Taking Care of Everyone's Business: Interpreting Sicilian Mafia Embedment through Spatial Network Analysis*

Michele Battisti † Andrea Mario Lavezzi‡ Roberto Musotto§

January 31, 2022

Abstract

Mafia-type organisations often have a strong geographical and cultural entrenchment in the territory they belong. However, their analysis as a spatially networked social structure is still missing. A combined socio-spatial network analysis is presented here, through the demise of a large police operation called *Operazione Perseo* in 2008. This approach is developed in two ways. At first, a visual representation of the social network of this large group of *mafiosi* embedded in a geographical space is presented. Three main salient territorial features of the network are thus highlighted. A high density of links in some neighbourhoods, as well as connections across different *Mandamenti*, the territorial units where Mafia families operate, and the correlation of links and socio-economic determinants, like the unemployment rates. Secondly, a spatial econometric analysis of centrality measures of the group is suggested here. Findings show a positive spatial correlation in the Eigenvalue centrality scores.

Keywords: Spatial Network Analysis, Social Networks, Organised Crime, Sicilian Mafia, Cosa Nostra, Spatial Regressions.

---

*We are grateful to three anonymous referees for comments and criticisms. Usual caveat applies.

†Dipartimento di Giurisprudenza, Università di Palermo, piazza Bologni - Palermo, Italy.

‡Dipartimento di Giurisprudenza, Università di Palermo, piazza Bologni - Palermo, Italy.

§Security Research Institute, School of Science, Edith Cowan University, Joondalup Drive - Perth, Western Australia.

1 Introduction

The classic idea of the Sicilian Mafia is the one of a criminal organisation with its hands or tentacles over anything and anyone that dares to cross its path in the areas that it controls. Is it true? And if it is true, how does it take place? This paper proposes an empirical analysis of the internal structure of the Sicilian Mafia by looking at how space is related the organization of criminals themselves to answer these questions.

Specifically, a spatial network analysis of a large group of members of *Cosa Nostra* is suggested here.[1] The group was operating in the town and province of Palermo, Sicily, identified by the police operation named *Operazione Perseo*. The novel aspect of this analysis with respect to existing social network analyses of organised crime is the geo-localisation of members of the criminal network, which allows to consider geographical space as an additional variable affecting the way connections are created.

Space is, indeed, a fundamental aspect for the functioning of the Sicilian Mafia, but the role of space in shaping its networked structure still needs to be fully analysed in the literature. Space for *Cosa Nostra* matters for at least two reasons. First of all, the control of the territory is one of its crucial goals (see, e.g. Lavezzi, 2008). Secondly, the Sicilian Mafia operates through a sharp division of the territory into *Mandamenti*, i.e. areas in which one or more families rule (Paoli, 2003, p. 45), without major interferences of other groups within the organisation. Therefore, spatial network analysis is much needed to add another layer of understanding on how such a control of the territory takes place. The result is to show not just how members are intertwined one to the other, as standard social network analysis does, but also how these connections can be characterised geographically.

In particular, in this paper we give two novel contributions. First of all, we provide a visual localisation in the geographical space of the identified network, along the lines suggested by Andris (2016). This is followed by a spatial econometric analysis of the network. Specifically, the spatial econometric analysis exploits information about the geographical distance of criminal members to identify spatial effects in the characterisation of their network centrality. This approach follows Mastrobuoni and Patacchini (2012), as it studies how individual characteristics correlate with measures of network centrality. Mastrobuoni and Patacchini (2012), however, study the American Mafia and do not consider spatial effects.

The main results of this article are the following. First of all, the embedment of the social network structure is shown through the identification of a high density of contacts that take

[1] *Cosa Nostra* is the most popular name that identifies the criminal Mafia-type organisation operating in Sicily, Italy. Along the paper the names of Sicilian Mafia and *Cosa Nostra* are used interchangeably.

place locally, confirming the importance of the *Mandamento* as the main place where *mafiosi* hold their contacts and interests. However, connections across *Mandamenti* highlight the necessity of some level of coordination for the well-functioning of the organization as a whole. Secondly, we find that the Eigenvalue centrality scores are positively spatially correlated. This means that members of the organisation that are connected to popular individuals tend to be located close in space. From a policy perspective, this result offers a novel point of view. The territorial embedment of the Mafia is the result of a perception of "being in charge" by its members that stick together, pick members from the same area they live in and victimise those that are the closest. It follows that policing the Sicilian Mafia should not look too far either when it is about finding other affiliates, in particular those with a high number of connections, and therefore a high knowledge of the network.

The paper is structured as follows: Section 2 describes the related literature and clarifies the contributions of this article; Section 3 presents the dataset and briefly outlines the main features of a spatial network analysis (Section 3.1); Section 4 presents the visualisation of the network on a map and the information that can be derived from this approach; Section 5 contains the results of the econometric analysis; Section 6 concludes.

2 Related Literature: Social Networks, Space and Crime

This paper is related to a vast and heterogenous body of literature. Three main areas can be identified: works that focus on the relation between space and crime; those on the social network analysis of organized crime, which include some studies of the Sicilian Mafia as well, and those works that introduce space into an analysis of social networks, proposing a spatial network analysis of crime. A succinct overview of these areas is provided in this section and it is clarified how this article relates to them.

Spatial analysis of crime is a well-established field of research. Its methodologies are detailed in Rossmo (1999), Chainey and Ratcliffe (2013), and Santos (2016). Some specific topics that are close to this paper include the analysis of crime hot spots, i.e. specific locations where a relatively high number of crimes occur (see, e.g., Levine, 2006, Chainey et al., 2008, Eck and Weisburd, 2015); the study of routine activities, i.e. whether a specific place attracts or inhibits crime during various moments of the day or over time (see, e.g. Felson and Cohen, 2008); the Risk Terrain modeling, where socio-economic spatial variables and past crimes can predict the occurrence of other crimes in one area (see, e.g. Dugato et al., 2020), and evaluations of how space can affect criminal behaviour, such as spatial proximity to other criminals (see, e.g., Calvó-Armengol and Zenou, 2004, for details and references). With respect to this strand of literature, this paper is not strictly concerned with the occurrence of offences or co-offences and the related role of space. This paper aims to shed light on the relationship between the geographical location of a member of a criminal organization (proxied by its address, see Section 3 for details) and its position into the network structure of the criminal group.

Social network analyses of organised crime include, among others, Morselli (2009), Varese (2012a,b) and Calderoni (2012). The work of Morselli (2009) shows how to analyse a criminal network, discusses several case studies and addresses a number of issues in understanding networks' characteristics, such as how personal connections are important to climb within the organization or the security-efficiency trade-off. Varese (2012a,b) provides an in-depth social network analysis of a Russian criminal group operating in Italy, shedding light on its topological characteristics (e.g. degree distribution, centrality), its organizational features (also by a content analysis of wiretapped conversations), and its capacity to take advantage of globalization by relocating to foreign countries. Finally, Calderoni (2012) offers a similar analysis of criminal groups involved in drug trafficking, belonging to *'Ndrangheta*, a powerful criminal organization originating in the Italian region of Calabria. Calderoni (2012), in particular, finds that the criminals most involved in the drug trafficking operations are not always the most central in the network, offering a useful perspective in the definition of the optimal disruption strategies that an external authority can implement (see also Calderoni and Superchi, 2019, on leadership in the *'Ndrangheta*). While offering remarkable insights, these studies are not explicitly focused on the Sicilian Mafia as this work, and do not consider the role of space in criminal networks.

