Proceedings of CONF-MLA 2024 Workshop: Semantic Communication Based Complexity Scalable Image Transmission System for Resource Constrained Devices
DOI: 10.54254/2755-2721/83/2024GLG0075

Strategy Selection Using Multi-Armed Bandit Algorithms in Financial Markets

Yushi Guo

School of Information and Communication Engineering, Hainan University, 58 Renmin Avenue, Haikou, Hainan, 570228, P.R. China

20223004700@hainanu.edu.cn

Abstract. This paper aims to evaluate the effectiveness of Multi-Armed Bandit (MAB) algorithms in choosing the optimal trading strategy among the sub optimal ones within financial markets. The research aims to addresses the challenge of adapting to dynamic market conditions. By introducing a Composite Trading Strategy that integrates trend-following, mean reversion, and momentum strategies, the study investigates whether increased trading frequency enhances the performance of profitability of various MAB algorithms, including UCB, Thompson Sampling, and epsilon-greedy. The experimental results indicate that while the introduction of complex strategies greatly improves returns in favorable market conditions, MAB algorithms still face limitations in adverse market environments. The findings highlight the potential of MAB algorithms in financial strategy selection and suggest directions for future research to enhance their adaptability in adverse markets.

Keywords: Multi-Armed Bandit Algorithms, Algorithmic Trading Optimization, Financial Strategy Selection.

1. Introduction

1.1. Research Background

Multi-Armed Bandit (MAB) algorithms have garnered increasing attention due to their effectiveness in addressing the exploration-exploitation trade-off in various decision-making scenarios. In financial markets, traders are constantly faced with the challenge of selecting the most profitable strategy in rapidly changing environments, which often involve significant uncertainty. The traditional trading strategies, although effective under certain conditions, typically lack the flexibility to adapt dynamically to these ever-changing market conditions, leading to suboptimal performance.

MAB algorithms are well-suited to this context as they enable continuous learning and adaptation by exploring new strategies and exploiting known profitable ones. While there has been considerable research on the application of MAB algorithms in finance, most of it focuses on portfolio management and risk optimization. However, the application of MAB algorithms to real-time strategy selection in dynamic financial environments remains largely unexplored.

This research aims to address this gap by investigating the use of MAB algorithms to optimize strategy selection in financial markets. Specifically, we introduce a Composite Trading Strategy that integrates trend-following, mean-reversion, and momentum strategies, designed to increase trading opportunities and enable MAB algorithms to make more frequent and informed decisions. By testing various MAB algorithms such as Upper Confidence Bound (UCB), Thompson Sampling, and epsilon-greedy, this study evaluates their effectiveness in different market conditions.

1.2. Research Objectives

The main objective of this research is to evaluate the effectiveness of MAB algorithms in identifying and selecting the optimal trading strategy from a set of initially unclear options. Specifically, this study compares the performance of different MAB algorithms across two strategy combinations, including both simple and complex strategies. The introduction of the Composite Trading Strategy serves as a key experiment to test whether increasing the trading frequency can enhance the performance of MAB algorithms.

This research seeks to answer the following key questions:

1.Can MAB algorithms effectively identify and exploit the best strategies in a trading environment with low-frequency strategies?

2.How does the introduction of the Composite Trading Strategy impact the performance of MAB algorithms compared to using all simpler, more traditional strategies?

3.What are the implications of using MAB algorithms for real-world financial strategy selection, particularly in terms of maximizing returns?

2. Literature Review

2.1. Applications of Multi-Armed Bandit Algorithms in Finance

MAB algorithms have been successfully applied in various areas of finance, primarily in portfolio selection and risk management. For instance, Shen et al. (2015) introduced MAB models in portfolio selection, demonstrating their utility in optimizing asset allocation under uncertain conditions [1]. Chanca (2022) extended this research by incorporating transaction costs into the decision-making process, emphasizing the need to account for real-world factors when applying MAB algorithms in financial markets [2].

Additionally, Huo and Fu (2017) developed a risk-aware MAB framework that focuses on balancing returns and risks in portfolio selection. Their research highlighted the importance of risk-awareness in financial decision-making, but their approach was limited to static market environments [3]. These studies have laid the foundation for the application of MAB algorithms in finance, but they are primarily concerned with optimizing long-term investment decisions rather than real-time strategy selection.

Recent advancements have further explored MAB algorithms in more complex market environments. De Curtò et al. (2023) proposed an integration of MAB algorithms with large language models (LLMs) to improve decision-making in non-stationary markets, highlighting the adaptability of these algorithms in changing environments [4]. Meanwhile, Cannelli et al. and Zhu et al. studied the benefits of adding multiple trading strategies to the MAB framework, and found that such combinations can boost returns, albeit with increased complexity and computational demands [5, 6]. Additionally, Bernasconi et al. explored strategies for dark-pool smart order routing, showing the potential of MAB algorithms in handling complex market order flows, especially in non-public market environments [7]. Liu and Cartlidge studied nonstationary continuum-armed bandit strategies and examined the performance of MAB algorithms in terms of dynamic adaptability [8].

However, these studies do not focus on the dynamic selection of trading strategies in real-time, a crucial element for optimizing decision-making in high-frequency trading.

