Manuscript accepted for publication in *Global Crime* 14(1), pp. 61-81, doi: https://doi.org/10.1080/17440572.2012.746940

Do all the pieces matter? Assessing the reliability of law enforcement data sources for the network analysis of wire taps

Giulia Berlusconi

Abstract Law enforcement agencies rely on data collected from wire taps to construct the organizational chart of criminal enterprises. Recently, a number of academics have also begun to utilise social network analysis to describe relations among criminals and understand the internal organisation of criminal groups. However, before drawing conclusions about the structure or the organisation of criminal groups, it is important to understand the limitations that selective samples such as wire taps may have on network analysis measures. Electronic surveillance data can be found in different kinds of court records and the selection of the data source is likely to influence the amount of missing information and, consequently, the results. This article discusses the impact that the selection of a specific data source for the social network analysis of criminal groups may have on centrality measures usually adopted in organised crime research to identify key players.

Keywords criminal networks; law enforcement data; social network analysis; sampling; centrality.

1. Introduction

In the last decades, the network perspective has become more and more influential in organised crime research, and social network analysis (SNA) has become popular among scholars interested in describing informal relations among criminals and understanding the internal organisation of criminal groups.[1] One of the common applications of SNA in a criminological context consists in the representation of criminal groups and their activities, mainly in a descriptive way.[2] These studies generally include only persons who are involved in criminal activities and are part of a specific criminal group; however, differences may exist in the way network boundaries are identified by the researcher. Another typical feature of these studies is the reliance on law enforcement data – especially, electronic surveillance data – as their primary source of information on the relations among criminal groups members.

Electronic surveillance data have been increasingly used by scholars interested in describing the structure of relations within criminal groups by capturing the conversations among their members, as they occur every time there is the need to communicate to organise illegal activities, exchange

[1] Sparrow, "Application of Network Analysis," 251–274; McIllwain, "Organized Crime," 301–323; McAndrew, "Structural Analysis," 51–94; Natarajan, "Drug Trafficking Organization," 273–298; Klerks, "Network Paradigm," 53–65; Bruinsma and Bernasco, "Transnational Illegal Markets," 79–94; Morselli, *Contacts, Opportunities and Criminal Enterprise*; Morselli, *Inside Criminal Networks*; Natarajan, "Heroin Distribution Network," 171–192; Varese, "Criminal Connections"; von Lampe, "Human Capital," 93–100; and van der Hulst, "Introduction to Social Network Analysis," 101–121.

[2] Carrington, "Social Network," 244.

information, or arrange meetings.[3] However, little attention has been paid to issues related to the accuracy, validity, and reliability of these data, with a notable exception, namely a recent article by Federico Varese and Paolo Campana addressing some of these issues and suggesting the use of a combination of different data analysis techniques.[4] This article deals with the problem of accuracy and reliability of electronic surveillance data available in different court records and how the source of information might affect the SNA of criminal groups. Two separate issues will be addressed: group coverage, namely the extent to which different court documents provide an accurate description of the overall network, and the identification of key players.

The next section introduces the problem of missing data in criminal network research, presents some of the possible causes of missing information beyond and within the final representation of criminal networks, and reviews and discusses the different sources of law enforcement data. After a description of the data and the methodology adopted for the study in Section 3, the fourth section presents the main results and Sections 5 and 6 discuss the limitation of this study and the main implication of the results for criminal network analysis, respectively. The last section concludes.

2. Missing data issues in criminal network research

The usefulness of SNA for the study of criminal groups has been largely discussed. Among the numerous benefits, this methodology allows to conduct a detailed analysis of the internal structure of criminal groups, capturing the informal relations among their members and avoiding any assumption about the organisational structure of the network prior to its actual analysis.[5]

Several studies have also highlighted a number of limitations and methodological problems connected to these applications of SNA.[6] Some of these limitations are well known by network analysts, such as the problem of capturing network dynamics; some others, although being shared with other fields of research, are magnified by the specific application of network analysis concepts and tools to covert networks. Among them, the problem of missing information – and thus network completeness – is particularly relevant in analysing criminal networks, since it affects the scope and structure of the network under consideration.[7] Missing nodes or edges can indeed create a sort of domino effect and heavily affect the results, creating problems of inference also in the case of descriptive purposes.[8]

The issue of missing information beyond the final representation of the network is usually associated with the specification of network boundaries, which refers to the definition of rules for including specific

[3] Natarajan, "Drug Trafficking Organization"; Natarajan, "Heroin Distribution Network"; Varese, "Criminal Connections"; Morselli, *Inside Criminal Networks*; Morselli, "Vulnerable and Strategic Positions," 382–392; Calderoni, "The 'Ndrangheta Through the Lens"; and Campana, "Eavesdropping on the Mob," 213–228.

[4] Campana and Varese, "Listening to the Wire," 13–30.

[5] Morselli, *Inside Criminal Networks*, 41; Ianni and Reuss-Ianni, "Network Analysis," 81–82; Sparrow, "Application of Network Analysis," 272; McIllwain, "Organized Crime," 319; von Lampe, "Human Capital," 95; and van der Hulst, "Introduction to Social Network Analysis," 104.

[6] Sparrow, "Application of Network Analysis," 262; McAndrew, "Structural Analysis," 62; Spapens, "Macro Networks," 193; von Lampe, "Human Capital," 95; van der Hulst, "Introduction to Social Network Analysis," 110; Morselli, *Inside Criminal Networks*, 47; and Malm, Bichler, and Van De Walle, "The Ties That Bind," 70.

[7] von Lampe, "Human Capital," 95; van der Hulst, "Introduction to Social Network Analysis," 110; Morselli, *Inside Criminal Networks*, 47; and Malm, Bichler, and Van De Walle, "The Ties That Bind," 70.

[8] Ianni and Reuss-Ianni, "Network Analysis," 70; van der Hulst, "Introduction to Social Network Analysis," 112; and Sparrow, "Application of Network Analysis," 262.

actors and relations in the network under investigation.[9] In organised crime research, the boundaries of the overall network, as identified by law enforcement agencies, do not necessarily coincide with those of the criminal group, and the researcher may identify internal boundaries on the basis of either theoretical or practical considerations.[10] A possible solution might be to include actors in the final representation of the network on the basis of their legal position across criminal justice stages, namely, whether the actors have been monitored, targeted, arrested, accused, or sentenced. Depending on the criminal justice stage data refer to, their scope, as well as their precision, varies. Indeed, the number of network members decreases through the various stages within the criminal justice system, since only a small percentage of monitored individuals is accused and sentenced; whereas the precision of information increases, because actors are included in the network on the basis of judicial evidence and not just because of their participation in phone conversations with alleged criminals.[11]

Nonetheless, the boundary specification problem cannot be completely handled by the researcher, since the definition of the external boundaries of the network lies where the court files end and it is thus dependent upon available information, rather than a precise choice of the researcher.[12] Scholars usually rely on law enforcement data sources for the collection of information used for their analyses, but law enforcement agencies have a partial vision of the network under investigation, so that their data contain incomplete information.[13] Indeed, investigators have to rely on data-gathering methods that result in incomplete information, such as observations, archives, informants, or witnesses. Hence, information from prosecutorial transcripts, interrogation material, or criminal investigation files from completed cases, as well as interviews with offenders or law enforcement officers, presents limitations in terms of data accuracy and completeness.[14]

Even transcripts of wiretapped conversations include only a sample of all the conversations that occurred among the members of a criminal group, although the sample is not random but rather a purposive one, since the police usually transcribes all conversations that pertain to criminal activities in a broad sense, whereas conversations on personal or unrelated matters are discarded.[15]

Different court documents might also be available to the researcher, and the external boundaries of the network, as defined by law enforcement agencies, might vary across these documents. The electronic communication transcripts can be found in different kinds of court records, such as wiretap records, police reports, arrest warrants, and sentences.[16] In jurisdictions that allow to use wiretaps as evidence in trials, all the relevant conversations are fully transcribed in wiretap records and made available to the prosecutor and the judge. The arrest warrant usually includes only a selection of the electronic communication transcripts, as well as other relevant information on the suspects from informants and other investigative activities. A part of the electronic surveillance data is also usually reported in the

[9] Marsden, "Network Data and Measurement," 439; Laumann, Marsden, and Prensky, "Boundary Specification Problem," 62; and Kossinets, "Effects of Missing Data," 249.

