UNIVERSITÀ DEGLI STUDI DI MESSINA

DOCTORAL THESIS

Social and spatial network analysis of organised crime

Author:
Roberto MUSOTTO

Supervisors:
Prof. Michele BATTISTI
Prof. Andrea Mario LAVEZZI

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

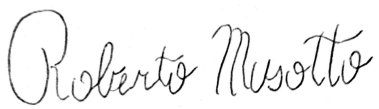
in the

Facoltà di economia

Declaration of Authorship

I, Roberto MUSOTTO, declare that this thesis titled, "Social and spatial network analysis of organised crime" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: 

Date: December the 7th, 2016

"Follow the money and you will find the Mafia."

Giovanni Falcone

Acknowledgements

I would like to express my gratitude to my supervisors Professors Michele Battisti and Andrea Mario Lavezzi for their continual support and invaluable guidance. If I become half the thinker they are, it will surely be one of my greatest accomplishments. I would also like to thank Dr. Giuseppe Musarra. He was the best supervisor I could possibly have in the Business School of Leeds University and helped me in understanding how to create the network.

I am incredibly thankful to my friend Víctor Vicente Palacios, co-founder of DataLab USAL at the University of Salamanca, who helped me choose the best software for geostatistic analysis and his helpful suggestions on network visualization. This thesis would not have been written in proper English if it was not for the invaluable help and patience that my friend Corinne McAlary offered to me proofreading all my drafts. Ievą, ačiū už viską, šita disertacija yra už tavę.

Part of the results of the thesis were presented in NetSciX 2017 conference and I would like to thank, among previously acknoledged people, Dan Cimpean, Enrico Sansone and the School of Law of the University of Leeds for its founding. I would like to thank David S. Wall for precious insights and comments on the whole work. I would have starved to death in the process of writing this thesis, if it was not for the moral and culinary support I got from Caroline Anderton for which I am extremely grateful.

Lastly, I wish to thank my family too. Even if they never understood why I decided to start a Ph.D., they supported me anyway. This thesis is for them and for their endless patience.

I would also acknowledge the helpful suggestions and comments from Nicola Persico, Luigi Balletta, Vincenzo Ruggiero and Deirdre McCloskey.

Contents

List of Figures

List of Tables

To the ones I love

Chapter 1

Introduction

This thesis is about social and spatial network effects of organised crime. I propose an analysis of the internal structure of the Sicilian Mafia, applying spatial network analysis to an original dataset on a network of Cosa Nostra members. The aim is to clarify the structure of Mafia-type organizations, where spatial network analysis highlights novel aspects on how such a criminal group manages to pervade a geographical area.

Therefore, in the first part the related literature is going to be analysed and it will be followed by a social network analysis of a Mafia related trial. In the last part, spatial effects will be studied. This study focuses on network and spatial implications in a criminal organisation such as *Cosa Nostra* which is something that was never tried before on organised crime groups.

I have built a dataset from judicial source documents about the Mafia. I have worked on a specific act of the process obtained from a real investigation and I propose a strategy in order to identify nodes and links. This is prodromal to the social and spatial network analysis that follows. The main problem dealing with criminal investigations is that there is no complete nor reliable information. I suggest a system based on source selection and data organisation that allows me to diminish the impact of missing data. Once the dataset is created, it is analysed with social network tools in order to show similarities with other criminal groups.

Then, each node is attached to its geographical coordinates and the network is analyzed on a map with spatial tools. Geolocalization of nodes and links allows us to assess a strong correlation between behavioural patterns and geography. This will be done in detail for the peculiar topology of the town of Palermo. Proximity and transitivity among nodes explains why specific areas are more targeted by extortion or criminal-led activities.

Previous work has proposed a SNA of organized crime or other criminal

networks. However, none of the existing works have analyzed organised crime in a spatial network context, in particular focusing on Cosa Nostra. No previous work has highlighted the internal structure of a Mafia group along with the geographical pattern of connections across members belonging to different mandamenti, as it is done here.

The thesis is structured as follows: Chapter 2 introduces organised crime and its internal structure in the light of the most recent findings in literature; Chapter 3 explains how to build the dataset and analyzes it as a social network; in Chapter 4 spatial network analysis is performed on the dataset.

Chapter 2

Literature review of Organised Crime

This chapter is going to focus on a definition of organised crime, followed by a socio-economic analysis of its origins, migrations and hierarchy. Afterwards the most relevant economic literature that focused on the internal structure of criminal organisations is presented, including what has been done so far in the field of social and spatial network analysis.

First of all, it is essential to define what organised crime is and its *nature* for the sake of clarity throughout the examination. Although in literature there is no general consensus about its definition that could enlighten both social and economical aspects, it "is viewed as a set of stable, hierarchically organized gangs which, through violence or its credible threat, have acquired monopoly control of certain major illegal markets. This control has produced enormous profits, which have been used to bribe public officials, thus further protecting the monopolies. These funds have also been invested in acquiring legitimate businesses in which the racketeers continue to use extortion and threats to minimize competition" (Reuter and Rubistein, 1978)[1].

According to Varese (2001) organised crime "aspires to obtain a monopoly over the production and distribution of a certain commodity in the underworld" and such a group does not refer to "every crime that is organised". Here, a definition will be employed that might be radical for sociologist literature, but which is certainly not new among organised crime scholars; that was employed by Leopoldo Franchetti. He argues that a mafia is a specific organisation which exercises abuses and seeks illicit earnings[2], working as a

[1] In order to get an insight on the complexity of this phenomenon, the best source and collection of more than 180 different definitions of organised crime can be found here

[2] "associazioni destinate all'esercizio della prepotenza e alla ricerca di guadagni illeciti", my translation.

real *industry of violence*.

This definition is also consistent with Gambetta's Sicilian Mafia (1996) description where such a group has the *monopoly of violence*. I.e. Mafia is willing to provide illegal governance services to anyone who might need them, acting as a quasi-governmental structure.

2.0.1 A socio-economic analysis of the origins of Organised Crime

In this part, mafia-type organisations will be investigated in the light of the most relevant economic and sociological literature. There are classic works which have wondered how it is possible that a mafia originates or moves into a specific area from other territories. There are peculiar conditions spotted in the literature to explain why the surfacing of organised crime in a specific market is allowed.

Gambetta (1996) shows that scholars assigned three possible explanations to the origin of the Sicilian Mafia: the latifundia, the urban markets and local political conflicts. However, he considers the Mafia as a "set of firms supplying protection in whatever context and to whichever customers they find profitable" (Gambetta, 1996, page 76). This is so because there was a *demand* for protecting fragile and conflicting transactions. Therefore, a *supply* of people, grouped into firms, that made sure that every contract was secured, rose and took advantage of the poor and inefficient protection of property rights that was granted by the local government. He argues that the lack of trust of Sicilians in local and national government and institutions could have been one of the potential causes that led people to look for protection elsewhere. This is why trust was imperfectly substituted with protection.

It started with the end of feudalism, when property rights on the octroyed land were not fixed anymore and it was possible to buy and to sell it and use it as a guarantee. Bandiera (2003) demonstrated "how the division of landholdings fostered the development of the mafia" (page 1). Starting from parliamentary surveys about Sicily published between 1881 and 1886 by Abele Damiani, she developed a single agent model where landlords pay for protection and when the fragmentation of the land augments there are more landlords willing to pay, incrementing the mafia surplus.

This market for protection extended from landowners to different markets too. Buonanno et al. (2015) investigated the most important export commodity in the XIX century in Sicily: sulfur. They found out that Mafia levels rose in the districts were sulfur mines concentrated, showing that "when public law enforcement institutions are weak or absent, a boom in the value of natural resources significantly contributes to the emergence of mafias, because it creates both a demand for private protection and opportunities to extract rents through extortion, two dimensions along which mafia-type organizations have a competitive advantage, thanks to the coordinated use (or threat) of violence" (page 3). Therefore, inefficient institutions, that could not meet the demand of protection, were replaced by other providers of a similar service in exchange for a rent.

Another theory about the emergence of the Mafia is found in Dimico et al. (2017). They analysed the market of oranges and lemons in Sicily in the period between 1881 and 1886, proving that high market revenues were the main reason that led to the insurgence of the Mafia in that specific market rather than land reforms and the broadening of property rights as stated by the rest of the literature. This is so because lemon producers hired the Mafia for private protection in order to create a barrier to entry for other producers rather than actually secure their rights.

Mafias can also originate because of migration (Buonanno and Pazzona, 2014) or transplant themselves to other areas (Varese, 2011). This has happened in the past due to the massive migration flows from southern Italy to the northern regions, northern countries and towards America and Australia, and because of *confino* policy that forced people convicted of Mafia charges to relocate to different provinces for a specific period of time. According to Varese (2011), this actually helped mafias to expand their control to new markets. In the regions where *Mafiosi* were relocated crime rates associated with Mafia activities augmented and had a strict causality with such a policy (Buonanno and Pazzona, 2014).

Varese (2011) studied the causes that could explain how organised crime can settle in a specific area. The research focused only on a specific way in which this radication could happen: the *transplant*. Members of criminal organisations were moved from their original place to unfamiliar environments. This new environment was supposed to be *healthier* and not penetrable because of the high civic engagement of people living in the northern part of Italy. However, such a transplant helped criminal organisations to prosper

creating more "demand for Mafia". This transplant success for the Mafia was possible thanks to several features:

Trust. Trust bonds together "organised criminals" (Von lampe and Ole Johansend, 2004). It helps such networks to reduce uncertainty in a field where there can not be law enforcement intervention. It also gives to its members a competitive advantage in business, thanks to the tight social relations.

Trust plays an important role from the point of view of collective action (Paoli, 2002) , because it is a strong non economic tie which allows the Mafia to be better organised than any other social group. This peculiar feature of organised crime makes the employment of collective action more difficult for every other social group which may not have such strong social ties.

Size of *locale*. It is easier for the Mafia to penetrate a market if the martket has not so many competitors. The smaller it is, the easier it will be penetrated. Once the Mafia breaks into the market, it will offer "efficient and convenient" services in order to patronize the largest stake of firms.[3] However, the Mafia clan designs to conquer the monopoly over protection in a specific neighborhood or market[4] and it dictates different or deferred payment solutions according to the size of the firm.[5]

In order to successfully settle in one specific market, firms should not be export-oriented and they should compete in the same place. The reason why stands on the fact that in such a situation it might rise a demand for a ruler or protector. Such a third party actor might help in sharing clients or revenues. That is the case of Bardonecchia example in Varese (2011). This is what happens in Sicily (Lavezzi, 2008), where "a relatively high share of traditional sectors and economic activities [are] strongly related to the territory, like the Construction sector; (ii) a relatively small firm size, (iii) a relatively low technological level and, (iii) a large dimension of the public sector" (page 2).

Large illegal/legal **Markets**. The intuition that these conditions offer is that expanding or illegal markets require a high demand for protection of

[3]This is the case that happened some time ago in Palermo where Mafia bosses lowered the price of *pizzo* in order to encourage more firms to join the service (La Repubblica, 24th June 2014).

[4]This is confirmed by Confindustria which alleged that 90% of businesses in Palermo pay pizzo for protection (Giornale di Sicilia, 27th December 2014) .

[5]This is the case where a car retailer was forced to pay *pizzo* with cars (Meridionews, 2014) .

Balletta and Lavezzi (2014) analyzed how the amount of *pizzo* moderately increases with the firms' size.

property rights where the State or the law can not help. In illegal markets, there is the need for an alternative institution that provides protection because the State can not deliver such a service. In a legal market, cartels are an efficient measure to exclude newcomers from transactions and insure that the markets will be equally shared.

There is a collection of causes whose combination allows organised crime to penetrate into a specific market in order to offer its protection services, defined by the term ”entry conditions” (Varese, 2011). The analysis of all these causes is directed to find a theoretical foundation about the birth and development of organised crime in a market or territory.

In literature, Gambetta and Reuter (1995) and Varese (2011) discuss incidentally about "entry conditions" too, highlighting several features that a market or territory should have in order to favour organised crime operations.

First of all, **product differentiation must be low**. Differentiation is a process that allows us to distinguish a product which is released by a firm according to its characteristics or qualities from another similar product offered by others (Kotler, Keller, Ancarani and Costabile, 2014). Differentiation is low when for a consumer or end user it is *almost* indifferent to pick a firm product or another from its competitors. An example might be a taxi service, since each enterprise offers a car and a driver. Although the kind of car and driver skills might differ, the purpose of every taxicab is to carry passengers from one place to another in exchange for an amount of money.

There must be **no barriers to entry** or at least they must be low (McAfee, Mialon and Williams, 2003). The definition of a barrier which is preferred is that from Stigler (1968). He considers barriers "a cost of producing that must be borne by firms seeking to enter an industry but is not borne by firms already in the industry". Such a description is a strict one compared to what Bain (1956) stated about barriers:"an advantage of established sellers in an industry over potential entrant sellers". The difference between those definitions stands on the fact that Stigler’s excludes that economies of scale and capital can be a barrier. Firms can freely enter or exit one market if barriers are low.

Barriers to entry can be varied, like, for example, intellectual property laws, inelastic demand, zoning or tariffs. Low barriers imply that specific markets are easy to penetrate because the cost of entry is low or close to zero, making it affordable for every enterprise that might want to compete there.

Technology must be low too. Technology is meant as the knowledge and information that society has acquired concerning the use of resources to produce goods and services and it affects the technical efficiency with which resources are combined in production. Low technology is referred to as that which requires non-automated and labour-intensive processes. The reason why technology might be considered low is because it involves unskilled labour and there is no need, or it is not possible to convert to a higher technology.

Unskilled labour: a workforce which requires no specific education or experience. The "advantage" that such a labour force gives is in the fact that each worker can be easily swapped with another without suffering any additional cost for their education. Varese (2011) agrees on the fact that their presence can help organised crime to penetrate a market and to settle.

Inelastic demand of products: a product is inelastic if the price increasing does not influence the quantity demanded. This definition implies that end users will always buy that specific product despite its price. Concerning organised crime services, inelasticity could be found in illegal products (this could be the case of the drug market) or in legal markets (see *Solncevo petrol oil interests* in Varese, 2001). (Albanese, 2008).

Large number of firms involved: according to Gambetta and Reuter (1995), Mafia controlled cartels tend to have a large number of members, although this pushes them to leave the agreement and "to take advantage of restrictions" (Scherer and Ross, 1990). However, Varese (2011) confuted such a requirement. In fact, an illegal organization that aims to penetrate into a market will face more problems because the number of people that will have to be intimidated, bullied, threatened and colluded will be greater.

The size of the firms should be small. The reason for such a requirement is straightforward. A big firm which has multiple branches around the world can easily resist, avoid to collude and they might eventually move elsewhere.

However, big firms and enterprises can also benefit from agreements with third parties. This is the case described by Saviano (2006). Parmalat milk, Cirio canned sauce and Bauli panettone had an agreement with Camorra clans. Those firms were paying a weekly fee in order to be sure that every food retailer in the Napoli area would sell only their products. Clans also set a very low and fixed price for such products in order to clear all the stocks quickly. Hence, although these colluding companies were losing profits in

the market, the volume of sales allowed them to gain credibility and high quotations on the stock exchange market.

Labor unions. Organised crime might use union power for personal benefit in order to promote licit and illicit activities. Unions are labor force representatives that are supposed to negotiate and rule workers rights. A more formal definition can be found in Marx (1867): trade unions wish to prevent the price of labour-power from falling below its value. Such an organisation can lead into collusion with management level (Blakey and Goldstock, 1980). Union corruption is functional for organised crime purposes (Block and Chambliss, 1981) because it creates a controlled mass which is good for being employed as a vote bank or pressure group.

Gambetta and Reuter (1995) wrote on *Mafia controlled Cartels* and about how such agreements could be drawn up. A *cartel* is an agreement among some or every producer of goods or services in a specific market which aims to limit, restrict and rule competition in order to get higher profits than the ones that they could get in a free market status. A cartel is contracted by competitors in a market, but such contractual parts might need the help of a third party which is going to get that agreement undertaken by everyone and which is going to punish any violation. Such a "third party" could be an organised crime group. Especially if it is not possible for cartel members to bargain with an institutional organisation that could effectively strengthen the bonds of contractors.

They suggest the idea that Mafia can not rise and settle if there is no demand for a specific (illegal) service. Such a particular service can be only supplied by organised crime and the State can not offer it because of legal or public order reasons[6].

However, the activity they analyse is one of those that pushes towards a "demand for Mafia". Such a demand is prodrome and essential in a market where organised crime wants to offer its services (Lavezzi, 2014 and Varese 2006).

[6]There is one interesting exception: in Indonesia after the national revolution against Dutch colonization, organised crime groups, composed of *preman*, took power creating a mixed situation of legalised felonies and apparent order (Aspinall and Mietzner, 2010).

2.1 On the Internal Structure of Criminal Organizations

2.1.1 Structure of Cosa Nostra

The analysis in the following chapters will be centered on the *Cosa Nostra* organisation. Therefore, how this specific criminal group is shaped is going to be outlined.

Cosa Nostra is the specific name for the Sicilian Mafia. It differentiates the identity and the operational area of this Mafia group from other criminal organizations operating in the peninsula like the *'ndrangheta* in Calabria or *Camorra* in Campania.

The structure of Cosa Nostra (Our Thing) was unknown for a very long time. It was only during the investigations for the Maxiprocesso that it was actually revealed and afterwards disclosed to the public (Dickie, 2006).[7] Differently from Mexican Cartels it is not possible to be directly hired by the criminal organisation and job vacancies are not advertised[8], but specific people recommend individuals to be affiliated to the organisation. Cosa Nostra imposes therefore a "status contract" upon their members (Paoli, 2001). A specific ritual allows the appointed person to become an honorable man and he will have this status until his death.[9] Those "men of honor" are devoted to silence and to antepone the interests of the criminal organisation to any other personal need.[10] The organisation is also multifunctional like the Chinese Triads or the Japanese Yakuza, in the sense that it simultaneously pursues different tasks (Paoli, 2001), but one is predominant: "the exercise of a political dominion" (Paoli, 2004, page 22). This is done by applying "their will to exercise political power and their interest in exercising sovereign control over the people in their communities" (Paoli, 2004 page 24). Therefore, the

[7] There is a brief analysis of Mafia related trials and the Maxiprocesso investigation in the following chapter.

[8] This is a peculiar characteristic that is found in the Sinaloa Cartel, Haitian and Guatemalan Cartels (Wainwright, 2016).

[9] In 'ndrangheta families this status is also reached just by being part of the natural family. Newborns are *baptised* and members are affiliated at a very young age (Paoli, 2004). At the same time their structure is fostered by multiple family ties (Calderoni, 2012).

[10] This is what the "decalogue" pizzino (a small piece of paper that Mafiosi use for communication), seized when Salvatore Lo Piccolo was arrested, stated. Members of the organisation had to be available for work even if their wives were pregnant. (La Repubblica, 7th November 2007) .

dangers of Cosa Nostra are potentially higher than any other criminal organisation that takes over illicit activities driven by their profitability. This is so, because part of the illicit profits are invested in the control of political power.

The Hierarchy

Criminal groups have many different ways to design their internal organisation. Mafia-type organisation have a vertical structure with a strict hierarchy outlined in multiple levels.

The first level in the hierarchy is the *soldato* (soldier) or *uomo d'onore* (honorable man). It has multiple tasks and ten or more soldiers are headed by a *capo-decina* (head of ten). All the soldiers are also part of a family, which is also called *cosca*. The family is controlled and directed by a family leader. Usually three or more families are grouped into a *Mandamento*, which is the turf in which the families operate. The mandamento is controlled by a capo-mandamento and he takes part in the district commission (*Commissione Provinciale*), together with other representatives.[11] On the top of each district commission there is a *Cupola* in which all the district representatives take part. It has directive roles of the entire organisation (Dickie, 2005).

2.1.2 Economic Analysis on the Internal Structure of Criminal Organisations

Economic literature (Baccara and Bar-Isaac, 2008 and Polo, 1995) has focused on the internal structure of the Mafia too. It tries to explain what bonds a group of criminals and how the group and single members behave under specific requirements. Polo (1995) developed a competition model, concentrating on the internal structure of the Mafia, using a principal-agent approach, which is based on trust and wages. He shows that a criminal organisation tends to hire affiliates offering a specific wage and threatens them from joining another illicit group with a very high punishment. If there is another criminal organisation that does so, hiring costs will be higher, punishments will be harder and less people will be joining each group.

[11] In the next chapter 3 there is a deeper analysis of the commission, because the arrested members tried to reorganise a district commission in Palermo.

Baccara and Bar- Isaac (2008) created a model where they "consider the trade-off between the enhancement in internal cohesion derived by exchanging internal information and the increase in vulnerability to detection that this exchange implies" (page 3). Their paper argues that inside a criminal group there are two objectives that are partially conflicting: the need for secrecy and the internal efficiency for the organisation. A perfect flow of information between each member means that 'everyone knows everything', therefore if one member is arrested the whole group is endangered. Since secrecy is vital for the survival of the organisation, information flow is strictly limited. Therefore if a peripheral member is detected and arrested, it would not undermine the entire group because the information at his disposal is limited. According to the organisation of the illicit group two detection strategies could be implemented. In the agent-based detection there is an authority that focuses on each agent independently. While in the cooperation-based detection, "in which the probability of detection is an increasing function of the cooperation level of the agents" (page 32).

2.1.3 Social Network Analysis of Organised Crime

Definition and contents

There is another part of literature that is related to social network analysis and crime. SNA is useful for investigations of kinship patterns, community structure and interlocking directorships , because it investigates the way they work through network and graph theories (Scott, 1988).

Specifically social network analysis looks for "patterns or regularities among interacting units" (Wasserman and Faust, 1994 page 3). It originally started in social sciences and eventually its perspective migrated to many other research fields. Networks are helpful in order to explain the structure, the environment and the relationships between actors.

SNA is often used in order to perform activities that concern criminal intelligence (Sparrow, 1991).It has the advantage of offering insights on the internal structure of the analyzed group. In fact, it is not always possible to freely gather information on criminal groups, given the sensitivity of the problem.

Social network analysis has had multiple applications in the criminal sector and specifically in the organised crime field. Mastrobuoni and Patacchini (2012) focused on the Italian-American Mafia network in the period of 1950-1960. They showed how the network has a strong hierarchy and also which qualities or characteristics a Mafia leader needs to have. Therefore an individual who is around 59 years old and not too tall, preferably born in Sicily, but residing in New York, married (possibly to a wife connected to another Mafia family), with a very extended family, many children, working in a Mafia related business and possibly arrested at a young age for a violent crime has a higher probability of being in a central position of a Mafia-type criminal organisation.

The Italian-American Mafia was the focus point of another work by Mastrobuoni (2015) where "gangster's network centrality influences his economic prospects" (page 3) taking into consideration the place where they were born and the value of the house where they lived once captured.

