Drug Trafficking in Relation to Global Shipping Network

Leibbrandt, Louise ; Zhang, Shilun; Roelvink, Marijn; Bergkamp, Stan ; Li, Xinqi; Bisschop, Lieselot; Wingerde, Karin van; Wang, Huijuan

DOI
10.1007/978-3-031-21131-7_52

Publication date
2023

Document Version
Final published version

Published in
Complex Networks and Their Applications XI - Proceedings of The 11th International Conference on Complex Networks and Their Applications

Citation (APA)
Leibbrandt, L., Zhang, S., Roelvink, M., Bergkamp, S., Li, X., Bisschop, L., Wingerde, K. V., & Wang, H. (2023). Drug Trafficking in Relation to Global Shipping Network. In H. Cherifi, R. N. Mantegna, L. M. Rocha, C. Cherifi, & S. Micciche (Eds.), *Complex Networks and Their Applications XI - Proceedings of The 11th International Conference on Complex Networks and Their Applications: COMPLEX NETWORKS 2022—Volume 2* (pp. 675-686). (Studies in Computational Intelligence; Vol. 1078). Springer. https://doi.org/10.1007/978-3-031-21131-7_52

Drug Trafficking in Relation to Global Shipping Network

Louise Leibbrandt, Shilun Zhang, Marijn Roelvink, Stan Bergkamp, Xinqi Li, Lieselot Bisschop, Karin van Wingerde, and Huijuan Wang

Abstract This paper aims to understand to what extent the amount of drug (e.g., cocaine) trafficking per country can be explained and predicted using the global shipping network. We propose three distinct network approaches, based on topological centrality metrics, Susceptible-Infected-Susceptible spreading process and a flow optimization model of drug trafficking on the shipping network, respectively. These approaches derive centrality metrics, infection probability, and inflow of drug traffic per country respectively, to estimate the amount of drug trafficking. We use the amount of drug seizure as an approximation of the amount of drug trafficking per country to evaluate our methods. Specifically, we investigate to what extent different methods could predict the ranking of countries in drug seizure (amount). Furthermore, these three approaches are integrated by a linear regression method in which we combine the nodal properties derived by each method to build a comprehensive model for the cocaine seizure data. Our analysis finds that the unweighted eigenvector centrality metric combined with the inflow derived by the flow optimization method best identifies the countries with a large amount of drug seizure (e.g., rank correlation 0.45 with the drug seizure). Extending this regression model with two extra features, the distance of a country from the source of cocaine production and a country's income group, increases further the prediction quality (e.g., rank correlation 0.79). This final model provides insights into network derived properties and complementary country features that are explanatory for the amount of cocaine seized. The model can also be used to identify countries that have no drug seizure data but are possibly susceptible to cocaine trafficking.

Keywords Drug trafficking · Drug seizure · Shipping network · Network method

L. Leibbrandt · S. Zhang · M. Roelvink · X. Li · H. Wang (✉)
Faculty of Electrical Engineering, Mathematics and Computer Science,
Delft University of Technology, Delft, The Netherlands
e-mail: H.Wang@tudelft.nl

S. Bergkamp
Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

L. Bisschop · K. van Wingerde
Erasmus School of Law, Erasmus University Rotterdam, Rotterdam, The Netherlands

H. Cherifi et al. (eds.), *Complex Networks and Their Applications XI*,
Studies in Computational Intelligence 1078,
https://doi.org/10.1007/978-3-031-21131-7_52

1 Introduction

Complex networks have been widely used to represent real-world complex systems, where nodes denote the components and links represent relations or interaction between components. Significant contributions have been made to characterize complex networks and to understand the effect of a network on a dynamic process unfolding on the network. Diverse nodal centrality metrics [1, 2] have been proposed to measure various topological properties of a node. Nodal centrality metrics have been applied in general to estimate the importance of nodes in their function, e.g., to identify nodes with high spreading capacity and to select nodes to be immunized when a virus is prevalent [3–5]. Spreading models such as Susceptible-Infected (SI) model and Susceptible-Infected-Susceptible (SIS) model have been intensively studied [6, 7] to model the spread of epidemic and information on networks. Deep understanding has been achieved regarding how the underlying network topology affects a spreading process, how to predict and control a spreading process on a network [8].