[10] Campana and Varese, "Listening to the Wire," 21.

[11] Morselli, *Inside Criminal Networks*, 44.

[12] Campana and Varese, "Listening to the Wire," 20; van der Hulst, "Introduction to Social Network Analysis," 111; and Ianni and Reuss-Ianni, "Network Analysis," 71.

[13] Morselli, *Inside Criminal Networks*, 41; von Lampe, "Human Capital," 95; Malm and Bichler, "Networks of Collaborating Criminals," 20; and Campana and Varese, "Listening to the Wire," 17.

[14] Baker and Faulkner, "Social Organization of Conspiracy," 847; and Malm, Bichler, and Van De Walle, "The Ties That Bind," 55.

[15] Campana and Varese, "Listening to the Wire," 20.

[16] Ibid., 18.

sentence, along with other sources of evidence.[17] Hence, as shown in Figure 1, the closer we are to the judgment, the more likely we are to lose a relevant subset of electronic surveillance data.

The purposive sample of electronic surveillance data includes all telephone conversations which were wiretapped by the police and considered relevant for the criminal case. Conversations on matters unrelated to criminal activities are not included in the sample, although some of them might have been misinterpreted by law enforcement agencies and included in the wiretap records as well. At the same time, not all criminally related conversations might have been wiretapped and reported in the court documents. As we move from the wiretap records to the judgment, the amount of electronic surveillance data included in the judicial documents decreases. It is plausible to think that information on actors who have been arrested and accused is maintained, whereas the information on individuals who fell into the surveillance net but were not brought to court is discarded.

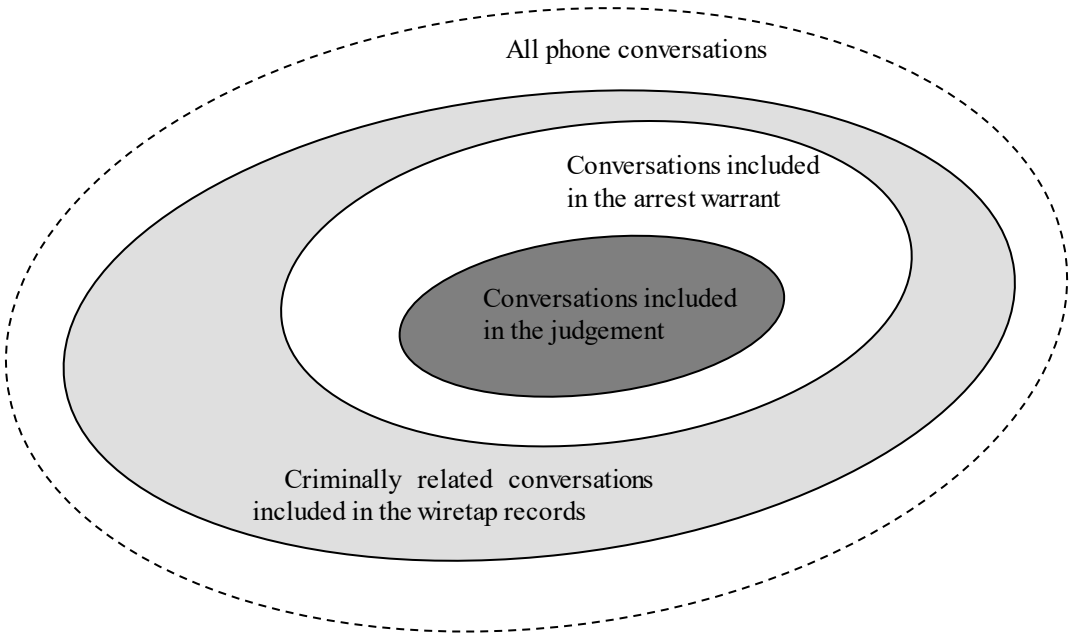

Figure 1. Samples of telephone conversations among criminal network members.

Source: Author's adaptation of Campana and Varese 2012.

However, the amount of missing data might be larger, since only a sample of conversations among arrested and accused individuals might be included in the arrest warrant and the judgment.

Although the conversations wiretapped by the police already constitute a sample of all the conversations that occurred among criminal network members during the investigation, the use of

[17] The United States, Canada and most European countries (e.g. Italy, Germany, the Netherlands) are among the jurisdictions in which wiretaps can be used as evidence. In Italy, for example, wiretapped conversations pertinent to the investigation are fully transcribed; these include only communications which are considered relevant by the public prosecutor and the defence attorneys, whereas conversations on personal or unrelated matters are discarded. The prosecutor's request for pre-trial detention and the arrest warrant usually include a selection of the electronic communication transcripts, as well as other relevant information (e.g., from informants and other investigative activities). A part of the electronic surveillance data is usually reported in the sentence, along with other sources of evidence.

sources of information other than wiretap records is likely to cause a further loss of electronic surveillance data and might lead to misleading results. For this reason, Campana and Varese suggest to avoid the judgment as the main source of information for statistical analysis, because the set of conversations is likely to be small and biased, and the results might not be reliable.[18] On the other hand, arrest warrants and sentences are more likely to be accessible to researchers aiming at studying criminal groups, and thus might constitute a useful source of data, provided they include sufficient information on relations among network members.[19]

Along with missing nodes, missing links or edges might also constitute a problem. Considering that criminals involved in organised crimes often use different telephones and lines to communicate with different individuals or categories of individuals, and that the name of the holder does not necessarily correspond with the actual user, the police might not be able to wiretap all the telephones and lines used to deal with criminal activities. If law enforcement agencies fail to intercept a specific telephone or line, this might result in missing links among specific actors. Also, after the end of the investigations, there might be evidence that a portion of electronic surveillance data gathered by law enforcement agencies was collected adopting an illegal/unauthorised procedure. This might result in the discard of those data from court documents and, as a consequence, in missing data within the final representation of the network.

Previous research in the field of network analysis mainly focused on the identification of key players and studied the effect of missing information on commonly used network measures, such as centrality.[20] The work by Bolland constitutes one of the first attempts to examine the robustness of centrality under conditions of random and systematic errors.[21] More recently, Costenbader and Valente assessed the stability of eleven centrality measures by randomly sampling 59 empirical networks at eight different sampling proportions.[22] This work has been extended by Borgatti and his colleagues, who examined the robustness of centrality measures under conditions of incomplete or imperfect data, including node and edge removal.[23] These studies proved that centrality measures are quite robust, especially under small amounts of error, and that their stability is influenced by the type of study and the characteristics of the network.[24] In all these cases, data were randomly sampled, i.e. nodes and links were randomly extracted from the networks, and social network measures were estimated and compared across different samples of the same network. However, 'the results could be quite different for practical settings in which the data collection methodology makes systematic errors (such as more readily losing nodes with low degree)'.[25]

A recent study by Xu and Chen analysed the effect of missing links and edges in dark networks.[26] Whereas different percentages of edges were randomly removed from the networks, confirming the validity of findings under small amounts of error, node removal was systematic. In order to test the

[18] Campana and Varese, "Listening to the Wire," 19.

[19] Bright, Caitlin, and Chalmers, "Illuminating Dark Networks," 152.

[20] Bolland, "Sorting Out Centrality", 233–253; Costenbader and Valente, "Stablity of Centrality Measures," 283–307; Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures." 124–136; and Xu and Chen, "Topology of Dark Networks," 58–65.

[21] Bolland, "Sorting Out Centrality."

[22] Costenbader and Valente, "Stablity of Centrality Measures."

[23] Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures."

[24] Costenbader and Valente, "Stability of Centrality Measures," 305; and Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures," 31.

[25] Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures," 132.

