Journal of Complex Networks, 2025, **4**, cnaf016
https://doi.org/10.1093/comnet/cnaf016
Advance access publication 4 July 2025
Article

OXFORD

Exploring the link between urban topology and street crime using complex networks: a case study from Southeast Brazil

Matheus de Andrade Flausino[1,*], Eric Araújo[2], and Angélica Sousa da Mata[3]

[1]Computer Science Department, Universidade Federal de Lavras, Caixa postal 3037, Lavras, MG CEP:37200-900, Brazil
[2]STEM, Computer Science Department, Calvin University, 3201 Burton SE, Grand Rapids, MI, 49546, United States
[3]Physics Department, Universidade Federal de Lavras, Caixa postal 3037, Lavras, MG CEP:37200-900, Brazil

[*]Corresponding author. Computer Science Department, Universidade Federal de Lavras, Caixa postal 3037, Lavras, MG CEP:37200-900, Brazil. E-mail: matheusdeandradeflausino@gmail.com.

ABSTRACT

Crime represents a complex social challenge, and understanding it is a crucial step for implementing effective measures to mitigate this problem. Public security agencies in the state of Minas Gerais, in the southeast of Brazil, currently document the spatiotemporal details of crimes, particularly street crimes, in their reports, enabling comprehensive analysis of these incidents. The aim of our study is to examine, using the provided data by the Military Police of Minas Gerais, the relationship between the urban topology of cities of similar size and the occurrence of street crimes. These cities also share the presence of a medium-sized university population, driven by the presence of significant federal university hubs in each of them. An approach using complex networks and their centrality measures is employed to understand the spatial concentration of street crimes in these cities over the period from 2014 to 2022. The results show, among other findings, that the closeness centrality measure exhibits the highest correlation with areas of criminal concentration. These findings may offer valuable insights into the potential application of complex network analysis for informing public policies and urban planning strategies aimed at enhancing public safety and addressing street crime in similar urban contexts.

KEYWORDS: complex networks; spatial analysis; urban topology; street crime.

1. INTRODUCTION

The advancement of geoprocessing technologies in recent years has made it possible to record more precise information about crime occurrences by law enforcement agencies, particularly street crimes [1]. Diverse and numerous data have been collected by various capture technologies applied to specific fields to generate new perspectives on criminal behavior and initiate new criminal justice practices. Some criminologists and law enforcement agencies seek to explore these capabilities to reach general or specific diagnostic conclusions about the connections between crime and

Received: 18 December 2024. **Editorial decision:** 28 May 2025. **Accepted:** 12 June 2025

geography, criminal trends and personal traits/history, or between past and (predicted) future crimes [2].

Criminology has already presented several models, including computational ones, to try to explain crime occurrences and the factors involved in this social phenomenon. These models range from biological aspects to social elements, as described, e.g. by Gerritsen [3], who formalizes and analyzes different theories in the criminological spectrum, and by Baal [4], who assesses the influence of sanctions on criminal deterrence.

Understanding crime patterns is an important step in developing effective crime prevention policies and strategies. With the increasing availability of data and the modernization of law enforcement, there is great potential for advancements in criminological studies, supported by real data, allowing for more accurate adjustments and the development of more realistic models. Furthermore, the results of this type of work can be widely used to support decision making by public security agencies, which are always managing limited human and material resources for their operations.

A growing body of research has applied tools from statistical physics and complex network theory to understand the spatial and temporal dynamics of urban crime. In their comprehensive review, D'Orsogna and Matjaž [5] explore crime phenomena through the lens of self-exciting processes, game theory, and network structures, highlighting the broad sociological implications of such approaches. Building upon this foundation, several studies have proposed methods to represent and analyze urban environments as complex networks, aiming to detect criminal hotspots and understand crime propagation mechanisms [6–8]. These works emphasize the potential of street-level network modeling to uncover patterns of criminal behavior and support data-driven public security policies.

Motivated by these advancements, we provide a focused overview of key contributions that further explore the relationship between street network structure and crime incidence. Studies such as [9, 10] propose novel methodologies to link network centrality measures with spatiotemporal crime patterns, revealing the influence of physical infrastructure and social context on criminal activity. Together, these findings highlight both the promise and the limitations of current analytical frameworks. In this context, we summarize key studies and advancements in the field, identifying gaps in existing knowledge, and highlighting the specific contributions of our work to enhancing understanding and improving practices within this area of study.

Jiang *et al.* [11], transform urban topology abstraction into a *named-street*, where nodes represent entire streets, and edges represent the connections each street has with others. In some way, is approach provides an adequate representation of street connectivity. However, it loses the geographic aspect, due to the fact that intersections are eliminated from the representation, which is of extreme importance for our work.

Mao *et al.* [12] conducted a literature review on the relationship between the urban environment and crime, considering the urban environment at three spatial levels: road networks, streets (paths), and nodes. The nodes here refer to the representation of facilities in cities, such as bus stops, hotels, etc. Thus, this literature review is closely related to our work. According to Mao *et al.* [12], morphological elements included in existing research can be divided into two groups: geometric and topological. The former deals with easily describable characteristics of morphological composition, such as density, length, width, and hierarchy, while the latter deals with underlying characteristics of composition such as permeability, accessibility, and connectivity. Our research falls into the category of studies observing the influence of topological elements on crime occurrences.

Yue *et al.* [13] presents a methodological approach similar to ours: it uses the concept of betweenness centrality as a measure of street permeability. This concept of permeability was applied locally, considering a 500 meters radius for pedestrian journeys and a 5000 meters radius for driver journeys. Thus, betweenness centrality is not considered for all points in the graph, only from one point i to another point j within the previously described radius areas. Finally, the authors combine the calculated permeability with other variables, such as the percentage of unemployed, to calculate the relationship with crimes using the binomial regression method. The results show that local permeability is negatively correlated with crime (count), i.e. the greater the observed permeability, the lower the expectation of crimes.

