Proc. 2018 IEEE International Symposium on Technologies for Homeland Security (HST'18),Woburn, MA, Oct. 23-24, 2018
received the Best Paper Award in the Land/Maritime Borders and Critical Infrastructure Protection Track

Integrative Analytics for Detecting and Disrupting Transnational Interdependent Criminal Smuggling, Money, and Money-Laundering Networks

Ashwin Bahulkar
Rensselaer Polytechnic Institute
Troy, NY, USA
bahula@rpi.edu

N. Orkun Baycik
Shenandoah University
Winchester, VA, USA
nbaycik@su.edu

Thomas Sharkey
Rensselaer Polytechnic Institute
Troy, NY, USA
sharkt@rpi.edu

Yeming Shen
Rensselaer Polytechnic Institute
Troy, NY, USA
sheny15@rpi.edu

Boleslaw Szymanski
Rensselaer Polytechnic Institute
Troy, NY, USA
szymab@rpi.edu

William Wallace
Rensselaer Polytechnic Institute
Troy, NY, USA
wallaw@rpi.edu

***Abstract*—In this paper, we describe a framework that integrates descriptive, predictive, and prescriptive analytics that aids detecting and disrupting a transnational criminal organization (TCO) operating as interdependent contraband smuggling, money, and money laundering networks. This type of TCO will smuggle contraband across the U.S. border, generate revenues from illegal sales within the U.S., and then use the money laundering network to send the money out of the U.S. Law enforcement may have partial information about the underlying social network of the TCO but this may be missing important, intentionally hidden connections between the criminals. The proposed framework predicts the missing links in the social network data and then algorithms are applied to the augmented data to detect the communities of the TCO. Each community serves a different role in the TCO and thus are necessary in modeling the operations of the organization. Once the communities are identified, we prescribe actions that allocate resources to disrupt the TCO operations optimally in terms of law enforcement criteria.**

***Index Terms*—community detection, transnational criminal organizations, interdependent networks, network interdiction**

This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number, 2017-ST-061-CINA01 and the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053 (the ARL Network Science CTA). The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

I. Introduction

The focus of this paper is on describing an integrative framework across descriptive, predictive, and prescriptive analytics that will help to detect and disrupt a transnational criminal organization (TCO) which operates through interdependent contraband smuggling, money, and money laundering networks. This type of TCO will smuggle contraband across the U.S. border, generate revenues from illegal sales within the U.S., and then use the money laundering network to send the money out of the U.S. This framework leverages incomplete social network data about the TCO in order to make predictions about missing links within the data. These predictions help to augment the social network data on the TCO in order to apply algorithms that better describe the roles and to detect communities of different criminals within the organization. Law enforcement can then use this description to investigate a subset of the criminals to understand the smuggling, money, and laundering flows between them. This understanding of the TCO will then allow us to prescribe decisions regarding prioritizing the allocation of resources to optimally disrupt the operations of the TCO according to law enforcement criteria. An overview of the integrative framework is provided in Figure 1. In this paper, we discuss how these

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

three phases of analytics can be integrated in order to detect and disrupt interdependent criminal smuggling, money, and money laundering networks as well as providing an overview of recent advances in detecting and disrupting TCOs. To the best of our knowledge, we are the first to describe an integrative framework across predictive, descriptive, and prescriptive analytics for detecting and disrupting TCOs. This framework could be useful to law enforcement agents and analysts within the homeland security enterprise, especially ones that are focused on border security.

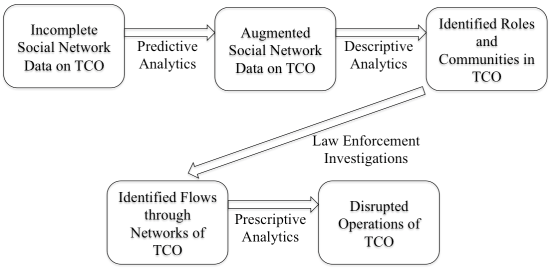

Fig. 1. An overview of the integrative framework.

