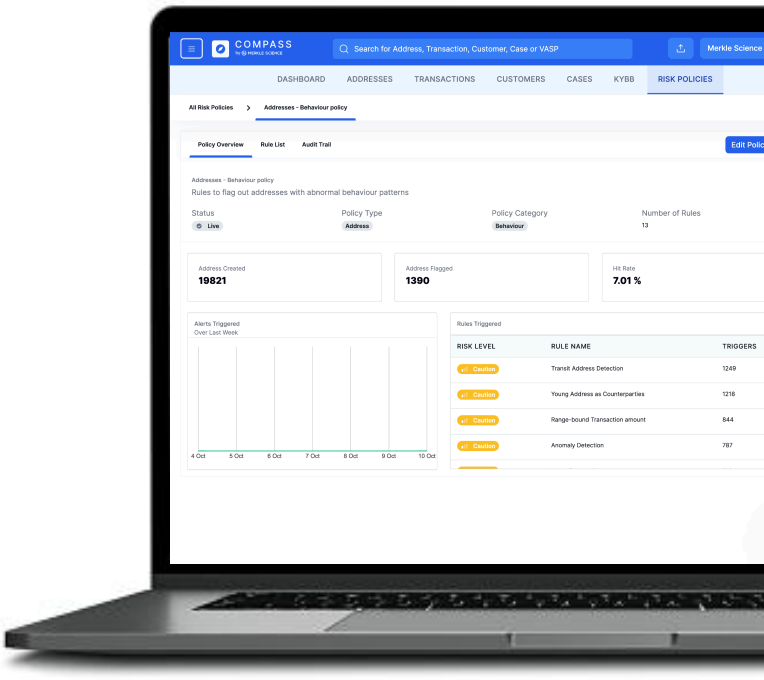
MERKLE SCIENCE

Behavior-Based Monitoring Results

A Proof-of-Value Evaluation for Behavior Monitoring of Cryptocurrency Addresses

Prepared Nov 13th, 2022



Executive Summary

At Merkle Science, we see flagging behavior on the blockchain along with labeling malicious entities as the way forward in building an effective compliance program and a way to meet the ever-evolving regulatory standards.

Providing an increasing number of real-time actionable insights to financial institutions enables them to take the best course of action possible. Mitigating risk and providing on-chain intelligence is the main objective of this experiment.

We sought to validate that on-chain patterns for illicit activity can be detected and provide predictive probabilities. Behavioral analytics can in fact be used to identify high-risk activity in many cases, months in advance.

In this report we identify the optimal Catch Rate to False Positive Ratio, the types of high-risk or illicit activities models detect best, and demonstrate the predictive capabilities behavioral analysis can bring to cryptocurrency investigations.

Objective

An exercise to demonstrate and validate the effectiveness of behavior-based monitoring alerts to flag illicit transactions and activity by analyzing a 6-month snapshot of the Ethereum blockchain as a test case.

Scope & Deliverables

Merkle Science will apply its behavior rules on available blockchain data to identify illicit wallets and transactions. For this exercise the data set will be initially limited to approximately 6 months, beginning in Apr 2022 and ending Oct 2022.

Questions that we aim to provide answers to

1. How accurate is on-chain behavioral analytics in identifying high-risk activity?
2. How much faster on-chain behavior analytics is at identifying high-risk activity when compared to attribution-based alert systems?

Results Overview

As part of this exercise we've run a total of 31 models corresponding to each week between Apr 3rd, 2022, and Oct 29th, 2022.

Metrics from our last model classifying illicit and benign behavior for the test week ie. between Oct 23rd, 2022, and Oct 29th, 2022 are considered and are shown below.

The Area Under Curve of ROC = 0.98

1. **How accurate is on-chain behavioral analytics in identifying high-risk activity?**
   The optimal catch rate determined by our models is 90%, at which point for each correctly tagged risky address, there will be 0.6 addresses falsely flagged as risky.