[26] Xu and Chen, "Topology of Dark Networks."

robustness of dark networks against attacks by law enforcement agencies, node removal involved either hubs, characterised by a high degree of direct contacts, or bridges, i.e. nodes positioned along the shortest path between two other network actors.[27] The finding that dark networks are more vulnerable to the removal of bridges rather than highly connected actors is relevant for intelligence purposes, but it does not provide insights into the effect that missing data due to the purposive sample of information by law enforcement agencies might have on SNA measures.

All these studies focused on different measures of the centrality of actors within a network. Centrality measures are particularly relevant in criminal network analysis, because they allow to identify critical nodes, namely those actors who are particularly involved in criminal activities and whose removal would maximise network disruption.[28] Degree and betweenness centrality are the two most common measures. The former estimates the number of actors every criminal network member is connected to; the latter measures the number of times every actor falls on the shortest path between two other members, allowing for the identification of those individuals who facilitate the exchange of information or resources in the network.[29]

Assuming that a researcher aiming at analysing a criminal group might want to collect information about all links among all individuals that fall within the surveillance net, and then eventually, decide to focus only on arrested or convicted ones; this article seeks to understand the influence that the reliance on a specific source of electronic surveillance data might have on the missing data problem and, as a consequence, on the results of the SNA of criminal groups. Considering the widespread use of centrality measures in criminal networks research, this article focuses on the two most common measures described above and assesses their robustness when moving from one kind of court record to another. The next section describes the data and the methodology adopted in this study.

3. Data and methods

When assessing the relationship between the focus of criminal network research – namely, one or more criminal groups – and the measures used for the analysis, the general aim of the research needs to be considered. If the researchers aim at providing a precise description of the criminal group, then the accuracy of available data is a fundamental prerequisite. On the other hand, if the focus is on differences between individuals in terms of their position within the criminal group, or on differences across networks, then the robustness of network measures to measurement errors needs to be addressed.[30] In this second case, complete information on each criminal network member is not necessary, as far as the sampling still provides reliable results.

The first part of the analysis aims at assessing whether and to what extent information on actors and their relations might differ across different judicial documents, and thus whether data from court documents other than wiretap records can guarantee the overall group coverage and provide an accurate description of organised crime networks. To achieve this aim, empirical data from an Italian criminal

[27] Ibid., 64.

[28] Morselli and Petit, "Law-Enforcement Disruption," 110; Schwartz and Rouselle, "Using Social Network Analysis," 189; and Williams, "Transnational Criminal Networks," 159.

[29] Wasserman and Faust, *Social Network Analysis*; Freeman, "Centrality in Social Networks," 223–258; and Bonacich, "Power and Centrality," 1170–1182.

[30] Marsden, "Network Data and Measurement."

investigation, for which different court records were available, are used, and networks based upon information from different judicial documents are compared.

Operation Oversize is a criminal investigation against a 'Ndrangheta group mainly operating in the Italian province of Lecco.[31] The investigation was conducted from 2000 to 2006 by the Police Offices of Lecco and Milan and the Investigation Group on Organised Crime (Gruppo d'Investigazione sulla Criminalità Organizzata, GICO) of the Guardia di Finanza of Milan and coordinated by the Antimafia Prosecutor's Office of Milan. The suspects were involved in homicides, robberies, and, above all, drug trafficking. The trial started in 2007 and lasted until 2009, when the judgment was passed.[32] The Antimafia District Directorate (Direzione Distrettuale Antimafia, DDA) of Milan provided access to several judicial documents related to this investigation. These documents include, among others, wiretap records containing the transcripts of 933 wiretapped conversations, the arrest warrant, and the trial judgment.[33]

The transcripts of intercepted communications contain the transcription of wiretappings and audio surveillance of monitored individuals. For each conversation, information on time (date and hour), as well as the full transcription, is reported. The warrant of arrest contains a selection of the transcripts, as well as other relevant information on both targeted and arrested individuals from informants and other investigative activities. Informants' statements are given prominence and several fragments of the interrogations are reported. The trial judgment summarises the evidence that came to light during the hearings and describes the illegal activities perpetrated by each convicted individual. The trial judgment includes information from several sources of evidence, including wiretappings and audio surveillance.

The three judicial documents are treated as separate sources of electronic surveillance data and used to create three different networks referring to the same criminal investigation. These networks are then analysed to assess whether and to what extent information on actors and their relations differ from one

---

[31] The 'Ndrangheta is a mafia-type organization mainly operating in the Calabria region (Italy), but with ramifications both in other Italian regions and abroad. Its affiliates are organized in *cosche* or '*ndrine*, whose members are usually bonded by kin ties and exercise their influence over a specific territory. A two-level hierarchical structure is present; low-level affiliates are referred to as members of the *società minore* (lower society), whereas bosses are part of the *società maggiore* (higher society). Within both the *società* a long series of different roles and tasks can be identified. The presence of a hierarchical structure, along with the presence of a set of formal rules, an extensive use of rituals and symbols, and strong barriers to entry, guarantees the internal cohesion and minimizes the risk of defections. Extortion still remains one of the main activities conducted by the 'Ndrangheta, especially in its territory of origin. This illegal activity is particularly relevant to exercise control over a territory rather than to gain large profits, which are instead generated from drug trafficking.

For further information on the 'Ndrangheta, cf. Paoli, "An Underestimated Criminal Phenomenon," 212–238; Paoli, *Mafia Brotherhoods*; Varese, "How Mafias Migrate," 411–443; Ciconte, '*Ndrangheta*; and CPA, *Relazione Annuale Sulla 'Ndrangheta*.

[32] DIA, *Relazione Del Ministro dell'Interno*, 105–106.

[33] Wiretap records report all wiretapped telephone conversations which were considered pertinent to the investigation by the public prosecutor and the defence attorneys. From a comparative perspective, these wiretap records can be considered consistent with any judicial document reporting the set of wiretap conversations transcribed by the police and made available to the prosecutor to be used as evidence in court.

In the Italian criminal justice system, before the actual conclusion of the investigation, the prosecution wraps up the evidence and formulate the request for remanding suspects in custody or pre-trial detention. The court order issued by the preliminary investigation judge (*giudice per le indagini preliminari*) upon this request is similar to arrest warrants in other jurisdictions.

The Italian criminal justice system enables up to three grades of judgment by different courts. For this study, the first grade judgment, which is issued by a tribunal composed by a panel of three judges, has been used. From a comparative perspective, this court document is similar to first grade judgments in other jurisdictions (cf. Calderoni, "Structure of Drug Trafficking Mafias," 325).

court document to another. The number of actors included in the networks and the number and proportion of links among them provide an insight into the overall group coverage when moving from data from wiretap records to those from the trial judgment, although a thorough analysis of the effects of the loss of information on overall network measures, such as density or cliques, is not provided and would require further research. However, the aim of the first part of the analysis is not limited to the assessment of whether a loss of information occurs. Node and edge removal might account for a different proportion of missing data. Also, node and edge removal might not be random, but actors or links with specific characteristics might be more likely to be excluded by the purposive sample after the investigation phase. Evidence of any pattern in nodes and edges removal is then gathered before moving to the second part of the analysis.

The second part of the analysis aims at examining the robustness of some of the most commonly used centrality measures when data are incomplete due to purposive sampling by law enforcement agencies. A set of five criminal networks based on empirical data from wiretap records are used for the analysis. This set includes the Oversize network and four criminal groups analysed by Carlo Morselli in his previous work and whose network data are publicly available.[34] The main requirement for the networks to be included in the analysis was the availability of data from wiretap records, whereas criminal groups whose networks are based on information from arrest warrants or judgments were excluded. However, some criminal networks analysed in the literature and based on information from wiretap records were excluded as well, since the data are not publicly available.

Table 1. Networks included in the study.

