysis techniques such as hotspot detection, crime comparison and crime pattern visualization.

The number of industry domains concerned with crime analytics has also increased significantly. Swanson (Swanson, 2015) explored crime analytics in various industry domains from customer insight, healthcare, insurance and finance through the use of briefing charts. The author showed that crime analytics could also impact manufacturing, education, telecommunications and cognitive services.

Many authors have studied crime patterns and attempted to predict crimes based on data from different heterogeneous sources. Mookiah and coworkers (Mookiah, Eberle, & Siraj, 2015) surveyed the research in crime analytics and demonstrated that in some cases, information commonly accepted as influencing the crime rate has little or no effect.

Recently, modeling techniques such as the Risk Terrain Modeling technique (Kocher, & Leitner, 2015) were proposed for localities where the probability is high that a crime event will take place. Contrary to other techniques, Risk Terrain Modeling does not focus on previous events, but on risk factors that have an influence on the environment. As a test on the city of Salzburg, Austria, Kocher and coworkers predicted assaults, auto thefts, burglaries, and robberies for 2013 and 2014. Out of 27 predictions, the best predictive accuracy was reached for assaults and for robberies. The authors concluded that the availability and quality of the risk factor data are critical for prediction accuracy.

Research has also been conducted on the visual analytics of heterogeneous data. Hatton and coworkers (Hatton, Zhao, Gorantla, Chae, Ahlbrand, Xu, Ko, 2015) analyzed the heterogeneous data provided by the 2015 VAST Challenge and developed a visual analytics system that combined analytic models and visualization techniques. However, in an experiment to process large volumes of data in real time, the proposed data models and clustering techniques suffered some limitations and scalability remains a challenge for real time interactivity and analysis.

Crime analytics can also benefit from the insights gained in other fields. For example, in preventive detection of product defects, Schuh and coworkers (Schuh, Rudolf, Doelle, & Riesener, 2015) identified hot spots within the deviation probability map to anticipate deviations from the target process and thus inefficiencies within development projects. This approach is comparable to the anticipation of crimes in that deviations in time, costs and quality can be considered a result of waste and therefore a dimension for inefficiencies or defects. In this analogy, the deviation probability map was a part of the superordinate methodology to intuitively identify hot spots and enhance preventive control.

Intelligence and analytic technologies can provide opportunities for police department to conduct information gathering and risk analysis. Sanders and coworkers (Sanders, Weston, & Schott, 2015) used 86 interviews and observations to explore how intelligence-led policing is used and integrated in the Canadian police. The authors attempted to identify the police cultures, the type of organization and the pace of policing that constrain their intelligence gathering and their definition of success.

The crime rate in the United States varies depending on the type of community. For example, the crime rate in metropolitan areas tends to be above the national average, including higher crime rates in cities such New York, Los Angeles, Chicago, and Houston. In particular, Chicago has one of the highest records for crimes committed and the tops the list of the cities where the chance of being a victim is 41 out of 1000, far ahead of other major cities. The Chicago crime data are analyzed for the purpose of this project.

Crimes are classified into different types varying from the place it occurred to the type of crime. This paper focuses on the correlation of crimes committed in the neighborhoods of Chicago and time series analysis on a monthly or quarterly basis. To understand and analyze the crime data, a valuable approach is strategic crime analysis, which is directed towards the development and evaluation of long-term strategies, policies and prevention methods. Strategic crime analysis includes long-term statistical trends, hot-spot analysis and also problem analysis. The data is often collected from the police record systems, which includes the primary data comes from both quantitative and qualitative methods. Qualitative analytical methods are generally performed on non-numeric text data, while quantitative modeling performed on numeric data can be used to interpret the location, time and frequency of the occurrence, and help deciphering the incidence that those values depict. In this study we concentrate on finding the correlation between the various districts and the crime hot spots in the city of Chicago using spatiotemporal data.