Our research distinguishes itself from the main works mentioned above in several aspects. Firstly, our study relies exclusively on crime records that, despite their differences—such as theft of valuables from passersby and vehicle or vessel theft—were all documented as having occurred on the streets. Additionally, we assume that representing urban roads as complex networks may reveal structural characteristics that help to understand crime occurrences. While many studies in the literature develop methodologies based on the topological features of urban roads from a managerial perspective, such as categorizing streets as arterial, collector, or local [14–16], but our approach emphasizes purely structural measures of centrality within the network. In particular, unlike Yue *et al.* [7], we do not directly calculate street permeability nor integrate socioeconomic factors such as unemployment into the analysis. Instead, we focus on how different types of network centrality correlate with crime distribution, without incorporating external variables like economic status or mobility patterns.

To this end, we implemented a complex network approach, using data from the state of Minas Gerais, in the southeast of Brazil, provided by the Military Police of Minas Gerais (PM-MG) to evaluate street robbery crimes committed in medium-sized cities in Minas Gerais that have a significant university population presence. Our hypothesis is that the urban structure of cities can be modeled as complex networks and that analyzing metrics associated with the obtained data can provide valuable insights into the locations of crime. We explore the characteristics of geographical distribution, represented as complex networks and the occurrence of crimes. Our main objective is to examine how the topological characteristics of the cities in Minas Gerais influence their attractiveness to offenders. Thus, our goal is to address the following research questions (RQ):

- RQ1: How can urban structure be represented as a complex network, and how does it relate to the geographic distribution of crime occurrences?
- RQ2: Is there a correlation between the road network and the spatial clustering of crime occurrences in the cities under study?

We hope that our findings will provide valuable insights into the potential use of complex network analysis to inform public policies and urban planning strategies focused on improving public safety and tackling street crime in similar urban contexts. The paper is structured as follows: Section 2 outlines the methodology, Section 3 presents our results, and, finally, Section 4 provides a discussion.

2. METHODOLOGY

This research used data on street crimes from eight cities of Minas Gerais state, located in southeastern Brazil: Alfenas, Itabira, Itajubá, João Monlevade, Lavras, Ouro Preto, São João Del Rei, and Viçosa. This data was provided by the PM-MG.

These cities were selected because they have similar populations, allowing scientific comparisons between them. The municipalities are considered similar due to their federal university hubs and significant student populations. Quantitatively, these cities have between 70 000 and 110 000 inhabitants and are all the municipalities with the largest populations in their immediate geographic region. All these municipalities are located in the countryside of the State of Minas Gerais. Table 1 presents some demographic data from the Brazilian Institute of Geography and Statistics (IBGE) for the cities under study [17].

Population size is generally the most considered dimension in public policies and sciences, with many studies attempting to define what constitutes small, medium, and large cities [18]. There is a somewhat arbitrary threshold where various studies group cities with more than 100 000 inhabitants into a different category from those with fewer than 100 000 inhabitants. In this scenario, there are two cities in this study that would belong to a different group from the others: Lavras

Table 1. IBGE demographic data for the cities under study covering the years 2022 to 2023

Cities	Population	Territorial Area	Density	GDP (R$)[a]
Itabira	113 343	$\sim 1254\,\text{km}^2$	$\sim 90\,\text{inh/km}^2$	56
Lavras	104 761	$\sim 565\,\text{km}^2$	$\sim 186\,\text{inh/km}^2$	25
Itajubá	93 073	$\sim 295\,\text{km}^2$	$\sim 316\,\text{inh/km}^2$	34
São João Del Rei	90 225	$\sim 1452\,\text{km}^2$	$\sim 62\,\text{inh/km}^2$	28
João Monlevade	80 187	$\sim 99\,\text{km}^2$	$\sim 809\,\text{inh/km}^2$	41
Alfenas	78 970	$\sim 850\,\text{km}^2$	$\sim 93\,\text{inh/km}^2$	34
Viçosa	76 430	$\sim 299\,\text{km}^2$	$\sim 255\,\text{inh/km}^2$	22
Ouro Preto	74 824	$\sim 1246\,\text{km}^2$	$\sim 60\,\text{inh/km}^2$	50

[a]GDP per capita per year [2020] in R$1000. Source: [17].

and Itabira. However, since the 100 000 inhabitants value is arbitrary and these cities have populations very close to 100 000, we will consider here that all the cities are part of the same population group. The data on street crimes are from the same historical period, covering the years from 2014 to 2022, and apply to all the municipalities mentioned above.

2.1 Data description

The database was provided by the PM-MG and consists of a spreadsheet as street crime records. It is important to emphasize that only this type of crime was made available for this project. Each row in the file represents an incident, and initially, there are 8364 records of these crimes in the database. For each record, there are 35 attributes (columns) with relevant and specific information about the occurrence.

The more relevant attributes are *Longitude* and *Latitude*. As the objective of this work is to assess and compare the geographic concentration of crime occurrences with the characteristics of the street structures in the analyzed cities, these two pieces of information are essential. Therefore, all records in the database lacking Latitude or Longitude information were discarded. In total, 453 records with invalid latitude or longitude coordinates were identified in the database, representing 5.42% of the total records.

For the execution of the research, only the information about longitude, latitude, and municipality of occurrence is necessary. However, additional information from the records can be used to visualize some characteristics of the offenses. These information can be seen in the Supplementary Material.

2.2 Network structure

A central part of research development is having the complex network representation of the road network for each municipality. The road networks consist of the streets, avenues, and highways that are part of each municipality.

The methodological framework adopted in this study was entirely conceived and implemented by the authors. The project was primarily developed using the Python programming language and its libraries. For the creation of each municipality's graph, geographic information of the urban roads was collected from the OpenStreetMap [19] platform using the OSMnx [20] library and converted into graphs. The OSMnx library automatically processes the raw topology of OpenStreetMap data so that the edges of the graphs represent street segments and the nodes represent intersections between segments or even dead-ends.

The *graph_from_point* method from the OSMnx library is responsible for returning a graph that represents the area specified in the parameters. To define the location of each municipality, three parameters were used:

Figure 1. The road topology of Alfenas as a graph.
Alt text: Spatial graph of the road topology in Alfenas, Brazil.

