Social Networks 48 (2017) 1–9

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

The structure and vulnerability of a drug trafficking collaboration network

George Wood*

Department of Sociology, University of Oxford, UK

ARTICLE INFO

Article history:
Available online 31 August 2016

Keywords:
Collaboration
Criminal network
Degree distribution
Assortativity
Vulnerability

ABSTRACT

Prior research has established that collaboration networks exhibit heavy-tailed degree distributions, assortative degree mixing, and large clustering coefficients. Using court record data, we assess these properties in a collaboration network among heroin traffickers. Consistent with prior research, we find an exponential degree distribution and strong local clustering. However, the traffickers mix dissortatively by degree rather than assortatively. Using a graph sampling method, we show that a consequence of dissortative mixing is that targeted vertex removals have a greater impact on the connectivity and cohesion of the trafficking network. We also note the importance of degree mixing for characterizing and identifying topological weaknesses.

1. Introduction

Many criminal endeavors are complex and demand a large degree of collaboration and coordination. Criminals rely on each other for tasks that cannot be performed alone, such as the trafficking of drugs and humans, money laundering, fixing sports events, and rigging markets. In this regard, however, criminals are not unique. After all, many non-criminal endeavors are just as complex. What makes criminal collaboration extraordinary is that, in most cases, criminals operate under the constant threat of detection and the punitive consequences that this entails.

Much research has focused on how these extraordinary conditions shape the structure of collaboration in criminal networks. For instance, it has long been proposed that, in order to remain covert, criminal networks adopt particular topological features, such as greater sparsity and decentralization (Baker and Faulkner, 1993; Krebs, 2002; Kenney, 2007; Williams, 2001). Further work has addressed the trade-off between efficiency and security in criminal networks (McCormick and Owen, 2000; Morselli and Petit, 2007; Lindelauf et al., 2009). This work has, in part, motivated research into the vulnerability and disruption of criminal networks (Ayling, 2009; Duijn et al., 2014; Morselli, 2010; Malm and Bichler, 2011).

Meanwhile, there has been a great deal of interest in the structure of collaboration networks in non-criminal domains, such as academia, music, and business. It is recognized that such collaboration networks typically share three properties in common (Ramasco et al., 2004). First, the vertex degree distribution is heavy-tailed (Newman, 2001b; Barabási et al., 2002; Guimerá et al., 2005) such that a minority of actors have a disproportionately large number of collaborators. Second, the degree of connected vertices is positively correlated (Newman, 2005; Chang et al., 2007). Well-connected actors tend to work with other well-connected actors, while the poorly-connected work with the poorly-connected. Third, collaboration networks exhibit strong local clustering. Actors who share at least one mutual partner are more likely to become partners themselves, whether as a matter of choice or chance, than would be the case if vertices were randomly connected (Watts and Strogatz, 1998; Newman, 2001a; Moody, 2004).

However, while these properties are routinely examined in collaboration networks outside of the criminal networks literature, whether collaboration among criminals shares the same three properties has not been addressed directly. Although studies on criminal networks often consider degree and clustering coefficients, they seldom characterize the degree distribution formally, and degree mixing is typically overlooked entirely.

Yet, there are good reasons to be interested in whether collaboration between criminals exhibits the same topological properties as collaboration elsewhere. First, such comparison provides a systematic means of identifying irregularities in the structure of criminal networks. Research on clandestine networks has drawn criticism for lacking empirical benchmarks against which to assess whether the observed properties are, in fact, exceptional (Varese, 2012; Crossley et al., 2012). By drawing on empirical results in the wider literature on collaboration networks, such benchmarks are made explicit.

* Correspondence to: Nuffield College, New Road, Oxford OX1 1NF, UK.
E-mail address: george.wood@sociology.ox.ac.uk

http://dx.doi.org/10.1016/j.socnet.2016.07.001
0378-8733/

Second, the combination of the topological properties, and in particular degree mixing, are interesting in their own right. It has long been known that assortatively mixed networks are more robust to the removal of high degree vertices, whereas this strategy is considerably more effective in dissortative networks (Newman et al., 2002). In light of growing research interest in disrupting criminal networks, degree mixing is therefore a property of particular relevance.

In this paper, we examine the degree distribution, degree mixing, and clustering of a heroin trafficking network. We remain agnostic about how these properties should manifest in the network, and instead simply evaluate each property in turn. We find that the trafficking network has a heavy-tailed degree distribution, which may be classified as either exponential or log-normal. Additionally, the network exhibits strong local clustering. However, unlike other collaboration networks, the traffickers mix dissortatively by degree. Traffickers with a large number of collaborators tend to be connected to traffickers with relatively few, and vice versa.

The dissortative mixing has substantial consequences for the vulnerability of the network to targeted vertex removals. We demonstrate this by analyzing the impact of vertex removals on the heroin trafficking network under different mixing configurations using a graph sampling method. In variants of the trafficking network which are more assortative, but where the degree distribution is preserved, the impact of targeted vertex removals on the cohesion and connectivity of the network diminishes. Conversely, targeted vertex removals have a greater impact in more dissortative variants of the network. Thus, the property on which the trafficking network deviates from other types of collaboration network appears to be critical for its robustness to attacks.

The remainder of this paper is structured as follows. In Section 2 we briefly review previous research on the structure of collaboration networks and criminal networks. In Section 3, we describe the data. In Section 4 we outline the methods for evaluating each property in the network, along with the graph sampling method for generating networks with varying degree mixing configurations, and the measurements used to assess the impact of vertex removal. In Section 5 we present the results, before concluding in Section 6.

2. Previous research

2.1. Topological features of collaboration networks

Studies on the structure of collaboration networks abound. Here, we focus on three topological features that feature regularly in these studies: clustering, a heavy-tailed degree distribution, and assortative degree mixing. In an early study, Watts and Strogatz (1998) used data on the co-appearance of actors in movies and found that when an actor i has appeared in a movie with another actor j, and i has also appeared with a third-actor k, it is far more likely than one would expect by chance that j and k will have also appeared in a movie together. This clustering phenomenon is now a routine finding in social networks, and is a feature of co-authorship among scientists (Newman, 2001a,b) and social scientists (Moody, 2004), in collaboration among jazz musicians (Gleiser and Danon, 2003), and in the Broadway musical industry (Guimerá et al., 2005).

Second, collaboration networks often exhibit a heavy-tailed distribution of degree. Newman (2001a) demonstrated a heavy-tailed distribution in co-authorship networks among physicists, biomedical researchers, and computer scientists. Across disciplines, a minority of authors has a disproportionately large number of co-authors, while the majority has few. The heavy-tail in degree persists outside of scientific co-authorship. For example, collaboration among jazz musicians and in the Broadway musical industry also exhibit a heavy-tailed distribution of vertex degree. Moreover, this property holds for collaboration networks among organizations. Butts et al. (2012) show that a network of inter-organizational collaboration, developed in response to Hurricane Katrina, has a characteristic right-skewed distribution of degree, while collaboration among firms in the biotechnology industry is characterized by a small number of high-degree firms and much larger group of poorly-connected firms (Gay and Dousset, 2005).