II. Analytics for Detecting TCO Operations: Link Prediction and Community Detection

Criminal activities are carried out by groups of collaborating individuals; we will refer to such groups as communities. These groups will often take on different roles within a TCO; for example, one group may be responsible for smuggling contraband across an international border, another group may be responsible for collecting revenue from these smuggling efforts, and a third group may be responsible for laundering the money so that it can move across an international border or enrich the TCO bosses. Membership in these communities may overlap when the same individual is involved in multiple crime activities. However, traditional clustering methods detect non-overlapping communities. The communities detected by traditional clustering algorithms may also be unstable and difficult to replicate, because traditional methods are sensitive to noise and attempts by criminals to hide their activities and connections. Therefore, in our detection analytics we first apply novel predictive capabilities to the incomplete social network data gathered about the TCO in order to predict hidden links about criminal connections; for more details on these types of algorithms, we refer the reader to [1]. We then apply the Louvain community detection algorithm [2] and SpeakEasy [3] in order to detect communities within the TCO.

For example, we consider the Caviar data set (see [4], [5] for details on this data set, which is also publicly available through the UCINET software website, along with other covert social networks) that describes criminals responsible for smuggling hashish and cocaine into Montreal, Canada in 1994-1996. This data set provides all the phone calls between different members of the TCO over 11 two-month periods. It should be noted that in each period, a drug seizure was performed by law enforcement *but no arrests were made*. We then run an intuitive link prediction scheme: amongst all pairs of criminals, we compute a measure, and then rank the criminal pairs according to this measure. For example, we could measure the number of common neighbors between the pair of criminals or measure the weighted number of edges to common neighbors (if each edge has a weight). Once we have computed a measure for every pair, we can then place edges into the network that have the n highest rankings. For this analysis, we have selected n to be equal to 25% of the original number of edges in the network. This intuitive link prediction scheme can have significant impacts on the ability to accurately capture the various communities within the TCO. We illustrate this impact for period 6 of the Caviar data set in [4] in Figure 2 (for the social network data before performing link prediction) and Figure 3 (for the network after performing link prediction and adding these edges to the data). In each of these figures, each detected community is identified with a separate color and, therefore, there are two communities detected on the incomplete network and four communities detected on the augmented network.

This TCO consists of four communities: cocaine smuggling, hashish smuggling, money operations, and transportation. As seen in Figure 2 and 3, the community structure of this criminal network becomes much more visible if we were to apply the community detection algorithm to the network data augmented with the predicted edges. More

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

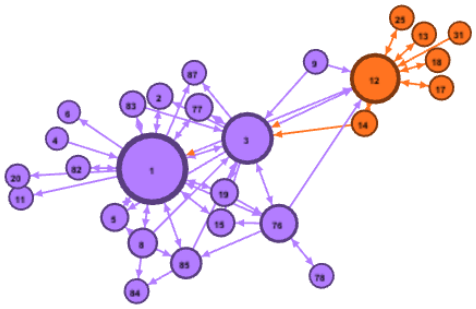

Fig. 2. The communities on the incomplete network.

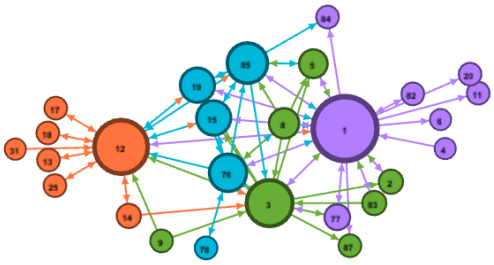

Fig. 3. The communities after performing link prediction.

detailed experiments and analysis on link prediction for community detection can be found in [1], [5], [6]. Note that both the prediction and detection algorithms run without knowledge about the number of communities within this data. As seen in Figure 3, the four communities of the TCO clearly emerge in the output of the detection algorithm run on the augmented network.

From a law enforcement perspective, the impact of this community detection is that it would require less investigative resources to understand which criminals have what roles within the TCO. For example, if we identified that two criminals in the same community both belong to the money network, then other members that were detected to be in the same community as those two criminals would likely have a role in the money network. Therefore, investigating future criminals within this community could focus on verifying their status as belonging to the money network. Once these roles are understood, along with the connections between the criminals within the TCO, we can then prescribe actions to best disrupt the TCO.

