From Criminal Spheres of Familiarity to Crime Networks

M. Oliveira, H. Barbosa-Filho, T. Yehle, S. White, and R. Menezes

Abstract We have never lived in a safer world. After peaking around 1985, both violent crime (homicide, robbery, assaut and rape) and property crimes (burglary, larceny and vehicle theft) are on a downward trend; from 1993 and 2012 crime activity has dropped by more than 40% (total number of crimes). Despite the good news, crime is still prevalent in most large cities. FBI reports that in 2013 there were about 3,098 crimes per 100,000 habitants in the USA, with 2,730 of them being property crimes and 367 violent. What most people can agree is that one preventable crime is one crime that should not have taken place. The unveiling of the structure of criminal activity can lead to a better understanding of crime as a whole which in turn can help us provide better protection to our citizens. We demonstrate in this paper that crime follows a very intersting spatial community pattern regardless of the type of crime, criminal activity aggregates in communities of well defined sizes. We believe the results of this paper is a first step towards a theory of crime modeling using network science.

Key words: Crime Networks, Crime Structure, Crime Analysis, Community Structure

Marcos Oliveira, Hugo Barbosa-Filho, Ronaldo Menezes
BioComplex Laboratory, Department of Computer Sciences, Florida Institute of Technology, Melbourne, Florida, USA e-mail: {moliveira,hbarbosa,rmenezes}@biocomplexlab.org

Tobin Yehle
School of Computing, University of Utah, Salt Lake City, Utah, USA e-mail: tobin.yehle@utah.edu

Sarah White
College of Arts and Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA e-mail: sarahw14@live.unc.edu

1 Introduction

The understanding of crime activity has for a long time puzzled government officials, law-enforcement officers, and researchers. A well-performed study on crime structure may have direct benefits to people's lives as it can lead to safer cities. According to the FBI Annual Crime Report [28], the USA is today much safer than it used to be in the 80s and 90s with about half of the number of crimes per 100,000 inhabitants, but still higher than the levels we enjoyed in the 60s and also higher than many countries in Europe. Indeed crime rate is dropping but the understanding of crime as a complex system can lead to further gains in public safety.

Law enforcement tends to be reactive and many times a step behind criminal activity. What if we could change this "game"? What if we could give the police an edge by making them understand criminal structure and perhaps prevent some activity before it takes place? This is becoming reality in this big-data world we live in. The change in crime rate from the 60s to today can probably be explained by a technology lag. In the 60s, we had a smaller population and hence crime was easier to understand and prevent with "manual" approaches. As the population grew, our ability to effectively keep track of what was going on diminished and consequently crime rate ramped up. More recently, we have seen technology catching up via the use of data analysis and mining. What if we could do more? Like many complex systems we believe there is a structure that governs the interactions of criminals. This paper is an initial step towards the understanding of this structure.

Most of the works in crime structure start from the premise that crime is a consequence of factors such as wealth (or lack of) [14], education levels [17], age [19], and many others. However, more recently we have seen scientist starting to look at structure in particular social networks, as a way to explain the existence of crime in certain neighborhoods [7, 8] but to our knowledge scientist are yet to look at the structure of spatial distributions of crime. Few have attempted to look at spatial data and analysis in the context of crime control [1] with most of the studies being related to understanding the emergence of hotspots of crime [11, 25]. In this paper we show that the use of hotspots to understand crime spatial structure misses important features that can be better represented and analyzed using networks. In fact, we show that crime networks built from spatial data about crime location appears to reveal social structures when the spatial resolution is high. Our results show that hidden in the distribution of crime (hotspots) is a social structure that may be related to the social network of criminals or the social network of people affected by crimes. In this paper we show how we can uncover this structure.

2 Related Work

Crime is a complex issue and many factors affects its occurrence including: sociological, economic, psychological, biological, philosophical and even religious factors [12].

