Haas and Ferreira *Security Informatics* (2015) 4:2
DOI 10.1186/s13388-015-0018-8

Security Informatics
a SpringerOpen Journal

RESEARCH Open Access

Federated databases and actionable intelligence: using social network analysis to disrupt transnational wildlife trafficking criminal networks

Timothy C Haas[1*] and Sam M Ferreira[2]

Abstract

Wildlife trafficking, a focus of organized transnational crime syndicates, is a threat to biodiversity. Such crime networks span beyond protected areas holding strongholds of species of interest such as African rhinos. Such networks extend over several countries and hence beyond the jurisdiction of any one law enforcement authority. We show how a federated database can overcome disjoint information kept in different databases. We also show how social network analyses can provide law enforcers with targeted responses that maximally disrupt a criminal network. We introduce an actionable intelligence report using social network measures that identifies key players and predicts player succession. Using a rhino case study we illustrate how such a report can be used to optimize enforcement operations.

Keywords: Rhino poaching; Criminal network disruption; Social network analysis; Federated databases; Actionable intelligence

Introduction

Wildlife trafficking is a key threat to biodiversity [1] and often associates with illicit arms deals [2], trafficking of people [3] and drug smuggling [4]. Recent links to terrorism [5] and well known links to conflict zones [6] serve to highlight the degradation of human society caused by illegal wildlife dealers.

Illegal trading in wildlife products is attractive to criminals because of the potential financial gains with relative little costs or risks [7]. Wildlife commodity prices depend on economic processes such as demand and supply [8]. Rare species by nature will have high value of associated commodities, sometimes accentuated by international trade-bans imposed by CITES [9].

Managing the threat associated with wildlife trafficking to biodiversity and human society alike collapses to three strategic actions [10]. First, protect wildlife assets through anti-poaching largely funded by non-government organizations trading in extinction anxiety as a commodity [11]. Protecting wildlife may ultimately resort to aggressive tactics such as shoot-on-sight policies [12]. The second associates with demand reduction strategies [13] also funded by the extinction anxiety trade. The final action associates with providing the desired commodity [14], usually in the form of a legal trade (*e.g.* [15,16]) that often elicits contentious debate and response inertia [9].

The nature of transnational crime syndicates [17] however, influences all three strategic actions. Poachers conflict with anti-poaching units; end-user suppliers come in conflict with demand reductionists particularly if campaigns are culturally insensitive; and syndicate supply chains compete with potential legal suppliers. For instance, when suppliers like transnational crime syndicates, have lower costs compared to legal producers and suppliers, imperfect economic competition results [7,18]. This nearly always leads to illegal suppliers outcompeting legal ones [19]. The disruption and even collapse of transnational crime may thus serve a key strategic contribution to the success of any of the potential actions available.

This realization increasingly imposes a need of knowledge on wildlife managers required to participate in the policing of wildlife crimes [20]. Wildlife trafficking criminal networks typically organize themselves into a hierarchical structure supporting the three basic tiers of a

* Correspondence: haas@uwm.edu
[1] Lubar School of Business, University of Wisconsin-Milwaukee Milwaukee, Wisconsin, USA
Full list of author information is available at the end of the article

normal economic process linking supply and demand (*e.g.* [21]) – producers (*e.g.* poachers and associated middle men or organizers), supply chains (*e.g.* shipping routes facilitated by exporters) and consumer retailers (*e.g.* sellers in markets supplying end-users).

Wildlife trafficking criminal networks often extend over several countries [22]. Transnational criminal network players and operations thus extend beyond the jurisdiction of a single law enforcement authority. Investigators within different jurisdictions need evidence gathered by investigators in other jurisdictions to define the full extent of a network's operations. In addition, investigators need to identify the roles particular players have in a transnational criminal network. Removing important role players destabilizes the resilience of wildlife crime networks [23] and may ultimately collapse the entire illegal supply chain.

Two particular challenges arise. First, the lack of trust, and absence of a secure, common and user-friendly protocol for transferring criminal intelligence data between jurisdictions can impede the sharing of evidence between investigators of different jurisdictions. We illustrate the use of a federated database approach [24] to address this challenge.

The second challenge associates with targeted responses that could have maximum disruptive effect on a crime network [25]. Authorities seldom have resources to investigate and target all players in large and complex transnational criminal networks [17-22]. They simply cannot arrest everybody. Law enforcers, however, can use evidence contained in a federated criminal intelligence database to derive actionable intelligence [26] on targeted individuals directed at providing law enforcement responses that disrupt the criminal network's operations.

We propose a specific step-by-step program that a heterogeneous group of law enforcement and wildlife conservation agencies should follow to create actionable intelligence. These steps are:

1. Invite into the group agencies whose jurisdiction contains suspected elements of a wildlife trafficking operation.
2. Invite all group members to collectively form and maintain a federated database of all criminal evidence items that each federation member collects within their respective jurisdictions.
3. Employ either a within-federation specialist or an outside consultant to perform social network analysis on queries from the federated database and prepare from those analyses actionable intelligence reports to be shared with all federation members.

To walk the reader through this program, the remainder of this article is structured as follows. First, we describe how such a group of agencies should form a federated database of all wildlife trafficker evidence items that they collect. Next, we illustrate the application of social network analyses [27] that make use of several forms of communication links between players in the federated database to identify key players and arrest sequences as an example of actionable intelligence. Then, we illustrate the creation of actionable intelligence through a case study on curbing rhino poaching in Kruger National Park, a stronghold for two rhino species threatened by illegal wildlife trafficking [14,28]. We reach conclusions in our final section.

Contextual setting

Wildlife trafficking syndicates have common generalized structures [17-22]. Within the illegal supply network, a first-tier intermediary (*i.e.* a receiver or courier) commissions a poaching hunting party and buys whatever wildlife commodities the party is able to poach [29]. Second-tier intermediaries (*i.e.* buyers) buy wildlife products from first-tier intermediaries and sell it to third-tier intermediaries (*i.e.* exporters). Exporters are the links in the supply chain connecting producers with consumers in end-user countries. Several intermediary tiers of vendors can also exist. For this article, we focus on a network composed of middlemen and the poaching parties that they sponsor – it's the basis of the supply part of a wildlife trafficking criminal network and contains four groups. Any member within any such group belonging to the network is referred to as a *player*.

Federated databases

Players at higher tiers in the network receive wildlife commodities from several different sub-networks of poachers and lower-tiered intermediaries [29]. Crime intelligence seldom originates from a single source [30,31] and hence also rarely does a single database capture all information. Combining data on these sub-networks into one database can reveal increasingly higher-tiered players (typically those in control of funds and with export capabilities), and the true extent of the wildlife trafficking network [17-22]. On the other hand, if investigators analyze sub-networks in isolation from each other, they may miss-identify high-level players as peripheral players. High-level players, for instance, often communicate with only a few members of a particular sub-network increasing the risk of pseudo-peripheral identification [32]. Accentuated risk arises when high-level players dictate poaching raids in widely geographically separated parts of a wildlife-hosting region. The use of a federated database allows investigators to discover long-distance links.

In this context a federated database [24] comprises the development of a virtual database of all evidence gathered on persons suspected of participating in wildlife trafficking (Table 1). This database holds data from

Table 1 Procedural steps required to establish a federated database

Step	Procedure
1	Separate investigation groups agree to a federated database approach to share information and become a federation member.
2	All federation members agree to include in their constituent databases a common core of attribute fields with agreed-upon names.
3	Each federation member creates a secure, local database that conforms to the Structured Query Language (SQL) for all evidence items that they directly collect using compatible software that constructs and processes this database *i.e.* the database engine.
4	Each federation member acquires a criminal intelligence software system. Commercial systems include Analyst's Notebook [37]. One free system is **id** [60]. Pajek, a free social network analysis program, can compute some measures typically used in the analysis of criminal networks [52].
5	Each federation member modifies its criminal intelligence software system so that it can read input from a report file generated by the local database engine.
6	Each federation member adopts a policy in which any new evidence collected is entered only once and only into the local database.

several investigation groups who may work within national parks, provincial governments, or national governments. The approach allows different investigation groups, or federation members, to share a virtual single database of wildlife trafficking criminals.

Implementing a federated database

Implementation requires investigators to run a query against a federated database composed of all evidence collectively held by all cooperating federation members. Typically, a federation member (the requestor) sends an email with a Structured Query Language (SQL), (see Appendix 1) query in it to each federation member. Upon receipt of an email query, a federation member may choose to ignore the query due to lack of trust in the requestor. Alternatively, the federation member may decide to run the query against their local database and send the query's result back to the requestor as an encrypted file attached to an email. Several free utilities are available on the web that encrypt files using a shared key (*e.g.* AxCrypt [33]). Federation members share this key amongst themselves by physically meeting at a central location and sharing a common key string.

