2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

# Monitoring Individuals in Drug Trafficking Organizations: A Social Network Analysis

Kaustav Basu
*NetXT Lab, SCIDSE*
*Arizona State University*
Tempe, USA
kaustav.basu@asu.edu

Arunabha Sen
*NetXT Lab, SCIDSE*
*Arizona State University*
Tempe, USA
asen@asu.edu

***Abstract*—The United Nations, in their annual World Drug Report in 2018, reported that the production of Opium, Cocaine, Cannabis, etc. all observed record highs, which indicates the ever-growing demand of these drugs. Social networks of individuals associated with Drug Trafficking Organizations (DTO) have been created and studied by various research groups to capture key individuals, in order to disrupt operations of a DTO. With drug offenses increasing globally, the list of suspect individuals has also been growing over the past decade. As it takes significant amount of technical and human resources to monitor a suspect, an increasing list entails higher resource requirements on the part of law enforcement agencies. Monitoring all the suspects soon becomes an impossible task. In this paper, we present a novel methodology which ensures reduction in resources on the part of law enforcement authorities, without compromising the ability to uniquely identify a suspect, when they become "active" in drug related activities. Our approach utilizes the mathematical notion of *Identifying Codes*, which generates unique identification for all the nodes in a network. We find that just monitoring *important* individuals in the network leads to a wastage in resources and show how our approach overcomes this shortcoming. Finally, we evaluate the efficacy of our approach on real world datasets.**

***Index Terms*—Drug Trafficking Organizations, Identifying Code, Unique Monitoring**

## I. Introduction

The 2018 World Drug Report released by the United Nations Office on Drugs and Crime (UNODC) [1], [2], highlighted several grave statistics pertaining to global drug use. 5.6% of the global population belonging to the 15-64 age group (about 275 million people worldwide), consumed drugs *at least once* in 2016. The 15-16 age group (almost 14 million individuals) was responsible for the highest consumption of Cannabis globally. About 450,000 people died as a consequence of drug use in 2015, according to the WHO. About 168,000 of these deaths were directly associated with drug overdoses. These *recorded* numbers are indicative of the ever growing demand for illicit drug consumption, and various Drug Trafficking Organizations (DTOs) are responding by increasing the production rates of their respective drugs. The production of drugs such as Opium, Cocaine, Cannabis, etc all observed record highs in 2016-2017 [2].

To counter this mass production of drugs worldwide, law enforcement agencies spend significant sums of money to just monitor individuals associated with DTOs. The US alone spends about $40 billion a year investigating drug offenses [4]. The investigation of drug offenses usually generates a list of suspects. With drug offenses increasing by the passing year, the suspect database also continues to grow. Hence, the amount of technical and human resources required to monitor a suspect in the database also grows. After a while, monitoring all the suspects in the database may become an impossible task.

Nonetheless, authorities around the world have busted several DTOs, by analyzing transcripts obtained from the interrogation of suspect individuals, to identify key actors and their relationships, to create a social network [4]–[6]. Such networks have been extensively studied by various researchers, to determine the importance of network members, presence of sub groups, number and nature of key facilitators, etc. [7]–[9]. In this paper, we propose a novel Drug Network Monitoring (DNM) approach, based on the mathematical concept of Identifying Codes, that not only reduces resource requirements on the part of the authorities, but also provides the capability of uniquely identifying a suspect, when the suspect becomes "active" in drug related activities. Our approach has the following assumption: when an individual in a DTO becomes "active" in drug related activities, his/her friends/associates will have some knowledge of the individual's plan. Accordingly, even if the individual is not under direct surveillance by the authorities, but the individual's friends/associates are, then *the individual involved in drug related activities can be uniquely identified.* Interestingly, we show that monitoring individuals using standard centrality metrics - *(i) does not necessarily guarantee unique identification of every individual, and (ii) leads to a wastage of resources on the part of the authorities.* To this end, the key contributions of the paper can be summarized as follows:

- We utilize the mathematical concepts of Identifying Codes to uniquely monitor individuals in DTOs, in order to reduce law enforcement requirements.
- We show that just monitoring *important* individuals in

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
*ASONAM '19*, August 27-30, 2019, Vancouver, Canada
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6868-1/19/08/$15.00
http://dx.doi.org/10.1145/3341161.3342938

the network may not be an effective strategy, and how our approach overcomes this shortcoming.

