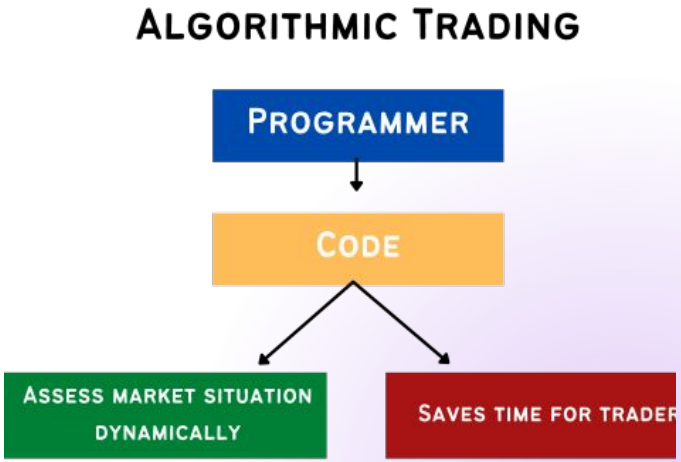
A Proximal Policy Optimization Approach to Detect Spoofing in Algorithmic Trading

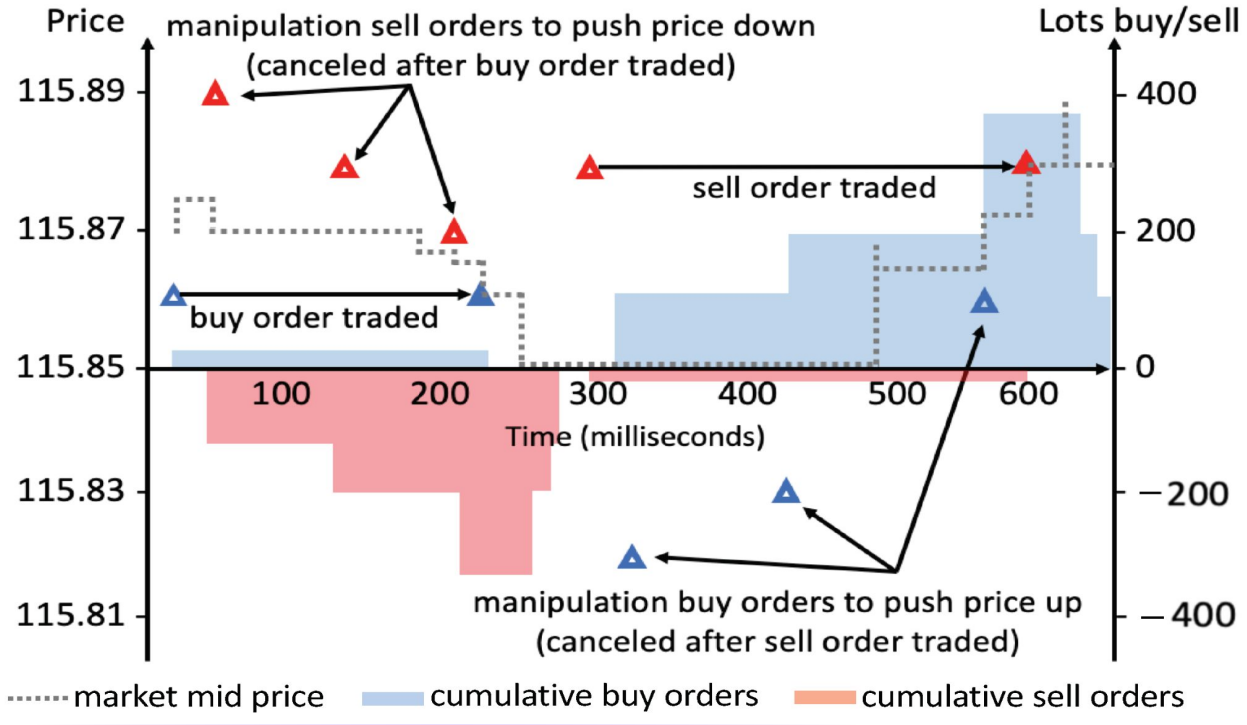
Author: **Iulia-Diana Groza**
Scientific Coordinator: **Diana-Lucia Miholca**
Computer Science, Babes-Bolyai University

Algorithmic Trading



The **liquidity** of markets is **improved** by ruling out the involvement of human emotions and execution delays specific to **traditional market-making**.