To visualize the correlation patterns of crimes in Chicago, we used three techniques. i) heat maps based on the number of crimes per district, ii) time series of the number of crimes per day throughout the year and iii) correlation heat maps, where the crime correlation between districts is plotted as shown below. We also need to recognize the importance of different factors like type of crimes, general location of crime (street, shops etc.) and the type of neighborhood for a better assessment of the underlying relationships between crime activities. This multi-dimensional approach may also include factors such as population density and socioeconomic background. In this paper, we used the three-step technique described above to analyze theft. Theft had the highest incidence among all types of crimes reported and provided us with a starting point for our analysis.

Identifying important factors and using them to understand, build and analyze patterns of crime activity should provide us with knowledge of how space and time influence crime and help develop theories and methodologies for better crime prevention programs. In this paper, we focused on correlations between crimes and districts and performed experiments to verify and measure the relationships between crimes and neighborhoods at different times in the city of Chicago. The Data section describes the data sets that we used for this study, while the Method section describes the correlation analysis method that we developed. We conclude with a discussion of some emerging trends in crime analytics including Big Data and how the results we obtained might be useful for resource allocation.

DATA

Collection of data and dissemination is an on-going and recurring process. A database can refer to anything from spreadsheets having a handful of cases and few variables to entire records management systems. These databases play a major role in crime analytics and predicting patterns and trends and it is important to establish the actual and potential data sources within the law enforcement agencies that can be used for analysis. The law enforcement information falls into two distinct categories, viz. tabular data and geographic data. Tabular data refers to a list of records in a table. For law enforcement, this can include information about the record, place of the crime (including addresses) or other types of geographic variables, calls for service, field information, suspects' information and arrests. Geographic data refers to data that is inherently geographic, i.e. they give a description of the geographic features. Examples of geographic data include streets, districts, census tracks, bus stop or route data and jurisdictions.

In law agencies, officers, clerks and dispatchers perform the data entry tasks. They often do not know how others will use the data they enter into the database. This can result in untrustworthy data. Each agency has a particular format for recording a crime and a standard procedure to summarize how crime analysis is performed. These procedures are important for an analyst to understand the data source and model trends based on the crime reported.

For the purpose of analysis in this paper we identified each state's available data sources and searched for crime data, which is open source. After thoroughly exploring data sources from each state, we concluded that Chicago's crime data is highly structured and available in different formats like Excel, CSV, JSON, Xml etc. The city also provided APIs to input the data into any application.

Chicago crime data consists of categorical and quantitative data with numerous column entries that help in describing the crime, including time of occurrence, county, place, and even latitude and longitude. Other variables include the scene or area where the crime occurred, population density of the district, frequency of crime, economic factors and demographics (e.g. age, ethnicity, and gender). The rich data source provided us with many opportunities for crime analytics and we analyzed Chicago to find where the crime rates were highest.

It was essential to consider the time and date when the crime occurred so that we could determine the frequency of the crime and whether there was a pattern that predicts when the crime is likely to re-occur. The area or location can be used to determine whether the crime rate is high or low in a particular district, while the time of day when a crime can be committed also plays a role. For example, an empty alley can be a perfect place for robbery at any point of time during the day. We explored these factors to predict crime patterns and trends and used correlation analysis to determine the major crime hot spots in Chicago.

For correlating crimes between the districts of Chicago we calculated the total number of crimes committed in every district on a daily basis and added these to the data set. When no crimes were reported, 0's were added to these dates to ensure that there were no time lags between districts with unreported crimes and those where crimes had been reported. After eliminating any time lags we were able to calculate the correlations easily and accurately.

After categorizing the type of crimes committed, theft was determined as the most common crime in the districts of Chicago. Analysis in this paper is based on calculating the correlation of the various districts for this highly committed crime. To achieve this, we needed to filter the records to obtain only theft-based crimes prior to calculating the correlation between districts based on crime numbers.

The data used in this paper was extracted from City of Chicago data portal. The data that is publicly available in the Chicago data portal, in turn extracts the data from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. Most of the crime data is kept confidential and is not for public use. This site provides the temporal, spatial and crime data required to identify the crime hot spots in Chicago for various crime types. This dataset reflects reported crime incidents (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present. Each recorded crime incident contains information on case ID, case number, date/time, type, location, description, addresses (shown at block level only), district and geo-coordinates. It was observed that over 5 million crimes were reported from 2001 through 2015 in 22 different districts of Chicago.

