Exploring the Landscape: A Comprehensive Review of Computational Methods in Criminological Research

Group 9: Mehmet Firat Dundar, Pratibha Alagh, and Xu Huang

Introduction

Computational methods in criminological research have gained a great focus recently and have attracted widespread attention due to their significant advantages over traditional methods. These methods can rapidly process vast and intricate crime data at a relatively low cost, providing a more comprehensive and profound perspective for research. And these could be extremely useful where empirical research encounters obstacles due to challenges associated with field experiments.

“With the belief that criminological knowledge should be established on the principles of observation, the study of analyzing the nature of a crime has existed since the 1800s. Surveys, interviews and manually statistics analysis were used for criminal research. With the development of computer science, computers were used for data storing, processing and analysis, which could be considered as the start of computational criminology research. Advanced technology like data mining and machine learning now allowed researchers study crime pattern, crime hotspots, and the relationship between crime cases." (Caccavale, 2013)

We conducted a literature review on several articles describing computational methods in criminological research, including agent-based modelling (ABM), statistics analysis, and social network analysis (SNA). We discuss these methods and their merits and disadvantages, and share personal reflections based on the course content. The key papers we have evaluated are as follows.

The first paper by (Caccavale, 2013) explores different computational methods including statistical techniques and agent-based simulations to study delinquent behavior and criminal policy. It clearly lays down the advantages of using computational techniques for synthesizing knowledge in an interdisciplinary field like criminology. The second paper by Gong et al. (2023) delves further into unravelling the complexity of human behavior within different social contexts. They propose an agent-based model CARESim to predict the possibility of high-risk individuals committing street violent crimes. This simulation environment evaluates how various factors, such as the neighborhood, the weather, police patrol strategies, and the population, impact street-level violent offenses.

The third paper by Corina et al. (2021) applies statistics methods to analyze how the spread of disadvantages through employment-based networks impacts crime rates in corresponding neighborhoods. Finally, our fourth paper by Berlusconi (2017) and fifth paper by Sierra-Arévalo (2015) delve into SNA and its intersection with criminology, exploring how social connections shape criminal behaviors and crime prevention strategies.

Overview of methods and results

Our selection of papers provides a general overview of the current state of computational methods in criminological research, supported by detailed examples of agent-based simulation, statistical analysis and social network analysis.

Agent-based simulation

Agent-based simulation is a method of simulating various possible situations in real life by controlling variables in a virtual environment, discovering how individual individuals or groups interact to produce complex social phenomena (Caccavale, 2013). For example, when simulating forest fires, variables such as tree density, wind direction, air humidity, etc. can be adjusted to simulate the extent of forest fires under different conditions. The same approach can be applied to simulate crimes that might occur in real life.

Caccavale (2013) suggests that official crime databases are highly restricted in their coverage, as they only capture crimes that have been reported to government agencies. However, there exists a substantial portion of criminal activity that goes unreported and consequently unrecorded in these databases. This unreported criminal behavior is often referred to as the "dark number" of crime, indicating offenses that have occurred but have not been officially documented. (Caccavale, 2013)

To demonstrate agent-based modelling, Gong et. al. (2023) proposes a simulation environment, called CARESim, to predict the risk of high-risk individuals (HRIs) in committing street violent crimes. They set up the experiment in Hampton, VA, US, with an actual population of 135,000. The numbers of HRIs, police, and citizens is 100, 16, and 3,000, respectively. Each HRI has three delinquent peers and HRIs are constantly motivated to go out and seek crime opportunities. Each simulation run outputs a complete list of all simulated crimes and their details, including the HRI who is likely to commit the offense, the exact location, the time, and the values of each factor group. The model simulates the city environment in the form of layers, including geographical information support, census data, crime data, transportation data, and weather data. The variable factors such as social networks of HRIs, neighborhoods, patrol strategies, and composition of citizens can be altered to understand different patterns and predict crime occurrence. Alongside actual geographical and public data, the model utilizes strain theory to categorize agents into different roles based on their employment status and lifestyle.