Table 1
The clustering coefficient c, degree assortativity r, and degree assortativity error σ_r in several collaboration and criminal networks.

Network		c	r	σ_r
a	Network Science	0.638	0.462	0.072
b	Condensed Matter Physics	0.636	0.186	0.004
c	High Energy Physics	0.442	0.294	0.019
d	Astrophysics	0.639	0.235	0.005
e	Jazz	0.617	0.020	0.026
f	Students	0.636	0.366	0.113

Barabási and Albert (1999) propose that preferential attachment is part of the explanation for this pattern of inequality. Preferential attachment is often associated with a power law in the tail of the degree distribution (Barabási and Albert, 1999; Albert et al., 1999), such that the probability that an actor has precisely x collaborators abides by $P(x) \sim x^{-\alpha}$, where the α scaling parameter relates to the extent of inequality in the number of collaborators per vertex. This work, along with Huang et al. (2008) and Moody's (2004) research on academic collaboration, point to a power law in the tail of degree. Moreover, using dynamic data, Newman (2005) shows that the growth of a co-authorship network is consistent with preferential attachment.

The third property typical of collaboration networks is the positive correlation in the degree of connected actors. Newman (2003a) demonstrates that the partners of academics, movie actors, and businesspeople who have large vertex degree are likely to have many partners themselves, calling this property "degree assortativity." Chang et al. (2007) and Ramasco and Morris (2006) show the degree–degree correlation holds in the movie actor network for both weighted and unweighted degree, while Ramasco et al. (2004) finds assortative degree mixing in company co-directorships. We reanalyze some of the aforementioned networks and report the local clustering coefficient c, assortativity r, and standard error for the assortativity coefficient σ_r in Table 1.[1]

All of these networks are, of course, qualitatively different in kind. The scientific collaboration networks are made up of thousands of actors partitioned into smaller co-authoring teams, where the high-performing teams cluster into a large component (Newman, 2001a; Guimerá et al., 2005). In contrast, those in the Broadway musical industry are divided into somewhat larger teams on a show-by-show basis, where actors cross-over or belong to several teams (Guimerá et al., 2005). Movie actors and jazz musicians assemble around records and movies, respectively. However, in spite of these differences, all of the networks exhibit the clustering, heavy-tailed degree, and assortativity characteristics. Whatever the specific purpose of collaboration, whether producing jazz records, co-authoring papers, or operating a business, similar patterns emerge. A minority of actors represents a large portion of the collaborative activity, well-connected actors tend to work with other well-connected actors, and cooperation is more likely between actors who share a mutual partner.

[1] The network data used in Table 1 were obtained from the Network Data Repository (Rossi and Ahmed, 2015).

2.2. Topological features of criminal networks

There has been a concerted effort to understand the structure of criminal organizations (Sparrow, 1991; Coles, 2001; von Lampe, 2009). Past research demonstrates that skewed degree distributions are also a characteristic feature of criminal networks. Varese (2012) finds a heavy-tailed degree distribution in a Russian Mafia group, while Krebs (2002) and Morselli et al. (2007) show that degree varies considerably in a terrorism network. In addition, Duijn et al. (2014) shows that degree follows a power law distribution in a large co-offending network of cannabis cultivators, and Qin et al. (2005) finds a truncated power law in a terrorism network. In comparable studies on drug trafficking networks, Natarajan (2006) finds that a small number of disproportionately high degree traffickers among a majority with low degree, and Bright et al. (2015) finds a skewed distribution of degree.

However, it is common for criminal network research to focus on identifying key players, rather than classifying network topology. This focus is perhaps unsurprising given the policy implications and substantive interest in understanding the roles of influential criminals. Nevertheless, since key player analyses are less common elsewhere in the social networks literature, many of the empirical findings on criminal networks are not directly comparable to their social network counterparts. Thus, although degree is nearly always part of the analysis, it is often tabulated at the actor-level or truncated to include only highly central actors. When the distribution of degree is visualized, it is rarely classified formally. Similarly, aggregated structural statistics, such as clustering, are less frequently considered.

Most notably, degree mixing is typically overlooked as a property of interest. This may owe to the focus on key players, where statistics such as eigenvector centrality are more often used to account for the connectivity of actors' neighbors. However, there are several reasons to be interested in degree mixing. Firstly, it provides a relatively simple summary of the tendency for actors to collaborate with other members of the network with similar degree, and is therefore indicative of the internal organization of criminal groups. Secondly, it has been proposed that assortative mixing is one of two properties, the other being high levels of clustering, that distinguishes social networks from other kinds of networks (Newman, 2003b). Should there be criminal networks that do not exhibit this property, it may inform our understanding of the underlying mechanics of tie formation in criminal organizations. Thirdly, degree mixing has long been associated with the vulnerability of networks to actor removal (Newman et al., 2002; Holme and Zhao, 2007), which is particularly relevant to criminal networks. Below, we begin to this rectify this omission.

3. Data

A major challenge in studying criminal networks lies in obtaining reliable and informative data. Standard tools like name generators are impractical for criminal networks, while observing clandestine groups is, by definition, not ordinarily an option. However, law enforcement agencies dedicate considerable resources to the surveillance of criminal groups, and parts of the information these agencies collect are often documented in court records. Archival court records have therefore become a substantial source of data on criminal networks (Natarajan, 2006; Morselli and Petit, 2007; Varese, 2012; Campana and Varese, 2012, 2013; Berlusconi, 2013; Campana, 2015).

As we describe below, the materials used in this paper were produced during the trial of a group of actors involved in a heroin trafficking operation. The traffickers, who were active throughout the period 1979–1984, sourced heroin in Turkey, arranged the movement of the narcotics through Europe and into Sicily, from where it was subsequently transported into the United States. The majority of the heroin was sold in large, wholesale quantities to smaller distributors in New York, New Jersey, Philadelphia, and the Midwest. The proceeds from the distribution were laundered primarily through cash deposits in Swiss bank accounts. Here, we describe the document selection and coding process used to recover the actors and edges.

3.1. Document selection

The source document for the network data was selected from a large corpus of documents produced during the trial. This corpus is publicly available through the United States National Archives and Records Administration. The source document is the Government's Sentencing Memorandum (GSM): a 107-page summary of the prosecution's evidence against the defendants produced towards the conclusion of the trial. The evidence is a mix of intercepted telephone conversations, witness testimony, and direct observations conducted by the FBI.

The purpose of selecting one document from the wider corpus is to twofold. First, the corpus is extremely large and almost entirely non-digitized. Coding the entire corpus would therefore involve manually working through hundreds of thousands of pages. Even automating the coding process using, for instance, researcher-generated search terms or natural language processing, would first require digitizing the documents. Selecting a subset of the documents is therefore the only practical way to make the data extraction manageable.