Analysis the crime data was categorized into three types:

- Spatiotemporal data (crime locations and type of occurrence)
- Crime natural specifications (crime scene description, offender's behavior)
- Offender profile (age, sex, ethnicity)

Our paper uses the spatiotemporal data to investigate whether crimes happening in one neighborhood are related to crimes happening in other neighborhoods.

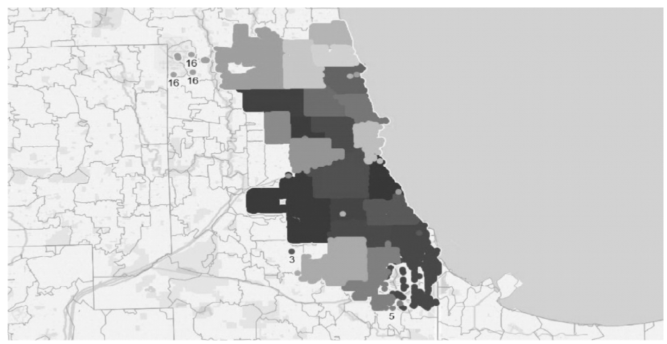

Figure 1: Crime events in the city of Chicago from 2001 to 2015. Geo-coded events reveal features of the city, such as street, alleys and grocery food stores

Figure 1 indicates a higher crime rate in the southern part of Chicago. Plotting geographical overlay of the city makes it possible to analyze the crime data at district or neighborhood level.

METHOD

We developed our analysis using the Python programming language and the Numpy numerical library. We focused our analysis on the distribution of crimes across Chicago, the identification of districts with the highest crime rates and the relationships between the data. Plotting the heat map of crimes on the Chicago map based on number of crimes in that district did not reveal any clear relationships. We than analyzed how crimes were distributed throughout the year. We observed that on Christmas Eve, the rate of crimes in every district was low compared to all other days in the year. This seemed reasonable as there might be less police working on a public holiday or that people were engaged in celebrations. A second observation was that the rate of crimes in winter was lower than in spring or summer, although the difference was not large. This was also observed in all districts of Chicago.

The temporal and spatial characteristics available in the crime data make it easy to show different patterns in crimes. With time and location details in each crime report known, the crime rate on a daily, weekly, or monthly scale can be measured for each district within city of Chicago. A time series for the number of crimes in the city for 2001 displays seasonal trends on a monthly scale. Figure 2 shows seasonal trends that indicate peaks in crime at the beginning of summer and around the midpoint of fall. It is important to note that there is a decrease around the holiday season time.

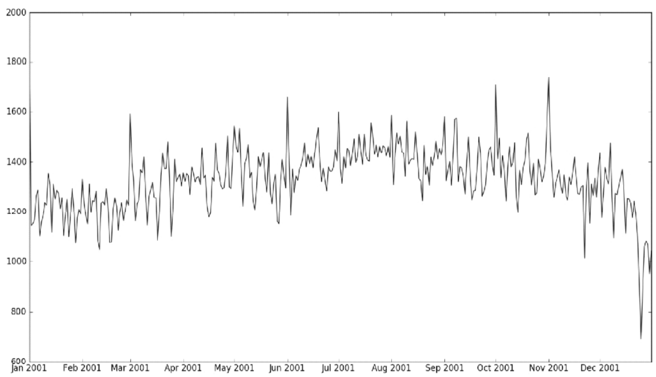

Figure 2: The figure represents a monthly times series plot against number of crimes in the city of Chicago in 2001.

There are around 6 million crimes recorded in Chicago from 2001 to 2015 and among these there are 33 different types of crimes involved. Due to complexity in calculating multi-dimensional correlations between the number of crimes, types of crimes and districts of Chicago, we decided to consider the crime of theft and calculate the correlation for districts and number of crimes.