Catch Rate	False Positive Ratio*
100%	5.25
98%	2.5
90%	0.6
50%	0.04

Performance of the model is described through a combination of (Catch-Rate and False Positive Ratio).

Catch-Rate: The % of addresses in the validation dataset that is high-risk (based on ground truth), and have been identified by the model as high-risk.

False Positive Ratio: Number of Benign addresses falsely marked as Risky for every Risky address correctly identified as Risky.

2. **How much faster on-chain behavior analytics is at identifying high-risk activity when compared to traditional attribution-based alert systems?**

   On average, the model was able to identify high risk addresses 55 days earlier than incumbent attribution based alerting.

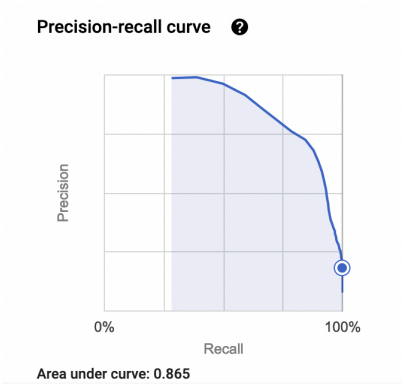
   A detailed address-level breakdown of detection time through behavior analytics detection and attribution-based detection is described in Appendix 1.

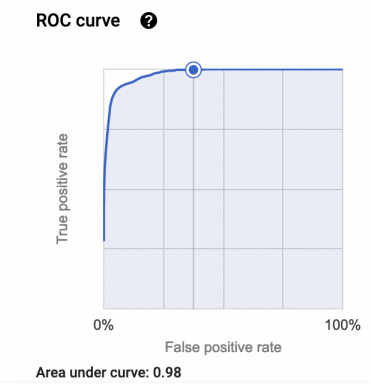
Results Detailed

To correctly detect all illicit behaving addresses (100% catch rate) from a sample, at worst we would flag 5 benign addresses as illicit for each correct illicit address we flag (16% precision).

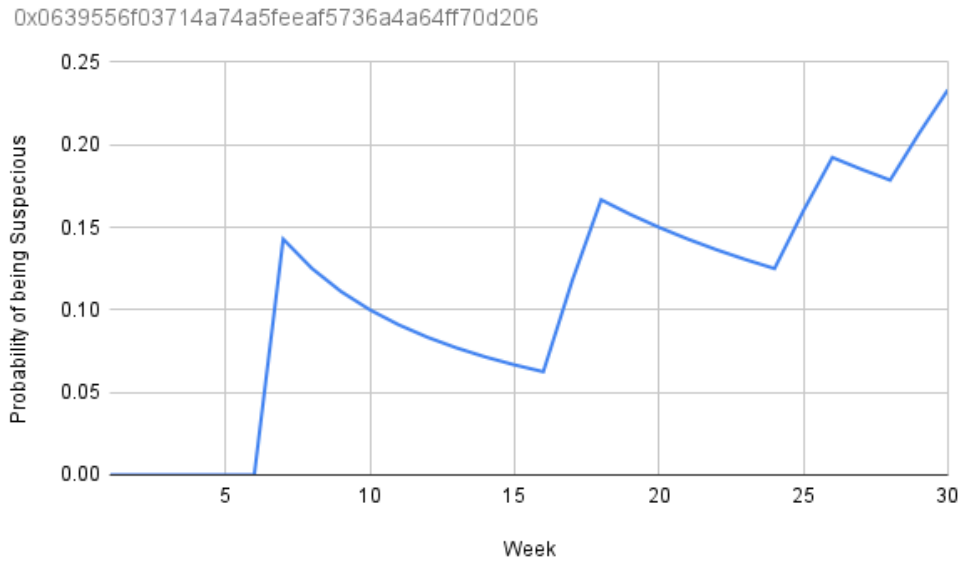
Similarly to correctly detect 98% of all illicit addresses, we flag 2 addresses illicit wrongly for every address flagged illicit correctly (29% precision).