With regards to crime structure two theories in criminology can be highlighted: the *routine activity* and *social disorganization* theories. The former argues that criminal activity occurs at the convergence of three things: a potential offender, a lack of guardianship or supervision, and a target [5]. The latter contends that criminal activity is the result of the social and physical environments of the neighborhood at hand [32]. Both theories seek to model crime phenomena using spatial and geographical context.

The aforementioned opportunistic nature of the routine activity theory supports that criminal activity typically occurs in the sphere of familiarity of the criminal. Despite this sphere of familiarity being peculiar to the individual, areas of high traffic, such as downtown areas, lie within the sphere of familiarity of many individuals; it is feasible that these criminals with the same sphere of familiarity are aware of each other. This aspect is also related to the fact that criminals typically commit crimes within a short distance from their home [15].

Metropolitan areas are typically organized by regions of different land uses such as: residential, commercial, and industrial use. The presence of types of crimes differs between these land uses; neighborhoods with residential housing and no commercial businesses are perceived as *safe* and non-residential land uses are correlated with an increase in criminal activity [10]. Non-residential areas are typically found to have higher traffic in comparison to residential areas, consequently they witness to more crime [31]. Non-residential land use, such as shopping centers or public parks, coincides with an increase in *foreign* or non-residential presence. This presence of such *strangers* negatively impacts a neighborhood's social structure [24].

Street network (from layout) are not only correlated with an increase in crime incidence but additionally have a relationship with the typical *journey-to-crime* length of an offender [15]. Roadways and public transportation link together different areas of a criminal's sphere of familiarity and facilitate travel outside of a criminal's immediate neighborhood. The type of crime can affect the *journey-to-crime* length. For example, violent crime trips are shorter in length than property crime trips [15].

Despite the understanding we have of crime activity, its causes and consequences, recent studies continue to look at spatial crime analysis using approaches related to the formation of hotspots [20]. Additionally, there has been many efforts that tries to analyze crime activity in light of the existence of social networks. Many studies have looked into characteristics of ties such as their strength [23], the frequency of ties [6], and the race and gender of those with more ties [30, 22]. Yet, these studies rarely consider the structure of the overall network and they assume the existence of some information regarding the social structure of criminals. However, this is not always possible and, in fact, such structure may not be available. Law enforcement datasets rarely include information about criminals acquaintances and when they do, the reliability of such information is doubtful.

The approach we propose then is to focus on the *journey-to-crime* [15] and build networks out of the distance between crimes. Rather than social networks we have crime networks where nodes represents actual crimes and links between crimes related to a distance (or sphere) between the crimes. Our results are important because we demonstrate that a spatial networks of crime appear to contain information about the social structure of the people involved in the criminal activity.

3 Constructing Network of Crimes

The network creation mechanism is based on the geographical proximity between crimes. Two events are *connected* if they occurred within a certain distance. This network creation model is as simple as possible. In fact, the mechanism is the same as used to generate random geometric graphs [9].

Notice that the connection definition we are using here is basically the same as in the context of geometric graphs and does no presume any actual relationship between the events other than their proximity. Therefore, the network structures are going to be fully determined by the spatial distribution of the crimes. Each point in our dataset can be seen as the location of a person—in this particular case, an offender— at a given moment, in a similar manner as checkins in geolocated social networks [21] or mobile phone activities in Call Detail Records datasets [29, 26]. The main difference however is that in our data there is no individual-level identification.

Although such data could evidently yield a higher-resolution analysis, we decided to focus on the coarse-grained spatial distribution of crimes. The rationale for this is twofold: (1) in this paper we aim to uncover network structures (possibly) embedded within the spatial distribution of crimes; (2) for practical reasons, we based our analysis exclusively on publicly available data and hence we do not use any individual-level information.

That said, the theoretical basis supporting our approach are grounded mainly on two principles, both very well documented in the criminology literature:

1. For most crime trips, the distance from the offender's home to the crime location is relatively short and the probability of an offender committing a crime decays with the distance from their home [13, 15];
2. Offenders tend to live near to their associates and long-distance ties are rare [18].