We used a pie chart to identify the types of crimes and their frequencies. Many of the types of crime, such as like obscenity and intimidation are rare, and make a negligible contribution to the pie chart. These crimes were grouped together so that 33 types of crimes were minimized to 13 types and plotted in a pie chart. The rare crimes among all the types of crimes have the smallest share of 7.13%. The share of drug-related crimes such narcotics possession and sale was 11.26%. Sex offenses like public indecency and prostitution were 11.4% of total crimes. Vandalism type crimes, including arson and criminal damage were 18.03%. Violence-related crimes involving assault, homicide and weapons was 24.8%. The major share of crimes was theft relate, such as robbery, burglary and auto theft, accounting for 35.04% of all crimes. When analyzed as individual crime types, theft held the highest share with 20.69% followed by battery with 18.23% (Figure 3 and Table 1). Theft holding the highest shares of crimes in Chicago, in our paper we use theft rates for calculating correlations between districts.

We also tested the relationship between theft committed in the city over a period of years with the various neighborhoods. Figure 3 shows that theft has the highest count and is the most commonly committed crime in the city through the years 2001 to 2015. To investigate the existence of a relation between the crimes taking place in different districts, we used the time series analysis and a correlation matrix analysis to obtain the most significantly correlated districts. Our data was converted into matrix form, where each row represents the district and each value in each row represents the total number of crimes for give period of time. For example if we are calculating a monthly correlation for each row (district), the values will be the number of theft-based crimes recorded for each month from 2001 to 2015. As described above, when crimes were not reported in one district but were reported in another we sometimes had a time lag between the districts. To resolve this issue, we marked down all the missing dates and added them into the dataset with the value 0 for each. We built these matrices for calculating daily, weekly, biweekly, monthly, bimonthly, quarterly, half-yearly and yearly correlations and used Python to calculate the correlation matrix of each of the data matrices as a heat map, so that the higher correlation between districts the darker the color in the map (Figure 4). Each district's correlation with itself will have the highest correlation 1, so the diagonal will be the darkest.

We performed a correlation matrix analysis to investigate the dependence of multiple variables at the same time. Following the notation of Toole et al (Toole, Eagle, & Plotkin, 2011), we built a correlation matrix Y with dimensions KxT, where K is the number of time series taken from the location and T is the length of each time series. Since each time series corresponds to

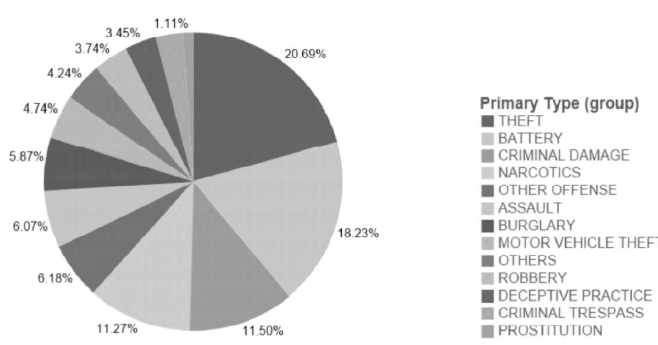

Figure 3. Percentage breakdown of crime types. The highest percentage of crimes in Chicago is Theft followed by Battery

a location in the matrix, we are able to associate locations in Chicago with correlations. The result for this analysis is a table that has the correlation coefficients of each variable with the other variables. The higher the value of the correlation coefficient, the stronger the significance of the relationship between the districts will be. Using the time series and the crime rate, we can find if the correlation exists between the different neighborhoods for theft-based crimes on multiple time scales.

Table 1: Categorical Groupings of Different Crime Types from 2001-2015

Category	*Offenses Included*	*Crimes (%)*
All	all reported offenses	5973040 (100%)
Theft	burglary, robbery, auto	2092785 (35.04%)
Violent	assault, homicide, gun	1481233 (24.8%)
Vandalism	deceptive practice, arson, criminal damage	1077077 (18.03%)
Sex Offenses	public indecency, prostitution	683602 (11.4%)
Drug	narcotics, possession, sale	672899 (11.26%)
Other Offenses	obscenity, intimidation, involving children	425967 (7.13%)

RESULTS

After the required fields and data to find the significance between the districts were established, we sought to analyze the data. We calculated correlations based on time-series with each district of Chicago. We created an MxN matrix where M is the number of time series and N is the number of districts. In certain instances, no crimes were reported for a particular day within a certain district, leading to the introduction of time lags between different districts.