Second, there is no straightforward way to sample the documents that would provide both sufficient coverage and consistency across sources. The indexing of the corpus does not describe the contents of the documents in detail. Many of the documents, particularly in the early stages of the trial, are stenographers minutes that describe various procedural arrangements such as juror selection and contain no information relevant to the network. In order to sample the documents therefore, one would have to draw from the entire corpus either randomly, on the basis of date, or document type. In either case, the documents may have been produced by different parties, such as the defense, prosecution, or the court itself, and would thus be subject to different types of bias or may simply be irrelevant.

The GSM has the advantage of being both concise and consistent, and it is therefore practical to extract a large quantity of relevant data and employ a coding method that is fine-tuned to the document. Finally, any bias in the data has greater tractability insofar as it is reasonable to expect sources of bias to be consistent within-document.

3.2. Actor and relational coding

The coding of the data proceeded in two steps. The first step involved a census of the actors named in the GSM. The GSM refers exclusively to actors involved in trafficking activities and does not refer to witnesses or other individuals related to the case. In total, 82 traffickers were identified in the document.

The second step involves extracting the collaborative edges among the actors. Collaboration is defined as an event in which two or more actors interact with each other for the purpose of a drug trafficking task, which broadly fall into three types: sourcing and arranging the supply of heroin, distributing the heroin, and laundering cash proceeds. Time resolution is unavailable for most collaboration events. Although the GSM reports that some actors were involved in several collaboration events together, as discussed below, weighting the edges may induce bias due to the non-random inclusion of collaboration events in the GSM. Thus, the edges are unweighted. In total, there are 432 edges in the network.

Similar coding methods are used in other studies on criminal networks (Morselli et al., 2007; Natarajan, 2006; Morselli and Petit, 2007; Varese, 2012; Campana and Varese, 2012, 2013; Berlusconi, 2013; Campana, 2015; Papachristos and Smith, 2014) and for other types of social network that are difficult to observe (Butts et al., 2012).

3.3. Limitations and bias

The data derive from a law-enforcement investigation. As such, the coverage of actors and edges depends, in part, on the focus of the investigation and may therefore be affected by sample selection bias. This could manifest in two ways. First, due to the nature of criminal investigations, the sampling of actors may more closely approximate a snowball sample than a random sample. In a criminal investigation, it is typical for a subset of actors to be placed under surveillance whom subsequently lead investigators to other actors as the investigation expands. This may affect the findings in the current study to the extent that the degree of initial targets is inflated relative to later targets of the investigation. Second, the inclusion of collaboration events may be influenced by the extent to which each event supports the prosecution's case against the defendants. It is possible that some collaboration events were omitted, leading to the degree of some actors being underrepresented.

The selected investigation and the coding strategy are designed to mitigate these possible sources of bias. First, a central plank of the investigative strategy was to secure prosecutions through the Racketeer Influenced and Corrupt Organizations (RICO) Act. Under RICO, it is necessary to demonstrate the existence of a long-term structure (Lynch, 1987, p. 703).[2] Accordingly, the investigation sought to collect evidence of connections between members of the trafficking network so that prosecutors could tie these individuals to the broader criminal enterprise. Hence, while early targets of the investigation may initially be placed under greater investigatory focus, the RICO strategy depended on uncovering the collaborative ties of actors who were identified at later stages of the investigation. This focus on identifying ties throughout the investigation should reduce any inflation in the degree of initial targets vis-a-vis latterly identified targets.

Second, we restrict the coding of edges in the GSM to collaboration events that relate directly to drug trafficking activities and we treat the edges as unweighted. Thus, we are concerned only with events that are eligible for inclusion in the GSM, and omitted events would not affect the findings unless that event is the only evidence of a tie between two actors. In principle, removing such events would weaken both the evidence that an enterprise exists and the evidence tying these actors to that enterprise, thus jeopardizing RICO prosecution. Moreover, in order for event omission to substantially affect the shape of the degree distribution or assortativity coefficient, it is necessary that events including particular actors are systematically selected against. It is not clear that this would benefit the prosecution. Nevertheless, we cannot rule out the possibility that collaboration events involving actors viewed as a priority for prosecution are selected for, which could affect the shape of the degree distribution. Unfortunately, non-random sampling is a feature data from criminal investigations, which remain one of the few detailed sources of information on criminal networks.

4. Methods

The analysis is formed of two parts. First, we examine the degree distribution, clustering, and degree mixing in the network. Second, we analyze the vulnerability of the network to vertex removals (i) in its observed state, and (ii) under several degree mixing configurations.

4.1. Structural analysis

4.1.1. Classifying the degree distribution

In order to classify the degree distribution in the trafficking network, we fit power law, log-normal, exponential, and Poisson models to the distribution using maximum-likelihood methods (Clauset et al., 2009). The fits proceed as follows. Each observed vertex degree, k, in the trafficking network represents a candidate threshold value, k_{min}, above which the scaling behavior associated with each model may plausibly begin. Then, for each model, we find the associated parameter values which maximize the fit of that model to the degree distribution above each candidate threshold value. We then select the power law, log-normal, exponential, and Poisson model with the best overall fit, given its corresponding threshold and parameter values.

To test whether each model is plausible given the data, we synthesize 2500 degree distributions from a "true" version of each model with the threshold and parameter values equal to those estimated for the observed degree distribution. We then refit the models, independently, to each of the synthetic degree distributions. The models are ruled out if 10% (i.e. $p < 0.10$) or fewer of the fits to the synthetic data are poorer than the best fit to the real data.[3] Finally, we compare the log-likelihood ratio of the degree distribution, given each model, in pairwise fashion. We thus arrive at a classification of the degree distribution by considering several competing models.

4.1.2. Clustering

There are several measures of clustering in networks. Here, we use the clustering coefficient, c, defined by Watts and Strogatz (1998). In short, the coefficient is based on the ratio of the edges that exist between a vertex and its neighbors to all of the possible edges that could exist. Averaging across all vertices gives the coefficient, which varies between 0 and 1, and captures the average local connectivity or "cliquishness" (Watts and Strogatz, 1998, p. 440) of the network.

4.1.3. Degree mixing

The final property we examine is the vertex–vertex degree correlation. A network is assortatively mixed when this correlation is positive, i.e. high degree vertices are connected to other high degree vertices, and dissortative when the correlation is negative. Here, I calculate the assortativity coefficient, r, defined in Newman (2003a). The intuition behind this coefficient is that, denoting the degree distribution of the trafficking network p_k, we select an edge and then take the degree of one of the vertices, i, incident to that edge. We then calculate the excess degree of i, which is the degree of i relative to our expectation given p_k. If the excess degree of i is frequently large as we move from one edge to the next, then a relatively large number of edges must end at a high degree vertex. Conversely, if the excess degree is often negative, then edges frequently terminate at low degree vertices. Such networks are assortative $r \rightarrow 1$ and dissortative $r \rightarrow -1$, respectively. A network is neither assortative nor dissortative when $r = 0$.