Not surprisingly, these characteristics conform to two behaviors largely observed in general human dynamics: (1) most of our trips are for short distances and very long jumps are less likely to occur [27, 29] and (2) the probability of finding a social tie between two individuals decays as a power function of the distance [16, 2, 29]. Therefore, it is plausible to assume that patterns on spatial distribution of crimes should emerge from the convolution of both the individual and social level dynamics.

4 Experimental Results

4.1 Spatial distribution of crimes

Hot spots of crimes do not occur uniformly in a region. This aspect of the criminal activity can be visualized in the heatmaps depicted in Figures 1(a-c) from the Los Angeles area. This type of map, which shows the places where most crimes were committed, is widely used as a tool to understand the emergence of hot spots as well as to elaborate law enforcement strategies. These heatmaps are geolocalized histograms that allow a prompt analysis of the crime frequency in a specific region. For example, as stated, the aforementioned maps show that there are certain sub-regions with high criminal activities placed across the Los Angeles area. These maps in Figure 1(a), 1(b) and 1(c) depict the placement of the hot spots regarding assaults, burglaries and thefts, respectively. Their analyses suggest that these types of crime have particular arrangements in the region and that they may occur due to different kinds of crime activity dynamics.

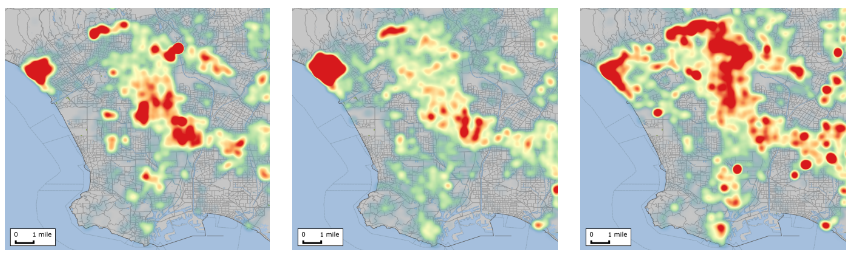
However, such maps do not allow analyses beyond the criminal activity frequency of a region. An example of this insufficient data description is that although these visualizations make possible to see many different hot spots together, there is no information about their relationships nor the overall structure that may enable the emergence of the hot spots. Actually, this structural analysis is carried out more by the viewer of the map than brought by the heatmap as a tool. Nevertheless, this structural information can be useful to understand underlying mechanisms in criminal phenomena. For instance, although the examination of Figure 1(a), 1(b) and 1(c) suggests that these particular kind of criminal activities have different dynamics across the region, this comparison may neglect similar underlying mechanisms related to the emergence of hot spots.

In order to capture the similarity of different types of crimes to subsequently analyses, Figure 1(d) is elaborated in such way that only the hottest spots of each kind of crime are considered. The hottest spots are the ones that the crime frequency is two standard deviations higher than the average frequency of this type of criminal activity, thus the map does not present any intensity interval. The intersections between these spots are shown in the map by different colors, as described in the map legend. The rationale of this visualization is to understand the hot-spot mixing in the region, *i.e.* the places where different crimes are concentrated. In the Los Angeles metropolitan area the mixing of the criminal activities coexistence are related to the colors in the map and their percentage is shown in Table 1.

