IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

HFT Inside the Blackbox that Beats the Market

[1]Kunal Jain, [2]Ashish Shinalkar, [3]Rohit Naidu

[1]Student, [2]Student, [3]Student

[1]Enggineering,

[1]MIT ADT University, Pune, India

Abstract: This paper focuses on how financial giants make use of automated trading and investment recommendation platforms to transact many orders at high speeds and provide recommendations and analytics to other AMCs. Their systems use algorithms to analyze markets and spot emerging trends in a fraction of a second. Due to this, the existing system has its possible Sharpe ratio ten times greater than the outdated buy-and-hold strategies. The effects of algorithmic and investment engines are the subject of ongoing research but, that does not mean we should not care about making it a little less taxed. Perhaps the biggest issue facing these trading engines is that they need a lot of computing resources and fast access to data to effectively make recommendations.

Index Terms **- high frequency, arbitrage, irrational exuberance, colocation, auto-scaling, load balancing.**

I. INTRODUCTION

With the advent of technology, automated trading machines and investment recommendation engines have started to manage the financial markets with remarkable consistency and utmost reliability. In the recent past, AI and machine learning techniques have led to interesting developments in the domain of economics and finance. On behalf of the financial institutions, the AI algorithms, developed by the quants, figure out the investment and trading strategies on their own. These engines use algorithms or scripts in specialized analytics software to trade and recommend investments automatically within the time frame of a few milliseconds. [1]

It is estimated that these accounts for about 40% of all the equity trading volume. Within a short span of time, these systems have managed to become an important part of the financial ecosystem. Large institutions employ the strategies coming out of their recommendation engines to transact trades and advice. The infrastructure behind these systems play one of the primary features that has led to the dominance of these systems in the financial realm. Most trading engines can be hooked up with expensive and powerful computing power and storage resources but what if we can integrate these resources directly inside the engine, automating infrastructure provisioning on the go which can also save costs and don't require much technical knowledge to work with. And what if you have changing demands in your recommendation infrastructure that can hamper your customer experience due to time consuming provisioning and over the roof costs? [2]

II. ARBITRAGE

The rapid access of information in the form of market prices and news-based data can present multiple arbitrage opportunities. The time span for arbitrage is extremely short and hence speed is of paramount importance. There are various arbitrage strategies such as triangular, covered interest, cross-broker but among them triangular arbitrage is a popular strategy. [3]

A simple example of triangular arbitrage:

NZD/USD = $0.9 [buy 1500 NZD units for US$1350]

AUD/NZD = $1.2 [buy 1500 AUD units with NZ$1800]

AUD/USD = $1.3 [sell 1500 AUD units and receive US$1950]

The trader started with US$1350 and ended with US$1950, making an arbitrage profit of US$600. This is an unrealistic example. In the real world, the difference between the currency pairs might be only a tiny fraction of 1 cent. But if HFT makes a large enough trade and does it regularly it can lead to large profits. The opportunity exists because of the discrepancy between currencies- the market is not in equilibrium for a very small amount of time. Arbitrage profits are theoretically available to all traders in a market. The reason why HFT have faced criticism is because it appears as if they are exploiting the market; they have a greater opportunity to take these near risk-free profits. Arbitrage is not completely risk-free. There is the risk that the quoted prices change between trades, bid/ask spreads (basically transaction fees to broker) and other costs/factors. [4]

III. ULTRA-LOW LATENCY DMA

The trades are executed at extraordinarily high speeds which is achieved through colocation and advanced proprietary hardware. In colocation, the HFT firms place their computer servers in the same place where the stock exchange's servers are located or as close to them as possible. [5]

HFT firms spend billions of dollars to be as close to the stock exchange as possible and they go to great lengths to achieve this goal. They dig tunnels for laying fiber optic cables in order to connect with the stock exchange's servers. With time, this need for low latency is becoming even more pressing and hence the elite financial firms are spending substantial money for the research and development of new and advanced technologies such as experiments with proprietary microwave, laser, and satellite technology which remain classified. This science is referred to as Ultra-low latency direct market access (ULLDMA). [6]

DMA (Direct Market Access) is a process which connects buyers and sellers together through the stock exchange. ULLDMA combines DMA and algorithmic trading for the execution of trades on the given trading platform and bypasses the clandestine malpractices employed by the stockbroker. Large volumes of orders are handled in less than a second. Typically, 5000 orders can be executed in less than 100 microseconds. [17]

3.1 NEWS BASED TRADING

Data from various news sources including commercial news such as Bloomberg, Financial Times and similar publications as well as public news websites. Social media trends are also taken into consideration and so the data feed from Twitter is analyzed and then turns social media streams into actionable trading signals. The trading platform identifies micro-trends in the stream of data and provides unique insights for investment predictions. [7]

