# School of Informatics


## Informatics Research Review
## Anti-Money Laundering (AML) Efforts in the Bitcoin Network

[redacted]

**January 2021**

**Abstract**

Anti-money laundering is a persistent question which keeps on refreshing when a new technology emerges. Currently it is the Bitcoin that makes the turbulence. Efforts is a broad word, which in our case includes justifying the feasibility of money laundering (ML) using Bitcoin, discovering the incentives behind criminals using Bitcoin ML, concerning adopting and coining regulations and utilizing modern techniques to study AML using unprecedentedly huge public data.

Date: Friday 29th January, 2021

**Supervisor:** [redacted]

# 1 Introduction

Criminals always act fast facing the emergence of new technology in an attempt to discover regulation loopholes, adopt novel methods and maximize their illicit profits. Since the launch of Bitcoin [1], the attention of both legal and ill-willed internet users has been drawn towards its alleged anonymity. Among a plethora of questions, we focus on the money laundering (ML) feasibility and try to discover what has been done in favor of anti-money laundering (AML) in the Bitcoin network.

Both traditional ML/AML and Bitcoin are huge topics. I will only refer to their special patterns or traits to introduce what AML means in the Bitcoin network. To start with, is this really feasible for money laundering to take place with the help of Bitcoin thus worth our time exploring corresponding AML methods? Some scholars once identify such concern as "virtual threat" [2] since that at the early years not much evidence is present to justify its possibility. That is why exploratory experiments were carried out to mimic and analyse the money laundering process.

Along with the experimental support, real life cases of conducting actual crime including ML with Bitcoin are revealing [3] in accordance, whose presence and arrest also boost the AML research. Interestingly though, existing arrest cases mostly attribute to criminals' sloppiness which can be captured via traffic interception etc. However, I will not discuss such efforts which drift away from Bitcoin itself. Given the feasibility of ML in the Bitcoin network, incentives for criminals are then studied by asking whether this method is lucrative enough to take the risk of exposing all transaction history in public even though they're under "pseudonym".

What shall be considered as AML methods in the Bitcoin network? Know-Your-Customer (KYC) principle is always mentioned along with AML. However, conventionally, it won't work without an intermediary and a complete procedure from profile authentication to the filing of suspicious activity reports (SARs). Regulations then won't fit Bitcoin into the existing AML scheme. So shall exists change in regulations or change in bitcoin or, first of all, shall gather focus on the entities involved yet bounded by the regulation, namely exchange services.

Despite the conflicts, diversities of regulations across the world, what can individuals do? Some attempt to de-anomynize the system by tying real person to its possessed public key. Some further try to discover yet-unidentified possession using machine learning algorithms. Apart from utilizing real world information, others focus on the anomaly attack. Compared to regular transaction or users, those involved in ML are assumed to adopt certain behaviours. Some try to characterize those behaviours in the Bitcoin network. Some use unsupervised or, unexpectedly, supervised learning algorithm to capture such abnormality. Such research is in a way mostly discussed during the last few years. However, I still notice lack of sharing, complete dataset and well-acknowledged standard i.e. in evaluation method. Two other interesting perspectives lie on 1) investigating real life suspect to determine the usage of Bitcoin in ML; 2) jumping out of one-cryptocurrency scenario and testing the "crypto-to-crypto" exchange services for potential ML.

Since that the discussion on AML in the Bitcoin network is relatively new, some researches on intriguing aspects such as de-mixing mixing services are not yet available for pre-print. However, from the papers I reviewed, I have several future suggestions. 1) If we look aside from utilizing Bitcoin public data, it's recommended that more research on the vulnerabilities of Bitcoin and, most of all, relevant ML tools such as mixing service should be conducted. Such vulnerabilities may leave criminals traceable in a short cut. Once vulnerabilities are publicly known, wise criminals will avoid using such risky tools and fall back into the traditional regime under AML regulation. 2) Universal consensus over Bitcoin regulation is hardly possible to be

reached, always leaving cracks a.k.a less strictly regulated areas for criminals to elude. Before the absolute acceptance or end of cryptocurrency, it's wise for authorities to keep it a balance and focus on more important combat towards financial crime. 3) Meanwhile, scholars can explore more after the detection of an abnormality to discover more patterns of money laundering either manually or better utilizing machine to predict. It is also suggested to develop robust system that can deal with acute change in the Bitcoin network, which can be seen as incorporating context information.