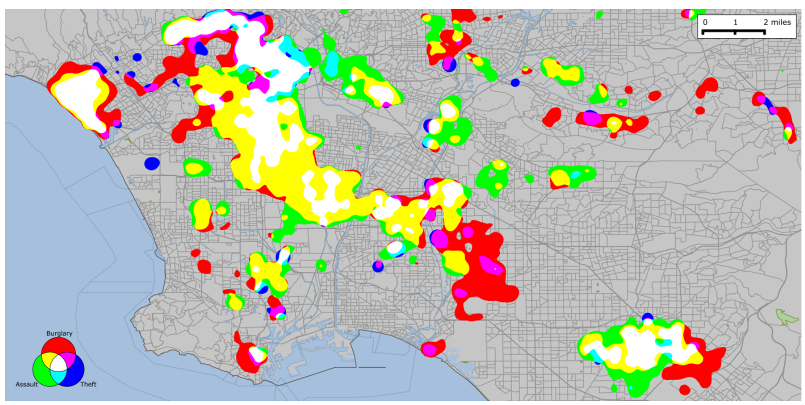
The hot-spot mixing indicates that assaults, burglaries and thefts tend to coexist as hot spots in Los Angeles area. This finding may hint to the existence of some similarities in possible underlying mechanisms that lead to the emergence of these hot spots. Conversely, assault and burglary do present some particularities that allow them to occur more independently across the area considered. In other words, these results suggest that these different types of crime seem to have a core behavior as well as particular behaviors. Regardless of this analysis, heatmaps look at crime frequency and are not enough to assess such underlying mechanisms.

Table 1 The criminal hot-spot mixing in the Los Angeles metropolitan area presents the coexistence of different criminal activities regions. This mixing reveals that the hot spots of thefts usually happen in companion to burglaries and assaults. On the other hand, the other two seem to be less linked to other crimes, resulting in more independent hot spots. In the table below the letters indicate the types of crime, hence A & B represents the region of the overlap of assauts and burglaries.

Assault (A) (green)	Burglary (B) (red)	Theft (T) (blue)	A & B (yellow)	A & T (cyan)	B & T (magenta)	A & B & T (white)
20%	24%	3%	26%	3%	5%	19%



(a) (b) (c)



(d)

Fig. 1 The places where crimes occur are not uniformly distributed in a region. The heatmaps of these events, in the Los Angeles metropolitan area, for different types of crime, assault (a), burglary (b) and theft (c), help the realization that hot spots of crime exist, but the approach is not adequate to carry out structural analysis of the crimes. These heatmaps together can help us visualize the different placements of the hot spots when different crimes are taken into account. The coincidence map (d), an overlap of the hottest spots from these heatmap, shows that thefts tend to happen in places where other crimes are also intensively happening, while burglaries and assault may occur more independently.

4.2 Microinteractions and the spatial distribution of crimes

Complex networks of a particular class often share several common topological features. For example, a social network is expected to have a high coefficient of clustering while having a short average path length. On the other hand, technological networks such as the Internet tend to have a hierarchical topology.

The existence of such structural and topological patterns plays a central role in order to have a better understanding of the various phenomena and real dynamics driven by one or more network structures. This is especially important when the complex network underlying a particular phenomenon can not be observed directly.

In this section, we seek to extract and identify possible network structures beyond the spatial geometric network itself that we built. When we build a network by simply connecting points geographically close to a distance d, this network will have features of a spatial or geometric network.

4.2.1 Clustering Coefficient

Many complex networks are characterized by a high clustering coefficient. That is specially true for geographically constrained networks where the characteristic link length is bounded up to a distance d. In a spatial network, the global clustering coefficient is expected to increase as a function of the distance d from fully disconnected nodes (when $d = 0$) to a single clique of size N (when $d \to \infty$) where N is the population size. However, neither of these two extremes are of much help in understanding a complex phenomenon such as the dynamics behind criminal activities. Hence, there must be a characteristic radius d (or a function $f(d)$) where the underlying networks unvail themselves.

In such spatial geometric networks, the clustering coefficient is a function of the connection threshold d and should increase monotonically with it. To test this hypothesis we analyzed the changes in the structure of the network for small increments in d starting from $d = 0.02$ miles to $d = 3.2$.

What was unexpected however is a gradual decrease observed in the clustering coefficient for a particular range of d (as in Figure 2), deviating from the characteristics of a spatial network [4, 33] whose clustering coefficient should increase monotonically with d once the spatial boundaries are growing and the longer links are becoming more frequent.