Spoofing

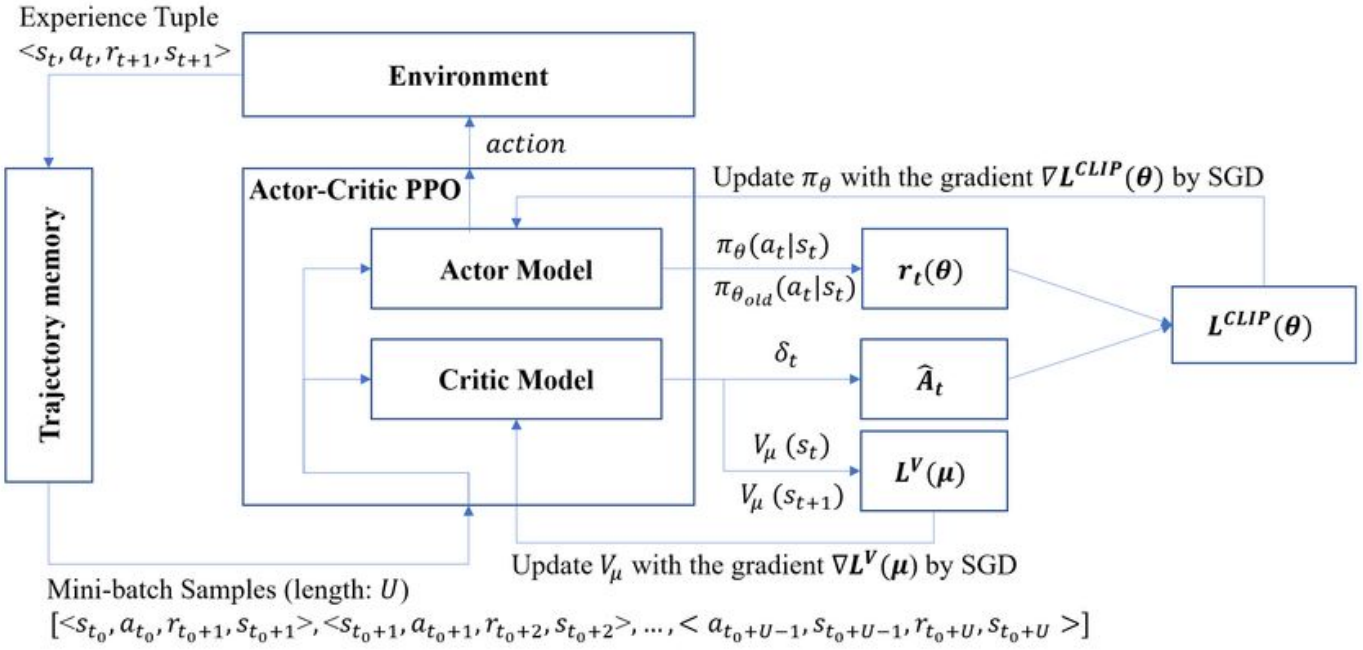

Price
manipulation sell orders to push price down
(canceled after buy order traded)
Lots buy/sell
115.89
115.87
115.85
115.83
115.81
400
200
0
−200
−400
sell order traded
buy order traded
100
200
300
400
500
600
Time (milliseconds)
manipulation buy orders to push price up
(canceled after sell order traded)
market mid price
cumulative buy orders
cumulative sell orders

Related Work

	Approach	Tools	Results
Tuccella et al. (2021)	Supervised	Gated Recurrent Units	Accuracy: **0.75**
Cao et al. (2015)	Unsupervised	Adaptive Hidden Markov Model with Anomaly States	Best AUC: **0.89**
Li et al. (2023)	Statistical Physics	Motion of Particles on Level 3 LOB Data	Largest Anomalous Deviation: **126.12**

OpenAI

Proximal Policy Optimization (PPO)



Research Contribution

01 Prove PPO is feasible in market surveillance

02 Harness Level 3 LOB data

03 Feature Engineering: Rolling statistics for price and size movements

Data Collection and Preprocessing

Li et al. (2023)

Historical Level 3 LOB Data on wLUNA/USD from the LUNA flash crash from May 2022 (11/05/2022, 16:00-20:00).