## 2 Literature Review

Bitcoin is a decentralised, peer-to-peer payment network [1] which has an identity called pseudo-anonymity. It shares quite the contrary traits with conventional fiat money transaction. In a traditional transaction, if not using cash, normally an intermediary such as a bank is involved. The bank knows every confidential detail of the user and the transaction, yet the transaction is only otherwise known between the two users. While in Bitcoin network, the user is not in the unit of a person but an address that can so easily be created without informing anyone his real identity. All the transactions though are visible on the blockchain, a public ledger of transactions, where anyone has access. For further review, graph neural network was proposed at almost the same time as the idea of Bitcoin. [4] It is one of the novel ideas used to process Bitcoin block chain data.

Money laundering is criminals' attempt of washing the ill-gotten money clean, making it appearing from a clean source. It involves three steps of 1) introducing the money into the financial system 2) try to disconnect money from its dirty origin 3) have the cleaned money reenter the system and prepared for use. To get to the last secure stage, criminals always try to minimise the risk of exposing 'washing' procedure and minimise the cost along the way. This is why when a new payment method such as Bitcoin emerges, both the criminals and investigator would look into the ML risk of using Bitcoin which may give better security and profit.

AML control is the main way of catching money launderer. It's considered to have five stages of 1) internal training of a cooperative entity 2) KYC procedures 3) user activity monitoring 4) manual review on suspicious user or transactions 5) filing SARs for further investigation. [5] Current conventional AML controls are considered complete. As the basis of further review, there exists machine learning utilization on money laundering anomaly detection [6] and a first implementation of scalable graph convolutional neural networks for AML [7]. They both inspired the technique to be used for AML in the Bitcoin network.

However, Bitcoin is bypassing those regulations without effort. Firstly, there was difficulty fitting Bitcoin into the definition of money, which has three functions being "medium of exchange, store of value, and unit of account", due to its volatile value and never a unit. [2] Furthermore, it can't even be called "electronic money". Money Laundering Regulations 2007 therefore is of no use. Secondly, anyone can create an address on the blockchain and perform transaction without identifying themselves. This lack of KYC regulation on the blockchain makes every transaction a mystery until we move our attention to the origin of ML. Consider that illicit money can come from two origins — another address in the blockchain or exchange service via buying in. For the common ultimate goal of money laundering, Bitcoin shall always be exchanged into a kind of fiat money to be of use. Under each combination of ML procedure, exchange service is always used which deals with fiat money, leaving it under the current regulation. KYC requires that "financial institutions and other regulated companies must perform to ascertain relevant information from their clients for the purpose of doing business with them"

[8]. Thus, so long as we maintain an explicit flow of dirty money inside the block chain, we can know of the identity of the criminal at the exchange service end. It's also supported by [9]. Most advocators recommend introducing Bitcoin currency into the previous AML frameworks. [10][11] The recent notable change is in May 2019, when "the Financial Crimes Enforcement Network (Fin-CEN) of the United States issued new guidance on how the Bank Secrecy Act (BSA) of 1970 applies to cryptocurrency". [12]

As we said that the money flow inside the blockchain is also important. In the study on cryptoasset [8], statistics has been given, showing that "half of all cryptoasset-only entities perform KYC/AML checks in some capacity" and "the vast majority of entities conducting KYC/AML checks inspect every account", which shows a good practice of KYC implementation in the cryptoworld. However, just like exchange services, internal services vary on their code of conduct. Exchange services in less regulated countries give the criminals a second chance. So do the other half of internal services. Among which, mixing services, gambling sites offer a disturbance in the tracking of illicit asset.