Shipping networks play a crucial role in world trade as around 80% of global trade by volume is carried by sea.[1] From the aspect of network topology, prior literature explored the overall structure of the global shipping network, revealing its scale-free property [9] and modular structure [10]. Li et al. investigated the relationship between the centrality of nodes in the global shipping network and the economy of corresponding areas [11]. Global shipping network has been found to have "economic small-world" characteristic [12], i.e., with high transportation efficiency and low wiring cost. Some other efforts have been devoted to model dynamic processes on shipping networks, e.g., marine species invasion process [13, 14]. Nonetheless, how illicit trafficking, like drug trafficking, is linked to the shipping network from the angle of network science remains unexplored.

In this paper, we investigate how to explain and predict drug (e.g., cocaine) trafficking using the global shipping network. The amount of drug seizure in each country is used as an approximation of the amount of drug trafficking to evaluate our methods. We propose three types of network-based methods. These methods are evaluated via their capability to predict the ranking countries in drug seizure (amount) thus to identify countries with a large drug seizure. The first method uses traditional nodal centrality metrics of a country in the shipping network to estimate the volume of drugs seized in the country [15]. Secondly, we employ the SIS spreading model on the shipping network. The infection probability of a node (country) in the metastable state is derived to indicate a country's drug seizure. In the third method, we formulate drug trafficking as the optimal flow on the shipping network, where the number of links to route the traffic from countries that produce drugs and countries that consume drugs is minimized. The inflow to a country is used to estimate the drug seizure of that country. We finally combine the above three methods using linear regression, showing a better prediction of the ranking of countries in drug seizure and identifying key factors that explain the amount of drug seizure per country. This

[1] https://unctad.org/webflyer/review-maritime-transport-2018.

linear regression model has also been extended by including two extra country-level properties, namely the distance of a country from the source of cocaine production and a country's income group/level. We find that these extra features could further improve the prediction quality and the essential role of network-based properties.

The paper is structured as follows. Section 2 introduces the construction of the shipping network and drug seizure data. Section 3 describes and evaluates our methods. Section 4 summarizes our key findings.

2 Datasets

Shipping Network Construction. The Global Liner Shipping Network has been derived from service routes data of the world's top 100 liner shipping companies in 2015, by mapping each service route as a complete graph where any two ports in the service route were connected via a link [16, 17]. It is composed of 977 unique ports and 16,680 inter-port connections. We construct the country-level shipping network as follows. Based on the country code of each port extracted from the Marine Traffic ports database [18], each port can be mapped to the country it belongs to. In the unweighted shipping network, nodes are the countries, and two nodes i and j are connected by a link, i.e., $a_{ij} = 1$ in the adjacency matrix A if at least two ports from the two countries respectively have an inter-port connection, otherwise $a_{ij} = 0$. A weighted network can be further constructed by having the same topology as the unweighted network and associating each link with a weight w_{ij} that equals the total number of inter-port connections between countries i and j. The weight $w_{ij} = 0$ if there is no inter-port connection between i and j. Both networks have $N = 174$ countries and $L = 2743$ links in 2015 and are relatively stable over time. The unweighted network is visualized in Fig. 1.

Drug Seizure Data. The drug seizure data is from the United Nations Office on Drugs and Crime annual drug seizures report [15]. The data contains reports for 144 unique countries between the years 2012–2016. We extract all entries that pertain to the drug group of *Cocaine-type* and can reliably be converted to kilogram equivalents. The average amount of drug seizures per year per country over 2012–2016 is considered. There is an overlap of $N = 110$ countries between the shipping network and the drug seizure data.

Hence, we consider the sub-shipping weighted and unweighted networks, that contain these 110 common countries as nodes and their connections in this paper. Both shipping networks contain $N = 110$ countries and $L = 1794$ links and the average (standard deviation) of link weights is 6.1 (13.4) in the weighted network.