Likewise, it is 63% precision for a 90% catch rate, and 96% precision for a 50% catch rate.





The below graph depicts the change in the probability scores over considered weeks for the address - 0x0639556f03714a74a5feeaf5736a4a64ff70d206. This particular address was first detected as suspicious in the week of 5/15/2022 while its entire 30 week detection history looks like {'0', '0', '0', '0', '0', '1', '0', '0', '0', '0', '0', '0', '0', '0', '0', '1', '1', '0', '0', '0', '0', '0', '0', '1', '1', '0', '0', '1', '1'}



Behavior Detected - Hacked

The following are addresses where we detected behavior similar to a hacked address:

Address	Asset	Reported Date	Probability	Earliest Detected Date	Date Delta
0x65760288f19cff476b80a36a61f9dedab16bab49	ETH	2022-08-06	1	2022-05-01	-97
0x6b88d0f4e94013b38e7c49ddc24135bfb0e2d49b	ETH	2022-10-29	1	2022-10-16	-13
0x0de41bdc58ffaf4a8c1b7084a544fbbe10a9de56	ETH	2022-07-09	1	2022-07-03	-6

Behavior Detected - Phishing

The following is an address where we detected behavior similar to phishing:

Address	Asset	Reported Date	Probability	Earliest Detected Date	Date Delta
0x4ad64fd7ca6d6150614179b9bce4094bc18f29cb	ETH	2022-10-08	1	2022-10-02	-6

Next Steps

As part of next steps we'd like to propose a validation study. We're requesting a sample set of data that your organization has already tagged and dispositioned.

- A minimum of 100 addresses is required, there is no upper limit on the number of addresses you can provide.

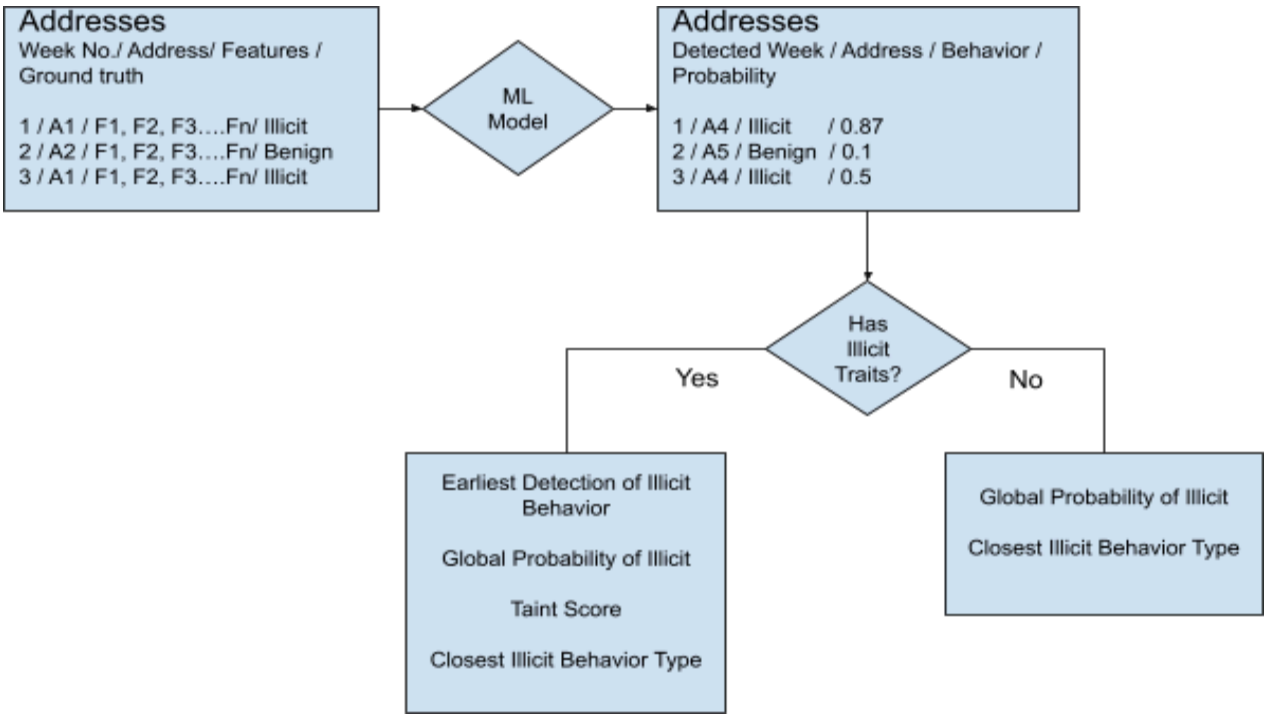
- Addresses should be a combination of pre-tagged good/bad addresses, scrubbed before providing to Merkle Science within the given date range of Apr 2022 to Oct 2022