The requestor collects all received query responses into a single data file. All of these different, local databases taken together thus form a single, virtual database against which an investigator may run a query. This is particularly advantageous because when data change frequently, executing queries against member databases is more efficient and cost effective than first building a master database before querying it [34].

The requestor then applies criminal social network analyses [35] to this single data file to produce actionable intelligence that can inform tactical responses directed at disrupting wildlife crimes. Investigators may automate parts of this process such as receiving the incoming query request email, running the query against the local database, and generating the outgoing email message containing the query's results [36]. This approach removes exposure risks of any federation member's local database, and does not need specialized software.

Alternatives to the zero-cost implementation highlighted above include a low-cost implementation. In this case each federation member purchases MS Office. This software bundle contains the database engine, MS Access. Database managers construct and maintain the local database in MS Access. The approach carries two advantages: dedicated support from Microsoft Corporation, and the availability of several online forums due to its large user base.

A more expensive implementation allows federation members to achieve secure transmissions without having to encrypt/decrypt individual files by jointly purchasing a Virtual Private Network (VPN). Several web-based database systems may offer greater opportunities for data integration such as a solution based on a set of distributed MS SQL Servers. These solutions however, are more expensive and require each federation member to have access to strong Information Technology (IT) support.

De-identifying data

Although federated databases allow enrichment of the database available to generate actionable intelligence, commercial criminal intelligence software systems may not provide the necessary analysis options to achieve that. Often investigators require specialist analyses of a particular body of evidence. This specialist may not be on the staff of one of the federation members. This introduces considerable risks associated with information confidentiality.

A data file, sent to an outside specialist should thus not contain any classified, private or confidential information. De-identifying or de-classifying the database requires replacing classified, private or confidential information in a database with encrypted or random identifiers. Some commercial software systems (*e.g.* Analyst's Notebook [37]) contain de-classifying options within their report creation capabilities.

If the criminal network changes frequently, specialist analysts may need to run actionable intelligence analyses every week. Automatic de-identification facilitates this process. Within the use of a federated database approach, the specialist analyst needs the encryption algorithm to create a unique codename for each unique player name regardless of the local database that the algorithm is run on.

To address these requirements, the specialist analyst provides a SQL query script to each federation member. This script, when run against a federation member's local database, de-identifies each suspect's name by replacing the name with an encrypted name (hereafter called a codename). Because all federation members use the same encryption key, persons without the key (such as an outside specialist) cannot decrypt codenames back to the original names.

We use an example to illustrate the process. A query script asks for (a) each pair of players that are linked through an intercepted phone call, (b) how many phones, and/or cars, and/or guns each player has, and (c) for each pair of players, the number of evidence items that mention both of them. Each federation member creates an email message containing this de-identified data file and sends it to the outside specialist for creation of actionable intelligence. Upon receipt of all data files from cooperating federation members, the outside specialist aggregates information into one file, analyzes the aggregated network, and then shares the created actionable intelligence with all federation members.

Data aggregation issues

A common encryption key across all federation member databases will result in a unique codename for a unique player as long as that player's name is spelled exactly the same in each federation member's database. But, this cannot always be assumed in practice. To address this and other concerns pertaining to the combination of data from different federation member databases, federation members agree on a common core of attributes and their names, e.g. suspect name, suspect address, suspect contact number 1, …, suspect contact number 10. Next, the specialist eliminates duplicate records caused by spelling differences of a suspect's name across the databases of the federation with an algorithm given by [38] that operates on the address and primary phone number fields only. Note that the suspect's name cannot be used here because the process of name encryption cannot be relied upon to be a monotonic mapping of distances between names in the original name space to distances between names in the encrypted space. See Appendix 1 for an implementation of the algorithm in [38].

Finally, the agreement among the federation members to use a common core of database attributes eliminates the need to disambiguate attribute fields that could arise from so-called *schema heterogeneity* (see [39]).

Creating actionable intelligence

Actionable intelligence is the fundamental ingredient of Intelligence-Led Policing (ILP) [31]. Such analysis should inform policing, from tactical to strategic levels and beyond to government policy. It serves as a model that uses intelligence to guide and shape policy, strategy and operations, rather than simply solving or supporting singular investigations [40].

Criminal network resilience

The appearance of organized transnational crime networks [17-22] stimulated the development and applications of network analytical approaches as part of developing actionable intelligence [31]. Strategically, law enforcement authorities seek the most cost effective action to disrupt or collapse a crime network.

Typically, crime networks carry trade-offs between efficiency and security that reflects in the architecture of a criminal network – a sparsely connected network is more secure, but less efficient [41]. A terrorist network would thus place a higher priority on security than a criminal enterprise such as a wildlife trafficking network. This predicts that a wildlife trafficking network will strive to raise its efficiency (connectedness) after a network disruption action rather than increase its security by remaining sparsely connected. This creates opportunities as well as challenges – networks may recover quickly following a disruptive legal action [23], but at the same time such responses allow opportunities to obtain more information when hidden network players come to the fore in the process [42].

Targeting efficiency or connectedness of wildlife trafficking networks is a high priority for law enforcement authorities. Three challenges arise. First, how often should law enforcers disrupt a network? Which network player should authorities target? How would a network recover if disrupted? The last question typically provides law enforcement guidance to task investigators to focus on key future suspects. Social network analytical approaches [35] provide ways to address these questions.

Social network analysis

Criminal networks have key players [35] identified through social network analytical metrics [43,44]. Eigenvector centrality helps to identify the most connected intermediary, while the betweenness centrality measure indicates the intermediary with the most control over information. Removing these two players would be the most effective strategy for disrupting the network's operations. In addition, the brokerage score flags influential players (Table 2).

To help in interpreting measures, we provide some network notation. Let $G = (V, E)$ be a graph where V is a list of its g vertices (players), and E is a list of its edges or links (the graph's connectivity). Label vertices with v_i, $i = 1, . . ., g$.

Table 2 Social network analysis metrics that allow the identification of key players in wildlife trafficking networks

Metric	Description
Degree centrality	The number of links directly connected to that player
Betweenness centrality	v_i is defined as follows. Consider the j^{th} pair of players $(v^{'}, v^{''})_j$ for which $v' /= v_i$ and $v'' /= v_i$. Let T_j be the number of shortest paths between $v^{'}$ and $v^{''}$. Let n_j be the number of these paths that contain player *i*. The betweenness centrality of v_i is $\sum_{j=1}^{m} \frac{n_j}{T_j}$ where *m* is the number of player-pairs that connect by at least one path [35]. The measure incorporates by re-defining the length of a path between players v_i and v_j to be the sum of the link weight inverses across all of the links in the path [44]. A player with high betweenness centrality has control of information propagation in a network
Eigenvector centrality	Let W be the weighted-link adjacency matrix. Let **e** be the first eigenvector of W. The eigenvector centrality of player v_i is the i^{th} component of **e** [35]. This metric measures a player's influence by measuring how easily information can flow between a player and all other players regardless of the path taken [43]
Network connectedness	The largest eigenvalue of W. Let CI be this index
Gould-Fernandez total brokerage score	The GF total brokerage score for *vi* is the number of player pairs, *v*′, *v*″ for which (*v*′, *vi*) and (*vi*, *v*″) are both in *E* but the link (*v*′, *v*′) is not in *E* [35]. A player with a high brokerage score functions as an intermediary (or broker) for many pairs of players not directly connected to each other

Denote an undirected edge with $\{v_i, v_j\}$, and a directed edge where v_i influences v_j with (v_i, v_j). Let A be the graph's adjacency matrix. The *i*, j^{th} component of A, $a_{i,j}$ is set to the value 1.0 if the corresponding edge is in *E*, and 0, otherwise. Thus, A is a square, $g \times g$ matrix.

Because first-tier intermediaries sponsor a poaching party, there is some justification for modeling the relationships between players with directed edges. In this case, betweenness centrality (Table 2) and proximity prestige score [32] would be the most relevant measures to use to identify key players.

Law enforcers can use social network analysis to help combat wildlife trafficking crimes by reconstructing the trafficking network from data on messages passed between players, followed by the identification of key players (Table 3). Law enforcement can then focus legal actions on these key players to disrupt the network's operations based on the actionable intelligence report that analysts provide (Figure 1).