Masías et al. (2016) analysed the Canadian-based drug network called *Caviar* (which was already used by Morselli in 2009) and also the Watergate scandal. It used a machine learning approach combined with social network analysis in order to generate a more accurate classification of models according to the final verdict. In such a case, different measures for measuring the centrality of the network[12] were compared with the final verdict in order to see which are the one that approximate the sentence better.

There are also other papers that apply social network analysis tools to groups similar to organised crime. This is the case of Krebs (2001 and its revised version of 2002) that used the terrorist network responsible for 9/11 in order to show how it was possible to map and link all the relational data and evaluate them in order to find the leader.

Methodology

The data is usually accessed through law-enforcement agencies or from criminal justice systems (Morselli, 2009). It usually consists of acts and evidence from the investigation or the trial that follows (Berlusconi et al., 2016, La Spina, 2011 and Natarajan 2000). The majority of networks are built from

[12]See 3 for a definition of centrality of the network.

a single source of information; Berlusconi et al. (2016) created different networks from three court proceedings about the same trial. This judicial access allows us to gather data with a certain degree of validity and reliability (Campana and Varese, 2011) if there is *no self censorship*; there is a *wide group coverage*; there is a *large sample of conversations*.

There are multiple risks once the sample is selected since it could be a false negative or a false positive and, most of the time information is just missing (Morselli, 2009). Sometimes, it depends also on the point of view of the investigation, which focuses on specific parts of the network (Berlusconi, 2013).These difficulties in the kind of data gathered are furthermore translated into unclear boundaries, since it is not possible to decide who to include or not, along with imprecise dynamics inside the group (Krebs, 2002). The *modus operandi* and the network structure are useful tools that can help prosecutors to understand how the organisation works and how to fight it (Mastrobuoni and Patacchini, 2012).

Kinship, violence and trust (Von Lampe and Johansen, 2003, Campana and Varese, 2013 and Campana 2016), which were simply considered as exogenous variables that kept the organisation together now have different perspective to be explored: the network perspective. It has been discussed if criminal organisations could be described as networks (Varese, 2010) and despite contrary opinions (Powell, 2000), criminal organisations have a less formal structure than firms and they might be considered as a form of of organisation (Williams, 2001), despite the spatially pointed bias that this kind of analysis faces.

Once the data is selected and targeted the network is constructed. The first step goes through the construction of a matrix. The matrix can indicate the presence (with 1) or the absence (with 0) of a contact (Morselli, 2009). It can also give more information about the direction and the symmetry of the relationship. E.g. the information is directed when Node 1 sends a letter to Node 2, but it is not directed when it is not possible to give a sense of direction to the communication: this is the case when Node 1 and Node 2 are having a conversation in front of each other. The symmetry explains how and how much the directed communication flows between nodes. Therefore it tells if a Node contacts another one more often.

After the matrix is created, the network is represented through a graph, which is a graphical representation of the nodes and their links on a plan (Jackson, 2008). The way the nodes are displayed and visualized is essential

in other to facilitate key points of the analysis to other people. In 3 there is a description of the graph visualization techniques that were employed in the analysis.

Once the graph is created statistics can be extracted from the network. Density describes the "texture" of the network (Scott, 1988), since it tells how complete the graph is and how connected compared to its highest potential. Other properties like reciprocity and transitivity of the network are also extracted.[13] Another piece of information that is usually extracted from the network is about the centrality (Morselli, 2009) of nodes. In 3 there is a definition of the most important centrality tools. The reason why they are employed in such analysis is because they allow better understanding of how the individual is placed in the network in comparison to other members according to different values assigned to their connections (Carrington et al. 2005).

There are methodological similarities of this work with Krebs (2001) Morselli (2009) and Berlusconi et alia (2016) papers. Krebs focused on terrorist networks explaining how it was possible to map a covert network. Morselli used a case study on the criminal group of Hells Angels conducting social network analysis over it. Berlusconi et al. (2016) did social network analysis on one *'ndrangheta* group[14] in order to get more insights on link prediction.

The way the network was accessed, gathered and completed follows their steps, as was stated already. First of all, an adjacency matrix was built. This was done in order to understand which node was in contact with which. This matrix was also reshaped in order to know how this relationship was directed and if links were symmetrical. A directed link is that one that goes from one node to the other and the contrary is not true. As previously shown, this might be the case when a node writes a letter to another: in this situation there is a clear sender and one receiver. If the sense of the information is not available or it is not possible to specify the direction the link is undirected. The relationship between nodes is symmetrical when they contact each other equally. Vice versa if one node contacts another node more often, the relationship is asymmetrical. Secondly, the matrix was employed to build the network representation and relational data. Differently from previous literature that employed ucinet software, this was done through the igraph package on R software.

[13] See chapter 3 for their definition and uses.

[14] A mafia-type organisation that originated in Calabria and it has a different structure from the other criminal groups in Italy, like Cosa Nostra and Camorra.

2.1.4 Spatial Network Analysis of Organised Crime

Definition and contents

Spatial networks are "networks for which the nodes are located in a space equipped with a metric" (Barthelemy, 2011, page 3). In the case of social networks, the space is added in order to assess, for example, if the probability of a link between two nodes will be lower if the distance increases. Spatial information gives more accurate information on the strength, density and distance of nodes (Barrat et al., 2005).

This kind of analysis has two main uses: to explain (Papadias et al., 2003) values of specific variables in the network or to predict values of variables (Meyers et al., 2005) in order to clarify the potential of the network.[15] This kind of analysis allows the effects of scale in dispersed networks and the context where nodes operate to be shown. In fact, every spatial network has nodes located in a space (Barthelemy, 2011). The space could be a Cartesian space or a geographical map.

Despite the fact that there are already applications of space into social networks and crime (Anselin et al, 2000), this is something that was never tried before with organised crime groups.

In fact, the most recent literature focuses on applications of social network analysis to criminal groups (Berlusconi et al., 2016; Morselli, 2009). However, there were already applications of this technique in the works of French sociologists in the 19^{th} century. One of the earliest can be traced back to Durkheim's essay on suicide (1897). He found out that suicidal rates cluster in geographic space. The relational data between this activity and geography led to multiple applications in other criminal sectors like assaults and robberies. This is the kind of analysis that, in the 20s, the Chigago School performed on the above mentioned offences (Anselin, 2000). In the literature, spatial analysis was applied to crime in order to get insights about its routinary activity (Levine, 2006) and in order to map hot spots where criminal offences thrive or they are more likely to happen in the future (Chainey et al., 2008).[16]

[15] Examples of this kind can be found in the traffic models or pipe networks and in the predictive model of SARS infectious disease.

[16] Routine activites study if a specific place attracts or inhibits crime in a limited period of time (Andersen, 2006). This theory is based on the fact that there is an offender, a potential victim and no-one to guard. Hotspots find patterns between illicit activities and concentration of places where they occur the most (Chainey et al., 2008).

Methodology

Alizadeh et al. (2016) explain how to transform classic networks like 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. Therefore, given an m X m Cartesian space and placing the nodes on the plan, the nodes are connected according to the assumptions of the network model which is considered each time. So, two nodes are connected if they reach a specific radius (random network), if they are the closest to a node (small-world network) or if the nodes are close and well connected with the others (scale-free network).

As it is shown in chapter 4, in order to conduct this analysis one more matrix is needed to show the probability of connections. This is the distance matrix that shows the length of the shortest path between two vertices (Anselin, 2013).

Chapter 3

Operazione Perseo: a Social Network Analysis of Sicilian Organised Crime

3.1 Introduction

Operazione Perseo is the code name of the police operation that took place on the 16th of December 2008.[1] This operation is named in such a picturesque way because of one of the most famous myths that surrounds the image of the legendary founder of Mycenae, the son of *Zeus* and *Danaë*: Perseus. According to the legend, while wearing winged sandals to fly and with a polished shield, he cut off the head of the gorgon Medusa.[2] In a very romantic comparison, this was the objective of police investigations: to cut the head of the gorgonic Mafia.

This chapter focuses on the topology of the non-spatial network and on econometric implications of the relevant variables extracted, while the following chapter 4 will include in the analysis spatial coordinates. After an introduction on the trial and on the type of document chosen for the extraction, the network and its properties will be analysed. Results show how mafia actors tend to cluster and how such a network is different in size and centrality from other criminal networks.

[1] It is possible to check here that its *official* name is N. 18038/08 R. mod. 21 D.D.A.
[2] It is possible to get the full story about Perseus here.

The context of the trial: Operazione Perseo and the Mafia trials in Palermo district

Operazione Perseo was not the first process of the Sicilian Mafia.[3] The first one dates back to 1863. This was the same year that, for the first time, a term referring to the Mafia was coined. The successful play *I Mafiusi di la Vicaria* (the Mafia men from the Vicaria) was released.[4] The first of October 1862, 13 men were killed and, after investigations, in 1863 the material executors were punished with death and forced labor as a part of a bigger cabal against the Bourbonic kingdom. Almost 30 years later, in 1883, in Favara 200 people were arrested for being part of *La Fratellanza* (the brotherhood). The process took place in the nearby town of Agrigento in 1885 for 107 members of the organisation (Dickie, 2006). This can be considered as the very first Mafia trial that took place in Sicily.[5]

Another essential investigation is the *Rapporto Sangiorgi*.[6] The *Rapporto* is a collection of communications that the public procurer Ermanno Sangiorgi sent to the Ministry of Internal Affairs between November 1898 and January 1900. Sangiorgi listed and mapped 218 members of the criminal organisation, their strategy and the techniques that were deployed to control the territory. This led to the 28th of April 1900 and in the following months to a warrant of capture issued for hundreds of members of the organisation. In May 1901, 89 members were formally accused and the judgement stated criminal accusations for 32 members. From its investigations in 1899 and in 1903 there was the famous Notarbartolo trial.[7] Only 2 people were investigated, but it was the first trial that put the Mafia on a public level.

During the fascist period, the political regime tried to suppress any criminal organisation in the island with *hard fist* policies that Cesare Mori, the

[3] All the police operations and trials described up to 2005 are described by Dickie (2006). All the information of other trials and operations after 2005 until 2016 was collected from local newspapers. There is a list in Italian here about all the major police operations against Mafia in Italy from 1945.

[4] Even if it is largely documented that there were episodes of organised crime before that year on the island (e.g. the report of the Baron Turrisi Colonna on public security in Sicily), it is not really possible to state if the criminal organisation considered itself as the Mafia (Dickie, 2006).

[5] There were smaller trials against similar organisations in Palermo regarding criminal groups operating in the *Uditore* neighborhood and in the Monreale area in 1883. However, those trials had no relevant impact, because of the environmental pressures that happened.

[6] The official name is *aa.gg.rr. Atti Speciali (1898-1940),b.1,f.1.*

[7] The marquis Emanuele Notarbartolo was killed in 1893 in a train cabin between the train stations of Termini Imerese and Trabia, near Palermo. The first trial had to be repeated because of a defect of form.

local prefect, was required to follow by Mussolini to cauterize the plague of organised crime (Mori, 1932). The result was that from the first of November 1925 until May 1929 "11000 people were arrested, of whom 5000 were in the province of Palermo" (Dickie, 2006 page 326). However, many regime opposers are included in this count (Lodato and Travaglio, 2005).

The next Mafia process was held almost 40 years after, in 1968, because of the massacre of Ciaculli.[8] 117 members were accused and condemned for the truces between Greco families and the members that had to take part in the first provincial *Commissione*.

In 1985, the investigation called *Pizza connection* traced back the network between the American and Sicilian Mafia. 19 members of the organisation were arrested and condemned.

On the 10th of February 1986, the *Maxiprocesso* started with 474 people accused of being part of the organisation, thanks to the declarations released by the *pentito* (repented) Tommaso Buscetta.[9] In the end, 346 members were condemned.

On the 24th March 1994, the *Petrov* operation cut off the heads of the Mafia based in Trapani, issuing 74 arrests for members of the organisation. The Mafia based in that district had many connections and links with the Palermo based Mafia.

In January 2005, the police operation *Grande Mandamento* arrested 46 members of the Mafia close to the boss of the bosses Bernardo Provenzano. His arrest happened on the 26th of April 2006.

Two months after this event, in June 2006, 52 Mafia members were arrested in the police operation *Gotha*. This uncovered two factions that wanted to take over the control of Palermo after the capture of Bernardo Provenzano. One group was held by Antonino "Nino" Rotolo and the other was headed by Salvatore and Sandro Lo Piccolo. The trials that followed the investigation are usually referred to as *Addiopizzo* trials from the name of the organisation that acted as a plaintiff.

Right before the Perseo operation in February 2008, the operation *Old*

[8]This episode put an end to the first Mafia war and it involved a car bomb that was meant to kill a rival boss but instead it hit several policemen.

[9]Buscetta became an informant and he spoke about the internal structure of the mafia, instigators and executors of killings.

Bridge arrested 90 members between Italy and United States of America unveiling the plans of the Inzeerillo family to take control of Palermo. In December 2008 operation *Perseo* took place, that cut off the heads of the organisation in the Palermo district.

In 2011, operation *Pedro* cut off the heads of the Porta Nuova mandamento arresting 28 people. That same year operation *Hydra* arrested 16 members of the San Lorenzo mandamento. In 2012 operation *Atropos* arrested 41 members of the Noce mandamento.

In 2013, operation *Nuovo Mandamento* arrested 37 members of the organisation that tried to create a new mandamento outside Palermo. That same year operation Alexander arrested 24 people from the Porta Nuova mandamento. In November 2014 operation *Zefiro* reached 18 members of the mandamento of Brancaccio.

In 2014 operation *Apocalisse*, followed by operation *Apocalisse 2* arrested 97 members of the San Lorenzo and Tommaso Natale clans. The four operations *Grande Passo* that same year captured 28 people of the Corleone mandamento. Simultaneously, 16 members of the Brancaccio clan were arrested in the *Eden* operation.

In 2015 operation *Panta Rei* arrested 39 members, and for the first time there was a proven role of females in the criminal organisation. That same year, operation *Verbero* arrested 39 people from the Pagliarelli mandamento and operation *Stirpe* hit 5 people from the Santa Maria di Gesú mandamento. 7 members of the organisation were arrested in the *Jafar* operation in March 2015.

In January 2016, operation *Cicero* arrested 9 members of the acquasanta mandamento. That March of the same year, operation *Brasca-quattro.zero* hit 62 members of the Villagrazia and San Giuseppe Jato mandamenti. In the October of 2016, operation *Monte Reale* arrested 16 members because they tried to reorganize the mandamento of Monreale.

A list of the processes and police operations are summed up in the following table 3.1.

Such a table 3.1 helps to contexualize operation Perseo in the light of the other trials and investigations about the Mafia.

Three investigations were not included in the analysis. The one that involved the 7 times prime minister Giulio Andreotti, because it was around only one person and it did not manage to prove a link between the Mafia and

Year	Name	Arrests
2016	Monte Reale	16
2016	Brasca-quattro.zero	62
2016	Cicero	9
2015	Jafar	7
2015	Stirpe	5
2015	Verbero	39
2015	Panta Rei	39
2014	Apocalisse 1 and 2	97
2014	Grande Passo 1, 2, 3 and 4	28
2014	Eden 1 and 2	16
2013	Nuovo Mandamento	37
2013	Zefiro	18
2012	Atropos	41
2011	Pedro	28
2011	Hydra	16
2008	Perseo	99
2008	Old Bridge	90
2006	Gotha	52
2005	Grande Mandamento	46
1994	Petrov	74
1986	Maxiprocesso	474
1985	Pizza connection	19
1968	Strage di Ciaculli	117
1925-1929	Mori arrests	11000
1898-1900	Sangiorgi Reports	218
1883	La Fratellanza	200
1862		13
Total	around 31 investigations	12860 arrests

TABLE 3.1: Table of trials and investigations about the Mafia in the district of Palermo and Sicily

politics.[10] The ones connected to the Maxiprocesso and the operation Vespri Siciliani because they were strictly involving the same people that were investigated in the first trial. The *Addiopizzo* trials because they are strictly connected with the Gotha investigations.

The analysis of Operazione Perseo benefits from the fact that it hit almost all the families and turfs in Palermo district during the same period of time. Despite its importance, there is only one academic study that talks about this

[10]However, the same sentence states that there might have been links until 1980, but it was not possible to punish anyone for them.

Aspects considered	Data
Kind of material	Arrest Warrant
Criminal Organisation	Cosa Nostra
Criminal Groups	25
Investigated members	99
Involved members	77
Place of investigations	Palermo
Period of investigations	from 2006 until 2008
Dimension of the file	1398
Wiretaps	226
Phone calls intercepted	95
Pizzini	15

TABLE 3.2: Empiric Material Employed

operation (see La Spina ,2009).[11]

3.2 The Operazione Perseo dataset

3.2.1 The data

Before introducing the dataset, a few words will be spent on the arrest warrant which is the source of the document.

The description in the table 3.2 of the empiric material employed is helpful to explain why this dataset is important and it will set the background story behind all the numbers and data. *Operazione Perseo* is the second biggest trial after the *Maxiprocesso* for number of arrests since the Italian Republic was created. Such a trial with all its data it is worth to be analysed. Two clarifications are necessary. The actual number of *Mandamenti* is around 11. However, in the file some families are considered to have had influence in such a big area that is is to be considered as a mandamento. Therefore the actual number of Mandamenti is reported to be various by different sources and scholars (among all, see Dickie, 2006). The file refers also to the existence of pizzini. Pizzini are the information or orders written on a (usually small) piece of paper.

This police investigation allowed detectives to arrest 99 mafia members. Similar numbers are considered in literature for social network analysis (La

[11] Its referral though, it was incidental.

Spina, 2009 and Berlusconi et al., 2016).[12] These *uomini d'onore* (honorable men, as they like to be called) were not only simple soldiers, but most of them were the heads of their *mandamenti* (i.e. the turf controlled by a mafia family). The main reason why they were targeted was to block the attempt to replicate for a third time the so-called *Commissione provinciale* (i.e. the district commission). Such a commission is the equivalent of a body of directors of a company, where members are elected or appointed, and they jointly oversee the activities of all the affiliated families. In Sicily, each district has a commission and all their chiefs are part of the highest board which is defined as the regional commission or the *Cupola*.

It might be evident already that the district commission in Palermo has a complicated history. In 2008, it was reconstructed for a third time.[13] The main function was to bring "harmony" between the families operating in the same area and, in the meanwhile, to be a common ground where "serious" decisions would be made.[14]

The existence of the commission is a distinguishing feature of *Cosa Nostra*. In fact, there is no a direct equivalent in Italian criminal organisations.[15] This sort of *democratic* tool that the Mafia created had in the past (and still has nowadays) the functions of a ruling and judicial body inside the organisation. In fact, all the matters that involved more than one family were supposed to be discussed there. This was a source of pride for *Cosa Nostra* members, since it allowed them to shape a better organisation than the Neapolitan *Camorra*.[16]

So, this operation cut off the strategic direction of the criminal organization in the whole district of Palermo. At the same time all these arrests provided a sort of map because of their quantity and quality. In fact, arrests hit almost all the *mandamenti* that are located in Palermo and around the town. Specifically, 25 of them[17], where the most represented is the Belmonte Mezzagno *mandamento*, as can be seen from the following picture 3.1.[18]

[12]Respectively, La Spina considered two criminal groups made out 78 and 73 members. Berlusconi et al. a group of 182 members.

[13]The first commission was created by American and Sicilian Mafia in 1957. This one is usually considered to have ended in 1963, because of the awakening of the first Mafia war. A second one was created in 1974 after a short period of triumvirate.

[14]See page 44 of the arrest warrant from Operazione Perseo.

[15]However, this peculiarity can be found also in Italian-American and Jewish mobs.

[16]See page 44 of the arrest warrant from Operazione Perseo.

[17]There is a conceptual error recurring in the document where multiple times families are considered mandamenti.

[18]Mandamento is the turf controlled by a mafia family.

FIGURE 3.1: Mafia distribution in Belmonte Mezzagno

The picture 3.1 is exemplary of the control and spread of the territory that the Mafia has in smaller villages. It is the graphical representation of the town of Belmonte Mezzagno where a heatmap of criminal members' locations and the local network ties are overlaid. The location of affiliated members is the last place where they were living before being captured. An actor that is or has been interacting with another gets a link from him to the other. As it will be shows further in this chapter, actors here are linked according to a specific set of documented contacts (such as calls, meetings and letters) that were gathered from the trial.

It is evident how each member of the organisation was extremely close to the others. It was actually possible to spot an arrested member of the organisation every couple of blocks. Each node has a blue halo that shows its vicinity. Yellow haloes show an extreme proximity of members (less than one block), while red spots show members of the organisation that were living in the same building. Green lines are the ties that link each node to the other. A quick glance displays how dense and tight the network was. In fact, criminal members covered almost the entire town and all of them had close ties to each other and with their fellows in the district. This map is a preliminary example of how it was possible to shape the network starting from the arrest warrant.

The image of Belmonte Mezzagno is supported by a geographical theory of crime that shows how crimes cluster. The actual proximity of criminals residences creates the false feeling of a safe environment where illicit actions can be carried out without many consequences. This feeling is the one that attracts more criminals around a specific area and it influences the composition of the community in the town. In fact, the ecology of a criminal network in a small village contributes to a *no escape* impression in victims' mind as their choice stands between surrending to the abuses or to join in multiple ways the criminal group.

There are multiple reasons why the arrest warrant is a good act to be chosen for such an analysis. In fact the *ratio* and the juridical nature of the documents are very unique (Tonini, 2012).

The related literature (Scaglione, 2011) successfully deployed arrest warrants as a starting point for social network analysis. The arrest warrant discloses for the very first time the fact that there is an ongoing investigation. Usually, the arrest warrant is one of the very first documents issued in any criminal process in Italy by the Public Prosecutor (i.e. the Italian *Pubblico*

Ministero) and, most of the time, the last document of every investigation, together with the notification of its end. Ideally, every investigation leads to an arrest warrant, if there are enough proofs to support the idea of a criminal offence. In the case of *Operazione Perseo*, an order was issued to take into custody (*provvedimento di fermo*) according to article 384 of the Italian criminal procedure code.[19]

The order of custody can be issued by the Public Prosecutor under a series of conditions:

- heavy proofs towards the person under investigation;
- that the collected proofs might raise the risk that the person under investigation might try to escape justice;
- the investigated crime might be punished with life in prison or a jail term with at least two years of jail time as minimum and at least six years as maximum or it concerns terrorism or democratic subversion (as Mafia crimes are).