Timestamp	Price	Order Type	Side	Size
16:31:39.50	1.62	limit	buy	111939.762
16:32:37.06	1.62	cancel	buy	111939.762

(a) 11/05/2022 16:00-17:00

Timestamp	Price	Order Type	Side	Size
18:03:32.66	1.10	limit	buy	29000
18:12:52.87	1.36	cancel	buy	29000
18:23:14.53	1.25	limit	buy	29000
18:24:27.68	1.24	limit	buy	29000
18:27:53.69	1.25	cancel	buy	29000
18:28:11.94	1.24	cancel	buy	29000
18:42:14.77	1.06	limit	buy	20000
18:54:15.66	0.80	limit	buy	29000
18:54:54.42	0.80	cancel	buy	29000
18:59:42.93	1.01	cancel	buy	20000

(b) 11/05/2022 18:00-19:00

Timestamp	Price	Order Type	Side	Size
19:28:34.77	0.73	limit	buy	535665.177
19:29:07.48	0.73	cancel	buy	535665.177

(c) 11/05/2022 19:00-20:00

~1.8 Million

Full Channel Records

~ 100,000

Ticker Records

Feature Engineering

Rolling Statistics: (Price + Size)

	Mean (μ)	Standard Deviation (σ)	Variance (σ^2)
Window Size:	5	10	15

Marking our **contribution**. They provide a detailed view of the **central tendency** and **dispersion** of the data over different time periods, allowing the identification of **abnormal fluctuations** in price and order sizes. Sudden increases create a **false impression of market depth**.

Order Flow Imbalance (OFI)

Reflects **market pressure**, caused by the **discrepancy** between buy and sell orders.

$$OFI(t) = \sum_{i=t-w}^{t} size_i \cdot side_i$$

Market Spread

Provides insight into the **market liquidity** and the **aggressiveness** of the trading activities.

$$spread = best\_ask - best\_bid$$

Cancellation Ratio

Evident indicator of spoofing.

$$CR = \frac{reason\_canceled}{type\_received\_adjusted}$$

Market Simulation Environment

Agent Playground

Provides a **controlled setting** where our PPO model can interact with **simulated** historical LUNA flash crash **market data**.

Reward Structure

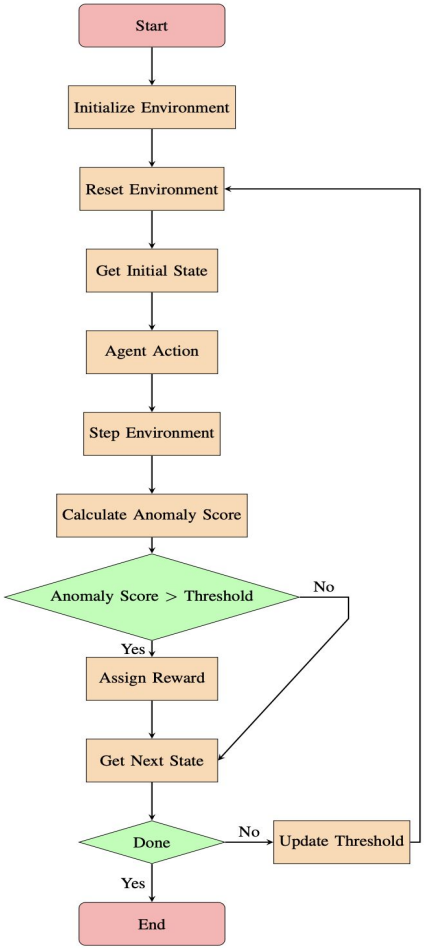
Reinforce correct detections and **penalize** incorrect actions. The time series are processed **sequentially**.

Anomaly Detection

Unsupervised system for labeling legitimate spoofing attempts. Anomaly **score** computed based on empirical **feature weights**.

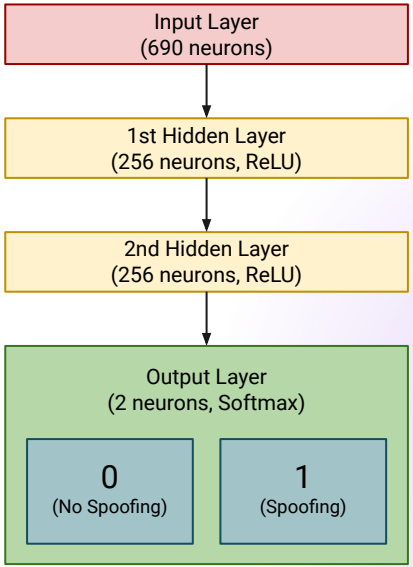
Adaptive Spoofing Threshold

Dynamically updated to the **75% most recent anomaly scores**. Ensures the model remains responsive to **market behavior shifts**.