The network structure of the Sicilian Mafia has been analysed in some recent works. Agreste et al. (2016) study a large *Cosa Nostra* network operating in North-Eastern Sicily. They identify two types of networks: one based on wiretapped conversations, and one based on different sources of information (bank transactions, co-offences, etc.). Interestingly, they find that the topological properties of the two networks are very different, and derive implications for policies aimed at disrupting the networks (see Bichler, 2019, Ch. 8, for a detailed discussion on mapping networks from different data sources). Cavallaro et al. (2020) analyse data from the operation "Montagna", which disrupted a *Cosa Nostra* group operating in the province of Messina around 2007, bringing to the arrest of 39 people. The focus of that work is on disrupting policies, pointing out that targeting agents with the largest betweenness centrality (see Section 5 for a discussion of centrality indices) is an effective policy as it strongly reduces the largest connected component of the network. Calderoni et al. (2020) consider the same dataset of Cavallaro et al. (2020), focusing on the problem of link prediction. Findings show that global measures of the network topology (such as the Katz centrality score) perform well in predicting links. Tumminello et al. (2021) study a large dataset from recent investigations on the Sicilian Mafia. They find that Mafia syndicates have a strong territorial characterization, as syndicate members are likely to live in the same territory, that Mafia members become more and more specialized as their criminal career proceeds and, finally, that women, individuals that are not part fo the criminal group, often play a crucial role in connecting different syndicates. None of these works, however, considers the role of space, as it is done in the present article.

Finally, few works carried out a spatial network analysis of organised crime. In particular, Tita and Radil (2011) study violence among gangs in Los Angeles, and suggest different definitions of distance beyond geographical proximity as "neighbour" gangs do not need to be geographically close. By looking at different spatial weight matrices, proximity among gangs based on their rivalry can indeed reveal interesting patterns of spatial influences in the occurrence of homicides (see Faust and Tita, 2019, for further discussion and references on the spatial network analysis of gangs' violence). Scalia (2020) analyses the the high rise building construction industry between 1950 and 1980 in Palermo operated by the Sicilian Mafia, making the point of how land fragmentation facilitated Mafia dominance. This work, however, takes a different perspective from ours, as it focuses on the organisation of a crime by a criminal organisation, while this paper is centered on the organisation of criminals, i.e. on the spatial network structure of the Sicilian Mafia. To the best of our knowledge, this is the first work that carries out a spatial network and econometric analyses of a Mafia-type criminal organization such as *Cosa Nostra*.[2]

3 Dataset and Methods

Data were extracted from the arrest warrant of the so-called *Operazione Perseo*,[3] a police operation that took place in December 2008 to disrupt a large criminal group, mostly operating

[2]The *Cosa Nostra* network in the social sciences has been considered as an expression of power against the State, as a result of cultural shifts in the society (Ruggiero, 2019). This "cultural context" includes families, friends and those legal actors which intersect the web of power, criminal and legal interests that surround the organisation (Allum et al., 2019; Sciarrone, 2019). The way cultural shifts have been influencing the Mafia network over time are evident. While in the early years of the XX century, *Cosa Nostra* thrived in an agricultural based society and a much stereotyped culture of silence (Sciarrone, 2019), the anti-mafia movements have perched on the intricate balance of power and crime of the Sicilian society at large (Santino, 2015; Santoro, 2021). Although not directly related to these strands of literature, the approach of this paper can nonetheless be seen as complementary to these ones.

[3]There are multiple reasons why the arrest warrant is a good document for this type of analysis. In the current Italian criminal procedure system, it conveys all the evidence in favor and against each person charged of a crime. It also provides a less biased legal document if compared to other official legal and judicial documents, that is helpful in building a network with a discrete reliance on validity and reliability (see, e.g.,Musotto, 2020).

in Palermo, that tried to establish a new *Cupola*, i.e. the highest hierarchical level of *Cosa Nostra*.[4] The operation covers a two-year period, from late 2006 to 2008. *Operazione Perseo* resulted in the arrest of 99 people, the largest number of arrests against *Cosa Nostra* in recent years (see Figure 1) in the district of Palermo. In particular, *Operazione Perseo* is the second biggest trial on Mafia after the *Maxiprocesso* ("maxi-trial") of 1986 for number of arrests.[5] *Operazione Perseo* hit almost all the families and turfs in the province of Palermo. This operation therefore offers an ideal opportunity to investigate how the organisation was networked over the territory.

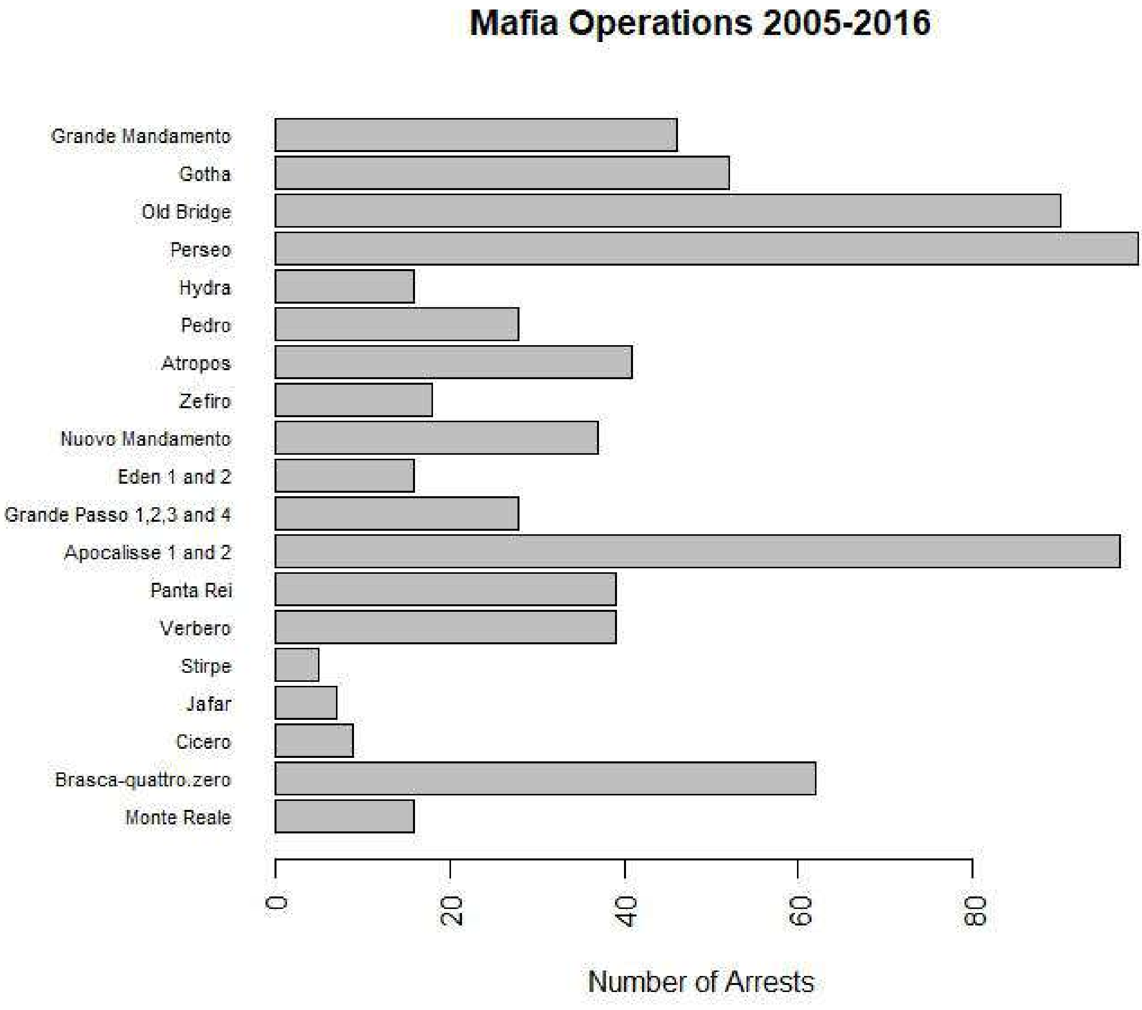

Figure 1: Police operations against the Mafia in the province of Palermo, 2005-2016

[4]Research has been started after ethical clearance, which has been granted in 2015 by the Università degli Studi di Messina.