2.2. Limitations of Existing Research

Despite the progress made in applying MAB algorithms to portfolio optimization and risk management, several limitations remain. First, most existing research focuses on long-term investment decisions and static market conditions, which do not reflect the fast-paced nature of financial trading environments

DOI: 10.54254/2755-2721/83/2024GLG0075

where strategies need to adapt in real-time. This limits the applicability of these studies to high-frequency trading, where decision-making must be quick and responsive to changing market conditions.

Second, while MAB algorithms have been effective in portfolio selection, they have not been extensively studied in the context of strategy selection. Most of the existing literature treats strategies as fixed and focuses on optimizing the selection of assets within a portfolio. However, financial markets are dynamic, and the effectiveness of a strategy can vary significantly based on market conditions. The lack of research on applying MAB algorithms to dynamically select and switch between strategies in response to changing market signals represents a significant gap.

Lastly, many studies incorporate MAB algorithms without fully addressing the computational complexity involved in applying these algorithms in real-time trading environments. The integration of multiple strategies into a composite framework, which allows for more frequent and diverse trading signals, adds complexity to the decision-making process. Yet, little research has explored how MAB algorithms handle this increased complexity in high-frequency trading environments.

2.3. Research Gap and Novelty

To the best of our knowledge, no existing studies have applied MAB algorithms specifically for real-time strategy selection in financial markets. This represents a major gap in the literature, as the ability to dynamically switch between different trading strategies in response to changing market conditions is crucial for optimizing trading performance.

This research addresses that gap by introducing a novel application of MAB algorithms for strategy selection. The proposed Composite Trading Strategy combines trend-following, mean-reversion, and momentum strategies, which allows for more frequent trading opportunities and better optimization of decision-making under varying market conditions. By evaluating the performance of several MAB algorithms across different market scenarios, this study not only expands the application of MAB algorithms in finance but also provides insights into their adaptability in high-frequency trading environments.

The novelty of this research lies in its focus on strategy selection rather than portfolio optimization, providing a new perspective on how MAB algorithms can be used to dynamically respond to market fluctuations. This approach has the potential to significantly improve trading outcomes by allowing for more flexible and adaptive decision-making processes.

3. Methodology

3.1. Overview of Multi-Armed Bandit (MAB) Algorithms

This section provides an overview of the Multi-Armed Bandit (MAB) algorithms utilized in this study. These algorithms are designed to tackle the exploration-exploitation trade-off, where a balance must be struck between trying out new strategies (exploration) and using the strategies that have performed well in the past (exploitation). The following MAB algorithms were selected for their effectiveness in various decision-making scenarios and their potential applicability in financial markets.

3.1.1. UCB Algorithm

The Upper Confidence Bound (UCB) algorithm is a widely used approach in the MAB problem. It selects strategies based on both the average observed rewards and the uncertainty associated with those rewards. The UCB algorithm calculates an upper confidence bound for each strategy, and the strategy with the highest bound is chosen. This method inherently balances exploration and exploitation by favoring strategies that either have high observed rewards or have been less frequently tried, thus having higher uncertainty.

The UCB value for each strategy i at time t is given by:

$$UCB_i(t) = \hat{\mu}_i(t) + c\sqrt{\frac{2\ln t}{n_i(t)}} \quad (1)$$

where:

- $\hat{\mu}_i(t)$ is the average reward for strategy i up to time t.
- $n_i(t)$ is the number of times strategy i has been selected up to time t.
- c is a parameter that controls the degree of exploration.

3.1.2. SW-UCB Algorithm

The Sliding Window UCB (SW-UCB) algorithm is an adaptation of the UCB algorithm designed to respond more effectively in environments where recent rewards may be more relevant than older ones. Instead of considering all past rewards, SW-UCB uses a sliding window to focus on the most recent rewards. This allows the algorithm to adapt more quickly to changes by giving more weight to recent data.

The SW-UCB value for each strategy i at time t is calculated similarly to the UCB algorithm, but only includes the rewards and counts within a fixed sliding window W:

$$SW_{UCB_i(t)} = \hat{\mu}_i^W(t) + c\sqrt{\frac{2\ln t}{n_i^W(t)}} \tag{2}$$

where W is the size of the sliding window.

3.1.3. EWMA-UCB Algorithm

The Exponentially Weighted Moving Average UCB (EWMA-UCB) algorithm is another variant of UCB, tailored to scenarios where more recent rewards should be given greater importance. This algorithm applies an exponentially decreasing weight to past observations, placing more emphasis on recent rewards. This approach helps the algorithm adapt to shifts in strategy performance more quickly than the standard UCB.

The EWMA-UCB value is given by:

$$EWMA_{UCB_i(t)} = \hat{\mu}_i^\lambda(t) + c\sqrt{\frac{2\ln t}{n_i^\lambda(t)}} \tag{3}$$

where λ is the decay factor that controls how quickly the weight of past observations diminishes.

3.1.4. Thompson Sampling (TS) Algorithm

Thompson Sampling is a Bayesian approach to the MAB problem, where the probability of selecting a strategy is proportional to the probability that it is the best option. For each strategy, a posterior distribution of the expected reward is maintained and updated after each trial. The strategy with the highest sampled reward from its posterior distribution is selected. This method inherently balances exploration and exploitation according to the uncertainty in the reward estimates.