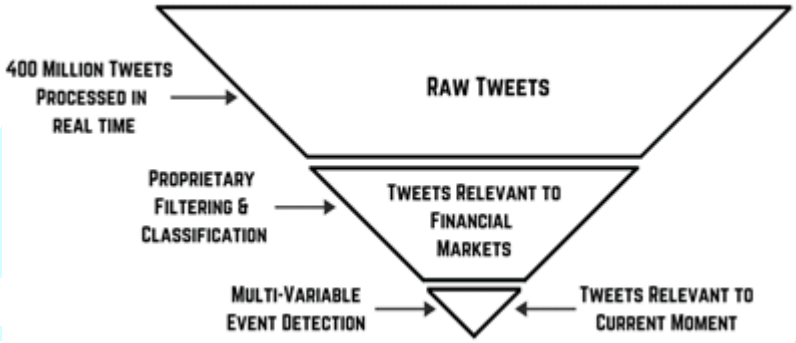

Figure 1: News Based Trading

In the above figure, raw Twitter data consisting of roughly 340 million tweets is analyzed and linguistic patterns are detected which are then aggregated using a smart linguistic real-time analytics tool. Live Twitter feeds are also used by trading platforms such as Bloomberg Terminals. [8]

Specialised algorithms are implemented for interpretation of the news. The news articles are scanned by these algorithms based on keywords and then the data is processed due to which the underlying meaning is identified, and its importance is assessed based on which the trades are executed. Quantification of news reports and articles is the most important step in news-based trading. This is achieved by assigning a score to each news article for the interpretation of the underlying sentiment. These sentiments can be positive, negative or neutral. The relevance and the source of the news article based on assigned keywords is taken into consideration. This is then quantified which enables the algorithm to make trading decisions.

3.2 Scaling and Load Balancing

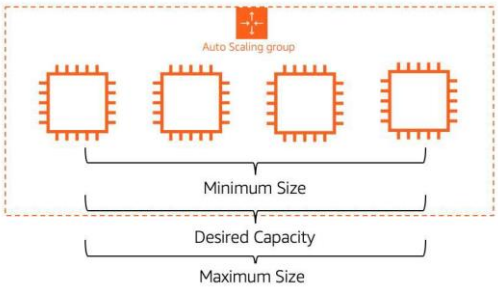

Figure 2: Scaling & Load balacning

This feature will analyze your access patterns on the AWS infrastructure and smartly adjust the capacity according to your required workload to maintain zero production downtime keeping in mind the expense. Along with that, Elastic Load balancing ensures that you have zero downtime based on latency, geographic location and in times of abrupt compute power termination. It can directly route users to appropriate compute power resource based on geography, latency and access. It can direct users to different availability zones based and if paired if scaling, it can commission new compute resources in seconds when it detects that one of its fleet members has gone down. [20][21]

4. Market Making

HFTs generally provide liquidity to the market and so in turn, they get paid to trade and it's a win-win situation for both the HFT trader and the free market participants. Sometimes they conduct arbitrage or directional trading. In those instances, they take liquidity but they're also driving price discovery, which is good for the consumers as well as the financial market. The best way to understand market making is to understand a situation where there are no market makers. [9]

Bid	Quantity	Offer	Quantity
$320.1	100	$324.6	200
$315.5	100	$326.54	26
$313.43	200	$327.53	282
$303.29	111	$330.3	314
$302.16	500	$335.2	100

Now any trader/investor will be able to see the evident risk in the above market depth for Ticker: PLTR.

Say if you are a buyer, the best price for which you can buy a PLTR stock is $324. Say if you want to buy 500 shares, then you will have to buy at approx. $327/stock which is $2 off the mark per share. Similarly, if you are a seller, the best price for which you can sell the stock is $320, and if you want to sell around 400 shares, then you will have to go almost $4.75 down per share. Even for the average trader/investor out there, these are the two highest risks here. First the gap between the best bid buyer and best offer is $4. In liquid stocks its almost never more than 10-20 cents (decent market conditions). This is where market makers come in. Market makers will give buy and sell quotes in such a way that the liquidity will automatically get created in the market. Market makers provide buy and sell offers in sophisticated ways so as to create liquidity, reduce uncertainty and contract manipulation and speculation. [10]

Bid	Quantity	Offer	Quantity
$321.1	150	$321.3	150
$320.05	100	$321.05	26
$320.01	200	$321.10	282
$319.30	111	$323.3	314
$319.40	500	$323.35	100

In the above market depth of PLTR, the gap between the best bid and best offer Is $0.20 cents and the gap between subsequent quotes is just $0.05 - $0.10 cents, this reduces the risks involved significantly.