### 2.1 Feasibility and Incentives

In 2013, [13] experimented to tell the feasibility of large-scale illicit transaction with Bitcoin. The data is self-scraped and information provided is not sufficient for reproducing the result. Some of them are labeled definitely to be some known identity due to solid knowledge and others are semi-labeled as the source is less reliable like forums. It involves two phases of 1) re-identification attack and 2) public key clustering. The public keys of certain Bitcoin merchants are exposed due to a large number of interactions with our experimenters. Those addresses are tainted and further used to taint the whole cluster of keys. The cluster is based on manually determined heuristics 1) the addresses as the inputs to the same transaction are deemed as the same user. 2) the concept of change address is introduced that "if an address looked like a one-time change address at one point in time (where time was measured by block height), and then at a later time the address was used again, we considered this a false positive", which makes the heuristic more conservative. It's said that mixing service may not be able to handle large volume of transaction but small-scale illicit transaction is possible. Authors explained the concept of peeling-chain and claim that they're able to track flow using heuristic two mentioned above. However, they seem to be deliberately avoiding the involvement of mixing service. Each peel can use a mixing service to find itself a clean return. It is considered a necessary reduction to focus on the topic and proposal. Yet, more complicated conditions should be considered.

In the same year, there was a study on different Bitcoin anonymizers [9]. Authors interacted with those anonymizers and then extracted corresponding transaction graph from the block chain. Bitcoin Fog, BitLaundry and the Send Shared functionality of Blockchain.info. are analysed. They found clear patterns of operation for certain services and proposed utilizing context and linkability information to attack those mix networks. Since it was eight years ago, when the exchange rate of Bitcoin hasn't soared. It is considered reasonable to test mixer behaviour by first interacting with them. However, It will be a costly experiment. Nevertheless, the patterns of mixers won't always stay the same, which makes it lack sustainability.

In 2017, authors [14] started with a mixer. They tried to characterize all the Bitcoins related to this mixer and also used a taint analysis to measure whether the mixer is effective in disconnecting corresponding input and output. They studied all sorts of workflow of different mixers including DarkLaunder, Bitlaunder and CoinMixer and other services like Helix and Alphabay. It is clear that there is always a flaw in the pattern no matter how well-known the service is. They spotted certain vulnerabilities like storing personal data and all transaction relevant

information. It is not considered safe because any potential arrest of the owner of the service can lead to the overall leakage. Poor design and outdated techniques used for establishing the server also make it vulnerable. Just like the design criterion of proper randomness in block chain contract, any tag not randomly selected can be calculated and foreseen easily. The usage of a central node and a constant interval in the design can be detected and made use of. This experiment shares the same concept of experimental research and suffers from the same shortcomings. It is encouraged to give a comprehensive analysis of mixer behaviours.

In 2018, [15] performed several cash-out experiments on five mixing and five exchange services. It's on the premise that ML build heavily on mixers and exchange services, which [16] may beg to differ, since that statistics show mixers and online gambling services receive the most of the illicit Bitcoins. However, exchange services still earn a place in the AML scheme. Consistently, taint analysis is again the evaluation method in such illicit flow detection. It points out that reviews are very important in choosing the honest service, considering that three out of five services in the experiment are scams and the outcome matches the review. If a criminal lack technique knowledge, it is easy for them to make revealing mistakes like not using Tor browser, reusing an address etc. It also agree with [13] that larger amount of ML may not be successful. This argument expose one defect that such experimental tests hardly meet the real needs of criminals. It is frustrating for criminals to have to test for themselves to understand what's really concerned. Authors estimate that 15% of the funds can be the cost for this ML strategy, which is considered much lower and intriguing.

According to [5], several factors contribute to the incentives behind ML with Bitcoin. Bitcoin has low acceptability worldwide, which means, except for the popularity within the DarkNet, criminal may eventually have to use the exchange service to convert Bitcoin into fiat money or have no chance of using it. The price volatility fluctuate over time. The Bitcoin market is deemed to be volatile, in which an acute change of exchange price may easily draw the unwanted attention. [10] It leaves performing ML in large volume not practical. The attention of cybercrime has focused on Bitcoin since that it is "accounting for over 40 per cent of all identified criminal-to-criminal payments". [5] To name just a few setbacks, plenty of benefits may be irresistible. The administration into the Bitcoin network is easy and lax. No AML control yet is available to discover their transaction patterns. There are various tools like mixers to hide their trace and it is not systematically studied yet. There is no border in the Bitcoin network and the transaction is instantaneous and cheap.