3.1.5. Epsilon-Greedy Algorithm

The epsilon-greedy algorithm is a simple yet effective approach for balancing exploration and exploitation. With a small probability ϵ, the algorithm explores by randomly selecting a strategy, and with probability $1-\epsilon$, it exploits by selecting the strategy with the highest estimated reward. This method is straightforward to implement and can perform well in a variety of settings, although it may not be as efficient as more sophisticated algorithms like UCB or Thompson Sampling, particularly in environments where optimal strategies need to be identified quickly.

3.2. Strategy Design

This section explores trading strategies ranging from simple, traditional methods to more advanced approaches, setting the stage for an in-depth examination of Multi-Armed Bandit (MAB) algorithms in different trading environments.

3.2.1. Initial Strategy Set

The initial strategy set comprises simple, well-known trading strategies that serve as a baseline for evaluating the performance of MAB algorithms. These strategies are selected for their simplicity and widespread use in various trading systems. They include:

DOI: 10.54254/2755-2721/83/2024GLG0075

- **SMA Crossover**: A trend-following strategy that generates a buy signal when a short-term Simple Moving Average (SMA) crosses above a long-term SMA, and a sell signal when the short-term SMA crosses below the long-term SMA.
- **RSI Strategy**: A momentum-based strategy that uses the Relative Strength Index (RSI) to identify overbought and oversold conditions. The strategy generates buy signals when the RSI falls below a certain threshold (indicating the asset is oversold) and sell signals when the RSI exceeds a certain threshold (indicating the asset is overbought).
- **Bollinger Bands**: A volatility-based strategy that uses Bollinger Bands to identify price extremes. The strategy generates buy signals when the price touches the lower band and sell signals when the price touches the upper band.

These strategies, while effective in certain market conditions, often result in low trading frequencies due to their reliance on specific market signals, making them ideal candidates for testing the limitations of MAB algorithms.

3.2.2. Composite Trading Strategy

To address the limitations observed with low-frequency trading strategies, the Composite Trading Strategy was introduced. This strategy is designed to increase trading frequency by integrating multiple trading approaches into a single strategy. Specifically, the Composite Trading Strategy combines elements of trend-following, mean reversion, and momentum strategies, allowing it to generate more frequent trading signals across various market conditions.

- **Trend-Following Strategy: Moving Average Crossover**
  - This component generates buy signals when a short-term SMA crosses above a long-term SMA, indicating a potential upward trend, and vice versa for sell signals.
- **Mean Reversion Strategy: Bollinger Bands**
  - This component generates buy signals when the current price approaches the lower Bollinger Band, anticipating a price rebound, and sell signals when the price approaches the upper Bollinger Band, anticipating a price correction.
- **Momentum Strategy: RSI Filter**
  - The momentum component uses the RSI to filter signals, generating sell signals when the RSI exceeds 70 (indicating overbought conditions) and buy signals when the RSI falls below 30 (indicating oversold conditions).

While MAB algorithms offer a dynamic approach to strategy selection, their performance may be limited when applied to simple trading strategies with low trading frequency. These simple strategies, triggered by single conditions, may not generate enough trading opportunities, thus restricting the MAB algorithms' ability to efficiently explore and exploit the strategy space. To address this limitation, this study introduces a more complex trading strategy—the Composite Trading Strategy. By integrating elements of trend-following, mean reversion, and momentum strategies, this Composite Strategy aims to increase trading activity, providing MAB algorithms with more opportunities to optimize decision-making.

The Composite Trading Strategy is expected to subject the MAB algorithms to more rigorous testing by increasing the number of trading opportunities, thereby enabling the algorithms to better demonstrate their optimization capabilities. In this way, not only can the adaptability of MAB algorithms be validated across diverse market conditions, but their performance limits when faced with complex strategies can also be explored, ultimately advancing their broader application in real-world financial markets.

3.3. Application of MAB Algorithms in Financial Markets

MAB algorithms offer several advantages when applied to strategy selection in financial markets. Their ability to dynamically balance exploration and exploitation makes them well-suited for environments where quick adaptation to market changes is critical. By leveraging real-time data and continuously updating their strategy selections, MAB algorithms have the potential to outperform static strategies that rely solely on historical performance.

DOI: 10.54254/2755-2721/83/2024GLG0075

4. Experimental Design

This chapter outlines the design of the experiments conducted to evaluate the performance of Multi-Armed Bandit (MAB) algorithms in selecting optimal trading strategies. The experiments were designed to test the algorithms across various strategy combinations, including both simple strategies and the more complex Composite Trading Strategy. The experiments aim to assess how effectively these algorithms can identify and exploit the best strategies to maximize returns.

4.1. Data Source Selection

To create a controlled environment for testing, simulated market data were generated using a Geometric Brownian Motion (GBM) model[9-11]. The GBM model is commonly used in financial modeling to simulate the random behavior of asset prices over time. The key parameters for the GBM model, such as drift (μ), volatility (σ), and the initial price, were varied across different simulations to create datasets with different market characteristics, including varying levels of volatility and trend direction.

The use of simulated data enables for the testing of MAB algorithms in a controlled setting where the underlying market behavior is known. This is particularly useful for evaluating how well the algorithms can adapt to different market conditions and whether they can successfully identify profitable strategies.