III. Analytics for Disrupting TCO Operations: Interdicting Interdependent Networks

In terms of disrupting TCO operations, it is the goal of law enforcement to make interdiction decisions that optimally limit the ability of the TCO to operate effectively. *Interdiction decisions* are broadly defined to be any activity that alters the behavior of the TCO which could include, but are not limited to, arresting criminals, installing checkpoints on border crossings that are believed to be smuggling hotspots, or better monitoring of financial records of certain corporations that are suspected to be linked to money laundering. The analytics for disrupting TCO operations can be used to understand how to best use scarce law enforcement resources to disrupt TCO operations while specifically accounting for the fact that disrupting one network (e.g., money laundering) can cause cascading disruptions to other networks (e.g., by disrupting the flow of money back across the border, we limit the ability of the TCO to smuggle contraband across it in the future). Further, these analytics can be used to understand the amount of resources required to reach a desired level of disruption to the TCO.

In order to prescribe how to disrupt the TCO, we must first discuss the operations of the TCO. The focus of this study is on a TCO whose operations rely on multiple interdependent networks: a contraband smuggling network, a money network, and a money laundering network. The contraband smuggling network models the flow of illegal goods between the criminals, the money network models the flow of money between the criminals, and the money laundering network models the activities of criminals and organizations who legitimize the money earned through these operations. It is important to note that the laundering network may represent the only way to move money generated by the smuggling operations across international borders. The operations of these three networks are interdependent: (i) the money flowing into the money network is from the revenue of the physical smuggling network; (ii) the money flows back to support the operations of the physical smuggling network; (iii) money can flow from the money network to the laundering

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

network; and (iv) the laundered money flow back to criminals in the other networks. By studying these interdependencies, law enforcement can better understand the structure of the TCO and how interdictions within one network can cause cascading disruptions in the other networks. This is similar to the fact that interdependencies in infrastructures can cause cascading disruptions when components in one infrastructure fail [7]. However, the vast majority of the research on network interdiction [8] focuses on a single network and thus ignores the interdependencies that may exist within the operations of the multiple networks of a TCO. We illustrate these networks in Figure 4.

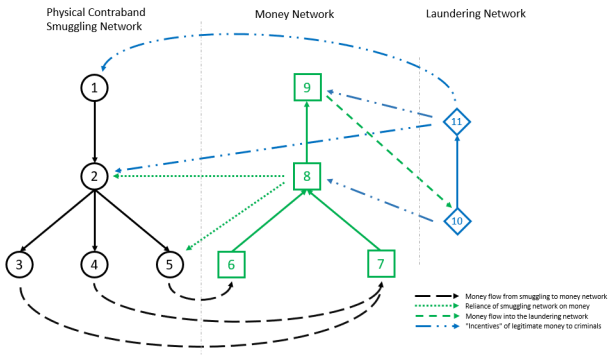

Fig. 4. The relationship between the smuggling, money, and money laundering networks. The nodes represent the criminals or criminal groups, and the edges represent the connections between them.

Figure 4 depicts the relationships and direction of flow between the criminals and between the networks. Each node represents a criminal or a group of criminals, and each edge represents the connection between the criminals or criminal groups. Nodes 1 to 5 operate in the smuggling network, nodes 6 to 9 operate in the money network, and node 10 and node 11 operate in the money laundering network. In the smuggling network, node 1 is the highest ranked criminal and nodes 3 to 5 are the lowest ranked criminals, and the flow of illegal drugs is from the highest ranked criminal to the lowest ranked criminal. In the money network, node 9 is the highest ranked criminal, and node 6 and node 7 are the lowest ranked criminals, and the flow of money is from the lowest ranked criminals to the higher ranked criminals unlike the flow direction in the smuggling network. This pattern follows for the money laundering network, as well. The edges from the nodes in the smuggling network to the money network represent the flow of revenue into the money network. The edges from the money network to the money laundering network represent the flow of illegal money that the money laundering criminals will legitimize in the money laundering network. Finally, the edges from the money laundering network to the other two networks represent the legal money that the criminals will use in smuggling operations or for other purposes. Note that the criminals in the smuggling network can carry out trafficking operations only if node 1 receives money, and the only way to provide money to this criminal is through the laundering network (i.e., through the connection from node 11 to node 1).