It is also noteworthy the fact that the clustering coefficient reached its minimum for $0.4 \leq d \leq 0.8$, for different cities and crime types suggesting that the networks are undergoing a phase transition for some critical value of $d \approx 0.6$. This behavior could be related to the case in which for very small values of d, the spatial constraints does not play a role anymore and therefore the remaining network structure could result from some other dynamic factor. To test such hypothesis, we investigate what other structural characteristics are also changing with d by comparing their properties for $d < 0.4$ and $d > 0.4$.

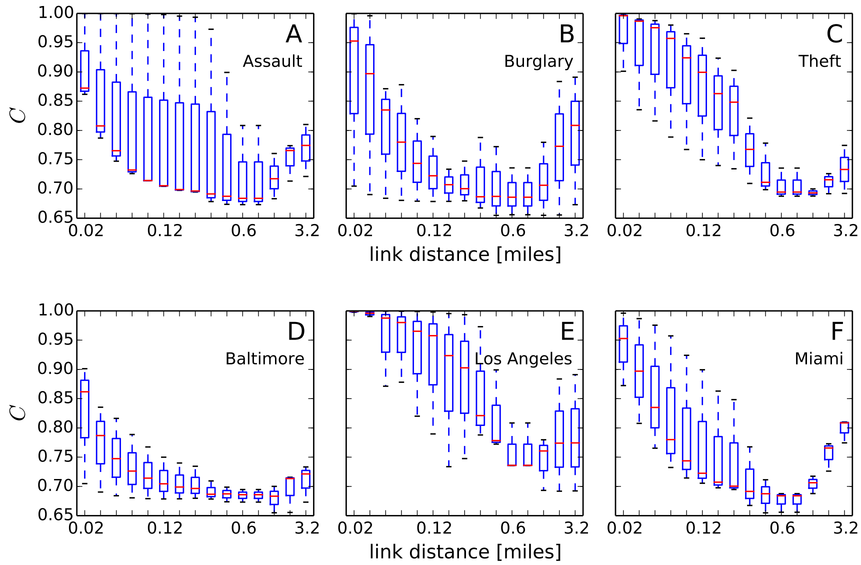


Fig. 2 The plots depict the evolution of the global clustering coefficients by the different linking distance threshold d. The top row shows the clustering coefficient for each type of crime, *assault*, *burglary* and *theft*. The bottom row shows the clustering coefficient for each of the three metropolitan areas. The correlation between d and the clustering coefficient suggest a marked structural change in the network with a critical point $0.4 \leq d \leq 0.8$ miles. Even though the actual shape of the curves varies over different networks, in all of them, the minimum clustering degree was reached in the region close to $d \approx 0.6$.

4.2.2 Degree Distribution

One next natural step would be an analysis to the degree distribution of the networks, assessing how good they fit to a heavy-tailed distribution. The rationale here is that a heavy tailed degree distribution is a key signature of some interesting complex networks such as social networks [3]. On the other hand, this property does not hold for other classes of networks, including spatial networks [4] which could indicate that the networks are not just undergoing structural transformations but also their signatures are transitioning from of one class of network to another.

From Figure 3, the linking threshold capable of producing networks with heavy-tailed degree distribution happens when $d = 0.1$. Another interesting result is that the power-law exponents of most of the networks have an exponent $\alpha \approx 2.1$ in agreement to the characteristic exponent of scale-free networks.