## 2.2 De-anonymization

In 2015, [17] tried to incorporate Bitcoin public key with its real user. They assume that given the information about the details of a transaction, they would be able to find an exact match in the blockchain. Therefore, two parts of information is needed: 1) context information regarding the known detail of the transaction; 2) blockchain data in the corresponding time period. Both the information are scraped by themselves and not publicly available. The first step is to scrape from public sites such as forums to find the connection between a transaction and a real person, which most of the time is another ID or nickname in the real world. The tough part if how to match the time and value of the transaction into one among millions. Therefore they developed a consensus to calculate expected output value and a time window to allow tolerance in matching. Apart from applying the same heuristic one for [13], they introduced a novel idea of using Page Rank in seeing the resemblance of Bitcoin graph and search engine graph. It is reasonable to consider nodes with larger traffic as important. The effort made can imply which entities certain addresses have involved with. The result of discovering the FBI address related

with Silk Road looks promising, yet the lack of evaluation and metrics makes the experiment insufficient. It is one of the many differences among all the experiments we reviewed. We can notice that certain papers invented metrics or used common ones. Some use the help of experts.

In 2018, a supervised measure is proposed to de-anonymize entity types.[18] Different from using rare context information like [17], this paper is satisfied with a broader granule of anonymization. The data used is from Chainalysis, a Bitcoin analysis company. All the information scraped from the blockchain is manually classified, which both gives the author a neat dataset and limits the author with certain features including cluster types and size. One of the clustering methods called Co-spend clustering implies the same as the above-mentioned heuristic one. Other two kinds of clustering are Intelligence-based clustering, which is similar to [17], and behavioural clustering, which is based on the transaction behaviour including transaction structure to identify a wallet. Like most of the similar experiments, minimising false positives is the main target. The supervised learning algorithms are k-Nearest Neighbours, Random Forests, Extra Trees, AdaBoost, Decision Trees, Bagging Classifier, Gradient Boosting. Notably, in this imbalanced dataset, Synthetic Minority Over-Sampling Technique (SMOTE) is used to oversample the minority classes, which can be of comparison to [19] which refuse to use SMOTE. In this paper, Random Forest, Bagging and Gradient Boosting are the best model with or without using SMOTE, the accuracy of which are 73%, 74% and 77% respectively. The downside though is the fact that not only is the dataset imbalanced, the dataset is not complete. Since this experiment is to study entity behaviour, the subset of its behaviour can possibly cut off the important transactions indicating its identity. Still, the success of the experiment shows the benefit of flagging potential suspects to discover new entities.

### 2.3 Anomaly Attack

In 2013, [20] used unsupervised learning to detect anomalous behaviours. The dataset is from a trusted source. K-means is used to cluster users, the result of which is then applied with another unsupervised learning algorithm named RolX (Role eXtraction) to assign "roles" to nodes. Even though the experiment is successfully conducted, no validation method is used to justify the anomalous behaviours. The author claims that it's due to the lack of labeled data, which is understandable since its early year. However, there lies a hidden trouble of dataset utility. Even though in the later years when data is sufficient and public awareness is high, not much comprehensive dataset is available. Most of the researchers download and preprocess data by themselves which leads to tons of repetitive work and lack of standard. It is something the community will be aware of.

In 2016, Pham posted two papers regarding unsupervised anomaly detection. [21] In addition to detecting the behaviour normality of a user, whether a transaction is suspicious is also considered. Therefore, there are two graph regarding user graph and transaction graph, where the character of the node differs. The first experiment used dataset given by the University of Illinois Urbana-Champaign. Three methods are considered: k-means clustering, Mahalanobis distance, and Unsupervised Support Vector Machine (SVM). They are each implemented on two graphs. There is a bad adjustment though due to the time-consuming SVM. The author had to limit the dataset to 100000 for all methods which undoubtedly can cause huge information loss. The other dataset is from Stanford Network Analysis Project. They used k-means as a baseline and use the laws of power degree and densification and local outlier factor (LOF) method. Both experiments can detect apparent anomalies and known cases. Especially for the LOF method, a Trojan keylogger theft was detected. It was a success when we look at the probability of detecting rare anomaly among huge dataset. They also proposed new evaluation methods to

evaluate the LOF consistency. However, the accuracy of which hasn't been justified by others. The limited use of the first data set shall be reexamined, possibly using parallel computation.

In the same year, [22] used similar approaches for unsupervised fraud detection. It is still considered relevant since it's part of the evolve of anomaly detection which helps in AML as well (ML is associated with financial fraud). The data used is from the Laboratory for Computational Biology at the University of Illinois. It could be the size of dataset or the application method that makes the author claim the "LOF does not scale well in large dataset". It remains a doubt about what's key difference between it and [21]. Not only classical k-means is used, trimmed k-means is also implemented and presents a more robust clustering result. Due to the different dataset, we can not compare its effectiveness with former works. Although five detection out of thirty existing cases sounds promising, the total number detected are not revealed, which also matters especially in the crime scene. Unlabelled dataset presents the same concern as in the difficulty to validate.

In 2017, another three methods are used. [23] Notably, those three methods tackle the anomaly detection from different angles: 1) Isolation Forest (IF), which considers that an anomaly must be easier to "be separated from the rest of the random subdivisions of the feature space"; 2) One Class SVM (OCSVM) is an unsupervised SVM with a boundary separating normal and anomalous cases; 3) Gaussian Mixture Models (GMM) takes the presumption that the data is generated from a combination of gaussian distributions. Apart from those three separate method, the score of each one of them are averaged after proper scaling, getting a fourth score, which is supposed to average out the disadvantages each method resides. The data is self-collected. They even use Pearson correlation coefficient to drop extra columns. They took active valuation method with experts, who can analyze the results manually. However, the use of method like GMM presume the data distribution which is counter-intuitive with the fact that we desire to learn the unknown pattern. Also, the outcome of each learning method is too obscure to be interpreted thus leaving us no further insight into the actual anomaly behaviour to be learned.

In 2019, [24] tried to characterise ML graph using four types of graph features: immediate neighbours, curated features, deepwalk embeddings, and node2vec embeddings. Among which, the node2vec-based classifier yields the best result of an average accuracy of 92.05% and an average F1-measure of 0.94. The data is self-collected. They also label with several known service addresses. The two clustering heuristics are also similar to the ones in [13]. It's found out that laundering transactions are always related with "high in-degree/out-degree ratio". There is one novel prospect of using transductive graph embeddings to allow the transplantation from one graph to another. It is a feature that node2vec can not give. But Graph convolutions such as GraphSage may give a chance to predict unknown nodes in a different graph.

In the same year, a major contribution is made. [12] Different from multi-classification like [18], it uses binary classification. It contributes the Elliptic Data Set which resides a sudden change in the Bitcoin network at the 34th time step, which raises the awareness of the robustness of systems in the face of context gap. It is also why most experiment using this dataset is splitting the train and test set at the 34th time step. The Elliptic contains features both local and in the immediate neighbourhood. It uses "variations of Logistic Regression (LR), Random Forest (RF), Multilayer Perceptrons (MLP), and Graph Convolutional Networks (GCN)" to predict illicit transactions. In the face of minority class, it uses a GCN model to pass a weighted cross entropy loss for assigning more importance to the illicit transactions. Random Forest gives the best result of 0.971 in precision and 0.796 in f1. Worth mentioning, the enriched information which involves the graph information, always helps in improving the accuracy for each method. The most important consideration is what can be done to the system to have a robust prediction

when the environment suddenly changes.

In 2020, a novel supervised version of illicit transaction detection is proposed. [19] It is a first of its kind and all former experiments implementing unsupervised learning methods recognise implementing supervised method is infeasible due to the lack of labeled data. The data set used is the Elliptic data set mentioned above. In which, 2% of the transactions are illicit and 21% are normal, all others being unknown. Methods like Random Forest, Extra Trees, Gradient Boosting, Bagging Classifier, AdaBoost and k-Nearest Neighbours are used. Ensemble learning method is also used to combine the best performing models, which turn out to be Random Forest, Extra Trees and Bagging classifiers. The best result is such an ensemble learning with the accuracy of 98.13% and F1 of 83.36%. If we recall [18] in the de-anonymization efforts, the methods used are alike. However, in processing imbalanced dataset, this paper doesn't consider it suitable to use resampling methods for the Elliptic. Undersampling can cause the loss of information, which could be only worse considering that the data set itself is a reduced version from the entire Bitcoin blockchain. On the other hand, oversampling like SMOTE won't work on the aggregated data. Aggregated features are graph features extracted from the transactions that are one step away from the current transaction. Using SMOTE can cause an interpolation that confuse the predictor. What the paper presented is the prediction result on the already-labeled illicit and normal transactions. It can be more complete to predict what's in the non-labeled data and give an meaningful evaluation method.

In the same year, another supervised model detecting illegal entities is proposed. [25] It uses DT and RF as baseline and proposes a tree-based classifier. The data it uses is also self-collected and pre-processed, which is then contributed to the community. More contribution made is the evaluation metrics it used sensitivity, specificity, accuracy and prevalence. They are relatively new compared to the most-often used accuracy and f1. However, more justification shall be needed to prove the effectiveness of those evaluation metrics. Also, since relevant dataset is available, it is better to include the prediction outcome on those dataset, so as to make a relatively comprehensive comparison.

### 2.4 Out of The Box

One of the novel AML effort is considered to be from the police's perspective, which asks how can we tell whether criminals are using cryptocurrency. [26] introduced relevant tools and methodologies which includes red flags and indicators. The relatively more important effort is the attempt to trace transactions across cryptocurrency ledgers.

All the former experiments focus on one of the aspect of the Bitcoin like users, transactions, mixers, exchangers. But it is necessary to look into the problem in a little larger scope. What can other cryptocurrency act as in one part of the ML using Bitcoin. The number of cryptocurrencies is increasing rapidly. Automated trading platforms such as ShapeShift and Changelly help users perform cross-currency trades without effort. The role of which looks like a mixer and exchanger, a "trusted" third-party whom you send money to and receive money from "under an agreement". However, it remains a question whether it can serve as a useful tool in ML.

The author [27] collected cross-currency data from three different sources: 1) the cryptocurrency blockchain, which in our consideration is Bitcoin; 2) data got from interacting with the trading platforms; 3) information available from those platforms' public API. The author uses a thirteen month period of data from ShapeShift and eight different blockchains (out of sixty-five different currencies), which accounts for 60.5% of the transactions within ShapeShift and try to characterise its possible ML pattern. The author tried to download transaction information

from ShapeShift website which offers top latest transaction data in a five seconds interval. However, the author discovered that among the limited amount of interactions with the platform himself, certain transactions are not visible in the collected data. Considering the probability, many important transaction can be lost thus leaving the remained data less informative. It is a terrible design if it is indeed as the author claimed that it's the website who underestimated the volume of customers therefore data is missing. We would have to nevertheless deal with such trouble because it can't be changed by us. The mission target is to find the corresponding output to the target currency network of the source input. Therefore, the main factors to look into is the time and value. The two transactions shall have a reasonable time interval and an expected value. It is similar to the [17] in de-anonymization. So long as no complex transaction method is adopted like the ones studied in the [9], in which mixers can set time interval and the number of output addresses etc. for each transaction, we can deduce the approximate time and value according to the context information. The time can be checked via API and the value can be calculated using the exchange rate and standard fee. It's notable that the exact exchange rate used in the transaction is not available. When we're calculating the value, a tolerable range shall be considered. The common relationship heuristic is quite similar to other heuristics mentioned: multiple inputs address sending to the same output address or vice versa, then these addresses are related.

The author made multiple usage of the public API, which gives out many useful information. For address identification, each recipient address is queried using the API to get a confirmed response, which means that the address belongs to the platform. It leads us to the transaction in the other cryptocurrency network. When we got multiple matches, simply querying the API of all the recipient addresses. In the limited interactions with ShapeShift and Changelly, long delays are observed, which can assist setting the tolerable time range. It also gave an analysis of three types of transaction patterns: 1) pass-through transactions; 2) U-turns; 3) round-trip transactions. It supports the claim that complex cross-currency transactions are traceable.

Tricky problem still remains since that ShapeShift reuses deposit address, which can cause chaos in the transaction analysis, since that the two adjacent transactions using the same deposit address may not match because we delayed in the data collection and it's only the latest transaction detail that is kept. Due to the display of transaction, the analysis is only possible when all data are collected real-time, which potentially makes old history harder to trace.

# 3 Summary & Conclusion

Anti-money laundering is a persistent question which keeps on refreshing when a new technology emerges. Currently it is the Bitcoin that makes the turbulence. Efforts is a broad word, which in our case includes justifying the feasibility of ML using Bitcoin, discovering the incentives behind criminals using Bitcoin ML, concerning adopting and coining regulations, utilizing modern techniques to study AML using unprecedentedly huge public data.

It is supported both by experiment and reality that AML is a real concern regarding Bitcoin. The history goes with the suggestion of fitting Bitcoin into the current regulation and recent achievements have been made.

Experimental study on Bitcoin laundering tools does gave us insight into ML measures involving mixer, exchanger, Tor, etc. However, the experiment shall not be considered an overall analysis, which should go beyond the scope of several test cases and discover the ever-growing patterns of behaviour in a clever way.

Same can be said in other experiments. Just like AML ideally requires the cooperation of countries or like block chain needs miners to race, compete and record, the community of AML in the Bitcoin network shall cooperate instead of doing their own work. Among de-anonymization, anomaly attack and techniques out of the box, we can see a great deal of resemblance and difference in several ways. The heuristics used are similar in finding the owner of several addresses. The studying object is the user, the transaction or both. Context information is always necessary for data relation outside of a blockchain.

Experiments all face dataset problems, meaning that there is no relatively sufficient labeled dataset yet. Some contribute and some complain. Those who considered extra all agree on that neighbourhood information is helpful in providing graph information. The majority of them understand the importance of characterizing graph features. Experiments all face evaluation problems, upon which some restrict themselves to the safe accuracy and f1, others bravely introducing new metrics. The approach is always evolving with the new learning methods. From unsupervised learning methods to supervised ones. Introducing new ideas like GCN and GraphSage.

In the future, emphasizing more on the characteristics of transaction graphs, the dynamic analysis of laundry tools, the interpretable learning of huge data and the context-incorporated robust system would be my suggestion.

## References


[1] Satoshi Nakamoto. A peer-to-peer electronic cash system. *Bitcoin.–URL: https://bitcoin. org/bitcoin. pdf*, 4, 2008.

[2] Malcolm Campbell-Verduyn. Bitcoin, crypto-coins, and global anti-money laundering governance. *Crime, Law and Social Change*, 69(2):283–305, 2018.

[3] John Bohannon. The bitcoin busts. *Science*, 351(6278):1144–1146, 2016.

[4] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80, 2008.

[5] Christian Brenig, Günter Müller, et al. Economic analysis of cryptocurrency backed money laundering. 2015.

[6] Zhiyuan Chen, Le Dinh Van Khoa, Ee Na Teoh, Amril Nazir, Ettikan Kandasamy Karuppiah, and Kim Sim Lam. Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review. *Knowledge and Information Systems*, 57(2):245–285, 2018.

[7] Mark Weber, Jie Chen, Toyotaro Suzumura, Aldo Pareja, Tengfei Ma, Hiroki Kanezashi, Tim Kaler, Charles E. Leiserson, and Tao B. Schardl. Scalable graph learning for anti-money laundering: A first look. *CoRR*, abs/1812.00076, 2018.

[8] Michel Rauchs, Apolline Blandin, Kristina Klein, Gina C Pieters, Martino Recanatini, and Bryan Zheng Zhang. 2nd global cryptoasset benchmarking study. *Available at SSRN 3306125*, 2018.