	No. nodes	No. edges	Max. edges	Density	Mean Degree
Oversize	182	247	16471	0.015	2.714
Caviar	110	205	5995	0.034	3.727
Siren	44	103	946	0.109	4.682
Togo	33	47	528	0.089	2.848
Ciel	25	35	300	0.117	2.800

Table 1 lists the networks used for the study, along with their size, density, and mean degree.[35] The choice of relying on empirical data, rather than generating random networks, depends on the acknowledgment that criminal networks generally present structural characteristics different from random ones, such as a power-law degree distribution where a large number of actors has only one or a few links and a small percentage of actors is highly connected.[36] Conversely, the reliance on empirical data means that it was not possible to control for the original network size and density, so that their

[34] The four networks whose data are available in Morselli's book are: Ciel, Caviar, Siren, and Togo. The Ciel and Caviar networks refer to police investigations that targeted groups involved in drug importation. Projects Siren and Togo were instead operations against stolen vehicle exportation networks. For each case study, Morselli did not simply described the organizational structure of the criminal groups, but rather focused on a specific feature of the networks, such as partnership configurations in drug trafficking groups (Ciel network), the contribution of legitimate players in criminal operations and the trade-off between efficiency and security (Caviar network), or brokerage and criminal network flexibility (Siren and Togo networks). For more details on the criminal groups analysed by Morselli and their network features cf. Morselli, *Inside Criminal Networks*.

[35] All analyses presented were conducted using binary, undirected matrices, since valued, directed data are not available for all Morselli's criminal networks.

[36] Xu and Chen, "Topology of Dark Networks."

influence on measurement error could not be fully addressed. Also, the original sample of networks was very limited, so that the replication of this study with a larger and/or different set of networks could provide further insights.

After computing centrality, the five selected networks were sampled in order to simulate the loss of information that was experienced by the Oversize network when moving across different court documents; centrality was then computed again and the results were compared to the centrality scores of the original network. Based on what was observed for the Oversize network, the measurement error introduced to construct the samples combined nodes and edge removal, so that the removal of nodes accounted for the 80% of missing links and the remaining 20% of missing edges concerned links between actors who were not affected by the node removal. Also, whereas the removal of edges was random, the probability for a node to be removed from the network varied according to its degree centrality, so that highly connected actors were less likely to be removed and peripheral nodes were more likely to be excluded from the sample.[37]

Different proportions of error were introduced, starting with the removal of 5% of the nodes (error I), and then 10%, 20%, 30%, and 50% (respectively, error II, III, IV, and V). For each of the five networks, and for each of the proportions of error, the sampling procedure was replicated 1000 times. Degree and betweenness centrality were then computed, and their robustness was assessed using the five measures of centrality robustness proposed by Borgatti and colleagues and presented in Table 2.[38] Scatterplots were then created to analyse the robustness of centrality measures as a function of the amount of measurement error added due to nodes and edges removal.[39]

Table 2. Measures of centrality robustness.

Measure	Description
Top 1	Proportion of times that the most central node in the sampled network is also the most central node in the empirical network.
Top 3	Proportion of times that the most central node in the sampled network is among the top three most central nodes in the empirical network.
Top 10%	Proportion of times that the most central node in the sampled network is among the top ten percent in the empirical network.
Overlap	Number of nodes in both the top 10% of the empirical network and the top 10% of the sampled network, divided by the number of node in either.[a]
R^2	Square of the Pearson correlation between the centralities of the empirical network and the centralities of the sampled network, taking only nodes found in both networks.

Note. [a]The overlap between the top 10% of the empirical network and the top 10% of the sampled network is computed as $|S \cap E| / |S \cup E|$, where S is the subset of nodes of the sampled network and E is the subset of nodes of the empirical one.
Source. Borgatti, Carley, and Krackhardt 2006.

4. Results

The first part of this section focuses on the differences across the three Oversize networks in terms of the amount of information on actors and their relations available in different judicial documents. To

[37] $p_i = \frac{deg_i}{\sum_{i=1}^{N} deg_i}$ where p_i is the probability of the node i to be included in the sample, and deg_i is its value of degree centrality.

[38] Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures."

[39] The analyses were performed using the sna and network packages in R (Butts, 'sna'; Butts, Handcock, and Hunter, 'network: Classes for Relational Data'). For further information on the sna and network packages, cf. Butts, "Social Network Analysis with sna"; Butts, "network."

assess whether and to what extent the arrest warrant and the judgment can provide an accurate description of the criminal group, Table 3 presents some descriptive statistics of the three Oversize networks.

Table 3. Descriptive statistics of the Oversize networks.

	No. Actors	% Actors	No. Links	% Links	Density	Mean Degree
Wiretap records	182	100	247	100	0.015	2.714
Arrest warrant	146	80.22	189	76.52	0.018	2.589
Judgment	89	48.90	106	46.15	0.029	2.562

Wiretap records report telephone conversations among 182 individuals. The total number of phone conversations wiretapped by the police and reported in the document are 933; however, data are here treated as binary, leading to 247 links between the actors, equal to the 1.5% of all possible links (density = 0.015). When the same network is based on the information available in the arrest warrant, the number of actors decreases to 146, equal to the 80% of those listed in the wiretap records. The number of links shows a similar trend, since it also experiences a drop of about 20% (from 247 to 189). The density of the network slightly increases (1.8%), probably because of the reduction in network actors. When moving from the arrest warrant to the judgment, the drop in actors and links is even lager. The remaining network members are indeed only 89, whereas the links among them are 114 (46.15% of the total number of observed relations).

Both the number of actors and the number of edges strongly decrease when moving from the network based on data from the wiretap records to the networks based on information from the arrest warrant and then the judgment. However, this drop does not seem to influence the measures of cohesion of the network. The average degree is indeed stable across the three networks and the density remains low, although its value has to be carefully compared across networks of different size.

A possible explanation of this phenomenon can be found by observing the degree distribution for the three Oversize networks in Figure 2. The three networks are characterised by the presence of a small number of actors involved in a large number of relations with other network members, and a large number of actors with only one or a few links. This confirms the results from previous studies, which found criminal networks to be characterised by the presence of a few highly connected individuals and a large number of peripheral actors.[40] As we move from the investigation phase to the judgment, the number of peripheral individuals decreases, whereas the highly connected nodes are still in the network, although the number of links to other criminal network members has dropped. This suggests that, when moving to the judgment, the remaining actors are those with relatively higher degree, while most of those who were excluded after the investigation phase were at the periphery of the network.

Simply looking at the visual representations of the three networks in Figure 3 provides a first confirmation of what observed above. Also, further evidence of the differences in terms of degree between actors who are present in all three networks and actors whose names appear only in the wiretap records or the arrest warrant is provided by the descriptive statistics reported in Table 4.

[40] Morselli, *Inside Criminal Networks*; Natarajan, "Drug Trafficking Organization"; and Natarajan, "Heroin Distribution Network."

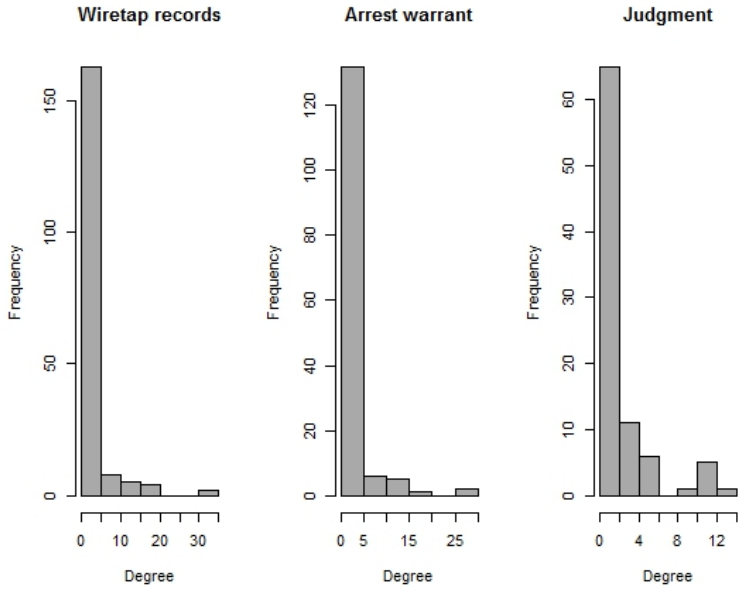

Figure 2. Degree distribution of Oversize networks.