Link weights

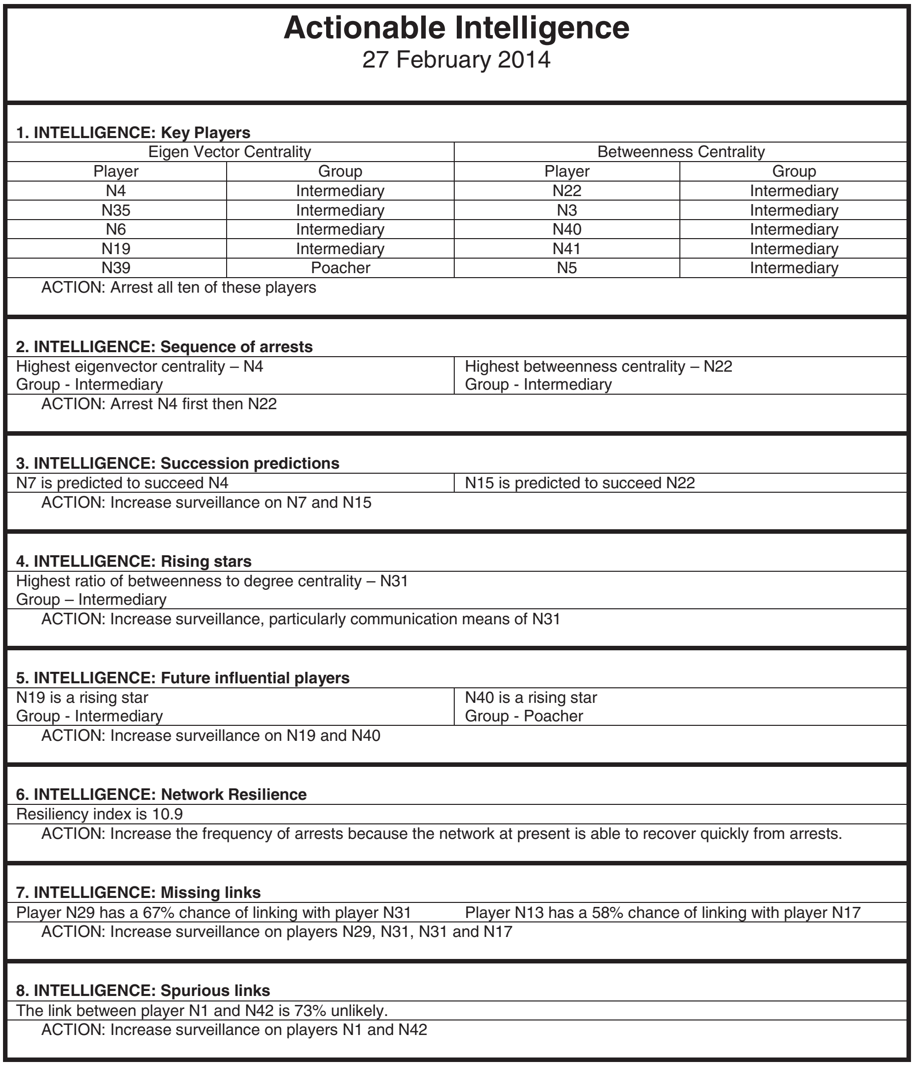
A key challenge arises in that different qualities of information comprise a criminal information database. Typically information needs verification to become intelligence. Crime database managers scale information based on reliability of the source and verification thereof. Intelligence that link players in social network analyses is thus of varying significance. In addition, a pair of players may be referenced on several pieces of evidence collected by investigators. For example, a mobile phone number that associates with both players (Table 4).

We capture this differential intelligence quality challenge by creating link weights. Information on player attributes is often available, but seldom incorporated to establish link weights [45]. For the player pair v_i, v_j, let b_{ij} be the number of these evidence items. Let $b_{\max}$ be the largest of these values across all player pairs in the network. In our rhino horn trafficking example, information on the numbers of mobile phones, vehicles, and guns owned by a player is available. Let c_i, and d_i be the number of mobile phones, and vehicles, respectively owned by player v_i. Let $c_{\max} = max\ \{c_1, \ldots, c_g\}$. Define $d_{\max}$ similarly. Steinhaeuser & Chawla [45] add the value $1 - \alpha|val_i - val_j\ |$ to the weight of the existing link between players v_i and v_j where val_i is the attribute value of player v_i, and α is a normalizing constant. The index notates the higher the attribute similarity, the higher the weight. One way to combine all three of these

Table 3 Actionable intelligence products provided through social network analysis of information

Product	Action
List of players who have high network centrality values	Remove all of these players from the network
Recommended sequence of players to remove	Remove these players in the recommended order
Prediction of players most likely to succeed removed players	Increase surveillance on potential successors
Prediction of influential players who are attempting to conceal themselves	Increase surveillance on these players as they are often highly influential and the sole connection to other networks
Predictions of rising stars	Increase surveillance on these increasingly influential players
Network resilience index value	Increase frequency of removals if this index is high

We also provide examples of response actions.



Actionable Intelligence

27 February 2014

1. INTELLIGENCE: Key Players

Eigen Vector Centrality		Betweenness Centrality	
Player	Group	Player	Group
N4	Intermediary	N22	Intermediary
N35	Intermediary	N3	Intermediary
N6	Intermediary	N40	Intermediary
N19	Intermediary	N41	Intermediary
N39	Poacher	N5	Intermediary

ACTION: Arrest all ten of these players

2. INTELLIGENCE: Sequence of arrests

Highest eigenvector centrality – N4 Group - Intermediary	Highest betweenness centrality – N22 Group - Intermediary

ACTION: Arrest N4 first then N22

3. INTELLIGENCE: Succession predictions

N7 is predicted to succeed N4	N15 is predicted to succeed N22

ACTION: Increase surveillance on N7 and N15

4. INTELLIGENCE: Rising stars

Highest ratio of betweenness to degree centrality – N31
Group – Intermediary

ACTION: Increase surveillance, particularly communication means of N31

5. INTELLIGENCE: Future influential players

N19 is a rising star Group - Intermediary	N40 is a rising star Group - Poacher

ACTION: Increase surveillance on N19 and N40

6. INTELLIGENCE: Network Resilience

Resiliency index is 10.9

ACTION: Increase the frequency of arrests because the network at present is able to recover quickly from arrests.

7. INTELLIGENCE: Missing links

Player N29 has a 67% chance of linking with player N31	Player N13 has a 58% chance of linking with player N17

ACTION: Increase surveillance on players N29, N31, N31 and N17

8. INTELLIGENCE: Spurious links

The link between player N1 and N42 is 73% unlikely.

ACTION: Increase surveillance on players N1 and N42

Figure 1 Example of an actionable intelligence report. The example illustrates a report generated for law enforcement authorities. Players have internal identification numbers that remove risks when investigators use external specialist analysts. Note that predicted group membership plays a role in actions responding to actionable intelligence items 2 and 4.

Table 4 Examples of data types and sources connecting players in a criminal network investigated by law enforcement authorities

Sender	Recipient	Message/Action type	Data source
A	D	Conversation	Cellphone intercept
B	D	Payment	Field evidence/confession
C	D	Conversation	Cellphone intercept
D	E	Conversation	Cellphone intercept
E	G	Conversation	Cellphone intercept
F	G	Conversation	Interrogation
G	H	Conversation	Cellphone intercept
H	I	Conversation	Cellphone intercept
I	J	Conversation	Cellphone intercept

contributions to a link's weight is as follows. For players v_i and v_j, let

$$w_{ij} \equiv 2 + \frac{b_{ij}}{b_{\max}} - \frac{|c_i - c_j|}{2c_{\max}} - \frac{|d_i - d_j|}{2d_{\max}}$$

if $b_{ij} > 0$. Otherwise, set w_{ij} to zero (no link). A set of players in the network will hold the value $c_{\max}$. Therefore, this value will force the third term of the w_{ij} definition to take on values between 0 and 0.5 for all player-pairs in the network. If a pair of players have similar values of c_i this third term will be close to zero – resulting in a small penalty for dissimilar c_i values. On the other hand, for a pair of players for which one holds the value $c_{\max}$ and the other holds the value zero, the penalty will take on its maximum value of 0.5. This explanation applies to the fourth term as well. If the pair of players have the same c_i values and the same d_i values, their link weight will be between 2.0 and 3.0. At the other extreme, when $b_{\max}$ is large, use of this definition yields a link weight of about 1.0 for a player-pair that have been linked through a single evidence item and who have maximally different quantities of phones and vehicles.

This approach assumes a relationship between node attributes and a link's weight: high attribute similarity translates to high weight. If this relationship is unknown, the stochastic blockmodel [46] may be extended by adding attribute affinity matrices [47]. The entries in these matrices are model parameters and hence need estimation.

Dynamic social networks

Responses of networks to law enforcement actions require some understanding of how criminal networks may recover [48]. Investigators require a dynamic model of the criminal network, which helps to define an arrest sequence and the network resiliency index. Yuan *et al.* [49] use a sequence of networks across discrete time, $G(i\Delta t)$, $i = 1, \ldots,$ where Δt is the discrete time step. Communication events observed between players after time $i \times \Delta t$, but before time $(i + 1) \times \Delta t$ are taken to be links in the network, $G(i\Delta t)$. The value of Δt is set to a value large enough so that for each i, the sample used to reconstruct $G(i\Delta t)$ contains at least one communication event to or from each player active in that network. Let $C\, I_i$ be the connectedness index value of network $G(i\Delta t)$.