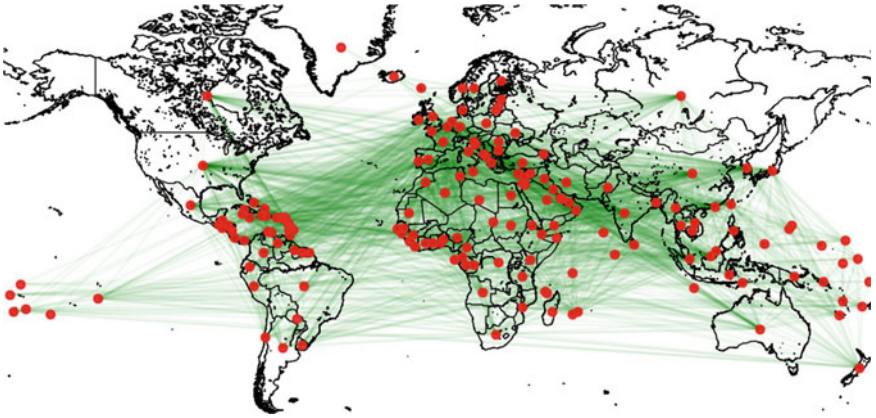

Fig. 1 Unweighted global liner shipping network. Each node represents a country (or district)

3 Methods

Firstly, we propose three distinct network based approaches to model the drug seizure. Each approach uses the shipping network as main ingredient and creates an output nodal property that is meant to be representative of the drug seizure of a country.

We evaluate the quality of all approaches in predicting the ranking of countries in drug trafficking/seizure amount via two measures: Spearman's rank correlation and the recognition rate between an output nodal property and drug seizure of a country. The top f recognition rate $r_{\phi v}(f)$ is defined as

$$r_{\phi v}(f) = \frac{|R_f^{\phi} \cap R_f^{v}|}{|R_f^{\phi}|},$$

where R_f^{ϕ} is the set of f percentage of countries that have the highest drug seizure amount and R_f^{v} is the set of f percentage of countries that have the highest (lowest[2]) value in property v. The recognition rate $r_{\phi v}(f)$ equals the size of the overlap between R_f^{ϕ} and R_f^{v}, or the number of common nodes, normalized the size of each set $|R_f^{\phi}| = |R_f^{v}| = Nf$. We consider $f = 10$, 20 and 40% as examples. Using these two measures, we aim to understand the capability of our approaches in identifying the countries with a high drug seizure.

[2] When drug seizure and the nodal property are supposed to be negatively correlated.

3.1 Topological Metrics

Our first approach uses each nodal centrality metric derived from the shipping network to estimate the drug seizure per country. We are interested in which metric has the highest correlation with drug seizure and is therefore most representative of drug seizure. We will briefly introduce our four chosen centrality metrics, our reasoning and hypotheses for their correlation with drug seizure.

- *Degree* d_i of a node i is the number of links incident to the node i. The degree of a country represents the number of countries directly connected to it in the shipping network.
- *Betweenness Centrality* b_i of a node i in the unweighted shipping network is the number of shortest paths that traverse the node i between all possibly node pairs [19, 20].
- *Clustering Coefficient* c_i of a node i equals the number of links among the neighbors of the node normalized by $\binom{d_i}{2}$. It tells the link density among the d_i neighbors of node i.
- *Eigenvector Centrality*. The unweighted (weighted) eigenvector centrality of a node is the principal eigenvector component and the principal eigenvector is the eigenvector corresponds to the largest eigenvalue of the unweighted (weighted) adjacency matrix A (W). A country with a high eigenvector centrality tends to connect to many countries who themselves have a high connection (eigenvector centrality).

We suspect that countries with a high degree, betweenness or Eigenvector centrality are susceptible to large amount of drug trafficking due to their good network connection, infrastructure and the corresponding flexibility to change modus operandi (e.g., shifting drugs to another port). In contrary, countries with a large clustering coefficient could be less susceptible since they do not have a large degree and drug trafficking can be through their mutually connected neighbors without the need of traversing the country itself.

Results. The Spearman's rank correlation and recognition rate between the drug seizure and each chosen centrality metric is given in Table 1. We find that all metrics have a statistically significant ($p < 0.01$) rank correlation with drug seizure, and the sign of each correlation is in line with our hypothesis for the corresponding metric. All centrality metrics can contribute to the identification of countries with large seizure since their recognition rate $r_{\phi v}(f) > f$, thus is better than that of a random guess. We find that the unweighted eigenvector centrality metric provides the highest correlation in strength, however the weighted eigenvector centrality and betweenness leads to the highest recognition rate when $f = 20\%$ and $f = 40\%$ respectively.