Table 2. Table with the central points of each municipality

City	Latitude	Longitude
Itabira	−19.644693	−43.228580
Lavras	−21.244698	−44.996468
Itajubá	−22.426959	−45.463703
São João Del Rei	−21.123135	−44.245778
João Monlevade	−19.820029	−43.154901
Alfenas	−21.420327	−45.948580
Viçosa	−20.748481	−42.878345
Ouro Preto	−20.391875	−43.502574

- *center point*: The central point around which the graph will be constructed. It is a tuple with latitude and longitude values.
- *dist*: The radius size, in meters, from the central point that will determine the entire area in which the graph will be constructed. That is, all elements of the road network that are within the distance (*dist*) from the central point will be included in the graph construction.
- *network_type*: The type of road network to be considered. In our case, drivable public roads will be returned.

For the *dist* parameter, we used a value of 5000 meters for all municipalities. This value was sufficient to satisfactorily cover the urban roads of all municipalities. Furthermore, it was possible to construct an area where the vast majority of occurrences in each municipality could be geographically encompassed. Figure 1 shows the complex network representing the city of Alfenas. The *center point* value was obtained using the *Google Maps* tool. For each municipality, we arbitrarily chosen a central point, i.e. a geographical point that should be the gravity center of each city, and its latitude and longitude values were collected. The central point of each city can be seen in Table 2.

2.3 Measures and strategies

From the networks of each city, two distinct types of graphs were constructed: *street graph* and *occurrence graphs*. Street graph refer to all graphs that represent the street network of each city. From these graphs we will take some characteristic of the network such as degree centrality, closeness centrality, betweeness centrality, eigenvector centrality, PageRank, and distance to the university. Thus, each city will have *one* street graph and *six* measurements taken.

All of these measures are well-known in the literature [21], except for the distance to the university, which will be addressed here. Leveraging the university characteristic of the municipalities in this study, information on the distance to the university was added, which involves calculating the distance from each vertex (intersection) in the graph to the city's university. Only the main campus or hub of the federal university of the municipality was considered. Details on the construction of these graphs are explained below. On the other hand, occurrence graphs refer to all graphs that represent some strategy for incorporating crime occurrences into the network of each municipality.

To calculate the various centrality measures of the networks for each municipality, the *networkx* library [22] was used. The OSMnx library [20], previously mentioned for extracting each city's network from the OpenStreetMap platform, is built on top of the *networkx* library, and therefore, the obtained networks are already fully compatible with *networkx*, requiring no additional conversion or configuration. *Networkx* already includes functions for calculating closeness centrality, degree centrality, betweenness centrality, eigenvector centrality, and PageRank, simplifying development. Another important library that we used is the *haversine* that contains the implementation of the *haversine* formula, which calculates the distance between two distinct points on Earth using their latitude and longitude values. It is a special case of a more general formula in spherical trigonometry, the law of haversines, relating the sides and angles of spherical triangles [23]. This library is very useful, since all methods for calculating occurrences in the graphs depend on calculating the distance between two points on Earth. More detailed descriptions of these tools are given in the Supplementary Material.

Here, we will describe all three distinct strategies for incorporating occurrences into the networks of each city. The first strategy to incorporate occurrences into the networks of each city is: for all occurrences of crime in a city, one must find the street intersection that is geographically closest to each occurrence, and then add the value of an occurrence to the vertex representing that intersection. So, the goal of this first method is to associate an occurrence with the closest intersection or vertex.

The second strategy is very similar to the first one, with the addition of counting an occurrence not just for the closest intersection but for the four closest intersections, in a weighted manner. There are two objectives for this strategy. The first one is to increase score for the intersections that are also very close to occurrences but are not necessarily the closest and often end up not counting a single occurrence in the *first strategy*. In this way, this approach registers a score not only for a single intersection but also for the immediate area of crime. Therefore, the association of occurrence with intersections (vertices) is also modified. Each intersection will receive a fraction of the entire occurrence, with the closest intersection receiving 40% (0.4), the second 30% (0.3), the third 20% (0.2), and the fourth 10% (0.1), totalling 100% of the occurrence. We will refer to this as weighted association. The second objective is related to how the association of occurrences and intersections happens in this method. Since intersections receive fractions of the total occurrence value, statistical possibilities are expanded, allowing intersections to have more distinct values.

The third and final method is very similar to the second. It uses the same mechanism of ordering the four closest intersections but it differs from the second method in one part: the calculation of the intersection-occurrence association. The goal here, in addition to the objectives mentioned in the second method, is to transform the weighted association into a continuous distribution. The formula used for this is an inverse proportion, meaning that the farther the intersection is from the occurrence, the smaller its association value with that occurrence. This association can be calculated as follows:

$$\text{continuous\_association} = \frac{1000}{\text{distance\_between\_occurrence\_and\_intersection}}$$

The value 1000 was chosen arbitrarily, as it is only necessary to perform the inverse proportion of the distance between the occurrence and the intersection, i.e. the farther away, the less influence the vertex receives from the occurrence. There is an expectation that the continuous distribution of association between intersections and occurrences will expand the correlational statistical possibilities between the occurrence graphs and the measurements from the street graphs.

Finally, it is important to mention that we used these three distinct approaches to ensure the robustness of our results in the face of variations that may be somewhat subjective regarding the choice of the location of the occurrences.

2.4 Correlation measure

We performed the correlation between the six measurements in the street graph and the three occurrence graphs, resulting in *eighteen* associations for each municipality. It is worth noting that correlation coefficients do not imply causality. Nor it is the objective of this research to attribute causality to the urban structure of municipalities in crime occurrences, but rather to investigate what other characteristics can help us understand and explain the phenomenon of street crimes, even if partially.

The most commonly used correlation coefficient for statistical analysis is Pearson's coefficient [24]. However, one of the assumptions of this technique is that the variables have a continuous nature. As presented earlier, we have three strategies for incorporating occurrences into the graphs, and only the third yields a continuous variable; the others are of a discrete nature. Thus, we discarded this technique as it is not recommended in all our scenarios.

This limitation does not apply to Spearman's rank correlation coefficient [24], and therefore, we will adopt this technique. Another advantage of using Spearman's rank correlation coefficient in our research is that, unlike Pearson, Spearman's coefficient does not require a strictly linear monotonic relationship. The present study, in turn, does not require a linear relationship between crime occurrences and network measures, so Spearman's coefficient appears to be a more suitable technique.

