PAF 593 Final Project Report - Co-Offending Networks in Phoenix (2018-2023)

Jacob T.N. Young

Degree Program: Program Evaluation and Data Analytics

Instructor: Gerber

Title: Co-Offending Networks in Phoenix (2018-2023)

Abstract

This project shows how co-offending networks can be created from open-access data. Specifically, using data from the City of Phoenix Open Data Portal (https://www.phoenixopendata.com/dataset/arrests), two-mode networks representing individuals who are arrested together are built and examined. Attention is dedicated to examining reachability and clustering of the networks, as these are important properties identified in the research literature. Examining the *reachability* and the *clustering* in the co-offending networks, indicated that there is a low level of reachability and clustering for all years. However, the results did show an increase in clustering from 2018 to 2023. Clustering and trust are important in co-offending networks as these structures may facilitate group solidarity and greater criminal organization. The policy recommendation is that police agencies focus greater attention on the mechanisms that may lead to the observed increase in clustering.

Introduction

By its nature, crime tends to be a group activity (Warr, 2002). Nearly a century of research points to the important role that social relationships play in criminal behavior, indicating that crime is a "network" phenomenon (Papachristos, 2014). While this body of research has covered diverse networks such as gangs (Grund and Densley, 2014), drug distribution (Bright, Koskinen, and Malm, 2018), and mafia syndicates (Andris, et. al., 2021), the bulk of the work focuses on co-offending networks. That is, the relational structure by which individuals participate in crime together. Simultaneously, as police departments have moved toward more intelligence-based efforts at policing, examining co-offending networks has become a key tool in crime reduction efforts (Rostami and Mondani, 2015).

Studies of co-offending have shown that two structural characteristics are relevant for developing crime reduction policy: reachability and clustering. This project begins by reviewing literature on these two structural properties (reachability and clustering) and then moves toward an analysis of reachability and clustering among arrests of individuals in Phoenix between January 2018 and November 2023. The project finishes with a discussion of the value of the "open-science" approach used here and emphasizes that collaboration among researchers and practitioners in the area of crime reduction and network analysis remains a fruitful area for synergistic activity.

Project Goal

This project seeks to develop an open-science approach by examining adult arrest data from the City of Phoenix Open Data Portal (https://www.phoenixopendata.com/dataset/arrests) to construct the co-offending network of individuals arrested in Phoenix between January 2018 and November 2023. This project makes two important contributions. First, the structure of these networks will be analyzed, and policies examined as they pertain to reachability and clustering. Second, it will document how to wrangle and analyze the open-access data.

Literature Review

Studies of co-offending have shown that two structural characteristics are relevant for developing crime reduction policy: *reachability* and *clustering*. These structural characteristics define important properties of the network and help provide insight about the processes that drive co-offending networks. This section reviews these studies.

Reachability and Diffusion

Networks that have high reachability are those that connect many individuals to each other. For example, while Jim and Mike may not be directly linked, Jim may be linked to Kate, who is linked to Steve, who is linked to Mike. Thus, Jim can "reach" Mike through this chain. Such a structural property facilities the spread of information, diseases, or risk. In contrast, a network with low reachability makes it difficult for content to spread. For example, if Jim cannot reach Mike through some chain of relationships, then Jim cannot indirectly influence Mike. This feature of a network facilitates "diffusion" or the spread of content.

Diffusion views crime as a consequence of an epidemic like process where some causal aspect is transmitted through the network. Networks that are more amenable to diffusion will influence which policies are adopted to prevent crime. For example, Green, Horel, and Papachristos (2017) examined whether individuals are more susceptible to gun violence if they are connected to other individuals who experience gun violence. Using co-arrests networks among individuals in Chicago from 2006 to 2014, they found that the likelihood of being fatally shot or injured by a firearm increased for an individual if someone whom they had been arrested with experienced this event. They argued that this suggests gun violence diffuses through co-offending networks like a virus. Similar processes of diffusion have been found in diverse outcomes from gang violence (Brantingham, Yuan, and Herz, 2020) to police misconduct (Ouellet, Hashimi, & Vega Yon, 2022).