“Agent-based simulation appears to be extremely beneficial when there is a need for complementarity between experimental and simulation methods, especially in situations where conducting experiments becomes extremely difficult. These methods can serve as a means of evaluating variations in different crime prevention programs before they can be effectively tested empirically. Besides, its low cost, low resource requirements, short execution time, and ability to easily handle large amounts of data also make it a very attractive option.” (Caccavale, 2013)

Statistics analysis

“Statistics analysis uses statistical methods such as regression analysis to analyze data and uncover patterns, relationships, and trends. To study how network spillovers within employment-based networks of neighborhoods affect crime rates within those neighborhoods, Corina et. al. (2021) proposed four hypotheses with the factors of the level of disadvantage within the employment-based network of neighborhoods, internal disadvantage within neighborhoods, spatial and network spillovers of disadvantage and the remaining effect of network disadvantage on crime which might influence the crime rate. Then they applied negative binomial regression method to find out the association between each factor and the crime, and computational test, leave-one-out cross-validation and permutation tests to compare the accuracy of different models, and to assess the robustness and statistical significance of the models.” (Corina et. al., 2021)

Social network analysis (SNA)

SNA refers to a methodological approach that explores how social interactions among individuals shape their behaviors. In the field of contemporary criminology, SNA has become a valuable tool, as emphasized by Berlusconi (2016). By examining these social connections, SNA helps in developing theories that aim to uncover the impact of personal and neighborhood networks on criminal activities. Moreover, SNA provides insights into the structures and operations of organizations like terrorist groups and street gangs. This enhances the understanding of researchers and law enforcement officials on these organizations leading to enforcement strategies.

In a similar manner, the paper by Sierra-Arevalo (2015) provides detailed examples of applications of SNA in the study of gang structures and interactions. The author provides an insight into the complex relationships within criminal organizations by giving detailed examples and doing network mapping using administrative data. This emphasis not only enriches the methodological approach but also highlights the importance of considering the complexities of social networks in the context of criminal behavior.

Achievements in learning

In this section, we evaluate the strengths and limitations of our chosen methods for criminological research.

Agent-based modelling

Based on an evaluation of the CARESim model proposed by Gong et. al. (2023), we can surmise that it captures factors like neighborhood and weather in experiments, which are usually untested, to determine how they can have an impact on crime occurrence. Also, its ease of adaptability can assist in high-level policy evaluations, as seen through different patrol strategies in the

experimental setup. Lastly, changing the social composition of a society can provide wider theoretical understanding of how varying lifestyles affect crime rates.

The simulation environment also underpins some well-known criminological theories. It presupposed the routine activity theory, according to which a crime is more likely to be committed when there is a willing perpetrator, a suitable victim, and a perceived opportunity (Goldstick & Jonathan, 2022). It also provides a reflection on the crime pattern theory, proposed by Brantingham & Brantingham (1984). The theory offers an explanation into why crimes are committed in certain areas. Based on the theory, a crime occurs when the activity space, which is the normal everyday path taken by a target and offender, intersects. These theoretical assumptions provide firm grounding for building agent-based models, such as CARESim, and predicting crime occurrence.

However, the CARESim model had some limitations in its application. It heavily relies on individual data elements that may need to be updated frequently, such as new crimes, HRIs location updates, and nuanced geographical data. Also, replicating the model for bigger cities, which normally have a higher crime rate (Oliveira, M, 2021), can mean adding more layers like highly complex public transportation systems. Finally, assigning correct places, for instance, for HRIs home locations or citizens' work or school locations, can be extremely challenging due to lack of reliable data. Also, lack of detailed personal data of potential criminals can restrict the scope of prediction, leading to a fairly simplistic understanding of crime patterns. But using this model in combination with other computational techniques, such as social network analysis, can better inform on crime patterns in a certain area.

Social network analysis

The utilization of SNA in studying gang structures and interactions, as exemplified in Sierra-Arevalo's research (2015), has led to significant achievements in understanding the social dynamics within criminal organizations. By employing SNA methodology to map out network structures using administrative data, researchers have been able to uncover the intricate relationships that define gang hierarchies and subgroups. This detailed analysis has challenged conventional notions of gangs as homogeneous units, revealing instead a network of smaller, highly connected cliques with distinct compositions and activities. Such findings have not only enriched our understanding of gang organization but have also provided insights into the varying roles and influences of individuals within these networks.

However, it is important to acknowledge the limitations of SNA in this context, such as potential biases in data collection and the challenge of accurately defining boundaries within fluid gang memberships. Despite these limitations, the application of SNA has advanced our knowledge of gang behavior and organizational dynamics, offering a more nuanced perspective on the complexities of criminal networks.