[5]The *Maxiprocesso* was the first, and so far unparalleled, large-scale trial against the Sicilian Mafia, that resulted in the arrest of 474 *mafiosi*.

In order to map the network (Bichler, 2019), the arrest warrant was manually examined in order to identify the nodes of the network (i.e. the individuals involved in the operation), and their links, i.e. the connections among them. The total number of individuals identified is 172, but the document reports the address, a crucial piece of evidence for geo-localization, only for the 99 individuals arrested in the operation. For this reason, the present work focuses only on these 99 agents.[6]

To identify a link between any two nodes of the network, every piece of information on a connection between any agent was extracted and recorded from the document. The aim with this procedure is to minimise biases in the detection of links as, for example, resorting only to wire-tapped conversation, which can underestimate the number of links as some agents, notably Mafia bosses, tend to avoid such type of communication (see Bichler, 2019, Ch. 9, for a thorough discussion of this aspect). Specifically, it has been assumed that a link between two agents exists if: i) prosecutors pointed at it in the arrest warrant;[7] ii) there is documented evidence of a connection between two individuals (i.e. a reported wire-tapped conversation, a connection mentioned in a recorded conversation or in other parts of the document, the joint participation to a meeting, etc.); iii) a direct (first-degree) or indirect (second-degree) family tie is documented. This procedure allowed to collect 1410 pieces of evidence on connections, and to identify from them 470 links among the 99 agents.[8] In the next section we highlight the main feature of a spatial network analysis that can be carried out on such data.

3.1 Spatial and Social Networks Methods

Spatial networks are those networks 'for which the nodes are located in a space equipped with a metrics (Barthélemy, 2011, p. 3). In these types of networks, the space dimension is added in

[6] In this article, as in other works that resort to court documents, the extent of the network is limited by the evidence collected by the investigations. As such, it includes cursorily attention to individuals outside the criminal network who, nonetheless, have interactions with the core group of *mafiosi* under arrest. In particular, it may miss the connections with individuals formally outside the organization, but still involved in its operations. Sciarrone (2019) defines this as the "grey area" between the criminal sphere of *mafiosi* and the legal sphere, in which legal actors such as professionals, entrepreneurs, etc. interact with *mafiosi* to make profits and/or to provide specialized skills to the organization (see also Lavezzi, 2008, p. 179). The investigation of this interesting aspect is left for future work.

[7] The first part of the arrest warrant presents a summary of the group activities, pointing out the main connections of individuals arrested, as resulting from the whole investigation.

[8] In the vast majority of cases, the amount of information available was not sufficient to identify either the direction or the intensity of the connection. For this reason, the network is assumed as undirected and unweighted. In addition, the network identified is static, because the available information did not allow for a rigorous identification of the dynamic dimension of network formation over time.

order to assess, for example, if the probability of a link between two nodes is inversely related to the distance. In terms of the relationship between crime and space this suggests that there are situations where the organisation of crimes or the organisation of criminals is affected by distance (Smith, 1986).

While the spatial dimension of these networks does not mean necessarily that they have to be set up in a geographical space, as pointed out by Tita and Radil (2011), geographical space is one of the various ways in which distance can be expressed and quantified. The advantage of this technique applied to networks is that spatial information gives more accurate information on the strength, density and distance of nodes (Barrat et al., 2005) compared to non-spatial ones. This is so because strong ties in a criminal organisation could be related to the fact that affiliates are geographically closer. This type of analysis has two main uses: to explain (e.g. Papadias et al., 2003) values of specific variables in the network or to predict values of variables (e.g.Meyers et al., 2005) in order to clarify the potential dynamics of the network.[9] This kind of analysis allows for effects of scale in dispersed networks while giving context as to where nodes operate.

Mixing social networks with spatial networks is still an uncharted theoretical territory. The seminal methodological contribution of Alizadeh et al. (2016)[10] explains how to transform classic-style networks such as random, small-world and scale-free networks into spatial ones. In order to create a spatial network it is assumed that nodes have geographical coordinates,[11] such as a latitude and a longitude or that there is a way to gather these coordinates (e.g. an address, a dead reckoning, etc). Therefore, given an $(m \times m)$ Cartesian space, nodes are placed in a plan. Afterwards, nodes are connected following the modeling of the network. For example, two nodes are connected if they reach a specific radius (in the case of a random network), if they are the closest to a node (in the case of a small-world network) or if nodes are close and well connected with others (in the case of a scale-free network). However, social networks in a geographical space might not follow these rules at all and connect if they have one or more interest in common, such as committing crimes together (Humphreys, 2010).

An adjacency matrix is then needed to build the network, as in standard network analysis.[12] However, to gather spatial information one more matrix is needed, i.e. the distance matrix, showing the length of the shortest path between two vertices (Anselin, 2013), with a diagonal

[9] Examples of this kind can be found, for example, in the traffic models or pipe networks and in the predictive model of SARS and Covid-19 (e.g. Franch-Pardo et al., 2020).

[10] See also Musotto (2017) for an early application of a similar method on organized crime.

[11] That is if the intent is to create a social network embedded in a geographical map. Otherwise there must be some other types of coordinates that allow nodes to be placed in a definable space, such as an euclidean space.

[12] See Musotto (2021) for a step-by-step guide on how to build a spatial social network.

equal to zero, as the distance between one node and itself is zero. One challenge for this type of analysis is to find a programming language and, additionally, a software where social network features would coexist with geographical measures.[13]

To sum up, Spatial Network Analysis is a specific set of techniques where a graph is placed in a metric space and distance functions are applied to get more information about the nodes and the links. A spatial network brings more information than a regular network. Once a network is placed into a space, it gives a more accurate description on the strength, distance and density of the nodes (Andris, 2016). In the next section spatial network analysis will be applied to the network emerging from *Operazione Perseo*, i.e. a Mafia-type criminal network. We will first of all focus on building the network on a map, on the variables that have been considered, and on whether there appear "spatial effects" from the visual representation.

4 *Operazione Perseo* on a Map

Spatial network analysis requires that nodes are geometrically located in a space. The *Operazione Perseo* network as it has been created and analysed in Battisti et al. (2021) does not allow inferences about the way it is embedded in the area of Palermo district or about neighbours of each node.[14] Figure 2 from Battisti et al. (2021) represents the social network structure in a standard way, i.e. without reference to the geographical space.[15]

Figures 3 and 4, from the annual report of DIA presented in Ministero dell'Interno (2016), show how Palermo and its districts are subdivided into different *Mandamenti*.[16] DIA investiga-

[13]While the spatial network is sourced from judicial documents and coded with R language and *ggmap* package (Kahle and Wickham, 2013), network visualisation presents the additional challenge of creating a network that is not simply laid over a geographical map, but that it is embedded in it. This has been achieved with the software CartoDB and the use of SQL language, as it has not been possible to achieve a satisfactory result with R (see Section 4 for details).

[14]The only way in which space is considered in Battisti et al. (2021) is by considering the Euclidean distance among nodes of the network as a possible determinant of the existence of a link between them in the estimation of dyadic regressions.