Table 1 Evaluation of prediction performance of topological metrics, SIS infection probability, and inflow (derived from the flow optimization model), via rank correlation and recognition rate

Metric	Corr.	p-value	$r_{\phi v}(10\%)$	$r_{\phi v}(20\%)$	$r_{\phi v}(40\%)$
Degree	0.39,	< 0.01	0.27	0.41	0.57
Betweenness	0.34,	< 0.01	0.27	0.36	**0.61**
Clustering	−0.35,	< 0.01	0.27	0.36	0.60
Eigenvector (unweighted)	**0.40**	< 0.01	0.27	0.41	0.59
Eigenvector (weighted)	0.39	< 0.01	0.27	**0.50**	0.52
SIS infection prob. (unweighted net) $\tau = 0.045$	**0.41**	< 0.01	0.27	0.41	0.59
SIS infection prob. (weighted net) $\tau = 0.021$	0.40	< 0.01	0.27	0.41	0.59
Inflow	0.33	< 0.01	0.36	0.36	0.57

The best performance of each category is in bold

3.2 SIS Spreading Process

We have shown recently that the SIS epidemic spreading model can be generalized to model the contagion of traffic congestion at an airport on an airline network and the infection probability of an airport can be used to estimate the probability of congestion at the airport [21]. Inspired by this, we propose our second approach: model the contagion of drug trafficking as an SIS spreading process on the shipping network and uses the infection probability of a country in the meta-stable state to estimate the ranking of countries in drug seizure. We will briefly introduce the SIS model, method to derive infection probabilities and evaluate this approach.

SIS Model. The homogeneous SIS model is defined as follows. At any time t, a node is either susceptible or infected. A susceptible node can be infected by each of its infected neighbors with an infection rate β and each infected node recovers to be susceptible again with a recovery rate δ. Both the infection and recovery processes are independent Poisson processes. For a given network upon which the SIS process is deployed, a critical epidemic threshold τ_c exists. When the effective infection rate $\tau = (\beta/\delta) > \tau_c$, a non-zero fraction of infected nodes persists in the meta-stable state. When $\tau < \tau_c$, the epidemic dies out.

Mean-Field Approximation of SIS Model. We derive nodal infection probabilities in the meta-stable state via N-Intertwined Mean-Field Approximation (NIMFA) [22]. NIMFA is chosen as it preserves the network topology in its governing equations, coupling the infection probability of neighboring nodes. It assumes that the states of neighboring nodes are uncorrelated. Under NIMFA, the governing equation for a node i in our heterogeneous SIS spreading model is

$$\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = -\delta v_i(t) + (1 - v_i(t)) \sum_{j=1}^{N} \beta w_{ij} v_j(t), \tag{1}$$

where $v_i(t)$ is the infection probability of node i at time t, and βw_{ij} is the infection rate associated to the link (i, j) with weight w_{ij}. The time derivative of the infection probability $v_i(t)$, is determined by two competing processes (a) while node i is infected, node i is cured at rate δ and (b) while node i is susceptible, each infected neighbour j infects node i with a rate βw_{ij}. In the meta-stable state, $\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = 0$, for any $i \in \mathcal{N}$. Hence, the infection probability of each node in this case, $v_{i\infty}$, can be derived by solving the equations $\delta v_{i\infty} + (1 - v_{i\infty}) \sum_{j=1}^{N} \beta w_{ij} v_{j\infty} = 0$. The trivial all-zero solution corresponds to the absorbing state where all nodes are susceptible.

Both the unweighted and weighted shipping networks will be considered. When the underlying network is the unweighted, the governing Equation (1) should be updated by replacing the weighted adjacency matrix element w_{ij} by the unweighted adjacency matrix element a_{ij}.

