Engineering Applications of Artificial Intelligence 160 (2025) 112021

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Research paper

Utilizing modularity to construct a trading strategies network for enhancing the diversification of group trading strategy portfolio by grouping genetic algorithm ☆

Kudakwashe Chideme [a], Chun-Hao Chen [b],*

[a] Department of Electronic Engineering, National Kaohsiung University of Science and Technology, No. 415, Jiangong Rd, Sanmin District, Kaohsiung, Taiwan
[b] Department of Computer Science and Information Engineering, National Kaohsiung University of Science and Technology, No. 415, Jiangong Rd, Sanmin District, Kaohsiung, Taiwan

ARTICLE INFO

Keywords:
Portfolio optimization
Grouping genetic algorithms
Genetic algorithms
Network modularity
Grouping problems
Trading strategy portfolio

ABSTRACT

Portfolio diversification is a perennial problem in finance. Identifying assets that offer uncorrelated returns is important for minimizing risk. To solve the problem, this paper proposes the Q-Group Trading Strategy Portfolio Optimization (QGTSPO) approach to increase the diversity of trading strategies for generating a QGTSP. QGTSPO extends the previous approach by modeling relationships between trading strategies using both correlation and mutual information, integrating a threshold component to determine optimal connectivity in the adjacency matrix. In other words, QGTSPO conceptualizes trading strategies as nodes in a graph and utilizes the grouping genetic algorithm to group similar nodes using modularity as a fitness measure. It first applies technical indicators and multiple financial datasets to generate candidate trading strategies with differentiated risk factors. The selected trading strategies are then used to generate chromosomes to initialize the population. Every possible QGTSP is encoded by the threshold, the take-profit and stop-loss, the grouping, and the weight parts. Next, the fitness value of a chromosome is evaluated by the Sharpe ratio, the modularity of the grouping structure, and the group balance. Extensive experiments on financial datasets from the New York Stock Exchange and Taiwan Stock Exchange demonstrate that QGTSPO achieves superior performance compared to the previous approach in terms of volatility reduction, crisis management, and consistency risk-adjusted returns, making it ideal in high-volatility markets and for investors with stringent risk management goals.

1. Introduction

The investment management field has been burdened with generating consistent good returns and, at the same time, sheltering investors from adverse conditions in the financial markets. A common strategy to accomplish this is to follow the principle "Don't put all your eggs in one basket". This implies that an investor should invest in a variety of assets, which is also referred to as diversifying a portfolio. To address this issue, the seminal work of Markowitz (Markowitz, 1952) laid the foundation for diversification with the Modern Portfolio Theory (MPT).

MPT outlined that investors can be sheltered from adverse conditions by spreading their capital across different assets and how various factors influence assets, resulting in diverse behaviors regarding their returns and risks. As a result, this leads to consistent returns because when one asset performs poorly, the other asset compensates by performing well. Therefore, the challenge to the MPT is to select uncorrelated assets. This problem has been explored widely. Solutions have ranged from mathematical to evolutionary methods (Kumar et al., 2023; Asawa, 2022; Gunjan and Bhattacharyya, 2023). In particular, Chen et al. (2019) proposed the Group Trading Strategy Portfolio optimization (GTSPO) algorithm. In the GTSPO framework, a portfolio of trading strategies, Group Trading Strategy Portfolio (GTSP) is optimized while providing a way for investors to make choices regarding the portfolio constituents.

Although the previous method presented in Chen et al. (2019) provides a way to categorize trading strategies in groups, in their approach, there are a number of challenges that hinder the realization of a well-diversified portfolio. Firstly, the fitness evaluation process focuses on the return and maximum drawdown (MDD), not the relationships

☆ This is a modified and expanded version of the paper "A diverse group trading strategy portfolio optimization algorithm based on network modularity" published at the *IEEE International Conference on Systems, Man, and Cybernetics*, 2024, Malaysia.

* Corresponding author.
E-mail addresses: I112152111@nkust.edu.tw (K. Chideme), chench@nkust.edu.tw (C.-H. Chen).

https://doi.org/10.1016/j.engappai.2025.112021
Received 9 August 2024; Received in revised form 15 March 2025; Accepted 11 August 2025

among the trading strategies. Even though MDD is a well-used risk measure, it still fails to provide an assessment of how the assets in a portfolio mirror each other's performance. The proposed approach, the Q-Group Trading Strategy Portfolio Optimization (QGTSPO) approach, introduces a modularity component to enhance the diversification process of the generated portfolios. Additionally, the trading strategies utilized in the previous approach rely on a single stock series, meaning they all share the same underlying risk factors. Since their performance is tied to this single stock, there is little diversification benefit. Hence, in this paper, an extension to the GTSPO framework is made by enhancing the quality of the underlying trading strategies, with each trading strategy operating on a unique stock.