[15]Battisti et al. (2021) highlight the main topological characteristics of the *Perseo* network, as represented in Figure 2. In particular, the network has few interactions with respect to the maximum possible number (i.e. it has a low density), but within the same sub-group connections are frequent (i.e. the clustering coefficient, i.e. the average fraction of agents' contacts which is also connected, is relatively high), and it is relatively quick to reach out to members that are at the opposite side of the network (i.e. the diameter and average path length, respectively indicating the largest distance between any two nodes and the average value of the shortest path between any two nodes, are short). In other words, the *Perseo* network is a small-world network (Watts and Strogatz, 1998). See Battisti et al. (2021) for more details.

[16]DIA (*Direzione Investigativa Antimafia*) is the Italian agency specialised in countering organised crime.

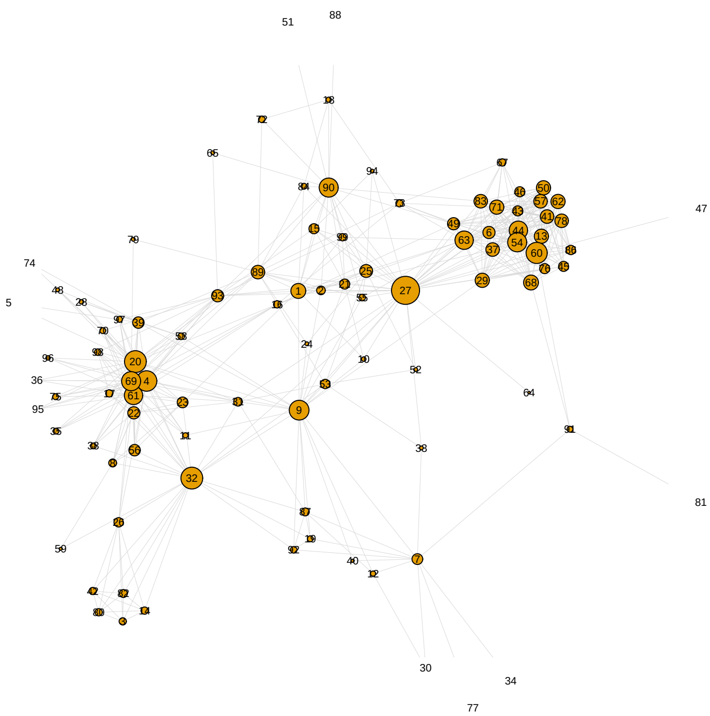

Figure 2: The *Perseo* Network from Battisti et al. (2021)

.

tions usually show a tree graph that draws connections inside and outside the organisation.[17] Is there a way to combine the two graphs? The analysis conducted here shows how this result can be achieved through spatial coordinates of affiliated members.

Figure 3: Mandamenti in Palermo municipality. Source: Ministero dell'Interno (2016)

Figure 4: Mandamenti in Palermo province. Source: Ministero dell'Interno (2016)

[17]An example is not shown here as it would reveal personal information.

The procedure suggested here follows three steps:

1. extraction of the relational data from the data source;

2. creation of the map;

3. juxtaposition of geo-localised nodes on the map according to the edge list.

As anticipated in the previous section, multiple approaches have been combined and softwares used to create the network map. Nodes and geographical coordinates have been analysed with *R* to begin with. Afterwards the dataset has been applied to different geographical visualization softwares: ArcGIS, GoogleEarth and Carto DB, to code the map and the network together. These softwares use a different language: Structured Query Language (SQL) and geographic information system (GIS), which have been applied to manipulate relational and geographical data. Spatial cooordinates were extracted from individuals' addresses, as reported in the judicial document.

The challenge of designing a criminal spatial network carries two major issues: working around software limitations and decide which type of information to display. To this regard three different software, languages and techniques have been considered: the function `geocode()` from the package `ggmap` in *R*, geocoding with Google Maps and overlaying the network in *CartoDB*. The main objective was to embed a social network in a map that would not be just a static picture, but that it could be explored, studied and that could be flexible in the type of information that it would display.

The first two approaches proved unsatisfactory for a number of limitations.[18] The choice made for this article was therefore to utilize the software *CartoDB*. The network has been rebuilt from a link dataset and the spatial information for each ID. The advantage of this software was in the fact that it required SQL language to be build and one of its functions `Thegeomwebmercator` is common in web mapping systems. Therefore it was possible to integrate every point in the

[18]In particular, the main problem of using `geocode()` was the incompatibility between the shape file for the town of Palermo, released in 2018, and the dataset created in *R*. On the other hand, using Google Maps API (now Google Maps Platform), proved problematic for the following reasons. Google Maps API functionalities allowe to connect one user or location to the other, crating a map, for example, of different crimes. Several layers could be added, so that it would be possible to distinguish the type of crime or combine time information with the network. While it has been possible to connect one node to another, this mainly worked in tracing routes from one address to the other. Another attempt was made by overlaying a geolocated network of addresses over the map. While this result was closer in spirit to creating a network of *mafiosi* over the geographical map, it still presented multiple disadvantages. Address location was not precise and areas of the network or subgroups could not be focused on. The map and the network *de facto* behaved like a picture over another, with all the disadvantages coming from the fact that the network overlay was not embedded in the map.

map with other features and overlays, including a shapefile for each zipcode in the Palermo area. There are several features in the following images that made the choice of the software preferred. Links from a node to the other could be overlayed and they appear darker in color where this happens. A density map could be added with a detail precision of about 100 meters to show how close some of these members where. By including the shapefile of zipcodes for the district of Palermo, it has been possible to visually integrate macro information about some socio-economic indicators in the different districts, e.g. unemployment levels. In what follows we illustrate the results of the application of this procedure.[19]

Figure 5 is based on a heatmap representation of the density of links among the network's nodes in the Palermo province. It clearly highlights a territorial clustering, with highly-connected clusters of individuals in the city of Palermo and in the the surrounding towns of Belmonte Mezzagno, Bagheria, Montelepre, Altofonte, Monreale, San Giuseppe Jato and San Cipirello. Darker red areas show a higher clustering of affiliated members. This is higher in the city of Palermo, which is the largest one in the west of Sicily. However, there are also many links between Palermo and other smaller centers where there is a very high clustering too. Some of these villages present a very high clustering of *mafiosi* too, such as Belmonte Mezzagno, which in turn looks more significant because there is a smaller number of people inhabiting the more rural areas. While affiliates are highly clustered and linked within their own town and village, the fact that there are many connections between different villages suggests that while the bulk of Mafiosi socialize within their area, links with other clans, families and *Mandamenti* are important too.[20]

[19]While many works in the literature use GIS to visualise social networks (see for example Andris, 2016), our network has been visualised through web mapping, through the use of SQL language, discarding GIS. Three reasons for this choice. Available packages in GIS, in lay terms, do nothing more than plotting a picture of an area with one or more layers added to it. While granularity and network features could be changed, they did force to plot a new graph every time. The network was small, but it still carried a significant amount of information that would have lengthened the process considerably. Second, there was a need to be able somehow to manipulate the data and its visualisation instantly so to be able to highlight relevant areas and part of network, this is only achievable through web mapping applications. Third, visualisation of the network should account for a certain degree of exploration of the network and area itself, which was easily achievable with the other softwares and languages chosen.

[20]This result is consistent with the findings on the territorial proximity of members of the same syndicate of Tumminello et al. (2021).