Merkle Science will provide an output similar to the tables above, minus the "Date Reported by Chainalysis or Cipherblade". Using the "Earliest Date Detected" field to see the effectiveness of the models.

Appendix 1. Methodology

We compiled a dataset of addresses tagged by Merkle Science into a weekly broken set of features against labels of illicit or benign. A supervised learning approach is considered, where multiple models broken by week predict whether an address was illicit or benign in that week. For a given week, we train its model with all illicit features from all previous weeks and benign features from the current week.

Our end goal is to calculate the taint and the earliest detection date of illicit behavior for any given address. The taint here is the ratio of the number of weeks the address was detected as illicit and the number of weeks it appeared in.



In the above figure, A1, and A2 are the addresses that we consider in training our ML model that uses Features F1 to Fn. Using the trained model we then predict the probability for address A4 and A5 and present different metrics based on whether it is suspicious or not.

Appendix 2. Additional Evaluation

We followed ~100 addresses from Chainalysis, Cipherblade, and Substack (via our subscription to these services) when they were flagged and compared it with how soon our model flags and its probability of being flagged.

Here, the probability for an address is taken as the no of weeks it was flagged as illicit by the ratio of the weeks it appeared in our experiment. We use a Scale of 0 to 1, where 0 is no illicit activity detected, and 1 is the illicit activity detected in all temporal snapshots.

Address	Asset	Date Reported by Service	Probability of being suspicious	Earliest Detection Date *(by Merkle Science)*	Detection Date Delta *(in days)*
0xb3860b2f2ed2df3fc0a925f2c3e9a89749442292	ETH	2022-11-15	1	2022-04-03	-226
0x557f4f61314252f3359762dc434d15c93c4dd1cd	ETH	2022-11-09	1	2022-04-03	-220
0x7e7d8a1cc7a9aefa1cecefaf093a24fd938aa876	ETH	2022-11-09	1	2022-04-03	-220
0xb3860b2f2ed2df3fc0a925f2c3e9a89749442292	ETH	2022-11-09	1	2022-04-03	-220
0xc8d328b21f476a4b6e0681f6e4e41693a220347d	ETH	2022-11-09	1	2022-04-03	-220
0x7e7d8a1cc7a9aefa1cecefaf093a24fd938aa876	ETH	2022-11-07	1	2022-04-03	-218
0xc8d328b21f476a4b6e0681f6e4e41693a220347d	ETH	2022-11-07	1	2022-04-03	-218
0xadbab4f38ff9dcd71886f43b148bcad4a3081fb9	ETH	2022-10-15	1	2022-04-03	-195
0xa6439ca0fcba1d0f80df0be6a17220fed9c9038a	ETH	2022-09-17	1	2022-04-03	-167
0xae85f6fdfc13a86e1a0e9b05a5186ce70b14e483	ETH	2022-11-09	1	2022-06-05	-157
0x68e8198d5b3b3639372358542b92eb997c5c314a	ETH	2022-10-15	1	2022-05-15	-153
0x6f1022c1be12a0d4523f9fd610611b45ccf9efda	ETH	2022-11-09	1	2022-06-19	-143
0x6f1022c1be12a0d4523f9fd610611b45ccf9efda	ETH	2022-11-07	1	2022-06-19	-141
0xa8c83b1b30291a3a1a118058b5445cc83041cd9d	ETH	2022-08-06	1	2022-04-10	-118