However, the most important feature of the warrant is that the Public Prosecutor is legally obliged to present **every proof** against and in favour of the person that is currently under investigation. This means that technically the Prosecutor does not have any real power to select proofs.[20] This is so the only proofs he is actually allowed to discard are the ones that are redundant or not relevant to the investigation.[21]

All proof must be presented because the prosecutor has to put the deciding judge in the position to evaluate the entire position of the person under

[19]It is not very accurate to write about an arrest warrant in such a context. For the sake of completeness, the difference is going to be quickly outlined here. However, during the rest of the paper, both terms will be used, because it will not affect the final result. According to the Italian criminal procedure code, there is a difference from arrest and custody. The first is the power to guarantee criminals to justice and to block them from committing more illicit actions. It can be executed only when the offence is *in flagrante* or *in quasi flagrante*. It is executed by police or, sometimes, by private citizens and it can not be issued. The second is a preventive detention that is executed in all the other situations, that do not allow for an arrest to take place, and it is issued by the *Pubblico Ministero*.

[20]This is so after the Law number 332 from 1995 and the Law number 47 from 2015. This task is really hard to achieve because, after all, the Public Prosecutor must prove the foundation of the alleged accusations.

[21]However, the Public Prosecutor is obliged to present the proof in its entirety. He is not allowed to discard redundant parts, like in that part of a conversation, that can be found at at page 295 of the arrest warrant, where Mafia Members discuss their favorite Neapolitan singer, Mario Merola.

In this conversation Mafia members discuss which song to pick. They usually do that in order to cover their voices in the case someone is tapping their conversations.

investigation. Therefore, proofs in favor of the investigated person are act as a counterbalance to the alleged accusations. The body that makes a selection between the proofs is the judge. The verdict is actually the result of this selective activity.

So, the network that could be displayed by the judge is *per se* limited to a smaller fraction of the list of nodes and links that was first presented to the judgement (Berlusconi et al., 2016). The latter network is the result of what it is worth to be punished according to criminal laws. This is why the sentence is less apt for such an analysis than the arrest warrant. In this latter, proofs are rawly displayed and they have not yet been filtered through the grinding procedure of the judgement. Consequently, the arrest warrant has interesting features in order to gather quality data for network analysis concerning criminal networks. This is so, because, despite the fact it is a secondary source that might have "incorrect attributions in the characteristics of nodes and ties" (Campana, 2016 page 5), the way the warrant is written might reduce the issues of missing data. In fact, all the available evidence about the organisation is set up for judgement. The peculiarity of this source is that the amount of available data for each node and link reduces the risks of missing data inside the investigated network. Even if the entire list of links inside the organisation may not be evident to investigators, all the observable information about every link is gathered together with the help of wiretap records, audio and video collections and family trees. The information which is collected should be the least biased possible. Afterwards, during the judgement process the judge will create a smaller network of criminally punishable nodes and links. The fact that the arrest warrant is more complete than the other judicial acts in the entire process and investigation is the reason why the network analysis in such a field should consider first of all the arrest warrant as a source.

However, all the other dangers related to data collection stay intact. The major problems are caused by the boundaries of the network. The boundary of a network is "defined by recording who is interacting with whom in a certain context" (Kossinets, 2006, page 249). Therefore, it is not always certain who and which ties might be really included in the network. In the specific case of *Operazione Perseo*, the problem is that, even if there might not be omissions in the data, the investigation focused only on one part of the Mafia around the district of Palermo. It is not possible to state that all the members in that area were captured. Also, all the relational ties that lead outside this geographical area are partially excluded from the analysis. This is the

case for the relationships between the Mafia network in Palermo and the one in Trapani. For example, in the warrant, there are constant references to the major boss of *Cosa Nostra* Matteo Messina Denaro[22], but the links with that part of the organisation are treated in the investigation as if they are outside the network. The relationships are linked, but it is not deeply analyzed as to how they are linked. In the previous case of Matteo Messina Denaro, for example, this means it is not possible to specify if the link is directed[23] or not[24] nor its quality[25] (Mastrobuoni and Patacchini, 2012). This translates on a partial vision of the network under investigation (Campana and Varese, 2012). Given the structure of the Mafia, which is devoted to secrecy, and criminal investigations that have a limited amount of funds and time this can not be avoided.

The dataset and its network were built through a complete analysis of the judicial text. All personal and sensitive information are not reported in the network in order to protect the privacy of anyone involved in the investigation and to share publicly its results.

3.2.2 Method

In this section the variables of the dataset will be described. This is the result of the information that could be extracted from the warrant. The dataset is a collection of nodes and links from the arrest warrant. The individuals involved are the nodes of the network. The link is a relation between two nodes.[26]

[22]See page 63 of Operazione Perseo. In this conversation each member boasts about the fact they keep in touch with Matteo Messina Denaro in order to give themselves more relevance in the Mafia meeting, because of such an endorsement.

[23]If each node is directly linked to the other. I.e. if actor A calls actor B, there will be a link which goes from A to B.

[24]If the nodes are connected through one or more nodes. This is particularly frequent in Mafia organisations. In the Mafia decalogue that was found when the Sicilian Mafia boss Salvatore Lo Piccolo was arrested in 2007 one of the rules was that no-one can present himself directly to another member, but there must be a third person to do that. In a Network Science context this means that nodes are indirectly connected (Barzel and Barabasi, 2013). I.e. it is a transitive relation: given the set X with three members $\forall n_1, n_2, n_3 \in X$ and the link R, member 1 and member 3 are linked if member 1 is linked to member 2 and member 2 is linked to member 3: $n_1 R n_3 \iff n_1 R n_2 \bigwedge n_2 R n_3$

[25]The nodes are connected because of family or business relations.

[26]The dataset was extracted by hand and then analysed thanks to the software R.

Nodes

In the dataset, the nodes are the actors that appear in the arrest warrant. 22 variables were collected for each actor. Like in the rest of the literature (Mastrobuoni and Patacchini, 2012; Fafchamps and Gubert, 2007), these variables involved individual and group characteristics.

ID: Identifier of the actor. This is how each node is identified in the dataset. The ID is kept constant in the actors' list and in the list of links. In total, 176 actors were collected. This is a common procedure for the description of the network that helps to facilitate transcription of links and nodes (Morselli et al., 2007). The number of the actors is also consistent with the one of other criminal networks analysed in similar works (Morselli et al. 2007 and Berlusconi et al. 2016).

Name of the actor. This was collected in order to facilitate data extraction. Names are not disclosed in the network, as it would make possible to trace back the actors of the network. However, the most common names were Giuseppe (23 entries), Salvatore (16 entries), Giovanni (12 entries), Francesco (9 entries), Gaetano and Antonino (7 entries).

Surname of the actor. Surnames are not disclosed in the network.

Sex of the actor Male or Female. Precisely: 166 Males; 9 Females and 1 not available. Literature often uses these variables in order to assess the relationships between gender (Snijders, 2011). The result is consistent with the nature of a criminal organisation like *Cosa Nostra*. In fact, women are not allowed inside the group *per se*. However, they have a different place in mafia related crimes. They give support within the family in their day-to-day activities.[27]

Nationality of each actor. 149 were Italians and the information was not available for 27. Even if it is not the case here, this variable is useful for criminal groups operating abroad (Varese, 2006) in order to show how transnational organisations move and operate. The nationality is also a key indicator for the heterogeneity of the members in the society, which is, according to early studies of ecological crimes from Shaw and McKay (1969) an indicator for the relationship between communities and crimes.

[27]Fiandaca (2007) reports how the Italian legislation and trials until 1999 were more favourable for women involved in Mafia related activities. In fact, since *Cosa Nostra* accepts only men and women do not have any membership nor formal role in the organisation they were considered not punishable. It is easy to confute such a conclusion looking at how women as much as men had benefits from being around the operating area of the organisation.

City of birth of each actor. Actors were born in 17 different cities: 88 actors were born in Palermo, 18 in Belmonte Mezzagno, 8 in Monreale, 7 in Bagheria, 2 in San Giuseppe Jato, 2 in Montelepre and 2 in Corleone. Information was not available for 39 actors. Therefore, half of the actors involved were born in Palermo and another 10% in Belmonte Mezzagno. City of birth and residency are the variables that are employed for understanding connections between members (Restrepo et al., 2006) and they often dictate one's mandamento affiliation (Mcillwain, 1999). This variable is different from other illicit organisations where the place of birth is not really relevant. This is the case of Hell's Angels affiliation for example where owning an Harley Davidson motorcycle is a paramount feature in order to join the gang.

Province of birth of each actor: 131 members were born in the Province of Palermo. Again, the data was not available for 39 members.

Date of birth of the actors. Birthdates are not disclosed in the network. However the age range is very wide: the eldest actor was born in 1918 and the youngest in 1981. This is an indicator that helps to define the quality and the strength of ties between actors (Granovetter, 1973; Petroczi et al., 2007).

City where the actor lives. This was not available for 63 actors. 57 were currently living in Palermo, 22 in Belmonte Mezzagno, 8 in Monreale, 6 in San Cipirello, 4 in Montelepre, 3 in Bagheria. As stated before, this is an indicator of one's party affiliation (Mcillwain, 1999).

Province where the actor lives. 110 lived in the province of Palermo, 63 were not available, 2 in Reggio Emilia (in Emilia Romagna) and one in Catania.

Address of the actor. Addresses are not going to be shared in the network, but they were essential for the geolocalisation of the nodes in the spatial network analysis. This is the starting point for the creation of a spatial social network (Barthelemy, 2010).

Mafioso. This is a dummy variable that it considers whether the actor is part of the organisation or not. 131 actors are allegedly supposed to be part of the organisation while there is no such an information for the other 45. This variable is identified in literature as the ability of the criminal justice to identify the perpetrator (Restrepo et al., 2006).

Trial. If the actor was arrested because of the accusation or he was just involved in the trial for other reasons. 99 actors were arrested because of the forthcoming trial. 77 were involved in the trial for other reasons. In

the literature it measures the ability to accuse and to put a person on trial (Restrepo et al., 2006).

Date of death. This was collected where applicable. It might refer to actors that died during the investigation or that died during the trial. This date is not going to be shared in the network either. This is relevant in the literature because of the quality of ties (Petroczi et al., 2007).

Actors' occupation. This variable refers to the work occupation the actor had in his everyday life. Similarly to other research conducted in the same field (Gambetta and Reuter, 1995) most of the actors that were involved in the Mafia had low skilled jobs or worked in sectors where the insurgence of Mafia controlled markets is more favourable. It was possible to gather this information only for a very limited number of members of the criminal group: 6 actors were builders, 5 fruit sellers, 3 handymen, 2 farmers, 2 restaurant owners and 2 drivers.

Mandamento affiliation. This is the turf controlled by one or more Mafia families. Usually every Mafia member is part of a specific family and they have control over a specific area. All the members live inside the mandamento and its extension or areas of influence might vary with time. Generally historical studies showed that there are 8 mandamenti in the town of Palermo and 7 within the Province of Palermo (Dickie, 2006). In the analysis of the arrest warrant 25 Mandamenti were found. This a conceptual error of the document and it is so because most of the time the investigators considered that one family had influence in such a wide area that it could be considered as a Mandamento. In any case mandamento areas should be considered in a broad sense and multiple times mafia families joined together or separated mandamenti.[28]

Position inside the Mandamento. This refers to the structure that the single actor had inside the organisation. This information was edited by the people that wrote the arrest warrant. The actor could be a family member, a head of the family, a mandamento leader or offer support as an external member. 75 actors were family members, 20 mandamento Leaders, 10 external members, 9 family leaders, 4 leaders.[29] The role, the directing activities

[28]This is the case of what Lo Piccolo family did when they joined the mandamento of San Lorenzo and Villagrazia together.

[29]The leaders' label refers to those affiliated members that were in such a high position not to be considered part of any family, because of their directive role in the entire organisation.

and the actors' tasks are useful to shape back the internal discipline and the actual roles of the criminal organisation (Varese, 2012).

Direction. This variable tells if, according to the investigation, the actor had leading activities inside the organisation. This happens when there are documented links that show how the actor was taking decisions on behalf of others and taking part in high level meetings. In the criminal code this is an aggravating factor that might lead to a higher punishment if the facts are proved to be right. 15 actors had this aggravating fact contested. This fact was not available for 14 actors. It was not contested for the other 85 suspected.

416-bis: If the actor was accused of being part of a criminal organisation that acts with methods and techniques that are alleged to Mafia-type organisations according to the Italian criminal code.[30] The article 416-bis here was imposed by the investigators and judges that wrote the document. It was employed as a dummy while gathering the data in order to identify the people that had enough proofs against themselves to be reasonably considered part of the criminal organisation. Among the 99 arrested actors, 83 were accused of being part of the organisation and 17 were not. All the 15 actors that were contested to have a directive role in the Mafia were accused also of being part of it.[31] All the 10 external members did not have the article 416-bis contested. 7 family members were not contested to be part of the criminal group. This group was made out of members that had minor roles in the organisation and they were accused of taking part in very few activities. These are peripheral nodes of the network (Morselli, 2009 and Newman & Girvan, 2004). Most of them were linked because of the strict family ties and wiretaps that connected them to specific criminal episodes without a real involvement with the rest of the network.[32]

[30]The article identifies specific behaviours that are typical of such an organisation. The third subparagraph states that there must be a group of people that uses intimidation, subjection and rule of silence in order to commit crimes or to control economic activities.

[31]This is so because the aggravating fact of Mafia direction is the second subparagraph of the article 416-bis.

[32]At the moment when an arrest warrant is usually issued, investigators face a problem concerning which kind of accusation to make. They have to evaluate the role of a node through data which might be inconsistent. So, the risk is that the position of each node might be over or underestimated. Simultaneously, these actors set the boundaries of the network. Since it is harder to define exactly the role of peripheral nodes, it is not possible to specify network boundaries clearly (Laumann, Marsden and Prensky, 1989).

Actors' tasks. This variable collected the kind of operations that the actor performed in the criminal group. They were collected through the warrant considering the specific kind of accusation that was alleged against each member.[33] 6 different tasks were collected. *Linking* if the actor was dealing with connections inside the mandamento. *Linking2* if the actor was dealing with connections outside the mandamento. *Meeting* if the actor attended meetings in person with other members of the organisation. *PubProc* if the actor tried or managed to influence public procurements. *Pizzo* if the actor was extorting money for protection. *Guns* for smuggling guns. *Drugs* for drug production and smuggling. This information was available only for the 99 arrested actors and for a few other actors that were previously arrested in other or connected trials. Each actor performed multiple actions. Linking and meeting activities could be considered *public relations* activities (Wainwright, 2016).[34] Extortion activities, guns and drugs smuggling are technical output activities because they have repercussions outside the criminal group. 44 actors were involved with *pizzo* activities, 8 with guns smuggling, 3 with drugs smuggling and production, 7 with public procurement frauds, with 31 keeping links outside the mandamento, 46 kept documented links inside the mandamento, and 35 took part in meetings.

Document label and **document page** had the function to order all the references of the actors all along the arrest warrant.

Links

In the dataset, links are any documented contact between nodes. Links were collected in three moments because of the very high amount of contacts encountered. In the first part of the document there was a description of the most important links and connections. The document reported connections among affiliated people for each actor which was charged with criminal allegations. All those contacts were extracted in a separate sheet. Afterwards,

[33] In fact, for every actor that was arrested, the warrant stated the personal data of the person, the general charge and the specific reason why it was believed that he or she deserved the allegation.

[34] An illicit action is the one that is against the law and offends somehow society (Voltaire, 1766). According to the Italian criminal law, meetings and links are assumed to be part of that set of actions that allows a criminal organisation to thrive. Therefore, not every link or meeting is punishable, because it depends on the kind of meeting. All along the arrest warrant organisational meetings and "work" contacts are the ones that are contested to the actors.

all the other documented contacts were independently listed. Eventually, all the links were merged together. 9 variables were collected.

Idlink: 1404 links were collected from the document. These are the ties of the network (Morselli et al, 2007).

Idlink1: This identifies the first actor. It shows who started the contact, if the link is directed. The analysis of this variable already shows a feature that is common in criminal networks (Morselli, 2009, Berlusconi et al. 2016, Xu and Chen, 2008).

Idlink2: This identifies the second actor. It shows who received the contact, if the link is directed. The same conclusions are applicable for what it was stated for idlink1 variable.

Directed: If the link has a direction. Two values were assigned. 0 if there was no direction or it was not applicable. 1 if the link was directed from idlink1 towards idlink2. This is useful above all to give a sense of nodes that are hubs, authorities (Kleinberg, 1999) or those who get or give more links (Klemm and Eguiluz, 2002).

Typelink: This shows the type of link. Different labels were assigned according to the quality of the link. Both weak and strong ties were considered (Patacchini and Zenou, 2008). The link was *generic* if it was generically defined by the prosecutors in the first pages of the document. 506 were identified as generic. It was the most common. This type of link shows that there are evidences of contacts between two actors but it is not possible to gather any more information than this. The link was labelled as *single* if it referred to a specific episode of a contact (for example during a meeting). Every specific episode had to be documented. It could have had a date or not. It could have been documented in many different ways: through letters, wiretaps, photo and video data. 616 single links were documented. *Family1* direct family links were labelled. I.e. sons, fathers, husbands, wives, daughters were classified under this category. 15 links of this kind were found. All the other family links were categorized under the label *family2*. I.e. grandfathers, nephews and cousins. 11 links of this kind were found. These might be some of the strong ties of the network. In fact for a link to be strong a family tie can be a sufficient condition, but not a necessary one as there might be other types of connections that are strong as well. It was not possible to document friendship ties since they were not reported in the warrant. The last label was about documented contacts through *meetings*. Meetings were

considered each contact in the same place between two or more people documented through wiretaps, audio and video data. 255 meeting contacts were collected.

Medium: This is the variable that describes how the link was documented. The way the link is collected tells exactly if the data is valid and reliable (Campana and Varese, 2012). Only two mediums were collected All along the warrant: wiretaps and phone calls. *Wiretaps* are the written record of a conversation between two or more people. They offer good reliability as they usually can be introduced in court as evidence (Varese, 2011). Sometimes, they might be misleading because they do not transmit any sentiment or the emotion attached to the words.[35] All the conversations in the warrant are wiretaps unless it was differently specified. 95 *phone calls* between actors were intercepted too. Calls were intercepted and their content transcribed in the act. 15 links were reconstructed through *pizzino* communications. Pizzino is a (generally) small piece of paper that it is often employed for communications between members of the organisation. Its content can be an order or a prescription and usually it comes from the heads of the organisation towards peripheral nodes.

Datelink: This variable collects the date in which the contact occurred. Literature (Carrington et al., 2005) considers dates as a key element in longitudinal networks that change over time. Links date from 2005 until November 2008.

Doc and **Docp**: these variables had the function of referring in which part of the document the information about a specific node was gathered.

3.3 Results

The variables were analysed through R software after they were collected and cleaned of errors. First of all they were considered in order to get its descriptive results. In order to get its relational data an adjacency matrix 3.2 (nxm) was built from the 1404 links and 176 actors. This is an adjacency matrix with weighted links, where the weight is the number of registered contacts between two actors. In order not to lose the data from the amount of links per node, the matrix was rebuilt as a heatmap. Black areas show

[35] However, most of the time, all the surrounding sounds in the conversation (e.g. laughs, shouts, music) are transcribed too.

a connection between nodes while white areas do not show any connection at all. Darker shades of black indicate a higher number of contacts between members.

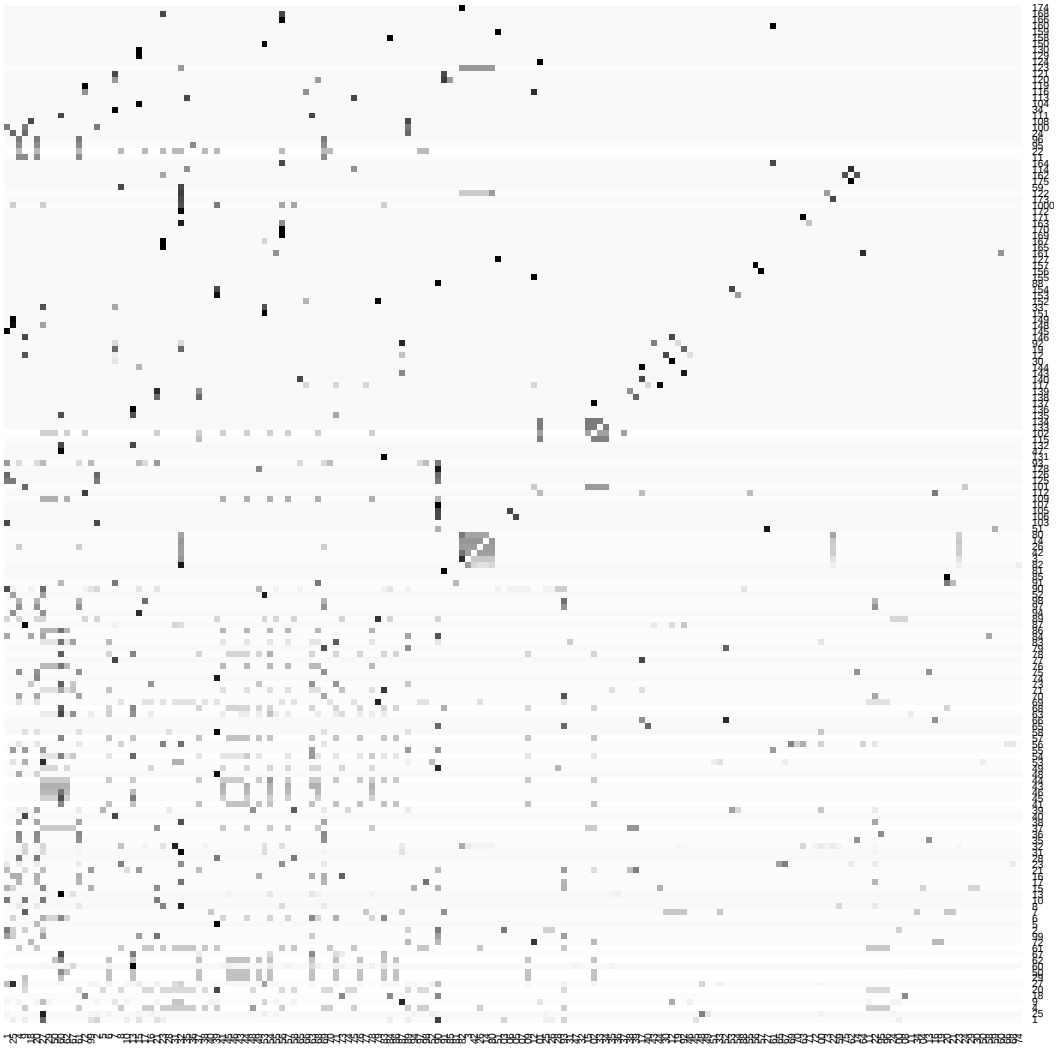


FIGURE 3.2: Network Matrix

3.3.1 Network Topology

Even if the statistics do not allow us to assess which position is held by every actor, they outline who has had a more active and, therefore, relevant role in the network. The topology of the network becomes relevant in such a context. In fact, the way the network is displayed on a plan gives an idea of its multiple ties and potential relationships that can be bond all through the organisation.