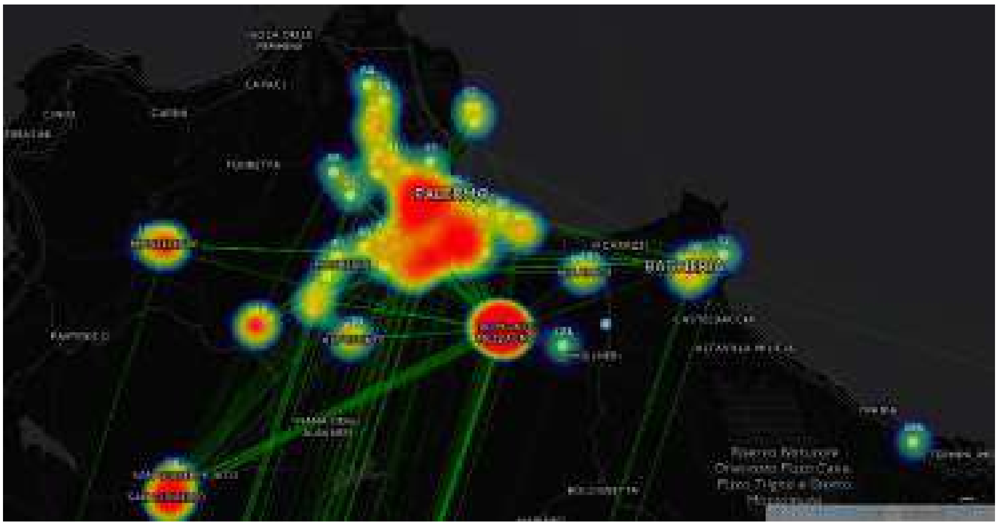

Figure 5: Network Visualization in the province of Palermo of the *Perseo* network with *Carto*.

Figure 6 presents a picture of the distribution of connections among members of the *Perseo* network on the city of Palermo, in which areas with darker colors represent higher unemployment levels, while lighter ones show lower unemployment.[21] It is possible to notice that many *mafiosi* are located in areas with medium to high levels of unemployment in Palermo.[22] This is a visual confirmation that the Mafia thrives in areas where unemployment is higher and, thus, poorer (Becucci, 2011).[23]

[21] Unemployment data come from the 2011 Census, the most recent available census data for Italy.

[22] With the exception of the ZEN neighbourhood, a district notoriously plagued by the presence of organized crime. The justification is that a previous investigation *Hydra* heavily hit that *Mandamento*. During *Operazione Perseo* investigation, Mafia bosses agreed to share the control of that area until someone could take control of it.

[23] A thorough analysis of the socio-economic conditions of the areas in which network members are located is beyond the scope of this paper and remains an interesting topic for further research.

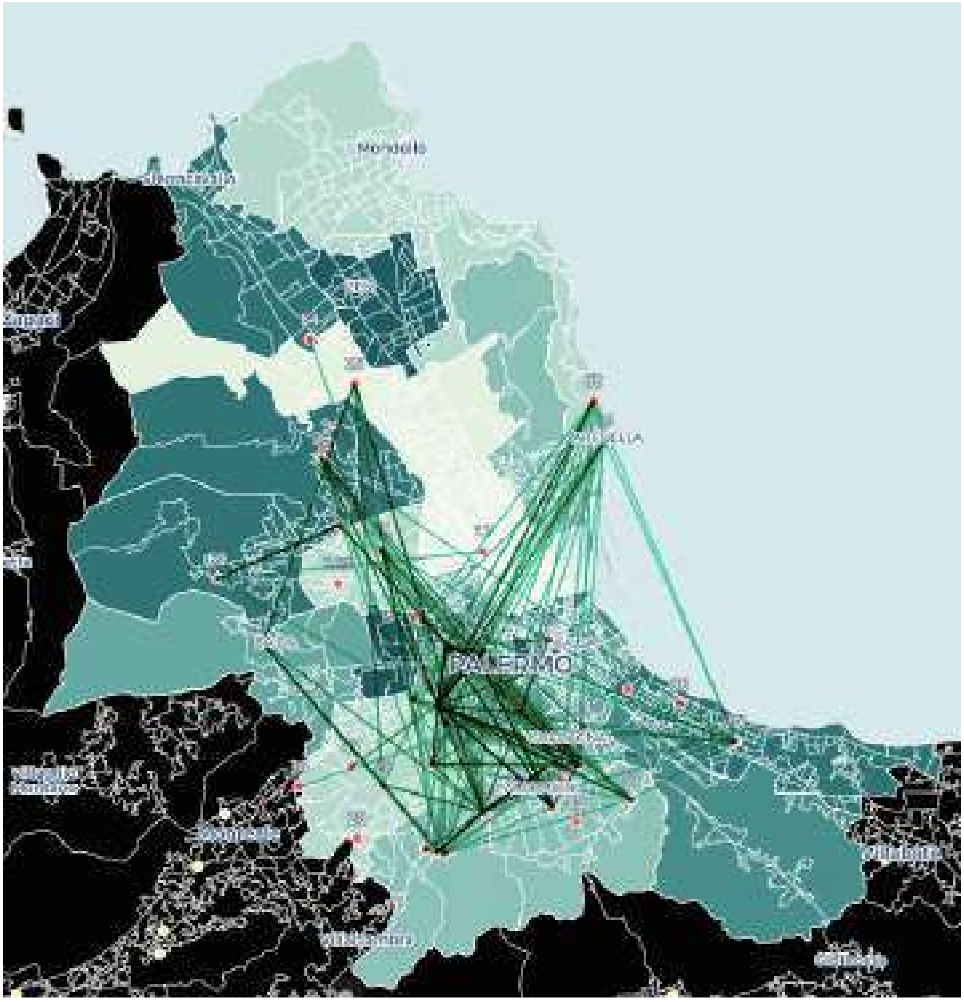

Figure 6: Network visualisation in the city of Palermo of the *Perseo* network with *Carto*.

Figure 7 shows a smaller portion of the city where the majority of links have been detected. From Figure 7 it is possible to draw some remarks just from the visualisation of the network as it is embedded in the map. First of all, it is possible to notice that nodes are heavily clustered together, but the network is not planar.[24] This is especially true for those that are located near the city centre.

Secondly, there are many links among affiliates that run not just within their own *Mandamento*, but also around town. Figure 7 shows that it is possible to recognise and classify different sub-groups among the affiliated members, according to their geographical location. This is more evident in smaller towns around Palermo, where actually most of the streets had

[24]A node that is allowed to be connected with others that are not the closest ones means that the spatial network is not planar. A classical example of a planar network is a road or pipe network because each node is connected in a way that link intersections (i.e. one link going over the other) are not possible (Barthélemy, 2011).