2.5 Concentration of occurrences

The concentration of occurrences can also be described in the form of numerical data. According to what has been discussed so far, occurrences will be associated with the closest vertex (in some of the approaches presented, an occurrence may even be associated with more than one vertex). Thus, it is possible to present a relation of occurrences per vertex to enhance understanding of their concentration within the urban structure.

We will refer to *total vertex concentration* (TVC) as the ratio of the number of vertices with at least one associated occurrence to the total number of vertices that make up the city's urban structure. This ratio can be described as follows:

$$TVC = \frac{num\_vertex\_with\_occurrence}{num\_total\_vertex}$$

Finally, we will also refer to *total occurrence spread* (TOS) as the ratio of the total number of recorded occurrences to the total number of vertices that make up the city's urban structure, associating each occurrence with a single vertex and each vertex with a single occurrence. This way, we can understand how well the total occurrences "cover" the city's structure.

$$TOS = \frac{num\_total\_occurrences}{num\_total\_vertex}$$

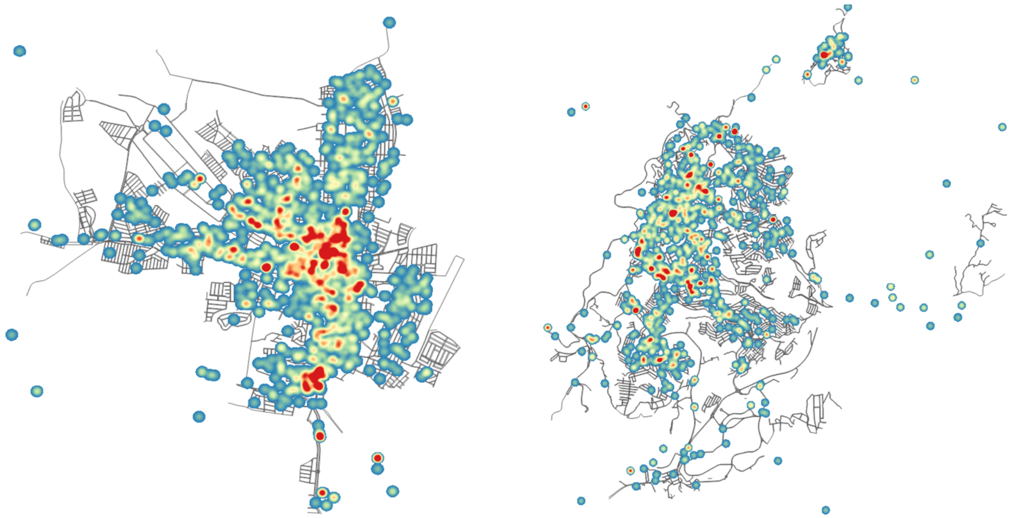

Figure 2. Heat map of crimes in Alfenas (left panel) and Itabira (right panel).
Alt text: Heat maps showing the spatial distribution of crimes in Alfenas (left) and Itabira (right), Brazil.

3. RESULTS

The occurrence of crimes is often observed as being spatially concentrated, as suggested by Rosser *et al.* [25]. This phenomenon has even been adopted as a basis for predictive models of crime occurrence. Thus, the first step in this work is to visualize the spatial concentration of crime occurrences in each of the cities of this study, during the period from 2014 to 2022.

Moreover, the spatial visualization of crime occurrences will be important as it will give us an advance indication of how the occurrence graphs for each city should look. Constructing these occurrence graphs is a significant step in our methodology. To allow for the visualization of occurrences, we used the *QGIS* tool (see Supplementary Material), and the municipal layers were obtained through the *OpenStreetMap* platform.

In Fig. 2 there are images representing the visualization of crime occurrences in two cities, Alfenas and Itabira, as an example. We used the heat map technique for visualization because it helps us identify which regions of each city are experiencing higher concentrations of crime occurrences through color indications. Regions with colors closer to red are areas with higher crime occurrence intensity, while regions closer to blue are areas with fewer crime records. Areas without colors are those with no crime records. The radius used to determine the clustering of occurrences, i.e. to decide the concentration of occurrences at a given point, is 50 meters on the real scale of the map. This 50-meter radius was chosen arbitrarily for practical visualization purposes only, and it does not affect the network construction or any of the statistical analyses.

Table 3 consolidates the TVC and TOS (2.5) values for each of the involved cities and illustrates the spatial concentration of crime occurrences. The goal was to evaluate whether occurrences were indeed being "clustered" in some vertices or regions. Thus, the lower the TVC value and the higher the TOS value, the greater the concentration occurring in that municipality. The cities of São João Del Rei and Lavras, which have the lowest crime concentration, are also those with the fewest occurrence records which implies that the concentration of crime can be better observed as the amount of crime also increase. Therefore, larger cities with a greater number of crime occurrences can be more useful in this type of analysis.

As previously mentioned, the greater the difference between TOS and TVC, the higher the concentration of crimes according to the methodology used in this study. Table 4 presents the difference between TOS and TVC for each city, along with the demographic information presented

Table 3. Result of TVC and TOS calculations for the cities

City	TVC	TOS
Itabira	23%	49%
Lavras	10%	15%
Itajubá	15%	26%
São João Del Rei	7%	9%
João Monlevade	29%	57%
Alfenas	29%	63%
Viçosa	22%	68%
Ouro Preto	20%	43%

Table 4. Comparison of crime concentration along with GDP and population

City	TOS—TVC	GDP(*)	Population	Density
Viçosa	46	22	76 430	255 inh/km^2
Alfenas	34	34	78 970	93 inh/km^2
João Monlevade	28	41	80 187	809 inh/km^2
Itabira	26	56	113 343	90 inh/km^2
Ouro Preto	23	50	74 824	60 inh/km^2
Itajubá	11	34	93 073	316 inh/km^2
Lavras	5	25	104 761	186 inh/km^2
São João Del Rei	2	28	90 225	62 inh/km^2

in Table 1. We observe that the increase in crime concentration is not accompanied by an urban escalation, either in terms of GDP per capita or population size.