Results. In Fig. 2, the Spearman's rank correlation and recognition rate between the infection probability and drug seizure amount are plotted as a function of the effective infection rate τ, when the underlying network is the unweighted shipping network. The Spearman's rank correlation varies only slightly when the effective infection rate is small around the critical epidemic threshold $\tau_c = 0.020$. This is in line with the theoretical and empirical finding in [23] that the ranking of nodal infection probability tends to change more with τ, when τ is small. The top f recognition rate is insensitive of the effective spreading rate τ. The same trends have been observed when the underlying network is the weighed shipping network. The maximal rank correlation and recognition rate that can be reached by choosing τ (just above the threshold) are summarized in Table 1. We also find the infection probability of a country derived from the unweighted (weighted) shipping network at the optimal effective infection rate, is strongly correlated with the unweighted (weighted) principal eigenvector component derived in Sect. 3.1, with a correlation coefficient around 0.99. This is in line with the theoretical proof that when the effective infection rate is just above the epidemic threshold, the meta-stable state infection probability of a node, obtained by NIMFA is proportional to the principal eigenvector component of the adjacency matrix A (W) [24]. By tuning the effective infection rate β to optimize the evaluation metrics, we improve the performance marginally compared to the principal eigenvector component.

3.3 Flow Optimization Model

For this method we view transnational drug trafficking as an economic process where drugs go from production countries to consumption countries through a chain of intermediaries [25]. This is motivated by two main factors that contribute to Cocaine trafficking. The first is that the price of Cocaine is largely attributable to

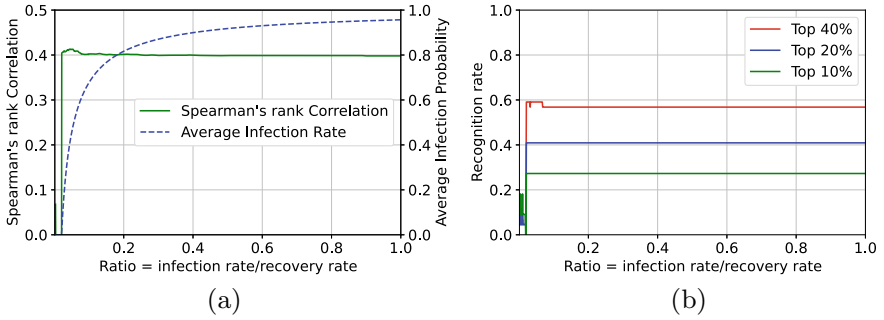

Fig. 2 **a** Spearman's rank correlation (and the average infection probability over all nodes) and **b** recognition rates between nodal infection probability and drug seizure as a function of the effective infection rate. The underlying network is the unweighted shipping network

risk compensation[25]; in an economic system, distributors may try to minimize the number of links used to transport drugs from supply countries to consumption countries. The second factor is that a large portion of Cocaine trafficking uses the shipment network as way of transportation [26]. We therefore create an optimization problem in which we minimize the number of links used to transport drugs in our shipping network whilst adhering to constraints that model the supply and demand for each country. The total inflow at each country will be used to predict the ranking of countries in drug seizure.

The following steps will be taken to derive the (normalized) supply and demand per country. Firstly, we extract the amount of cocaine production s_i of each country i in 2012 from [27] and the drug usage u_i(as % of population) of each country i from [28]. We define the normalized supply of a country i as $s_i^* = \frac{s_i}{\sum_{i=1}^{N^s} s_i}$, where N^s is the number of supply countries. Similarly, the normalized demand of a country i is defined as $u_i^* = \frac{u_i * m_i}{\sum_{i=1}^{N^u} u_i * m_i}$, where m_i is the population of country i and N^u is the number of consumption countries. The total normalized supply (or demand) of all countries is one.

Countries of supply or consumption may not have any port, thus not exist in our shipping networks. Hence, we extend our unweighted shipping network with 174 nodes by including all the supply and consumption countries listed in the data [27] and [28] and adding directed link(s) from (to) a supply (consumption) country that has no port to (from) countries that have a port and share a border with it, using country border dataset [29]. In total, 49 nodes and 130 directed links have been added.

We assume that the flow from the multiple supply countries to the consumption countries follows the paths on the shipping network with the minimal number of links, i.e. $\sum_{i,j} \mathbf{1}_{\{A_{ij} \cdot F_{ij} > 0\}}$, where F_{ij} is the flow from country i to j and the indicator function $\mathbf{1}_{\boldsymbol{x}}$ is one (zero) when the condition $\boldsymbol{x}$ is true (false). For each country, the following constraint must hold:

$$s_i^* + \sum_{j=1}^{N^*} A_{ji} F_{ji} = \sum_{j=1}^{N^*} A_{ij} F_{ij} + u_i^* \tag{2}$$

The constraint ensures the total incoming and outgoing flow combined with the supply and demand in each node is balanced. After the optimization, 68 nodes have a positive inflow or outflow.