Although the cumulative degree distributions were consistent with the findings about the clustering coefficients in Section 4.2.1, this analysis is not sufficient to assess the correlation between the value of d and the goodness of fit of a power law to the degree distribution. For this task we used the Kolmogorov-Smirnov test to check for which ranges of d the power-law distribution presents a good fit to the

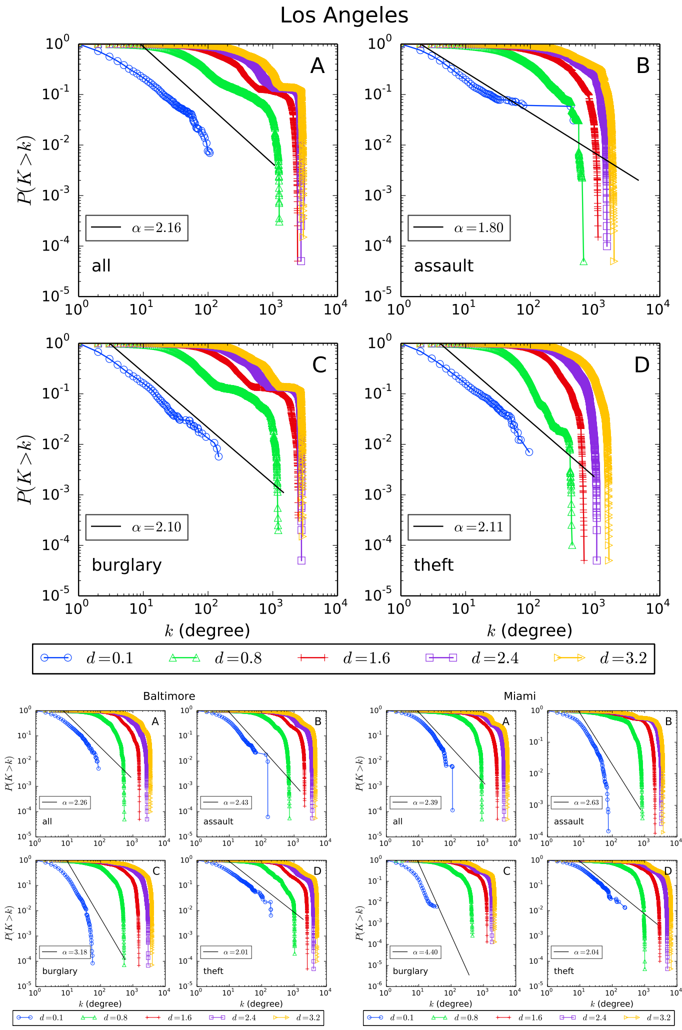


Fig. 3 The cumulative degree distribution of the networks exhibit a strong pattern accross different cities and crime types. In all the networks we investigated, the degree distribution exhibited a heavy tail but only up to a critical value of $d = \delta$. Beyond this point the heavy tail vanishes. On the other hand, for $d < \delta$ almost all the networks had degree distribution in agreement with a power law with exponent $\alpha \approx 2.1$. Straight lines are shown as a guide.

degree distribution. Figure 4 depicts the KS distance from nodes degree cumulative distribution function to a theoretical power-law distribution. However, it is important to emphasize that our focus is not to determine whether the degree distribution is indeed a power law but rather to assess the intervals for the parameter d for which the degree distribution agrees to or deviates from a heavy tailed.

The KS test confirmed our hypothesis that the degree distribution for values of d beyond a certain point have no interesting feature. Based on the test results with KS, for $d > 0.8$ we witness an abrupt increase in the distance from the degree distribution to the power law, in agreement with the results previously found.

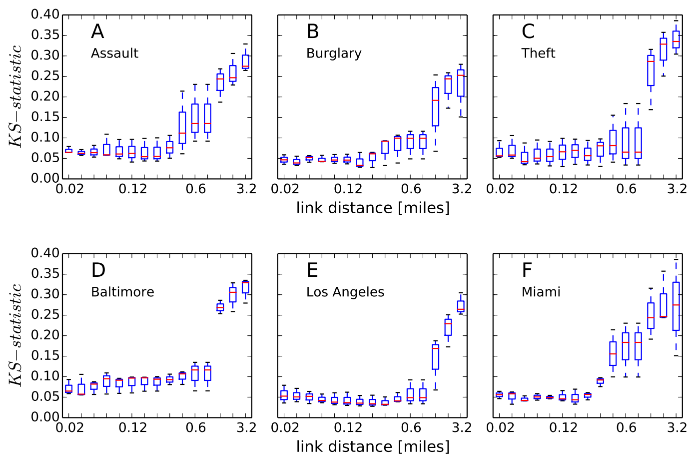


Fig. 4 Kolmogorov-Smirnov distance from the empirical degree distribution to a theoretical power law for different metropolitan areas and crime types. The KS statistic supports our claim that the degree distributions of the crimes networks follow a power-law distribution up to a certain value $d = \delta$.