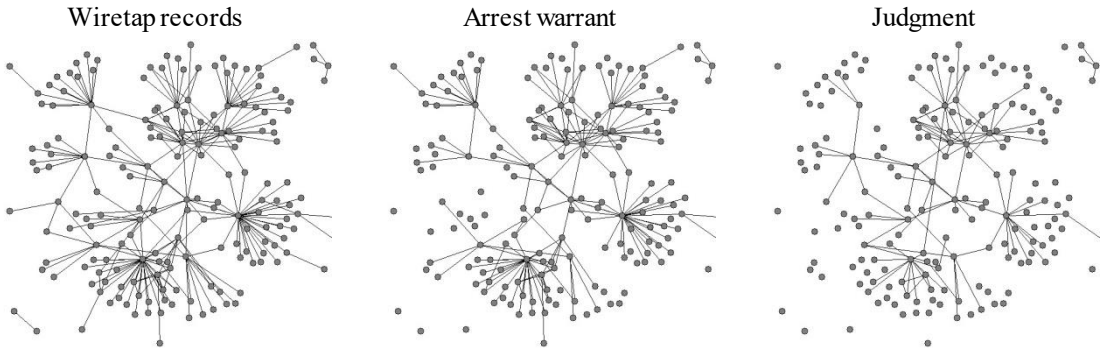

Figure 3. Sociograms of Oversize networks.

Table 4. Degree and density between and within subgroups.

	Arrest warrant		Judgment	
	Yes (n=146)	No (n=36)	Yes (n=89)	No (n=93)
Mean degree	3.12	1.08	4.37	1.13
Max degree	32	2	32	3
Min degree	1	1	1	1
Density within groups	0.020	0.002	0.037	0.001
Density between groups	0.007	0.007	0.012	0.012

Note. All measures were estimated on the network based on wiretap records, distinguishing between actors whose information was also present in the arrest warrant and the judgment and those whose conversations were discarded after the investigation phase.

Actors who are present both in the wiretap records and in the arrest warrant have a mean degree three times bigger than those whose phone conversations are reported in the wiretap records only. The same consideration can be done when comparing actors whose phone conversations are reported in both the wiretap records and the judgment with those who are included only in the largest network; the former have a mean degree about four times bigger than the latter. Most individuals excluded from the smaller sample have indeed only one link to other members of the criminal group, and a maximum of 2 for the network based on the arrest warrant and 3 for the network based on data available in the judgment. Finally, the peripheral position of actors whose information was discarded in the later phases of the judicial process also emerges from the comparison of density scores within and between groups (Table 4). In both cases, actors who are present only in the larger network have less contact among themselves than with those who can be found in all three networks, who have a higher density within group and form the core of the criminal network.

Peripheral and poorly connected actors are thus more likely to be excluded from the network when moving from the investigation to the trial phase. This trend can be explained in two different and non-mutually exclusive ways. On the one hand, actors with higher degree are probably those more involved in criminal activities and more active during the investigations. On the other hand, they simply might be those whose phones or lines were actually wiretapped, or those who were targeted by the police in the early stages of the criminal investigation. Being the main focus of the investigation, the police was able to gather more information about them, despite their actual position in the criminal group, although long investigations should be less affected by the goals of law enforcement agencies and the way the investigation was conducted, especially in its early stages.[41] In both cases, law enforcement agencies could collect more evidence on these individuals and consequently arrest and charge them. Actors with one or a few contacts were instead either less involved in criminal activities or not directly wiretapped, and this might have led the police to clear them and to discard the conversations they were involved in from further court documents.

An analysis of the outcome of the trial shows that actors who appeared to be central at the end of the investigation were more likely to be prosecuted and convicted, confirming results from previous studies.[42] A significant positive correlation was indeed found between degree centrality and being arrested ($r = .586$, $p < .01$) and convicted ($r = .256$, $p < .01$). Similar results were obtained after correlating betweenness centrality with being arrested ($r = .531$, $p < .01$) and convicted ($r = .305$, $p < .01$). Conviction was also positively correlated with the affiliation to the 'Ndrangheta ($r = .660$, $p < .01$). Actors who had a formal role within the mafia group were also those with the highest centrality scores. They were indeed particularly involved in the organisation of illegal activities and in group management, whereas peripheral individuals often participated in the activities of the group only in specific situations or with a marginal role. Most individuals whose conversations were only reported in wiretap records were involved in the retail or purchase of the drug or in support of traffickers. 'Ndrangheta members, on the other hand, were directly involved in drug trafficking and other illegal activities, and thus needed to participate in more conversations with other network members.

The exclusion of peripheral actors from the arrest warrant and the judgment leads also to the loss of their links. However, missing edges for reasons other than node removal might also occur. In the Oversize network, the exclusion of peripheral actors accounted for the 65.5% of missing edges when moving from the network based on the wiretap records to the network based on the arrest warrant, and

[41] Campana and Varese, "Listening to the Wire," 16.
[42] Morselli, "Vulnerable and Strategic Positions."

for the 75.9% of missing edges in the network created with information from the judgment. The remaining missing links (34.5% and 24.1%, respectively) instead concerned the loss of information about actors who were still part of the criminal network when moving from the investigation to the judgment, but not linked to each other.

The analysis of the differences across the three Oversize networks leads to a first partial conclusion that the reliance on data from court documents such as the arrest warrant and the judgment can affect the results when the aim of the research is a precise description of a criminal group and the relations among its members. Although wiretap records also report only a selection of the electronic communication transcripts, they can provide more accurate information of criminal networks, rather than the arrest warrant and the judgment, in which half of the actors and links might be missing. However, the researcher might seek to obtain indicators reflecting differences among actors in terms of network position or differences across networks in terms of structural properties, rather than a precise description of the criminal group. In this case, the exclusion of peripheral actors, along with the loss of a small percentage of other links among network members, might not have a significant impact on group coverage or on network analysis measures of the criminal group, such as centrality measures.

The aim of the second part of the analysis is thus to assess the robustness of some of the most common centrality measures when data are incomplete due to purposive sampling by law enforcement agencies. The impact of random measurement error in random networks on the robustness of centrality measures has already been tested and the measures of centrality have proved to be quite robust under small amount of error.[43] However, criminal networks are not randomly constructed, nor is the loss of information systematic.

Table 5 presents the accuracy scores for the five criminal networks considered in this study. Consistently with the results of Borgatti and colleagues the two centrality measures behave similarly across increasing measurement error due to purposive sampling by law enforcement agencies, although betweenness tends to be less robust than degree in two cases (Oversize and Ciel). This is also supported by Figure 4, which presents robustness as a function of error level and allows to compare the different measures of centrality robustness across the five criminal networks. The Overlap measure is among the measures of centrality robustness that declines more rapidly. Only in the Ciel network, its value does not experience a sharp drop, probably because of the small size of the network. However, contrary to what happened with the random removal of nodes and edges from random networks, here the value of the Overlap measure is always above 0.25 and its curve usually becomes steeper after the removal of 10% of the nodes (error II).

But the most interesting result is probably the fact that the Top 3 and Top 10% curves show only a small deviation from 1, also when half of the nodes are removed from the original network. This result could be partially expected, considering the sampling procedure adopted, with actors with high degree being less likely to be removed from the network; however, this observation is consistent across all five networks and there are no significant differences between degree and betweenness centrality. The largest decrease is indeed the decrease of 0.14 when half of the nodes are sampled, experienced by betweenness for the Oversize network. Hence, also when 50% of nodes are missing (error V), the actor who appears to be central in the sampled network can be found among the top 3 actors and the top 10% actors of the original network almost 90% of times.

[43] Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures."

Table 5. Results for all networks.