Networks characterized by high reachability, where causal factors can diffuse, require different interventions than those that are characterized by less reachability. Networks in which individuals are reachable may require identifying individuals or sets of individuals who serve as "bridges" between other individuals. These individuals, when targeted for intervention, may disconnect the network reducing the spread of the causal factor in question. In contrast, networks that are not highly reachable will have disconnected groups, requiring a different intervention approach. Therefore, it is important to examine the reachability of a co-offending network.

Clustering and Trust

Another important feature of co-offending networks for policy considerations is the level of clustering in the network. Networks that have high clustering are those in which two individuals share a common third person. For example, Jim and Mike are friends and Mike and Kate are friends. If Jim and Kate are friends this forms a “triadic” structure. The more triadic structures in the network, the more clustering. For a co-offending network, there is clustering if Steve, Trevor, and Ross all engage in offending together. Such a structural property facilities the spread of information and experience with individuals leading to the development of trust relationships. When individuals know about the behavior of others, they are better positioned to make trust judgments about those individuals. In contrast, a network with low clustering makes it difficult for actors to develop trust.

Clustering and trust are important in co-offending networks as these structures may facilitate group solidarity and greater criminal organization. Moreover, networks where there is high clustering tend to create modularity or sets of individuals that are only connected to members of that set. For example, Grund and Densley (2014) examined the level of clustering among gang members in London. They found that racial homogeneity among gangs was largely explained by clustering in co-offending networks. That is, individuals of the same race tended to engage in crime together, leading to groups of racially segregated gangs. Clustering has also been shown to be important for future offending. Charette and Papachristos (2017) found that individuals were more likely to engage in an offense together if they had similar co-offenders in the past. Therefore, it is important to examine the level of clustering in a co-offending network.

An Open-Science Approach

Collaboration among researchers and practitioners in the area of crime reduction and network analysis remains a fruitful area for synergistic activity. However, a key barrier to synergistic activity of this sort is accessing data and sharing information between researchers and departments. As Bright, Brewer, and Morselli (2021: 54) note, negotiating data access is “one of the greatest challenges” for advancing research and policy in this area. Many of the studies on co-offending networks are the consequence of long-term relationship building between researchers and police departments. Such limited access to data on co-offending networks has resulted in researchers using the same data for different projects and multiple studies. The consequence is that the knowledge base for studies examining co-offending networks, and the policies that are gleaned from those studies, rely on a limited or single source of data when making inferences about the processes of interest. Moving toward an approach to research where data are available and accessible would be beneficial to help overcome this barrier between researchers and agencies and generate better policy and better crime reduction efforts.

In the last quarter century, there has been a monumental shift toward making scientific practice more "open". A key component of this so-called open science movement is access to data collected and held by government agencies, often referred to as open government. The open science and open government movements have created a watershed moment for scholars, policymakers, and citizens concerned about public safety. A key feature of the embrace of open government is in how agencies think about data. Specifically, agencies view their data as a strategic asset to achieve their missions. The availability of large amounts of digital data ("big data") amassed in agencies marks a transition toward "governance by data" which is the "gathering, assemblage, formatting, analysis, transmission, and storage of digital data to have governance effects" (Johns, 2021: 54-55). Specific to co-offending networks, an "open science" approach would aid in forming collaborations among researcher and agencies, but would also provide a roadmap for individuals working with other agencies to follow and perhaps even coordinate.

Project Methodology

Data Acquisition and Workflow

Data were downloaded from the City of Phoenix Open Data Portal. A workflow was created whereby the data were downloaded, cleaned, and then analyzed. The workflow and all code used for this project are available at the following GitHub repository: https://github.com/jacobtnyoung/PAF-593-project.