0x65760288f19cff476b80a36a61f9dedab16bab49	ETH	2022-08-06	1	2022-05-01	-97
0x55ac7ff467902067bfe8082371f090c6d9b22ef6	ETH	2022-11-09	1	2022-08-07	-94
0xa546073567d0396ecb2e7d3b856e91a9085ab50d	ETH	2022-11-09	1	2022-08-07	-94
0x55ac7ff467902067bfe8082371f090c6d9b22ef6	ETH	2022-11-07	1	2022-08-07	-92
0x9fc8265f2b376084423a1a348a89ecd894a9d106	ETH	2022-10-29	1	2022-08-07	-83
0x78f05acd03b4dc51db68527afde64eb2f07938e4	ETH	2022-09-10	1	2022-06-26	-76
0xca92077acd49b523045754c1fe3ccc1d7710b119	ETH	2022-10-22	1	2022-08-07	-76
0x1d371cf00038421d6e57cfc31eeff7a09d4b8760	ETH	2022-10-15	1	2022-08-14	-62
0x84527b5949d479c879b8dd71cd8f79048cdf6fb8	ETH	2022-10-29	1	2022-08-28	-62
0x6df149efb195bab15533a6582028edbf3c2c7381	ETH	2022-11-09	1	2022-09-11	-59
0xc2c6fc1f75b745e803229894a942b600b13e6070	ETH	2022-11-09	1	2022-09-11	-59
0xe1c12148c6aeb50469d91cbdf4b07e9baf9062b0	ETH	2022-11-09	1	2022-09-11	-59
0x6df149efb195bab15533a6582028edbf3c2c7381	ETH	2022-11-07	1	2022-09-11	-57
0x180ea08644b123d8a3f0eccf2a3b45a582075538	ETH	2022-06-25	1	2022-05-01	-55
0x2249ee0bee00bfa86eb2deee983fd241f47a1242	ETH	2022-11-15	1	2022-09-25	-51
0x56d8b635a7c88fd1104d23d632af40c1c3aac4e3	ETH	2022-08-06	1	2022-06-19	-48
0xb5c55f76f90cc528b2609109ca14d8d84593590e	ETH	2022-08-06	1	2022-06-19	-48
0x2249ee0bee00bfa86eb2deee983fd241f47a1242	ETH	2022-11-09	1	2022-09-25	-45
0x49d3ff895bdad96f9c13051311dfa1f6ab068a1a	ETH	2022-11-09	1	2022-09-25	-45
0x49d3ff895bdad96f9c13051311dfa1f6ab068a1a	ETH	2022-11-07	1	2022-09-25	-43