3.1 Occurrence graphs and measurement graphs

Before presenting the analysis of the spatial correlation between crime occurrences and centrality measurements from in the street graphs, it may be useful to visualize each of them. The goal of this visualization is also to visually identify the regions where there is a higher concentration of the observed phenomenon. In this case, the phenomena are both the strategies adopted to assimilate the occurrences into the city graph and the network characteristics (centrality measures) that represent such municipality. Due to the large number of city data, we show here the visualization of Alfenas networks. The visualization of the other network cities can be found in the Supplementary Material. A visual comparison between the areas of higher centrality (Fig. 3) and the regions with greater crime concentration (Fig. 4) can be observed.

3.2 Correlations

We calculated the correlations for each of the municipalities involved in the study. The variables quantifying the association of crimes in the first two strategies (see Section 2.3) have a discrete nature and low value diversity. For these reasons, the use of *Spearman's* correlation, which is a robust non-parametric technique for these data characteristics. This study uses conventional stratification for interpreting the Spearman correlation coefficient. This convention is shown in Table 5.

In Table 6, we show the Spearman correlation coefficients for each municipality. The notation first, second and third refer to the three strategies that we used to incorporate the crime for each node of the networks as explained in Section 2.3. And each column refers to the centrality measures: degree(DG), closeness(CL), betweeness (BW), eigenvector (EV), and pagerank (PR) centrality, and the last column corresponds to the distance from the university (UD).

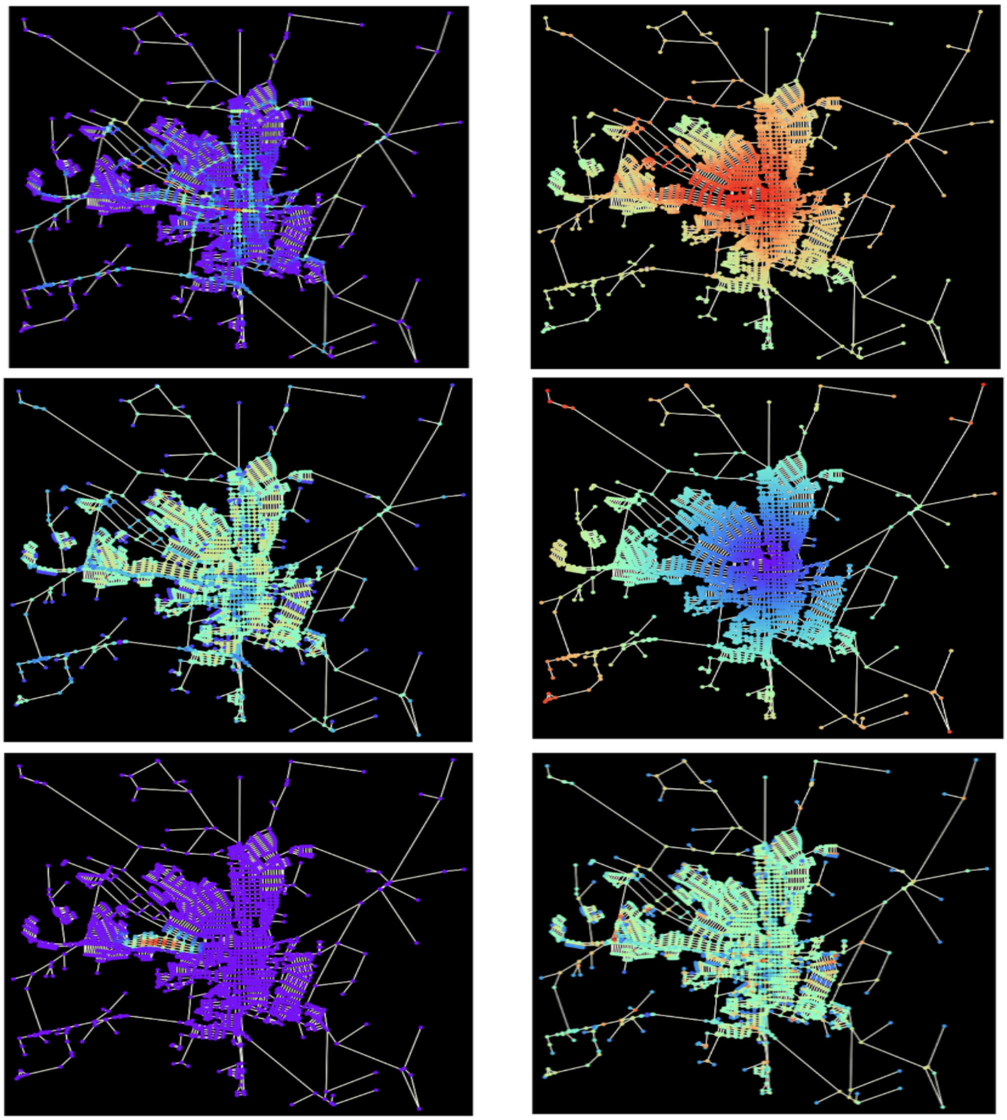

Figure 3. From the left to the right and from the top to the bottom: Betweeness centrality, Closeness centrality, Degree centrality, Distance to university, Eingenvector centrality and Page Rank centrality for Alfenas. Areas in shades of red represents the higher values of the evaluated measure calculate in each picture while areas in shades of blue represents the lower values of the assess measure.
Alt text: Six colored maps of Alfenas, Brazil, showing spatial distributions of betweenness, closeness, degree, eigenvector, and PageRank centralities, and distance to the university. Red areas indicate higher values; blue areas indicate lower values.

Considering the conventional approach for interpreting Spearman's correlation coefficients, we observe that in Itabira all correlation coefficients showed a negligible or weak magnitude of correlation. However, it should be noted that the closeness centrality and distance from the university measures had the highest magnitude of correlation.

In Lavras, no measure reached a moderate correlation. However, the closeness centrality measure presents the highest magnitude of correlation among all network characteristics analyzed.