The way a graph is displayed shows the sense in which there are the most relevant social connections. Two nodes that are connected are represented next to each other. This is so because physical space is not taken into account yet. At the same time, the closer the nodes are, the higher the probability is that they might be connected. The reason why two nodes are closer to each other is because they might be *similar* to each other. Here similar does not mean that nodes have the same characteristics, but that they might have similar connections. This is the definition of *homophily*. It shows tendency that "individuals often interact with others similar to themselves" (Kossinets and Watts, 2006 p. 88). Therefore, similar actors are represented closer to each other as they might end up having a connection. If two nodes are close and if they share one or more contacts together they create a triad. Triads give the idea of nodes that are clustered (Wasserman and Faust, 1994), which means that if two nodes have a specific behavior, the third connected node has a similar behavior (Centola, 2010).

The difference between non-spatial and spatial network stands on the fact that networks are usually represented on an Euclidean space, which can have two or three dimensions. Ties measure the distance between real connections in the euclidean space. Two nodes are closer in the Euclidean space if their distance is short. This short distance might be the result of the share of the same set of connections. The topology of spatial networks adds something more to be displayed. In fact, spatial networks have nodes which are geolocalised in a geographical map. Once the nodes are distributed in the network according to their geografical coordinates, connections are overlayed. The fact that nodes are embedded in a geographical space is a feature that non-spatial networks do not have.

Nodes in a spatial network have also a (more) limited number of potential connections that can be made, since they are limited by the space which is necessary to connect each other. This is the case of road networks for example or networks of isolated groups of people.

The cost of connecting two geographically distant nodes is higher and less likely to happen compared to two closer nodes (Boccaletti et al, 2005). This feature is not considered in non-spatial networks where the distance is cost-free.

According to Ravasz and Barabasi (2003) non-spatial networks describe better hierarchy power between nodes because the lenght of the ties reflects

Actors Number	Links	%
32	84	5,9
27	79	5,6
60	74	5,2
13	57	4
9	57	4
25	39	2,7
31	36	2,5
1	28	1,9
90	27	1,9
69	27	1,9

TABLE 3.3: Actors' distribution according to the number of links

the importance of the nodes and it "is absent in networks with strong geographical constraints, as the limitation on the link length strongly constraints the network topology" (Ravasz and Barabasi, 2003 page 4).

The proximity, homophily and degree distribution of the network give information about the local and the global structure of the network.

The following table 3.3 lists the first 10 nodes according to the number of links they had, which is an aspect of the more general degree distribution. The first ten nodes cover 508 links which is 36% of the total amount of links. Reading the table it is be evident that actor 32 had 84 links which is almost 6% of the total links. The following actors in the top 5 cover 25% of the links. Even if it is not shown in the table, there are plenty of actors with very few or no links. As stated before, this is a feature that it is expected to be common in other very covert illicit networks similar to *Cosa Nostra* such as the Italian *Camorra* and *'Ndrangheta* and not in other organisations like Hell's Angels, because in this latter group members are actively invited to socialise. This distribution is also consistent with online criminal social networks (Yip et al., 2012) as there are few members that are highly connected between each other in online illegal trading that tend to have consistent contacts from other affiliated people that are less active. In the context of *Cosa Nostra* organisation this leads to a strategic management configuration where the administration of the entire organisation is left in the hands of few actors that are very active, but they tend to engage on tasks and to address orders directly to few individuals rather than hosting collegial meetings (Baccara, 2008).

The network was built thanks to the igraph package (Csardi and Nepusz,

2006) in R which was built for network analysis. Its simplest representation has got loops and multiple arrows between nodes removed. It clearly shows in the following figure 3.3 the strong relationships between the most connected nodes.

This is the case even if the algorithm is changed from the standard Barabasi-Albert model (Barabasi-Albert, 2002) that is actually meant to show hubs as opposed to a random network.It was checked with spring forced graphs 3.4 like the Kamada-Kawai (1989) and the Fruchterman-Thomas (1991). Spring forced graphs have the advantage that they behave like physical systems and in order to be drawn they have to reach an equilibrium between all the nodes. The graph was also displayed in figure 3.5 with a simulated annealing algorithm (Davidson and Harel, 1996) that minimizes the energy of peripheral nodes. The simulation here is on the probability that nodes that have a higher density are closer to each other.[36] The advantage of having different displays of the network stands in the fact that there are multiple nuances to be caught for each picture. This helps to offer alternative when studying a covert network like Perseo.

Another classical approach was checked, building the network, with the multidimensional scaling (MDS) (Wickelmaier, 2003) that aims to get more "stimulated" nodes closer and further those that are not.

Networks properties (density, reciprocity and transitivity) are now showed and compared with other criminal networks. Afterwards diameter and degree centrality results are going to be discussed. Results show that the network is not highly connected if not for a denser core of nodes, with a high presence of peripheral ones.

Density

The density of a network, which is the portion of the potential connections in a network that are actual connections, is extremely low, at 4,8 % since there are only 1404 links. This is so also because there is a high number of actors and many of them had just one link in the organisation rather than 175 possible connections. In a context of organised crime this might be so because, if everyone was connected to anyone, it would be easier to unveil the entire organisation. The density gives the idea of a certain "halo" that surrounds the organisation: mafia is there, but none knows where exactly.

[36]It does not simulate the existence of any other node or link.

FIGURE 3.3: Perseo Network

FIGURE 3.4: Network shaped according to Kamada-Kawaii (up), Fruchterman-Reingold (down) algorithms

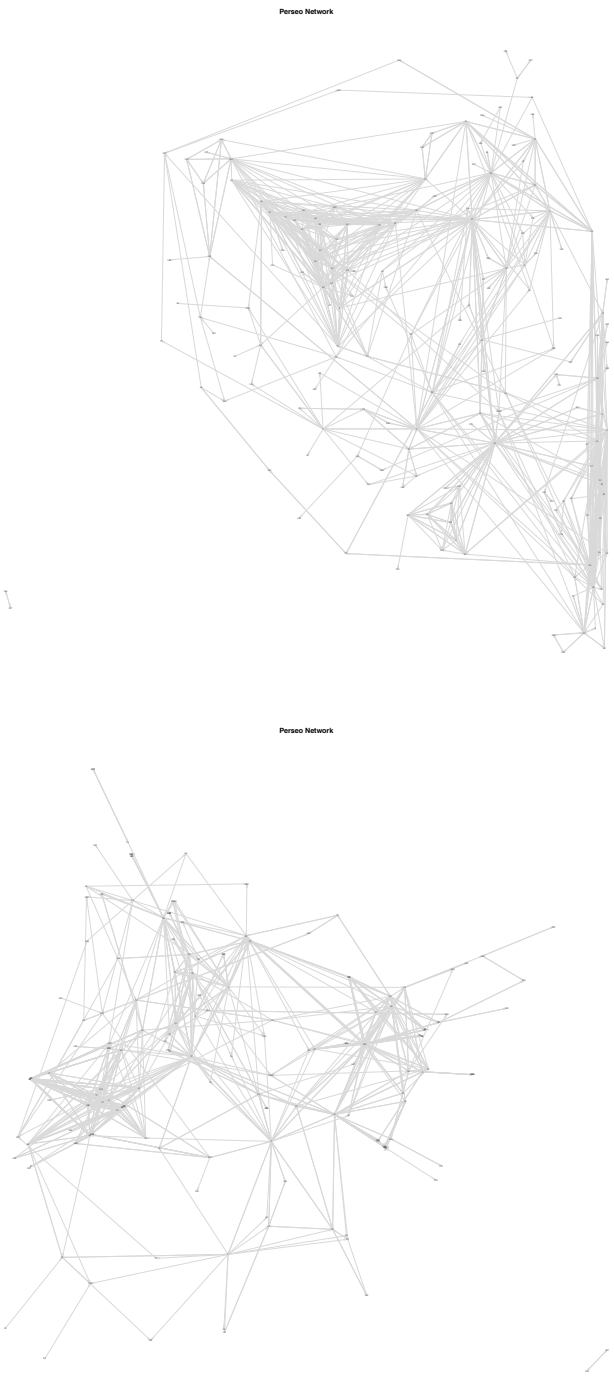


FIGURE 3.5: Network shaped according to Davidson-Harel (up) algorithm and multidimensional scaling (down)

This is a different result from Morselli's Hell's Angels where the network is more dense, which means it might be easier to penetrate than the Mafia related one. Low network density in this context is a valid indicator of the secrecy of the network. The less secret the network is, the easier it will be to be infiltrated.

Reciprocity

The reciprocity of the network is the likelihood of vertices in a directed network to be mutually linked. The reciprocity is equal to the number of links pointing in both directions divided by the total number of links. In the Perseo Network this value is above the set at 58%. This value tells that there are 406 mutual node pairs and 586 other connections are asymmetrical, which means that some links are not reciprocated. In a criminal organisation this is shown as someone that send orders to other actors which few times need a reply. Indeed reciprocicated links happened only in phone conversations and in documented meetings. This is another element that tells about the covertness of the network, because there are more hints to identify members of the network if links are reciprocicated.

Transitivity

The transitivity of the network is the measure of the degree to which nodes in a graph tend to cluster together. It is also known as the *clustering coefficient* and, in other words, it represents a number of closed triads within the network. It can be global, if it tells the number of closed triplets over the total number of triplets, or local, if it tells how close its neighbours are in order to be a complete graph. In the case of the network the global transitivity is rather low and equal to 35%. As a relative result this gives however an interesting outcome. Every two random members inside the organisation there is a 35% chance that they share a link with a common member. Transitivity allows the specification and computation of all the possible triadic relations inside the network (Davis and Leinhardt, 1967). Every triple of vertices (A,B,C) are classified into the 16 possible states. Empty graphs A,B,C: 706537; single directed edge A->B,C: 61456; A<->B,C one mutual connection between two vertices: 29825; A<-B->C the *out-star*: 1821;A->B<-C the *in-star*: 879; A->B->C

directed line: 692; A<->B<-C 896; A<->B->C 856; A->B<-C,A->C 494; A<-B<-C,A->C 4; A<->B<->C 531; A<-B->C,A<->C 164; A->B<-C,A<->C 85; A->B->C,A<->C 75; A->B<->C,A<->C 90; A<->B<->C,A<->C the complete 135. In the terms of the darkness of the network this tells that even if arrested and willing to collaborate with detectives, a member of the organisation might just reveal a small portion of the entire network, but it might be fairly easy to connect members to other affiliated ones.

This result can be comparable with the structure of covert networks inside Massively-multiplayer online games (MMOGs) (Keegan, 2011) where players act in a virtual world while *drifting* their behavior in another social context. High transitivity is found when players give or take resources (for free or in exchange of real/virtual money) expecting to get other resources from a third party. In the virtual world, high transitivity is associated with antisocial behaviors like stealing are tolerated or encouraged while gaming. In the terms of the criminal organisation this is comparable with Provenzano rules where none can join the organisation if he is not presented by a third party that is already a member.

Diameter and Node Degree

A measure of the trade-off between the efficiency and security of the network is analysed through its diameter and degree of the network.

The diameter of the network is the length of the longest geodesic distance between nodes. It looks for the length of the shortest path between two nodes. In this network it is equal to 8 and it means that the network is fairly wide, as it might take for two random members of the organisation up to 8 common members to find a connection.

The diameter gives the sense of how the network is *long*, i.e. spread around multiple actors. This implies that spreading any information could be time consuming and given the actors' distribution it has to pass through specific actors that then convey the information to the specific recipients. This implies that if the information takes long time to reach the recipient it is less efficient, but it is harder to breach it at the same time. The reason stands on the fact that, even if there is the chance to intercept communication, it makes less evident the direction of the message.

A practical example is provided looking at the network. The furthest distance in the network is between nodes 77 and 175. This means that every

message from one to the other has to pass this sequence (as it is considered only the shortest path possible): 77->117->112->101->9->1->23->162->175. The following picture 3.6 shows the diameter of the network with a yellow line. It must be taken into account that in order to make sure the message can be reached it has to pass through two "hubs"; 9 and 1 which are among the most connected actors in the network.

This is the case when the boss Bernardo Provenzano had to be captured. Detectives followed the path of a box of clean laundry to reach his hands. This process took months and multiple times this same box went back and forth from multiple affiliated members before getting to its destination. Therefore, even if it was not the most time-effective measure, it proved to be efficient in mantaining the security of the network for almost 45 years.

The degree of the vertices is the number of adjacent vertices, i.e. the number of connections it has. The most connected nodes are 32 and 60, which have 167 connections each. The graphical representation 3.7 of the network demonstrates which are the most connected nodes. From the degree of the network 3.9 it is possible to compute the degree distribution, which is the fraction of nodes in the network with that same degree. The histogram below 3.8 shows that the majority of the nodes have a low degree with few highly connected nodes. The degree of vertices and the degree distribution give exactly the picture of who is the most essential member for the organisation. Those are the targeted members during the investigation. Their arrest potentially can break the network and it affects the efficiency of the entire organisation. The adjacency gives signals also of a specific information protocols. As stated before, if an affiliated has to give a message to someone inside the organisation that he does not know the most efficient way is to send it to the most connected actors he is linked with. This analysis on the network degree leads towards a broader discussion about centrality of a network.

Centrality

The centrality degree of the network shows the number of links that are incident upon a node. There are many ways to compute centrality and all of them give different nuances to the network.

It is possible to compute it in a general way, using the indegree or outdegree which is the number of ties directed to the node and the second is the number of ties that a node directs. In other words, centrality degree in a

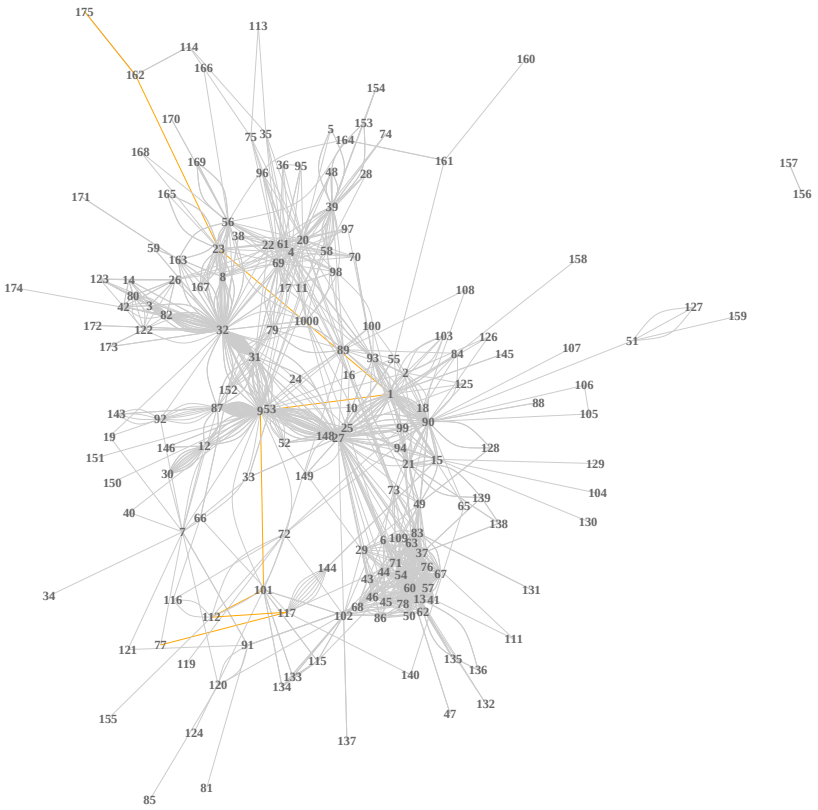


FIGURE 3.6: Diameter of the network

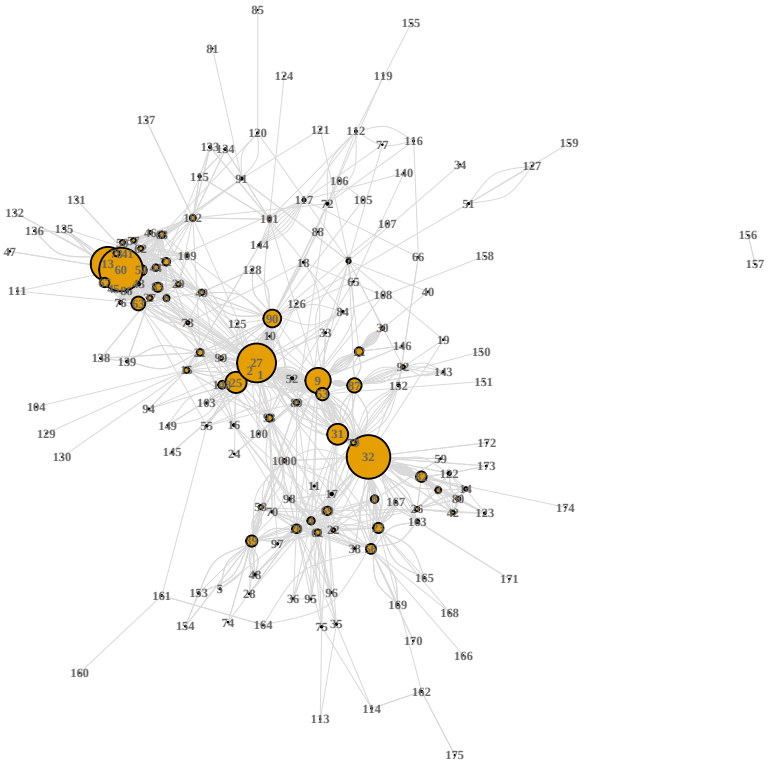


FIGURE 3.7: Degree of the nodes

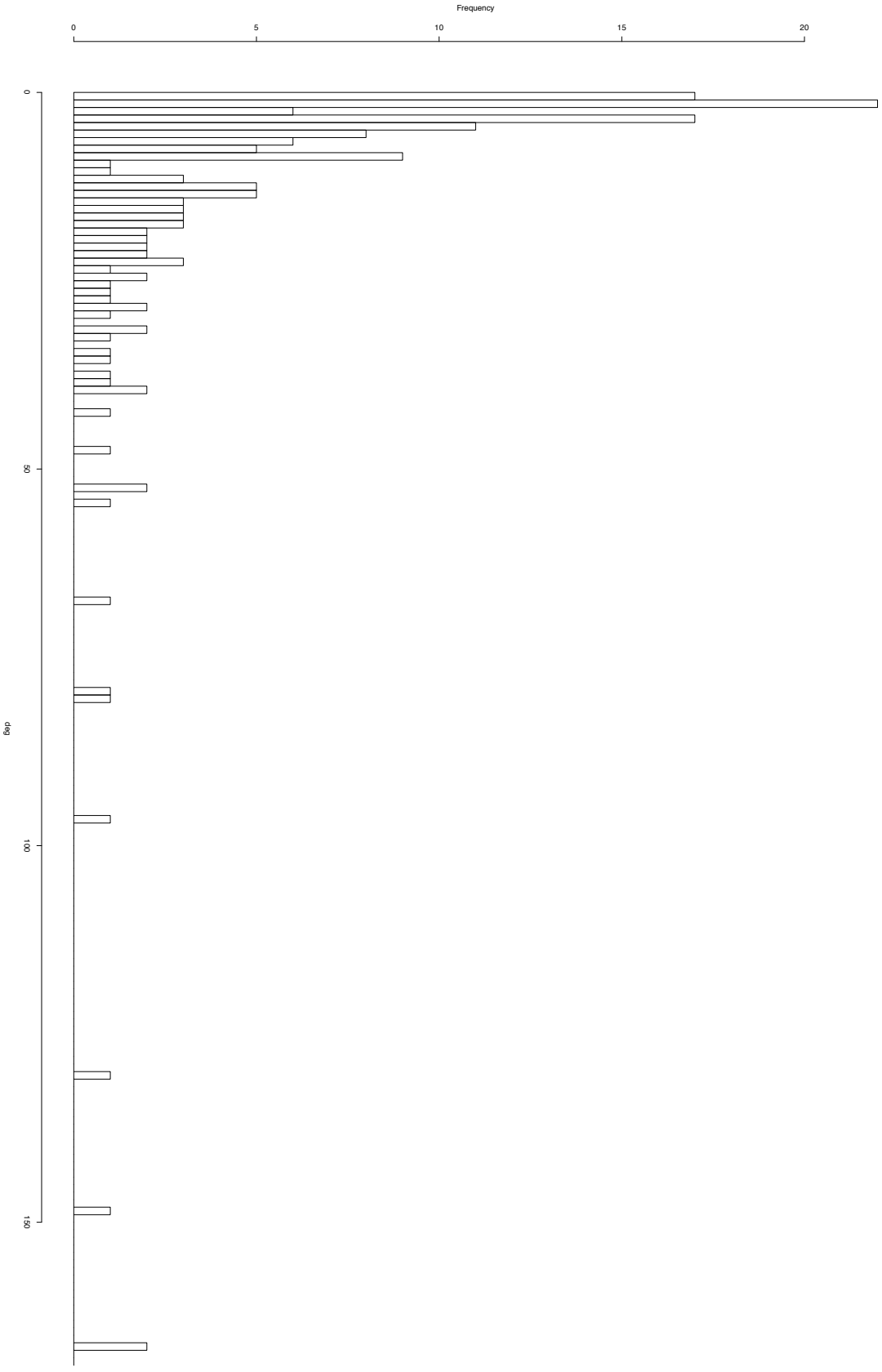


FIGURE 3.8: Histogram of nodes

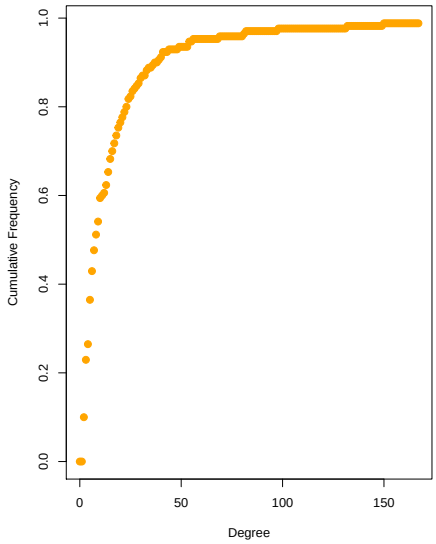


FIGURE 3.9: Degree distribution

criminal network shows who got the most contacts (total centrality degree), who was more active (centrality degree out) and who got more information as it was the most contacted (centrality degree in). This table does not tell however how many actual contacts there were between each node. This is an information that can be retrieved looking at table 3.3 where it is considered the number of actual links that involve each node. Therefore looking the distribution shows the overall activity that each node had. This activity can be broken down considering centrality degree and it is done through considering the direction of the link. Considering the incoming links, actors 60, 32,13,27,31,25,90,9,63,87 are the most contacted ones, with 50% of links received by these nodes. Similar results can be shown with the outcoming links where the most "active" actors are 32,27,60,13,9,25,31,1,90,69, with 44% of links sent by the nodes. Those results can be compared with Morselli's network which had a smaller degree of centrality than Perseo network. This difference is justified by the different degree distribution that the network had.