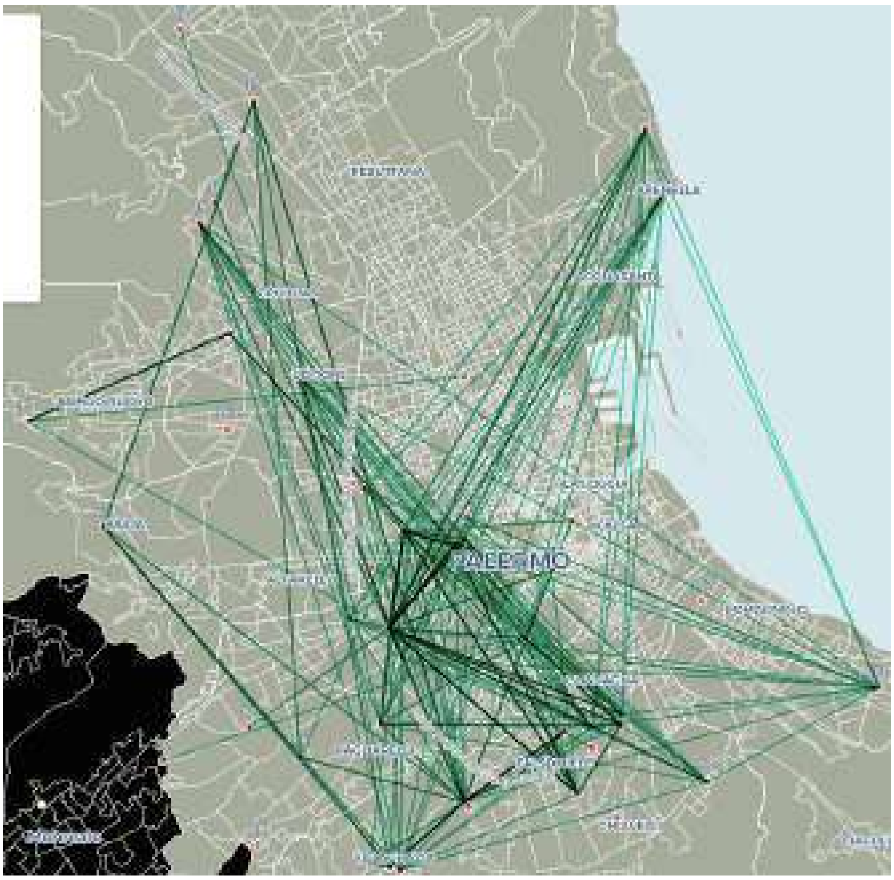

Figure 7: Network Visualisation in the city of Palermo (zoomed) of the *Perseo* network with *Carto*.

affiliated members that were arrested.[25] The fact that there are many links around different areas of town shows that *mafiosi* are not operating completely disconnected from other families, suggesting that some interaction across the different *Mandamenti* is essential for the organization to operate.

Thirdly, some of these affiliates have many links between each other. This is shown by darker colored links. The network is very dense in the city centre, where few stronger triangulations are highlighted. These triangulations, however, do not behave like Delaunay triangulations.[26] From a centrality point of view, it seems that there are few well connected individuals that are quite close to each other. Could it be an expression of eigenvector centrality? This is one of the types of centrality considered afterwards.

These three pieces of evidence show how embedding a social network in a geographical space add more information than the classic non-spatial graph in Figure 2. The other major advantage from this analysis is that there are non-network related variables, such as unemployment levels, that can be meaningfully added to the representation of the network to characterize link formation, centrality and the structure of the organization. Once it is visually assessed that *mafiosi* are heavily clustered together and that they are well-linked with their neighbours, in the next section a spatial econometric analysis of one specific aspect of the network structure, i.e. the centrality of its members, is explored.

5 Spatial Network Econometric Analysis of Centrality Measures in the *Perseo* Network

The previous section highlighted some salient features of the territorial organisation of the *Perseo* network. In this section a spatial econometric analysis looks at the position of central agents. Centrality in criminal networks has been discussed, among others, by Calderoni and Superchi (2019) and Cavallaro et al. (2020). One relevant aspect of this analysis is understanding, in particular, whether Mafia leaders have a central role in the network as well. This is not always obvious because, as pointed out for example by Agreste et al. (2015), bosses tend to hide and, therefore, they might reduce the number of contacts with other members (see also the mentioned results of Calderoni, 2012 in Section 2). Another related problem is the identi-

[25] This was the case for the social network map of Mafiosi in Belmonte Mezzagno or San Cipirello.

[26] Those are triangulations where the inner part of the triangle is an empty circle (Okabe, 1992). It is evident that links are more dense in the network, because they show which members where in contact with the others. The reason is that this spatial criminal network is first of all a social one. Therefore it is possible that a node is connected also with others that are not the closest ones.

fication of which agents are more "central", assuming that their removal can be disruptive for the group. In this case, different measures of centrality can be considered (see Borgatti, 2006 for an insightful discussion of this topic). In this section our aim is to identify whether, given different centrality measures, there is any significant spatial pattern among agents. In other words, the underlying question looks at whether individuals with similar centralities tend to be connected and located closer.

Given a network different centrality measures can be computed, which can highlight different roles in the network. Five different types of centrality are considered here. In particular: i) the *degree* of the agent, which simply measures the number of links of each node; ii) the *closeness centrality*, which measures how many steps are required to access every other node from a given node of the network; iii) the *betweenness centrality*, which quantifies the number of times a node acts as a bridge along the shortest path between two other nodes; iv) the *eigenvector centrality*, which takes into account who each node is in contact with and the connections of those that are connected to it; v) the *Bonacich centrality*, which measures the power of a vertex which is recursively defined by the sum of the power of its connections.

Calderoni and Superchi (2019), in an analysis of the Italian *'Ndrangheta*, find that leadership in a criminal network is often associated to betweenness centrality, while Ballester et al. (2006) demonstrate that the "key player" in a network, i.e. the player whose removal from the network could substantially reduce network's productivity, can be associated to a measure such as the Bonacich centrality.

The econometric analysis of this section takes an initial step which is similar to the approach taken by Mastrobuoni and Patacchini (2012). That is, to identify spatial effects among agents with different centralities, first of all we consider a simple cross-section OLS model such as:

$$y_i = \beta \mathbf{X}_i + \epsilon_i \tag{1}$$

where y_i is centrality measure of agent i in the network, $\mathbf{X}_i$ is a vector of individual characteristics of agent i including a constant term, and ϵ_i denotes standard iid errors.

The arrest warrant of the *Perseo* operation allows to extract for each agent information on the following individual characteristics, that are useful for the econometric analysis: age, place of residence (city and address), the indication on whether, according to the investigators, the agent had a directive role in the organization, the agent's tasks as identified by the prosecutors. In particular, eight different tasks have been identified, and coded in this analysis by dummy variables: *LinkingIn*, if the agent was dealing with connections inside the *Mandamento*; *LinkingOut*, if the agent was taking care of keeping connections outside the *Mandamento*; *Meeting*, if the agent met with other members of the organisation in person; *PubProc*, if the agent tried or managed to influence the adjudication of public works through public procurement; *Pizzo*, if the agent was in charge of collecting ”protection money” (known as *Pizzo* in Sicilian); *Guns* if the agent was in charge of smuggling arms; *Drugs* if the agent was in charge of drug production and smuggling; *host* if the agent provided accommodation to other members.[27] As pointed out in previous sections, the place of residence of each agent has been associated to their latitude and longitude, a key aspect that first allowed a visual representation in Section 4 and then the spatial econometric analysis here.

A way to assess whether spatial correlation exists among the characteristics of each individual is to run a test on the residuals of a regression as specified in Eq. (1). This is assessed through the Moran index of autocorrelation that is computed as:

$$I = \frac{\epsilon' \mathbf{W} \epsilon}{\sigma^2 [trace\,(\mathbf{W}' + \mathbf{W})\mathbf{W})]^{\frac{1}{2}}} \tag{2}$$

where $\mathbf{W}$ is the distance matrix with elements given by the inverted distance among individuals at a given cut-off, so that, above a certain threshold, the higher distance is assumed not to exert any influence on other agents. Therefore, two individuals that are closer to each other would receive a higher “score” from matrix $\mathbf{W}$ and, ideally, they should be able to influence each other more than agents further away in space. In the present case, in absence of regular lattices (like the districts of a city or the regions of a country), distance-based and not contiguity-based matrices will be employed.

The null hypothesis of this test is that the residuals are not spatially autocorrelated (that is, they are iid from a spatial point of view), so that the rejection of this hypothesis means that estimation of Eq. 2 is biased, due to the omission of the consideration of spatial nuisance parameters (as this is a specific case of omitted variables bias).