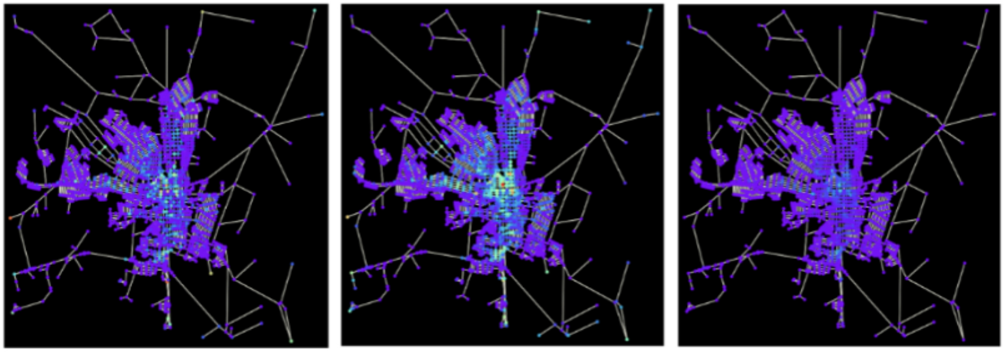

Figure 4. From the left to the right: First, second and third strategies applied in Alfenas, respectively. Areas in shades of red represents the higher values of the evaluated measure calculate in each picture while areas in shades of blue represents the lower values of the assess measure.
Alt text: Three colored maps of Alfenas, Brazil, showing the spatial distribution of the first, second, and third strategies. Red areas indicate higher values; blue areas indicate lower values.

Table 5. Table with conventional interpretation for correlation values[a]

Absolute value of the coefficient	Interpretation
0.0–0.1	Negligible correlation
0.1–0.39	Weak correlation
0.4–0.69	Moderate correlation
0.7–0.89	Strong correlation
0.9–1.0	Very strong correlation

[a]Conventional intepretation from Patrick Schober and Lothar A. Schwarte [24].

Additionally, the eigenvector measure in Lavras stands out by showing a correlation score higher than most of the other municipalities.

In São João Del Rei, all measures were classified as weak or negligible according to the conventional interpretation approach. However, it is important to highlight that the closeness centrality measure shows the highest correlation value, almost reaching a moderate interpretation in the second and third strategies.

In Itajubá, most correlation coefficients were considered weak or negligible. However, the closeness centrality in the second and third strategies showed moderate correlation according to conventional criteria. The distance from the university also showed a correlation close to moderate in the second strategy and achieved moderate correlation in the third strategy.

In João Monlevade although the closeness centrality measure showed the highest correlation magnitude among the network characteristics, all scores were very low, not allowing for significant correlations beyond what is considered weak or negligible in this municipality.

In Alfenas the Page Rank measure has negligible correlation with crime occurrences in all adopted strategies. Meanwhile, the degree, betweenness, and eigenvector centrality measures have a weak correlation with crime occurrences in all strategies. On the other hand, we note that the closeness centrality and distance from the university measures show a weak correlation in the first strategy but a moderate correlation in the other strategies. Thus, we find that these measures are the most correlated with crime incidence in the city of Alfenas. These results are somewhat expected, as when analyzing the heat map of criminal occurrences presented at the beginning of this Section, showing

Table 6. Spearman correlation coefficient for all cities

City	Strategy	DG	CL	BW	EV	PR	UD
Itabira	First	0.03	0.16	0.06	−0.02	−0.01	0.13
	Second	0.02	0.26	0.11	−0.05	−0.01	0.22
	Third	0.03	0.32	0.13	−0.03	0.00	0.24
Lavras	First	−0.01	0.20	0.10	0.12	−0.01	−0.10
	Second	−0.02	0.32	0.16	0.18	0.01	−0.14
	Third	−0.01	0.34	0.17	0.20	0.01	−0.16
Itajubá	First	0.01	0.24	0.17	−0.01	0.03	−0.22
	Second	0.00	0.39	0.24	0.01	0.01	−0.36
	Third	0.01	0.44	0.25	0.03	0.01	−0.41
S. J. Del Rei	First	0.00	0.20	0.12	−0.08	0.00	−0.15
	Second	−0.01	0.34	0.16	−0.12	−0.01	−0.26
	Third	0.00	0.35	0.16	−0.11	−0.01	−0.27
J. Monlevade	First	0.03	0.09	0.12	0.07	0.00	0.03
	Second	−0.01	0.14	0.18	0.09	−0.01	0.05
	Third	0.00	0.19	0.19	0.10	−0.01	−0.01
Alfenas	First	0.14	0.35	0.18	0.21	0.06	−0.34
	Second	0.15	0.49	0.22	0.30	0.05	−0.47
	Third	0.16	0.53	0.22	0.35	0.05	−0.53
Viçosa	First	0.03	0.26	0.14	0.07	0.05	−0.16
	Second	0.03	0.34	0.18	0.10	0.05	−0.15
	Third	0.06	0.37	0.20	0.14	0.05	−0.23
Ouro Preto	First	0.01	0.35	0.20	0.04	−0.01	−0.15
	Second	0.00	0.48	0.23	0.05	−0.01	−0.19
	Third	0.03	0.52	0.26	0.07	0.01	−0.24

concentration of crimes in the central area of the city. Additionally, it is important to highlight that the Federal University of Alfenas is located in this same central region of the city.

The municipality of Viçosa followed a similar pattern to São João Del Rei, where all measures were classified as weak or negligible. Notably, the closeness centrality measure shows the highest correlation value, almost achieving a moderate interpretation in the second and third strategies.

In Ouro Preto, we observe a moderate correlation for the closeness centrality measure in both the second and third strategies. Additionally, in the first strategy, a value close to a moderate correlation is observed. The other measures have values considered weak or negligible.

4. DISCUSSION

In this research, we used data from the PM-MG regarding street crime occurrences to evaluate street robbery crimes committed in medium- sized cities of Minas Gerais state that have a significant university population presence. Using complex network measurements, we constructed graphs that could be related to the previously obtained occurrences, with the aim of identifying spatial concentration areas where these crimes occur within each city. For answering our *RQ1* we developed three strategies, discussed in our methodology section, to include the occurrences over complex networks that represent each city.

We assumed that the centrality measures of complex networks calculated from the road network of each municipality could be a technique that helps in understanding the phenomenon of spatial concentration of crimes. Using statistical techniques, we correlated the regions where crimes occurred in historical data with the regions of highest centrality in the road network. In the following paragraphs, we will discuss the correlation obtained between the road network and the spatial clustering of crime occurrences, addressing our *RQ2*.