			% Error - Degree							% Error - Betweenness		
Oversize	NA	I	II	III	IV	V	*Oversize*	NA	I	II	III	IV	V
Top 1	1.00	0.62	0.59	0.49	0.36	0.27	Top 1	1.00	0.28	0.27	0.22	0.12	0.05
Top 3	1.00	1.00	1.00	1.00	1.00	1.00	Top 3	1.00	0.93	0.92	0.91	0.89	0.86
Top 10%	1.00	1.00	1.00	1.00	1.00	1.00	Top 10%	1.00	1.00	1.00	1.00	1.00	1.00
Overlap	1.00	0.924	0.880	0.776	0.666	0.499	Overlap	1.00	0.917	0.874	0.772	0.664	0.496
R-Squared	1.00	0.983	0.979	0.967	0.949	0.886	R-Squared	1.00	0.943	0.942	0.931	0.911	0.849
Caviar	NA	I	II	III	IV	V	*Caviar*	NA	I	II	III	IV	V
Top 1	1.00	1.00	1.00	1.00	1.00	1.00	Top 1	1.00	1.00	1.00	1.00	1.00	0.99
Top 3	1.00	1.00	1.00	1.00	1.00	1.00	Top 3	1.00	1.00	1.00	1.00	1.00	1.00
Top 10%	1.00	1.00	1.00	1.00	1.00	1.00	Top 10%	1.00	1.00	1.00	1.00	1.00	1.00
Overlap	1.00	0.826	0.788	0.715	0.629	0.453	Overlap	1.00	0.841	0.777	0.700	0.620	0.449
R-Squared	1.00	0.991	0.989	0.984	0.976	0.952	R-Squared	1.00	0.989	0.985	0.979	0.970	0.943
Ciel	NA	I	II	III	IV	V	*Ciel*	NA	I	II	III	IV	V
Top 1	1.00	0.92	0.90	0.84	0.84	0.84	Top 1	1.00	0.66	0.66	0.61	0.62	0.62
Top 3	1.00	1.00	1.00	1.00	1.00	1.00	Top 3	1.00	1.00	1.00	1.00	0.99	0.99
Top 10%	1.00	1.00	1.00	1.00	1.00	1.00	Top 10%	1.00	1.00	1.00	1.00	0.99	0.99
Overlap	1.00	1.00	1.00	0.938	0.938	0.938	Overlap	1.00	0.956	0.938	0.938	0.938	0.938
R-Squared	1.00	0.966	0.958	0.932	0.910	0.910	R-Squared	1.00	0.946	0.932	0.913	0.904	0.904
Siren	NA	I	II	III	IV	V	*Siren*	NA	I	II	III	IV	V
Top 1	1.00	1.00	1.00	1.00	1.00	0.73	Top 1	1.00	1.00	1.00	1.00	0.99	0.62
Top 3	1.00	1.00	1.00	1.00	1.00	1.00	Top 3	1.00	1.00	1.00	1.00	1.00	1.00
Top 10%	1.00	1.00	1.00	1.00	1.00	1.00	Top 10%	1.00	1.00	1.00	1.00	1.00	1.00
Overlap	1.00	0.631	0.640	0.693	0.642	0.492	Overlap	1.00	0.783	0.793	0.699	0.674	0.495
R-Squared	1.00	0.983	0.979	0.947	0.903	0.743	R-Squared	1.00	0.989	0.986	0.964	0.935	0.814
Togo	NA	I	II	III	IV	V	*Togo*	NA	I	II	III	IV	V
Top 1	1.00	1.00	1.00	1.00	0.99	0.99	Top 1	1.00	1.00	1.00	1.00	1.00	0.99
Top 3	1.00	1.00	1.00	1.00	1.00	1.00	Top 3	1.00	1.00	1.00	1.00	1.00	1.00
Top 10%	1.00	1.00	1.00	1.00	1.00	1.00	Top 10%	1.00	1.00	1.00	1.00	1.00	1.00
Overlap	1.00	0.997	0.972	0.666	0.662	0.652	Overlap	1.00	0.980	0.959	0.657	0.652	0.652
R-Squared	1.00	0.976	0.972	0.958	0.937	0.885	R-Squared	1.00	0.975	0.971	0.963	0.945	0.905

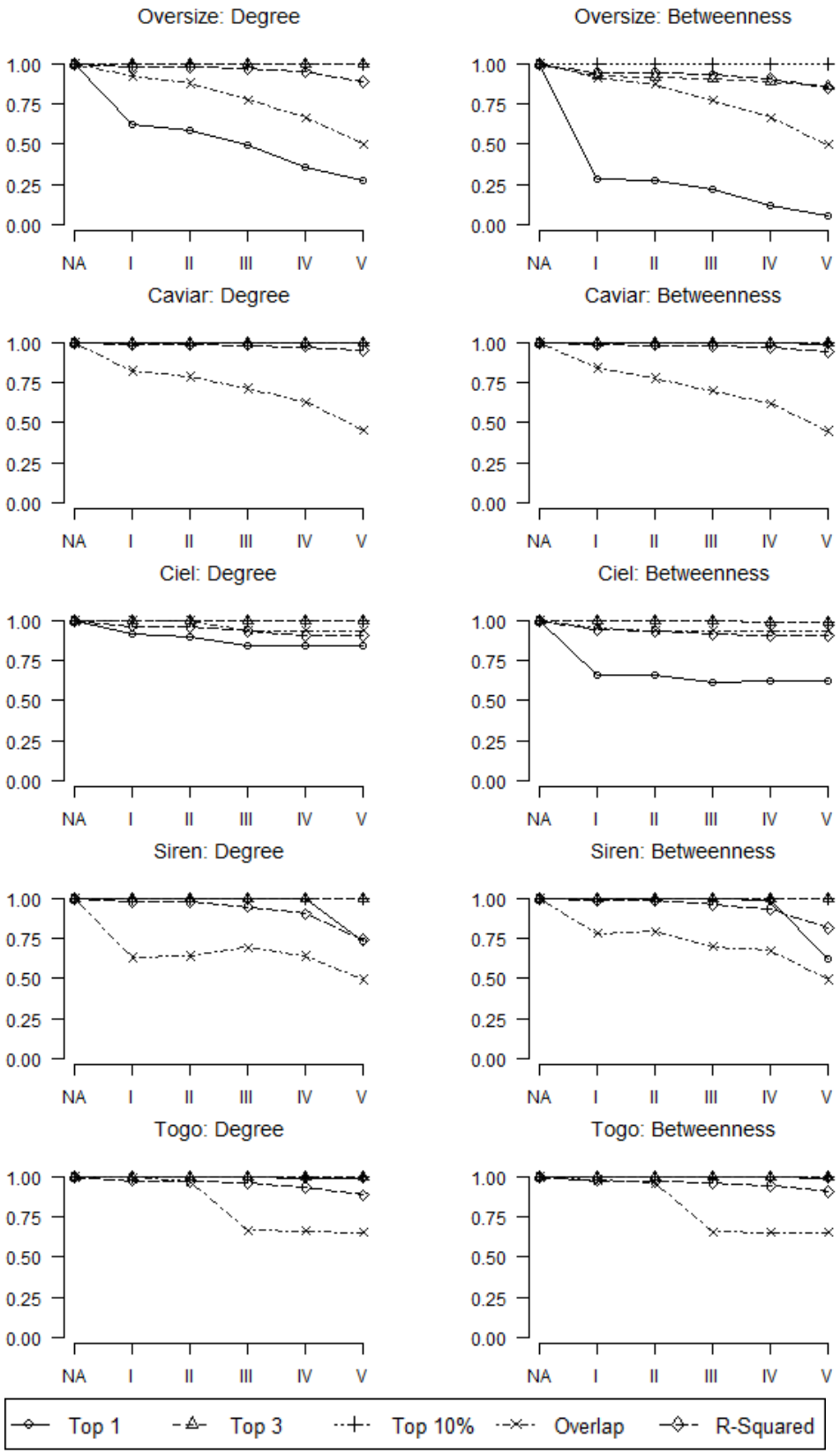

Figure 4. Degree and betweenness accuracy as a function of error level for all networks.