## II. Related Work

In the past few years, significant research on DTO networks have been conducted utilizing Social Network Analysis (SNA). Natarajan in [5], analyzed wiretap data to create an organizational structure and studied the roles of particular individuals. She analyzed over 2000 wiretap conversations and performed SNA of phone contacts, to reveal a large and loosely structured group of 294 individuals in [6]. Bright and Delaney in [7], utilized SNA to study the evolution of a drug trafficking network, based in Australia. They observed changes in centrality scores and the roles performed by particular individuals. Bright *et. al.* in [8] utilized individual attributes coupled with centrality measures to identify key actors in a drug trafficking network. Bright *et. al.* in [9] analyzed judges' sentencing comments to create a network of individuals involved in the distribution of methamphetamine in Australia during the 1990s. Carley in [10], explored the potential use of SNA and multi agent modeling for the purpose of destabilizing covert networks. Hughes in [3] analyzed court transcripts and identified key actors to create a social network to explore product diversification in three drug syndicates in Australia.

In addition to drug organizational structure research through SNA, the last few years have seen a significant amount of research on Identifying Codes and its applications to networks. Karpovsky *et. al.* introduced the concept of Identifying Codes in [11] and provided results for Identifying Codes for graphs with specific topologies, such as binary cubes and trees. Using Identifying Codes, Laifenfeld *et. al.* studied joint monitoring and routing in wireless sensor networks in [12]. Charon *et. al.* in [13], studied complexity issues related to computation of minimum Identifying Codes for graphs and showed that in several types of graphs, the problem is NP-hard. Ray *et. al.* in [14] presented a *minimal* algorithm for the computation of Identifying Codes.

SNA studies have usually utilized network centrality measures to identify key actors in the network. In this paper, we show that if these metrics were used to deploy agents for unique monitoring of individuals, then the authorities would end up wasting their resources. In our previous efforts, we have utilized Identifying Codes for, (i) monitoring the health of critical infrastructures and civilian structures [15], [19], (ii) monitoring the activities of terrorists [16], [18], (iii) propagation of misinformation on social networks [20] and, (iv) monitoring regions on the surface of the earth [17]. To the best of our knowledge, this is the *first* study to analyze the social network of a DTO and apply *Identifying Codes* to uniquely identify individuals in a DTO network.

## III. Identifying Codes

In this section, we formally define the mathematical notion of *Identifying Codes*. Subsequently, we present our novel approach which utilizes this notion to uniquely monitor a given drug network.

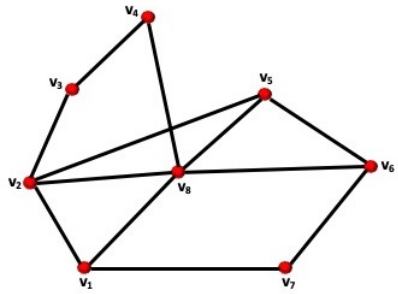


Fig. 1. Undirected graph with Identifying Code Set $\{v_4, v_6, v_7, v_8\}$

**Definition III.1.** Given an undirected graph $G = (V, E)$, the subset $V' \subseteq V$, is defined as an Identifying Code Set (ICS) for the vertex set $V$, if $\forall v \in V, N[v] \cap V'$ is unique, where, $N[v] = v \cup N(v)$ and $N(v)$ represents the set of nodes adjacent to $v$ in $G = (V, E)$. The Minimum Identifying Code Set (MICS) problem is to find the ICS of smallest cardinality.