It is clear that these tests are not sufficient to prove that the networks emerging for small d are actually the social networks of criminals. In fact, what we are arguing instead is that the dynamics that produced the spatial distribution of crimes result from a combination of influences of two complex systems: the social dynamics and spatial constraints. However, when analyzing the network of crimes in a high resolution where the characteristic edge length is very short, our results suggest that the observed network no longer behaves as a spatial network and starts to display characteristics observed also in social networks.

5 Conclusion and Future Work

In this paper we looked at the structure of crime in urban environments and demonstrated that one may be able to use spatial networks [4] to extract social information. This seems to be quite clear to case of crime. Our results show that in higher spatial resolutions (less than a mile), network of crimes appear to contain information of the social structure of the individuals involved in the criminal activity. One questions that arrises here is do other spatial networks could also contain social information. We are currently working on other datasets.

In addition to the contribution of showing that social information may be extracted from spatial networks, Our work may be used in the decision-making process of law enforcement officials. We have mentioned earlier that in many instances, the law enforcement agencies may not have in their datasets social information about the criminals and that sometimes the information is incomplete. We believe further work on our approach may lead to the ability of reconstructing these structures. As is, the work can already help decision making because theories from network science can tell us which nodes to focus if we want to disrupt the network; the social structure of crime can be used as a way to understand where the police should focus.

There are several points that need to be studied further. One of the main points is the possibility of defining a scaling law for different types of crimes. Our results appear to show that the social structure emerges at slight different scales depending on the type of crime. However one needs to understand the other variables that may play a role in this such as city demographics and city layout, to name a few.

The test on other cities may also be useful. We tested with 3 cities in the USA. We have not used any variable that is particular to the USA and we have no reason to believe the approach would not be applicable to other places. However we intend to apply the same approach to cities in South America and Europe.

References

1. L. Anselin, J. Cohen, D. Cook, W. Gorr, and G. Tita. Spatial analyses of crime. *Criminal justice*, 4(2):213–262, 2000.
2. L. Backstrom, E. Sun, and C. Marlow. Find me if you can: improving geographical prediction with social and spatial proximity. *Proceedings of the 19th international conference on World Wide Web*, 2010.
3. A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, page 11, Oct. 1999.
4. M. Barthélemy. Spatial networks. *Physics Reports*, 499(1):1–101, 2011.
5. D. Beavon, P. Brantingham, and P. Brantingham. The Influence of Street Networks on the Patterning of Property Offenses. *Crime prevention studies*, 1994.
6. P. E. Bellair. Social interaction and community crime: Examining the importance of neighbor networks. *Criminology*, 35(4):677–704, 1997.
7. C. R. Browning, R. D. Dietz, and S. L. Feinberg. The paradox of social organization: Networks, collective efficacy, and violent crime in urban neighborhoods. *Social Forces*, 83(2):503–534, 2004.
8. A. Calvó-Armengol and Y. Zenou. Social networks and crime decisions: The role of social structure in facilitating delinquent behavior. *International Economic Review*, 45(3):939–958, 2004.
9. J. Dall and M. Christensen. Random geometric graphs. *Physical Review E*, 66(1):016121, July 2002.
10. S. Foster, L. Wood, H. Christian, M. Knuiman, and B. Giles-Corti. Planning safer suburbs: Do changes in the built environment influence residents' perceptions of crime risk? *Social Science & Medicine*, 97:87–94, 2013.