Policy Network

Feedforward Neural Network

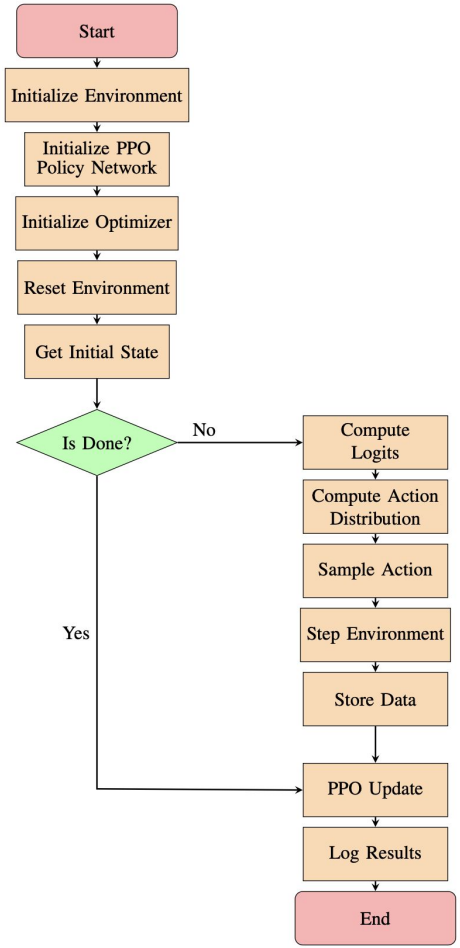


Network Weights

- Initialization: **Kaiming Normal**.
- Optimization: **Adam Optimizer with Learning Rate of 1 x 10^{-3}**.

PPO

- **Discounted Rewards:** emphasize immediate actions.
- **GAE:** smooth out advantage estimates ⇨ stability & reliability.
- **Clipped Surrogate Objective:** prevent drastic updates, ensure stable policy improvement.
- **Entropy Term:** encourage exploration.



Hyperparameter Tuning - Anomaly Detection

FEATURE WEIGHTS FOR ANOMALY SCORE CALCULATION

Feature	Weight
order_flow_imbalance	0.15
cancel_to_received_ratio	0.15
price_5_std	0.05
price_10_std	0.05
price_15_std	0.05
size_5_var	0.05
size_10_var	0.05
size_15_var	0.05
spread	0.10
last_size_5_var	0.05
last_size_10_var	0.05
hour_of_day	0.15
hour_15	0.05
hour_16	0.05
hour_17	0.05
hour_18	0.05
hour_19	0.05

FEATURE WEIGHTS FOR ANOMALY SCORE CALCULATION WITHOUT ROLLING STATISTICS

Feature	Weight
order_flow_imbalance	0.25
cancel_to_received_ratio	0.25
spread	0.10
hour_of_day	0.15
hour_15	0.25
hour_16	0.25
hour_17	0.25
hour_18	0.25
hour_19	0.25

Hyperparameter Tuning - PPO Parameters

PERFORMANCE METRICS FOR SELECTED PPO CONFIGURATIONS

Total Reward	Avg Reward	Std Reward	Learning Rate	Batch Size	Epochs	Spoofing Threshold
9500	0.317	0.120	1×10^{-3}	128	30	0.8
9200	0.307	0.115	5×10^{-4}	128	30	0.8
9000	0.300	0.110	1×10^{-3}	64	20	0.8
8900	0.297	0.105	5×10^{-4}	32	20	0.8
8500	0.283	0.102	1×10^{-3}	128	30	0.7
8300	0.277	0.099	5×10^{-4}	32	20	0.9
8000	0.267	0.094	1×10^{-3}	64	20	0.7
7800	0.260	0.091	5×10^{-4}	64	20	0.7
7500	0.250	0.088	1×10^{-4}	32	10	0.8
7300	0.243	0.085	1×10^{-4}	64	10	0.9