The vertices of the set $V'$ can be thought of as *alphabets* of the code, and the string made up with the alphabets of $N[v]$ can be viewed as the unique "code" for the node $v$. This is better explained with the help of the following example. Consider a graph $G = (V, E)$, as illustrated in Fig. 1. The MICS for $G$ is $V' = \{v_4, v_6, v_7, v_8\}$. As shown in Table I, in $G = (V, E)$, $\forall v \in V, N[v] \cap V'$ is unique. Hence, the node set $V' = \{v_4, v_6, v_7, v_8\}$ is an MICS of $G$.

**Definition III.2.** Given a directed graph $G = (V, E)$, the subset $V' \subseteq V$, is defined as an Identifying Code Set (ICS) for the vertex set $V$, if $\forall v \in V, N^{out}[v] \cap V'$ is unique, where, $N^{out}[v] = v \cup N^{out}(v)$ and $N^{out}(v)$ represents the set of out-neighbors of $v$ in $G = (V, E)$. As before, the Minimum Identifying Code Set (MICS) problem is to find the ICS of smallest cardinality.

**Definition III.3.** Two nodes $u, v \in V$ are "twins" if $N[u] = N[v] (N^{out}[u] = N^{out}[v])$ in an undirected (directed) graph.

**Observation:** Identifying Code Set (ICS) of a graph $G = (V, E)$ does not exist, if any two nodes $u, v \in V$ are "twins".

TABLE I
$N[v] \cap V'$ results for all $v \in V$ for the graph in Fig. 1

| | |
|---|---|
| $N[v_1] \cap V' = \{v_7, v8\}$ | $N[v_2] \cap V' = \{v_8\}$ |
| $N[v_3] \cap V' = \{v_4\}$ | $N[v_4] \cap V' = \{v_4, v_8\}$ |
| $N[v_5] \cap V' = \{v_6, v_8\}$ | $N[v_6] \cap V' = \{v_6, v_7, v_8\}$ |
| $N[v_7] \cap V' = \{v_6, v_7\}$ | $N[v_8] \cap V' = \{v_4, v_6, v_7, v_8\}$ |

**MICS and MDCS computation as a Graph Coloring with Seepage (GCS) Problem**: The MICS and MDCS computation problem can be viewed as a novel variation of the standard Graph Coloring problem. We will refer to this version as the *Graph Coloring with Seepage (GCS)* problem. In the standard graph coloring problem, when a color is *assigned* (or injected) to a node, only that node is colored. The goal of the standard graph coloring problem to use as few distinct colors as possible such that (i) every node receives a color, and (ii) no two adjacent nodes of the graph have the same color. In the GCS problem, when a color is assigned (or injected) to a node, not

only that node receives the color, the color also *seeps* into all the adjoining nodes. As a node $v_i$ may be adjacent to two other nodes $v_j$ and $v_k$ in the graph, if the color red is injected to $v_j$, not only will $v_j$ become red, but also $v_i$ will become red, as it is adjacent to $v_j$. Now if the color blue is injected to $v_k$, not only will $v_k$ become red, but also, the color blue will seep in to $v_i$, as it is adjacent $v_k$. Since $v_i$ was ready colored red (due to seepage from $v_j$), after color seepage from $v_k$, its color will be a *combination of red and blue, i.e., purple*. At this point all three nodes $v_i$, $v_j$ and $v_k$ have a color and all of them have distinct colors (red, blue and purple). The goal of the GCS problem is to inject colors to as few nodes as possible, such that *(i) every node receives a color, and (ii) no two nodes of the graph have the same color*. Suppose that the node set $V'$ is an ICS of of a graph $G = (V, E)$ and $|V'| = p$. If $p$ distinct colors are injected to the nodes of $V'$ (one distinct color to one node of $V'$ ), then as by the definition of ICS for all $v \in V$ if $N[v] \cap V'$ is unique, all nodes of $G = (V, E)$ will be colored and no two nodes will have the same color. Accordingly, computation of the MICS problem is equivalent to computation of the GCS problem.