[2] See Lynch (1987) for a review of the RICO act.

[3] Further details of this procedure are outlined by Clauset et al. (2009) and Virkar and Clauset (2014) for continuous and binned data, respectively. For the implementation of this process in R, see Gillespie (2015).

4.2. Vulnerability analysis

4.2.1. Vertex removal

For the second part of the analysis, we estimate the impact of targeted vertex removals on the cohesion and connectivity of the trafficking network. We implement a simple degree-based removal process, whereby traffickers are removed in descending order of degree (Albert et al., 2004). Three metrics are used to assess the impact of each vertex removal: (1) the number of components, (2) the size of the largest component, and (3) the number of communities. The first two metrics are straightforward, capturing the number of groups that are split-off entirely from the largest connected group of traffickers and the extent to which each split reduces the size of this largest component, respectively.

In order to capture the impact of vertex removals beyond "clean" breaks in the components, the third metric is based on community dissolution. To identify the communities, we use a community detection algorithm based on random walks. Each walk begins at a random vertex and takes a step to a randomly selected adjacent vertex, with the maximum number of steps per walk set to four. After multiple iterations, patterns in the sequences of visited vertices are used to classify each vertex into a community. Put simply, the more often the walks are trapped in certain regions of the network, the more likely those regions are classified as communities (Pons and Latapy, 2006). There may be many candidate classifications, so we select the classification that maximizes a quality function known as modularity (Newman, 2004). Thus, we capture community dissolution in addition to fractures in the network.

4.2.2. Graph sampling

To evaluate how degree mixing influences the impact of targeted vertex removals on the trafficking network, we generate synthetic variants of the network at several benchmark levels of assortativity, r_b. Following Newman and Park (2003), the synthetic networks are produced by swapping edges while preserving the degree distribution. Our algorithm can be described as follows. First, we calculate the assortativity coefficient, r, and clustering coefficient, c, for the trafficking network. Next, two randomly selected edges, (i_1, j_1) and (i_2, j_2), are toggled off. We propose two randomly selected replacement edges, (i_3, j_3) and (i_4, j_4), with the restriction that the degree sequence, or k-sequence, prior to the toggle must be preserved exactly under the proposed edges. Updated assortativity and clustering coefficients, r' and c', are then calculated under the proposed edges. The proposed edges are accepted if $r' \to r_b$ and $c' \geq c$. This is designed to preserve the degree distribution and clustering of the observed trafficking network, while reshuffling each synthetic network into a more assortative or dissortative state, as desired. This process is iterated until the desired level of assortativity is reached. We generate 1000 synthetic networks at each r_b (Fig. 1).

Of course, the $c' \geq c$ condition means that clustering can increase, but not decrease, in each iteration. This is due to limits on the assortativity coefficient, r_{min} and r_{max}, imposed by c.[4] Essentially, the problem is that as $c \to 1$ or $c \to 0$, r_{min} and r_{max} converge. Thus, it is not feasible to reshuffle networks with large clustering coefficients into highly assortative or dissortative states. Since the observed trafficking network has a large clustering coefficient, c tends downwards as we reshuffle towards such states. The $c' \geq c$ condition attempts to mitigate this downward trend, unless reducing c is necessary for reaching r_b. More precisely, we accept a proposed reshuffle if $r' \to r_b$ but $c' < c$ with probability 0.1. This acceptance rate strikes a balance between the computational burden of finding edge reshuffles which satisfy the stringent $c' \geq c$, on the one hand, and allowing c to decrease as required in order to reach r_b, on the other. We return to the reshuffling of r and c in Section 5.2.

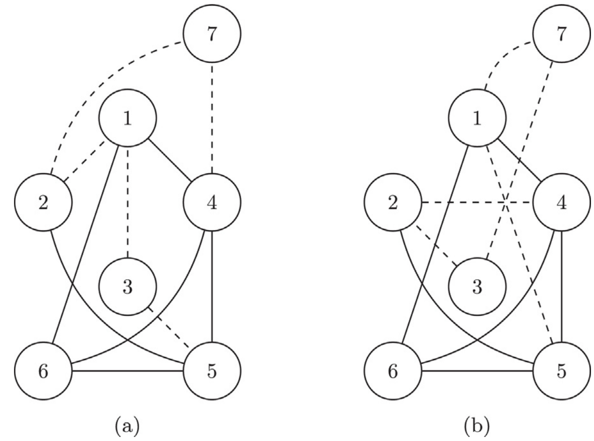

Fig. 1. Example of graph sampling. (a) Network prior to reshuffling. The dashed edges are ultimately toggled off. (B) The same network after reshuffling. The dashed edges have been toggled on. Note that each vertex has equal degree in both configurations, but the assortativity coefficient has increased.

5. Results

5.1. Structural properties

We begin by classifying the degree distribution in the trafficking network. Fig. 2 shows the power law, log-normal, exponential, and Poisson fits to the cumulative distribution function (CDF) of degree, k. As is clear in the CDF, the degree distribution is heavy-tailed, such that most traffickers have fewer than 10 collaborators while the top tenth of the traffickers fall in the range $25 < k < 50$.

Table 2 summarizes the parameter estimates, goodness-of-fit, and results of the hypothesis tests for each fit. The p-value exceeds 0.1 in all four models, thus the hypothesis test fails to rule out any of the competing classifications. However, n is small in both the power law and Poisson fits due to the selection of a large k_{min} threshold. The parameter estimates are therefore unreliable, while the hypothesis test has little power to discriminate against either fit.[5] Consequently, while a power law or Poisson model cannot be ruled out, neither represents a parsimonious classification of the degree distribution.

Table 3 shows the results of a direct comparison of the log-likelihood of the degree distribution for each model. Where the optimal k_{min} threshold differs between two models, we select the larger of the two, therefore "rewarding" the model which describes a larger section of the distribution. The log-likelihood ratio indicates that the log-normal and exponential models provide considerably stronger classifications of the degree distribution than the power law model. In addition, Fig. 2 shows that both the log-normal and exponential models are optimized at the lowest possible threshold value, $k_{min} = 1$, thus providing a parsimonious description of the degree distribution in the network as a whole.

Regarding whether the log-normal or exponential model offers the more compelling classification, the direct comparison favors neither. Although the sign of the log-likelihood ratio indicates that the exponential model is a better fit, the magnitude of the ratio does not decisively favor either model. However, one small advantage of the exponential model is that it rests on one less parameter. Overall,

[4] See Holme and Zhao (2007) for further details on the relationship between assortativity and clustering.

[5] Clauset et al. (2009, p. 669) recommend $n > 50$ for obtaining reliable parameter estimates.