Performance Analysis

63% → **89%**

Cancellation Records

0.24 → **0.13**

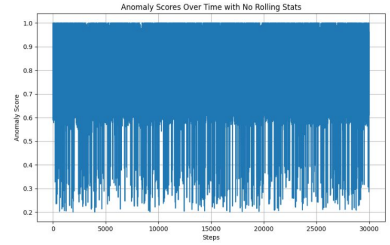
Training Loss

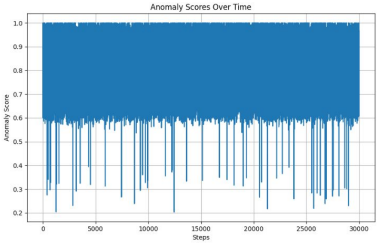
4,200 → **9,500**

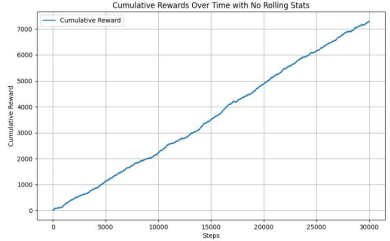
Total Reward

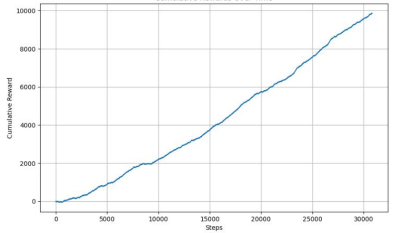
17k → **20k**

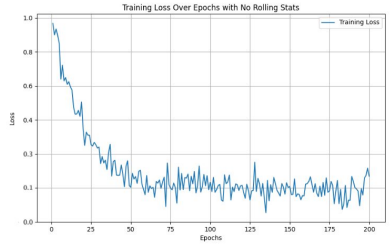
Maximum Frequency of Positive Reward

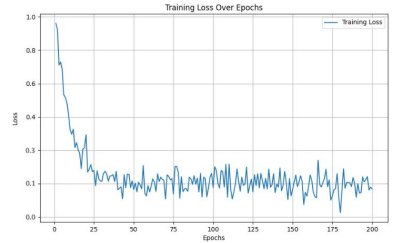


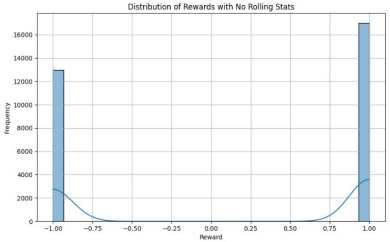


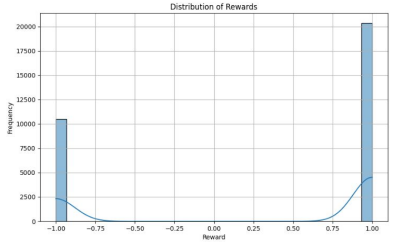












Future Considerations

01 Online Learning

02 Labeled Data for a Safer Detection

03 Advanced Hypertuning: Grid Search, Bayesian Optimization

Thank you!

Questions?

Bibliography

- https://github.com/iuliagroza/spoof.io
- Bachelor's Thesis
- https://www.5paisa.com/finschool/basics-of-algo-trading-and-examples/
- John D. Montgomery. Spoofing, market manipulation, and the limit-order book. May 2016.
- Jean-Noel Tuccella and Philip Nadler and Ovidiu Serban. Protecting Retail Investors from Order Book Spoofing using a GRU-based Detection Model. 2021.
- Cao, Yi and Li, Yuhua and Coleman, Sonya and Belatreche, Ammar and McGinnity, Thomas Martin. Adaptive Hidden Markov Model With Anomaly States for Price Manipulation Detection. 2015.
- Haochen Li and Maria Polukarova and Carmine Ventre. Detecting Financial Market Manipulation with Statistical Physics Tools. 2023.
- https://openai.com/research/openai-baselines-ppo
- John Schulman and Filip Wolski and Prafulla Dhariwal and Alec Radford and Oleg Klimov. Proximal Policy Optimization Algorithms. 2017.
- Hyun-Kyo Lim and Ju-Bong Kim and Youn-Hee Han. Federated Reinforcement Learning for Training Control Policies on Multiple IoT Devices. March 2020.

Demo