The QGTSPO draws inspiration from network theory, introducing a novel perspective on the GTSPO framework. In the proposed framework, trading strategies are conceptualized as nodes within a network, enabling a comprehensive analysis of their interactions and dynamics. In previous studies, network and clustering models have been used to group similar assets, like the work in Ioannidis et al. (2023), Giudici et al. (2020). Measures like correlation have been used to gauge the dependencies between assets. Portfolio building using network models is done by choosing assets with minimal connections, whereas in clustering models, assets are selected from different clusters. The Hierarchical Risk Parity (HRP) model, developed by Lopez de Prado (2016), is one of the most widely used clustering approaches in portfolio optimization. By leveraging machine learning techniques, they demonstrated that grouping assets leads to more diversified portfolios. In contrast to these machine learning approaches, this paper utilizes evolutionary algorithms, specifically the Grouping Genetic Algorithm (GGA), with modularity-based fitness evaluation for portfolio optimization. Modularity developed by Newman and Girvan (2004) is a well-used measure to gauge how well a network is clustered. Several other works have also looked at network partitioning methods (Fortunato, 2010; Vehlow et al., 2013; Raghavan et al., 2007; Šubelj and Bajec, 2012; Lou et al., 2013). However, the modularity (Newman and Girvan, 2004) has remained a widely adopted metric due to its effectiveness in assessing network partitioning quality and its use in the proposed approach.

Integrating the ideas in Newman and Girvan (2004) and the GTSPO approach (Chen et al., 2019), this paper proposed an approach, namely the Q-Group Trading Strategy Portfolio Optimization approach (QGTSPO) to obtain a QGTSP. The QGTSPO approach views the trading strategies as nodes in a graph. The GGA is then used to group similar nodes. QGTSPO leverages modularity as a fitness metric to evaluate the co-dependencies among trading strategies, with co-dependencies modeled using correlation and mutual information. Unlike in the previous approach (Chen et al., 2019), the QGTSPO develops a closer look at the relationship among trading strategies. As a result, the proposed algorithm offers more diversified trading strategy portfolios than the previous approach. Additionally, a threshold component is integrated into the chromosome structure to determine the optimal correlation/mutual information value at which two trading strategies are related, thereby establishing an edge between them. In other words, the difference between the proposed approach and the previous approach (Chen et al., 2019) is that the proposed approach incorporates both the Sharpe ratio and modularity into the fitness function, particularly leveraging modularity in the network representation of trading strategies, to obtain a QGTSP. The proposed approach also elevates the quality of the underlying trading strategies by ensuring that each trading strategy operates on a distinct stock series, thereby introducing more diversification to the generated portfolios. Datasets from the New York Stock Exchange (NYSE) and the Taiwan Stock Exchange (TWSE) have been used to assess the effectiveness of the proposed approach. The results show superior diversification by QGTSPO and optimal capital allocation within the portfolio, ultimately leading to consistent returns during the period under testing. Hence, the four contributions of the proposed QGTSPO approach are listed as follows:

- **Utilizing Modularity to Enhance Diversification of QGTSP**: Our strategy involves optimizing the grouping of assets to foster stronger relationships within each group. The GGA is used to maximize modularity. Ensuring that assets within the same group exhibit high interconnectivity enhances the robustness of generated QGTSP. This leads to enhanced portfolio performance stability under different market conditions.
- **Leveraging Multiple Financial Datasets to Enhance Risk Aversion Ability of Used Trading Strategies**: We harness the potential of diverse financial datasets as inputs to provide ample risk aversion opportunities for our trading strategies. By focusing each trading strategy on a particular stock dataset, we provide differentiated risk factors to the trading strategies, enabling more robust portfolios compared to the previous approach.
- **Optimizing the Threshold Value for the Association Matrix to Form Effective Trading Strategies Network**: Within QGTSPO, we introduce a component for threshold value in chromosome encoding. This feature aids in identifying the optimal correlation or mutual information values for constructing the association matrix of the trading strategies. Thus, using QGTSPO, the optimized threshold value is obtained and employed to calculate modularity, which is then used to find the effective trading strategies network.
- **Utilizing QGTSP for Users to Keep Crisis Period Performance Stable and Risk Management**: During crisis periods, such as 2008, QGTSPO shows superior drawdown protection in markets like the New York Stock Exchange and Taiwan Stock Exchange, offering more stable recovery, reduced volatility, and better risk-adjusted metrics than the previous method. Experimental results indicate that QGTSPO outperforms the previous method in achieving favorable risk-adjusted returns, with notable improvements in risk management and crisis resilience. The QGTSPO exhibits stable volatility patterns across various market conditions, which is crucial for investors with strict risk controls.