4.2. Experimental Procedures

The experiments were designed to compare the performance of MAB algorithms when applied to both simple and complex trading strategies. The procedures were structured to observe how each algorithm balances exploration and exploitation and how effectively it can optimize strategy selection in different environments.

4.2.1. Baseline Experiment

In the baseline experiment, the MAB algorithms were tested using the initial set of simple trading strategies. These strategies, which include SMA Crossover, RSI Strategy, and Bollinger Bands, were chosen for their simplicity and because they represent commonly used techniques in trading.

- **Procedure**:
  - Each MAB algorithm was initialized with equal probabilities assigned to each strategy.
  - Over multiple trading periods, the algorithms selected strategies based on their respective decision rules (e.g., UCB, Thompson Sampling).
  - The performance of each strategy was tracked in terms of cumulative regret, final returns, and strategy selection frequency.
  - The baseline experiment was repeated multiple times (at least 8 iterations) with different market data scenarios to ensure the robustness of the results.
- **Objective**: The objective was to establish a benchmark for the algorithms' performance when applied to simple strategies and to identify any limitations related to low trading frequency or insufficient exploration opportunities.

4.2.2. Composite Strategy Introduction Experiment

The second set of experiments introduced the Composite Trading Strategy into the pool of available strategies. This strategy was designed to generate more frequent trading signals by integrating multiple trading techniques (trend-following, mean reversion, and momentum). The purpose of these experiments was to determine whether the increased trading activity provided by the Composite Strategy would enhance the MAB algorithms' ability to optimize strategy selection.

- **Procedure**:
  - The Composite Trading Strategy was added to the existing pool of simple strategies.
  - The same MAB algorithms were run under similar market conditions as in the baseline experiment, with the addition of the Composite Strategy as an option.

DOI: 10.54254/2755-2721/83/2024GLG0075

  - The algorithms' performance was again tracked over multiple trading periods, focusing on cumulative regret, final returns, and the frequency of strategy selection.
  - These experiments were also repeated multiple times (at least 8 iterations) to ensure consistency and reliability of the results.
- **Objective**: To evaluate whether the introduction of a more active trading strategy would improve the overall performance of the MAB algorithms, particularly in terms of increasing returns and reducing cumulative regret.

4.3. Evaluation Metrics

To assess the performance of the MAB algorithms, several key metrics were used. These metrics provide a comprehensive view of how well each algorithm performed in terms of strategy optimization and overall profitability.

4.3.1. Cumulative Regret

Cumulative regret is a standard measure in MAB problems that quantifies the difference between the actual returns obtained by the algorithm and the returns that would have been obtained by always choosing the best possible strategy. Lower cumulative regret indicates better performance, as it means the algorithm has made fewer suboptimal choices.

- **Calculation**: For each algorithm and strategy combination, cumulative regret was calculated at each time step and then summed over the entire trading period.

4.3.2. Strategy Selection Frequency

This metric tracks how often each strategy was selected by the MAB algorithm over the course of the experiment. It provides insights into the algorithm's behavior—whether it is heavily exploring (trying out different strategies) or exploiting (sticking with a known profitable strategy).

- **Analysis**: By examining the selection frequency, we can assess whether the algorithm is biased toward certain strategies and how this affects overall performance.

4.3.3. Final Returns

Final returns measure the total profit generated by the algorithm at the end of the trading period. This metric is crucial for evaluating the practical success of the strategy selection process, as it directly reflects the profitability of the chosen strategies.

- **Comparison**: Final returns were compared across different algorithms and strategy sets to determine which combinations yielded the highest profits.

5. Experimental Results

This section analyzes the performance of MAB algorithms across diverse market scenarios using simple and complex trading strategies. We begin by assessing basic strategies like SMA Crossover, RSI, and Bollinger Bands through key metrics such as cumulative regret and final returns, then explore the enhanced effectiveness of composite strategies in varied market conditions.

5.1. Results of the Initial Strategy Set

- **Cumulative Regret**: Present the performance of MAB algorithms in the simple strategy set, including selection frequency and final returns.
- **Final Returns**: Analyze the profitability of MAB algorithms in the initial strategy set, particularly whether they can identify relatively superior strategies.