The Top 1 measure of centrality robustness shows instead varying results across the five networks considered in this study. The actor with the highest degree centrality in the sampled network appears to be the one with the highest degree score in the original network most of times, with one exception, namely the results for the Oversize network. Further analyses would be necessary to support any conclusion, but a possible explanation might be related to the degree distribution and the presence, in the Oversize network, of two actors with similarly high centrality scores, so that variations in the number of nodes and edges might easily lead to the exchange of the first and second position in the network, in terms of degree. When the difference between the most central actor and the other network members is instead heightened, any position exchange might be less likely to occur and the Top 1 measure should be more robust after nodes and edges removal.

The Top 1 measure of robustness shows a similar trend also for betweenness centrality. The Oversize network still presents values significantly lower than all other criminal groups, although the four remaining networks present more variability in terms of the robustness of the Top 1 measure than they do in the case of degree centrality. The robustness of the Top 1 measure for the Caviar network, for example, experiences a substantial drop after the removal of 5% of nodes (error I), although the accuracy scores do not decrease further after the first drop.

To conclude, both centrality measures appear to be more robust in case of purposive sampling by law enforcement agencies than in case of random errors in random networks. Also, when half of nodes are missing, there is a high probability that the actor with the highest centrality score in the sampled network is also among the first three most central nodes in the original, overall network. The overlap between the top actor in both the sampled and the original network is instead less likely to occur, as the measurement error increases and it is likely to be influenced by the degree distribution within the network, along with its size and density. The next two paragraphs discuss the implications of these results for criminal network analysis and present the limitations of the study, whereas the last paragraph draws the conclusions.

5. Limitations

This study comes with a series of limitations that hinder the generalisation of the results. First, to create the networks samples it has been assumed that the loss of information from the wiretap records to the judgment experienced by the Oversize network is likely to occur to the same extent in other criminal investigations. It is not possible to know whether this is actually true, although it is plausible to expect a similar trend, namely the exclusion of peripheral actors rather than central ones as the investigation proceeds.

Second, node and edge removal might account for a different proportion of missing data in different networks, although the sampling procedure adopted in this study is more realistic than random removal. Also, only nodes were sampled according to their degree, whereas another 20% of edges was randomly removed. Replicating the same study with valued and directed data might provide an insight into the characteristics of this 20% of links who were missing from the network for reasons other than node removal. We might expect, for example, that reciprocated ties are less likely to be missing. Or that links based only on one or a few telephone calls are more likely to be excluded than those based of a larger number of conversations.

Finally, the reliance on empirical data, rather than generating random networks, did not allow to control for network size and density, so that their influence on the measurement error could not be fully

addressed. Also, the original sample of networks was very limited, mainly because of the small number of publicly available data on criminal networks. The replication of this study with a larger set of networks, as well as a different one, might provide further insights into the robustness of centrality measures in this particular setting.

Despite these limitations, some conclusions on the reliability of different court records to analyse criminal networks and identify central actors can be drawn.

6. Discussion

The first aim of this study was to assess whether data from court documents other than wiretap records can provide an accurate description of criminal networks. The analysis of the three Oversize networks allowed the identification of differences among sets of relational data extracted from different court documents. Both the number of actors and the number of links among them decrease, and in the judgment, only about 50% of both actors and edges are reported. Due to this substantial drop, both the arrest warrants and the judgment do not seem to provide an accurate description of the network, although the two measures of cohesion of the network – density and mean degree – remain fairly stable across the three Oversize networks. However, the lesser accuracy of judgments and, to some extent, of arrest warrants, compared to wiretap records, may lead to an incomplete picture of the network and, consequently, to misleading results for the overall network measures, as suggested by Campana and Varese.[44]

The first part of the analysis also provided evidence of the purposive sampling conducted by law enforcement agencies when moving from the investigation to the trial phase and discarding part of the electronic surveillance data. Peripheral and poorly connected actors were more likely to be excluded from the network after the investigation phase, either because they are less involved in criminal activities or they do not contribute directly to the issues targeted by the police.

The second aim of this research was instead to assess the reliability of data found in different judicial documents to identify key players. The results seem to suggest that degree and betweenness centrality are quite robust under conditions of missing nodes and edges. These findings are consistent with those from previous studies, which proved the accuracy of centrality measures under small amounts of random measurement error.[45] They also show that the two most commonly used centrality measures are even more robust in organised crime research, where networks are not random, and neither is the measurement error, since actors with low degree are more likely to be excluded from the networks based on information from arrest warrants and judgments. Degree and betweenness centrality have indeed proved to be particularly robust, also under large amounts of error. Whereas the correct identification of the most central actor might not always be accurate, especially when a large number of nodes and edges is missing, widening the scope to the first three actors is likely to provide reliable results in terms of the identification of the key players within the criminal group.

On the basis of these results, some considerations on the implications that the reliance on a specific source of wiretap data might have on the SNA of criminal groups can be made. Despite the probability of missing information, arrest warrants and judgments seem to be reliable data sources to identify key

[44] Campana and Varese, "Listening to the Wire."

[45] Bolland, "Sorting Out Centrality"; Costenbader and Valente, "Stability of Centrality Measures"; Borgatti, Carley, and Krackhardt, "Robustness of Centrality Measures"; and Xu and Chen, "Topology of Dark Networks."

players regardless of whether a large proportion of peripheral nodes is missing. The purposive sampling conducted by law enforcement agencies guarantees a low probability of relational information on key players to be discarded when moving from the investigation phase to the judgment. Also, contrary to other judicial documents, judgments are publically available and contain information other than conversations among network members. Provided that they include sufficient information on relations among the members of a criminal group, they thus may be a useful source of data for the SNA of criminal organisations.[46]

Considering the amount of information present in the different types of court documents, wiretap records should probably be the first choice for researchers interested in the analysis of criminal networks. However, given the public availability of judgments, as opposed to wiretap records and arrest warrants, and the robustness of centrality measures under conditions of imperfect data, the reliance on judgments might further facilitate the spread of SNA methods in organised crime research.

7. Conclusions

This study has addressed the issue of the accuracy and reliability of different law enforcement data sources for the analysis of criminal networks through SNA and the identification f key players. Despite several limitations, the results suggest that degree and betweenness centrality are particularly robust when the available information constitutes a purposive sample of all wiretapped telephone conversations among criminal network members. This findings have both research and policy implications.

First, they suggest that researchers aiming at applying SNA in a criminological context should be confident about relational data collected from different kinds of court records, including judgments, especially when the focus of the research is on differences between individuals in terms of their position within the criminal group, rather than a precise description of the group itself. This is particularly relevant considering that some types of judicial sources, such as judgments, are publicly available once the case is closed, whereas wiretap records and arrest warrants might not be accessible to the researcher.

Second, evidence of the robustness of centrality measures under conditions of imperfect data means that the adoption of network analysis measures during criminal investigations may help law enforcement agencies to identify key players within the criminal group under investigation, in order to arrest them and achieve the maximum of disruption of the network.

The positive findings with respect to the reliability of data from different court documents should however be supported and strengthen by further studies involving different criminal investigations in different jurisdictions, a larger set of criminal networks, or relational data where the information on the direction and amount of telephone calls is preserved.

Acknowledgements I thank Ferdinando Pomarici and Galileo Proietto, state prosecutors of the Antimafia District Directorate (Direzione Distrettuale Antimafia, DDA) of Milan, who provided access to the judicial documents for operation Oversize. I am also grateful to Francesco Calderoni (Università Cattolica del Sacro Cuore, Milano), George Tita (University of California, Irvine), and the two anonymous reviewers for their valuable comments. My gratitude also goes to Zack Almquist and Adam Boessen (University of California, Irvine) for their help with the sna and network packages in R.

[46] Bright, Caitlin, and Chalmers, "Illuminating Dark Networks," 174.

References

Baker, W. E., and R. Faulkner. "The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry." *American Sociological Review* 58, no. 6 (1993): 847.

Bolland, J. M. "Sorting Out Centrality: An Analysis of the Performance of Four Centrality Models in Real and Simulated Networks." *Social Networks* 10, no. 3 (1988): 233–253.

Bonacich, P. "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92 (1987): 1170–1182.

Borgatti, S. P., K. M. Carley, and D. Krackhardt. "On the Robustness of Centrality Measures Under Conditions of Imperfect Data." *Social Networks* 28 (2006): 124–136.