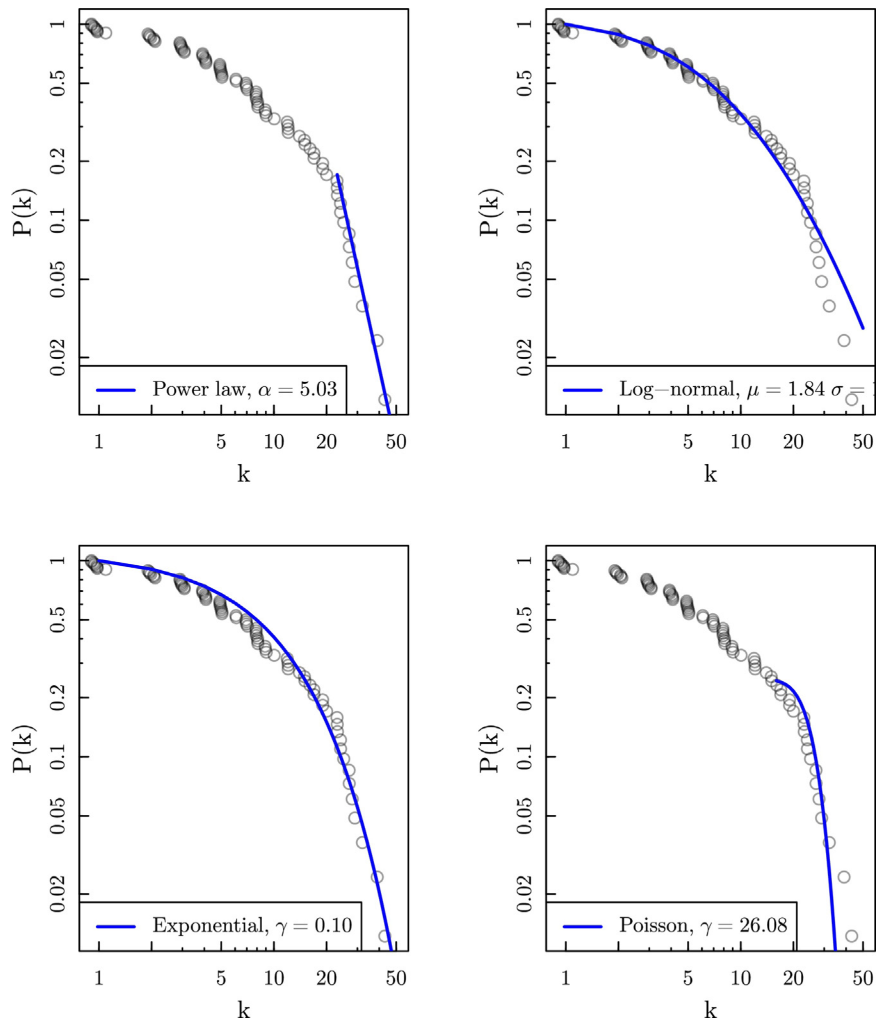

Fig. 2. Top-left: Power law fit to the cumulative distribution function (CDF) of degree in the trafficking network. Top-right: Log-normal fit. Bottom-left: Exponential fit. Bottom-right: Poisson fit.

Table 2
Parameters for the competing models of the degree distribution. The k_{min} represents the threshold above which the fit applies to the distribution, where n is the number of traffickers above that threshold. The goodness-of-fit (GOF) is the Kolmogorov–Smirnov statistic. The p-value represents the proportion of fits to synthetic data drawn from a true version of each model that are poorer than the fit to the real data. A fit is ruled out when $p < 0.10$.

Model	n	k_{min}	α	γ	μ	σ	GOF	p
Power law	14	23	5.03				0.099	0.751
Log-Normal	82	1			1.84	1.07	0.053	0.915
Exponential	82	1		0.10			0.517	0.489
Poisson	20	16		26.08			0.177	0.628

the exponential decay in degree is consistent with the heavy-tailed degree distribution observed in other collaboration networks.

The exponential classification implies a probability distribution on the formation of collaborative ties in the network. For an actor with degree k, the probability of gaining y collaborators is given by $P(y) = \gamma^{\frac{y}{k}}$. If the network were to continue operating, traffickers with larger degree would therefore be expected to increase their degree at a faster rate than poorly connected traffickers, as in a preferential attachment process.

Turning to clustering, the trafficking network exhibits a large local clustering coefficient $C = 0.801$. This reflects the large proportion of complete triads in the network in comparison to two-paths and triads containing only one edge. The clustering coefficient also implies that around 80% of connected trafficker pairs share at least one mutual collaborator. This finding is consistent with other collaboration networks where it has been observed that actors have a strong tendency to collaborate with the partners of their partners (Newman, 2001a), as noted in Section 2. However, this tendency appears to be particularly strong in the trafficking network, indicating that the network is highly cliquish and suggesting that mutual partners may play an important role in brokering new collaborative relationships.

Table 3
Comparison of alternative models of the degree distribution. A positive log-likelihood ratio indicates that model 1 provides a better fit to the data than model 2. Neither model is preferred when $p > 0.1$.

Model 1	Model 2	k_{min}	LLR	p
Log-Normal	Power law	1	5.57	<0.001
Exponential	Power law	1	5.45	<0.001
Poisson	Power law	23	−1.63	0.101
Exponential	Log-Normal	1	0.58	0.560

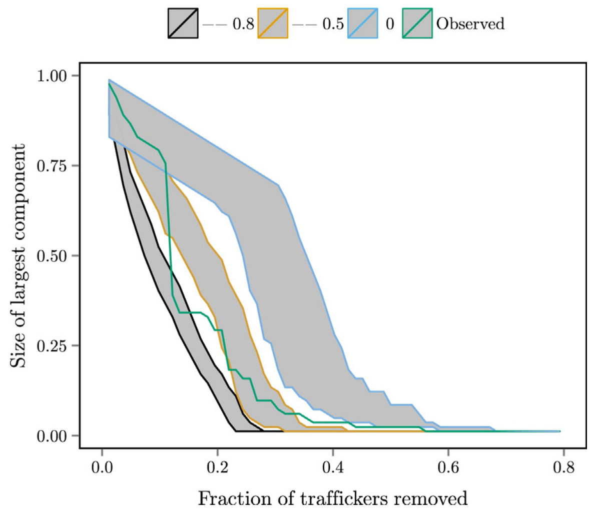

Fig. 3. Size of the largest component, S, as a function of the fraction of vertices removed, f. Each band represents a reshuffled version of the trafficking network to a specific average level of degree dissortativity, where the degree distribution of the observed network is preserved exactly. The green line represents the observed trafficking network. Since the rewiring process is stochastic, the bands cover the range of values of S given f in 1000 rewired networks.