Results. We use the final inflow, $F_i = \sum_{j=1}^{N^*} A_{ji} F_{ji}$, per country to estimate ranking of the same set of 110 countries in drug seizure as in previous analysis. The rank correlation and recognition rate between the amount of in-flow and drug seizure of a country can be found in Table 1. We find that whilst this method produces a lower rank-correlation than the methods described in Sects. 3.1 and 3.2 likely due to the zero inflow of many countries, it achieves a significantly higher recognition rate at $f = 10\%$. This implies the complementary nature of the three methods and the potential synergy when combining them.

3.4 Regression Model

We have shown that all three methods (their output nodal properties) contribute to the estimation of countries with the highest amount of drug seizure. In order to explore their synergy and identify the key nodal properties in explaining drug seizure, we build a regression model that uses aforementioned network derived features in combination with other country features. We fit multiple Gaussian linear regression models on the logarithm of the drug seizure data, taking different features as independent variables. We split our analysis into three parts; in Analysis 1 we investigate our network derived features, in Analysis 2 we investigate extra country features, and in Analysis 3 we combine the results from analyses 1 and 2 to propose a final model.

Table 2 provides an overview of the results for each analysis. The adjusted R^2 value indicates the predictive power of each regression model. We also provide Spearman's rank correlation and recognition rate between the drug seizure amounts predicted by each regression model and the actual drug seizure amounts. The correlation and recognition rate in Table 2 for the single feature regression in $A.1$ are in line with the results produced by each method in Sects. 3.1, 3.2 and 3.3 and serve as a baseline for the subsequent models.

Analysis 1. We start with regression models that use each output property derived in Sects. 3.1, 3.2 and 3.3 as the single feature. We find that the unweighted Eigenvector Centrality (EC), derived in Sect. 3.1, results in the highest scoring R^2 value and therefore provides the best fit over all network based methods. Betweenness and clustering coefficient perform worse than EC and are not considered in this analysis. The Infection Probability (IP), Eigenvector Centrality (EC) on weighted and unweighted networks are strongly correlated and the unweighted EC performs the best among these four properties. Hence, we consider the combination of the unweighted EC and

Table 2 Results for regression analyses. Country level features considered include: principal eigenvector component of the unweighted shipping network (EC unweighted), of the weighted network (EC weighted), infection probability (IP) in unweighted and unweighted network respectively, inflow (IF), distance to source countries (D) and the Income Group (IG)

	Formula for regression	Adj. R^2	Corr.	$r_{\phi v}(10\%)$	$r_{\phi v}(20\%)$	$r_{\phi v}(40\%)$
A.1	(3.1) EC unweighted	0.13	0.40	0.27	0.41	0.59
	(3.1) EC weighted	0.10	0.39	0.27	0.50	0.52
	(3.2) IP unweighted	0.11	0.41	0.27	0.41	0.59
	(3.2) IP weighted	0.11	0.40	0.27	0.55	0.55
	(3.3) IF	0.07	0.33	0.36	0.36	0.57
	EC_u + IF	**0.16**	**0.45**	**0.36**	**0.55**	**0.64**
A.2	D	0.27	0.55	0.64	0.59	0.64
	IG	0.20	0.41	0.27	0.41	0.43
	D + IG	**0.41**	**0.68**	**0.64**	**0.50**	**0.77**
A.3	**EC unweighted + IF + D + IG**	**0.55**	**0.79**	**0.73**	**0.82**	**0.82**

The model that performs the best in each of the three categories are in bold

the inflow (IF). Whilst the total country Inflow (IF), derived in Sect. 3.3, does not provide relatively high rank correlation, the combined regression of unweighted EC and IF performs better than that of a single network based property, revealing that these two features contain cumulative predictive power.

