The anatomy of a cryptocurrency pump-and-dump scheme
J. Xu and B. Livshits

Sajjad Heydari

Nov 2020

Pump and Dump

- Fraudulent activity in stock market
- Rely on supply and demand to momentarily increase price

Pump and Dump

- Fraudulent activity in stock market
- Rely on supply and demand to momentarily increase price
- Study pump and dump on blockchain
- Create model to predict which coins are going to be pumped
- Create a proof of concept trader around pump and dump scheme

Process

- Acquire penny-stock
- Announce the stock and **pump** it by encouraging others to buy it for quick profit
- **Dump** while the price is higher
- Tell others to start dumping their stock

Process

- Acquire penny-stock
- Announce the stock and **pump** it by encouraging others to buy it for quick profit
- **Dump** while the price is higher
- Tell others to start dumping their stock
- **Trade is zero sum game**, so organizers and early participants profit while others loose capital

Cryptocurrency Pump and Dump

- **Set-Up**: Create public group on Telegram or Discord, recruit as many members as possible
- **Pre-pump**: Announce details of pump a few days ahead, including time, exchange, and pairing coin
- **Pump**: Announce the coin in OCR-proof image, urge members to buy
- **Dump**: While the price is nearing its peak, start selling
- **Post-pump review**: Publish a one-sided review for future operations

Actors

- Organizers
- Participants
- Target exchange

Case Study

- November 14th 2018, on Cryptopia's BTC market
- At least four Telegram channels, the largest one being *"Official McAfee Pump Signals"* with 12,333 members prior to coin announcement
- At 19:30 GMT on November 14, 2018 announced BVB coin
- Dormant coin, announced on August 2016 with last git commit on August 10 2017
- First channel announced at 19:30:04
- Last channel announced at 19:30:23

Tick-by-tick movement

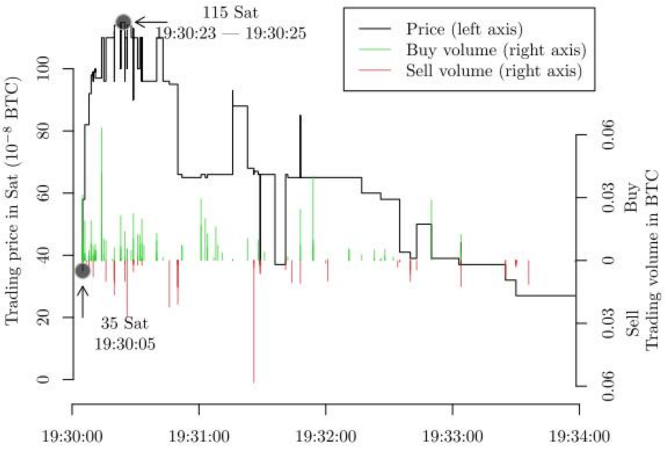


Figure 1: Tick by tick movement of BTC/BVB

Buy and Sell Volume

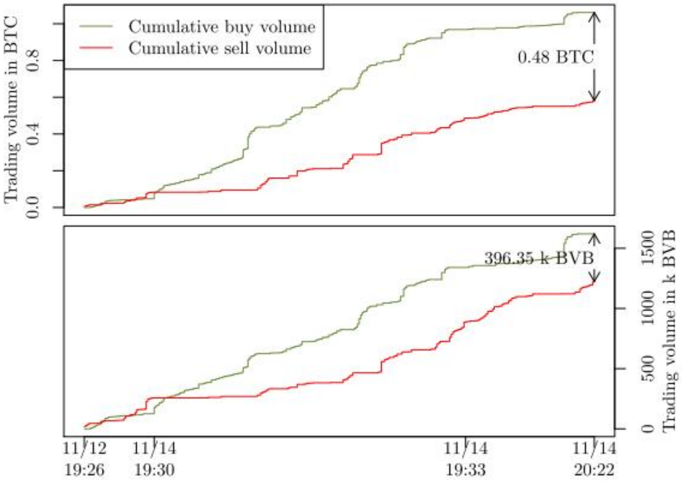


Figure 2: Buy and sell volume caused by pump-and-dump

Notes On Analysis

- Only Telegram was considered, since every pump and dump on Discord was also on Telegram
- Telegram channels were gathered[1] by searching related keywords on Telegram aggregators as well as cross-promotions
- Only consider pumps paired with BTC
- Traced 429 pump-and-dump actions after clearing and merging the data