## IV. Approach

In this section, we formalize the DNM problem utilizing two types of networks - undirected and directed. In the two types of networks, the nodes represent the suspect individuals associated with the respective DTO, and the edges represent the relationships between the individuals. Here, the objective is to monitor (surveillance of the suspected individual) a subset of the nodes $V' \subseteq V$, so that each and every node $v \in V$, can be *uniquely* identified, in case the individual represented by $v$ becomes active in drug related activities. Here, we assume that only *one* node $v$ becomes active at a time and each and every node $v \in V$ can be monitored. Recalling the objective of the GCS problem described in Section III, a natural connection between the GCS problem and the DNM problem can be made. The suspect individuals who have to be monitored, can be thought of as nodes in the network, where colors have to be injected in the GCS problem. In the GCS problem, injection of colors in a subset of nodes in the network (or graph), ensures that each node in the network receives a unique color. Thus, monitoring (assigning law enforcement authorities to) this subset of suspect individuals will result in the unique identification (or signature) for all of the suspect individuals (or all the nodes) in the network. It may be noted that in Section III, the GCS problem was shown to be equivalent to the computation of the MICS problem. By transitivity, the undirected/directed DNM problem is equivalent to solving the MICS of the corresponding drug network. We now provide solution techniques utilizing Integer Linear Programs (ILP).

***Instance:*** $G = (V, E)$, an undirected graph.

***Problem***: Find the smallest subset $V' \subseteq V$, such that injection of colors at these nodes, ensures that each node $v \in V$, receives a unique color (atomic/composite) through seepage.

We use the notation $N[v_i]$ to denote the closed neighborhood of $v_i$, $\forall v_i \in V$. Corresponding to each $v_i \in V$, we use an indicator variable $x_i$,

$$x_i = \begin{cases} 1, & \text{if a color is injected at node } v_i, \\ 0, & \text{otherwise} \end{cases}$$

*Objective Function: Minimize* $\sum_{v_i \in V} x_i$

*Coloring Constraint:* $\sum_{v_i \in N[v_j]} x_i \geq 1$, $\forall v_j \in V$

*Unique Coloring Constraint:*
$\sum_{v_i \in \{N[v_j] \bigoplus N[v_k]\}} x_i \geq 1$, $\forall v_j \neq v_k, \in V$

$N[v_j] \bigoplus N[v_k]$ denotes the Exclusive-OR of the node sets $N[v_j]$ and $N[v_k]$. It may be noted that the objective function ensures that the fewest number of nodes in $V$ are assigned a color. The Coloring Constraint ensures that every node in $V$ receives at least one color through seepage from the colors injected at nodes in its closed neighborhood. A consequence of the Coloring Constraint is that, a node in $V$ may receive more than one color through seepage from the colors injected at its neighborhood. The Unique Coloring Constraint ensures that, for every pair of nodes $(v_j, v_k)$ in $V$, at least one node in the node set $N[v_j] \bigoplus N[v_k] \subseteq V$ is injected with a color. This guarantees that $v_j$ and $v_k$ will not receive identical colors.

The ILP formulation for directed graphs is very similar to the undirected graph ILP formulation. In this case, we consider the closed out-neighborhood set of a node $N^{out}[v_i]$ instead of the closed neighborhood set $N[v_i]$.

## V. Experimental Results

To highlight the effectiveness of our algorithms, we executed the ILPs on various real world drug network datasets [21]. It may be recalled that for a network to have an Identifying Code, it must be "twin-free". For "twin-free" networks $G = (V, E)$, one trivial Identifying Code set solution is the node set $V$, although, $V$ may not be the Identifying Code set of *minimum cardinality*. This implies that, if monitors were placed on every node in the network, then all the nodes in the network would receive a unique identification. However, our algorithms show that unique identification for all the nodes in the network can be obtained by injecting colors in a subset $V' \subseteq V$. Since no additional benefits are realized by injecting more colors (deploying a higher number of monitors) in the network, there is absolutely no reason to deploy a larger number of monitors. In Table II, we highlight on the reduction in resource requirements brought about by our methods. As monitoring suspect individuals in drug networks can be a costly affair on the part of law enforcement authorities [4], a significant reduction in resources will be of great interest to the respective authorities. Furthermore, we also highlight the amount of resource wasted when standard network centrality metrics are utilized for monitoring individuals in a network.

The Juanes, Mambo, Heroin and DruNet networks were not initially "twin-free". By combining the nodes which form "twins" into super-nodes, one can ensure (i) the network

TABLE II
MICS CARDINALITIES FOR DRUG NETWORKS

| Networks | Type | Number of Nodes | MICS Cardinality | Reduction in Resources | DC MICS | Resource Wastage | BC MICS | Resource Wastage | EV MICS | Resource Wastage |
|---|---|---|---|---|---|---|---|---|---|---|
| Operation Juanes | Undirected | 50 | 22 | **56%** | 45 | **104.54%** | 45 | **104.54%** | 45 | **104.54%** |
| Operation Acero | Undirected | 25 | 13 | **48%** | 23 | **76.92%** | 23 | **76.92%** | 23 | **76.92%** |
| Operation Mambo | Undirected | 30 | 16 | **46.66%** | 29 | **81.25%** | 29 | **81.25%** | 29 | **81.25%** |
| Heroin Dealing | Undirected | 37 | 15 | **59.46%** | 36 | **140%** | 36 | **140%** | 36 | **140%** |
| Montreal Street Gangs | Undirected | 35 | 16 | **54.28%** | 35 | **118.75%** | 35 | **118.75%** | 35 | **118.75%** |
| Cocaine Dealers | Directed | 28 | 23 | **17.85%** | 25 | **8.69%** | 28 | **21.74%** | 25 | **8.69%** |
| Operation Jake | Directed | 38 | 29 | **23.68%** | 36 | **24.13%** | 36 | **24.13%** | 36 | **24.13%** |
| DrugNet | Directed | 281 | 207 | **26.33%** | 281 | **35.75%** | 281 | **35.75%** | 281 | **35.75%** |

becomes "twin-free" and, (ii) the computation of Identifying Codes. If the nodes $u, v \in V$ form "twins" we can create a super node $(u, v)$ by condensing nodes $u$ and $v$. However, as the modified network does not contain either $u$ or $v$ (it has the super node $(u, v)$), if the individual corresponding to nodes $u$ or $v$ were to become "active", Identifying Code will not be able to distinguish between these two individuals. Hence, further (lower level) analysis will be needed to find out whether node $u$ or $v$ is in the process of being "active".

For the ease of understanding, we describe the results of the Operation Juanes drug network in detail. Out of 50 nodes in the network, if colors were injected in *only 22 nodes*, then all the nodes in the network would receive a unique identification. In other words, if authorities monitored just 22 individuals, then all the individuals in the network, could be uniquely identified. This indicates a *reduction* of $\frac{(50-22)}{50} \times 100$ = **56%** in resource requirements. Furthermore, we compare the MICS cardinality following the Identifying Code approach with the MICS cardinalities obtained by following the standard SNA centrality approaches - degree centrality (DC), betweenness centrality (BC) and eigen vector centrality (EV). We observe that if we were to adopt these standard SNA metrics as monitoring strategies, then law enforcement agents will need to monitor 45 individuals with the *highest centrality scores*, to ensure that each node in the network has unique identification. Individuals with the highest centrality scores are an indication of the importance of the individual to the respective drug network [7]. Requiring 45 nodes instead of the MICS solution of 22, clearly results in wastage of resources. For this network, the wastage amounts to $\frac{(45-22)}{22} \times 100$ = **104.54%**. This is somewhat counter-intuitive in the sense that, authorities would end up wasting resources if they were to *monitor only important individuals*. Identifying Code overcomes this shortcoming as it takes the entire network (a global view) into account rather than the individual importance of respective nodes (a local view). Results for the other undirected/directed graphs are presented in Table II.

## VI. CONCLUSION

In this paper, we presented a solution technique for monitoring drug networks which results in a significant reduction in resource requirements on the part of the authorities. We conducted extensive experimentation to show that utilizing standard SNA centrality metrics for monitoring DTOs, leads to wastage in resources, on the part of the authorities and how Identifying Code can overcome this shortcoming.

## REFERENCES


[1] "Drug Trafficking" [Online]. Available: https://www.unodc.org/unodc/en/drug-trafficking/'

[2] "UNODC World Report" [Online]. Available: https://www.unodc.org/wdr2018/en/exsum.html

[3] C. E. Hughes, D. A. Bright, and J. Chalmers. "Social network analysis of Australian poly-drug trafficking networks: How do drug traffickers manage multiple illicit drugs?" Social Networks 51 (2017): 135-147.

[4] "How Drug Dealers Get Caught" [Online]. Available: https://www.robertkinglawfirm.com/blog/2013/november/how-drug-dealers-get-caught/

[5] M. Natarajan, "Understanding the structure of a drug trafficking organization: a conversational analysis." Crime Prev. Studies 11 (2000): 273-298.

[6] M. Natarajan, "Understanding the structure of a large heroin distribution network: A quantitative analysis of qualitative data." Journal of Quantitative Criminology 22.2 (2006): 171-192.

[7] D. A. Bright, and J. J. Delaney. "Evolution of a drug trafficking network: Mapping changes in network structure and function across time." Global Crime 14.2-3 (2013): 238-260.

[8] D. A. Bright, et al. "The use of actor-level attributes and centrality measures to identify key actors: A case study of an Australian drug traffic network." J. of Contemp. Criminal Justice 31.3 (2015): 262-278.

[9] D. A. Bright, C. E. Hughes, and J. Chalmers. "Illuminating dark networks: A social network analysis of an Australian drug trafficking syndicate." Crime, law and social change 57.2 (2012): 151-176.

[10] K. M. Carley, et al. "Destabilizing dynamic covert networks." Proceedings of the 8th international Command and Control Research and Technology Symposium. 2003.

[11] M. G. Karpovsky, K. Chakrabarty, and L. B. Levitin. "On a new class of codes for identifying vertices in graphs." IEEE Transactions on Information Theory 44.2 (1998): 599-611.

[12] M. Laifenfeld et al. "Joint monitoring and routing in wireless sensor networks using robust identifying codes." Mobile Networks and Applications 14.4 (2009): 415-432.

[13] I. Charon, O. Hudry, and A. Lobstein. "Minimizing the size of an identifying or locating-dominating code in a graph is NP-hard." Theoretical Computer Science 290.3 (2003): 2109-2120.

[14] S. Ray et al. "Robust location detection in emergency sensor networks." IEEE INFOCOM. Vol. 2. IEEE, 2003.

[15] K. Basu et al. "Health Monitoring of Critical Power System Equipments Using Identifying Codes." International Conference on Critical Information Infrastructures Security. Springer, Cham, 2018.

[16] A. Sen. et al. "Terrorist Network Monitoring with Identifying Code." International Conference on SBP-BRiMS. Springer, Cham, 2018.

[17] A. Sen et al. "On upper and lower bounds of identifying code set for soccer ball graph with application to satellite deployment." In Proceedings of the 20th ICDCN, pp. 307-316. ACM, 2019.

[18] K. Basu et al. "A Novel Graph Analytic Approach to Monitor Terrorist Networks." IEEE SocialCom, 2018.

[19] K. Basu et al. "Sensor Networks for Structural Health Monitoring of Critical Infrastructures Using Identifying Codes." In 2019 15th IEEE (DRCN), pp. 43-50., 2019.

[20] K. Basu. "Identification of the Source(s) of Misinformation Propagation Utilizing Identifying Codes". In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19), pp 7-11.

[21] "Covert Networks" [Online]. Available: https://sites.google.com/site/ucinetsoftware/datasets/covert-networks