These were removed by identifying which days had no crimes reported and introducing zeroes at the appropriate intervals in the time series.

Initially, we computed the correlation from one day to the next. However, we found no significant correlation between districts. Hence we performed the correlation analysis on different time scales: weekly, biweekly, monthly, bimonthly, quarterly and yearly. From this analysis, we observed that a significant correlation exists between districts for bimonthly and quarterly theft-based crimes as shown in Figure 4. There are four districts where the highest correlations were found over two of the time intervals we tested and two districts were found to be common to both.

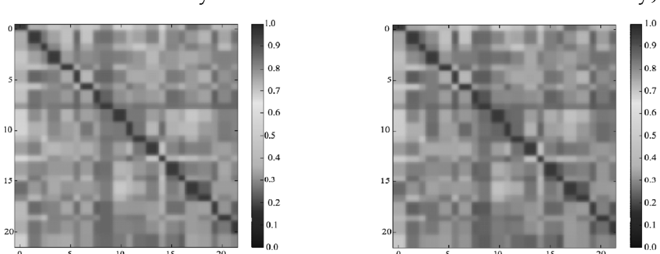

Figure 4: This figure represents the results of our correlation analysis on theft records from year 2001 through 2015 on a bi-monthly scale (left) and on a quarterly scale (right)

DISCUSSION

Crime and its prevention is a global challenge. The large and complex crime data sets that are nowadays available offer an opportunity to gain insights into how crime is propagated and the possibility of new ways of combatting it. There are many datasets in many formats that are rapidly growing. These not only include government statistics and crime incidence reports, but can also include intelligence gathering resources such as social media postings, commercial transaction records, trading records and phone and text records. These datasets can only be handled by advanced Big Data technology to get productive results.

Law enforcement organizations are also under huge pressure due to daily increase in data. The tools which were previously reserved for large agencies and research groups are becoming available to state and law enforcement agencies to make better use of crime analytics and decision-making. The advancements in technology data mining tools no longer require large budgets for analyzing data. Many products are easy-to- use, easy to access and highly intuitive.

Data mining involves multiple processes that help in finding patterns and relationships using these Big Data. This can lead to accurate predictions of events based upon historical records and their patterns and relationships. Developing algorithms and the powerful technologies to explore the depths of data might help us to create new crime prevention programs and locating suspects. The advantage in data mining is that we do not need to know what we are looking for in advance. Instead, the key to using data mining is to discover something new. This data-driven approach starts with data and then builds patterns and theories based on the trends.

In our correlation analysis of the city of Chicago crime data, we observed patterns between the neighborhoods. Using the correlation coefficients from the matrices we found the two highest correlations among 4 districts from bimonthly and quarterly correlations. Plotting the highly correlated districts geographically gave the insight that crimes happening in one neighborhood were correlated with later crimes happening in another.

For the bimonthly correlation analysis, the top two pairs of significantly correlated districts are on one hand the Near North district and the Town Hall district, and on the other hand the Foster district and the Rogers Park district (Figure 5, left). For the quarterly data, the top two pairs of significantly correlated districts are on one hand the Chicago Lawn district and the Deering district and on the other hand the Foster district and the Rogers Park district (Figure 5, right). We observed that theft is significantly higher in the Foster district and the Rogers Park district in both the bimonthly and quarterly correlations.

One use from calculating these correlations is if any two districts are correlated in crime, the police department can allocate officers accordingly, and as there can be approximately equal crimes in both districts, officers can also be appointed in approximately equal numbers. With these kinds of results, we can expand our approach towards different types of crimes other than theft alone and it can be applied to crimes in other cities and states.