Network Construction and Measures

The co-offending data are used to construct two-mode networks. A two-mode network has two sets of nodes: incidents (i.e. arrests) and individuals (i.e. people arrested together) and edges only occur between these partitions (i.e. not within). The definition of a two-mode network is the following: $G = (N, M, L)$ where G is the graph and is defined by two node sets $N = \{n_1, n_2 \dots, n_g\}$ and $M = \{m_1, m_2 \dots, m_g\}$ and an edge set: $L = \{l_1, l_2 \dots, l_L\}$. In this definition, there are N nodes in the first set, M nodes in the second set, and L lines/edges in the network.

In a typical, one-mode network, adjacency is defined between two nodes, n_i and n_j as a link $l_k = (n_i, n_j)$. These data can be represented as an **adjacency matrix**, where each node is listed on the row and the column. The i_{th} row and the j_{th} column of X_{ij} records the value of a tie from *i* to *j*. In two-mode networks, an **adjacency matrix** can also be used to represent the data. The difference is that there are two different sets of nodes N and M. Two nodes from separate node sets, n_i and m_j are adjacent if there is a link $l_k = (n_i, m_j)$. The i_{th} row and the j_{th} column of X_{ij} records the value of a tie from n_i to m_j.

Reachability

Based on the adjacency matrix created above, a measure of reachability can be calculated based on how "reachable" individual **i** is to individual **j** where these nodes are connected through a co-arrest. Reachability is the inverse of the distance between the two nodes where distance is defined as the number of events that connect two individuals. In this analysis, the minimum, mean, and maximum reachabiilty will be examined.

Clustering

In a two-mode network, the extent to which individuals overlap in their co-offenders is capture by the clustering in the network. Specifically, clustering coefficient is:

$$\frac{4 \times C_4}{L_3}$$

where C_4 is defined as a cycle (i.e. individuals 1 and 2 are both connected to events A and B) and L_3 defines a 3-path structure (i.e. individual 1 is connected to events A and B). This ratio is bound between 0 and 1. The ratio represents the proportion of cycles that are closed. A value of 1 indicates that every 3-path is closed (i.e., is embedded in a cycle). Networks with values at or close to 1 will have considerable redundancy in ties. In other words, higher values of the clustering coefficient indicate greater overlap in shared co-offenders.

Analytic Approach

I will examine the reachability scores and clustering for the co-offending networks by year.

Results

Descriptive Information for Co-Offending Networks

Year	Events	Individuals	Edges
2018	7353	13044	15979
2019	7992	12962	16043
2020	6580	10118	12594
2021	5405	8201	10246
2022	4330	6422	7648
2023	3582	6048	7359

The table above shows the descriptive information for the networks. The average number of incidents per year was 5,874 and the average number of individuals involved in those events was 9,466. The table also shows a general decline in the number of co-offending events from 2018 to 2023.

Description of Reachability for Co-Offending Networks

Year	Mean	Minimum	Maximum
2018	5.941779	0.2542611	16.27271
2019	7.052026	0.2039933	21.41930
2020	6.983182	0.2571481	21.60044
2021	6.778950	0.2085411	21.37547
2022	6.222526	0.1499775	17.32240
2023	5.748357	0.2657048	17.35938

Reachability

Recall that **reachability** in a two-mode network is defined as the number of common nodes that two individuals connect through. In this case, it represents the number of incidents in which two individuals were arrest together. Across all years, the average reachability (i.e. the mean of the mean reachability) is 6.45. This means that the average individual can reach 6.45 other individuals through their co-arrest history. The trend over time has been for reachability to remain roughly similar from 2018-2023.

Description of Clustering for Co-Offending Networks

Year	Clustering
2018	0.0067226
2019	0.0079311
2020	0.0139256
2021	0.0186057
2022	0.0295114
2023	0.0252911

Clustering

The table also shows the scores for the level of clustering in the network. Recall that networks with values at or close to 1 will have considerable redundancy in ties. In other words, higher values of the clustering coefficient indicate greater overlap in shared co-offenders. Overall, the clustering is very low. Across all years, the average clustering is 0.02. This suggest that there is very little overlap in co-offenders. However, analysis of the trend across time shows that there has been an increase in clustering from 2018 (i.e. 0.007) to 2023 (i.e. 0.025). This increase in clustering over time may be indicative of an increase in the tendency for co-offenders to trust other co-offenders.