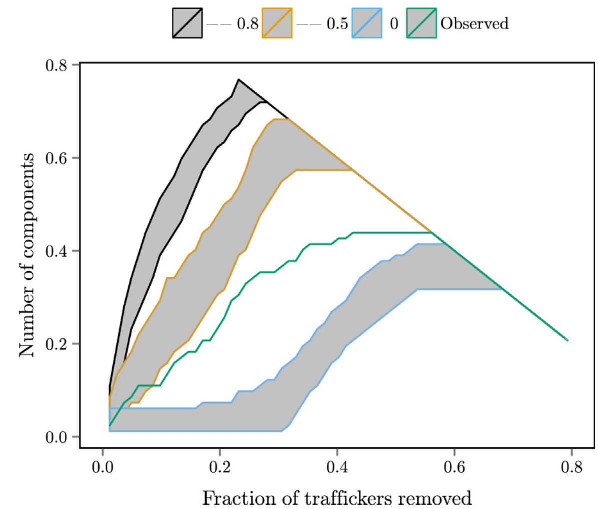

Fig. 4. The number of components, P, as a proportion of the maximum possible number, in response to the fraction of vertices removed, f. The vertices are removed in order of degree, from highest to lowest.

Despite the characteristic heavy-tailed degree distribution and strong clustering, the mixing pattern in the trafficking network is dissortative rather than assortative. The assortativity coefficient is r=–0.192 with a standard error of $\sigma_r = 0.052$, which we estimated via the bootstrapping method detailed in Newman (2002), and therefore the dissortativity is a 3.6σ result and statistically significant. Thus, while ordinarily in collaboration networks large degree actors are connected to other high degree actors and low with low, in the trafficking network those with many partners often collaborate directly with their poorly-connected peers.

5.2. Vertex removal

Having established the properties of the trafficking network, we now turn to the results of vertex removal on the topology and cohesiveness of the network. Fig. 3 shows the size, S, of the largest component in the observed trafficking network and in the reshuffled networks as a function of the proportion of traffickers removed, f, where the traffickers are removed in order of their degree. Each band covers the range of values of S given f across 1,000 synthetic networks, which each have an identical degree distribution to the observed network but have been vary in their average levels of degree dissortativity, $r_b = \{0, -0.5, -0.8\}$.

Varying the level of dissortativity changes both the shape and rate at which S falls as a function of f. For example, removing $f = 0.1$ of the highest-degree vertices in the network reduces the size of the largest component to $S = 0.79$ of its original size. This compares to $S = \{0.86, 0.70, 0.46\}$ for $r_k = \{0, -0.5, -0.8\}$, respectively. The results of vertex removal in the observed network overlaps with the r_k=–0.5 synthetic networks for small values and large values of f, but S decreases faster in the middle range of f. The reduction in S is slower in the $r_k = 0$ networks across the range of f. Conversely, in the most dissortative case, the size of the largest component steeply declines in response to high-degree removals; removing 13% of the traffickers reduces the size of the largest component to approximately half of its original size.

Fig. 4 presents the number of components, P, as vertices are removed in order or degree. Again, P increases faster in the trafficking network than in the reshuffled networks with benchmark assortativity $r_b = 0$. Indeed, removing the top $f = 0.1$ traffickers from the $r_b = 0$ networks leads to an average split of 2.3 components, while removing the same fraction from the observed network leads to a split into 9 components. Thus, the dissortative mixing in the trafficking network appears to be associated with an increased vulnerability to targeted removals, insofar as the network fractures into a larger number of distinct components across all values of f. The number of components increases faster still as reshuffled to more dissortative states: removing $f = 0.1$ vertices in the $r_b = 0.8$ networks leads to a split of 37 components, on average.

Finally, Fig. 5 shows the change in the number of communities, U, as vertices are removed in degree order. As above, varying the dissortativity of the degree mixing has a profound impact on the results of vertex removal. When $r_k = 0$, U remains relatively stable up to $f > 0.3$. However, in the observed network, U begins to increase as $f > 0$, albeit non-monotonically. In the more dissortative networks, the increase in U is faster still, but the modularity scores remain generally lower, indicating that the classification of the communities is not as clear cut. One explanation for this outcome is that, since the cohesiveness of the network is so radically reduced by targeted vertex removals in the most dissortative networks, the number of reasonable classifications of the remaining traffickers into communities remains relatively large, and ultimately approaches the point where all vertices are isolates at considerably smaller values of f. Thus, in addition to the faster reduction in cohesiveness in the more dissortative variants of the observed network, the communities that remain also seem to be less tight-knit.

It is important to note that, while the k-sequence is preserved during reshuffling, assortativity is not the only property that we manipulate. As discussed in the Section 4.2.2, the level of clustering in a network constraints the maximum and minimum assortativity that can be obtained through reshuffling. At the extreme, placing an exact constraint on c throughout reshuffling, such that edge proposals are accepted only if they maintain c precisely, places a stringent upper bound on the number of potentially acceptable proposals. Thus, we allowed c to increase or decrease, as required, in order to reach the benchmark assortativity coefficients above.

Fig. 6 shows the change in c during the graph sampling procedure used to reshuffle the observed network into the most dissortative configuration, $r_b = -0.8$. Note that this represents only one out of 1000 sampling iterations. To reach $r_b = -0.8$, the clustering is reduced to approximately $c = 0.2$ in this particular network. Although the precise change in c varies across sampling iterations due to the stochastic nature of the edge toggling process, this reduction is fairly typical. The rate of reduction in c per edge toggle

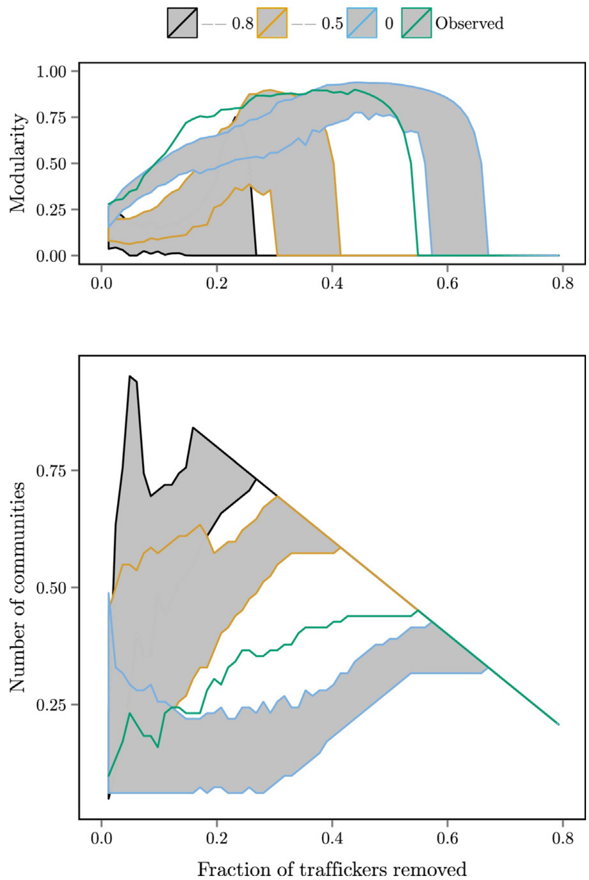

Fig. 5. Bottom: Change in the number of communities, *U*, as a proportion of the maximum possible number, in response to the fraction of vertices removed, *f*. Top: The modularity score associated with the classification of communities in the bottom panel. The vertices are removed in order of degree, from highest to lowest.