DOI: 10.54254/2755-2721/83/2024GLG0075

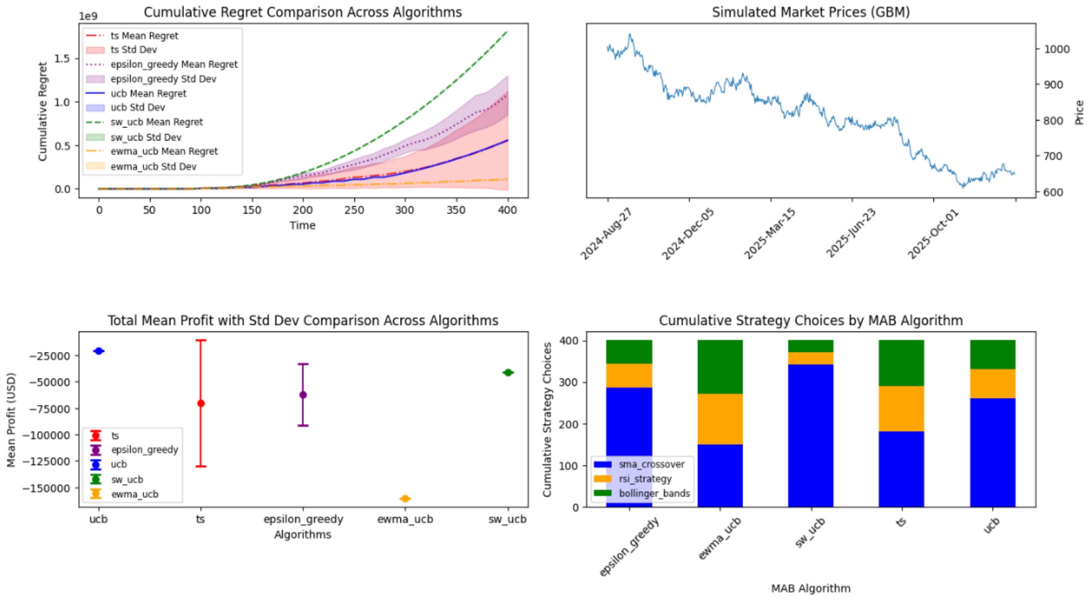

Figure 1. Performance in a down trend market.

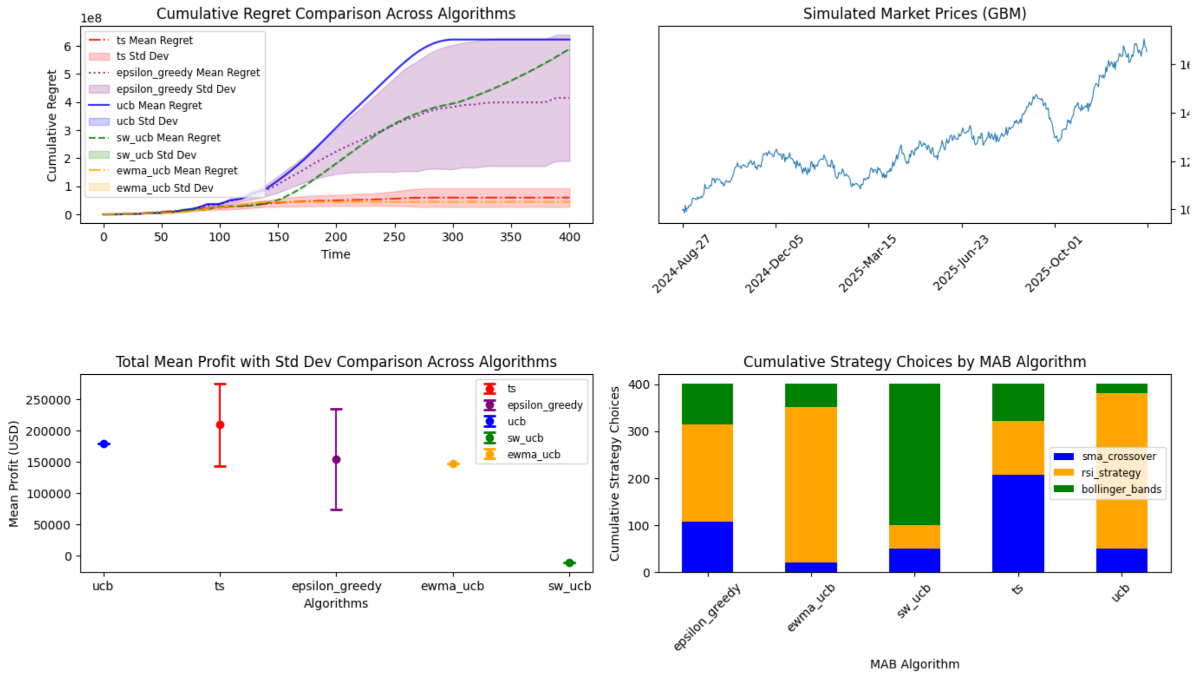

Figure 2. Performance in a up trend market.