[9] Malte Möser, Rainer Böhme, and Dominic Breuker. An inquiry into money laundering tools in the bitcoin ecosystem. In *2013 APWG eCrime researchers summit*, pages 1–14. Ieee, 2013.

[10] Robert Stokes. Virtual money laundering: the case of bitcoin and the linden dollar. *Information & Communications Technology Law*, 21(3):221–236, 2012.

[11] Danton Bryans. Bitcoin and money laundering: mining for an effective solution. *Ind. LJ*, 89:441, 2014.

[12] Mark Weber, Giacomo Domeniconi, Jie Chen, Daniel Karl I Weidele, Claudio Bellei, Tom Robinson, and Charles E Leiserson. Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. *arXiv preprint arXiv:1908.02591*, 2019.

[13] Sarah Meiklejohn, Marjori Pomarole, Grant Jordan, Kirill Levchenko, Damon McCoy, Geoffrey M Voelker, and Stefan Savage. A fistful of bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 127–140, 2013.

[14] Thibault de Balthasar and Julio Hernandez-Castro. An analysis of bitcoin laundry services. In Helger Lipmaa, Aikaterini Mitrokotsa, and Raimundas Matulevičius, editors, *Secure IT Systems*, pages 297–312, Cham, 2017. Springer International Publishing.

[15] Rolf Van Wegberg, Jan-Jaap Oerlemans, and Oskar van Deventer. Bitcoin money laundering: mixed results? *Journal of Financial Crime*, 2018.

[16] Yaya Fanusie and Tom Robinson. Bitcoin laundering: an analysis of illicit flows into digital currency services. *Center on Sanctions and Illicit Finance memorandum, January*, 2018.

[17] Michael Fleder, Michael S Kester, and Sudeep Pillai. Bitcoin transaction graph analysis. *arXiv preprint arXiv:1502.01657*, 2015.

[18] Mikkel Alexander Harlev, Haohua Sun Yin, Klaus Christian Langenheldt, Raghava Mukkamala, and Ravi Vatrapu. Breaking bad: De-anonymising entity types on the bitcoin blockchain using supervised machine learning. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

[19] Ismail Alarab, Simant Prakoonwit, and Mohamed Ikbal Nacer. Comparative analysis using supervised learning methods for anti-money laundering in bitcoin. In *Proceedings of the 2020 5th International Conference on Machine Learning Technologies*, ICMLT 2020, page 11–17, New York, NY, USA, 2020. Association for Computing Machinery.

[20] Jason Hirshman, Yifei Huang, and Stephen Macke. Unsupervised approaches to detecting anomalous behavior in the bitcoin transaction network.

[21] Thai Pham and Steven Lee. Anomaly detection in bitcoin network using unsupervised learning methods. *arXiv preprint arXiv:1611.03941*, 2016.

[22] Patrick Monamo, Vukosi Marivate, and Bheki Twala. Unsupervised learning for robust bitcoin fraud detection. In *2016 Information Security for South Africa (ISSA)*, pages 129–134. IEEE, 2016.

[23] R. D. Camino, R. State, L. Montero, and P. Valtchev. Finding suspicious activities in financial transactions and distributed ledgers. In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 787–796, 2017.

[24] Yining Hu, Suranga Seneviratne, Kanchana Thilakarathna, Kensuke Fukuda, and Aruna Seneviratne. Characterizing and detecting money laundering activities on the bitcoin network. *arXiv preprint arXiv:1912.12060*, 2019.

[25] Pranav Nerurkar, Yann Busnel, Romaric Ludinard, Kunjal Shah, Sunil Bhirud, and Dhiren Patel. *Detecting Illicit Entities in Bitcoin Using Supervised Learning of Ensemble Decision Trees*, page 25–30. Association for Computing Machinery, New York, NY, USA, 2020.

[26] Basic Manual on the Detection And Investigation of the Laundering of Crime Proceeds Using Virtual Currencies. Technical report, 2014.

[27] Haaroon Yousaf, George Kappos, and Sarah Meiklejohn. Tracing transactions across cryptocurrency ledgers. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 837–850, Santa Clara, CA, August 2019. USENIX Association.