("iterations" in Fig. 6) is considerably higher during early iterations of the sampling algorithm. Almost half of the total decrease in *r* is achieved using only 1000 toggle iterations. However, around 4000 toggles were required for the reduction $r = -0.7$ to $r = -0.8$. This is indicative of the small set of graph configurations that exist for

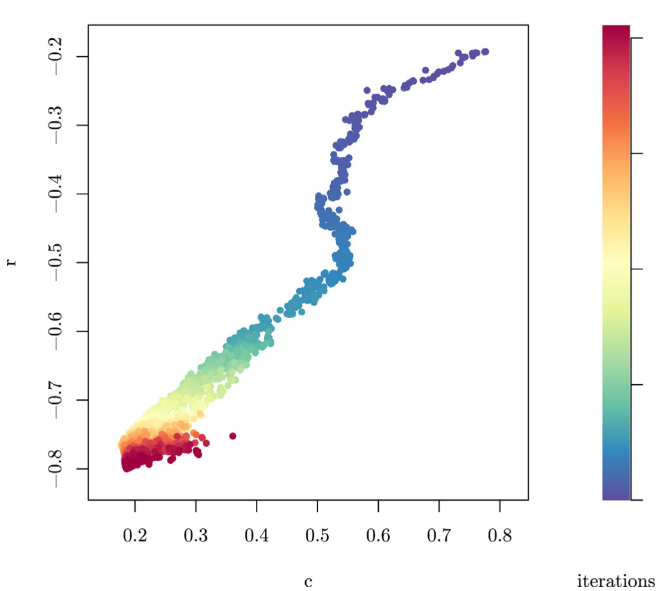

Fig. 6. The shift in local clustering, *c*, and assortativity, *r*, in the graph sampling algorithm used to reshuffle the network into a more dissortative configuration.

satisfy extreme *r* values given our constraint on the *k*-sequence, but it perhaps also suggests that our algorithm could be improved by fine-tuning the acceptance rate to be more stringent for reductions in *c* when *r* has not yet reached extreme values.

Due to the decrease in *c*, the increasing vulnerability of the network to targeted vertex removals in more dissortative states cannot be put down entirely to the degree mixing configuration of the network. Indeed, there is a complex dependence between assortativity, clustering, and vulnerability (Holme and Zhao, 2007), examination of which goes beyond the scope of this paper.

Nevertheless, there are intuitive reasons to suspect that dissortativity is driving much of the vulnerability of the trafficking network. In a dissortative network, there is a greater tendency for high and low degree actors to be connected. Removing a small number of high-degree traffickers can therefore disconnect a large number of low degree traffickers. Indeed, the unusually large clustering coefficient in the trafficking network, which indicates that most traffickers are members of complete triadic structures, may be a critical property for protecting the network from further fragmentation. Moreover, dissortative networks exhibit a different type of vulnerability not examined here. Due to the mixing of high and low degree actors, the detection of any low degree trafficker, of which there are many due to the heavy-tailed degree distribution, may lead investigators directly to highly central, more positionally important traffickers. As such, dissortative mixing places the entire network at greater risk of exposure. Indeed, this may have contributed to the eventual failure of the drug trafficking network.

6. Conclusion

It has long been proposed that criminal networks exhibit abnormal topological properties due to the conditions under which criminals operate. This paper investigates whether a collaboration network among heroin traffickers has a heavy-tailed distribution of degree, strong clustering, and assortative degree mixing – three characteristic properties of collaboration networks outside of the criminal domain. The trafficking network exhibits exponential decay in degree and strong clustering, but dissortative degree mixing, such that high-degree traffickers collaborate preferentially with low-degree traffickers, and vice versa.

The dissortative mixing, in addition to being unusual in a collaboration network and indeed in social networks more generally (Newman and Park, 2003), seems to increase the vulnerability of the network to degree-targeted vertex removals. Indeed, reshuffling the network into more dissortative states increased the impact of such removals, although part of this effect is likely explained by the decrease in clustering that accompanies highly dissortative states. However, degree mixing has received relatively little attention in the empirical literature on criminal networks, despite the fact it can be calculated, in most cases, without additional data. Since degree mixing has demonstrable consequences for the vulnerability of the network to targeted attacks, a growing area of research interest (Morselli, 2010; Malm and Bichler, 2011; Duijn et al., 2014), it is hoped this study encourages researchers to include degree mixing in future analyses of criminal networks.

To the extent that dissortativity is observed in other criminal networks, our findings raise several questions that were not explored here, particularly regarding the tie formation processes that underpin dissortative mixing, but also regarding the possible upsides of dissortativity for the coordination of criminal activity. One plausible upside lies in the capacity that dissortative networks confer for well-connected criminals to monitor and coordinate peripheral actors directly, for example.

The study is not without limitations. First, although our data source is internally consistent in its reporting of actors and collaboration events, the possibility remains that the data capture an unobserved nonrandom selection process. If the proclivity to include edges in the source document is influenced by which actors are incident to those edges, the assortativity coefficient could be affected. It would be overestimated if, for example, prosecutors wished to link peripheral members of the network with highly active members. On the other hand, it would be underestimated if peripheral members truly have a tendency to collaborate with high-degree actors, but are also less likely to be included in the sample. Second, while I preserve the degree distribution, rewiring the network to different levels of dissortativity effects other properties in the network such as clustering. Hence the results of vertex removal on the rewired networks do not derive entirely from degree mixing.

However, our results are consistent with previous research that demonstrates that dissortative networks are more vulnerable to targeted attacks than assortative networks. Further research is required to evaluate whether dissortativity is a common property of collaboration among criminals.

Acknowledgements

The author thanks Federico Varese, Christiaan Monden, and three anonymous reviewers for their comments and suggestions.

References

Albert, R., Albert, I., Nakarado, G.L., 2004. Structural vulnerability of the north American power grid. Phys. Rev. E 69, 025103.

Albert, R., Jeong, H., Barabási, A.-L., 1999. Diameter of the world-wide web. Nature 401, 130–131.

Ayling, J., 2009. Criminal organizations and reslience. Int. J. Law Crime Justice 37, 182–196.

Baker, W.E., Faulkner, R.R., 1993. The social organization of conspiracy: illegal networks in the heavy electrical equipment industry. Am. Sociol. Rev. 58, 837–860.

Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.

Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T., 2002. Evolution of the social network of scientific collaborations. Physica A 311, 590–614.

Berlusconi, G., 2013. Do all the pieces metter? Assessing the reliability of law enforcement data sources for the network analysis of wire taps. Global Crime 14, 61–81.

Bright, D.A., Greenhill, C., Reynolds, M., Ritter, A., Morselli, C., 2015. The use of actor-level attributes to identify key actors: a case study of an Australian drug trafficking network. J. Contemp. Crim. Just. 31, 262–278.