11. V. Furtado, A. Melo, A. Coelho, and R. Menezes. A crime simulation model based on social networks and swarm intelligence. In *Proceedings of the 2007 ACM symposium on Applied computing*, pages 56–57. ACM, 2007.
12. S. Guarino-Ghezzi and A. J. Treviño. *Understanding Crime: A Multidisciplinary Approach*. Elsevier, 2010.
13. K. Harries. Mapping crime: Principle and practice. Technical report, U.S. Department of Justice - Office of Justice Program, Washington, DC, US, 1999.
14. B. P. Kennedy, I. Kawachi, D. Prothrow-Stith, K. Lochner, and V. Gupta. Social capital, income inequality, and firearm violent crime. *Social science & medicine*, 47(1):7–17, 1998.
15. N. Levine and P. Lee. Journey-to-crime by gender and age group in manchester, england. In *Crime modeling and mapping using geospatial technologies*, pages 145–178. Springer, 2013.
16. D. Liben-Nowell and J. Novak. Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, pages 1–6, 2005.
17. L. Lochner and E. Moretti. The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. Technical report, National Bureau of Economic Research, 2001.
18. A. E. Malm, J. B. Kinney, and N. R. Pollard. Social Network and Distance Correlates of Criminal Associates Involved in Illicit Drug Production. *Security Journal*, 21(1-2):77–94, Feb. 2008.
19. T. B. Marvell and C. E. Moody Jr. Age structure and crime rates: The conflicting evidence. *Journal of Quantitative Criminology*, 7(3):237–273, 1991.
20. A. T. Murray and T. H. Grubesic. Exploring spatial patterns of crime using non-hierarchical cluster analysis. In *Crime Modeling and Mapping Using Geospatial Technologies*, pages 105–124. Springer, 2013.
21. A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo. A tale of many cities: universal patterns in human urban mobility. *PloS one*, 7(5):e37027, Jan. 2012.
22. P. W. Rountree. A reexamination of the crime-fear linkage. *Journal of research in crime and delinquency*, 35(3):341–372, 1998.
23. R. J. Sampson. Networks and neighbourhoods: The implications of connectivity for thinking about crime in the modern city. In H. McCarthy, P. Miller, and P. Skidmore, editors, *Network Logic: Who Governs in an Interconnected World?*, pages 105–124. London: Demos, 2004.
24. R. J. Sampson and W. B. Groves. Community structure and crime: Testing social-disorganization theory. *American journal of sociology*, pages 774–802, 1989.
25. M. B. Short, P. J. Brantingham, A. L. Bertozzi, and G. E. Tita. Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences*, 107(9):3961–3965, 2010.
26. F. Simini, A. Maritan, and Z. Néda. Human mobility in a continuum approach. *PloS one*, 8(3):e60069, Jan. 2013.
27. C. Song, T. Koren, P. Wang, and A.-l. Barabási. Modelling the scaling properties of human mobility. *Nature Physics*, 6(10):818–823, Sept. 2010.
28. The Federal Bureau of Investigation. Crime in the united states (cius). Technical report, Department of Justice, 2013. http://www.fbi.gov/about-us/cjis/ucr/ucr.
29. D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabási. Human mobility, social ties, and link prediction. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11*, page 1100, New York, New York, USA, 2011. ACM Press.
30. B. D. Warner and P. W. Rountree. Local social ties in a community and crime model: Questioning the systemic nature of informal social control. *Soc. Probs.*, 44:520, 1997.
31. G. F. White. Neighborhood permeability and burglary rates. *Justice Quarterly*, 7(1):57–67, 1990.
32. D. Willits, L. Broidy, A. Gonzales, and K. Denman. Place and Neighborhood Crime: Examining the Relationship between Schools, Churches, and Alcohol Related Establishments and Crime. Technical report, Institute for Social Research, Mar. 2011.
33. L. Wong, P. Pattison, and G. Robins. A spatial model for social networks. *Physica A: Statistical Mechanics and its Applications*, pages 1–23, 2006.