Our data also suggests it might be possible to predict an increase in crime where there is a correlation in crime between neighborhoods. We identified that the crimes in one neighborhood were correlated with the crimes in a different neighborhood at a later point in time. This correlation could represent a crime wave that originates in one neighborhood and predictably moves to another. Under these circumstances, such data would be useful for law enforcement as it would allow extra resources to be assigned to the latter neighborhood after an increase in crime was detected in the former. It would be interesting to see if this intriguing possibility is also inherent in crime data sets from other cities and whether it is a general feature of theft and other crimes in urban areas.

In future research, we would like to explore the relations inherent in other kinds of crime, such as battery, which also has a high incidence in Chicago. Some relevant questions include whether similar spatiotemporal correlations can be obtained for battery and whether there are interactions between the incidence of theft with battery or other kinds of crimes. Building a composite picture of these interactions should provide insights into how different crimes are related and aid in crime prevention strategies.

Figure 5: Neighborhoods on Chicago with the highest correlation based on the bimonthly correlation analysis (left) and based on the quarterly correlation analysis (right).

CONCLUSION

We analyzed the correlation of crimes that occurred in different neighborhoods of Chicago over different time intervals. We observed that no significant correlation was found for most of the crimes. However, after we narrowed down to the hot spots with the highest crime rates, a significant correlation could be observed between a small number of neighborhoods and a high incidence crime such as theft. The highest correlation was found between Foster district and Rogers Park district.

References

[1] Chen, H., Chung, W., Qin, Y., Chau, M., Xu, J.J., Wang, G., Zheng, R. and Atabakhsh, H. (2003). Crime data mining: an overview and case studies, in Proceedings of the 2003 annual national conference on Digital government research. Digital Government Society of North America, 2003, pp. 1–5.

[2] Estivill-Castro, V. and Lee, I. (2001). Data mining techniques for autonomous exploration of large volumes of geo-referenced crime data, in Proc. of the 6th International Conference on Geocomputation. Citeseer, 2001, pp. 24–26

[3] Hatton, W., Zhao, J., Gorantla, M. B., Chae, J., Ahlbrand, B., Xu, H., & Ko, S. (2015, October). Visual Analytics for Detecting Communication Patterns. In IEEE Conference on Visual Analytics Science and Technology (VAST), 2015, pp. 137-138.

[4] Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project, Government Information Quarterly, vol. 30, no. 4, pp. 508–513.

[5] Keyvanpour, M.R., Javideh, M., and Ebrahimi, M.R. (2011). Detecting and investigating crime by means of data mining: a general crime matching framework, Procedia Computer Science, vol. 3, pp. 872–880.

[6] Kocher, M., & Leitner, M. (2015). Forecasting of Crime Events Applying Risk Terrain Modeling. GI_Forum, 2015, 30-40.

[7] Mookiah, L., Eberle, W., & Siraj, A. (2015, April). Survey of Crime Analysis and Prediction. In The Twenty-Eighth International Flairs Conference.

[8] Nath, S.V. (2006). Crime pattern detection using data mining, in IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) Workshops. IEEE, 2006, pp. 41–44.

[9] Sanders, C. B., Weston, C., & Schott, N. (2015). Police Innovations, 'Secret Squirrels' and Accountability: Empirically Studying Intelligence-led Policing in

Canada. British Journal of Criminology, 55(4), 711-729.

[10] Schuh, G., Rudolf, S., Doelle, C., & Riesener, M. (2015, December). Preventive controlling of product development projects by aid of predictive analytics-identifying hot spots within the deviation probability map. In Industrial Engineering and Engineering Management (IEEM), 2015 IEEE International Conference on (pp. 651-655). IEEE.

[11] Swanson, B. (2015). Industry Use Cases and the Underlying Content Analytics Technology used in Big Data and Predictive Analytics (Briefing Charts). Dataskill Inc San Diego CA.

[12] Toole, J. L., Eagle, N. and Plotkin, J. B. (2011). Spatiotemporal correlations in criminal offense records, ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, no. 4, p. 38.

[13] Yu, C.-H., Ward, M.W., Morabito, M. and Ding, W. (2011). Crime forecasting using data mining techniques, in IEEE 11th International Conference on Data Mining Workshops (ICDMW), pp. 779–786.