0x0000003502aa61a5f1b1fdadff2cf947dfda526e	ETH	2022-06-11	1	2022-05-01	-41
0x44183fd1a79704f79e0986c6380dd9bfbbc7e6d2	ETH	2022-09-03	1	2022-07-24	-41
0xb88189cd5168c4676bd93e9768497155956f8445	ETH	2022-08-06	1	2022-07-03	-34
0xf57113d8f6ff35747737f026fe0b37d4d7f42777	ETH	2022-08-06	1	2022-07-03	-34
0x4206d62305d2815494dcdb759c4e32fca1d181a0	ETH	2022-10-22	1	2022-09-25	-27
0x92a26975433a61cf1134802586aa669bab8b69f3	ETH	2022-06-11	1	2022-05-15	-27
0x000000000000660def84e69995117c0176ba446e	ETH	2022-08-06	1	2022-07-17	-20
0x3245e504d75e4148a491baac5066b0f26965567d	ETH	2022-07-23	1	2022-07-03	-20
0x4e8d918118e00f049712bb8da2b42088909eeee7	ETH	2022-04-23	1	2022-04-03	-20
0x50f9202e0f1c1577822bd67193960b213cd2f331	ETH	2022-08-13	1	2022-07-24	-20
0x847e74d8cd0d4bc2716a6382736ae2870db94148	ETH	2022-08-06	1	2022-07-17	-20
0x87c828593984381e50d55f755b8462e074047cf7	ETH	2022-10-22	1	2022-10-02	-20
0xdc25df861f979a175bfe4f3737d1562d45cdc5cd	ETH	2022-09-03	1	2022-08-14	-20
0xdf31f4c8dc9548eb4c416af26dc396a25fde4d5f	ETH	2022-10-15	1	2022-09-25	-20
0x1d2677ed1b0815fab22368347723551a9dd1fb1b	ETH	2022-07-23	1	2022-07-10	-13
0x6b88d0f4e94013b38e7c49ddc24135bfb0e2d49b	ETH	2022-10-29	1	2022-10-16	-13
0x8c7934611b6ad70fbea13a1593de167a4689b9a9	ETH	2022-04-30	1	2022-04-17	-13
0xb0f5fa0cd2726844526e3f70e76f54c6d91530dd	ETH	2022-08-20	1	2022-08-07	-13
0x0248f752802b2cfb4373cc0c3bc3964429385c26	ETH	2022-09-24	1	2022-09-18	-6
0x0c9797805a22e507bf48f35c72a67f001b7418d0	ETH	2022-07-16	1	2022-07-10	-6

0x0d043128146654c7683fbf30ac98d7b2285ded00	ETH	2022-06-25	1	2022-06-19	-6
0x0de41bdc58ffaf4a8c1b7084a544fbbe10a9de56	ETH	2022-07-09	1	2022-07-03	-6
0x1079061d37f7f3fd3295e4aad02ece4a3f20de2d	ETH	2022-06-04	1	2022-05-29	-6
0x158f5cb7ab88e1c2418d5509d3fe43ae548ce345	ETH	2022-09-24	1	2022-09-18	-6
0x161cebb807ac181d5303a4ccec2fc580cc5899fd	ETH	2022-10-29	1	2022-10-23	-6
0x189e078ef2c61c2b11f6b0f6c6d6fe645d1ad995	ETH	2022-05-21	1	2022-05-15	-6
0x1c8465662caa8005ed41430e433e399c699cbce2	ETH	2022-05-21	1	2022-05-15	-6
0x282971ded7d0b8c5b0358ebebe3b2bc6a24a6b10	ETH	2022-07-09	1	2022-07-03	-6
0x2c177d20b1b1d68cc85d3215904a7bb6629ca954	ETH	2022-08-06	1	2022-07-31	-6
0x3497b57fe49e90a783cc7b1d62dbabf560785744	ETH	2022-08-20	1	2022-08-14	-6
0x4429abbf523bef0f1e934b04cff8584955c72548	ETH	2022-07-09	1	2022-07-03	-6
0x443cf223e209e5a2c08114a2501d8f0f9ec7d9be	ETH	2022-10-22	1	2022-10-16	-6
0x489a8756c18c0b8b24ec2a2b9ff3d4d447f79bec	ETH	2022-10-08	1	2022-10-02	-6
0x4ad64fd7ca6d6150614179b9bce4094bc18f29cb	ETH	2022-10-08	1	2022-10-02	-6
0x4c6731d49a8667fa5e853ef2f586e9c7f73c3d72	ETH	2022-04-23	1	2022-04-17	-6
0x510e4d61663be6a24d600aaf90f892dd8c8c61dc	ETH	2022-09-10	1	2022-09-04	-6
0x53b70c26a5ffa5da24c737b6109dd1f4e4d9d895	ETH	2022-06-11	1	2022-06-05	-6
0x58f4baccb411acef70a5f6dd174af7854fc48fa9	ETH	2022-06-25	1	2022-06-19	-6
0x5c95123b1c8d9d8639197c81a829793b469a9f32	ETH	2022-10-15	1	2022-10-09	-6
0x5f7848ec0286304dc5fe6497af4b3c0fead6a920	ETH	2022-10-22	1	2022-10-16	-6