The results have shown that the centrality measure that best correlates with crime occurrence regions is closeness centrality. This measure was the best in seven out of the eight cities in this study, being less significant only in the city of João Monlevade. Furthermore, closeness centrality was the only centrality measure that achieved a moderate correlation with crime regions, observed in the cities of Alfenas, Itajubá, and Ouro Preto. Still, in cities like São João Del Rei, Viçosa, and Itabira, the correlation magnitude was close to being considered moderate according to the conventional approach for interpreting Spearman correlation. The closeness centrality measure identifies the nodes best positioned within the network by evaluating their proximity to all other nodes. In essence, nodes with high closeness centrality have the shortest average distances to other nodes [26].

In the cities, these intersections nodes with high closeness centrality scores are generally associated with central areas. The central areas of cities are often those with high flows of people and commerce. These characteristics make these areas attractive for crime occurrences, as pointed out by Xiang *et al.* [27]. However, drawing the conclusion that areas with high closeness centrality necessarily correspond to central urban regions in the cities studied requires further detailed investigation. If this correlation is confirmed, it could significantly streamline criminological analyses, given that calculating closeness centrality for urban structures is a relatively simple task. In future studies, it would be important to gather more detailed data regarding the central and commercial areas of the cities, aiming to establish correlations between regions of high closeness centrality and areas with greater economic activity and population movement. Additionally, a complementary analysis exploring the relationship between crime occurrences and local socioeconomic indicators could provide deeper insights into spatial crime patterns. We acknowledge, however, that such analyses are often constrained by the quality, availability, and granularity of socioeconomic data, and that they may also be subject to potential biases inherent to the interpretation of social and economic inequalities. Nonetheless, we consider this a promising direction for future research, in line with recent findings that emphasize the role of urban morphology in shaping neighborhood safety [28] and the role of per capita rankings and wealth distribution in the scaling of crime concentration across cities [29–31].

We also incorporated the measure of distance to the university, which, as previously mentioned, assigns a score to each node based on its proximity to the state university within the respective city. This measure was included because all the municipalities are classified as "university cities." Notably, in some cities, such as Alfenas and Itajubá, this measure exhibited stronger (negative) correlation magnitudes, even reaching levels that could be considered moderate correlations. In cities with significant correlation magnitudes for the distance to the university, universities are located in regions of high closeness centrality. However, in cities with lower correlation magnitudes, universities are situated outside these high centrality areas as happen in Ouro Preto. This pattern underscores the varying relationships between crime hotspots, centrality, and university location within urban environments. This observation is noteworthy as it suggests that crime occurrences may not be inherently linked to the student population in these cities. Instead, the university's location could influence the perception of safety, potentially contributing to an improved sense of security among students.

The degree centrality measure showed negligible correlation values in all strategies and cities observed, which is likely due to the homogeneity of road networks concerning degree, while crime occurrences tend to be heterogeneous, with higher concentration in certain network regions. Other measures showed very weak correlation magnitudes, some even negligible, which is why we did not develop any further analysis on them. Page Rank and degree centrality were considered negligible in all three strategies and all cities observed. For eigenvector centrality, 50% of its correlation with crime regions in municipalities was negligible, while the other 50% was weak. For betweenness centrality, almost all (92%) was considered a weak correlation, very close to a negligible magnitude. In Ref. [32], the authors show that correlations between centrality measures, such as betweenness and closeness, can vary significantly across different network types. Although a positive correlation is common in many contexts, our analysis indicates that in our case, the correlation between betweenness and closeness is not consistently high across cities, likely due to the nearly grid-like structure of the street networks.

During the study, we identified several factors that could pose a threat to the validity of the research. The primary concern relates to the strategy used to integrate occurrences into the complex network. To address this, we employed three distinct strategies, each carefully designed to achieve specific objectives. These strategies were chosen for their coherence and methodological soundness, ensuring that the study was as robust and reliable as possible. However, it is important to note that these approaches are unique to this research, and alternative methods for incorporating occurrences into the graph might yield different results.

The second factor is related to the data. We believe that a dataset encompassing a larger number of occurrences, particularly in larger cities such as metropolitan areas, could provide stronger insights into the correlation between crimes and network centrality measures. A higher volume of occurrences would likely offer a clearer picture of the spatial concentration of crime. Additionally, we underscore the ethical considerations inherent in this type of research. Simplifying complex social issues, such as crime incidence, carries the risk of perpetuating biases and potentially exacerbating the problem, as highlighted by O'Neil et al. [33]. This emphasizes the importance of carefully interpreting results and acknowledging the broader social context in which these findings are situated.

The goal of this study was to identify a variable or factor that could contribute to a better understanding of criminal occurrences in the analyzed cities, with potential applicability in predictive models. However, it is important to note that this work is based on historical data provided by the PM-MG. As such, the results may reflect patterns inherent to the way the data was recorded and reported. This highlights the need for careful interpretation of the findings and consideration of the broader ethical implications associated with this type of research.

In summary, regarding *RQ1*, we represented the urban structure as complex networks using three complementary strategies to map crime occurrences onto the road network of each city. Regarding *RQ2*, we found a moderate positive correlation between crime spatial clustering and closeness centrality in most cities studied, indicating that crimes tend to occur in areas that are more central and accessible within the urban network structure.

For future work, it would be valuable to include larger cities with higher numbers of recorded crimes, as we identified notable trends in concentration as the number of crimes increases within a city. It would also be important to reassess the strategies for incorporating crime into the network. These aspects are crucial for determining the strength and direction of the correlations and should accurately reflect real-world dynamics as closely as possible.

ACKNOWLEDGEMENTS

The authors thank the Military Police of Minas Gerais (PMMG) for providing the data.

SUPPLEMENTARY DATA

Supplementary data is available at *COMNET Journal* online.

FUNDING

Financial Support from CNPq/FAPEMIG grant no. APQ-06591-24.

DATA AVAILABILITY

All the code developed for this study is publicly available at: complex-network-crimes.