In this case one solution is to model the spatial influence by considering the possibility that the centrality of an agent is influenced (positively or negatively) by the centrality of his neighbours. In this case, the model to estimate is a spatial autoregressive SAR model (LeSage and Pace, 2008), such as the following:

$$y_i = \beta \mathbf{X}_i + \rho \mathbf{W} y_i + \epsilon_i \tag{3}$$

which features the introduction of the term $\rho \mathbf{W} y_i$ in Eq. (1), where ρ (supposed to be lower than 1) is the spatial autocorrelation parameter, and $\mathbf{W}$ is the (inverted) distance matrix of the agents.

[27]Most of the agents were associated to more than one task because they committed more than one crime. The task *host* was indicated for only one member of the set of 99 agents, so it has been excluded from the main empirical analysis.

The model in Eq. (2) allows to estimate whether the centrality of an agent in the network is related to some observable characteristics, as for instance age or some specific tasks, and to the centrality of other agents, whose influence is weighted by matrix **W**.

To sum up, this procedure implies first of all estimating the model in Eq. (1) and test the presence of spatially correlated residuals by the Moran index of Eq. (2). Then, if the hypothesis of spatially id residuals is not rejected, the estimation of Eq. (3) is considered. In particular, in implementing this procedure we considered the different measures of centrality defined above, including in the regression dummies for the different tasks and agents' age.[28]

Table 1 shows the results, for the eigenvalue centrality, which turned out to be the only centrality measure with significant spatial effects, in specifications containing age and the dummy variable "Guns" indicating agents in charge of smuggling guns. The other dummies for other functions performed turned out to be non-significant (the entire set of results is available upon request).[29]

The first two columns of Table 1 show that age is positively related to eigenvalue centrality and the task "guns" is negatively related. This suggests that as age increases there is a higher chance that a member will have a central role, as expected.[30] In addition, dealing with smuggling of weapons implies that the agent has low centrality in the network, another significant result. Let us note that the dummy variable for the directive role did not turn out to be significantly correlated with the centrality index, a result consistent with the idea that Mafia bosses do not necessarily have central positions in the network (Calderoni and Superchi, 2019; Agreste et al., 2015).

The bottom rows of Table 1 report the p-values of the Moran test with a full distance matrix and a truncated distance matrix, in which the spatial influence is assumed to be zero for distances greater than 5 kilometers. In both cases the null hypothesis is strongly rejected (always at 5% and in one case at 1% confidence levels). This is a strong evidence for spatial correlation among eigenvalue centrality measures for members of the *Perseo* network.

For these reasons, Columns 3 and 4 report the results of the estimation of Eq. (3), testing whether neighbour clusters have similar or different centrality values (implying, respectively,

[28]While the original sample includes 99 members, it reduced to 91 in the estimations after dropping units with missing data and units with the same coordinate values, as it is the case of agents located very close to each other (in this case the matrix **W** would contain empty rows and the estimation of Eq. 3 would not be consistent).

[29]The cross-section analysis of this article has a relatively low number of observations, which can explain the low significance of many regressors considered in the estimations.

[30]Mastrobuoni and Patacchini (2012) also find that age has a significant positive correlation with eigenvector centrality for a sample of members of the American Mafia, albeit in a specification without spatial effects.

positive or negative autocorrelation).

Table 1: Network centrality regressions

	(1)	(2)	(3)	(4)
Intercept	-17.2	-13.90	-21.52^{**}	-18.22*
	(12.21)	(12.19)	(10.96)	(10.91)
Age	0.773***	0.736***	0.680***	0.643***
	(0.257)	(0.257)	(0.207)	(0.204)
Guns		-17.15***		-16.95*
		(3.00)		(9.32)
Rho			0.380***	0.382***
			(0.138)	(0.135)
Obs	91	91	91	91
R^2/Pseudo R^2	0.12	0.15	0.17	0.19
Moran test (full matrix)	(0.028)	(0.015)	-	-
Moran test (truncated matrix)	(0.014)	(0.008)	-	-

Notes: Robust standard errors within brackets;
***, **, * denote significance at, respectively, 1%, 5% and 10% confidence level.

Columns 3 and 4 in Table 1 show that there is a strong and visible "spatial effect". The spatial coefficient ρ is positive and strongly significant at 1%. As remarked, a significant spatial effect is only present for the Eigenvector centrality. In the *Perseo* network, agents with high eigenvector centrality tend to be spatially related (i.e. close) to other agents with similar levels of Eigenvector centrality.

Eigenvector centrality can be relevant in a criminal network, as pointed out by Malm and Bichler (2013, p. 373), in a study of a network of money launderers, as: "Individuals with high eigenvector centrality have more opportunity to interact with key players in the network. This means that these individuals may have only one or two connections, but they associate with the most popular individuals (key people with the most links)."

This result has interesting implications in understanding the organisation of a Mafia network. There are procedures by which a Mafia network is constructed and organised. In particular, recruitment takes place locally and there is a clear hierarchical order to be respected. This could imply that agents with high eigenvector centrality, i.e. with connections to highly connected individuals, tend to live close to each other. Assuming that individuals with high eigenvector centrality can put the organization at risk (see Baccara and Bar-Isaac, 2008, for an economic analysis of the risks faced by a criminal organization depending on its network structure), since they "know" individuals with many connections, this finding has interesting applications for policies aimed at combating organized crime, as the close physical distance among such members can increase the vulnerability for the group as a whole. A thorough analysis of the mechanism generating this spatial pattern as well as the implications for disruption policies are interesting topics for future research.

6 Conclusion

A criminal organisation like the Sicilian Mafia is defined by its strong structure embedded in the territory it controls. Network analysis tools are one of the possible ways to describe the way crime gangs organise themselves. Through these tools it is possible to get a better theoretical understanding of structures and possibly future developments of mafia-type groups. They are therefore an important part in devising effective response measures addressed against organised crime.

This article has two innovative aspects with respect to this analysis. The visual component of a social network in a geographical space with underlying territorial socio-economic variables, and the search for a spatial effect in some features of the network. By embedding social networks in a map it is possible to gather more information than the classic a-spatial graph alone, and it is possible to combine features that would not otherwise fit within that representation. As observed above, *Cosa Nostra* members are clustered in the geographical space according to their family belonging. Their *Mandamento* is not just where they live, but where they mainly bind their ties too. It is possible to understand how the network is built by also looking at contextual variables, such as unemployment. In such a context, closer members are not just linked together, but they also influence each other. This is especially true in villages around Palermo as shown in Figure 5. Most of the links are within each single village, in the realm of their own *Mandamento* and just few are tied outside it, but these ones keep the whole organisation operational. This finding is also consistent with one of the classic 'safety measures': members that do not know each other can reveal their affiliation only if a third party, known to both, introduces them (Morselli et al., 2007). This is an informal internal limit that ensures the criminal structure will remain sustainable despite arrest. Links between members are therefore very local: close to each other. Just few occasions allow selected members from different clans to meet, such as rebuilding its major representation body (the *Cupola*) or coordinating the smuggling of guns, which in this investigation was carried by multiple families from separate *Mandamenti*.

The spatial econometric analysis points out the presence of spatial autocorrelation of residuals of models containing network centrality measures. To the best of our knowledge this is a first attempt to model centrality network measures as related to distance and real space, through spatial models. The modeling of these parameters through spatial autoregressive settings confirms the presence of positive and strongly significant spatial relationships among the Eigenvalue centrality of an agent and his neighbours. There is therefore a spatial effect in the network of the Sicilian Mafia, which gives a nuanced characterisation to the findings of the social network analysis, but it is not a matter of tentacles over the areas it controls. Members with similar features and specialisations tend to be connected and live closer or, as the saying goes, birds of the same feather stick together.