So, the market maker for PLTR is a trader or arbitrageur, who will typically place a bid at $321.1 and also an offer at $321.3. The gap of $0.2 is the spread on which the market maker will trade and book profits. There may be some additional hidden exchange fees but the market maker operates on huge volumes not margins on a daily basis, however the risk here is that there's no guarantee that both of the bid and offer quotes will be executed, this is the only risk the market maker has to trade on. [11][12]

5. Dark Pools

Dark Pools are basically private exchanges where you "Trade" instruments and the trade information is not accessible to the initial public or individual investor until after the trade is executed. The reason they are called Dark pools is because of their lack of transparency and regulation/government intervention.

Say you own a Public Pension Fund, and you want to sell $4.1B worth of GE shares. Now, if you make this trade on the Public Exchange, this may create excess liquidity in the markets and may even be a signal for the Exchange Commission to scrutinize your trade and fund for manipulation and insider trading. This is where Dark pools come to the rescue, you can easily find a buyer for your shares without making this information public before execution, you have a better chance of finding a buyer here for your huge volume since dark pools are dedicated to large investors and you only have to publicize the trade details after the trade has been executed that too because you are a Public Pension Fund. If you had sold off your shares on the Public Exchange, your trade weightage would bring the GE market value down significantly creating fear and uncertainty among the individual investors which would eventually lead to a huge sell off.

This lack of transparency actually works in both the institutional and individual investor's favor since it may result in a better-realized price and won't have any effect on the public markets. Your trades can go unnoticed as trade details will only be reflected on the consolidated tape after a fair delay and even if the amount of trading in dark pools owned by market makers continues to grow, stock prices on exchanges may not reflect the actual market value of securities. Through Dark pools you may also have a chance to lower transaction costs on trades as you do not have to pay the exchange fees. However, the disadvantages of Dark Pools outweigh the advantages. As all trades are anonymous and out of public reach, this can sideline other investors not using these pools which may result in unfair and undue advantages and may create what we call as: unheard of and unseen information gap in the markets. This is the major reason Dark Pools are being viewed with suspicion by regulators in major financial nations. [13][14]

6. Effects of HFT in the financial markets

HFT Flash Crash occurred on 6th May 2010, when HFT was used to execute a trade worth $4.1 Billion. Due to an anomaly in the HFT strategy, the Dow Jones Industrial Average plummeted 1000 points causing the loss of $1 Trillion in market value. This software glitch was quickly detected and fixed within 36 minutes, but this caused a lot of chaos and panic in the financial markets. Following the flash crash, new regulations were introduced by various stock exchanges of multiple countries to avoid similar glitches in the future. However, flash crash is known to rapidly recover partial or total value lost during the flash.

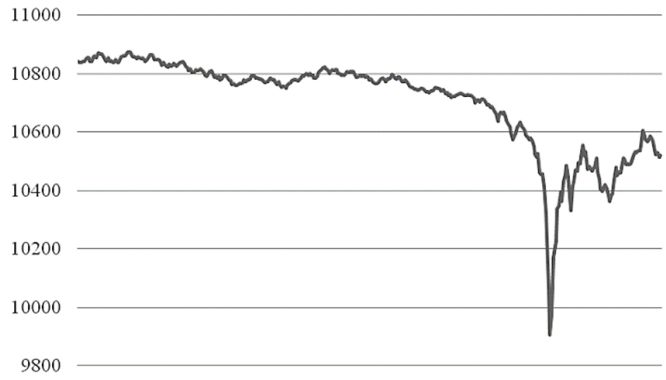

Figure 3. May 2010 – Dow Jones Industrial Average

During the Facebook IPO, Nasdaq faced a difficult technical issue for trading of the Facebook stock. This technical issue occurred because the software was unable to handle the rapid order execution and order cancellation placed by the HFT algorithms.[16] The orders placed and cancelled were substantially large positions. The algorithms were tremendously interested in the Facebook IPO. This continuous order placement and cancellation eventually got stuck in an infinite loop. This technical glitch led to inaccurate pricing of the Facebook stock by Nasdaq and caused a lot of trading disputes and all these problems eventually lost $460 million. [17]

7. Conclusion

Our solution to this problem is building a recommendation engine that is integrated with cloud computing resources in such a way that users of the engine only have to provide it with appropriate data or just simply define a path from where it can retrieve the data. The engine can then use the cloud resources such as computing power which will be available on the go. The resources will automatically be provisioned according to the usage and will also be decommissioned when the usage drops below a specified threshold. Moreover, the data that the engine works on will also be available in a dynamic environment that changes the underlying storage device according to how the data is being accessed. In a nutshell, we can build a Recommendation Engine that can automatically provision and decommission resources it needs to work on depending upon its own usage metrics.