[1] by PumpOlymp

Coin Data

- CryptoCompare API for hourly OHLC[2] and volume on 189 exchanges
- Non-financial feature including listing status, algorithm and total supply from the same source

[2] Open, high, low, close

Observation #1 Channel Activity

- Over half (168/315) Telegram channels have not been active for a month,

Observation #1 Channel Activity

- Over half (168/315) Telegram channels have not been active for a month,
- Cautious admins delete their history
- Hit-and-run characteristics due to ineffectivity in the long run

Observation #2 Exchanges

- Larger exchange such as Binance and Bittrex have large user base and abnormal price attract others
- Smaller exchange such as Cryptopia and Yobit host more easily manipulated coin prices

Exchanges contd.

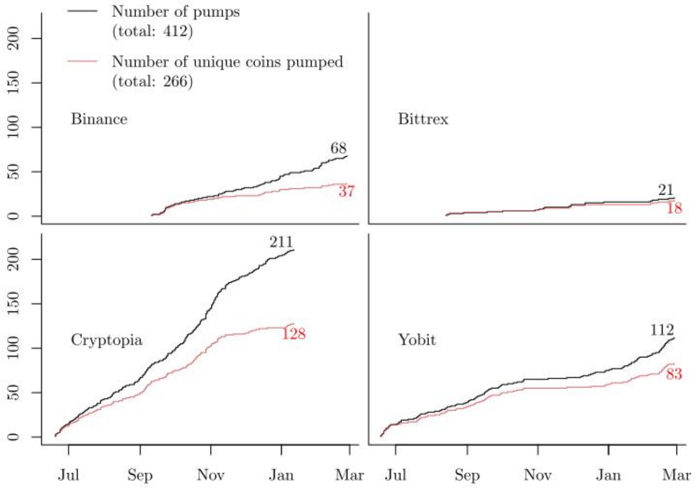


Figure 3: Cumulative count between June 2018 to February 2019

Observation #3 Admin's Profits

- **Assuming** admins purchase coins and enter sell orders only prior to pump,
- **Assuming** admins purchase immediately before pump
- **Assuming** during pump investors lift admins' offers and during drop investor transact with each other
- We **observer** that admins made a net profit of 199.52 BTC through 348 events, averaging over 18% return

Observation #4 Announcement Views

- Negative correlation (-0.162) between number of views and pump gain,

Observation #4 Announcement Views

- Negative correlation (-0.162) between number of views and pump gain,
- Not everyone who views participates
- Double counting views and views after pump
- Bot users views aren't counted

Prediction Overview

- **Objective**: Given the discussed features for multiple coins, predict the coin that will be chosen for pump-and-dump
- **Input**: One cryptocurrency at a time
- **Output**: Boolean (Whether the coin has a chance at selection) and the confidence
- **Justification**: Previous analysis shows that there is unusual market movement prior to the event and the model should be able to identify it

Feature Selection

- Only **Cryptopia**'s exchange info was used in modeling
- Great emphasis on features associated with market movement[3], 85% of all features
- Covering 296 candidate coins, with only one actually pumped at each pump

[3] such as price, returns, volatilities in various lengths of time

Dataset Preparation

Pumped	Training	Validation	Test	Total	
TRUE	60	60	60	180	0.3%
FALSE	17,078	17,995	18,135	53,028	99.7%
TOTAL	17,138	18,055	18,195	53,208	100%