DOI: 10.54254/2755-2721/83/2024GLG0075

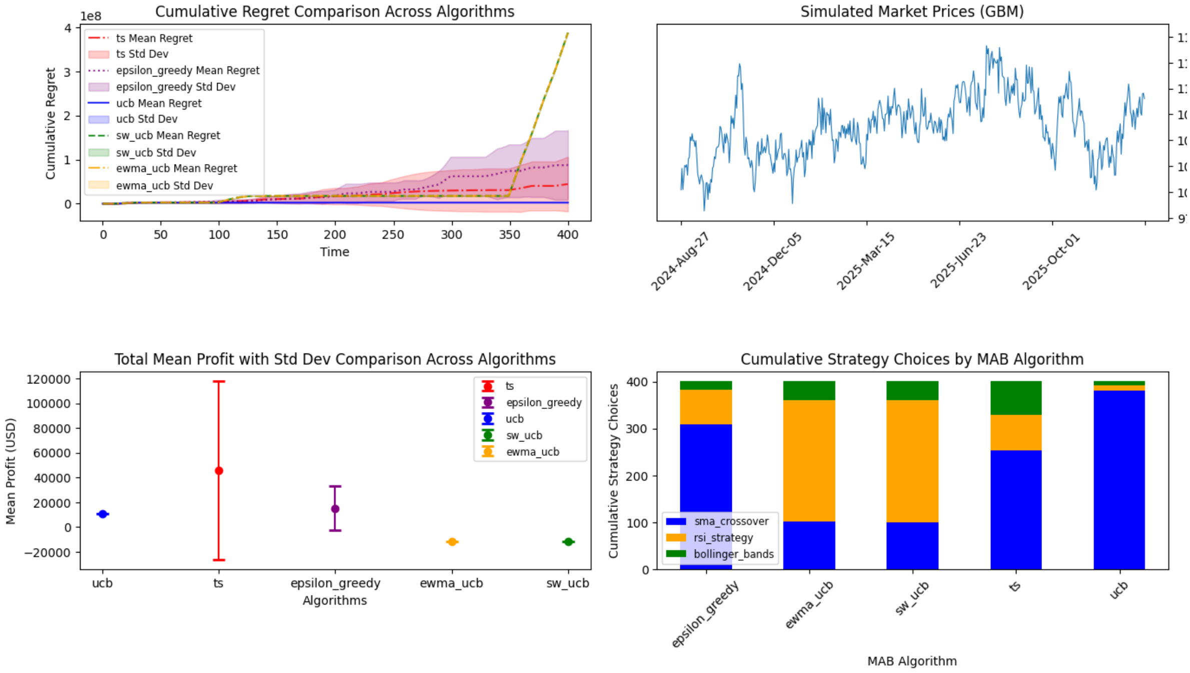

Figure 3. Performance in a range-bound market.

As Figure 1 shows, none of the algorithms achieved profitability, and the cumulative regret increased exponentially across all of them.

As Figure 2 shows, all algorithms except SW-UCB were able to achieve profitability.

As Figure 3 shows, EWMA_UCB and SW_UCB incurred losses, and their regret slopes were significantly steeper than those of the other MAB algorithms, indicating that the RSI strategy, which they selected most frequently, performed poorly in this context.

Table 1. Eight experiments were conducted for each type of market, with the profitability of each algorithm summarized in the table below

MAB algorithm	count of profitable times		
	up trend	Down trend	Range-bound
UCB	8	1	3
TS	7	1	3
Epsilon greedy	7	1	4
EWMA UCB	6	2	4
SW UCB	6	2	4

Table 1 shows that the Initial Strategy Set tends to generate relatively stable profits in uptrend markets. However, in range-bound markets, the profitability is less consistent, and in downtrend markets, it is difficult to achieve any profits.

5.2. Results of the Complex Strategy Introduction Experiment

- **Return Comparison**: Present the significant improvement in returns after introducing complex strategies and compare with the initial strategy set.
- **Cumulative Regret Changes**: Analyze whether the cumulative regret of MAB algorithms significantly decreases after the introduction of complex strategies, indicating that the algorithm effectively identified and exploited the advantages of the complex strategies.

DOI: 10.54254/2755-2721/83/2024GLG0075

- **Strategy Selection Frequency**: Show whether MAB algorithms increasingly selected complex strategies after their introduction.

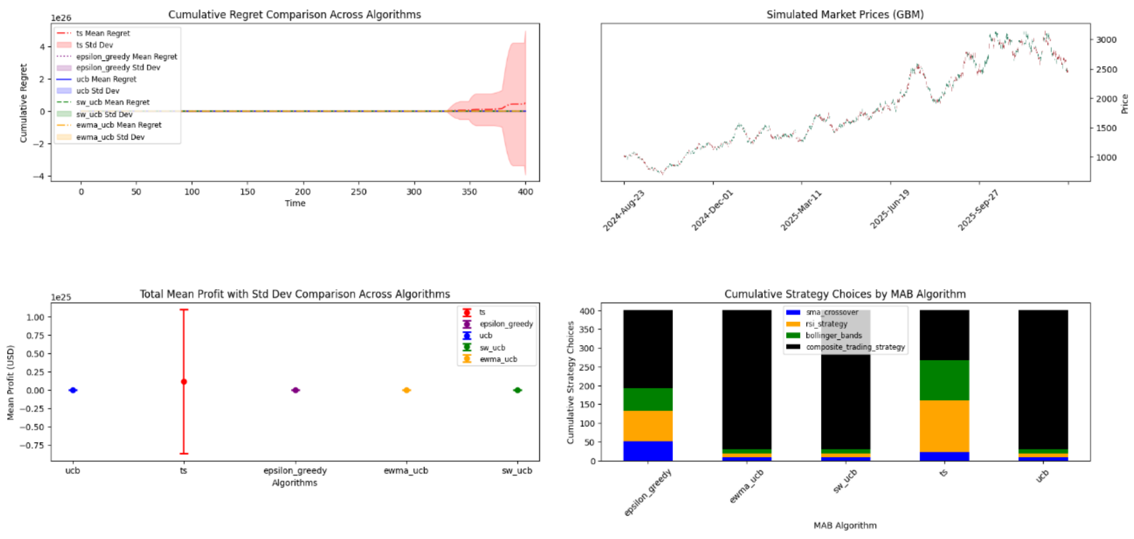

Figure 4. In this case simply justify the caption so that it is as the same width as the graphic.