Network disruption

Frequency of disruptive actions

Wildlife trafficking criminal networks are typically resilient – they are able to quickly reform themselves after a network disruption action [23]. Computing an index of resilience (*e.g. RI* - the inverse of the number of weeks to regain 90% functionality after a disruptive action) could help law enforcement units to adjust the frequency of their network disruption efforts to maximize crime disruption. One measure of a network's functionality is its overall connectedness – and one measure of that is CI, the largest eigenvalue of the matrix of link weights, W. In addition to tracking the network's resilience measure, by following changes in CI, law enforcement agencies can assess the degree to which their network disruption actions have affected network functionality.

Optimal sequence of arrests and rising stars

Yuan *et al.* [49] recommend two procedures for deciding the next arrest. We apply these recommendations using social network centrality measures in two ways. First, if a player's eigenvector centrality (Table 2) is increasing through time, law enforcement should arrest this player first because the player may be the next rising star. Second, if a player's betweenness centrality (Table 2) is increasing through time, law enforcement should arrest the player last if it is important for investigators to not alert other players in the network.

An additional network disruption action is to remove players with high ratios of betweenness centrality to degree centrality [50]. The reason for this guideline is that such players may actually be running the criminal operation while trying to remain inconspicuous. Therefore, they should be removed immediately even if their high ratio of betweenness to degree centrality is due to a high betweenness centrality value – contradicting the above recommendation of removing players with high betweenness centrality last.

These guidelines should only be applied to influential players in the network.

Succession

Which player is most likely to take up the role that an arrested player had in a network? Such a player should be the recipient of increased surveillance. Typically, network dynamics predict that when a player is lost, surrounding players quickly establish new connections and share responsibilities of the lost member [51]. Hence, one way to predict who will take up an arrested player's role is to identify the player that is closest to the removed player along three dimensions: 1) path length between the player and the removed player, 2) eigenvector centrality, and 3) betweenness centrality. An investigator could test this prediction by re-computing the network's centrality measures after a period has passed since the arrest.

The authors' **id** software system predicts who these people will be by first finding the player directly connected to the first (to be) arrested player who has the most similar eigenvector centrality. This player is predicted to succeed the first arrested player. In a similar manner, the player directly connected to the second arrested player who has the most similar betweenness centrality value to that of the arrested player is predicted to succeed the second arrested player.

Criminal network inference

Types of data used to observe the network

Investigators use several types of data to determine the groups that players belong to as well as how players link to each other. These include names mentioned while interrogating a poaching suspect, interceptions of mobile phone conversations, evidence gathered at the scene of a poaching incident and informant reports. These provide a data set composed of observations between pairs of players or links. In such a data set, some links may be unobserved, and some observed links might be spurious.

Often, observed communication events are between pairs of players whose identities are unknown. For such data, analysts assign to both players identifier labels (*e.g.* Table 4). Link software packages such as Pajek [52], (see Appendix 3) can draw the network formed from this data (Figure 2a).

Types of inference

To overcome hidden or spurious links in observed networks, some network inference may assist investigators to reconstruct a crime network. At least three network inference tasks arise [53]. The first is to assign the name of a player to each identifier label in the data set – entity resolution. Next is to predict the set of edges in the true network – link prediction. Finally, the analyst predicts the role of each player in the network – collective classification. We focus on two approaches that allow analysts to reconstruct a network from an incomplete data set on pairs of linked players. The graph-based approach [54] employs only basic graph-theoretic measures to predict the presence of a link and does not attempt to model the network in any way. The model-based approach [46]

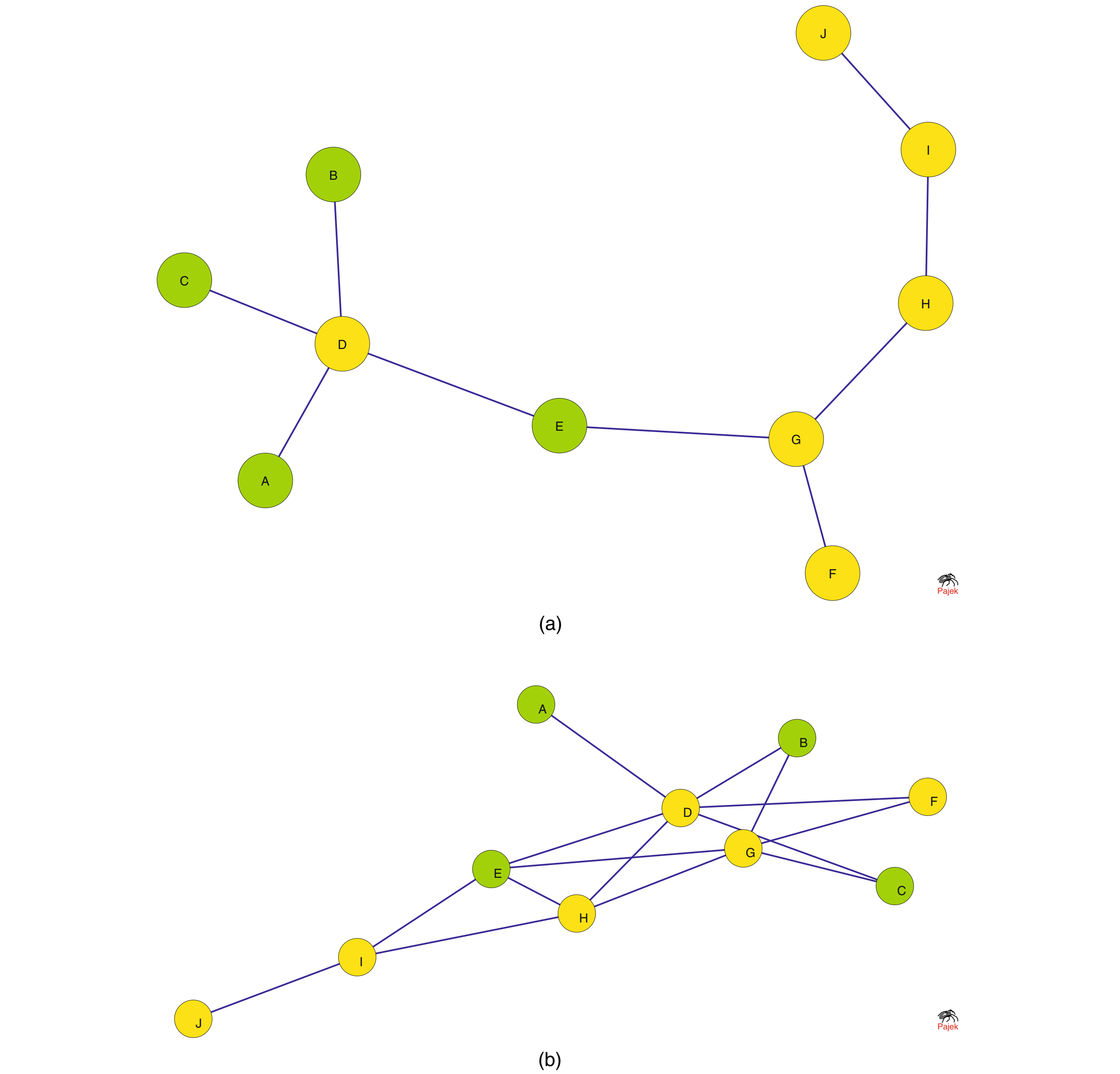


Figure 2 Construction of criminal networks from link data. a. Observed links. **b**. Network inference results using model-based predictions identifying missing and/or spurious links between players (node colors indicate group membership predictions). Graphs are drawn with the software package, Pajek [52].

assumes the network was generated by a probabilistic function. Under this assumption, analysts use the data to first estimate the model's parameters, and then to predict link presence or absence from this fitted model.

Graph-based link prediction and network reconstruction

For large networks, a statistical estimate of the network's parameters may be too time- consuming to compute on typical desktop computers. In these cases, analysts may employ a graph-based approach. If the network has many unobserved links, a graph-based link predictor should not depend on having information at both ends of a potential link. Rather, it is better to be robust to loss of information at either of the two ends of a potential link [54].