REFERENCES

[1]. Statistical Arbitrage Trading Strategies and High Frequency Trading - Thomas A. Hanson, Joshua Hall link DOI:10.13140/RG.2.2.20620.10889 September 2012SSRN Electronic Journal
[2]. What Do We Know About High-Frequency Trading? - Charles M. Jones DOI:10.2139/ssrn.2236201 March 2013SSRN Electronic Journal
[3]. High frequency trading and its impact on market quality - Jonathan A. Brogaard April 2011 Corpus ID: 159415842
[4]. Report on Direct Market Access and Ultra Low Latency Trading in India - Bhanu Chandar Udatha March 2011SSRN Electronic Journal DOI:10.2139/ssrn.1795782
[5]. News-based trading strategies - Stefan Feuerriegel, Helmut Prendinger July 2016 Decision Support Systems 90 DOI:10.1016/j.dss.2016.06.020
[6]. Trading Strategies To Exploit Blog and News Sentiment - Wenbin Zhang, Steven Skiena Conference: Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23-26, 2010
Comparison of different market making strategies for high frequency traders - Yibing Xiong, Takashi Yamada, Takao Terano
[7]. The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes, and the Probability of Informed Trading - David Easley, Marcos M. López de Prado and Maureen O'Hara The Journal of Portfolio Management, Vol. 37, No. 2, pp. 118-128, Winter 2011
[8]. What Do We Know About High-Frequency Trading - Charles M. Jones? Columbia Business School Research Paper No. 13-11 Date Written: March 20, 2013
[9]. Exchange-Traded Funds, Market Structure, and the Flash Crash - Ananth Madhavan October 2011Financial Analysts Journal 68 DOI:10.2139/ssrn.1932925
[10]. Benefits of AWS in Modern Cloud - Sourav Mukherjee March 2019 DOI:10.5281/zenodo.2587217
[11]. Cost comparison of running web applications in the cloud using monolithic, microservice, and AWS Lambda architectures - Mario Villamizar, Oscar Garcés, Lina Ochoa, Harold Castro, Lorena Salamanca, Mauricio Verano, Rubby Casallas, Santiago Gil, Carlos Valencia, Angee Zambrano & Mery Lang June 2017 Service Oriented Computing and Applications 11 DOI:10.1007/s11761-017-0208-y Project: Colombian geoscience data cube.
[12]. Storage options in the AWS cloud - Joseph Baron, Sanjay Kotecha July 2016 Conference: EIT New Zealand
[13]. Performance and cost analysis of the Supernova factory on the Amazon AWS cloud - Jackson, Keith R., Muriki, Krishna, Ramakrishnan, Lavanya, Runge, Karl J., Thomas, Rollin C. January 2011Scientific Programming 19(2-3):107-119 DOI:10.1155/2011/498542 SourceDBLP
[14]. Database security management for healthcare SaaS in the Amazon AWS Cloud - Fabio Bracci; Antonio Corradi; Luca Foschini Published 2012 Computer Science 2012 IEEE Symposium on Computers and Communications (ISCC)
[15]. Accelerating Memcached on AWS Cloud FPGAs - Jongsok Choi, Jason Anderson HEART 2018: Proceedings of the 9th International Symposium on Highly-Efficient Accelerators and Reconfigurable TechnologiesJune 2018 Article No.: 2Pages 1–8https://doi.org/10.1145/3241793.3241795
[16]. Scheming a Proficient Auto Scaling Technique for Minimizing Response Time in Load Balancing on Amazon AWS Cloud - M Arvindhan, Abhineet Anand International Conference on Advances in Engineering Science Management & Technology (ICAESMT) - 2019, Uttaranchal University, Dehradun, India

[17]. Infrastructure Cost Comparison of Running Web Applications in the Cloud Using AWS Lambda and Monolithic and Microservice Architectures - Mario Villamizar; Oscar Garcés; Lina Ochoa; Harold Castro; Lorena Salamanca; Mauricio Verano May 2016 DOI:10.1109/CCGrid.2016.37 Conference: 2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)At: Bogotá, Colombia Project: cloud architectures.

[18]. Comparison among cloud technologies and cloud performance - O Alzakholi, H Shukur, R Zebari, S Abas April 2020Journal of Applied Science and Technology Trends 1(2):40-47 DOI:10.38094/jastt1219 Project: RG Academic Publishers & Reviewerss

[19]. Machine Learning Best Practices in Financial Services – Stefan Natu, David Ping, Alvin Huang July 2020 First publication AWS Whitepapers.

[20]. Financial Services Grid Computing on AWS - Alex Kimber, Ian Meyers September 2019 Updates to services, diagrams, and topology AWS Whitepapers.

[21]. AWS & Cybersecurity in the Financial Services Sector - Rahul Prabhakar, Mark Ryland, Bill Shinn July 2019 First publication AWS Whitepapers.

[22]. Derive Insights from AWS Lake House - Raghavarao Sodabathina, Changbin Gong June 3, 2021, First publication AWS Whitepapers.