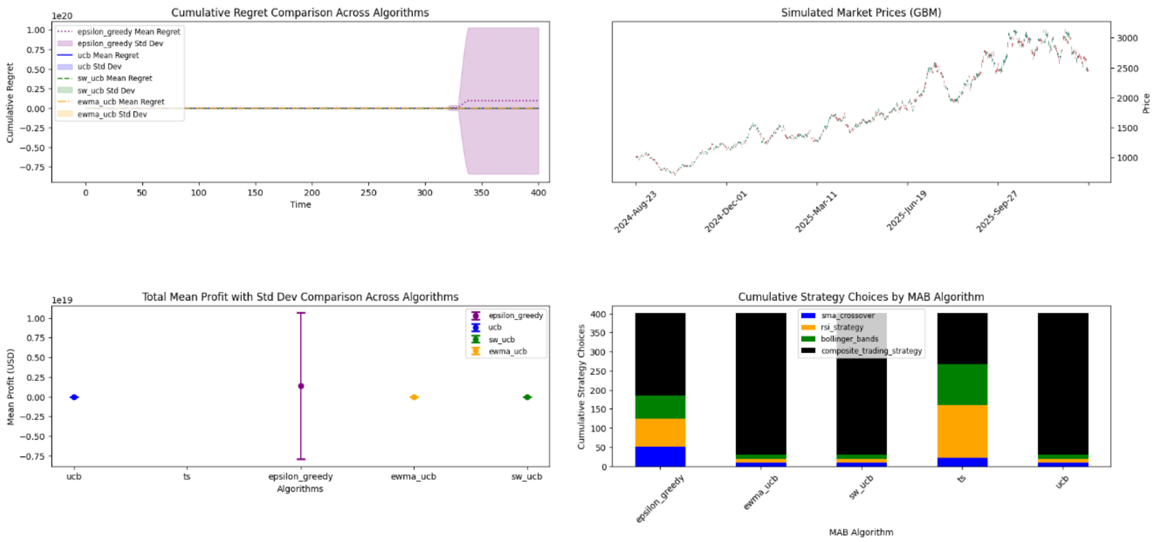

Figure 5. In this case simply justify the caption so that it is as the same width as the graphic.

Figure 4 and figure 5 are the same experiment, but for better visualization, figure 5 do not draw Thompson Sampling.

Figures 4 and 5 show that all strategies were very profitable in up trend market, with Thompson Sampling and Epsilon Greedy achieving the highest profits. However, due to their randomness, they also exhibited significant variation in profits across repeated experiments.

DOI: 10.54254/2755-2721/83/2024GLG0075

Table 2. Eight experiments were conducted for each type of market, with the profitability of each algorithm summarized in the table below

MAB Algorithm	count of profitable times		
	up trend	Down trend	Range-bound
UCB	8	1	3
TS	7	1	3
Epsilon greedy	7	1	4
EWMA UCB	6	2	4
SW UCB	6	2	4

Table 2 indicates that even with the introduction of a better-performing, higher-frequency Composite Strategy, there was no significant improvement in the profitability of MAB algorithms in adverse market conditions.

5.3. Comparison Experiment Results

These comparison experiments were designed to evaluate how the Composite Trading Strategy impacts MAB algorithm performance across different market conditions, focusing on returns and strategy selection frequency. The following sections detail the experiment setup and results.

5.3.1. Overview of Experiment Setup

This section presents the results of the comparison experiments, designed to evaluate the impact of introducing the Composite Trading Strategy on the performance of various MAB algorithms under different market conditions. The focus is on comparing the changes in returns and strategy selection frequency between the simple strategy group and the complex strategy group.

5.3.2. Comparison of Returns

In the simple strategy group, the algorithms were only able to achieve a high probability of profit in uptrend markets. In contrast, in downtrend and range-bound markets, they generally failed to generate profits in most instances. However, in the complex strategy group, the number of profitable trades was roughly the same across different market conditions. While the frequency of profitable trades was consistent, the magnitude of profits increased, particularly in uptrend markets.

Specifically, the Thompson Sampling (TS) and epsilon-greedy algorithms showed greater profitability after the introduction of the complex strategy. TS had the highest average returns, but it also exhibited a large standard deviation in its returns, indicating significant variability in its performance across different market conditions. The epsilon-greedy algorithm also performed well, with slightly lower average returns and standard deviation compared to TS.

5.3.3. Analysis of Strategy Selection Frequency

In terms of strategy selection, MAB algorithms in the simple strategy group were more inclined to choose higher-yielding strategies in uptrend markets, while in other market conditions, their strategy selection was more dispersed, with no clear advantage. After introducing the Composite Strategy, which integrates multiple trading approaches, the frequency of strategy selection became more balanced across different market conditions. This composite strategy increased trading frequency and improved profitability, particularly in uptrend markets, by providing a more diverse set of opportunities for the algorithms to exploit.

However, while the Composite Strategy demonstrated its strengths by enhancing trading frequency and returns in favorable market conditions, its limitations became apparent in adverse market conditions such as downtrends or range-bound markets. The MAB algorithms' performance in these unfavorable environments remained constrained, with no significant improvement in profitability, despite the increased trading opportunities offered by the composite approach.

This analysis suggests that while the Composite Strategy can optimize performance under certain market conditions, merely increasing trading frequency and diversifying strategies is not sufficient to overcome the inherent limitations of MAB algorithms in more challenging market environments. Further refinement of strategy design and adaptability is necessary to address these challenges effectively.