Based on these comments, we evaluated the use of both Common Friends (CF) and Total Friends (TF) in a method to score potential links as follows. Let $\Gamma(v)$ be the set of players that are directly connected to player v. Note that player v's degree centrality is $|\Gamma(v)|$. Hence, degree centrality can be interpreted as the number of friends that player v has, *i.e.*, the number of players that are directly connected to player v. The CF and TF scoring functions are

$$\text{score}_{CF}\left(v, v'\right) = \left|\Gamma(v) \cap \Gamma\left(v'\right)\right|,$$

and

$$\text{score}_{TF}\left(v, v'\right) = \left|\Gamma(v) \cup \Gamma\left(v'\right)\right|.$$

The CF score is the number of players that are directly connected to both of the players, v and v' – while the TF score is simply the total number of unique friends enjoyed by both of these players. To make link predictions, the analyst computes these scores for each pair of players not linked in the observed data. A predicted link for a pair of players results if their CF score is larger than the 90% quantile of the CF scores computed on all observed link pairs, or the TF score is larger than 99.99% quantile of the TF scores computed on all observed link pairs. This graph-based link prediction algorithm is computationally inexpensive allowing it to be applied to networks of any size.

Predicting group membership using a graph-based approach makes use of conditional statements (*e.g.* for the case study below, If a player has only one link into the network, assign that player to group 1 (poachers). Otherwise, assign the player to group 2 (intermediaries)).

Model-based link prediction and network reconstruction

We introduce a frequentist (non-Bayesian) method for model-based link prediction from link data in which (a) there may be unobserved links, and (b) an observed link is reliable, *i.e.*, there are no observed links that are spurious. This method assumes the true network is stochastic in its links and group membership. We model this random network as a stochastic blockmodel.

Let Y_{ij} be a Bernoulli random variable that, when holding the value 1, indicates the presence of a link between player i and player j. Let X_i be an m-valued discrete random variable that indicates the group that player i belongs to. The parameters of this random variable are $\theta_1, \ldots, \theta_m : P\ (X_i = k) = \theta_k$. Within the stochastic blockmodel, $P\ (Y_{ij} = 1) = \eta_{Xi,Xj}$. In other words, the probability of a link between players i and j is a function of only the group memberships of the two players.

Snijders & Nowicki [46] give the joint probability of the stacked random vector

$$\left(\mathbf{Y}', \ \mathbf{X}'\right)' :$$

$$P(\mathbf{Y} = \mathbf{y}, \ \mathbf{X} = \mathbf{x}; \ \boldsymbol{\theta}, \ \boldsymbol{\eta}) = \left[\prod_{i=1}^{m} \theta_i^{n_i}\right] \prod_{1 \le k \le l \le m} \eta_{kl}^{e_{kl}} \left(1 - \eta_{kl}\right)^{n_{kl} - e_{kl}}$$

where n_k denotes the number of players that belong to group k, and e_{kl} denotes the number of links that are between a player in group k and a player in group. The count

$$n_{kl} = n_k n_l I_{\{k \neq l\}}(k, l) + \binom{n_k}{2} I_{\{k = l\}}(k, l).$$

Let $\mathbf{y}_o$ be the set of links that are observed to be either "on," or "off". An investigator may suspect that censored potential links exist between certain players, *i.e.*, because present data do not allow observation of these potential links, it is unknown if they exist or not. Such censoring may occur because the investigator believes these players are using a method of communication not monitored by law enforcement. In other words, all that the investigator is sure of are the person-to-person communication events already observed. The vector $\mathbf{y}_c$ capture the unobserved potential links. Let n_c be the number of these unobserved potential links. The likelihood function comes from summing the joint probability over all possible values of the unobserved random vector $\mathbf{X}$, and the unobserved links, $\mathbf{Y}_c$ giving:

$$P(\mathbf{Y}_o = \mathbf{y}_o; \boldsymbol{\theta}, \boldsymbol{\eta}) = \sum_{\mathbf{x} \in \{1,2\}^g, \mathbf{y}_c \in \{0,1\}^{nc}} P(\mathbf{Y} = \mathbf{y}, \mathbf{X} = \mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\eta})$$

By maximizing this equation we estimate parameters given the data over all possible values of $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ (see Appendix 2 for further details). Using these estimated parameter values, we reconstruct the trafficking network by simulating the conditional expected values of $\mathbf{Y}$ and $\mathbf{X}$ given the data, $\mathbf{y}_o$. This simulation proceeds by randomly drawing many networks from the estimated model and retaining only those networks for which $\mathbf{Y}_o =$

$\mathbf{y}_o$. Link ij is predicted to exist if the average value of Y_{ij} over these retained networks is greater than 0.5. Player i is predicted to belong to group k if k is the closest integer to the average of the X_i values over these retained networks. The reconstructed network then, is the conditional expected network under the estimated stochastic blockmodel given the set of observed links. This simulation procedure is not trivial because each player does not necessarily keep membership in the same group from simulated network to simulated network. Note that under this frequentist approach to network reconstruction, we interpret the data to be a (possibly censored) realization from the stochastic network defined by the joint probability law given above.

As an example, we reconstruct the network from the data set of Table 4 with the above statistical method to produce a reconstructed network (Figure 2b). Table 5 collects centrality measures of the reconstructed network and delivers an optimal removal strategy based on this social network analysis.

Table 5 Predicted group membership and centrality scores of players derived from a reconstructed network

Player	Predicted group	Eigenvector centrality	Betweenness centrality	GF total brokerage
D	1	0.478	0.431	14
G	1	0.442	0.208	9
H	1	0.411	0.178	3
E	2	0.411	0.178	3
I	1	0.255	0.222	2
C	2	0.228	0.011	1
B	2	0.228	0.011	1
F	1	0.228	0.011	1
A	2	0.118	0.0	0
J	1	0.065	0.0	0

We order players by their eigenvector centrality value. Removal of players D, G and I should seriously disrupt the network's operations. Members of Group 1 are first-tier intermediaries having more influence than those of Group 2.

Spurious and missing links

Investigators can identify spurious links through a non-Bayesian approach with a measurement model. Let X be the Bernoulli random variable representing the true potential link, and Y be the observable potential link. Let $\gamma \equiv P\ (Y = 1|X = 0)$ be the probability of observing a spurious link. If for example $\gamma \approx 0.01$ there is little chance that a reported link is spurious. An alternative Bayesian approach due to Guimerà & Sales-Pardo [55] allows the estimation of the edge set of a network from noisy and incomplete data that is resistant to the effects of missing and/or spurious edges in the observed network. Use of either approach allows the identification of potential missing and/or spurious links to inform the actionable intelligence report that is provided to law enforcement authorities (Figure 1).

Case study: Rhino poaching

Kruger National Park, South Africa, is a stronghold of two sub-species of rhino, the southern white rhino (*Ceratotherium simum simum*) and the south-eastern black rhino (*Diceros bicornis minor*). In Kruger, both these species suffer a new poaching onslaught for their prized horns [56]. Despite intensified anti-poaching activities, the number of rhinos poached per day continued to increase since 2008 (Department of Environmental Affairs (DEA) South Africa, unpublished data). Although the outcome of such a poaching threat is not detectable at the population level in Kruger National Park for both black [28] and white rhino [14], continued trends in poaching predicts detectable declines by 2016.

Poaching compromises at least two conservation values: the opportunity to contribute to rhino range expansion programmes [57] as well as generating revenue from rhino sales [14]. An integrated framework [10] aims to manipulate demand and supply dynamics associated with the use of rhino horn and in the process alter the incentives and disincentives within the decision making of poachers [58]. The disruption of transnational organized crime syndicates is thus a key element that influences various aspects of the proposed integrated framework [10].

We used real-world network data on a rhino horn trafficking network operating close to Kruger National Park. The example comprises a 134-player network defined by data collected during September 2013. We used a modified form of the ForceAtlas2 graph-drawing algorithm [59] implemented in the **id** software system to illustrate a re-constructed network (Figure 3). The ForceAtlas2 algorithm reveals clusters of players and the connections between these clusters. We also used the **id** software system to produce an Actionable Intelligence Report for the end of September 2013 (Figure 4). Note that the network reconstruction algorithm added several links (see Figure 3). Player H100's high eigenvector centrality is illustrated in the graph: this person is able to transfer information in parallel across chains of other players rather than having to channel his/her information through one neighbor at a time, i.e. he/she is a player that can communicate with most other players through a multiplicity of short walks [35] and hence can propagate information quickly to many players in the network. Player H28 stands as a gatekeeper between large numbers of players distributed across widely dispersed sub-networks. Removing player H100 will damage the network's leadership, and removing player H28 will damage the operation's ability to communicate a message to all members of the network. Note that player H100 possesses both high eigenvector and high betweenness centrality.

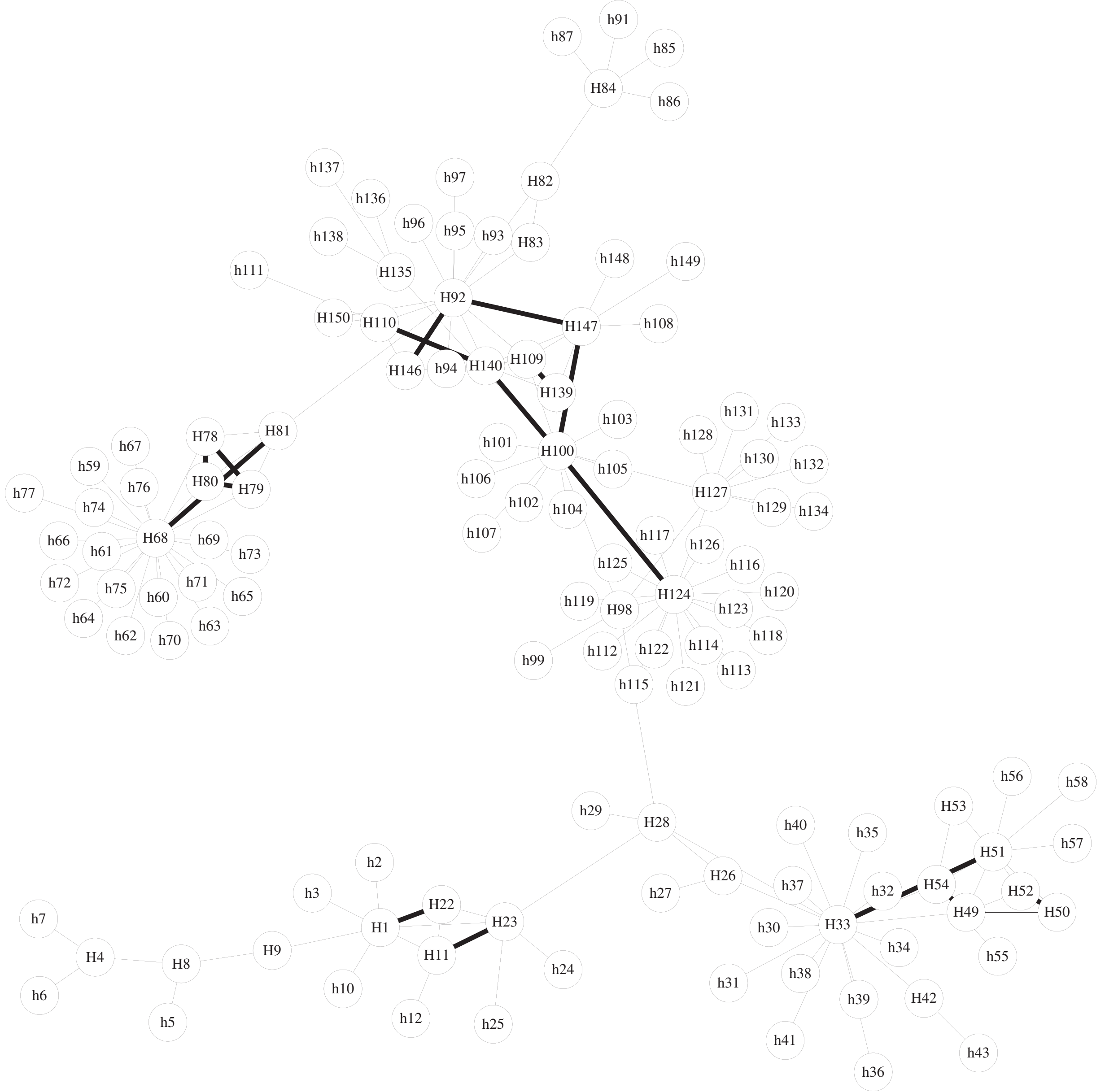


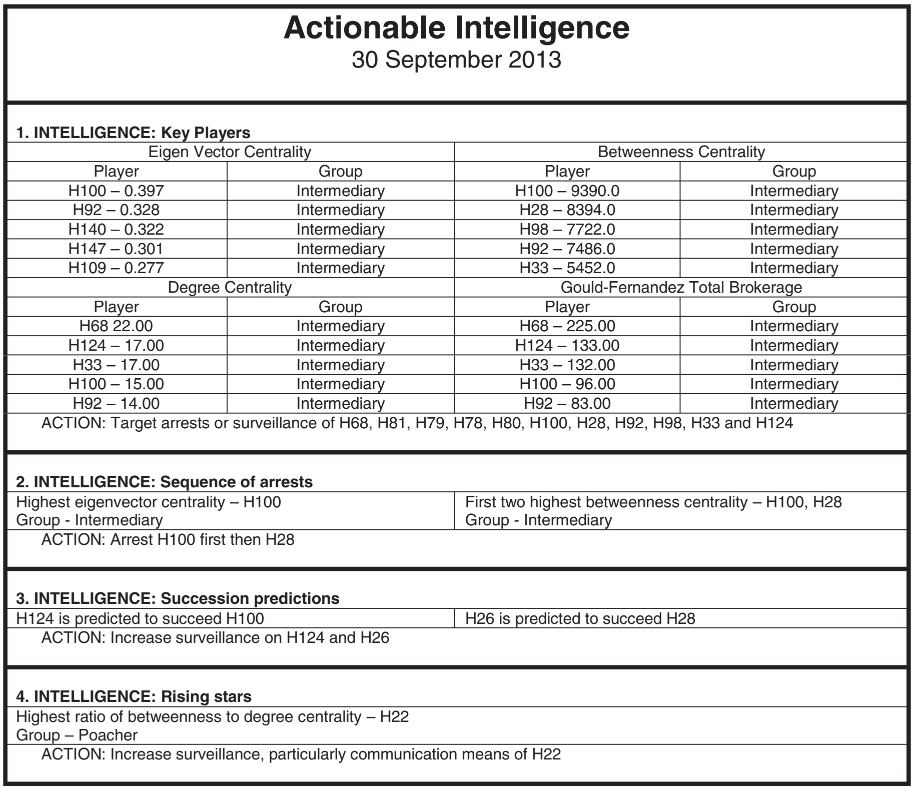
Figure 3 An example of the reconstructed criminal network associated with rhino poaching in Kruger National Park as of September 2013. Heavy lines indicate links added during network reconstruction.

The network is composed of several loosely-connected sub-networks. Each of these sub-networks consists of members of a local gang that all live within a town bordering the park that is typically densely populated and very poor.

Conclusions

This article gives (a) a practical, software-based guide to developing a federated database of criminal evidence items, (b) a non-Bayesian statistical estimation method of the true network, (c) generalized link weights, (d) a focused, simplified protocol for using disparate social network measures to identify key players in a criminal network, and (e) a practical, software-supported implementation of the here-to-fore vague concept of actionable intelligence.

Specific to our case study, the continued onslaught of illegal wildlife trafficking transnational crime syndicates on charismatic species such as rhino [56] forced authorities to consider different and integrated approaches. These range from intensified protection, to demand reduction and ways of providing products [10]. We argued that these contentious strategic responses share a common challenge – competition and conflict with wildlife



Actionable Intelligence

30 September 2013

1. INTELLIGENCE: Key Players

Eigen Vector Centrality		Betweenness Centrality	
Player	Group	Player	Group
H100 – 0.397	Intermediary	H100 – 9390.0	Intermediary
H92 – 0.328	Intermediary	H28 – 8394.0	Intermediary
H140 – 0.322	Intermediary	H98 – 7722.0	Intermediary
H147 – 0.301	Intermediary	H92 – 7486.0	Intermediary
H109 – 0.277	Intermediary	H33 – 5452.0	Intermediary

Degree Centrality		Gould-Fernandez Total Brokerage	
Player	Group	Player	Group
H68 22.00	Intermediary	H68 – 225.00	Intermediary
H124 – 17.00	Intermediary	H124 – 133.00	Intermediary
H33 – 17.00	Intermediary	H33 – 132.00	Intermediary
H100 – 15.00	Intermediary	H100 – 96.00	Intermediary
H92 – 14.00	Intermediary	H92 – 83.00	Intermediary

ACTION: Target arrests or surveillance of H68, H81, H79, H78, H80, H100, H28, H92, H98, H33 and H124

2. INTELLIGENCE: Sequence of arrests

Highest eigenvector centrality – H100
Group - Intermediary

First two highest betweenness centrality – H100, H28
Group - Intermediary

ACTION: Arrest H100 first then H28

3. INTELLIGENCE: Succession predictions

H124 is predicted to succeed H100

H26 is predicted to succeed H28

ACTION: Increase surveillance on H124 and H26

4. INTELLIGENCE: Rising stars

Highest ratio of betweenness to degree centrality – H22
Group – Poacher

ACTION: Increase surveillance, particularly communication means of H22

Figure 4 Kruger Rhino actionable intelligence report. An example of the output created for the rhino case study. Note that we observed the network only once; hence, we could not predict rising stars or compute the network's resilience.

crime syndicates. Because these are transnational in nature, law enforcers from a single jurisdiction have limited success in lasting disruption of crime networks because information is in different databases. We illustrated the use of federated databases to overcome this problem. Even so, considerable trust between different jurisdictions remains a key element. We also showed how several measures derived from social network analyses can provide several types of actionable intelligence to law enforcers. This will allow targeted legal responses or tasking of investigators. We conclude that the use of social network analyses applied to federated data may greatly assist law enforcers to disrupt or even collapse transnational wildlife trafficking criminal networks efficiently.

Appendix 1

MySQL Database of Wildlife Trafficking Criminals

Haas [60] created a free criminal intelligence database written in the freely available MySQL language. It follows the SQL criminal intelligence database described in [61]. MySQL version 5.2.3 or higher is needed and is available from [62].

The web resource of Haas [60] contains an example that consists of three files: createdatabase.sql, addtodatabase.sql, and querydatabase.sql. The first file creates a six-table database comprising of players, phones, cars, guns, random identifiers and encryption keys. The second file gives an example of adding three suspects to the database. The third file runs the query needed for creating actionable intelligence.

The query file in the example generates two output files that replace player names using two different coding approaches. The first uses internally generated random identifier numbers as player identifiers. Doing so relieves the data manager from having to maintain a log relating local suspect identifiers to local suspect names. Note however, that random identifiers are not a substitute for encrypted names in a federated database because there is no guarantee that the same suspect will have the same random identifier in the local databases of two or more federation members (see text).

The second output file is an example wherein suspect names are encrypted via the MySQL implementation of the AES encryption algorithm of Daemen & Rijmen [63]. This encryption algorithm is considered unbreakable.

The web resource [60] also contains an implementation of the deduplication algorithm in [38].

Appendix 2

Parameter estimation details

For $g > 20$, the stochastic blockmodel's likelihood function becomes computationally intractable. For these networks, maximum simulated likelihood (MSL) as implemented in the **id** software system [60] is used to fit the parameters. This implementation of MSL requires a distance metric between two networks. For this purpose, the metric of [64] is employed because its run-time is linear in g. The marginal, multivariate distribution of $\mathbf{Y}_o$ is approximated by simulating many networks, ($\mathbf{Y}$, $\mathbf{X}$) and computing the density estimate of $\mathbf{y}_o$ using only the simulated values on $\mathbf{Y}_o$ regardless of the values on ($\mathbf{Y}_c$, $\mathbf{X}$). One could argue that no metric is needed because with a large enough set of simulated networks, the probability of networks that contain the observed network is simply the fraction of these simulated networks that contain the observed one. The difficulty is that there is a large number of states: potentially 2^{g^2} or 2.58×10^{120} for $g = 20$. This means that a very large number of simulated networks might be needed in order to generate just one that contains the observed network when the parameters have been set to values that make networks containing the observed one unlikely. For the optimization's run-time to be reasonable however, only a modest number of simulated networks can be generated at each evaluation of the objective function (each setting of parameter values). In this case, it is unlikely that any network will be simulated that contains the one observed. But unless such networks are simulated, the probability of networks that contain the observed one under a particular setting of parameter values will be exactly zero. This in-turn, will produce large regions of the search space wherein the objective function does not change value; thus causing the optimization algorithm to fail.

Appendix 3

Software options for creating actionable intelligence

Pajek is a free package that can draw a network and compute centrality measures on it. Network inference via the stochastic blockmodel can be performed with the free package **id** [60]. The Actionable Intelligence Report described in this article is implemented in **id**. As mentioned above, a commercial package from IBM is Analyst's Notebook [37]. This system combines a non-SQL criminal intelligence database with extensive network drawing capabilities and SNA metrics. The Tartan plug-in for Analyst's Notebook from [65] allows unobserved links to be predicted via the Common Friends score function (see text). Neither Analyst's Notebook nor the Tartan plug-in have the capability to predict group memberships. Another commercial package is Sentinel Visualizer [66]. This package has many SNA metrics and, unlike Analyst's Notebook, an SQL-compliant criminal intelligence database. Sentinel Visualizer cannot infer links or group memberships.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SF posed the problem, potential solution, contributed to the development of link weights, and wrote the problem description. TH developed the network reconstruction algorithm, wrote the statistical sections, wrote the software system, developed the actionable intelligence report format, created the motivating example, processed the data, and created the figures. Both authors read and approved the final manuscript.

Author information

Timothy C. Haas is an associate professor in the Supply Chain and Operations Management/Quantitative Methods group in the Lubar School of Business, University of Wisconsin-Milwaukee. Dr. Sam Ferreira is Large Mammal Ecologist, Scientific Services, South African National Parks.

Author details

[1]Lubar School of Business, University of Wisconsin-Milwaukee Milwaukee, Wisconsin, USA. [2]Scientific Services, SANParks, Skukuza, South Africa.

Received: 18 September 2014 Accepted: 13 March 2015
Published online: 12 April 2015

References

1. GE Rosen, KF Smith, Summarizing the evidence on the illegal international trade in wildlife. Ecohealth. **7**, 24–32 (2010)
2. S Demetriou, R Muggah, I Biddle, *Small arms availability, trade and impacts in the republic of Congo* (International Organization for Migration and the United Nations Development Programme, Geneva, Switzerland, 2002)
3. SX Zhang, *Smuggling and trafficking in human beings: all roads lead to America* (Praeger Publishers, Westport, United States, 2007)
4. N South, T Wyatt, Comparing illicit trades in wildlife and drugs: an exploratory study. Deviant. Behavior. **32**, 538–561 (2011)
5. T Wyatt, *Green criminology and wildlife trafficking: the illegal fur and falcon trades in Russia Far East* (LAP Lambert Academic Publishing, Saarbrücken, Germany, 2012)
6. Duffy R, St John FAV. Poverty, poaching and trafficking: What are the links. Evidence on Demand. 2013. http://dx.doi.org/10.12774/eod_hd059.jun2013.duffy. Accessed 19 Jan 2015.
7. R Barone, D Masciandaro, Organized crime, money laundering and legal economy: theory and simulations. European. J. Law. Economics. **32**, 115–142 (2011)
8. R Damiana, EH Bulte, The economics of wildlife farming and endangered species conservation. Ecol. Econ. **62**, 461–472 (2007)
9. PH Sand, Enforcing CITES: The rise and fall of trade sanctions. Rev. European. Int. Environ. Law. **22**, 251–263 (2013)
10. SM Ferreira, B Okita-Ouma, A proposed framework for short-, medium- and long-term responses by range States to curb poaching for African rhino horns. Pachyderm. **51**, 60–74 (2012)
11. L Moore, The neoliberal elephant: exploring the impacts of the trade ban in ivory on the commodification and neoliberalisation of elephants. Geophys. J. Roy. Astron. Soc. **42**, 51–60 (2011)
12. KD Messer, Protecting endangered species: when are shoot-on-sight policies the only viable option to stop poaching? Ecol. Econ. **69**, 2334–2340 (2010)
13. JL Schneider, Reducing the illicit trade in endangered wildlife: the market reduction approach. J. Contemporary. Criminal. Justice. **24**, 274–295 (2008)
14. SM Ferreira, JM Botha, M Emmett, Anthropogenic influences on conservation values of white rhinos. Public. Library. Sc. ONE. **7**, e45989 (2012). doi:10.1371/ journal.pone.0045989
15. B Child, The sustainable use approach could save South Africa's rhinos. S. Afr. J. Sci. **108**, 21–25 (2012)
16. D Biggs, F Courchamp, R Martin, HP Possingham, Legal trade of Africa's rhino horns. Science **339**, 1038–1039 (2013)
17. P Williams, Transnational criminal networks, in *Networks and netwars: the future of terror, crime and militancy*, ed. by J Arquilla, D Ronfeldt (RAND, Pittsburg, 2001)
18. P Ferrier, Illicit agricultural trade. Agric. Resource. Econ. Rev. **37**, 273–287 (2008)
19. E Catullo, An agent based model of monopolistic competition in international trade with emerging firm heterogeneity. J. Artificial. Societies. Social. Simulation. **16**, 7 (2013)