Our goal is to allocate law enforcement resources to disrupt these interdependent networks. We use network interdiction approach (see [9] for details on the network interdiction problem) to allocate law enforcement resources to disrupt the TCO. In a network interdiction problem, there exist an attacker and a defender, which make sequential decisions on a network and have conflicting objectives. This problem has a 'bilevel' structure in which the upper-level models the attacker's problem and the lower problem models the defender's problem. In our study, the attacker is law enforcement, and the defenders are the criminals. The studies in [10], [11] apply network interdiction techniques to allocate scarce law enforcement resources against illegal drug trafficking networks where the criminals aim to maximize the amount of illegal drugs delivered to the drug users, and law enforcement seeks to minimize this maximum amount by arresting (interdicting) criminals. Malaviya et al. [10] focus on a city-level drug enforcement problem in which they model the hierarchical structure of the illegal drug trafficking networks. Their approach considers the physical aspect of trafficking operations and provides law enforcement a decision-making tool over multiple periods to disrupt the trafficking operations as much as possible. This paper also discusses a procedure to generate realistic instances of small city-level trafficking networks for crack-cocaine. Baycik et al. [11] extend their study by including the information flows in which the physical operations (i.e., the trafficking of illegal drugs) depend on receiving enough information in regard to trafficking opera-

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

tions. They also examine law enforcement decisions against criminals operating in different cities that are coordinated by the same criminal organization. Our prescriptive analytics will create a bilevel programming model of allocating interdiction resources to disrupt the interdependent smuggling, money, and money laundering networks. If these interdependent networks can be modeled as a linear program, we can then apply dual-based reformulation techniques to solve this bilevel program.

The output of these models will be how to disrupt the networks by specifically accounting for the cascading impacts of interdictions. For example, consider the networks in Figure 4. A strategy to eliminate the smuggling operations is to arrest criminals 3, 4, and 5 (assuming that criminals 1 and 2 operate outside the jurisdiction of law enforcement, for example, outside of the country). Alternatively, arresting cimrinals 6 and 7 would have the same impact since it would cut off the flow of money that supports the smuggling networks. Finally, law enforcement could have the same impact should it implement operations that disrupt the money laundering network and prevent money from reaching criminals 1 and 2 in the smuggling network. Law enforcement could then analyze the amount of required interdiction resources to implement these three strategies and select the one that requires the least amount. Note that operations to disrupt the money laundering network could have wide-reaching impacts, especially if multiple TCOs use the same laundering network.

IV. Conclusions

We have provided an overview of an integrative framework across descriptive, predictive, and prescriptive analytics to disrupt a TCO whose operations require interdependent networks. The predictive analytics help to provide missing information into social network data about the TCO and improves the descriptive capabilities that provide the different communities (roles) within the TCO. The combination of this community data and law enforcement investigations will help to understand the detailed operations of the TCO. This understanding can then be used by our prescriptive analytics to determine how to best disrupt the TCO with limited resources while specifically accounting for the cascading disruptions across the smuggling, money, and laundering networks resulting from interdictions.

There are some potential limitations of the integrative framework that should be discussed. The first limitation is that link prediction may not be necessary for more formally organized criminal networks. For example, [1] demonstrate that there is little room for improvement for Italian gang networks [12]–[14] in community detection. This is because these networks are more hierarchical than the Caviar network.

The second limitation is that the prescriptive analytics, i.e., the interdiction of the TCO, is focused on disrupting a "snapshot" of the activities without necessarily considering the adaptations of the TCO based upon the interdictions. Therefore, the current focus of the prescriptive analytics would be to disrupt the current status quo of the TCO and then law enforcement would need to reapply the integrative framework once the TCO has adapted their operations based on the initial interdictions. It would be of interest to address this limitation by predicting the response of the TCO to selected interdictions.