0x64e5ac2e59ccd85c02dede27d290f16d0ed5bf24	ETH	2022-07-23	1	2022-07-17	-6
0x6ae09ac63487fcf63117a6d6fafa894473d47b93	ETH	2022-09-17	1	2022-09-11	-6
0x6d3e6ba1b510287141b27f763a86e04c72a001d1	ETH	2022-10-22	1	2022-10-16	-6
0x701428525cbac59dae7af833f19d9c3aaa2a37cb	ETH	2022-04-30	1	2022-04-24	-6
0x84d3656163005ecdec0339b502068fc8e520feb1	ETH	2022-10-01	1	2022-09-25	-6
0x8a6b84a3f95ecf40f5991c99635f1f5467c4d2d9	ETH	2022-06-11	1	2022-06-05	-6
0x8ca72f46056d85db271dd305f6944f32a9870ff0	ETH	2022-10-29	1	2022-10-23	-6

Appendix 3. Feature List

Below is the list of features used for identifying illegal/illicit accounts based on previous attacks on blockchains and performing time series analysis. As new features, we propose temporal burst, degree burst, gasPrice burst, and attractiveness.

	Feature Name	Variable	Description
1.	transactedFirst	Block number	Identify the first Block Number the transaction address appeared.
2.	transactedLast	Block number	Identify the last Block Number the transaction address appeared.
3.	activeDuration	transactedLast-transactedFirst	Delta between the first and last blocks, typically illicit addresses are short lived.
4.	averagePerInBal	TotalIncoming/number of TXNs	Illicit addresses statistically have higher inbound transaction value and maintain higher balances or average balance per wallet.
5.	uniqueIn	Uniques in Address interacted with	Unique number of addresses the address under consideration interacted with. Interaction with a larger number of unique input addresses indicate illicit/suspicious activity.
6.	uniqueOut	Uniques out Address interacted with	Unique number of addresses the address under consideration interacted with. Interaction with a larger number of unique Out addresses indicate illicit/suspicious activity.
7.	numberOfBalanceburstTemporalOut	Number of blocks where address has more than X balance in the out txns	Identify the number of blocks where the total count of outbound transactions exceeds a threshold. A normal address behavior would be regular while a hacked address or an illicit address would show abnormal increase in transactions.
8.	numberOfBalanceburstTemporalIn	Number of blocks where address has more than X balance in the in transactions	Identify the number of blocks where the total count of inbound transactions exceeds the threshold. A normal address behavior would be regular while a hacked address or an illicit address would show abnormal increase in transactions.
9.	gasPrice__quantile__q_0.2	Quantile of gasprice time series	Each transaction will have a different gas price. Illicit actors lure a miner or validator to accept a transaction by paying a significantly higher gas price. This promotes selfish mining by miners where legitimate transactions are ignored or remain in pending state.
10.	In_attractiveness__mean	Average of distinct or new neighbors	Each address will have a different set of neighbors. This quantity identifies the number of unique and distinct neighbors it receives funds from. Usually ransomware addresses and hacker addresses have high attractiveness.
11.	InterEventDuration__pct_70	70th percentile of Duration between two consecutive transactions	The Frequency with which an address is involved in transactions can act as a key separator between benign and illicit behavior. Smaller Inter Event Durations are seen for Phishing/Scam behavior, while very large Inter Event Durations are seen for Money Laundering behavior.
12.	is_exchange	Pre Calculated from MS attribution	Reduces False Positives