Butts, C.T., Acton, R.M., Marcum, C.S., 2012. Interorganizational collaboration in the hurricane Katrina response. J. Soc. Struct. 13, 1–36.

Campana, P., 2015. The structure of human trafficking: lifting the bonnet on a Nigerian trafficking network. Br. J. Ciminol. Adv. Access.

Campana, P., Varese, F., 2012. Listening to the wire: criteria and techniques for the quantitative analysis of phone intercepts. Trends Organ. Crime 15, 13–30.

Campana, P., Varese, F., 2013. Kinship and violence as credible commitments in criminal organisations: an empirical test. Ration. Soc. 25, 263–289.

Chang, H., Si, B.-B., Zhou, Y.-P., He, D.-R., 2007. Assortativity and act degree distribution of some collaboration networks. Physica A 383, 687–702.

Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. SIAM Rev. 51, 661–703.

Coles, N., 2001. It's not what you know – it's who you know that counts: analysing serious crime groups as social networks. Br. J. Criminol. 41, 580–594.

Crossley, N., Edwards, G., Stevenson, R., 2012. Covert social movement networks and the secrecy-efficiency trade off: the case of the UK suffragettes (1906–1914). Soc. Netw. 34, 634–644.

Duijn, P.A., Kashirin, V., Sloot, P.M., 2014. The relative ineffectiveness of criminal network disruption. Sci. Rep. 4.

Gay, B., Dousset, B., 2005. Innovation and network structural dynamics: study of the alliance network of a major sector of the biotechnology industry. Res. Policy 34, 1457–1475.

Gillespie, C.S., 2015. Fitting heavy tailed distributions: the powerlaw package. J. Stat. Softw. 64.

Gleiser, P., Danon, L., 2003. Community structure in jazz. Adv. Complex Syst. 6, 565–573.

Guimerá, R., Uzzi, B., Spiro, J., Amaral, L.A.N., 2005. Team assembly mechanisms determine collaboration network structure and team performance. Science 308, 697–702.

Holme, P., Zhao, J., 2007. Exploring the assortativity-clustering space of a network's degree sequence. Phys. Rev. E, 75.

Huang, J., Zhuang, Z., Li, J., Giles, C.L., 2008. Collaboration over time: characterizing and modeling network evolution. In: International Conference on Web Search and Data Mining (WSDM), pp. 107–116.

Kenney, M., 2007. The architecture of drug trafficking: network forms of organisation in the Colombian cocaine trade. Global Crime 8, 233–259.

Krebs, V.E., 2002. Mapping networks of terrorist cells. Connections 24, 43–52.

von Lampe, K., 2009. The study of organised crime: an assessment of the state of affairs. In: Organised crime: norms, markets, regulation and research. Unipub, Oslo, pp. 165–211.

Lindelauf, R., Borm, P., Hamers, H., 2009. The influence of secrecy on the communication structure of covert networks. Soc. Netw. 31, 126–137.

Lynch, G.E., 1987. Rico: the crime of being a criminal, Parts I & II. Columbia Law Rev. 87, 661–764.

Malm, A., Bichler, G., 2011. Networks of collaborating criminals: assessing the structural vulnerability of drug markets. J. Res. Crime Delinq. 48, 271–297.

McCormick, G.H., Owen, G., 2000. Security and coordination in a clandestine organization. Math. Comput. Model. 31, 175–192.

Moody, J., 2004. The structure of a social science collaboration network: disciplinary cohesion from 1963 to 1999. Am. Sociol. Rev. 69, 213–238.

Morselli, C., 2010. Assessing vulnerable and strategic positions in a criminal network. J. Contemp. Crim. Just. 26, 382–392.

Morselli, C., Giguére, C., Katia, P., 2007. The efficiency/security trade-off in criminal networks. Soc. Netw. 29, 143–153.

Morselli, C., Petit, K., 2007. Law-enforcement disruption of a drug importation network. Global Crime 8, 109–130.

Natarajan, M., 2006. Understanding the structure of a large heroin distribution network: a quantitative analysis of qualitative data. J. Quant. Criminol. 22, 171–192.

Newman, M.E.J., 2001a. Clustering and preferential attachment in growing networks. Phys. Rev. E 64.

Newman, M.E.J., 2001b. The structure of scientific collaboration networks. PNAS 98, 404–409.

Newman, M.E.J., 2002. Assortative mixing in networks. Phys. Rev. Lett. 89, 208701.

Newman, M.E.J., 2003a. Mixing patterns in networks. Phys. Rev. E 67, 026126.

Newman, M.E.J., 2003b. The structure and function of complex networks. SIAM Rev. 45, 167–256.

Newman, M.E.J., 2004. Fast algorithm for detecting community structure in networks. Phys. Rev. E 69, 066133.

Newman, M.E.J., 2005. Power laws, Pareto distributions and Zipf's law. Contemp. Phys. 46, 323–351.

Newman, M.E.J., Forrest, S., Balthrop, J., 2002. Email networks and the spread of computer viruses. Phys. Rev. E, 66.

Newman, M.E.J., Park, J., 2003. Why social networks are different from other types of networks. Phys. Rev. E 68, 036122.

Papachristos, A.V., Smith, C.M., 2014. Crime and networks. In: The Small World of Al Capone: The Embedded and Multiplex Nature of Organized Crime. Routledge, New York.

Pons, P., Latapy, M., 2006. Computing communities in large networks using random walks. J. Graph Algorithms Appl. 10, 191–218.

Qin, J., Xu, J., Hu, D., Sageman, M., Chen, H., 2005. Intelligence and security informatics. In: Analyzing Terrorist Networks: A Case Study of the Global Salafi Jihad Network. Springer, Berlin, Heidelberg, pp. 287–304, vol. 3495 of *Lecture Notes in Computer Science*.

Ramasco, J.J., Dorogovtsev, S.N., Pastor-Satorras, R., 2004. Self-organization of collaboration networks. Phys. Rev. E 70.

Ramasco, J.J., Morris, S.A., 2006. Social inertia in collaboration networks. Phys. Rev. E 73, 016122.

Rossi, R.A., Ahmed, N.K., 2015. The network data repository with interactive graph analytics and visualization.

Sparrow, M.K., 1991. The application of network analysis to criminal intelligence: an assessment of the prospects. Soc. Netw. 13, 251–274.

Varese, F., 2012. The Structure and the Content of Criminal Connections: The Russian Mafia in Italy. Eur. Sociol. Rev., Adv. AC.

Virkar, Y., Clauset, A., 2014. Power-law distributions in binned empirical data. Ann. Appl. Stat. 8, 89–119.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small world' networks. Nature 393, 440–442.

Williams, P., 2001. Transnational criminal networks. In: Arquilla, J., Ronfeldt, D.F. (Eds.), Networks and Netwars: The Future of Terror, Crime, and Militancy. Rand Coorporation, Washington, D.C., pp. 61–97.