It is possible to compute node closeness too, (Boldi and Vigna 2014) which isconnected to this feature of the network. As for the centrality degree, it was computed the total closeness and the in and out degrees.

The closeness is the distance of one node to the others in the group. For this kind of network it means how many people it takes to reach the most

peripheral nodes. The outdegree considers how many steps it takes for that degree to reach the border of the network and the indegree considers the number of people necessary to reach that specific member. In such a case it sat generally on 11%, which means that the majority of nodes are not very close to each other. Network closeness shows a different nuance than transitivity. It states that if a node has more links it might have a bigger access to gather information from other sources but also somehow to influence the close ones.

The betweenness centrality is rather small too, around 27% for the whole network. It quantifies the number of times a node acts as a bridge along the shortest path between two other nodes (Freeman, 1977). Similarly for the edge betweenness which is the number of shortest paths going through an edge. Results considered the maximum number of triads possible in the network. Betweeness gives also the idea of which nodes are more apt to get closed triads between them.This value can show which nodes have a bigger influence over the information, which means that if those actors will not bridge links the network might be disrupted or less efficient. If compared with terrorist networks (Morselli et al. 2007), the overall betweeness value is much higher here. This is due because the structure of a terrorist network lacks of a real central operative core and this makes it harder to detect it in spite of its internal efficiency.

The eigenvector centrality index is the measure of the influence of a node in a network. Therefore, larger scores are assigned to nodes that are linked directly to more relevant nodes and lower scores are given to nodes linked to the least relevant ones. This index has already been employed in literature (Mastrobuoni and Patacchini, 2012, Berlusconi et al., 2016) in order to predict the leader of the network and to predict unobserved links in the network. In the network its value is on 46% which means that high scoring nodes are rather close to each other.

Node Distances

Once centrality measures were calculated, the analysis concentrated on node distances. As stated in R *igraph* package description it "calculates the length of all the shortest paths from or to the vertices in the network" (Csardi and Nepusz, 2006). Now, when this value is zero it means that either the

node is not reachable or that it is heading to itself.All the paths were calculated considering the distance from actor number 1 for example. Therefore actor 1 has zero distance from itself. All the other measures consider the distance from (distances out) or to (distances in) the actor number one. The total (distances all) considers the graph as undirected.

The measure of distances gives an idea on how the information flows in and out a network. In the case of this network it is evident that the closest nodes can transmit information in and out more quickly but given the direction of those links it might take more people for the news to be reached or received. This is an information that is justified for the kind of network shown here. It makes sense for a covert network to have different procedures in order to receive or send information. An investigator might follow the link from one contact to another and probably wait for a reply tat comes in that same direction. In this way, if one of the links or actors is wiretapped or controlled, it might be easier to send the information through a different mean or person.

In the search of the way information is spread and how quickly inside the network, the mean distance was computed too. This is the mean of the shortest distance between each pair of nodes in the network in both directions for directed graphs. This value gives an idea of the verticality of the network: i.e. how quickly the information flows from one side to the other. Despite the dimension of the network this value was relatively small and equal to 3.5 people.

Node Coupling

Another information extracted from the network is about cocitation coupling. This analysis takes two nodes and checks for each couple how many shared nodes they have (Csardi and Nepusz, 2006). Even if this is a kind of analysis that it is used for bibliographic references, it offers the advantage to tell how many "common friends" each couple of nodes have. Therefore the result,show the most popular nodes in the network. This kind of analysis can be run for each node and for istance it was run reporting the level of cocitation of each node with number one.[37] However, the drawback is that revealing the most *popular* nodes in the network is not *per se* significant of the

[37] In this example, actor number one shares 5 other "friends" with actor number 27, giving the idea that if they are not connected together there is the possibility they might be more than others with a lower cocitation number.

most important one in the hierarchy of the network. The control of the most relevant nodes is otherwise paramount to uncover less frequent and more periferal nodes, as well as those that are actually part of the deep core of the network.

Decomposition

One other check on the network, which somewhat similar to cocitation, is run through k-core decomposition. "The k-core of graph is a maximal subgraph in which each vertex has at least degree k" (Csardi and Nepusz, 2006). This k-core aims to show how inclusive is the graph to other nodes. This is a relevant feature given the fact that it allows to asses if inside the organisation there are different subgroups. Those groups are identified according to the functions they perform. Therefore, if two nodes have some level of similarity in the way they are linked to other nodes or in the way they spread or receive information they will be connected together in the same group.

In the case of this mafia network, there are 12 different subgroups and 6 groups for the in-core and out-core. Those last two show subgroups given the flow of information that is sent or reached. The intuition given from the k-core is that it is possible to create a functional hierarchy in a real network using those tools by showing the connectivity of nodes.

In a covert network it is relevant to show which nodes do most of the job carrying information. Although this algorithm was first developed for web pages, it has transmigrated to social sciences network applications too (Deguchi et al., 2014).

Hubs and Authorities

Through the hubs and authorities algorithm, also known as hyperlink-induced topic search, (Kleinberg, 1999) it is possible to compute and show in figure 3.10 who carries more information, having a large number of outcoming links. In the context of an organised crime network this might be the role of mid-low level members that have to set out tasks. Simultaneously, the authorities are the nodes that get incoming links from hubs. In such a network this is the role of low ranked members that receive orders and high ranked affiliates that get informations in this way.

These results however are highly biased by the way the information was collected. In fact, as stated before, many wiretaps were recorded in the same

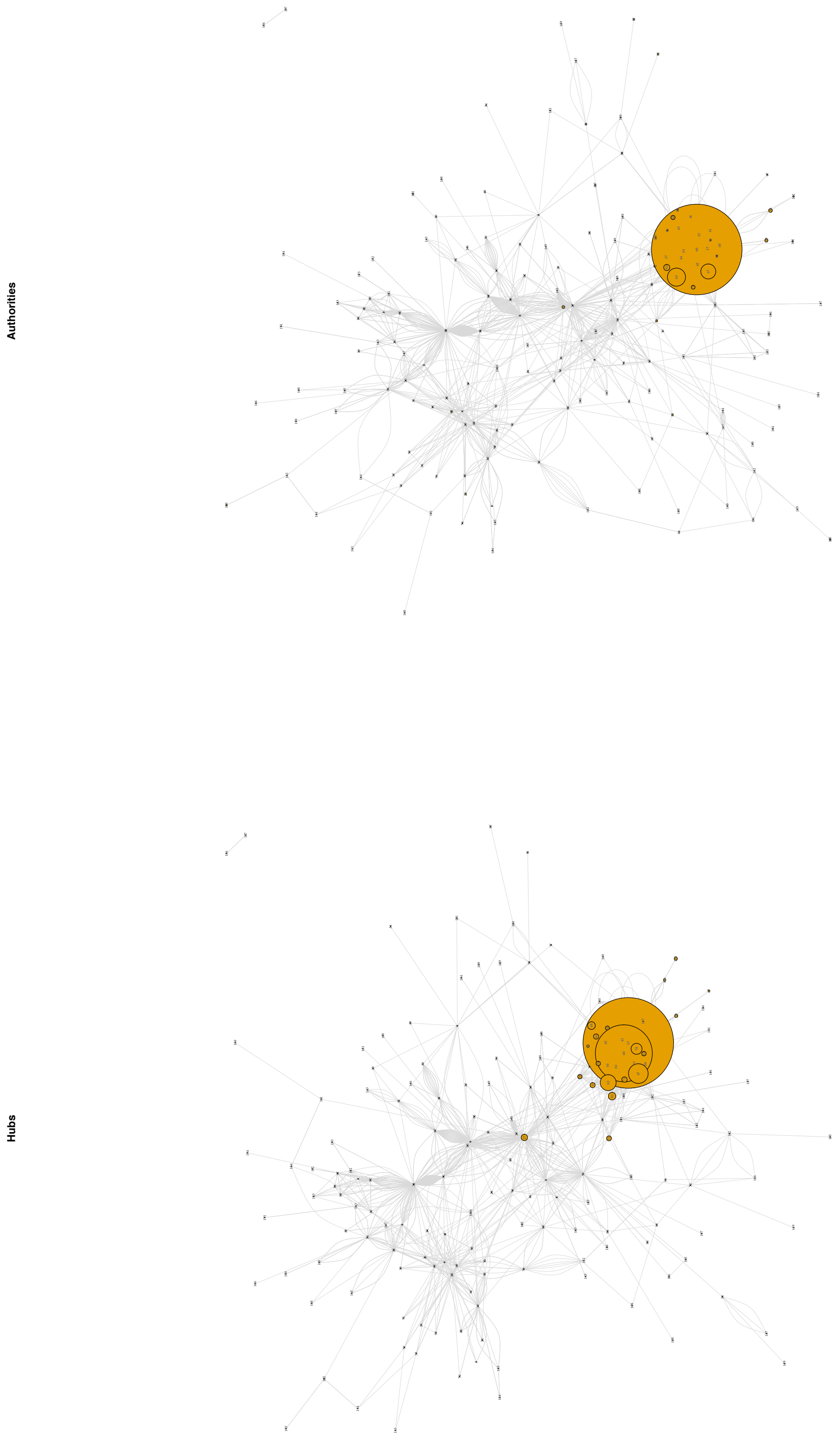


FIGURE 3.10: Hubs and Authorities

places. This means that possibly the same people had access to information continuously.

In mafia-type networks also the person that receives the most links is not likely to be the most important (Mastrobuoni and Patacchini, 2012 and Varese, 2006).[38] Therefore, even if results might not be accurate in the authorities section, there is a valid indication in the hubs graph. Actor 60 was the driver of a mandamento leader. The investigation tracked down his car and plenty of conversations inside or while driving. This allowed them to unveil the rest of the organisation.

Eventually, the nodes were collected in the graph according to their affiliation to a specific family. The visualization of the network from figure 3.11 allows the assessment that there was a high number of contacts of members between families. The legend colors each family membership differently. The fact that there are multiple contacts between families differentiates this network from Krebs' (2001) terrorist cells because they operated autonomously and without contacts between groups except for a few key members. This Mafia network is also distinguishable from the structure of Mexican Cartels where there are different and almost completely separated branches and its structure is more similar to franchise companies (Wainwright, 2016).

It is possible to conclude that in the social network analysis of Organised Crime the arrest warrant is an interesting source for data extraction because of its reliability and validity. In fact, the investigators are obliged by law to include in such an act everything that was gathered about every single actor during the period of the investigation. This reduces incentives to be reticent about findings.

However, this act is also the result of a limited number of observations. Therefore, the police agents, having a limited amount of time and resources, are the ones that decide which are the actors they will investigate, and thusly the network might be biased.

It is also possible that, in a legal comparative perspective, different judicial systems do not offer the opportunity of an act where all the proofs against and in favor are put together to build the case. This is the reason why it is

[38] E.g. During the last years before his capture, Bernardo Provenzano, one of the most important bosses of the Palermo district Mafia, had only one direct link to the outside world. He got messages and information only from one person and he was sending orders in the same way via pizzini through that same canal. This system made his arrest almost impossible for 45 years.

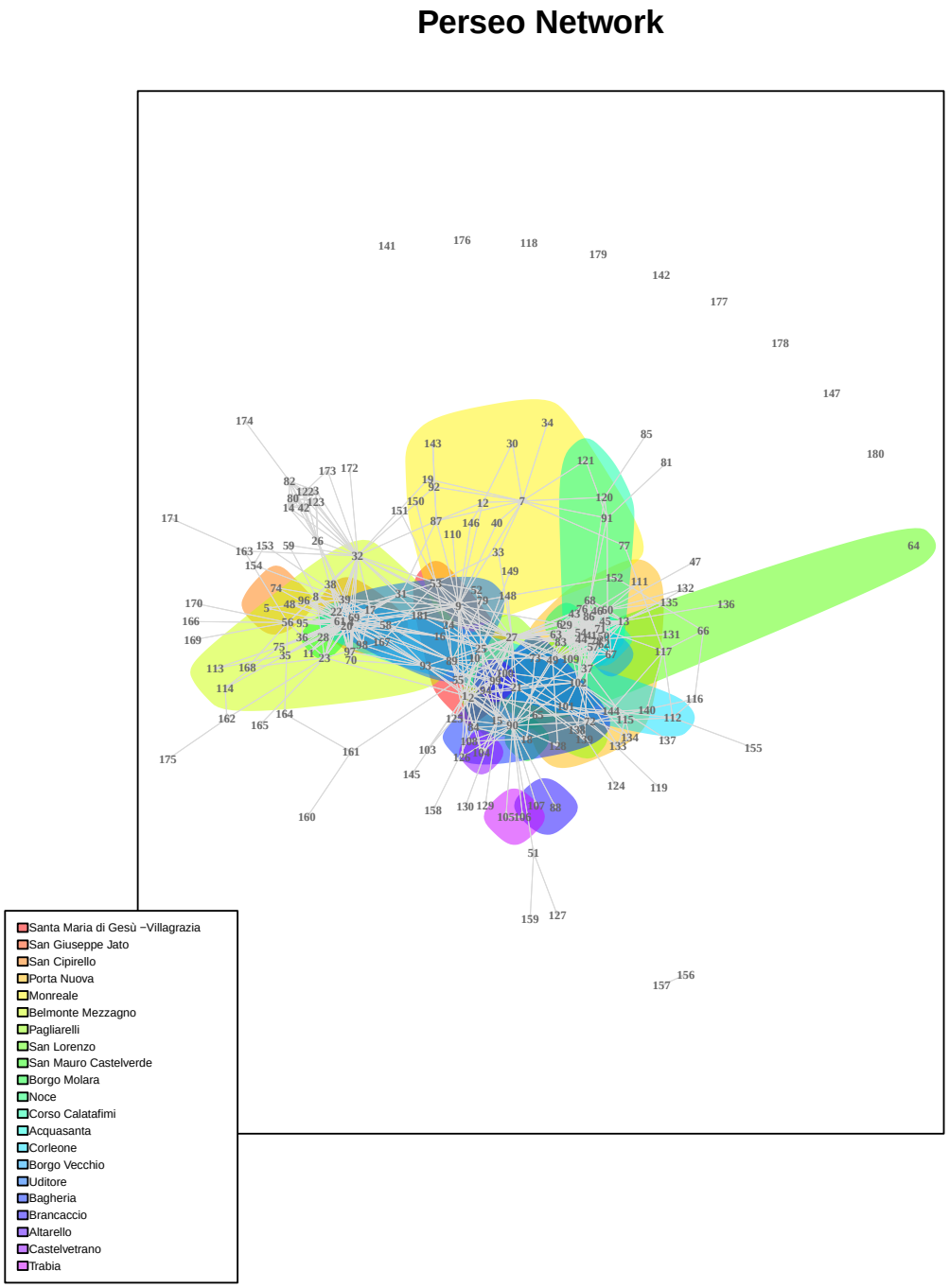


FIGURE 3.11: Perseo Network and Mafia families

not possible to leave the design of the data source out of consideration once social network analysis of criminal groups is carried through.[39]

3.3.2 Econometric Analysis

Specific features of the data highlighted for each node are now analysed with probit regressions. This is helpful to understand the marginal effects of each coefficient. *Id est* the regression aims at telling how much the probability of the outcome can change when the regressor increases or decreases of one value, while keeping other regressors costant.

For the avaliable data, a first regression was run using *R* software package *stats* and its function *glm*. In this research, a probit model is created to explain the probability that a member of the organisation would have a leading role. Every probit model needs a dependent variable that can only have two values (i.e. a binary variable). The dependent variable employed here is the direction. As it was already explained in the first part of this chapter this variable shows if the actor had leading activities inside the organisation. This fact was documented for 15 members of the organisation. Therefore, the parameters to be estimated are β_0 and β_1. Where β_0 is the directive role when x=0 and $\epsilon = 0$ and β_1 is a vector of regressor coefficients that quantify the increasing of one unit of age. The probability of having a directive role is equal to:

$$P_d = \frac{exp(\beta_0 + \beta_1 x + \epsilon)}{1 + exp(\beta_0 + \beta_1 x + \epsilon)} \tag{3.1}$$

In this model it is supposed that there is a factor that influences the outcome of the dependent variable. The year of birth of each mafioso is taken as independent variable for instance. Also the city of birth is used as a bounding variable. Results in table A.1 show that increasing of one unit on the year of birth of a mafioso, the probability of having a directive role in the organisation decreases. In other words, the older the mafioso is the higher the chances are that he will have a directive role. This same regression was run also considering the age of each actor, which gives the inverse result.

[39] This is so in the United Kingdom for example where one wiretap can not be used against that same person as this is considered to be a sort of self-incrimination. Only wiretaps that are just about a crime or taken from people that talk about other people can be used as an evidence in the trial. Interestingly enough, if the same evidence was gathered outside the borders of UK, EU laws force it to use a wiretap as an evidence even if it couldn't be considered as such according to their national rules.

This same test is now replicated for all variables that were contested as an assumption of being part of a criminal organisation. Therefore the checked variables are about smuggling of drugs, guns, hosting meetings, having connections outside the mandamento, having meetings with members of the organisation, asking for *pizzo* and influencing public procurements.

Out of all controls, results are significant at 10% level for having connections outside the mandamento (linking variable) in tableA.2 and for influencing public procurement (pubProc variable) in table. Results show that the older the mafioso is, there is a higher probability he will have connections outside his own *mandamento* with fellow members and be able to influence public procurements.

After analysing the results of the network, they were afterwards evaluated keeping fixed city effects and considerations where made on the area they were living. In fact, each member of the organisation lives in a area of neighborhood that could be taken as a proxy for all the other information that was not possible to gather from the act of the trial. Even if this is not a perfect proxy, this kind of shift is useful in order to understand and explain any other social determinant that turns around the members of the organisation.

Specifically, this was done for the members of the network living in the town of Palermo. In fact, it was possible to access coeval census data for the areas they were living in before the arrest. This led to exclude from the network all the other memvers that were living outside the town. Therefore the network is now about 53 people.

For each member living in Palermo their addresses were matched using the section they were registered in. The section is a variable used in census collection which describes the smallest area of the town. Differently from Italian postcodes, they picture more in detail the area as they manage to describe even one building. In total 47 different sections were gathered. This is a smaller number than the members living in the area as 6 members were registered to be living in the same address.[40]

[40]In 4 cases those members where even living together or in the same building. This explains why considering census data has a relative goodness of fit as a proxy. Sections most of the times consider single buildings or very small areas where relatively few people live. This is also based on the assumption that people with somehwat similar interest and or skills and qualities tend to cluster together to a certain degree. If this is true, then it is possible to assume that the information gathered from census is, to a greater extent, homogeneous to what it is the personal data of affiliated members of the gang.

Once the section number was gathered it was joint with the rest of the data in a bigger dataset with all the actors from Palermo. The following matrix 3.12 shows the correlation between all the census variables.

With the help of the correlation matrix and using a stepway regression significant variables were extracted and results are shown in tables A.4 A.2 and A.3. Here again variables about the directive role, links outside the mandamento and public procurements were employed as dependent ones.

Among all the variables, in table A.2 apartments were used also as a proxy for assessing the wealth of the area where criminal members used to live. The most significant result was about the number of household buildings. Results showed that the older the Mafioso is and if he has a directive role or keeps connections outside the mandamento there is a higher possibility that they will live in areas where apartment are slightly bigger, more comfortable and in areas which look more residential. This result shows a trade-off situation between the need for an affiliated member of having a direct control on the turf he is controlling with the need of living in a hidden place.

In table A.4, same dependent variables about members tasks were compared with population data. Results show that increasing the age of the members that are alleged to have a directive role, influence on public procurements and connections outside their mandamento, there are less chances that people in that area are divorced and mostly males.[41]

Table A.3 shows also that the areas where those member live tend to be more homogeneous as there is a smaller chance that foreigners are living in that same area.[42]

Now the analysis aims to assess the determinants about the degree of connectivity of each agent in the network and about the likelihood to be condemned as mafioso.[43] In the table A.7, it was wondered if being identified as a *mafioso* is correlated with specific activities carried out by the organisation and specifically identified and punished by the Italian criminal code as distinctive of a serious illicit organisation. Results show that the fact of meeting with other people identified as affiliated members of the organisation and having links inside the organisation is positively related with being identified as part of the organisation.

[41] This result was true also for families without children, but the value was less significant.

[42] More census variables were tried with those dependent variables, but results were not more significant.

[43] More probit regressions have been tried but they were not significant enough.