Bright, D. A., E. H. Caitlin, and J. Chalmers. "Illuminating Dark Networks: A Social Network Analysis of an Australian Drug Trafficking Syndicate." *Crime, Law and Social Change* 57, no. 2 (2012): 152.

Bruinsma, G., and W. Bernasco. "Criminal Groups and Transnational Illegal Markets: A More Detailed Examination on the Basis of Social Network Theory." *Crime, Law and Social Change* 41, no. 1 (2004): 79–94.

Butts, C. T. "Social Network Analysis with sna." Journal of Statistical Software 24, no. 6 (2008): 1–51

Butts, C. T. "network: A Package for Managing Relational Data in R." *Journal of Statistical Software* 24, no. 6 (2008): 1–36

Butts, C. T. "sna: Tools for Social Network Analysis." R Package Version 2.2-0, 2010.

Butts, C. T., M. S. Handcock, and D. R. Hunter. "network: Classes for Relational Data." R Package Version 1.7-1, 2012.

Calderoni, F. "The 'Ndrangheta Through the Lens of Social Network Analysis: A Tale from Calabria." Presented at the American Society of Criminology Annual Meeting 2010, San Francisco, CA, November 17, 2010.

Calderoni, F. "The Structure of Drug Trafficking Mafias: The 'Ndrangheta and Cocaine." *Crime, Law and Social Change* 58, no. 3 (2012): 325.

Campana, P. "Eavesdropping on the Mob: The Functional Diversification of Mafia Activities Across Territories." *European Journal of Criminology* 8, no. 3 (2011): 213–228.

Campana, P., and F. Varese. "Listening to the Wire: Criteria and Techniques for the Quantitative Analysis of Phone Intercepts." *Trends in Organized Crime* 15, no. 1 (2012): 13–30.

Carrington, P. J. "Crime and Social Network Analysis." In *Sage Handbook of Social Network Analysis*, edited by J. Scott and P. J. Carrington, 244. London: SAGE Publications, 2011.

Ciconte, E. *'Ndrangheta*. Soveria Mannelli: Rubbettino, 2008.

CPA (Commissione Parlamentare Antimafia). *Relazione Annuale Sulla 'Ndrangheta*, Doc. XXIII, N. 5, XV Legislatura, Roma: Commissione parlamentare di inchiesta sul fenomeno della criminalità organizzata mafiosa o similare, 2008.

Costenbader, E., and T. W. Valente. "The Stability of Centrality Measures When Networks Are Sampled." *Social Networks* 25, no. 4 (2003): 283–307.

DIA (Direzione Investigativa Antimafia). *Relazione Del Ministro dell'Interno Al Parlamento Sull'attività Svolta e Sui Risultati Conseguiti Dalla Direzione Investigativa Antimafia*. 1◦ Semestre 2007. Roma: Direzione Investigativa Antimafia, 2007, 105–106.

Freeman, L. C. "Centrality in Social Networks: Conceptual Clarification." *Social Networks* 1, no. 3 (1979): 223–258.

Ianni, F. A. J., and E. Reuss-Ianni. "Network Analysis." In *Criminal Intelligence Analysis*, edited by P. Andrews and M. B. Peterson, 81–82. Loomis, CA: Palmer Enterprises, 1990.

Klerks, P. "The Network Paradigm Applied to Criminal Organizations: Theoretical Nitpicking or a Relevant Doctrine for Investigators? Recent Developments in the Netherlands." *Connections* 24, no. 3 (2001): 53–65.

Kossinets, G. "Effects of Missing Data in Social Networks." *Social Networks* 28, no. 3 (2006): 249.

Laumann, E. O., P. V. Marsden, and D. Prensky. "The Boundary Specification Problem in Network Analysis." In *Research Methods in Social Network Analysis*, edited by L. C. Freeman, D. R. White, and A. Kimball Romney, 62. New Brunswick, NJ: Transaction Publishers, 1992.

Malm, A. E., and G. Bichler. "Networks of Collaborating Criminals: Assessing the Structural Vulnerability of Drug Markets." *Journal of Research in Crime and Delinquency* 48, no. 2 (2011): 20.

Malm, A. E., G. Bichler, and S. Van De Walle. "Comparing the Ties That Bind Criminal Networks: Is Blood Thicker Than Water?" *Security Journal* 23, no. 1 (2010): 70.

Marsden, P. V. "Network Data and Measurement." *Annual Review of Sociology* 16, no. 1 (1990): 439.

McAndrew, D. "The Structural Analysis of Criminal Networks." In *The Social Psychology of Crime: Groups, Teams and Networks*, edited by D. Canter and L. Alison, Offender Profiling Series III, 51–94. Aldershot: Ashgate Publishing, 1999.

McIllwain, J. S. "Organized Crime: A Social Network Approach." *Crime, Law and Social Change* 32, no. 4 (1999): 301–323.

Morselli, C. "Assessing Vulnerable and Strategic Positions in a Criminal Network." *Journal of Contemporary Criminal Justice* 26, no. 4 (2010): 382–392.

Morselli, C. *Contacts, Opportunities and Criminal Enterprise*. Toronto: University of Toronto Press, 2005.

Morselli, C. *Inside Criminal Networks*. New York, NY: Springer, 2009.

Morselli, C., and K. Petit. "Law-Enforcement Disruption of a Drug Importation Network." *Global Crime* 8, no. 2 (2007): 110.

Natarajan, M. "Understanding the Structure of a Drug Trafficking Organization: A Conversational Analysis." In *Illegal Drug Markets: From Research to Policy*, edited by M. Natarajan and M. Hough, 273–298. Vol. 11. Crime Prevention Studies. Monsey, NY: Criminal Justice Press, 2000.

Natarajan, M. "Understanding the Structure of a Large Heroin Distribution Network: A Quantitative Analysis of Qualitative Data." *Journal of Quantitative Criminology* 22, no. 2 (2006): 171–192.

Paoli, L. Mafia Brotherhoods: Organized Crime, Italian Style. New York: Oxford University Press, 2003.

Paoli, L. "An Underestimated Criminal Phenomenon: The Calabrian 'Ndrangheta." *European Journal of Crime, Criminal Law and Criminal Justice* 2, no. 3 (1994): 212–238.

Schwartz, D. M., and T. D. A. Rouselle. "Using Social Network Analysis to Target Criminal Networks." *Trends in Organized Crime* 12, no. 2 (2009): 189.

Spapens, T. "Macro Networks, Collectives, and Business Processes: An Integrated Approach to Organized Crime." European Journal of Crime, *Criminal Law and Criminal Justice* 18, no. 2 (2010): 193.

Sparrow, M. K. "The Application of Network Analysis to Criminal Intelligence: An Assessment of the Prospects." *Social Networks* 13 (1991): 251–274.

Varese, F. "The Structure of Criminal Connections: The Russian-Italian Mafia Network." Oxford Legal Studies Research Paper 21 (2006): 1–71

Varese, F. "How Mafias Migrate: The Case of the 'Ndrangheta in Northern Italy." *Law and Society Review* 40, no. 2 (2006): 411–443.

van der Hulst, R. C. "Introduction to Social Network Analysis (SNA) as an Investigative Tool." *Trends in Organized Crime* 12, no. 2 (2009): 101–121.

von Lampe, K. "Human Capital and Social Capital in Criminal Networks: Introduction to the Special Issue on the 7th Blankensee Colloquium." *Trends in Organized Crime* 12, no. 2 (2009): 93–100.

Wasserman, S., and K. Faust. *Social Network Analysis: Methods and Applications*. New York, NY: Cambridge University Press, 1994.

Williams, P. "Transnational Criminal Networks." In *Networks and Netwars: The Future of Terror, Crime and Militancy*, edited by J. Arquilla and D. Ronfeldt, 159. Washington, DC: RAND Corporation, 2001.

Xu, J. J., and H. Chen. "The Topology of Dark Networks." *Communications of the ACM* 51, no. 10 (2008): 58–65.