Analysis 2. Besides network based features, we investigate another two key country level features: the Distance (D) of a country from the countries where cocaine is produced and the Income Group (IG) of a country, because they are known to influence cocaine seizure rates [26]. For the Distance (D), we take the geodesic distance in kilometers between the coordinate of each country and the central coordinate of the source countries. The betweenness and D are weakly correlated (rank correlation 0.14). We also extract the Income Group (IG) of each country [30] in 2016; this assigns each country as either a *low*, *lower middle*, *upper middle* or *high* income country.

We find that both distance and Income Group have a high predictive power for drug seizure, and distance is an especially effective method to obtain high recognition rates, i.e. $r_{\phi v}(10\%) = 0.64$. The combined predictive power of the two country features is strong and this model is able to outperform the best model from Analysis 1 in all but one, $r_{\phi v}(20\%)$, of the evaluation metrics.

Analysis 3. For our final investigation, we combine the best performing models from Analysis 1 and 2. A correlation analysis reveals low rank correlation amongst EC, IF, D and IG (< 0.150). Therefore, we conclude that each feature provides unique information and should be used in the final model. This model outperforms all the other models. It is able to correctly predict 73% of the countries that are in the top 10% of the highest scoring countries based on drug seizure amounts, and 82% of the countries in both the top 20% and 40% lists. This observation highlights the essential role of both network based properties and the two country level properties in estimating drug seizure.

4 Conclusion

In this work, we propose different methods to explain and predict the amount of cocaine trafficking/seized per country using the global shipping network. We firstly propose three distinct network-based approaches: a centrality metrics based, an SIS spreading process based and a flow optimization model based approach. Correspondingly, the derived centrality metrics, infection probabilities, and inflow of drug traffic of nodes are used to estimate the ranking of countries in drug seizure. Furthermore, a linear regression analysis is designed to investigate the cumulative power of each approach with other country features.

We find that each approach with its output nodal property could contribute to the estimation of the ranking of countries in drug seizure. The eigenvector centrality of the unweighted shipping network seems to perform the best, also in view of the amount of data needed. Furthermore, the regression analysis reveals that inflow derived from the optimization model contains unique information when compared with the eigenvector centrality; combining in-flow and unweighted eigenvector centrality results in an evidently better estimation of drug seizure. We end our investigation by showing the benefit of combining our network-based results with other features, e.g., a country's distance to the source of cocaine and its income group.

The identification of countries that are susceptible to a large amount of cocaine trafficking is crucial in stopping the illicit substance from reaching its destination. Our final proposed model provides a starting point to tackling this problem. For countries that have no drug seizure data, the model can be used to predict their amount of drug trafficking. Furthermore, our methods can be applied or extended for drug trafficking of other drug groups such as amphetamine-type stimulants and opioids on a multi-layer network that includes diverse transportation modes. Our method could be further developed to distinguish between drug producing and non-producing countries, between domestic and trade/port related seizures and to understand whether the difference in the predicted and actual ranking of countries in drug seizure could be explained by the variance in law enforcement efforts of countries and other social and political factors.

References

1. Li, C., Li, Q., Van Mieghem, P., Eugene Stanley, H., Wang, H.: Correlation between centrality metrics and their application to the opinion model. Euro. Phys. J. B **88**(3), 1–13 (2015)
2. Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., Zhou, T.: Vital nodes identification in complex networks. Phys. Rep. **650**, 1–63 (2016)
3. Pastor-Satorras, R., Vespignani, A.: Immunization of complex networks. Phys. Rev. E **65**(3), 036104 (2002)
4. Kitsak, M., Gallos, L.K., Havlin, S., Liljeros, F., Muchnik, L., Eugene Stanley, H., Makse, H.A.: Identification of influential spreaders in complex networks. Nat. Phys. **6**(11), 888–893 (2010)
5. Borge-Holthoefer, J., Moreno, Y.: Absence of influential spreaders in rumor dynamics. Phys. Rev. E **85**(2), 026116 (2012)

6. Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A.: Epidemic processes in complex networks. Rev. Mod. Phys. **87**(3), 925 (2015)
7. Kiss, I.Z., Miller, J.C., Simon, P.L., et al.: Mathematics of Epidemics on Networks, p. 598. Springer, Cham (2017)
8. Zhang, S., Zhao, X., Wang, H.: Mitigate sir epidemic spreading via contact blocking in temporal networks. Appl. Netw. Sci. **7**(1), 1–22 (2022)
9. Kaluza, P., Kölzsch, A., Gastner, M.T., Blasius, B.: The complex network of global cargo ship movements. J. R. Soc. Interface **7**(48), 1093–1103 (2010)
10. Pan, J.-J., Bell, M.G.H., Cheung, K.-F., Perera, S., Yu, H.: Connectivity analysis of the global shipping network by eigenvalue decomposition. Maritime Policy Manage. **46**(8), 957–966 (2019)
11. Li, Z., Mengqiao, X., Shi, Y.: Centrality in global shipping network basing on worldwide shipping areas. GeoJournal **80**(1), 47–60 (2015)
12. Xu, M., Pan, Q., Muscoloni, A., Xia, H., Cannistraci, C.V.: Modular gateway-ness connectivity and structural core organization in maritime network science. Nat. Commun. **11**(1), 1–15 (2020)
13. Xu, J., Wickramarathne, T.L., Chawla, N.V., Grey, E.K., Steinhaeuser, K., Keller, R.P., Drake, J.M., Lodge, D.M.: Improving management of aquatic invasions by integrating shipping network, ecological, and environmental data: data mining for social good. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1699–1708 (2014)
14. Saebi, M., Xu, J., Grey, E.K., Lodge, D.M., Corbett, J.J., Chawla, N.: Higher-order patterns of aquatic species spread through the global shipping network. Plos One **15**(7), e0220353 (2020)
15. UNODC: Annual Drug Seizure Report. https://dataunodc.un.org/drugs/seizures, 2012–2016. Online. Accessed 20 Mar 2022
16. Kojaku, S., Mengqiao, X., Xia, H., Masuda, N.: Multiscale core-periphery structure in a global liner shipping network. Sci. Rep. **9**(1), 1–15 (2019)
17. Xu, M.: Resource. http://www.mengqiaoxu.com/resource.html, 2019. Online. Accessed 28 Mar 2022
18. Marine Traffic. Ports Database. https://www.marinetraffic.com/en/data/, 2022. Online. Accessed 28 Mar 2022
19. Anthonisse, J.M.: The rush in a directed graph. In: Stichting Mathematisch Centrum. Mathematische Besliskunde, (BN 9/71) (1971)
20. Wang, H., Hernandez, J.M., Van Mieghem, P.: Betweenness centrality in a weighted network. Phys. Rev. E **77**(4), 046105 (2008)
21. Ceria, A., Köstler, K., Gobardhan, R., Wang, H.: Modeling airport congestion contagion by heterogeneous sis epidemic spreading on airline networks. Plos One **16**(1), e0245043 (2021)
22. Van Mieghem, P., Omic, J., Kooij, R.: Virus spread in networks. IEEE/ACM Trans. Netw. **17**(1), 1–14 (2008)
23. Bo, Q., Li, C., Van Mieghem, P., Wang, H.: Ranking of nodal infection probability in susceptible-infected-susceptible epidemic. Sci. Rep. **7**, 9233 (2017)
24. Van Mieghem, P.: Performance Analysis of Complex Networks and Systems. Cambridge University Press (2014)
25. Boivin, R.: Drug trafficking networks in the world-economy. Crime Netw. (2013)
26. UNODC and EUROPOL. The illicit trade of cocaine from latin america to europe—from oligopolies to freefor-all. Cocaine Insights 1, Sept 2021
27. UNODC. Statistical Annex: 6.1.3 Cocaine manufacture. https://www.unodc.org/unodc/en/data-and-analysis/wdr2021_annex.html, 2019. Online. Accessed 20 Mar 2022
28. UNODC. Statistical Annex: 1.2 Prevalence of drug use in the general population, including NPS—national data. https://www.unodc.org/unodc/en/data-and-analysis/wdr2021_annex.html, 2019. Online. Accessed 20 Mar 2022
29. Geo Data Source. Country Borders. https://www.geodatasource.com/addon/country-borders, 2022. Online. Accessed 20 Mar 2022
30. The World Bank. World Development Indicators. https://data.worldbank.org/indicator, 2020. Online. Accessed 03 May 2022