REFERENCES

1. Tom-Jack QT, Bernstein JM, Loyola LC. The role of geoprocessing in mapping crime using hot streets. *ISPRS Int J Geo-Inf* 2019;**8**:540.
2. Smith GJD, Bennett Moses L, Chan J. The challenges of doing criminology in the big data era: towards a digital and data-driven approach. *Br J Criminol* 2017;**57**:259–74. https://doi.org/10.1093/bjc/azw096
3. Gerritsen C. *Caught in the Act: Investigation Crime by Agent-Based Simulation*. 2010, Vrije Universiteit Amsterdam, Amsterdam, Netherlands.
4. van Baal P. *Computer Simulations of Criminal Deterrence*. Den Haag: Erasmus Universiteit Rotterdam (EUR), Den Haag, The Netherlands, 2004. 171 p. ISBN 90-5454-443-0.
5. D'Orsogna MR, Perc M. Statistical physics of crime: a review. *Phys Life Rev* 2015;**12**:1–21.
6. Spadon G, Scabora LC, Araujo MV *et al.* Complex-network tools to understand the behavior of criminality in urban areas. In: Latifi, S. (eds) *Information Technology-New Generations. Advances in Intelligent Systems and Computing, vol 558: 14th International Conference on Information Technology*. Cham: Springer, 2018, 493–500. https://doi.org/10.1007/978-3-319-54978-1_63
7. Ferreira A, Rubiano G, Mojica-Nava E. Urban security analysis in the city of Bogota using complex networks. In: Morales, A., Gershenson, C., Braha, D., Minai, A., Bar-Yam, Y. (eds) *Unifying Themes in Complex Systems IX. ICCS2018. Springer Proceedings in Complexity. Springer, Cham,* 2018, 424–38. https://doi.org/10.1007/978-3-319-96661-8_44
8. Garcia-Zanabria G, Gomez-Nieto E, Silveira J *et al.* Mirante: a visualization tool for analyzing urban crimes. In: *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. Porto de Galinhas, Brazil: IEEE, 2020, 148–55. https://dx.doi.org/10.1109/SIBGRAPI51738.2020.00028
9. Garcia-Zanabria G, Raimundo MM, Poco J *et al.* CriPAV: street-level crime patterns analysis and visualization. *IEEE Trans Vis Comput Graph* 2022;**28**:4000–15.
10. Kim Y-A, Hipp JR. Pathways: examining street network configurations, structural characteristics and spatial crime patterns in street segments. *J Quant Criminol* 2020;**36**:725–52.
11. Jiang B, Claramunt C. Topological analysis of urban street networks. *Environ Plann B Plann Des* 2004;**31**:151–62. https://doi.org/10.1068/b306
12. Mao Y, Yin L, Zeng M *et al.* Review of empirical studies on relationship between street environment and crime. *J Plan Lit* 2021;**36**:187–202. https://doi.org/10.1177/0885412221992280
13. Yue H, Zhu X, Ye X *et al.* Modelling the effects of street permeability on burglary in Wuhan, China. *Appl Geogr* 2018;**98**:177–83.
14. Al Mushayt NS, Dal Cin F, Barreiros Proença S. New lens to reveal the street interface. A morphological-visual perception methodological contribution for decoding the public/private edge of arterial streets. *Sustainability* 2021;**13**:11442.
15. Marshall S, Jones P, Plowright I. A framework for classification and assessment of arterial streets. In: *European Commission Fifth Framework Programme Key Action: City of Tomorrow and Cultural Heritage*, London, UK, 2004.
16. Sarkar C, Webster C, Kumari S. Street morphology and severity of road casualties: a 5-year study of Greater London. *Int J Sustainable Transport* 2018;**12**:510–25. https://doi.org/10.1080/15568318.2017.1402972
17. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Panorama de Minas Gerais e Cidades: Censo Brasileiro de 2020. IBGE, Rio de Janeiro, Brazil, 2020.
18. Bitoun J. *Tipologia Das Cidades Brasileiras e Políticas Territoriais: Pistas Para Reflexão*. FASE, Observatório das Metrópoles; Rio de Janeiro, Brazil, 2009.
19. Haklay M, Weber P. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Comput* 2008;**7**:12–8.
20. Boeing G. OSMnx: a Python package to work with graph-theoretic OpenStreetMap street networks. *JOSS* 2017;**2**:215.https://doi.org/10.21105/joss.00215
21. Sousa da Mata A. Complex networks: a mini-review. *Braz J Phys* 2020;**50**:658–72. https://doi.org/10.1007/s13538-020-00772-9
22. Hagberg AA, Schult DA, Swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. In: Proceedings of the 7th Python in Science Conference, Pasadena, USA, 2008. https://dx.doi.org/10.25080/TCWV9851
23. Robusto CC. The cosine-haversine formula. *Am Math Monthly* 1957;**64**:38–40.
24. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg* 2018;**126**:1763–8. https://doi.org/10.1213/ANE.0000000000002864
25. Rosser G, Davies T, Bowers KJ *et al.* Predictive crime mapping: arbitrary grids or street networks? *J Quant Criminol* 2017;**33**:569–94. https://doi.org/10.1007/s10940-016-9321-x
26. Centeno EGZ, Moreni G, Vriend C *et al.* A hands-on tutorial on network and topological neuroscience. *Brain Struct Funct* 2022;**227**:741–62. https://doi.org/10.1007/s00429-021-02435-0

27. Xiang L, Sheng J, Liao P. Understanding the relationship between the spatial configuration and the crime rate of Downtown Eastside in Vancouver, Canada. *Habitat Int* 2023;**137**:102847.
28. Benita F. Associations between street connectivity and poverty. *Netw Spat Econ* 2022;**22**:181–201.
29. Bettencourt LMA, Lobo J, Strumsky D *et al.* Urban scaling and its deviations: revealing the structure of wealth, innovation and crime across cities. *PLoS One* 2010;**5**:e13541.
30. Oliveira M, Bastos-Filho C, Menezes R. The scaling of crime concentration in cities. *PLoS One* 2017;**12**:e0183110.
31. Oliveira M. More crime in cities? On the scaling laws of crime and the inadequacy of per capita rankings—a cross-country study. *Crime Sci* 2021;**10**:27.
32. Valente TW, Coronges K, Lakon C *et al.* How correlated are network centrality measures? *Connections (Toronto, Ont.)* 2008;**28**:16–26.
33. O'Neil C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.