Statistical analysis

This study of Corina et. al. (2021) proposed four hypotheses to systematically explore the impact of network spillover on crime rates, providing multiple perspectives for understanding this complex phenomenon. By using a variety of statistical methods, the study not only conducted an in-depth analysis of the relationship between various influencing factors and crime rates, but also enhanced the persuasiveness of the research conclusions by verifying the robustness and reliability of the model. However, despite the carefully designed statistical methods used in the study, the generalization ability of the results is limited because the data are limited to a specific region. In addition, there may be uncertainties in the measurement of some variables, which may affect the accuracy and interpretability of the results, and the research results need to be interpreted with caution.

Conclusion

Agent-based models, such as CARESim, provide some clear advantages in criminological research compared to the traditional methods. ABMs can capture nonlinear interactions between heterogeneous agents. These models can be scaled down to provide more granular predictions,

such as street segments. They also offer flexibility to adjust various theoretical assumptions by adjusting the environment and redesigning agents. In cases where real-life implementation can be difficult or expansion, ABMs can prove to be cost-effective when there are limited resources, such as limited police personnel. However, they can be oversimplified at times due to computational challenges, such as the inability to capture psychological elements when determining motivation of offenders. Therefore, using them in juxtaposition with other computational methods can prove to be more beneficial for deriving substantive results.

The incorporation of SNA in the field of criminology has been a method for examining the operations of networks. Although SNA has provided information on the structure of gangs and the existence of linked subgroups within them it is important to acknowledge the inherent difficulties and constraints in researching criminal networks. Addressing issues such as biases in data and the fluid nature of gang memberships will be essential for future researchers utilizing SNA in the study of criminal organizations. Despite these challenges, the application of SNA has enriched our knowledge of dynamics of criminal groups and their hierarchies.

In comparison to traditional statistical methods, SNA and ABMs offer unique advantages in studying complex social systems such as criminal organizations. While statistical methods focus on analyzing relationships between variables and estimating effects, SNA and ABMs delve deeper into the dynamics of interactions between individual actors within networks. SNA emphasizes the relationships between actors as the primary unit of analysis, providing insights into the patterns of connections and influence within social networks, including criminal groups. On the other hand, ABMs simulate the behavior of individual agents based on predefined rules and interactions, allowing researchers to visualize and simulate behaviors in near real time that may not be captured by traditional statistical approaches. By integrating these diverse methodological approaches, researchers can gain a more comprehensive understanding of the underlying mechanisms driving criminal behaviors, enhancing the effectiveness of interventions and policies aimed at crime prevention and promoting community safety.

References

Berlusconi, G. (2016). Social Network Analysis and Crime Prevention. Springer eBooks. https://doi.org/10.1007/978-3-319-27793-6_10.

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.

Brantingham, P., & Brantingham, P. (2021). Crime Pattern Theory. Oxford Research Encyclopedia of Criminology. Oxford University Press.

Caccavale, F. (2013). Informatica e diritto, XXXIX annata, Vol. XXII, 2013, n. 1, pp. 169-179.

Corina, G., Brittany, N. F., Yu-Hsuan, K., Hongjian, W., Zhenhui, L., & Daniel, K. (2021). Network Spillovers and Neighborhood Crime: A Computational Statistics Analysis of Employment-Based Networks of Neighborhoods. Justice Quarterly, 38(2), 344-374. https://doi.org/10.1080/07418825.2019.1602160.

Goldstick, J. E., & Jay, J. (2022). Agent-Based Modeling: an Underutilized Tool in Community Violence Research. Current Epidemiology Reports, 9(3), 135-141. https://doi.org/10.1007/s40471-022-00292-x.

Gong, Y., Dai, M., & Gu, F. (2023). CARESim: An integrated agent-based simulation environment for crime analysis and risk evaluation (CARE). Expert Systems with Applications, 214, 119070. https://doi.org/10.1016/j.eswa.2022.119070.

Oliveira, M. (2021). More Crime in Cities? On the Scaling Laws of Crime and the Inadequacy of Per Capita Rankings—A Cross-Country Study. Crime Science, 10.

Sierra-Arévalo, M. (2015). Social Network Analysis and Gangs. The Handbook of Gangs, 157–177. https://doi.org/10.1002/9781118726822.ch9.

Wright, C. R. E. (2021). Network Spillovers and Neighborhood Crime: A Computational Statistics Analysis of Employment-Based Networks of Neighborhoods. Justice Quarterly, 38(2), 344-374. https://doi.org/10.1080/07418825.2019.1602160.