Table 1: Datasets

Model Selection

- **Random Forest**
  - Handles large quantities of variables
  - Resilient to correlation, interaction or non-linearity of the features
  - Relies on voting between different trees
  - Time-consuming
  - Highlights important features
- **Generalized Linear Model**
  - Highly interpretable model
  - Uncover correlation between features and dependent variables
  - Efficient
  - Prone to overfitting

Hyperparameters

- **Random Forest**
  - Split the FALSE training sample size to 20k, 5k and 1k,
  - With tree sizes of 5k, 10k and 20k respectively
- **Generalized Linear Model**
  - Used LASSO[4] regularization,
  - With λ at 10^{-8}, 10^{-3} and 5×10^{-3}

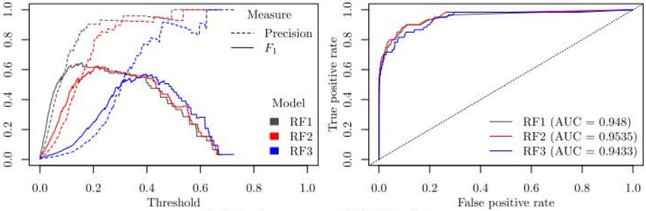
[4] Least Absolute Shrinkage and selection Operator

Measures

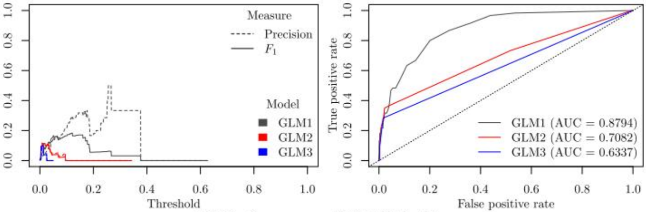
- Precision
- F1 measure
- Area under ROC[5] curve

[5] Receiver Operating Characteristic

Model Performance



(a) Performance of RF Models.



(b) Performance of GLM Models.

Figure 4: Model Performance

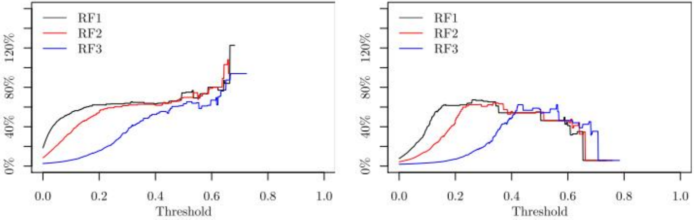
Investment Strategy

- Before each pump predict the coin and buy it
- Acquire k times the RF votes in BTC value
- Assume the gain is

$$\frac{\text{high price} - \text{open price}}{2 \times \text{open price}}$$

- Assume false positives generate a return of zero

Trader Performance



(a) Training sample. (b) Validation sample.

Figure 5: Trader Performance

Trader Final Test

- Used RF1 and threshold of 0.3,
- Consider the market depth in the uptick transaction as baseline investment quantity
- Discount it by prediction likelihood to get investment amount

Trader Final Test

- Used RF1 and threshold of 0.3,
- Consider the market depth in the uptick transaction as baseline investment quantity
- Discount it by prediction likelihood to get investment amount
- Trader suggests 9 coins
- With return of 60% over the test sample period of two and half months

Future Works

- Arbitrage's effect, i.e., price increase in one exchange results in price increase in others
- Data, including order book, tick by tick before pump and traders' account information could be useful
- Testing other models including classification and regressions
- Trader could be improved by considering coin price increase potential along with pump likelihood

Review

- The paper was very disorganized with related information scattered across different sections
- Trader would not be good in practice since the activity on the coin would alert the organizers to change their coin. They mention this in the last section but don't consider it in their calculations
- Model selection was not justifiable, despite the effort to show otherwise

Questions for the audience

- How does each members profit from a pump and dump event? (Admins, Exchanges, Participants)
- How can a cryptocurrency systematically prevent pump-and-dump usage on its coins?

QA

Thank you!