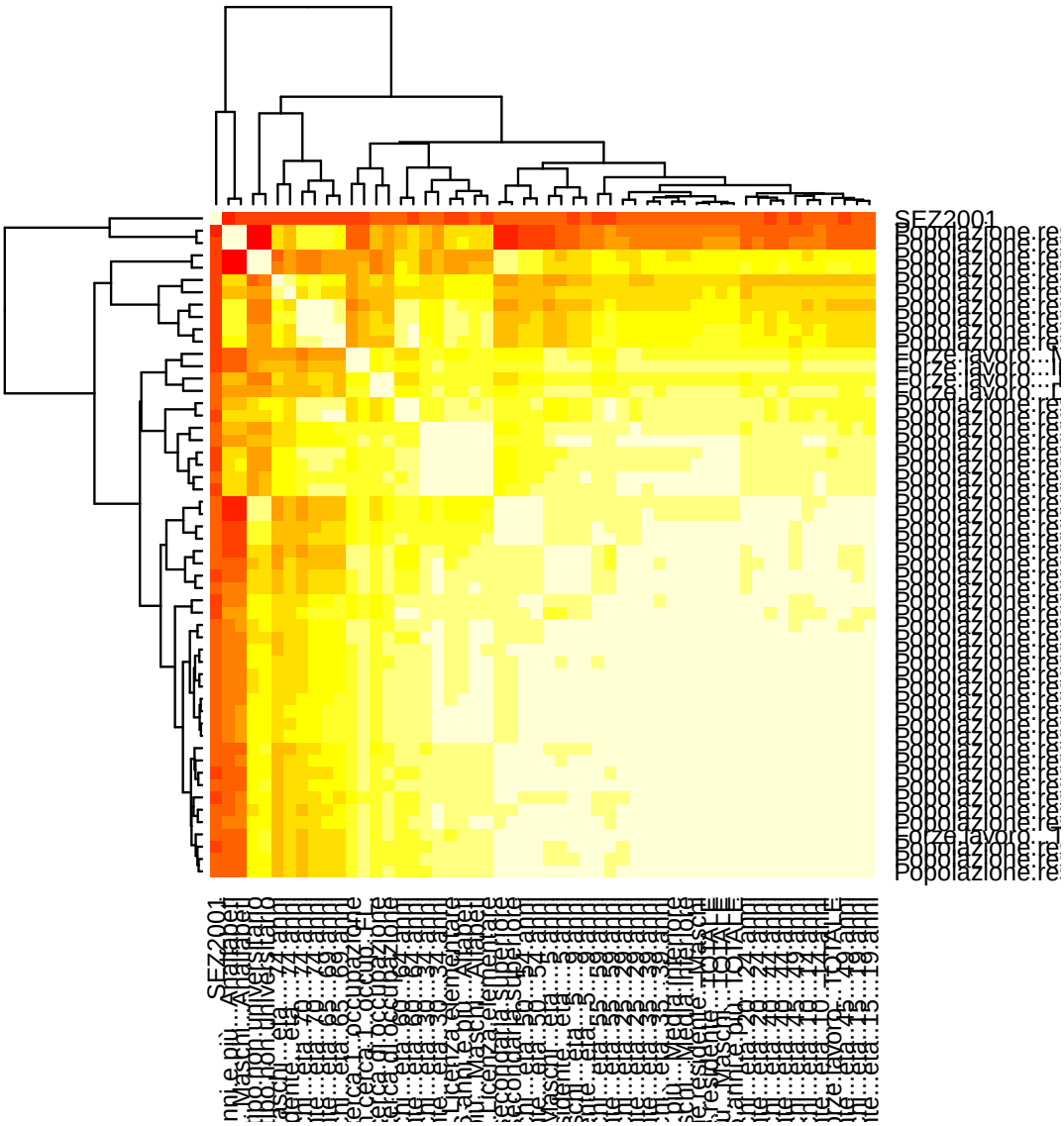


FIGURE 3.12: Matrix of the census

Age has a influence in the role of the organisation. As it is shown in table **??**, it was run a multinominal logit where the determinants of being inside the organisation are checked with the age and with the number of links of each node. With linkings outside the mandamento it is possible to notice that the degree of actors is negative and significant.

Eventually, it was extracted from the network information about the number of links per node. The total degree of nodes is measured as dependent term of a OLS regression with age, direction, influence of public procurements, meetings variable and links outside the mandamento. The intuition stands on the fact that the more links one node has the more important it is for the flow of information inside the network. Results in A.5 show that having a directive role inside the organisation and influencing public procurements is highly correlated with the total node degree of the actors. However, increasing the node degree of one unit the chance for that node to have a directive role decreases.

Those results will be rediscussed in 4 where the space dimension is going to be analysed.

Chapter 4

Spatial Network Analisys on Organised Crime

4.1 Introduction

Spatial Network Analysis is that specific technique where a graph is placed in a metric space and distance functions are applied in order to get more information about the nodes and the links.

A spatial network brings more information than a regular network, as was shown already in chapter 2. Once a network is placed into a space, it gives a more accurate description on the strenght, distance and density of the nodes.

The fact that the network operates in a "limited" space means that link predictions and explanations can be drawn from the number and length of links. Therefore, nodes that are closer have a higher probability of being linked together than those who are far from each other. I.e. the probability of two nodes being connected decreases with the distance. Furthermore, two nodes that have a long link and which are far from each other have a higher cost to maintain that relation. This "cost" is translated into a larger degree or into an higher traffic of the network (Hassan et al. 2017).

In this part, spatial network analysis is going to be applied to the Perseo network, which is a Mafia-type criminal organisation. As stated before, there is no previous literature about the application of this analysis on organised crime.[1] This chapter is focused on building the network on a map and it is going to be followed by a study of its spatial effects on the data collected.

In the beginning of the chapter it is shown how the network was put in a spatial environment, followed by the econometric analysis.

[1] However, geographical theories of crime trace back to the 20s with the School of Chicago.

FIGURE 4.1: Mandamenti in Palermo according to DNA

4.2 Operation Perseo on a Map

Spatial network analysis requires that nodes are geometrically located in a space. The Perseo network as it was created and shaped in 3 does not say much about the way it is embedded in the area of Palermo district nor about the neighbors of each node. This is why the network is now overlayed on a geographical map of the Palermo district.

The Italian National Crime Agency for Mafia-type crimes, the *Direzione Nazionale Antimafia* (or DNA), already created in the past network visualizations of the Mafia in Palermo district.[2] Figures 4.1 and 4.2, taken from the 2012 report of DNA, show how Palermo and its district are shared among different mandamenti. The analysis conducted here is going to show how this result can be reached considering spatial coordinates of affiliated members.

This operation follows three steps:

- Extraction of the relational data from the data source
- Creation of the map
- Juxtaposition of geolocalised nodes on the map according to the edge-list

In order to create the network map multiple approaches and softwares were used. Datasets were listed with Excel. R, Carto, ArcGis and GoogleEarth

[2]DNA coordinates multiple smaller and local agencies in the fight of organised crime and terrorism, which are called *direzione distrettuale antimafia* and are one per district of the court of appeal.

FIGURE 4.2: Mandamenti in Palermo district according to DNA

7 softwares were used to code the map and the network. Structured Query Language (SQL) and geographic information systems (GIS) were also applied to manipulate relational and geographical data.

4.2.1 Data Extraction

Once the data for the case study was accessed, coordinates were needed in order to trace back nodes on the map. Therefore, addresses were extracted by hand from the data source for each node of the network. As stated before, the addresses add information about the quality of the links. Another aspect to be considered is that the majority of the meetings was set in the same places. This might be an indication of a criminal pattern, but this is biased by the same fact that this was where the majority of events were observed by investigators (Brantingham and Brantingham, 2008; Campana, 2016). In total, out of all 176 nodes it was possible to gather the addresses for 111 actors only. When the address was not available nor was it possible to track it down on the map, it was assigned a "ghost" location on the map.

Space

A dataset containing space variables was added to the one described in 3 for spatial network analysis purposes. Literature has already focused on the procedures of analysing data in a spatial context (Bailey and Gatrell, 1995)

and applied that successfully in the analysis of crime (Haining, 2003). 9 variables were collected.

Id: This variable collects the ID of the people that participated in meetings. It also helped to identify isolated nodes of the network (Berlusconi et al., 2016). Here again, there is a small group of members that is highly connected and many peripheral nodes (Morselli, 2009). In particular, out of 176 members, 112 actors did not take part in meetings. This information gives the idea that there were more intense connections among 64 actors, which helps to contour the core of an organisation that sets out meetings only when it is strictly necessary.

This is higly consistent with other criminal networks (Morselli, 2009 and Berlusconi et al., 2016), but there are some differences. The Perseo network is very different from the Hell's Angels one analysed by Morselli. This is so in the way meetings are intended to be in the different organisations. In an organisation such as *Cosa Nostra* meetings tend to be avoided, reduced to a minimum because the social gathering reveals patterns and contacts that are vital for investigators. In the other organisation, meetings are more frequent as every part of their social life turns around them, with different levels of meetings where members are allowed to join.

Addrspat: Address of the location. The addresses are not going to be shared in the network but they were part of the social network analysis. The addresses of meetings added more information about the quality of the links. The majority of the meetings took place in the same places. This might be an indication of a criminal pattern but this is biased by the same fact that this was where the majority of events were observed by investigators (Brantingham and Brantingham, 2008; Campana, 2005).

Cityspat: City where the meeting took place. Almost all of them were in Palermo, followed by Monreale and Bagheria.

Lat and **Long**: Latitude and Longitude. These coordinates were extracted from the address thanks to R package through a personalised geocoding function. The geographical coordinates were useful for plotting the network on a map.

Timespat: This variable recorded the time of a localized event. The events come from early 2006 and end in November 2008.

Typespat: This is the type of the localized event. The event could be a meeting: i.e. a place where more actors gathered together; or an address: where the actor lived or used to work.

Doc and **docp** these variables had the function to take note where all the events and addresses were gathered in the document.

Map Design and Placing the Network on the Map

At the very beginning, the map was created using shapefiles of Italy, Sicily and Palermo district loaded in R. All the addresses were geolocalised customising the function `geocode()` from the package `ggmap`.

This function finds the location (i.e. latitude and longitude) using either (1) the Data Science Toolkit or (2) Google Maps.[3] The advantage of such a procedure is that it coded automatically a large set of addresses instead of doing them one by one. It also collected them in a separate list with all the details.

Figure 4.3 shows the result obtained on the shapefiles using R packages `maps` and `geosphere`. The advantage of preloaded maps in R was the extremely intuitive use of the function and the possibility to mold the network as needed. Unfortunately, the network visualization was poor from the point of view that it did not allowed a more detalied fit for the nodes. This implied that it was not really clear how all the relationships where displaced in the network.

Such a problem was solved once the map was designed with Google Maps system. However, the process was long, less intuitive and it did not allow to show the links on the map, as it can be seen from figure 4.4.[4] Even if the map was more detailed now it did not allow a proper specification for the network. Still, few operations were performed like the density of the network in palermo (Figure 4.5).

Eventually, this same task was done using Carto software. The result is shown in figures 4.6, 4.7, 4.8. The advantage of using this software was that the entire map and network was automatically designed and set up quickly and intuitively.

Once the network was loaded, SQL language was used in order to get connect the geolocalised nodes and connect them according to a specific edge list.[5]

[3]The geocoding function was similar to the one found here and it was used in respect of Google Maps APIs.

[4]It was possible though to trace the route from one node to the other.

[5]This is was translated in the following queries A.6.
Where `Coords dat` and `my links` are respectively the nodelist and the edgelist. `The geom webmercator` is the coordinate system adopted by the software that locates points

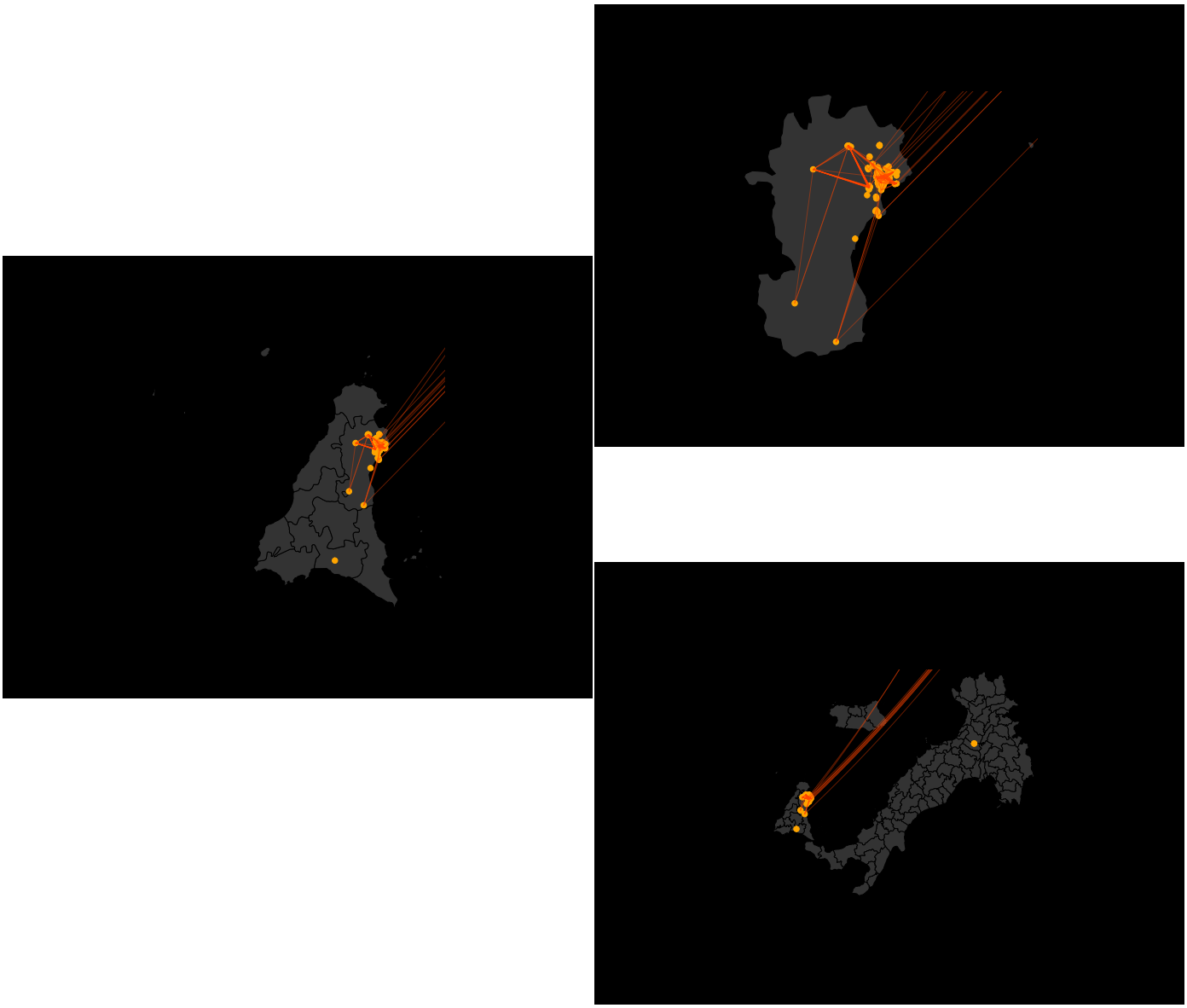


FIGURE 4.3: Network Visualization with R

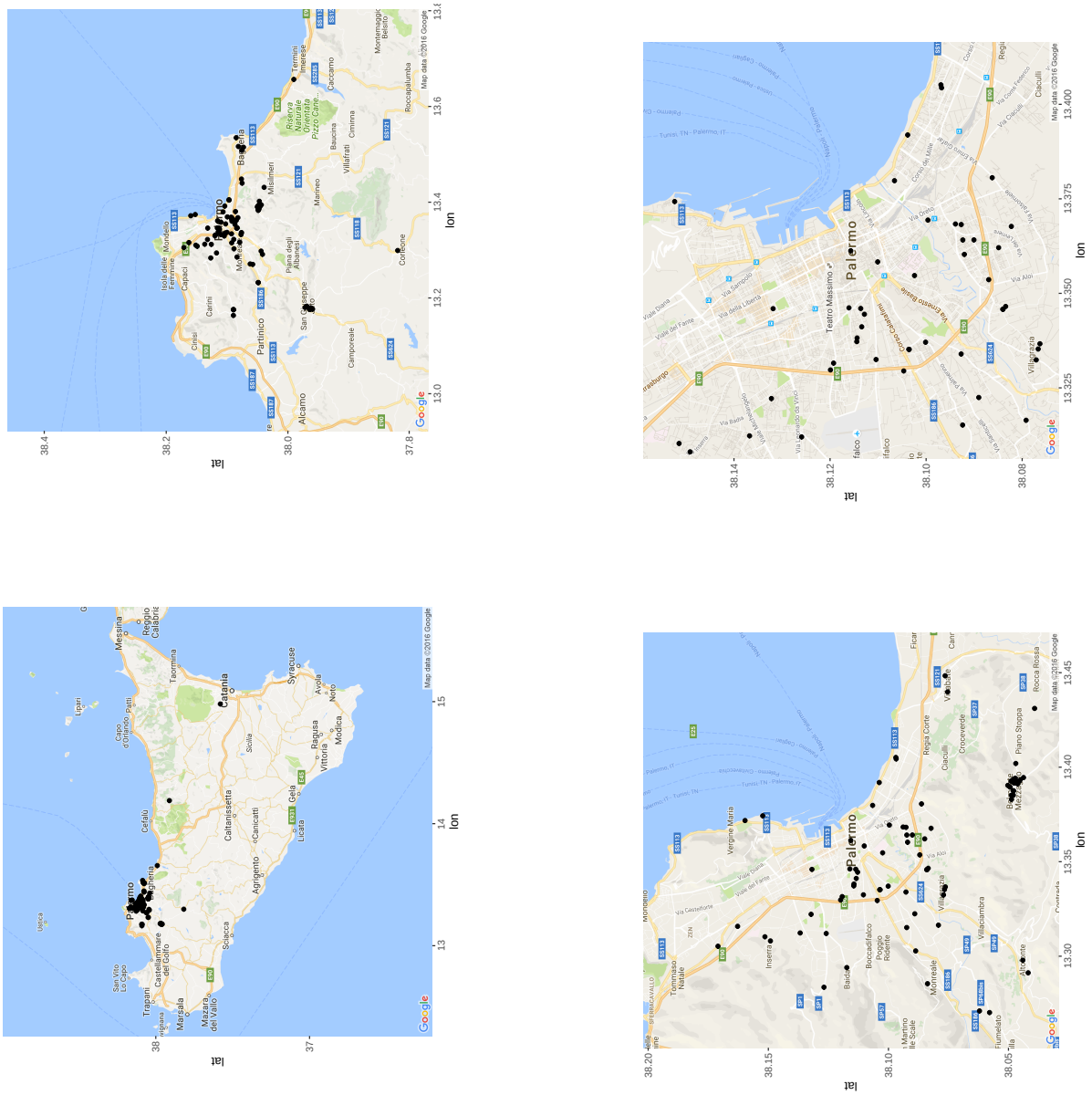


FIGURE 4.4: Network Visualization with Google Maps

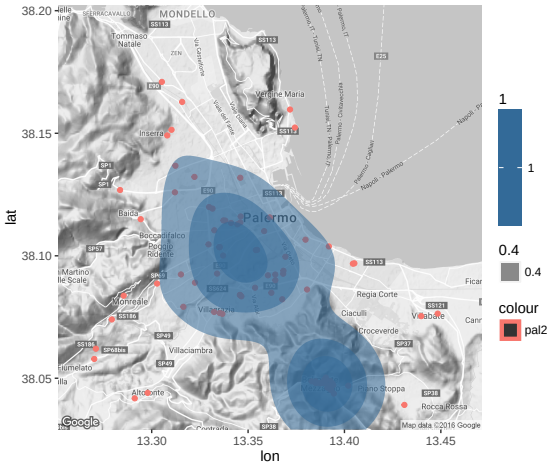


FIGURE 4.5: Mafia density in Palermo

Eventually, over the nodes and links it was drawn an heatmap that shows where there was a higher density of members of the mafia according to the police operation.

It is evident from figures 3.1, 4.6 and 4.7 that the operation concentrated only on the Mafia operating in the west side of Sicily. In the district of Palermo many members concentrated in the surrounding towns of Belmonte Mezzagno, Bagheria, Montelepre, Altofonte, Monreale, San Giuseppe Jato and San Cipirello.

In this sense it is possible to draw some remarks just from the visualisation of the network embedded in the map. It is possible to notice that nodes are heavily clustered together. This means that it is possible to recognise and classify different groups among the affiliated members according to their geographical location. This is more evident in smaller towns[6] around Palermo where actually most of the streets had affiliated members that were arrested.

in a sphere, which is common in web mapping systems. `The geom webmercator` is a variation of Mercator projections. More information can be found here.

[6]See for example the network map 3.1 of Belmonte Mezzagno in 2 or San Cipirello clustering in 4.8.