20. N Dudley, S Stolton, W Elliott, Wildlife crime poses unique challenges to protected areas. PARKS. **19**, 7–12 (2013)
21. T Milliken, J Shaw, *The South Africa-Viet Nam rhino horn trade nexus: a deadly combination of institutional lapses, corrupt wildlife industry professionals and Asian crime syndicates* (TRAFFIC, Johannesburg, 2012)
22. GL Warchol, The transnational illegal wildlife trade. Criminal. Justice. Studies. Critical. J. Crime Law. Soc. **17**, 57–73 (2004)
23. J Ayling, What sustains wildlife crime? Rhino horn trading and the resilience of criminal networks. J. Int. Wildlife Law Policy. **16**, 57–80 (2013)
24. AP Sheth, JA Larson, Federated database systems for managing distributed, heterogeneous, and autonomous databases. ACM. Comput. Surveys. (CSUR). **22**(3), 183–236 (1990). doi:10.1145/96602.96604
25. C Morselli, K Petit, Law enforcement disruption of a drug importation network. Global. Crime. **8**, 109–130 (2007)
26. United Nations, Police information and intelligence systems. Office on Drugs and Crime (2006). http://www.unodc.org/documents/justice-and-prison-reform/cjat_eng/4_Police_Information_Intelligence_Systems.pdf. Accessed 26 March 2015.
27. D Knoke, S Yang, *Social network analyses*, 2nd edn. (Sage Publications, London, 2008)
28. S Ferreira, CC Greaver, MH Knight, Assessing the population performance of the black rhinoceros in Kruger National Park. South. African. J. Wildlife. Res. **41**, 192–204 (2011)
29. Maggs K. Rhino poaching in South Africa. K2C Rhino Forum Workshop, Hoedspruit, South Africa, 30 September 2011. https://www.environment.gov.za/sites/default/files/docs/nationalrhino_conservation_hawks_sanparks_0.pdf. Accessed 26 March 2015.
30. M Innes, N Fielding, N Cope, The appliance of science? The theory and practice of crime intelligence analysis. British. J. Criminol. **45**, 39–57 (2005)
31. P Bell, M Congram, Intelligence-Led policing (ILP) as a strategic planning resource in the fight against transnational organized crime (TOC). Int. J. Business. Commerce. **2**, 15–28 (2013)
32. K Faust, S Wasserman, Centrality and prestige: a review and synthesis. J. Quantitative. Anthropol. **4**, 23–78 (1992)
33. Axantum. AxCrypt - Password protect files with strong encryption. 2014. www.axantum.com/AxCrypt/Default.html. Accessed 28 Jan 2014.
34. Haase P, Math a*β* T, Ziller M. An evaluation of approaches to federated query processing over linked data. I-SEMANTICS'10 Proceedings of the 6th International Conference on Semantic Systems. Article No. 5, September 1–3, Graz, Austria.
35. CT Butts, Social network analysis with sna. J. Stat. Software **24**, 1–51 (2008). 2010. http://dl.acm.org/citation.cfm?id=1839713. Accessed 8 Nov 2013
36. Mehta A, Williams D. SQL and Outlook: Enable database access and updates through exchange and any E-mail client. MSDN Magazine, Microsoft Corp. 2002. www.msdn.microsoft.com/en-us/magazine/cc301799.aspx. Accessed 8 Nov 2013.
37. IBM. i2 Analyst's Notebook. 2013. www.ibm.com/software/products/en/analysts-notebook. Accessed 24 December 2013.
38. Bhattacharya I, Getoor L. Relational Clustering for Multi-type Entity Resolution. Proceedings of the Fourth International Workshop on Multi-Relational Data Mining (MRDM-2005), Aug 21, 2005, Chicago, 2005. http://www.umiacs.umd.edu/~getoor/Publications/mrdm05.pdf. Accessed 26 March 2015.
39. Hull R. Managing semantic heterogeneity in databases: a theoretical perspective. PODS '97 Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems. 1997. p. 51–61. http://groups.lis.illinois.edu/amag/langevgroup/localpapers/hull-managing-semantic-heterogeneity-pods-1997.pdf. Accessed 17 Jan 2015.
40. G Wardlaw, J Boughton, Intelligence-Led policing: the AFP approach, in *Fighting crime together: The challenges of policing and security networks*, ed. by J Fleming, J Woods (University of New South Wales Press, Sydney, Australia., 2006)
41. C Morselli, C Giguère, K Petit, The efficiency/security trade-off in criminal networks. Social. Networks **29**, 143–153 (2007)
42. VE Krebs, Mapping networks of terrorist cells. Connections. **24**, 43–52 (2002)
43. SP Borgatti, Centrality and network flow. Social. Networks. **27**, 55–71 (2005)
44. T Opsahl, F Agneessens, J Skvoretz, Node centrality in weighted networks: generalizing degree and shortest paths. Social. Networks. **32**, 245–251 (2010)
45. K Steinhaeuser, NV Chawla, Community detection in a large real-world social network, in *Social computing, behavioral modeling and prediction*, ed. by H Liu, JJ Salerno, MJ Young (Springer US, New York, 2008)
46. TAB Snijders, K Nowicki, Estimation and prediction for stochastic blockmodels for graphs with latent block structure. J. Classification. **14**, 5–100 (1997)
47. M Kim, J Leskovec, Modeling social networks with node attributes using the multiplicative attribute graph model, in *Proceedings of the twenty-seventh conference on uncertainty in artificial intelligence, July 14–17, Barcelona, Spain*, ed. by FG Cozman, A Politecnica, A Pfeffer (AUAI Press, Corvallis, OR, 2011)
48. J Ayling, Criminal organizations and resilience. Int. J. Law. Crime. Justice. **37**, 182–196.31 (2009)
49. Yuan J, Cao J, Xia B. Arresting strategy based on dynamic criminal networks changing over time. Discrete Dynamics in Nature and Society. 2013;2013. www.hindawi.com/journals/ddns/2013/296729/. Accessed 29 Aug 2013.
50. C Morselli, Assessing vulnerable and strategic positions in a criminal network. J Contemporary. Criminal. Justice. **26**, 382–392 (2010)
51. D Penzar, A Srbljinović, About modeling of complex networks with applications to terrorist group modeling. Interdisciplinary. Description. Complex Systems. **3**, 27–43 (2005)
52. W de Nooy, A Mrvar, V Batagelj, *Exploratory social network analysis with Pajek*, 2nd edn. (Cambridge University Press. Software freely, Cambridge, U.K., 2011). available from http://mrvar.fdv.uni-lj.si/pajek/be2.htm. Accessed 26 March 2015
53. GMS Namata Jr, L Getoor, Identifying graphs from noisy and incomplete data. ACM. SIGKDD. Explorations. Newsletter. **12**, 33–39 (2010)
54. M Fire, R Puzis, Y Elovici, Link prediction in highly fractional data sets, in *Handbook of computational approaches to counterterrorism*, ed. by VS Subrahmanian (Springer, New York, 2013)
55. R Guimerà, M Sales-Pardo, Missing and spurious interactions and the reconstruction of complex networks. Proc. Natl. Acad. Sci. U. S. A. **106**, 22073–22078 (2009)
56. R Thomas, Surge in rhinoceros poaching in South Africa. TRAFFIC. Bulletin. **23**, 3 (2010)
57. RH Emslie, M Brooks, *African rhino: Status, survey and conservation action plan* (IUCN/SSC African Rhino Specialist Group, IUCN, Gland, Switzerland, 1999)
58. Ferreira SM, Pfab M, Knight M. Management strategies to curb rhino poaching: an exploration of alternative options. South African Journal of Science. 110, 5/6 (2014). Article #2012-0055, 8 pages. http://dx.doi.org/10.1590/sajs.2014/20120055. Accessed 26 March 2015.
59. Jacomy M, Heymann S, Venturini T, Bastian M. ForceAtlas2: A continuous graph layout algorithm for handy network visualization. 2012. http://www.medialab.sciences-po.fr/publications/Jacomy_Heymann_Venturini-Force_Atlas2.pdf. Accessed 2 Jan 2014.
60. Haas TC. Analyzing wildlife trafficking criminal networks. Online resource. 2014. http://www4.uwm.edu/people/haas/sna. Accessed 19 Jan 2015.
61. S Hick, C Grubb, *Regional crime analysis data-sharing with ArcIMS: Kansas city regional crime analysis GIS* (Proceedings of the Twenty-Third Annual ESRI International User Conference, San Diego, United States, 2003). http://proceedings.esri.com/library/userconf/proc03/p1079.pdf. Accessed 19 Jan 2015
62. Softonic. MySQL Free Download. 2014. http://mysql.en.softonic.com/download. Accessed 20 Jan 2015.
63. Daemen J, Rijmen V. The Rijndael Block Cipher. AES Proposal. 2003. http://csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf. Accessed 20 Dec 2013.
64. M Dehmer, F Emmert-Streib, Comparing large graphs efficiently by means of feature vectors. Appl. Mathematics. Comput. **188**, 1699–1710 (2007)
65. Ntrepid. Tartan Plug-In for Analyst's Notebook. www.ntrepidcorp.com/tartan/analysts-notebook-plugin.php. 2013. Accessed 2 Dec 2013.
66. FMS. Sentinel visualizer provides advanced link analysis, data visualization, geospatial mapping, and social network analysis (SNA). 2014. www.fmsasg.com. Accessed 6 Jan 2014.