References

Agreste, S., S. Catanese, P. De Meo, E. Ferrara, and G. Fiumara (2016). Network structure and resilience of mafia syndicates. *Information Sciences 351*, 30–47.

Agreste, S., P. De Meo, E. Ferrara, S. Piccolo, and A. Provetti (2015). Analysis of a heterogeneous social network of humans and cultural objects. *IEEE Transactions on Systems, Man, and Cybernetics: Systems 45*(4), 559–570.

Alizadeh, M., C. Cioffi-Revilla, and A. Crooks (2016). Generating and analyzing spatial social networks. *Computational and Mathematical Organization Theory*, 1–29.

Allum, F., I. C. Marinaro, and R. Sciarrone (2019). *Italian mafias today: territory, business and politics*. Edward Elgar Publishing.

Andris, C. (2016). Integrating social network data into gisystems. *International Journal of Geographical Information Science 30*(10), 2009–2031.

Anselin, L. (2013). *Spatial econometrics: methods and models*, Volume 4. Springer Science & Business Media.

Baccara, M. and H. Bar-Isaac (2008). How to organize crime. *The Review of Economic Studies 75*(4), 1039–1067.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: The key player. *Econometrica 74*(5), 1403–1417.

Barrat, A., M. Barthélemy, and A. Vespignani (2005). The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics: Theory and Experiment 2005*(05), P05003.

Barthélemy, M. (2011). Spatial networks. *Physics Reports 499*(1), 1–101.

Battisti, M., A. M. Lavezzi, and R. Musotto (2021). Organizing crime: an empirical analysis of the sicilian mafia.

Becucci, S. (2011). Criminal infiltration and social mobilisation against the mafia. gela: A city between tradition and modernity. *Global Crime 12*(1), 1–18.

Bichler, G. (2019). *Understanding criminal networks.* University of California Press.

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory 12*(1), 21–34.

Calderoni, F. (2012). The structure of drug trafficking mafias: the â ndrangheta and cocaine. *Crime, Law and Social Change 58*(3), 321–349.

Calderoni, F., S. Catanese, P. De Meo, A. Ficara, and G. Fiumara (2020). Robust link prediction in criminal networks: A case study of the sicilian mafia. *Expert Systems with Applications 161*, 113666.

Calderoni, F. and E. Superchi (2019). The nature of organized crime leadership: criminal leaders in meeting and wiretap networks. *Crime, Law and Social Change 72*(4), 419–444.

Calvó-Armengol, A. and Y. Zenou (2004). Social networks and crime decisions: The role of social structure in facilitating delinquent behavior. *International Economic Review 45*(3), 939–958.

Cavallaro, L., A. Ficara, P. De Meo, G. Fiumara, S. Catanese, O. Bagdasar, W. Song, and A. Liotta (2020). Disrupting resilient criminal networks through data analysis: The case of sicilian mafia. *Plos one 15*(8), e0236476.

Chainey, S. and J. Ratcliffe (2013). *GIS and crime mapping.* John Wiley & Sons.

Chainey, S., L. Tompson, and S. Uhlig (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal 21*(1-2), 4–28.

Dugato, M., F. Calderoni, and G. Berlusconi (2020). Forecasting organized crime homicides: Risk terrain modeling of camorra violence in naples, italy. *Journal of interpersonal violence 35*(19-20), 4013–4039.

Eck, J. and D. L. Weisburd (2015). Crime places in crime theory. *Crime and place: Crime prevention studies 4*.

Faust, K. and G. E. Tita (2019). Social networks and crime: Pitfalls and promises for advancing the field. *Annual Review of Criminology 2*, 99–122.

Felson, M. and L. Cohen (2008). Routine activity theory. *Environmental criminology and crime analysis*, 70–77.

Franch-Pardo, I., B. M. Napoletano, F. Rosete-Verges, and L. Billa (2020). Spatial analysis and gis in the study of covid-19. a review. *Science of The Total Environment 739*, 140033.

Humphreys, L. (2010). Mobile social networks and urban public space. *New Media & Society 12*(5), 763–778.

Kahle, D. and H. Wickham (2013). ggmap: Spatial visualization with ggplot2. *The R journal 5*(1), 144–161.

Lavezzi, A. M. (2008). Economic structure and vulnerability to organised crime: Evidence from sicily. *Global Crime 9*(3), 198–220.

LeSage, J. P. and R. K. Pace (2008). Spatial econometric modeling of origin-destination flows. *Journal of Regional Science 48*(5), 941–967.

Levine, N. (2006). Crime mapping and the crimestat program. *Geographical analysis 38*(1), 41–56.

Malm, A. and G. Bichler (2013). Using friends for money: the positional importance of money-launderers in organized crime. *Trends in Organized Crime 16*(4), 365–381.

Mastrobuoni, G. and E. Patacchini (2012). Organized crime networks: An application of network analysis techniques to the american mafia. *Review of Network Economics 11*(3).

Meyers, L. A., B. Pourbohloul, M. E. Newman, D. M. Skowronski, and R. C. Brunham (2005). Network theory and sars: predicting outbreak diversity. *Journal of theoretical biology 232*(1), 71–81.

Ministero dell'Interno (2016). Attività svolta e risultati conseguiti dalla direzione investigativa antimafia (dia).

Morselli, C. (2009). *Inside criminal networks*. Springer.

Morselli, C., C. Giguère, and K. Petit (2007). The efficiency/security trade-off in criminal networks. *Social Networks 29*(1), 143–153.

Musotto, R. (2017). *Social and spatial network analysis of organised crime*. Ph. D. thesis, Università degli Studi di Messina.

Musotto, R. (2020). From evidence to proof: Social network analysis in italian criminal courts of justice. *UNICRI Special Collection on Artificial Intelligence*, 38.

Musotto, R. (2021). Inside mafia networks. Proceedings of ASNAC2020.

Okabe, A. (1992). *Spatial tessellations*. Wiley Online Library.

Paoli, L. (2003). *Mafia Brotherhoods. Organized Crime, Italian Style*. Oxford University Press.

Papadias, D., J. Zhang, N. Mamoulis, and Y. Tao (2003). Query processing in spatial network databases. In *Proceedings of the 29th international conference on Very large data bases-Volume 29*, pp. 802–813. VLDB Endowment.

Rossmo, D. K. (1999). *Geographic profiling*. CRC press.

Ruggiero, V. (2019). *Organized crime and terrorist networks*. Routledge.

Santino, U. (2015). *Mafia and antimafia: a brief history*. Bloomsbury Publishing.

Santoro, M. (2021). *Mafia Politics*. John Wiley & Sons.

Santos, R. B. (2016). *Crime analysis with crime mapping*. Sage publications.

Scalia, V. (2020). The production of the mafioso space. a spatial analysis of the sack of palermo. *Trends in Organized Crime*, 1–20.

Sciarrone, R. (2019). The economic dimension of mafias, social capital and the grey area. In *Italian Mafias Today*. Edward Elgar Publishing.

Smith, S. J. (1986). *Crime, space and society*. Cambridge University Press Cambridge.

Tita, G. E. and S. M. Radil (2011). Spatializing the social networks of gangs to explore patterns of violence. *Journal of Quantitative Criminology 27*(4), 521–545.

Tumminello, M., F. Petruzzella, C. Ferrara, and S. Miccichè (2021). Anagraphical relationships and crime specialization within cosa nostra. *Social Networks 64*, 29–41.

Varese, F. (2012a). How mafias take advantage of globalization: The russian mafia in italy. *The British Journal of Criminology 52*(2), 235–253.

Varese, F. (2012b). The structure and the content of criminal connections: The russian mafia in italy. *European sociological review 29*(5), 899–909.

Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of small-world networks. *nature 393*(6684), 440.