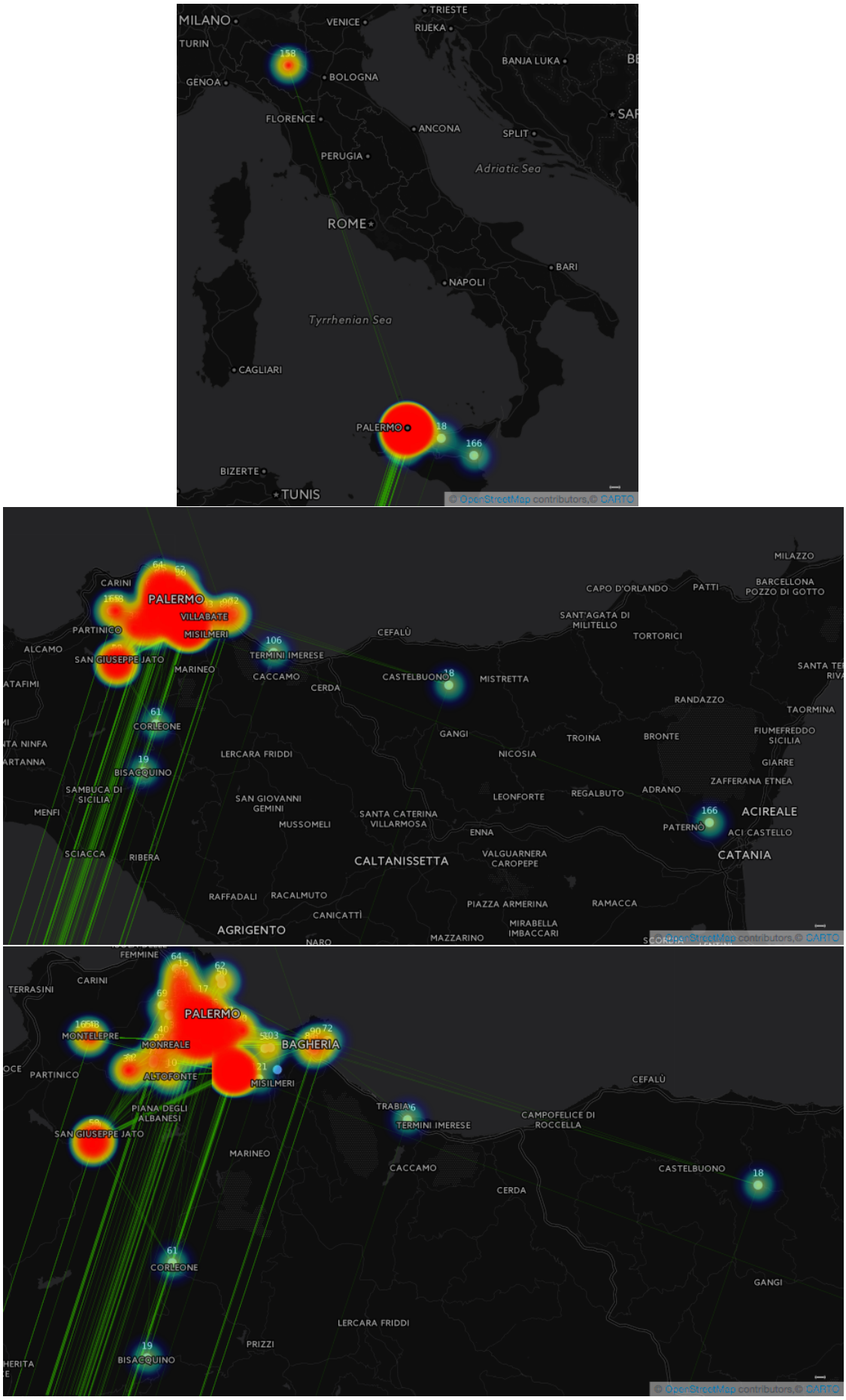


FIGURE 4.6: Network Visualization with Carto 1

FIGURE 4.7: Network Visualization with Carto 2

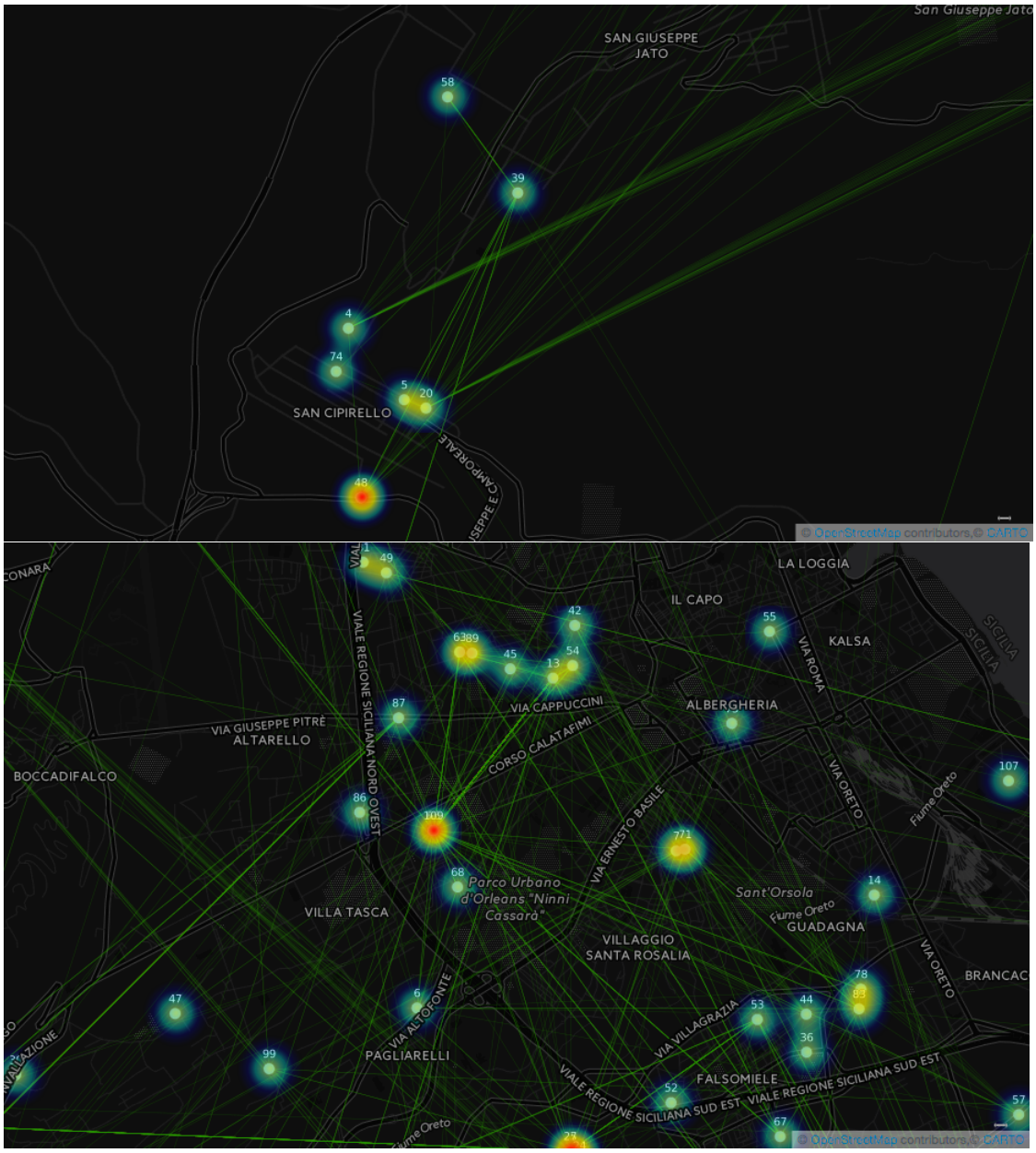


FIGURE 4.8: Network Visualization with Carto 3

The links between members do not behave like delaunay triangulations. Those are triangulations where the inner part of the triangle is an empty circle (Okabe, 1992). It is evident that links are more instricated 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[7] that a node is connected also with others that are not the closest ones. This spatial network is not a planar network. A node that is allowed to be connected with others that are not the closest ones is the sign that the spatial network is not planar. A classical example of a planar network is the road or pipe network because each node is connected in a way that link intersections are not possible (Barthelemy, 2011).

Next section focuses on the spatial effects of the network.

4.3 Spatial Network Analysis on Perseo

In this Section an econometric analysis of the determinants of link formation is carried out, following Jackson's (2008) book. In particular, the role of spatial effects along the links will be identified as suggested by Neumayer and Plümper (2010).

4.3.1 The Model

The model is taken from Le Sage and Pace (2008). It is a mixed spatial autoregressive SAR probit model. It is a mixed one because the spatially lagged value of the dependent variable, y is estimated with an additional predictor variable which is not spatially embedded. Therefore the model looks more like a linear regression like the one showed in 4.1, where *X(n k)* is the matrix of covariates and *W(n n)* is an inverted distance matrix of the actors.

$$y = X\beta + \rho W y + \epsilon \tag{4.1}$$

In such an equation terms ρ and β are the parameter vectors to be estimated and ρ is supposed to be $\rho < 1$.

[7] Even if the probability decreases with the distance.

4.3.2 The Distance Matrix

The distance matrix aims to represent the connectivities between observations. The matrix is built using an inverse scaled process in order to show proximity between nodes.

First of all the latitude and longitude of each actor are collected in a dataset. This coordinates are placed for each actor in a distance matrix *W(n n)* using *distm* function. However, such a distance matrix is not very functional for the analysis as it is. The way distance is measured, it considers objects that are far from each node. In a spatial network context, it is more useful to consider inverse distances, because of the higher probability that closer nodes have a bigger influence between each other. In order to make the data comparable, the inverse distance matrix was also standardised to 1. Thresholds were put at 0.2, 0.3 and 0.5, as it is shown in matrices 4.9 and 4.10 in order to get more information about the nearest neighbors.

For models' sake another inverted distance matrix *W(n x n)* standardized to 1 was built, as it is shown in 4.11. In such a matrix the treshold is set within the 5 nearest neighbors. Which means that each node is clustered together with its five closer neighbors. This gives information about the dependence structure between neighboring observations. In fact, each node that falls within this limit is joined together in order to create a new node. This matrix was afterwards employed in SAR probit regressions.

4.4 Results and Conclusions

The probit model is created hereby still to explain the probability that a member of the organisation would have a leading role. However this time it is going to be estimated with a matrix which is embedded in a space. The matrix emplyed here considers spatial clusters of 5 members together and compares it with other clusters that are close together. The idea is that neighbor clusters might have similar values (giving a positive autocorrelation) or completely different values which might not be very meaningful (resulting in a negative autocorrelation). The probit model needs a dependent variable that can only have two values (i.e. a binary variable), which is y in the formula 4.1. The dependent variable employed here are again direction (which looks if a member has a directive role), linking2 (if a member has connections outside the mandamento) and pubProc (if a member has influenced public

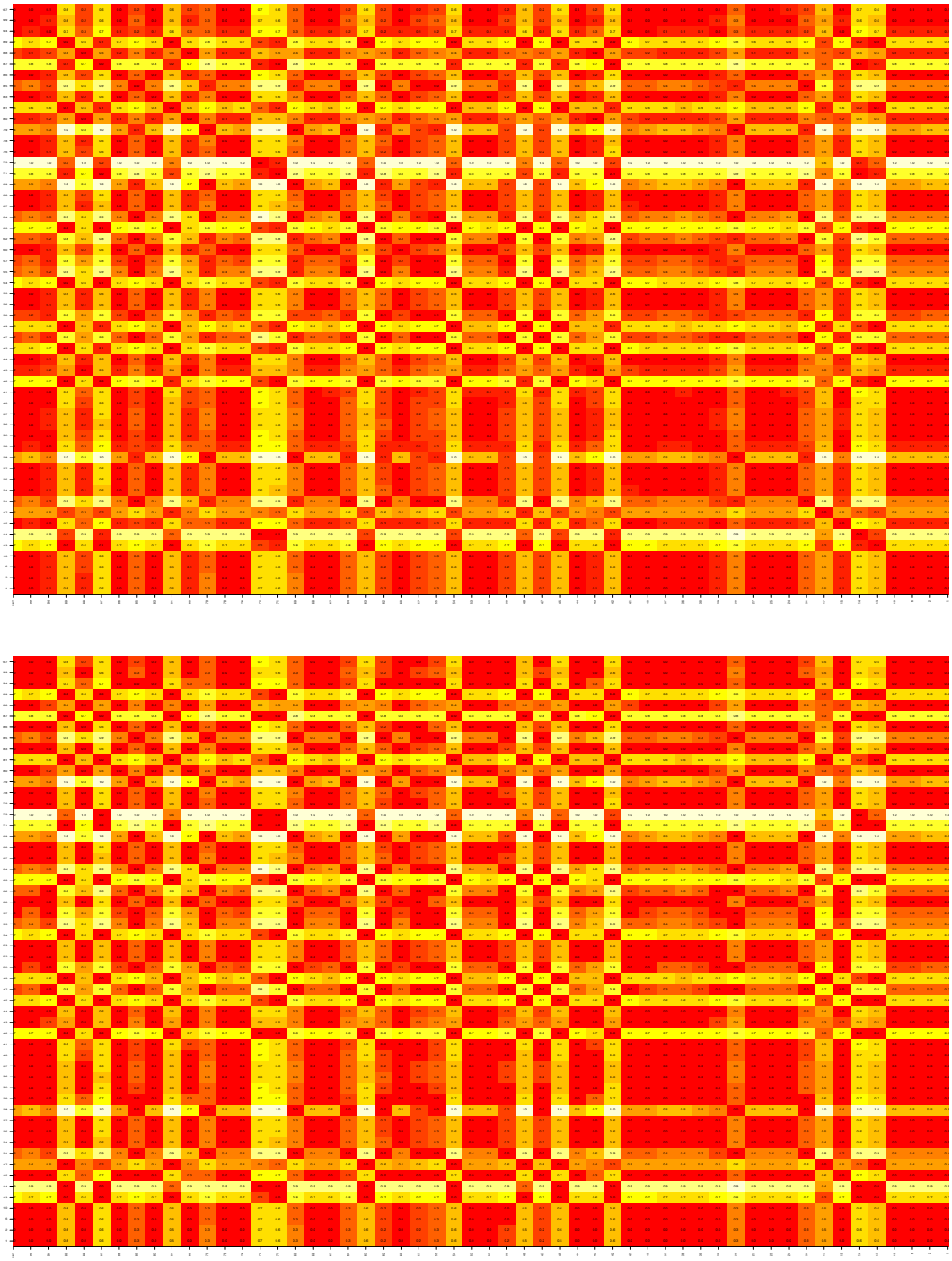


FIGURE 4.9: Distance Matrix of actors in Palermo and distance matrix with threshold 0.2

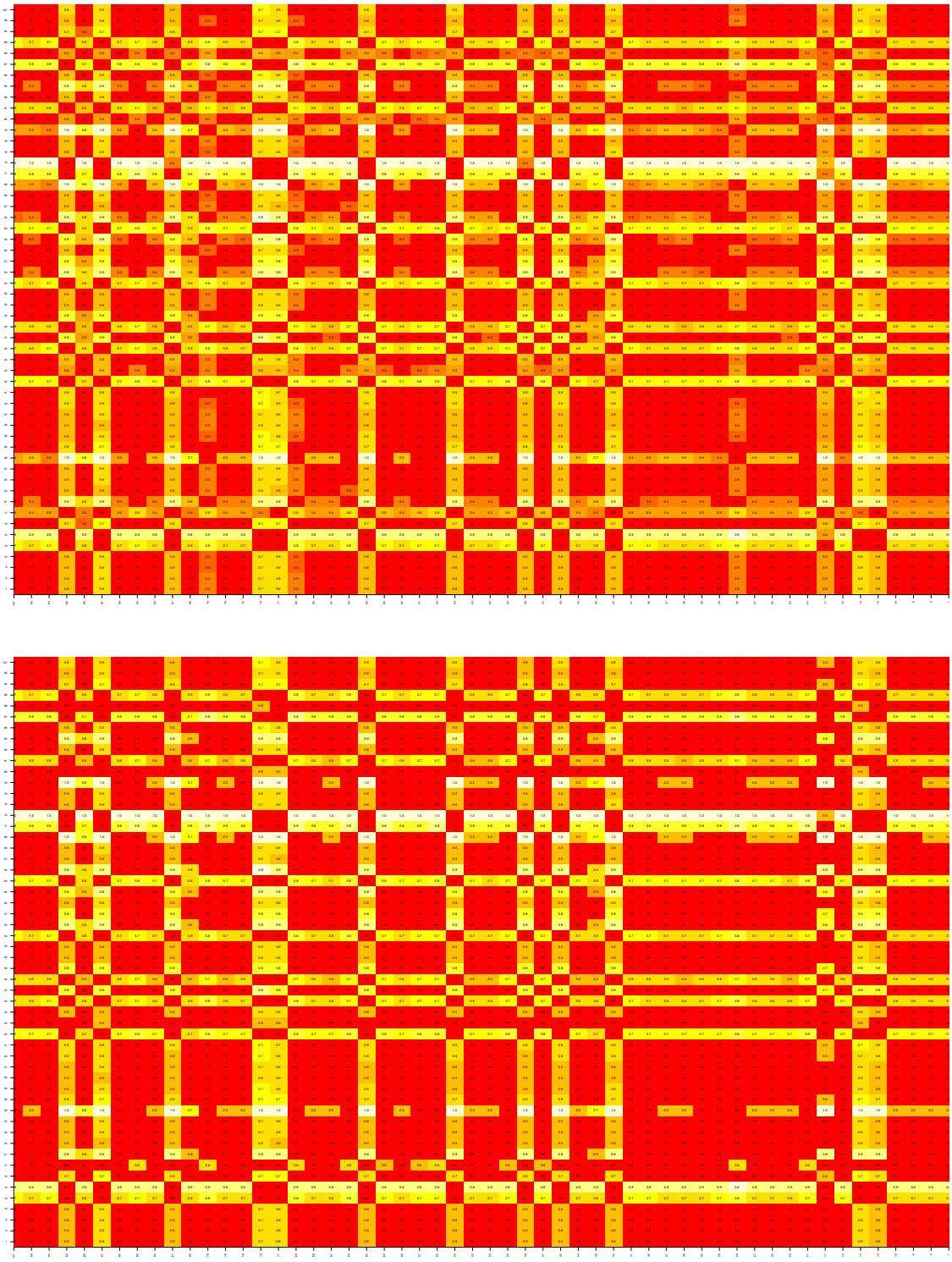


FIGURE 4.10: Distance Matrix of actors in Palermo with threshold 0.3 and 0.5

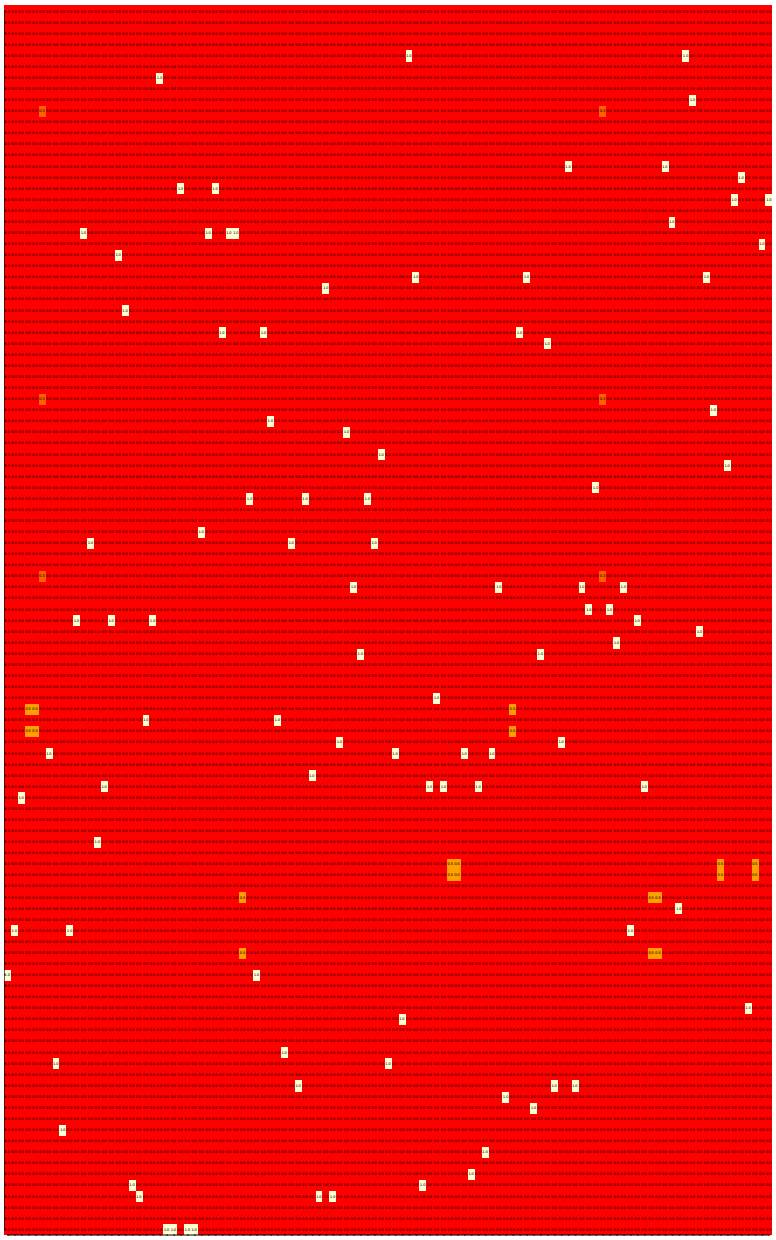


FIGURE 4.11: Distance Matrix of actors in the network with the five closest neighbors.

procurements. In the formula it was also considered a vector of regressor coefficients that quantify the increasing of one unit of age.

TABLE 4.1: SAR regressions

	direction	*linking2*	*Public Procurement*
	(1)	(2)	(3)
Intercept	-2.158^{****}	-1.603^{***}	-3.635^{****}
	(0.598)	(0.546)	(1.044)
Age	0.027^{*}	0.018^{*}	0.025
	(0.011)	(0.01)	(0.017)
Rho	0.427^{**}	0.226	-0.042
	(0.152)	(0.287)	(0.367)
N	111	111	111

Notes:
****Significant at less than 1 percent level.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Results of those regressions show that there is a "spatial effect" visible when such a dimension is considered. Once table 4.1 is compared with results shown in table A.1it is possible to see that there is still a higher probability that increasing of one unit the age of an affiliated member, with his closest four neighbors, there is a higher chance that the member will have a directive role, have links outside the mandamento and influence public procurements. However, this probability is slightly lower once the space that surrounds other members is not considered with fixed-effects anymore.

This is also visible once nodes' degree is considered, but it shows also that nodes with the highest degree have a lower probability of having a directive role, links outside the mandamento and influence public procurements. As shown in table 4.2, increasing of one unit the degree of one node, it decreases the probability of the dependent variable.

This results in a possibly more accurate outcome than general probit models which is also accounting for the role of space on the topological characteristics of the network (Barthelemy, 2011). In fact, the way members and neighbors are clustered is the result of how the network is actually shaped. Members of the same mafia families live in the same area or neighborhood, members operating in the same mandamento live in there. Commuting for

TABLE 4.2: SAR regressions on the degree of the network

	direction	*linking2*	*Public Procurement*
	(1)	(2)	(3)
Intercept	-0.541^{**}	-0.601^{**}	-2.03^{***}
	(0.01)	(0.046)	(0.001)
Nodes' degree	-3.146^{**}	-0.608^{*}	-3.422^{**}
	(0.022)	(0.051)	(0.013)
Rho	0.183	0.072	-0.059
	(0.146)	(0.782)	(0.779)
N	111	111	111

Notes:
****Significant at less than 1 percent level.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

work is almost non-existent as shown in 3 since members are embedded in a limited geographical space. Therefore it is possible to reason that closer members are influencing each other. Looking at the plots of small villages around Palermo 3.1 and 4.8 it is possible to notice such characterisation.

This framework aims at giving a better theoretical understanding of structures, dynamics and possibly future developments of mafia-type networks. Those features are necessary to a greater extent for adopting effective preventative and response measures addressed against organised crime.

This research proposed a judicial evidence-based development for explaining how a criminal network works and how it is different from other examples found in literature like terrorism and other criminal networks.

Appendix A

Tables

TABLE A.1: Probit regressions 1

	direction	linking	pubProc
	(1)	(2)	(3)
Intercept	43.994	31.792	71.421
	(2,530.692)	(1,568.678)	(4,087.664)
Birth Year	-0.025*	-0.019*	-0.040*
	(0.013)	(0.012)	(0.023)
N	111	111	111
Notes:	*Sign. at 10%.		