Discussion and Policy Recommendations

This project developed an open-science approach to examining adult arrest data from the City of Phoenix Open Data Portal (https://www.phoenixopendata.com/dataset/arrests) by constructing the co-offending network of individuals arrested in Phoenix between January 2018 and November 2023. This project makes two important contributions.

First, examining the *reachability* and the *clustering* in the co-offending networks, network properties shown to be important for understanding criminal organization, indicated that there is a fairly low level of reachability and clustering for all years. However, the results did show an increase in clustering from 2018 to 2023. Clustering and trust are important in co-offending networks as these structures may facilitate group solidarity and greater criminal organization. Moreover, networks where there is high clustering tend to create modularity or sets of individuals that are only connected to members of that set. The policy recommendation is that agencies focus greater attention on the mechanisms that may lead to the observed increase in clustering.

Second, this study showed how to wrangle and analyze open-access data. This study demonstrates the value of the City of Phoenix Open Data Portal by facilitating external analysis, evaluation, and recommendation beyond actors in city agencies. The ability to create working relationships among external actors and those who seek to reduce crime in Phoenix is a major advantage of making the data available to the public. The policy recommendation is to continue to make these data public and maintain consistency in the quality of the data made available.

References

Andris, C., DellaPosta, D., Freelin, B. N., Zhu, X., Hinger, B., & Chen, H. (2021). To racketeer among neighbors: spatial features of criminal collaboration in the American Mafia. *International Journal of Geographical Information Science*, 35(12), 2463-2488. https://doi.org/10.1080/13658816.2021.1884869.

Brantingham, P. J., Yuan, B., & Herz, D. (2020). Is Gang Violent Crime More Contagious than Non-Gang Violent Crime? *Journal of Quantitative Criminology*, 37(4), 953-977. https://doi.org/10.1007/s10940-020-09479-1.

Bright, D., Koskinen, J., & Malm, A. (2018). Illicit Network Dynamics: The Formation and Evolution of a Drug Trafficking Network. *Journal of Quantitative Criminology*, 35(2), 237-258. https://doi.org/10.1007/s10940-018-9379-8.

Bright, D., Brewer, R., & Morselli, C. (2021). Using social network analysis to study crime: Navigating the challenges of criminal justice records. *Social Networks*, 66, 50-64. https://doi.org/10.1016/j.socnet.2021.01.006.

Charette, Y., & Papachristos, A. V. (2017). The network dynamics of co-offending careers. *Social Networks*, 51, 3-13. https://doi.org/10.1016/j.socnet.2016.12.005.

Green, B., Horel, T., & Papachristos, A. V. (2017). Modeling Contagion Through Social Networks to Explain and Predict Gunshot Violence in Chicago, 2006 to 2014. *JAMA Intern Med*, 177(3), 326-333. https://doi.org/10.1001/jamainternmed.2016.8245.

Grund, T. U., & Densley, J. A. (2014). Ethnic Homophily and Triad Closure. *Journal of Contemporary Criminal Justice*, 31(3), 354-370. https://doi.org/10.1177/1043986214553377.

Johns, F. 2021. Governance by Data. *Annual Review of Law and Social Science*, 17(1), 53-71.

Ouellet, M., Hashimi, S., & Vega Yon, G. G. (2022). Officer Networks and Firearm Behaviors: Assessing the Social Transmission of Weapon-Use. *Journal of Quantitative Criminology*, 39(3), 679-703. https://doi.org/10.1007/s10940-022-09546-9.

Papachristos, A. V. (2014). The Network Structure of Crime. *Sociology Compass*, 8(4), 347-357. https://doi.org/10.1111/soc4.12147.

Rostami, A., & Mondani, H. (2015). The complexity of crime network data: a case study of its consequences for crime control and the study of networks. *PLoS One*, 10(3), e0119309. https://doi.org/10.1371/journal.pone.0119309.

Warr, M. (2002). *Companions in crime: The social aspects of criminal conduct*. New York, NY: Cambridge University Press.