The third potential limitation has to do with the scalability of the different phases of the integrative framework. For the predictive and the descriptive phases, the link prediction and community detection are computationally fast since the size of criminal networks is relatively small. However, for larger networks, link prediction can be sped up by only calculating prediction scores for nodes that are within a certain distance of one another. Further, the average case computational complexity of the Louvain community detection algorithm is $O(n \log n)$, which is comparable to several other community detection algorithms. For the prescriptive (interdiction) phase, the problem of disrupting a network that seeks to maximize its flow is NP-Complete [9] and disrupting the interdependent networks of a TCO is a generalization of this problem. For layered information and physical flow networks, Baycik et al. [11] demonstrate that empirically standard exact solution approaches tend to fail for networks with about 200 criminals; however, customized reformulation techniques can help to solve problems with up to 2000 criminals. Therefore, it may be necessary

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

to create customized reformulation techniques based on the type of operations of the TCO that is to be disrupted.

REFERENCES

[1] A. Bahulkar, N.O. Baycik, T. C. Sharkey, and B. K. Szymanski, "Community detection by augmentation of networks in criminal networks," Proc. IEEE/ACM Int. Conf. on Adv. in Soc. Net. Ana. and Min. (ASONAM), Barcelona, Spain, pp.1168-1175, 2018.

[2] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," Phys. Rev. E, vol. 70, pp. 1–6, Dec. 2004.

[3] C. Gaiteri, M. Chen, B. Szymanski, K. Kuzmin, J. Xie, C. Lee, T. Blanche, E.C. Neto, S.-C. Huang, T. Grabowski, T. Madhyastha, and V. Komanshko, "Identifying robust communities and multi-community nodes by combining top-down and bottom-up approaches to clustering," Sci. Rep., vol. 5, art. no. 16361, 2015.

[4] C. Morselli, Inside Criminal Networks. Springer, New York, NY, USA, 2009.

[5] Q. Zhang and D. Skillicorn, Social Networks with Rich Edge Semantics. CRC Press, Boca Raton, FL, 2017.

[6] M. Chen, A. Bahulkar, K. Kuzmin, and B. K. Szymanski, "Improving network community structure with link prediction ranking," Proc. 7th Work. on Comp. Net. (CompleNet), Dijon, France, Stud. in Art. Int., vol. 644, Springer, Cham, Switzerland, pp.145-158, 2016.

[7] S.M. Rinaldi, J.P. Peerenboom, and T.K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," IEEE Cont. Sys., vol. 21, no. 6, pp. 11–25, Dec. 2001.

[8] J.C. Smith, "Basic interdiction models," Encycl. of Ops. Res. and Man. Sci., 2011.

[9] R. K. Wood, "Deterministic network interdiction," Math. and Comput. Modeling, vol. 17, no. 2, pp. 1–18, Jan. 1993.

[10] A. Malaviya, C. E. Rainwater, and T. C. Sharkey, "Multi-period network interdiction problems with applications to city-level drug enforcement," IIE Trans., vol. 44, no. 5, pp. 368–380, May 2012.

[11] N. O. Baycik, T. C. Sharkey, and C. E. Rainwater, "Interdicting layered physical and information flow networks," IISE Trans., vol. 50, no. 4, pp. 316–331, Apr. 2018.

[12] F. Calderoni, D. B. Skillicorn and Q. Zheng, Quan, "Inductive Discovery Of Criminal Group Structure Using Spectral Embedding," Inf. and Sec.: An Int. Jour, vol. 31, pp. 49-66, Jan.2014.

[13] E. Ferrara and M. De, P. C. Salvatore and F. Giacomo, "Detecting criminal organizations in mobile phone networks," Exp. Sys.with Apps., vol. 41, no. 13, pp. 5733-5750, Oct. 2014.

[14] Q. Zheng and D. B. Skillicorn, "Analysis of criminal social networks with typed and directed edges," 2015 IEEE Int. Conf. on Int. and Sec. Inf. (ISI), 1-610.1109/ISI.2015.7165930, 2015.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE