Follow the money: Revealing risky nodes in a Ransomware-Bitcoin network

Adam Brian Turner
Macquarie University, Department of Security Studies and Criminology, Sydney, Australia
adam.turner@students.mq.edu.au

Stephen McCombie
Macquarie University, Department of Security Studies and Criminology, Sydney, Australia
stephen.mccombie@mq.edu.au

Allon J. Uhlmann
Macquarie University, Department of Security Studies and Criminology, Sydney, Australia
allon.uhlmann@mq.edu.au

Abstract

This paper demonstrates the use of network analysis to identify core nodes associated with ransomware attacks in cryptocurrency transaction networks. The method helps trace the cyber entities involved in cryptocurrency attacks and supports intelligence efforts to identify and disrupt cryptocurrency networks.

A data corpus is built by the unsupervised machine learning graph algorithm 'DeepWalk' [1]. DeepWalk evaluates the position of nodes within networks. It compares the relative position of different nodes (similarity) and identifies those whose removal would most affect the network (riskiness). This method helps identify on the blockchain the key nodes that are involved in the execution of a ransomware attack.

When applied to the ransomware "cash out" graph, the method derived "riskiness" scores for specific nodes. Analysing the derived "riskiness" at a community level (groups of nodes in the network) provides an enhanced granularity for identifying and targeting influential nodes. Such insight could potentially support both intelligence and forensics investigations.

1. Introduction

In 2019 over US$6.6 million was paid globally to cryptocurrency addresses related to ransomware, according to the 2020 Crypto Crime Report from blockchain analysis company Chainalysis [2]. This is emphasised by the fact that the United States Securities and Exchange Commission (US SEC) has seen over 1,000 documents submitted by companies between April 2019 and May 2020 that list ransomware as a critical risk factor to their businesses [3]. Companies face multi-million dollar outages such as those faced by the city of New Orleans in 2019. The city's Chief Administrative Officer, Gilbert Montaño, indicated that the ransomware attack will cost the city at least US$7 million [4]. There are plenty of opportunities for cyber criminals to cash out their booty. One of the most popular ways for ransomware attackers to do so between 2013 and 2016 was through the Russian based BTC-e exchange [2].

Identifying the magnitude and location of illicit funds throughout the blockchain is no easy endeavour, the cryptocurrency investigation companies Elliptic and Chainalysis provide their own powerful proprietary software platforms to do this. However, there are some open-source tools and techniques that allow us to analyse this evolving threat to confront ransomware attacks.

Throughout this paper, network and graph will be used interchangeably as we explore the utility of graph analysis for cyber financial crime prevention. Out of the hundreds of thousands of Bitcoin transactions on a blockchain, the first challenge is to isolate the relevant Bitcoin nodes used in a ransomware attack. We will further show how graph analysis reveals patterns and provides the capability to expose nefarious relationships between the Bitcoin transactions and addresses in the ransomware-Bitcoin network. In addition, DeepWalk [1] embeddings provide a machine-learning technique for graphs that sets up feature extraction from the ransomware-Bitcoin cash-out network. These features can be used in a similarity analysis that is based on Cosine Similarity to identify the risk posed by the removal of a node from the Bitcoin-ransomware cash-out network. We will apply the Cosine Similarity calculation comparing nodes with the ransomware seed address to isolate individuals and communities of risky nodes. Furthermore, our target network dataset can be enriched with contextual labels derived from other open source blockchain analysis tools setting up future research with more advanced machine learning prediction techniques.

URI: https://hdl.handle.net/10125/70801
978-0-9981331-4-0

2. Fighting financial crime with graph analysis

Tracing illicit flows of money through a network requires techniques that reveal patterns and provide the capability to expose nefarious relationships across vast amounts of data. The trails left behind by these financial flows provide a web of transactions interconnected by accounts and services to obfuscate identity on purpose by blending seamlessly into the economic system. In traditional banking, the transactions, accounts and services form a network and can be modelled as a graph. For example, De Marzi [5], uses credit card fraud as a case study, modelling where credit card holders make legitimate transactions at different services and in another graph showing where fraud actors with stolen credit card data test the stolen credit card numbers. By modelling this fraud scenario as a graph, it helps identify patterns where the credit card data may have been stolen or where stolen credit card data is being tested at certain services.

Voutila [6] uses the PaySim mobile money network financial dataset originally posited by Lopez-Rojas *et al* [7]. The graph model created contains transactions, merchants, clients and client identifiers in order to filter a large set of activity and perform graph analysis, such as weakly connected components, to identify fraud rings within the larger graph. Components, nodes, in a graph are said to be weakly connected if they are all connected or reachable from any other node in the same graph. Galler and Fischer [8], first revealed this algorithm and it has been used to understand how well connected networks are, how clusters of activity form and how well the network remains connected when nodes of certain authority are eliminated.

Furthermore, the case for revealing money laundering has an even stronger emphasis today. Anti-Money Laundering laws, regulation and compliance such as, The Anti-money Laundering and Counter-terrorism Financing Act 2006 (Cth) in Australia [9] and the 5th Anti-Money Laundering Directive of the European Union [10] provide a legislative framework for the prevention and detection of money laundering and terrorism financing. However, detecting money laundering networks still proves extremely difficult as seen in the 2017 royal commission into the Australian banking, superannuation, and financial services industry, where over 200 money laundering compliance failures were revealed with one bank alone [11, 12]. Data and Analytics firm Dun and Bradstreet put this difficulty down to the scale and complexity of the data that needs to be analysed to find the nefarious relationships within the financial transactions. As a firm they are using graph technology to meet the Anti-Money Laundering standards previously mentioned [13].

Whilst graph analysis has become an established tool for identifying and fighting financial fraud in the traditional economy, a question remains as to the utility of the method in the emerging space of cryptocurrencies, especially Bitcoin, which is the cryptocurrency of choice for most ransomware attacks today. Bitcoin uses Bitcoin addresses as a banking client would use their bank account number. Bitcoin value is sent and received between addresses via transactions. There are many Bitcoin addresses that make up a Bitcoin wallet. It is not uncommon for wallet users to create a new Bitcoin address for every new transaction to help preserve their anonymity [14]. The full balance of a Bitcoin address needs to be spent during a transaction and as such change addresses are often found, where the balance of the transaction is paid back to the originating Bitcoin address. Spotting irregular Bitcoin activity and unusual connections occurring in the blockchain at scale proves extremely difficult without the aid of graph visualization tools [14]. These reveal patterns and anomalies in intuitive and interactive ways. Therefore, the graphs derived for ransomware-Bitcoin behaviour provide a powerful analysis capability which leverages the Bitcoin ecosystem to build the scope of the ransomware-Bitcoin target network.

3. The Ransomware-Bitcoin target network

In line with Clark and Mitchell's target centric approach, we begin the intelligence process by defining a generic target model to guide intelligence collection and analysis [15]. The generic Target Network Model (TNM) for our ransomware-Bitcoin target network is represented by Figure 1.

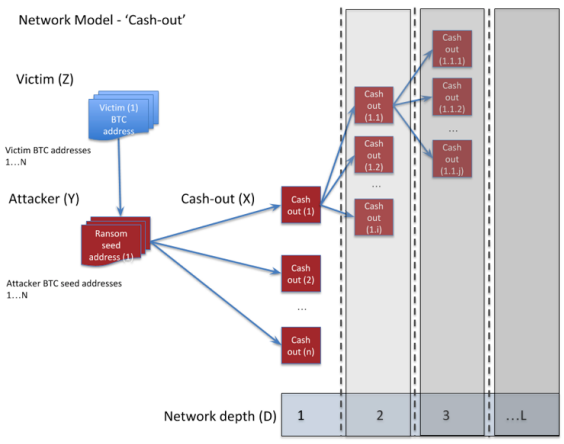

Figure 1 – Ransomware - Bitcoin Target Network Model ('Cash-out')

Figure 1 shows the representation of Bitcoin addresses and transactions at different levels of a target network in a model of a ransomware campaign. Due to the size and complexity of the overall ransomware campaign network the TNM is split between cash-in and cash-out models. Figure 1 only shows the cash-out side of the network.

The cash-out network models the proceeds of crime as they flow from the ransomware seed address that victims of the ransomware attack have paid into to other addresses in the Bitcoin universe. These ransom payments ultimately exit the network where they are exchanged for other cryptocurrencies or even fiat currency.

In order to demonstrate the method, we will collect data related to the cash-out network of the ransomware campaign, WannaCry 2.0 and populate the generic target model. This campaign was chosen because the findings from our network investigation can be validated against other sources. The next section will identify the data collection requirements and methods used and introduce the analysis system being applied to the populated TNM.

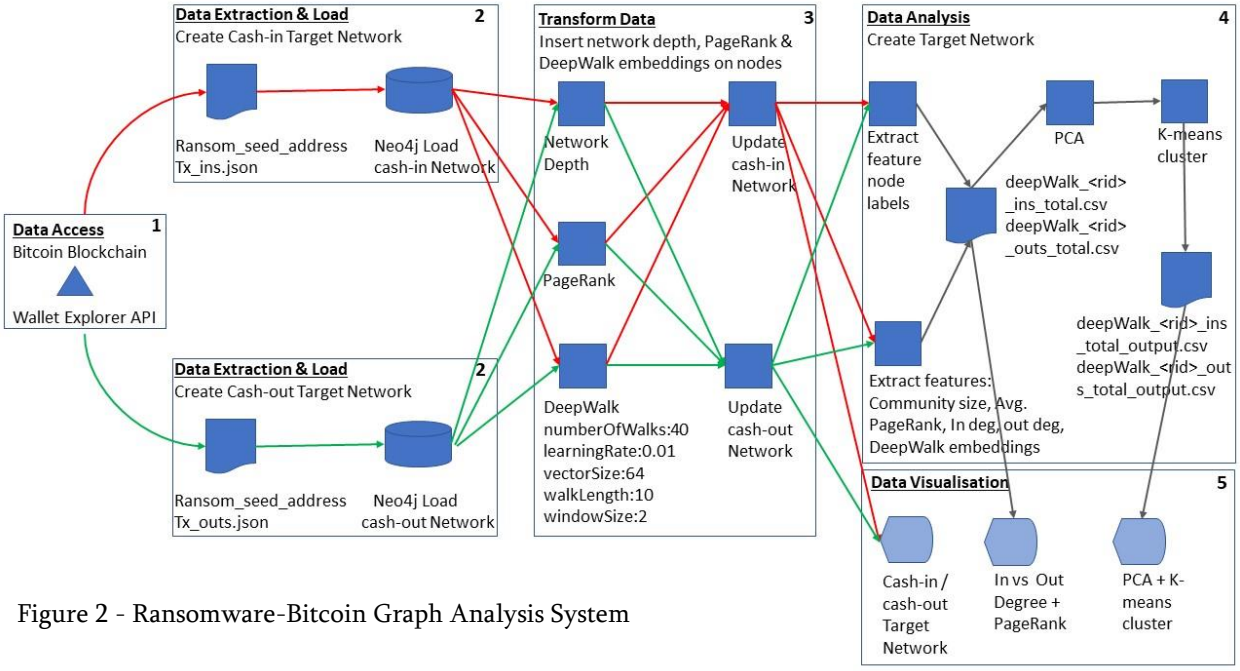

Figure 2 - Ransomware-Bitcoin Graph Analysis System

4. Data Collection

The Bitcoin blockchain contains the record of addresses and transactions involved in Bitcoin transactions. This 'on-chain' data along with the respective meta-data can be exploited.

Figure 2 shows a data pipeline with associated analysis techniques that are used to exploit Bitcoin blockchain data.

Step one – Extract data from the Bitcoin blockchain

- Extract transaction history relating to the ransom seed address from the walletexplorer.com Application Programming Interface (API).
- For each incoming and outgoing transaction from the seed address, build the input and output graphs respectively at 'D' levels deep away from the seed address (see Figure 1).

Step two – Load data into graph database (Neo4j)

- Load extracted input and output graph files setting input/output addresses as nodes; transactions as nodes; and Payments as a relationship between them.
- Post process address nodes to include corresponding depth 'D' of transaction nodes.

Step three – Transform data

- Run the PageRank algorithm and add this as a property on the nodes in the network.
- Run the DeepWalk algorithm on nodes and embed the results onto the nodes in the network

Step four – Data analysis preparation

- Run Louvain community detection algorithm using average in/out degree and PageRank. Aggregate results of communities.
- Run community detection and return non-aggregated results, returning all nodes in the

network with the respective in/out degree, PageRank, DeepWalk embeddings, labels, depth, timestamp. Export to Comma Separated Values (CSV).
- These network analysis algorithms were run from within the Neo4j graph database

Step five – Data visualisation

- Import the CSV into python script to:
- Visualise community detection profile
- Python was used to perform Principal Components Analysis (PCA) + K-means clustering on DeepWalk embeddings
- Output results to CSV for deep dive analysis into comparing communities and clusters across different ransomware by using Cosine Similarity.

The key transformation of the data we will focus on in this publication relates to the graph embeddings derived in steps three and four. The graph embeddings will become features for future graph machine learning applications. The PCA undertaken in step five is essential for managing the dimensionality of the embedding computations. It is key to the analysis to examine specific nodes and determine how influential they are within the Bitcoin-ransomware network. This would serve as an indicator of their relative importance in the transfer and circulation of ransom payments. For this reason, the PageRank algorithm was chosen as an appropriate centrality measure for this purpose. The subsequent sections elaborate on the proposed methodology.

5. Risky node analysis

The blockchain data should yield a network of wallets and transactions involved in the WannaCry ransomware attack, but the network data on its own does not provide sufficient context to identify the key nodes that are involved in the process. We propose to approximate the significance of each node in the network by measuring the effect the node's removal would have on the viability and function of the network. We conceptualise this effect as risk to the network, and the measurement involved as a measure of riskiness. Using the DeepWalk graph embeddings that encode the structure of a graph at each node relative to its position in the target network (see figure 1), we can leverage these embeddings as features into a Cosine Similarity calculation which provides an index of how 'risky' the nodes are relative to the ransomware Bitcoin seed address. Furthermore, analysing the risky nodes collectively forms target communities which could prove more effective as opposed to targeting these nodes individually.

5.1. Graph embeddings and features

Once the TNM has been created and populated with the extracted data, the graph itself becomes very large and dense making it difficult to detect any unusual behaviour at face value. There needs to be a simplified way of preserving the graph properties like the structure and the features on the nodes and edges [16]. This is achieved by the graph embedding algorithm that transforms all the information learned from a graph into a lower dimensional vector space representation. The graph embedding algorithm chosen for this analysis is DeepWalk by Perozzi *et al* [1].

DeepWalk learns structural representations of a graph's nodes by capturing its similarity in a neighbourhood of other nodes and allocating individual nodes to cliques we call communities [1]. By taking a graph as an input to the algorithm, latent representations are produced as an output. These representations become the input to a neural network. Operating a neural network on a graph structure allows for deep feature learning of nodes and edges for a graph [17]. DeepWalk uses deep learning for unsupervised feature learning, which means, the system learns the node's embeddings without any prior knowledge of the graph topology. Depending on what nodes are encountered and how often they are traversed during a random walk, the neural network makes a prediction about a node feature or classification and embeds that into the node as metadata. By sampling the graph via random walks, we build the data corpus for that graph. The data corpus is then used as the reference library for a node's purpose within the graph. For example, in the ransomware-Bitcoin TNM the ransomware seed address can be taken and its "context" predicted within the scope of the entire graph. This means embedding an understanding of a node's features, such as, transaction amount, connectivity to other nodes (how many input and output transactions there are from a node) and structural role (E.g. the root node of the network or a leaf of a weakly connected branch). Having these embeddings encoded into a node provides a basis for subsequent generalisation through various possible means. In this case, we chose the PCA and K-means clustering analysis (see figure 2) to reduce the dimensionality of the embeddings. PCA as a method of reducing large datasets whilst preserving as much information, or statistical variability, from the original data [18]. In this case we were able to reduce the relevant dimensionality from 128 down to a two-dimensional vector space. This two-dimensional representation of the graph embeddings will now be used in the next section as input

into a similarity analysis to ascertain which nodes in the TNM are riskier than others.

5.2. Concept of Similarity

Cosine Similarity is a measure used to identify how similar entities, or in this case nodes in a network, are irrespective of their magnitude [19]. In this case, the graph analysed by the DeepWalk algorithm is reduced to a series of variables which incorporate latent features of a node's community structure as an output for use in calculating the similarity between the vector representations of these features.

Seeing as the theory of this process has its roots in natural language processing, we use the analogy of finding meaning and context (similarity) in a text. The process allows us to analyse the similarity in words' meanings, while ignoring the words' location in the text. The procedure we are proposing would be equivalent to having a graph (the document), containing many random walks (sentences) from each of the nodes (words) in the graph. Ultimately arriving at a meaningful similarity of a node's context with respect to the other nodes in the graph. An illustration of this can be seen in figure 3.

5.3. Application

Instead of measuring the distance between two nodes, Cosine Similarity measures the cosine of the angle between them. Cosine similarity is superior to a simple measure of distance in identifying the common features of disparate nodes. Plotting the distribution of the cosine similarity, box labelled 5 in figure 3, assert that there are close similarities between nodes not purely related to the latent features derived from the DeepWalk embeddings. By taking the cosine similarity of these features we are not only considering the proximity of a node to the ransomware seed address, rather the context of the node in the whole graph being analysed. This can be seen in figure 3, box labelled 3, where node C is in close proximity to the ransomware seed address X, however in figure 3, box labelled 5, the angle between C and X is larger than the angle between A and X, where A is more distant from the ransomware seed address in box labelled 3.

There could be several reasons for this. The number of nodes directly connected to nodes C and A, or closely connected in the neighbourhood (two or three hops away). Additionally, how many times the particular

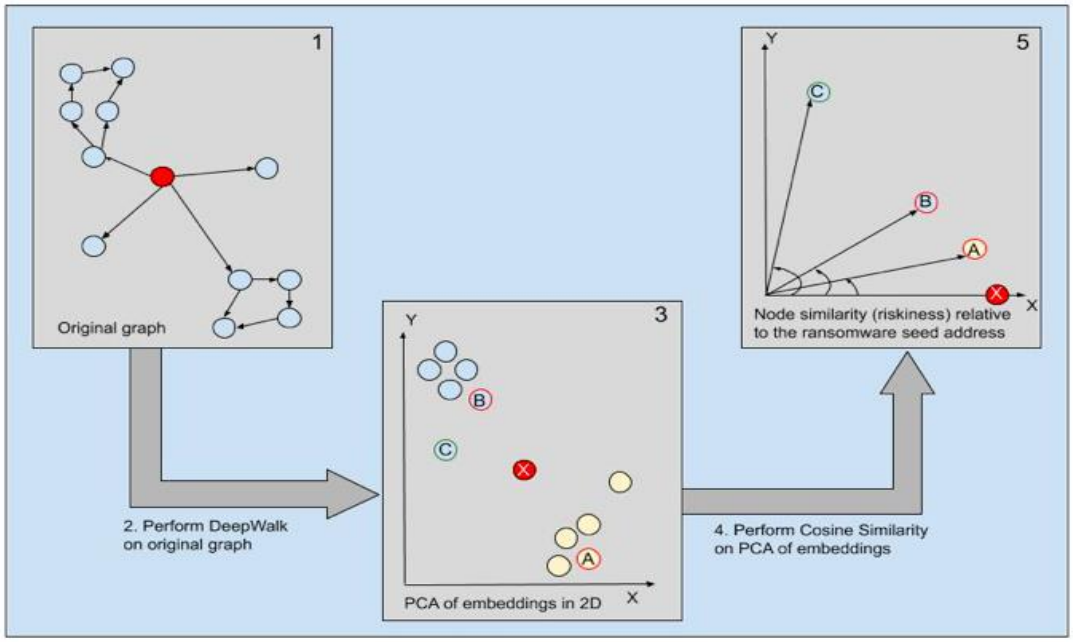

Figure 3 - Conceptual view of arriving at a measure of riskiness in the ransomware-Bitcoin graph

node occurs in context to other nodes in the generated corpus of the entire graph relative to the ransomware

bitcoin seed address in the network. For example, in the cash-out graph for WannaCry ransomware Bitcoin seed address, 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw, in figure 4, a node, 1BvTQTP5PJVCEz7dCU2YxgMskMxxikSruM, with a high similarity relative to the ransom seed address resides at a Bitcoin exchange Poloniex.com. This was identified as one of the cash-out exchanges used by the attackers in WannaCry (Bistarelli et al 2018).

Therefore, it follows that similarity scores relative to the ransom seed address might usefully serve as a proxy measure for riskiness. The higher the similarity calculated for a node with respect to the ransom seed address, the higher the risk score for that node. For example, if a high-risk scoring node was removed from the network the attackers would be unable to cash out their proceeds of crime. Therefore, risk used in this context refers to the risk imposed on the attacker fulfilling the objectives of the network.

5.4. Similarity as a measure of risk

Similarity can therefore be used as a proxy for riskiness. Using the mathematical calculation of Cosine Similarity, a proxy measure for riskiness is established relative to the ransomware seed address. Taking the ransomware seed address as the most significant node on a ransomware-Bitcoin cash-out network, (because if there was no seed address created, there would be no ransom collected), then using the graph embeddings to computationally calculate the similarity of every other node relative to this node we are able to derive a risk score. As a result of this analysis a similarity matrix is produced that can be used as a heatmap to target the risky nodes in the network. If another node in the ransomware cash-out network scores a high 'similarity' relative to the ransomware seed address, and if that highly scored node is removed from the network this node is the next critical to the network fulfilling its objective of cashing out the ransom collected. This would allow for the targeting of nodes with the high similarity scores and hence if we target or neutralise this node, it puts the network's objectives "at risk". For example, using a classification range the heatmap could be represented as follows: 'Very High' for risk scores ranging from 0.95 to 1, shaded in red; 'High' from 0.75 to 0.95, shaded in Orange-Red; 'Medium' from 0.5 to 0.75, Yellow-Orange; 'Low' from 0 to 0.5, Green-Yellow. Applying this concept to the Wannacry cash-out network produces the following results in table 1.

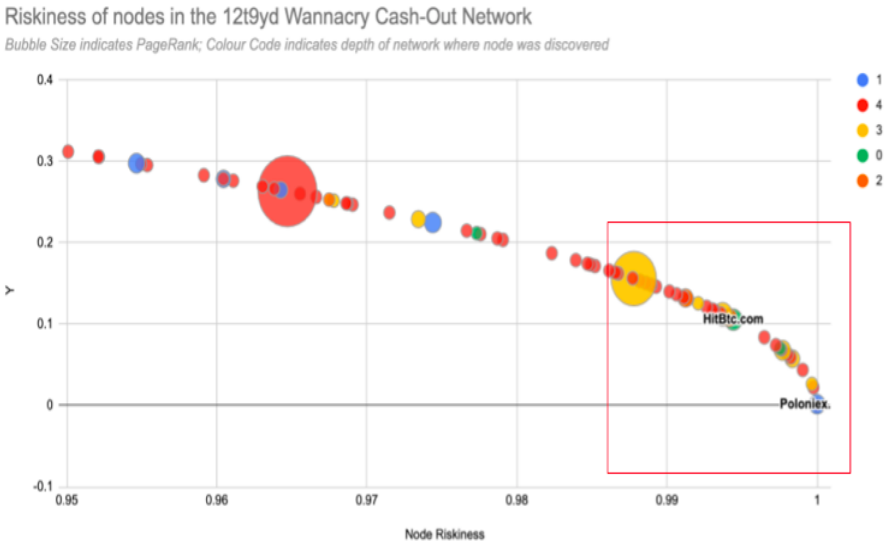

Figure 4 - Distribution of node similarity for WannaCry Ransomware seed address *12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw* cash-out network. Top 20% of nodes by risk score

ID	Node (Address / Transaction)	Community	Risk Score
1	12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw	1	1
2	1LZ9WozeiHEQWE3JQbikGHLXa6qiKLXJjN	4	0.9999999066
3	1BvTQTP5PJVCEz7dCU2YxgMskMxxikSruM	1	0.9999987018
4	1EqmnkqcN9MQzPLH8zVMGeJwicENe9PQhz	4	0.999752307
5	19JCSFRPyXnVn7ptXyqmhLKNBAmPcksZS6	3	0.999649391
6	1D3U69XBgqpvHwU5q6U2237AWWnGyFUVz7	4	0.9990430646
7	340b44c7a7857e36f81b2e8ba713911ea93e82afde6ea5590df1a35688845d16	3	0.9983512281
8	1BmnfremSpD7GBySqi4mqvz3Pc1JiVQvoG	4	0.9982444506
9	178NdniKXphgQifBKcuxbTQyU1ro42k3LP	4	0.9980235722
10	1FN5JaTm1EdMjou9Rydu4htvRHXDHK8Hzr	4	0.9977387825
11	1CZH527GEeR5WDyGac5WHrD6tnW5qJkFGR	3	0.9977009081
12	1P9MofjYUCquw5AzhdHT4uQsf2hrNuqojL	4	0.9976210255
13	1697wz4hwVXNqNa7WaAVyLQ8UAdLP3JyNA	0	0.9975391621
14	1Ph8BFiuxPWS22V24qBcdDMkLu5oq8vCqE	4	0.9972508058
15	1HY8jUGFg6DwE3fHEwVFViHjZLQwLwV4Tj	4	0.9964831496
16	1Dha5e1jbTtu4YGALQ3DnfTAk5yxzm4XSR	0	0.9944175344
17	1b2a3333f583ae54dba78ccc71f4fe24a22acd0991d364e75bcf099ce3a84759	2	0.9941592119
18	131551e35e7a644b76ea5366f744313bff3f959207c416f7b7b7f9b1cc90b0a3	3	0.9937028227
19	193C6ArL53dZ5k3wSSe8a4FjVx56Drir74	4	0.9935641946
20	1ANeJQbuuDxBATkWRaHL378EczzbA1zUyu	4	0.9931619272

Table 1 - Top 20 by risk score for nodes in the WannaCry Ransomware seed address 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw cash-out network.

Figure 4 shows the distribution of the risk scores and the respective node index (address or transaction) for the WannaCry cash-out graph. In this figure we concentrate on those nodes with riskiness ranging from 0.95 to 1, representing the top 20% of nodes by risk score. On the x-axis, node riskiness is represented by a score from 0 to 1 across the entire network dataset. Where a score closer to '0' shows little similarity relative to the ransomware seed address and can be interpreted as a node in the network that exhibits little risk when it comes to facilitating the cash-out of ransom collected. On the other hand, those scores closer to '1', show a similarity or closeness to the ransomware seed address. The actual riskiness should be viewed from both the X & Y measures as it is the cosine of the coordinate point of one node relative to the ransomware seed address which is always at position (1,0). Because the analysis is normalised the radius (or arc in this case) will not exceed a radius of 1 as it moves from (1,0) to (0,1). This calculation removes the emphasis on magnitude of the vectors and measures the angle between two nodes showing a relative importance to the network no matter how many levels deep in the network we move away from the ransomware seed address.

The ransomware seed address, 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw, sits at position Y=0 and X=1 on the chart and the next closest node 1LZ9WozeiHEQWE3JQbikGHLXa6qiKLXJjN represents an address that has a node riskiness of 0.9999 (see table 1) and is therefore deemed critical to the movement of funds out of the ransomware seed address.

The second most risky node with risk score = 0.9999987018 is another Bitcoin address 1BvTQTP5PJVCEz7dCU2YxgMskMxxikSruM. This is an address directly linked to the Poloniex.com exchange where the WannaCry attackers cashed out their proceeds of crime [20].

Looking into why these nodes are deemed risky in the context of this research, we could determine the Bitcoin address, 1LZ9WozeiHEQWE3JQbikGHLXa6qiKLXJjN, to be a false positive in our detection system as it seems to be part of a bigger cluster of nodes, centred around the transaction (ID: 29779df2e2a5a1f823b22e7e974a0082bdfd389edc1c11d1d4f6b290d8118d27) contributing a small amount of Bitcoin (0.0034398 BTC) taking place on 31st August 2017 at 16:32:00 UTC. Considering the WannaCry campaign cashed out on the 3rd August 2017 from the ransomware seed address [21], this has greatly exceeded the campaign time window and targeting this particular node might provide little impact on the risk of the network fulfilling the objectives. However, some forensic analysis might be warranted. This address is one out of 236 other addresses taking part in peeling activity which ultimately outputs to an address (1ETWkyQUY9nRpVMyGwha4vRhwKgMbomMQe) linked to another exchange, HitBTC.com which could be targeted for investigation for playing a part in soliciting illegal ransomware money flows (Neutrino, 2017). As previously mentioned for the next risky node, 1BvTQTP5PJVCEz7dCU2YxgMskMxxikSruM, it has a direct link to the exchange Poloniex.com. This can be interpreted as a true positive result from the analysis system shown in figure 2. Looking at the detail behind this address, it directly receives 17 BTC, the full amount of ransom collected from the WannaCry campaign on the 3rd August 2017 at 10:04:51 UTC. The same time a twitter bot known as @actual_ransom identified the first outflows from the WannaCry attackers' wallets. This bot was set up by journalist Keith Collins to monitor activity of the WannaCry ransom addresses [22].

5.5. Risk in communities

To complement the derivation of the risk score is the identification of additional data that has been extracted from the walletexplorer API and collected as part of the analysis system depicted in figure 2. This takes the form of 'labels', 'PageRank' and 'community'. The labels nominate what service the node belongs to and provide a strong indicator for the attribution of real world identification into the Bitcoin ecosystem. In figure 4, we can see HitBTC and Polinex on two of the highly ranked

nodes (1BvTQTP5PJVCEz7dCU2YxgMskMxxikSruM and 1Dha5e1jbTtu4YGALQ3DnfTAk5yxzm4XSR) indicating these could be cash-out exchanges used by the attackers. In addition, PageRank is represented by the size of the bubble in figure 4 and defines node influence in a network based on the frequency of its connections to other nodes [23]. That is the larger the bubble in figure 4, the larger the PageRank and the larger the influence of the node in the network.

which is a transaction, 29779df2e2a5a1f823b22e7e974a0082bdfd389edc1c11d1d4f6b290d8118d27 [PageRank=40.4775] occurring on 31st August 2017 having over 230 inputs with an output connected to an address (1ETWkyQUY9nRpVMyGwha4vRhwKgMbomMQe) controlled by exchange HitBTC.com. Despite these intricate connections these nodes only yield a risk score of 0.987784659211026 and 0.964703062973474

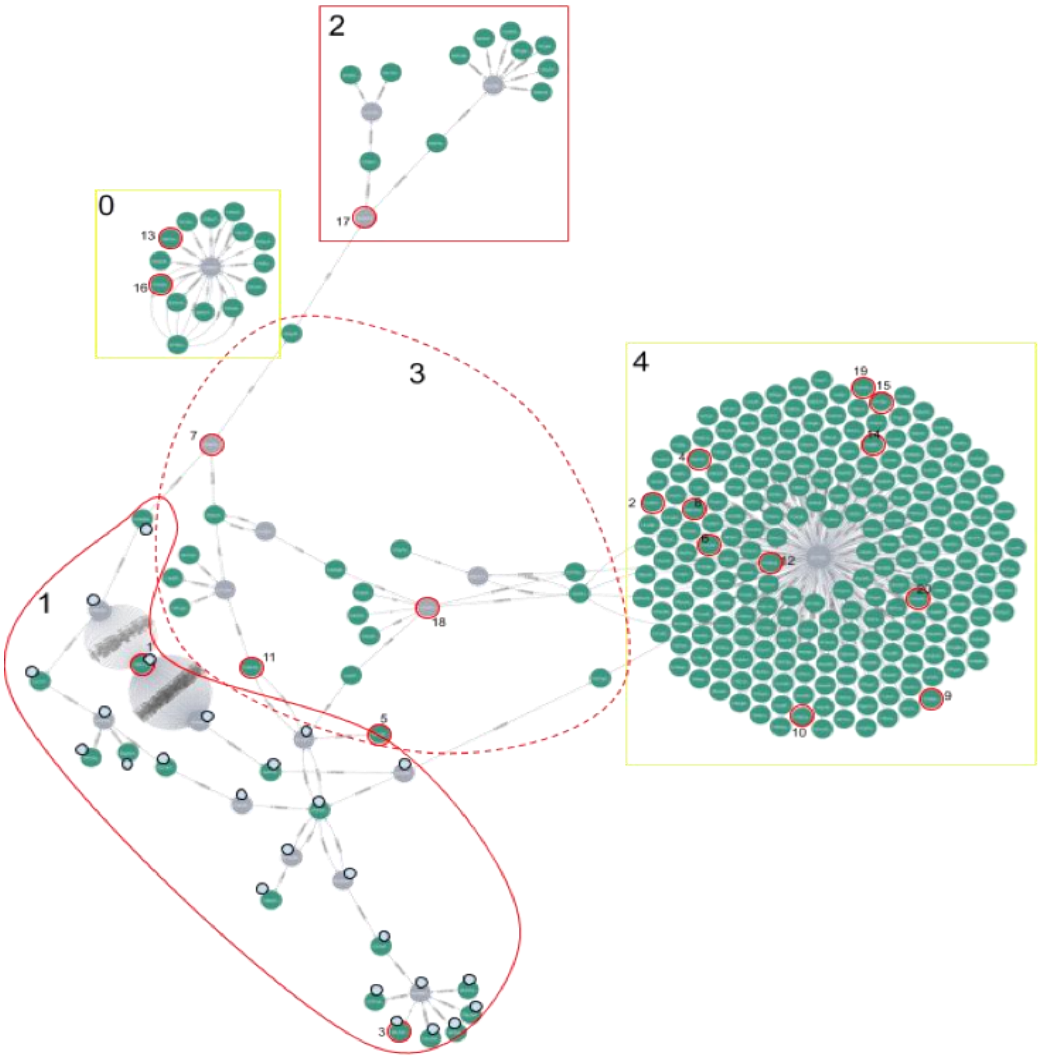

Figure 5 - Graph representation of the WannaCry Ransomware seed address *12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw* cash-out network

It is interesting to observe the node position relative to PageRank and the risk score. The second largest page ranked nodes, 1ArG3JwEbF4WrCiEnXQXUAgQumAVzqnQHD [PageRank=20.493] is used as a change address during the WannaCry campaign on 4th August 2017 and subsequently linked to the largest page ranked node respectively and are positioned well outside the top 20 risky nodes identified in table 1. Therefore, in this instance, little correlation can be derived between the risk score and page rank. However, using the combination of risk score on individual nodes and community detection it is possible to augment decision intelligence on what areas of the graph to monitor and investigate. This is illustrated in Figure 5.

Community detection is another common fraud detection technique used on networks to identify communities of nodes exhibiting anomalous behaviour that can be targeted for investigation [24]. Attributing a risk score to these communities on the aggregate we can determine which communities pose the greatest risk to the successful fulfilment of the network objectives. Table 2 demonstrates this.

Community	Node Count	Avg. Riskiness	Median Riskiness	Avg. PageRank
0	15	0.568214791	0.5539839493	0.3471999825
1	25	0.7281286781	0.7821614957	0.3968711842
2	14	0.7220384599	0.866469264	0.384656545
3	21	0.7797696805	0.8503920419	2.629416667
4	224	0.623651481	0.6648694238	0.3300336552
Grand Total	**299**	**0.6451774993**	**0.7116752424**	**0.5005359857**

Table 2 - Median risk score grouped by community for nodes in the WannaCry Ransomware seed address 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw cash-out network

By using the median riskiness for the communities, it is possible to see how high the middle score is in the ordered set of risk scores for that community. The higher that middle score the higher the concentration of risky nodes to go after. This can be further validated via the graph visualisation in figure 5. The nodes highlighted by the red circles indicate the top 20 risky nodes from table 1 and the groups of nodes encircled by the shapes highlight the communities of nodes from table 2.

It can be seen from table 2 and figure 5 that communities two and three share the highest community risk scores. Community three contains four of the top 20 risky nodes and visually plays a very central role in the facilitation of cashing out the proceeds of the WannaCry ransom. Node number 7 is a transaction within community three, 340b44c7a7857e36f81b2e8ba713911ea93e82afde6ea5590df1a35688845d16, that handles 8.715 BTC of the collected ransom and routes 6.877 BTC through community three on 3rd August 2017 and a further 1.8376 BTC splits off into community two. Community three acts as a mixing community to obfuscate this portion of the ransom with the transaction at node 18 (131551e35e7a644b76ea5366f744313bff3f959207c416f7b7b7f9b1cc90b0a3) combining the ransom cash-out with four other inputs to produce an output of 32 BTC on the 4th August 2018 to HitBTC.com owned address 1ETWkyQUY9nRpVMyGwha4vRhwKgMbomMQe. A considerable sized transaction heading to an exchange that would certainly raise suspicion. An interesting observation on community two is that even though it has a high community risk score it only contains one of the top 20 risky nodes, a transaction 1b2a3333f583ae54dba78ccc71f4fe24a22acd0991d364e75bcf099ce3a84759, ranked 17th in table 1, occurring on 3rd August 2017 which facilitates 1.8376 BTC of the cash-out for the WannaCry ransom via one input address and two output addresses. This is where the combination of the risk score and community detection provides further targeted analysis. If we were only to go on the list of risky nodes in table 1 the investigator could spend their time looking at community four where 11 of the nodes reside. However, examining the collective reveals the median riskiness of that community is only 0.66 (see table 2). Community four is also the largest community by membership and the relevance of the risk score dispels the myth that a more populated community would produce a higher concentration of risky nodes.

5.6. Targeted disruption

Now that there is a way of identifying risky nodes in the ransomware-Bitcoin network, intervention can be considered to target these nodes and disrupt or eliminate them. Looking at figure 6 which is a replication of figure 5 with one of the risky nodes, transaction 131551e35e7a644b76ea5366f744313bff3f959207c416f7b7b7f9b1cc90b0a3, node 18, removed.

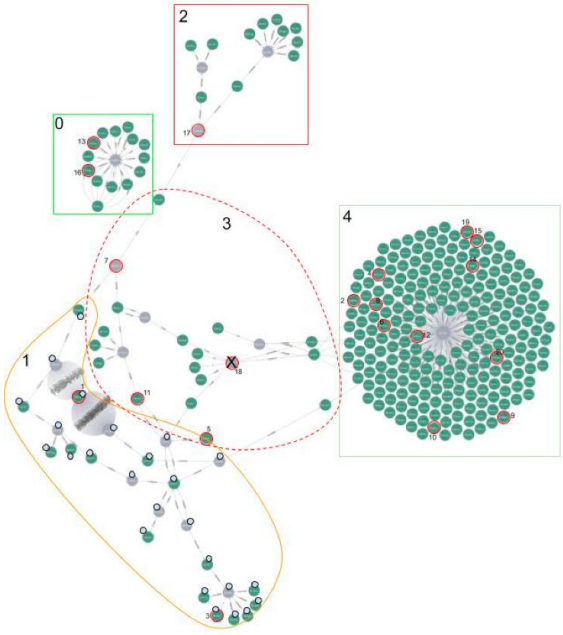

Figure 6 - Graph representation of the WannaCry Ransomware seed address

12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw cash-out network, disrupting node 18

At first glance this looks like a good tactic, severing the transaction at node 18 will inhibit the ability of the attackers to continue cashing out their proceeds of crime. However, the practical implications of doing this are not so simple. Node 18 represents a transaction with ID 131551e35e7a644b76ea5366f744313bff3f959207c416f7b7b7f9b1cc90b0a3, the details of this transaction can be seen in figure 7. If it were possible to disrupt this transaction, there would be significant impact to the attackers fulfilling their objectives. Tracing the amount from the ransomware seed address, this transaction receives 5.1309 BTC of the ransom from address 1HQiNjBRrHZpuyaWYXnCMhwcvJPqF5e97M. This amount is combined with inputs from four other addresses to send a total of 32.02476446 BTC to address 1ETWkyQUY9nRpVMyGwha4vRhwKgMbomMQe which belongs to exchange HitBtc.com.

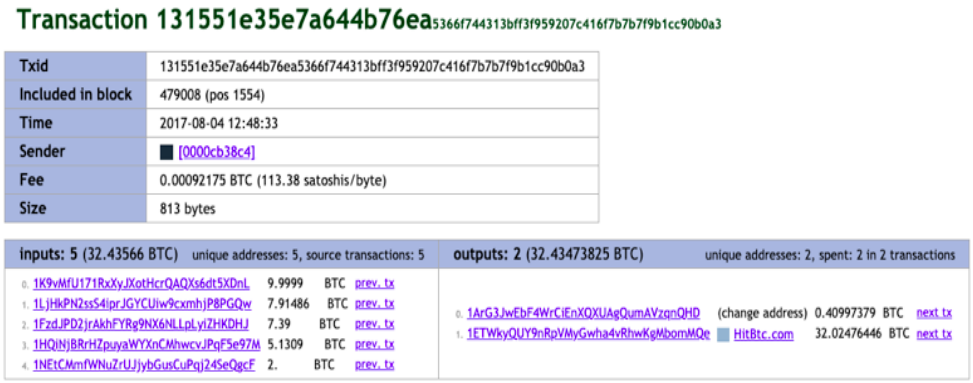

Figure 7 - Transaction details for transaction ID: *131551e35e7a644b76ea5366f744313bff3f959207c416f7b7b7f9b1cc90b0a3* (screenshot courtesy of walletexplorer.com)

It would take significant effort, knowledge and real time action to be able to disrupt this transaction. This would have to be done in near real time by corrupting the transaction script by hacking at the Bitcoin software as was the case when Mt Gox destroyed 2,609 BTC [25]. Alternatively, fictitious addresses can be simultaneously generated with their public and private keys, at the time of the transaction, to receive payments and sign the Bitcoin over to the next owner in the chain and divert the ransom funds away from being exchanged at HitBtc.com [26].

6. Limitations

The concept of the similarity analysis and the application to a ransomware-Bitcoin cash out network was only applied to one of the WannaCry 2.0 ransomware seed addresses (12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw). The analysis system is highly dependent on the quality of graph embeddings produced by the DeepWalk algorithm. Whilst preserving the structure of the nodes of the graph in relation to each other, the embedding algorithm used in this analysis was still only in development and not released in the Neo4j (graph database) production library for graph data science. Therefore, at this stage, validating the quality of the embeddings is difficult. In addition, using the output of the DeepWalk algorithm as input features to the cosine similarity score, the risk rating or recommendation on which nodes to attack in the network for intervention in illicit money flows has a dependency on knowing the ransomware seed address. Nevertheless, the system can still be used to initiate responses based on the 'riskiness' score obtained from the cosine similarity calculation and auto classify existing and new nodes coming into the network. To be effective in this manner the operation would need to be done in near or real-time. For example, we see the cash-out activity for the 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw ransomware seed address during the WannaCry campaign all happen within the space of six hours. The initial transactions out from the ransomware seed address, (409803bb5e124fd028c0482027c7722e84ce55b78204b279d3a44aba5e7c1698 and 35e5d5fe8c8128cfa6884f56be5817e4138c58c91b79d78d3e78a8d365b9d8a7), began at 03/08/2017 04:28:20 UTC. The transaction

(36ef488e59d719fb906254aed61bfe46e8f64778bc6cac97e56a68c241004c28) that facilitated a cash out at the exchange Poloniex.com occurred at 03/08/2017 10:04:51 UTC.

7. Future Research

Considering the most targeted pieces of information revealed from the similarity analysis are the identity of the address node and its risk score. There remain gaps in the available identity information from the raw data. Nonetheless, several features could be used in further machine learning techniques to predict the nature of nodes. A prediction algorithm could be built that would identify, for example, probable exchange services or other types of categories such as whether a node is involved in ransomware or not. This would allow analysts and investigators to estimate the location and ultimate owner of the address in the Bitcoin network removing significant barriers to the anonymity afforded to nefarious actors using cryptocurrencies. This would be an enormous improvement given the magnitude of the gap in the raw data. For example, in the data on the cash-out graph for WannaCry ransom seed address 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw only 3 out of the 280 addresses are labelled with exchange services (approx. 1%). Table 3 highlights Poloniex.com and HitBtc.com as the most prominent exchange used when cashing out the ransomware proceeds.

n.label	n.index
"HitBtc.com"	"1ETWkyQUY9nRpVMyGwha4vRhwKgMbomMQe"
"HitBtc.com"	"1Dha5e1jbTtu4YGALQ3DnfTAk5yxzm4XSR"
"Poloniex.com"	"1BvTQTP5PJVCEz7dCU2YxgMskMxxikSruM"

Table 3 - Available labels on the 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw WannaCry ransom seed address

Building a prediction engine at scale to assist attribution of anomalous nodes in the network is outside the scope of this research paper. However, the data collected from the analysis system paves the way for future research in this area. As an example, cryptocurrency forensic analysis firm Elliptic and researchers at IBM and Massachusetts Institute of Technology (MIT) have released a public data set of around 200,000 transactions partially labelled with illicit or non-illicit flags to identify suspicious transactions on the blockchain within the context of Anti-money Laundering (AML) [27].

Understanding a graph in the past helps create a baseline for what to look for in the future. In order to understand how a current scenario relates to that baseline it is important to know what has changed and what hasn't. This helps detect any anomalies, or unusual patterns, within the dynamic nature of a Bitcoin blockchain graph. More sophisticated algorithms such as, Microcluster-Based Detector of Anomalies in Edge Streams (MIDAS), are able to detect dynamic behaviours in graphs [28]. This lends itself well to the Bitcoin - blockchain environment as the graphs formed here are constantly being updated with new addresses and transactions. In addition, when it comes to discovering ransomware graphs in such an environment micro cluster detection helps detect sudden bursts of activity on nodes or edges, which are common to the behaviours of both the cash in and cash out graphs in ransomware / Bitcoin activity [29].

8. Conclusion

This research paper draws insights into using machine learning techniques combined with human interpretation to identify nefarious nodes in the WannaCry ransomware-Bitcoin cash-out network. The focus of this paper has been on using the Cosine Similarity calculation on DeepWalk embeddings to define a risk index that identifies what nodes, if eliminated from the network, carry the greatest risk to the attacker achieving their objectives, i.e. cashing out collected ransom payments. Using the Cosine Similarity as a risk index on an individual basis may not yield a targeted disruption of the network objectives. However, when the risk index was taken in combination with community detection a more powerful analysis emerged to isolate risky sections of the network. In particular, the practices of graph embedding and principal component analysis provide a truly reusable set of features for future machine learning applications. Furthermore, finding mechanisms to estimate the identities of nodes on the network will help attribute nodes with a particular Bitcoin service. However, limitations are evident with these techniques having only used a data set relating to the WannaCry ransomware-Bitcoin cash-out network. One broader benefit to the research community would be to open source multiple ransomware-Bitcoin network data sets for validation of analysis techniques.

Significantly, the entire approach remains predicated on identifying the ransomware seed address to build the target network.

9. References

[1] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations", In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, August 2014, pp. 701-710.

[2] Chainalysis, "The 2020 State of Crypto Crime", Chainalysis, January 2020. [Online] Available at: https://go.chainalysis.com/rs/503-FAP-074/images/2020-Crypto-Crime-Report.pdf (accessed 18 May 2020).

[3] C. Cimpanu, "Ransomware mentioned in 1,000+ SEC filings over the past year", April 30, 2020. [Online] Available at: https://www.zdnet.com/article/ransomware-mentioned-in-1000-sec-filings-over-the-past-year/? (accessed 18 May 2020).

[4] B. Sussman, "New Orleans Ransomware Attack Costs Climb to $7 Million", Wednesday, February 5, 2020. [Online] Available at: https://www.secureworldexpo.com/industry-news/new-orleans-ransomware-attack-update-cost (accessed 18 May 2020).

[5] M. De Marzi, "Finding Fraud", August 19, 2019. Available at: https://maxdemarzi.com/2019/08/19/finding-fraud/ (accessed 18 May 2020).

[6] D. Voutila, "Analyzing First Party Fraud with Neo4j", Available at: https://www.sisu.io/posts/paysim-part3/ (accessed on 18 May 2020).

[7] E. A. Lopez-Rojas, A. Elmir and S. Axelsson, "PaySim: A financial mobile money simulator for fraud detection", In: The 28th European Modeling and Simulation Symposium-EMSS, Larnaca, Cyprus. 2016.

[8] B. A. Galler & M. J. Fisher, "An improved equivalence algorithm", Communications of the ACM, 7(5), 1964, pp. 301-303.

[9] P. Reeves and G. Wilcock, "Anti-Money Laundering", 2019. [Online] Available at: https://gettingthedealthrough.com/area/50/jurisdiction/5/anti-money-laundering-australia/ (accessed on 18 May 2020).

[10] European Union, "The 5th Anti-Money Laundering Directive", 2018. [Online] Available at: https://ec.europa.eu/info/business-economy-euro/banking-and-finance/financial-supervision-and-risk-management/anti-money-laundering-and-counter-terrorist-financing_en (accessed on 18 May 2020).

[11] K. M. Hayne, "Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry", Australian Government, 2019. [Online] Available at: https://www.royalcommission.gov.au/sites/default/files/2019-02/fsrc-volume-1-final-report.pdf (accessed 18 May 2020).

[12] J. Frost, "Banking royal commission: CBA finds 45 new money-laundering flaws", 2018. [Online] Available at: https://www.afr.com/companies/financial-services/banking-royal-commission-cba-finds-45-new-money-laundering-flaws-20181120-h184fs (accessed 18 May 2020).

[13] G. Flood, "Graph proves an ideal way to answer the 'Who Owns Me?' anti-money laundering question", March 13, 2020. [Online] Available at: https://diginomica.com/graph-proves-ideal-way-answer-who-owns-me-anti-money-laundering-question (accessed 18 May 2020).

[14] C. Miles, "Detecting Cryptocurrency Fraud with Neo4j", March 30, 2020. [Online] Available at: https://neo4j.com/blog/detecting-cryptocurrency-fraud-with-neo4j/ (accessed on 18 May 2020).

[15] R. M. Clark & W. L. Mitchell, "Target-Centric Network Modeling: Case Studies in Analyzing Complex Intelligence Issues", CQ Press, 2015.

[16] F. Tong, "Graph Embedding for Deep Learning Graph Learning and Geometric Deep Learning — Part 1", 2019. [Online] Available at: https://towardsdatascience.com/overview-of-deep-learning-on-graph-embeddings-4305c10ad4a4 (accessed 18 May 2020).

[17] R. A. Rossi, R. Zhou & N. K. Ahmed, "Deep feature learning for graphs", arXiv preprint arXiv:1704.08829, April 28 2017.

[18] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments", Phil. Trans. R. Soc. A.37420150202, 2016. [Online] http://doi.org/10.1098/rsta.2015.0202

[19] J. Han, M. Kamber and J. Pei, "2 - Getting to Know Your Data", Editor(s): Jiawei Han, Micheline Kamber, Jian Pei, In The Morgan Kaufmann Series in Data Management Systems, Data Mining (Third Edition), Morgan Kaufmann, 2012, pp. 39-82, ISBN 9780123814791. [Online] https://doi.org/10.1016/B978-0-12-381479-1.00002-2.

[20] S. Bistarelli, M. Parroccini and F. Santini, "Visualizing bitcoin flows of ransomware: WannaCry one week later", ITASEC, ser. CEUR Workshop Proceedings, no. 2058, 2018. [Online] Available at: http://ceurws.org/Vol-2058/#paper-13 (accessed 6 November 2018).

[21] Neutrino, "WannaShift to Monero", Neutrino Research Team – September 2017. [Online] Available at: www.neutrino.nu/Research_WannaShift_to_Monero.html (accessed 7 November 2018).

[22] A. B. Turner, S. McCombie & A. J. Uhlmann, "A target-centric intelligence approach to WannaCry 2.0", Journal of Money Laundering Control, Vol. 22 No. 4, https://doi.org/10.1108/JMLC-01-2019-0005 October 7, 2019, pp. 646-665.

[23] M. Needham & A. E. Hodler, “Graph Algorithms: Practical Examples in Apache Spark and Neo4j”, O'Reilly Media, 2019.

[24] A. E. Hodler, “Financial Fraud Detection with Graph Data Science”, 2020. [Online] Whitepaper available from: https://neo4j.com/blog/financial-fraud-detection-graph-data-science-identifying-fraud-rings/ (accessed 18 May 2020).

[25] K. Sedgewick, “Bitcoin History Part 17: That Time Mt. Gox Destroyed 2,609 BTC”, September 29, 2019. [Online] Available at: https://news.bitcoin.com/bitcoin-history-part-17-that-time-mt-gox-destroyed-2609-btc/ (accessed 18 May 2019)

[26] P. Ducklin, “Serious Security: How to cut-and-paste your way to Bitcoin riches”, SophosLabs, 05 July 2018. [Online] Available at: https://nakedsecurity.sophos.com/2018/07/05/serious-security-how-to-cut-and-paste-your-way-to-bitcoin-riches/ (accessed 18 May 2020).

[27] M. Weber, G. Domeniconi, J. Chen, D. K. I. Weidele, C. Bellei, T. Robinson and C. E. Leiserson, “Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics”, Tutorial in the Anomaly Detection in Finance Workshop at the 25th SIGKDD Conference on Knowledge Discovery and Data Mining. 31 Jul 2019.

[28] N. Mishra, “Anomaly detection in dynamic graphs using MIDAS: An interesting approach to modelling network security”, March 8, 2020. [Online] Available at: https://towardsdatascience.com/anomaly-detection-in-dynamic-graphs-using-midas-e4f8d0b1db45 (accessed 18 May 2020).

[29] S. Bhatia, B. Hooi, M. Yoon, K. Shin & C. Faloutsos, “MIDAS: Microcluster-Based Detector of Anomalies in Edge Streams”, arXiv preprint arXiv:1911.04464, Feb 10 2020, pp. 3242-3249.