5.3.4. Summary of Findings

The comparison experiment results indicate that the introduction of the Composite Trading Strategy somewhat improved the profitability of MAB algorithms, particularly in uptrend markets. However, this improvement was mainly reflected in increased returns, while the performance of MAB algorithms under unfavorable market conditions (such as market downtrends or range-bound conditions) did not significantly improve. Although the Composite Strategy led to a more balanced strategy selection across different market conditions, it did not fully address the challenges faced by the algorithms in complex market environments.

These findings suggest that while the Composite Strategy increases trading opportunities and provides richer market signals, MAB algorithms still face limitations in dealing with adverse market conditions. This highlights the need for future research to further consider the diversity of market conditions and the adaptability of algorithms when designing strategies and optimizing their performance.

6. Conclusion

This study explored the application of Multi-Armed Bandit (MAB) algorithms in the dynamic selection of trading strategies within financial markets, a novel application that has not been addressed in existing literature. By introducing the Composite Trading Strategy, which integrates trend-following, mean-reversion, and momentum strategies, this research demonstrated how MAB algorithms can enhance decision-making in high-frequency trading environments.

The experimental results indicate that MAB algorithms such as Upper Confidence Bound (UCB), Thompson Sampling, and epsilon-greedy are effective in identifying optimal strategies under certain market conditions. In particular, the introduction of the Composite Trading Strategy led to increased trading opportunities, which enabled the algorithms to more effectively balance exploration and exploitation. However, the performance of these algorithms in adverse market conditions, such as downtrends or range-bound markets, remains limited. While the Composite Strategy improved profitability in favorable conditions, it did not fully overcome the inherent challenges faced by MAB algorithms in more complex and volatile market environments.

The primary contribution of this research is the application of MAB algorithms to real-time strategy selection, a field where these algorithms have not been previously explored. This study demonstrates that MAB algorithms can be applied beyond traditional portfolio optimization and risk management, offering new possibilities for improving trading performance in real-world financial markets.

Despite these findings, there are several limitations to this study. First, the experiments were conducted using simulated market data, which, while useful for controlled testing, may not fully capture the complexity and unpredictability of real-world financial markets. Second, the increased computational complexity introduced by the Composite Trading Strategy poses challenges for real-time application, especially in high-frequency trading scenarios where rapid decision-making is critical.

Future research should focus on further refining MAB algorithms to enhance their adaptability in adverse market conditions. One promising direction could involve integrating predictive models that can better anticipate market shifts, allowing the algorithms to make more context-aware decisions. Additionally, exploring the combination of MAB algorithms with machine learning techniques, such as reinforcement learning, could offer new avenues for improving decision-making under uncertainty. This

DOI: 10.54254/2755-2721/83/2024GLG0075

would not only address the challenges of strategy selection but also expand the practical applicability of MAB algorithms in increasingly complex financial markets.

In conclusion, this study has highlighted both the potential and limitations of MAB algorithms in financial strategy selection, underscoring the need for continued research to fully harness their capabilities in dynamic and uncertain trading environments.

References

[1] Shen W, Wang J, Jiang YG, Zha H. Portfolio choices with orthogonal bandit learning. InTwenty-fourth international joint conference on artificial intelligence 2015 Jun 23.
[2] Chanca, P.M.M., 2022. Multi-Armed Bandits for Selection of Risk-Based Portfolios with Transaction Costs.
[3] Huo X, Fu F. Risk-aware multi-armed bandit problem with application to portfolio selection. Royal Society open science. 2017 Nov 15;4(11):171377.
[4] De Curtò J, de Zarzà I, Roig G, Cano JC, Manzoni P, Calafate CT. LLM-informed multi-armed bandit strategies for non-stationary environments. Electronics. 2023 Jun 25;12(13):2814.
[5] Cannelli L, Nuti G, Sala M, Szehr O. Hedging using reinforcement learning: Contextual k-armed bandit versus Q-learning. The Journal of Finance and Data Science. 2023 Nov 1;9:100101.
[6] Zhu M, Zheng X, Wang Y, Li Y, Liang Q. Adaptive portfolio by solving multi-armed bandit via thompson sampling. arXiv preprint arXiv:1911.05309. 2019 Nov 13.
[7] Bernasconi M, Martino S, Vittori E, Trovò F, Restelli M. Dark-pool smart order routing: a combinatorial multi-armed bandit approach. InProceedings of the Third ACM International Conference on AI in Finance 2022 Nov 2 (pp. 352-360).
[8] Liu B, Cartlidge J. Nonstationary Continuum-Armed Bandit Strategies for Automated Trading in a Simulated Financial Market. arXiv preprint arXiv:2208.02901. 2022 Aug 4.
[9] Samuelson PA. Proof that properly anticipated prices fluctuate randomly. InThe world scientific handbook of futures markets 2016 (pp. 25-38).
[10] Black F, Scholes M. The pricing of options and corporate liabilities. InWorld Scientific Reference on Contingent Claims Analysis in Corporate Finance: Volume 1: Foundations of CCA and Equity Valuation 2019 (pp. 3-21).
[11] Hull JC, Basu S. Options, futures, and other derivatives. Pearson Education India; 2016.