TABLE A.2: Probit regressions 2

	d_pubProc.x	direction.x	d_linking2.x
	probit	*probit*	*probit*
	(1)	(2)	(3)
Intercept	-2.243^{**}	-3.116^{***}	-1.988^{**}
	(1.063)	(1.027)	(0.817)
Age	0.010	0.028^{*}	0.028^{**}
	(0.017)	(0.016)	(0.014)
# Household Buildings	0.006	0.010^{**}	0.001
	(0.005)	(0.004)	(0.004)
N	53	52	53
Log Likelihood	−15.647	−20.121	−31.732
Akaike Inf. Crit.	37.294	46.242	69.463

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

TABLE A.3: Probit regressions 3

	d_pubProc.x	direction.x	d_linking2.x
	probit	*probit*	*probit*
	(1)	(2)	(3)
Intercept	−1.281	-1.796^{*}	-1.486^{*}
	(1.006)	(0.962)	(0.788)
Age	0.006	0.027	0.027^{*}
	(0.018)	(0.018)	(0.014)
# Foreigners	−0.250	-0.573^{**}	-0.186^{*}
	(0.216)	(0.285)	(0.103)
N	53	52	53
Log Likelihood	−15.243	−18.948	−29.122
Akaike Inf. Crit.	36.486	43.895	64.243

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

TABLE A.4: Probit regressions 4

	d_pubProc.x	direction.x	d_linking2.x	d_pubProc.x
	probit	*probit*	*probit*	*probit*
	(1)	(2)	(3)	(4)
Intercept	−1.304	−1.830**	−1.640**	−1.246
	(1.036)	(0.898)	(0.778)	(1.009)
Age	0.010	0.023	0.030**	0.009
	(0.019)	(0.016)	(0.014)	(0.018)
# Divorced Males	−0.490*	−0.256	−0.240*	
	(0.290)	(0.174)	(0.143)	
# Divorced				−0.182*
				(0.110)
N	53	52	53	53
Log Likelihood	−14.394	−21.851	−30.191	−14.748
Akaike Inf. Crit.	34.788	49.702	66.382	35.496

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

TABLE A.5: OLS

	degree_actors			
	OLS			
	(1)	(2)	(3)	(4)
Intercept	0.249	0.237	0.235	0.218
	(0.185)	(0.187)	(0.188)	(0.188)
Age	0.003	0.003	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
Direction	−0.420***	−0.399***	−0.421**	−0.433**
	(0.136)	(0.140)	(0.174)	(0.174)
PubProc		−0.176	−0.156	−0.101
		(0.282)	(0.297)	(0.300)
Linking2			0.031	0.079
			(0.141)	(0.146)
Meetings				0.140
				(0.120)
N	98	98	98	98
R^2	0.091	0.095	0.095	0.108
Adjusted R^2	0.072	0.066	0.056	0.060
Residual Std. Error	0.458 (df = 95)	0.459 (df = 94)	0.461 (df = 93)	0.461 (df = 92)
F Statistic	4.747** (df = 2; 95)	3.275** (df = 3; 94)	2.444* (df = 4; 93)	2.236* (df = 5; 92)

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

TABLE A.6: Carto queries

```
SELECT a2.cartodb_id, r.idlink1, r.idlink2,
ST_Makeline(a2.the_geom_webmercator, a1.the_geom_webmercator)
as the_geom_webmercator
FROM coords_dat a1
JOIN my_links r ON r.idlink1 = a1.field_1
JOIN coords_dat a2 ON r.idlink2 = a2.field_1
```

TABLE A.7: Probit regressions: dependent variable "mafioso"

	mafioso			
	probit			
	(1)	(2)	(3)	(4)
Intercept	0.916***	0.810***	0.898***	0.586***
	(0.155)	(0.173)	(0.181)	(0.204)
Meetings	0.774	0.881*	0.793	1.105**
	(0.490)	(0.496)	(0.499)	(0.507)
Linking		0.525	0.437	0.750*
		(0.413)	(0.416)	(0.426)
Host			−6.109	−6.336
			(266.405)	(1,109.106)
Linking2				5.164
				(334.408)
N	111	111	111	111
Log Likelihood	−45.992	−45.112	−41.861	−36.229
Akaike Inf. Crit.	95.983	96.224	91.722	82.458

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Bibliography

.

Abbate, Lirio and Peter Gomez (2013). *I complici: Tutti gli uomini di Bernando Provenzano da Corleone al Parlamento*. Fazi Editore.

Acemoglu, Daron, Simon Johnson, and James Robinson (2003). "Understanding prosperity and poverty: Geography, institutions and the reversal of fortune". In: *Understanding poverty*, pp. 19–36.

Acemoglu, Daron and James A Robinson (2005). *Economic origins of dictatorship and democracy*. Cambridge University Press.

— (2006). "De facto political power and institutional persistence". In: *The American economic review*, pp. 325–330.

Albanese, Jay S (2008). "Risk Assessment in Organized Crime Developing a Market and Product-Based Model to Determine Threat Levels". In: *Journal of Contemporary Criminal Justice* 24.3, pp. 263–273.

Albert, Réka and Albert-László Barabási (2002). "Statistical mechanics of complex networks". In: *Reviews of modern physics* 74.1, p. 47.

Alizadeh, Meysam, Claudio Cioffi-Revilla, and Andrew Crooks (2016). "Generating and analyzing spatial social networks". In: *Computational and Mathematical Organization Theory*, pp. 1–29.

Alosi, Salvatore (1983). *Banca e latifondo nella Sicilia degli anni Trenta*. Vol. 100. Guida Editori.

Andresen, Martin A (2006). "A spatial analysis of crime in Vancouver, British Columbia: a synthesis of social disorganization and routine activity theory". In: *The Canadian Geographer/Le Géographe canadien* 50.4, pp. 487–502.

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

Anselin, Luc et al. (2000). "Spatial analyses of crime". In: *Criminal justice* 4.2, pp. 213–262.

Aron, Raymond (1950). "Social structure and the ruling class: part 1". In: *The British Journal of Sociology* 1.1, pp. 1–16.

Aspinall, Edward and Marcus Mietzner (2010). *Problems of democratisation in Indonesia: Elections, institutions and society*. Institute of Southeast Asian Studies.

Baccara, Mariagiovanna and Heski Bar-Isaac (2008). "How to organize crime". In: *The Review of Economic Studies* 75.4, pp. 1039–1067.

Bailey, Trevor C and Anthony C Gatrell (1995). *Interactive spatial data analysis*. Vol. 413. Longman Scientific & Technical Essex.

Bain, Joe Staten (1956). *Barriers to new competition, their character and consequences in manufacturing industries.* Tech. rep.

Balletta, Luigi and AM Lavezzi (2014). "Extortion, Firm's Size and the Sectoral Allocation of Capital"". In: *Manuscript. Universita di Palermo, Italy*.

Bandiera, Oriana (2003). "Land reform, the market for protection, and the origins of the Sicilian mafia: theory and evidence". In: *Journal of Law, Economics, and Organization* 19.1, pp. 218–244.

Barabási, Albert-László, Erzsébet Ravasz, and Zoltán Oltvai (2003). "Hierarchical organization of modularity in complex networks". In: *Statistical mechanics of complex networks*. Springer, pp. 46–65.

Barrat, Alain, Marc Barthélemy, and Alessandro Vespignani (2005). "The effects of spatial constraints on the evolution of weighted complex networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2005.05, P05003.

Barthélemy, Marc (2011). "Spatial networks". In: *Physics Reports* 499.1, pp. 1–101.

Barzel, Baruch and Albert-László Barabási (2013). "Network link prediction by global silencing of indirect correlations". In: *Nature biotechnology* 31.8, pp. 720–725.

Berlusconi, Giulia (2013). "Do all the pieces matter? Assessing the reliability of law enforcement data sources for the network analysis of wire taps". In: *Global Crime* 14.1, pp. 61–81.

Berlusconi, Giulia et al. (2016). "Link Prediction in Criminal Networks: A Tool for Criminal Intelligence Analysis". In: *PloS one* 11.4, e0154244.

Blakely, G Robert and Ronald Goldstock (1979). "On the Waterfront: RICO and Labor Racketeering". In: *Am. Crim. L. Rev.* 17, p. 341.

Block, Alan A and William J Chambliss (1981). *Organizing crime*. Elsevier New York.

Boccaletti, Stefano et al. (2006). "Complex networks: Structure and dynamics". In: *Physics reports* 424.4, pp. 175–308.

Boldi, Paolo and Sebastiano Vigna (2014). "Axioms for centrality". In: *Internet Mathematics* 10.3-4, pp. 222–262.

Brantingham, Paul and Patricia Brantingham (2008). "5. Crime pattern theory". In: *Environmental criminology and crime analysis*, p. 78.

Bueno De Mesquita, Ethan and Catherine Hafer (2008). "Public Protection or Private Extortion?" In: *Economics & politics* 20.1, pp. 1–32.

Buonanno, Paolo and Matteo Pazzona (2014). "Migrating mafias". In: *Regional Science and Urban Economics* 44, pp. 75–81.

Buonanno, Paolo et al. (2015). "Poor institutions, rich mines: Resource curse in the origins of the Sicilian mafia". In: *The Economic Journal* 125.586, F175–F202.

Calderoni, Francesco (2012). "The structure of drug trafficking mafias: the 'Ndrangheta and cocaine". In: *Crime, Law and Social Change* 58.3, pp. 321–349.

Campana, Paolo (2016). "Explaining criminal networks: Strategies and potential pitfalls". In: *Methodological Innovations* 9, p. 2059799115622748.

Campana, Paolo and Federico Varese (2012). "Listening to the wire: criteria and techniques for the quantitative analysis of phone intercepts". In: *Trends in organized crime* 15.1, pp. 13–30.

— (2013). "Cooperation in criminal organizations: Kinship and violence as credible commitments". In: *Rationality and society* 25.3, pp. 263–289.

Carrington, Peter J, John Scott, and Stanley Wasserman (2005). *Models and methods in social network analysis*. Vol. 28. Cambridge university press.

Centola, Damon (2010). "The spread of behavior in an online social network experiment". In: *science* 329.5996, pp. 1194–1197.

Chainey, Spencer, Lisa Tompson, and Sebastian Uhlig (2008). "The utility of hotspot mapping for predicting spatial patterns of crime". In: *Security Journal* 21.1-2, pp. 4–28.

Csardi, Gabor and Tamas Nepusz (2006). "The igraph software package for complex network research". In: *InterJournal, Complex Systems* 1695.5, pp. 1–9.

Dal Bó, Ernesto, Pedro Dal Bó, and Rafael Di Tella (2006). ""Plata o Plomo?": Bribe and Punishment in a Theory of Political Influence". In: *American Political Science Review* 100.01, pp. 41–53.

Davidson, Ron and David Harel (1996). "Drawing graphs nicely using simulated annealing". In: *ACM Transactions on Graphics (TOG)* 15.4, pp. 301–331.

Davis, James A and Samuel Leinhardt (1967). "The structure of positive interpersonal relations in small groups." In:

Deguchi, Tsuyoshi et al. (2014). "Hubs and authorities in the world trade network using a Weighted HITS algorithm". In: *PloS one* 9.7, e100338.

Diamond, Jared M and Doug Ordunio (2005). *Guns, germs, and steel*. National Geographic.

Dickie, John (2005). *Cosa Nostra: a history of the Sicilian Mafia*. Macmillan.

Dimico, Arcangelo, Alessia Isopi, and Ola Olsson (2017). "Origins of the sicilian mafia: The market for lemons". In:

Durkheim, Emile (1897). *Le suicide: étude de sociologie*. F. Alcan.

Fafchamps, Marcel and Flore Gubert (2007). "Risk sharing and network formation". In: *The American economic review* 97.2, pp. 75–79.

Federico, Varese and Federico Varese (2010). "Organized crime". In: *Organized Crime: Critical Concepts in Criminology*, pp. 1–33.

Fiandaca, Giovanni (2007). *Women and the mafia: Female roles in organized crime structures*. Vol. 5. Springer Science & Business Media.

Franchetti, Leopoldo and Sidney Sonnino (1877). *La Sicilia nel 1876*. Vol. 1. G. Barbèra.

Freeman, Linton C (1977). "A set of measures of centrality based on betweenness". In: *Sociometry*, pp. 35–41.

Fruchterman, Thomas MJ and Edward M Reingold (1991). "Graph drawing by force-directed placement". In: *Software: Practice and experience* 21.11, pp. 1129–1164.

Gambetta, Diego (1996). *The Sicilian Mafia: the business of private protection*. Harvard University Press.

Gambetta, Diego and Peter Reuter (1995). "Conspiracy among the many: the mafia in legitimate industries". In: *The Economic Dimensions of Crime*. Springer, pp. 99–120.

Geys, Benny and Gianmarco Daniele (2014). "Organized Crime, Institutions and political quality: empirical evidence from Italian municipalities". In: *Economic Journal, forthcoming*.

Granovetter, Mark S (1973). "The strength of weak ties". In: *American journal of sociology*, pp. 1360–1380.

Haining, Robert P (2003). *Spatial data analysis: theory and practice*. Cambridge University Press.

Harvey, Jackie et al. (2009). "Not only banks: Criminological models on the infiltration of public contracts by Italian organized crime". In: *Journal of Money Laundering Control* 12.2, pp. 151–172.

Hassan, Md Kamrul, Liana Islam, and Syed Arefinul Haque (2017). "Degree distribution, rank-size distribution, and leadership persistence in mediation-driven attachment networks". In: *Physica A: Statistical Mechanics and its Applications* 469, pp. 23–30.

Hernández, Anabel (2012). *Los señores del narco*. Grijalbo.

Jackson, Matthew O et al. (2008). *Social and economic networks*. Vol. 3. Princeton university press Princeton.

Kamada, Tomihisa and Satoru Kawai (1989). "An algorithm for drawing general undirected graphs". In: *Information processing letters* 31.1, pp. 7–15.

Kearns, Michael, Siddharth Suri, and Nick Montfort (2006). "An experimental study of the coloring problem on human subject networks". In: *Science* 313.5788, pp. 824–827.

Keegan, Brian et al. (2011). "Sic transit gloria mundi virtuali". In: *Promise and peril at the intersection of computational social science and online clandestine organizations. ACM Web Science*, pp. 14–17.

Kingston, Paul W (2000). *The classless society*. Stanford University Press.

Kleinberg, Jon M (1999). "Authoritative sources in a hyperlinked environment". In: *Journal of the ACM (JACM)* 46.5, pp. 604–632.

Klemm, Konstantin and Victor M Eguiluz (2002). "Highly clustered scale-free networks". In: *Physical Review E* 65.3, p. 036123.

Konrad, Kai A and Stergios Skaperdas (2012). "The Market for Protection and the Origin of the State". In: *Economic Theory* 50.2, pp. 417–443.

Kossinets, Gueorgi (2006). "Effects of missing data in social networks". In: *Social networks* 28.3, pp. 247–268.

Kossinets, Gueorgi and Duncan J Watts (2006). "Empirical analysis of an evolving social network". In: *science* 311.5757, pp. 88–90.

Kotler, Philip et al. (2014). *Marketing management 14/e*. Pearson.

Krebs, Valdis (2001). "Uncloaking terrorist networks". In: *First Monday* 7.4.

Krebs, Valdis E (2002). "Mapping networks of terrorist cells". In: *Connections* 24.3, pp. 43–52.

Kumar, Vimal, Stergios Skaperdas, and Nuno Garoupa (2008). "On the economics of organized crime". In: *prepared for inclusion in Gauropa, Nuno (ed), Criminal Law and Economics*.

La Spina, Antonio (2009). "La sociologia del fenomeno mafioso dopo il 2006". In: *Rassegna italiana di sociologia* 50.2, pp. 301–308.

Lagziel, David and Ehud Lehrer (2015). "Reward Schemes". In:

Lampe, Klaus von and Per Ole Johansen (2003). "Criminal networks and trust". In: *3rd annual meeting of the European Society of Criminology, Helsinki*. Vol. 29.

Laumann, Edward O, Peter V Marsden, and David Prensky (1989). "The boundary specification problem in network analysis". In: *Research methods in social network analysis* 61, p. 87.

Lavezzi, Andrea Mario (2008). "Economic structure and vulnerability to organised crime: Evidence from Sicily". In: *Global Crime* 9.3, pp. 198–220.

— (2014). "Organised crime and the economy: A framework for policy prescriptions". In: *Global Crime* 15.1-2, pp. 164–190.

LeSage, James P and R Kelley Pace (2008). "SPATIAL ECONOMETRIC MODELING OF ORIGIN-DESTINATION FLOWS". In: *Journal of Regional Science* 48.5, pp. 941–967.

Levine, Ned (2006). "Crime mapping and the Crimestat program". In: *Geographical analysis* 38.1, pp. 41–56.

Lodato, Saverio and Marco Travaglio (2005). *Intoccabili*. Bureau Biblioteca Univ. Rizzoli.

Marx, Karl (1867). *Capital, volume I*.

Masías, Víctor Hugo et al. (2016). "Modeling Verdict Outcomes Using Social Network Measures: The Watergate and Caviar Network Cases". In: *PloS one* 11.1, e0147248.

Mastrobuoni, Giovanni (2015). "The Value of Connections: Evidence from the Italian-American Mafia". In: *The Economic Journal* 125.586, F256–F288.

Mastrobuoni, Giovanni, Eleonora Patacchini, et al. (2012). "Organized crime networks: An application of network analysis techniques to the American mafia". In: *Review of Network Economics* 11.3, pp. 1–43.

McAfee, R Preston, Hugo M Mialon, and Michael A Williams (2003). "What is a Barrier to Entry?" In: *Available at SSRN 594601*.

McCloskey, Deirdre N (2010). *Bourgeois dignity: Why economics can't explain the modern world*. University of Chicago Press.

McIllwain, Jeffrey Scott (1999). "Organized crime: A social network approach". In: *Crime, Law and Social Change* 32.4, pp. 301–323.

Meyers, Lauren Ancel et al. (2005). "Network theory and SARS: predicting outbreak diversity". In: *Journal of theoretical biology* 232.1, pp. 71–81.

Mori, C. *Con la mafia ai ferri corti, Milano, 1932; A*.

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

Morselli, Carlo, Cynthia Giguère, and Katia Petit (2007). "The efficiency/security trade-off in criminal networks". In: *Social Networks* 29.1, pp. 143–153.

Morselli, Carlo and Katia Petit (2007). "Law-enforcement disruption of a drug importation network". In: *Global Crime* 8.2, pp. 109–130.

Mosca, Gaetano (1896). *Elementi di scienza politica*. Fratelli Bocca.

Natarajan, Mangai (2000). "Understanding the structure of a drug trafficking organization: a conversational analysis". In: *Crime Prevention Studies* 11, pp. 273–298.

Neumayer, Eric and Thomas Plümper (2010). "Spatial effects in dyadic data". In: *International Organization* 64.01, pp. 145–166.

Newman, Mark EJ and Michelle Girvan (2004). "Finding and evaluating community structure in networks". In: *Physical review E* 69.2, p. 026113.

North, Douglass C (1990). *Institutions, institutional change and economic performance*. Cambridge university press.

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

Olson, Mancur (2009). *The logic of collective action*. Vol. 124. Harvard University Press.

Paoli, Letizia (2001). "Crime, italian style". In: *Daedalus* 130.3, pp. 157–185.

— (2002). "The paradoxes of organized crime". In: *Crime, law and social change* 37.1, pp. 51–97.

— (2004). "Italian organised crime: Mafia associations and criminal enterprises". In: *Global Crime* 6.1, pp. 19–31.

Papadias, Dimitris et al. (2003). "Query processing in spatial network databases". In: *Proceedings of the 29th international conference on Very large data bases-Volume 29*. VLDB Endowment, pp. 802–813.

Pareto, Vilfredo (1916). *Trattato di sociologia generale...* Vol. 2. G. Barbèra.

Patacchini, Eleonora and Yves Zenou (2008). "The strength of weak ties in crime". In: *European Economic Review* 52.2, pp. 209–236.

Petróczi, Andrea, Tamás Nepusz, and Fülöp Bazsó (2007). "Measuring tie-strength in virtual social networks". In: *Connections* 27.2, pp. 39–52.

Pinotti, Paolo (2011). "The Economic Consequences of Organized Crime: Evidence from Southern Italy". In: *Bank of Italy*.

Polo, Michele (1995). "Internal cohesion and competition among criminal organizations". In: *The economics of organised crime*, pp. 87–108.

Porter, John (1957). "The economic elite and the social structure in Canada". In: *Canadian Journal of Economics and Political Science/Revue canadienne de economiques et science politique* 23.03, pp. 376–394.

Powell, Walter (2003). "Neither market nor hierarchy". In: *The sociology of organizations: classic, contemporary, and critical readings* 315, pp. 104–117.

Restrepo, Juan G, Edward Ott, and Brian R Hunt (2006). "Characterizing the dynamical importance of network nodes and links". In: *Physical review letters* 97.9, p. 094102.

Reuter, Peter and Jonathan B Rubinstein (1978). "Fact, fancy, and organized crime". In: *The Public Interest* 53, p. 45.

Romer, David (2003). "Misconceptions and political outcomes*". In: *The Economic Journal* 113.484, pp. 1–20.

Saviano, Roberto (2006). *Gomorra, Mondadori*.

Scaglione, Attilio (2011). *Reti mafiose. Cosa Nostra e Camorra: organizzazioni criminali a confronto: Cosa Nostra e Camorra: organizzazioni criminali a confronto*. FrancoAngeli.

Scherer, FM and David Ross (1990). *Economic Performance*. Boston: Houghton-Mifflin.

Schrecker, John E (2004). *The Chinese revolution in historical perspective*. Greenwood Publishing Group.

Scott, John (1988). "Social network analysis". In: *Sociology* 22.1, pp. 109–127.

Shaw, Clifford R and Henry D McKay (1969). *Juvenile delinquency and urban areas. Revised edition*.

Skaperdas, Stergios and Constantinos Syropoulos (1997). "3 Gangs as primitive states". In: *The Economics of Organised Crime*, p. 61.

Snijders, Tom AB (2011). *Multilevel analysis*. Springer.

Sparrow, Malcolm K (1991). "The application of network analysis to criminal intelligence: An assessment of the prospects". In: *Social networks* 13.3, pp. 251–274.

Stigler, George J (1968). "Price and non-price competition". In: *The Journal of Political Economy*, pp. 149–154.

Strom, Kevin et al. (2010). "The private security industry: A review of the definitions, available data sources, and paths moving forward". In: *Washington DC: US Department of Justice*.

Tonini, Paolo (2012). *Manuale di procedura penale*. Giuffrè Editore.

Van Dijk, Jan (2007). "Mafia markers: assessing organized crime and its impact upon societies". In: *Trends in organized crime* 10.4, pp. 39–56.

Varese, Federico (2001). *The Russian Mafia: private protection in a new market economy*. OUP Oxford.

— (2006a). "How Mafias migrate: the case of theNdrangheta in Northern Italy". In: *Law & Society Review* 40.2, pp. 411–444.

— (2006b). "The structure of a criminal network examined: the Russian Mafia in Rome". In:

— (2011). *Mafias on the move: How organized crime conquers new territories*. Princeton University Press.

— (2014). "Protection and extortion". In: *Oxford Handbook of Organized Crime*, pp. 343–58.

Voltaire (1766). *Commentaire sur le livre Des délits et des peines*.

Von Lampe, Klaus and Per Ole Johansen (2004). "Organized Crime and Trust:: On the conceptualization and empirical relevance of trust in the context of criminal networks". In: *Global Crime* 6.2, pp. 159–184.

Wainwright, Tom (2016). *Narconomics: How to run a drug cartel*. PublicAffairs.

Wasserman, Stanley and Katherine Faust (1994). *Social network analysis: Methods and applications*. Vol. 8. Cambridge university press.

Wickelmaier, Florian (2003). "An introduction to MDS". In: *Sound Quality Research Unit, Aalborg University, Denmark* 46.

Williams, Phil (2001). "Transnational Criminal Networks". In: *Networks and netwars: the future of terror, crime, and militancy* 1382, p. 61.

Xu, Jennifer and Hsinchun Chen (2008). "The topology of dark networks". In: *Communications of the ACM* 51.10, pp. 58–65.

Yip, Michael, Nigel Shadbolt, and Craig Webber (2012). "Structural analysis of online criminal social networks". In: *Intelligence and Security Informatics (ISI), 2012 IEEE International Conference on*. IEEE, pp. 60–65.