The structure of the paper unfolds as follows: In Section 2, we offer background knowledge and an overview of prior research on portfolio optimization. Section 3 elaborates on the proposed framework. Subsequently, Section 4 presents and discusses the experimental results. Finally, Section 5 outlines the study's conclusions and suggests future work directions.

2. Background knowledge & related work

In this section, the background knowledge underpinning this paper is introduced in Section 2.1. After that, the related work on portfolio management is discussed in Section 2.2.

2.1. Background knowledge

2.1.1. Group genetic algorithms

The grouping genetic algorithm (GGA) was introduced by Falkenauer (1998) as an extension of the classical genetic algorithms to solve grouping problems. In grouping problems, there is a set of instances U to be grouped into a collection of disjoints sets such that :

$\cup U_i = U$ and $U_i \cap U_j = \emptyset$, $i \neq j$.

When used for grouping problems, the classical genetic algorithm uses the objects to perform genetic operations. However, this may lead to offspring chromosomes with no meaningful structure. For example, assume we are tasked with grouping objects into groups. Each letter represents an object. In chromosome C_1 below, for example, object A belongs to group 1, B to group 2, and C to group 3:

$$\begin{aligned} C_1 &= A \quad B \quad C \quad B \quad B \quad C \\ C_2 &= M \quad N \quad O \quad N \quad O \quad M \end{aligned}$$

After crossover, the following chromosomes are generated:

$$\begin{array}{lcllllll} C_2|C_1 & = & M & N & O| & B & B & C \\ C_1|C_2 & = & A & B & C| & N & O & M \end{array}$$

The resulting offspring will contain many illegal groups if some constraints exist on the groups. GGA solves the inconsistencies by adding a grouping element to the chromosome structure. Genetic operations of crossover and mutation are then applied to the grouping element. The chromosome structure in GGA is as shown below:

$$\begin{array}{lcllllll} C_1 & = & A & B & C & B & B & C : ABC \\ C_2 & = & M & N & O & N & O & M : MNO \end{array}$$

2.1.2. Community detection

Newman (2006) introduced modularity, a measure of how well a network is clustered into modules. For a network with m edges, the expected number of edges connecting two nodes i and j, each with degrees k_i and k_j, is calculated as $k_i * k_j$ divided by twice the total number of edges, represented by $2m$. That is, the expected number of edges is given by $\frac{k_i * k_j}{2m}$. The difference between the actual number of edges and the expected number of edges is $A_{ij} - \frac{k_i * k_j}{2m}$. This is summed on all nodes, giving us the value of modularity:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(G_i, G_j) \tag{1}$$

The network partition algorithm introduced by Newman and Girvan (2004) sought to maximize the value of modularity. The intuition behind maximizing modularity is that a good network partitioning or grouping should have nodes in the same group that are more connected than others. That is, nodes in the same group have more connections (edges) and are related.

2.2. Related work

This section provides an overview of the relevant literature on evolutionary-based portfolio optimization and graph-based portfolio optimization approaches, focusing on recent advancements in methodologies and their applications.

2.2.1. Evolutionary-based portfolio optimization

In this section, studies that explore evolutionary algorithms for portfolio optimization are reviewed and discussed. This review also examines the different fitness evaluation measures, different chromosome encoding structures, and varying asset classes adopted to enhance portfolio performance under varying market conditions by these studies. A summary of the literature on evolutionary-based portfolio optimization is shown in Table 1.

For stock portfolio optimization, Chen et al. proposed approaches for finding a Group Stock Portfolio (GSP) using GGA (Chen et al., 2020, 2023). A GSP consists of many stock groups, and stocks in the same stock group have similar properties. After finding GSP, various stock portfolios can be generated by selecting stocks from stock groups according to the user's preference. Next, for trading strategy portfolio optimization, Chen et al. introduced the GTSPO framework (Chen et al., 2019, 2018), wherein the authors devised a mechanism for suggesting portfolios of trading strategies to investors based on portfolio constituent preferences, employing GGA for optimization. Both approaches used MDD as a risk measure and incorporated it into the fitness function to generate portfolios with little downside risk. While optimizing for MDD can help reduce downside risk, portfolios may still experience losses because the same underlying factors could influence the assets within the portfolio and behave the same way under the same market conditions. A good portfolio has constituents that behave differently under the same conditions. To this end, this paper introduces the QGTSPO approach. QGTSPO models the trading strategies as nodes in a graph and then uses modularity to group similar nodes in the same group using GGA. This differs from the path taken in the previous approaches (Chen et al., 2019, 2018), where GGA was implemented but without catering to the relationships between the trading strategies.

Guarino et al. (2024) proposed EvoFolio, a portfolio optimization system based on the Non-Dominated Sorting Genetic Algorithm II (NSGAII), which maximizes return, minimizes risk, and integrates user preferences, allowing investors to influence stock selection, making it more interactive and adaptable to individual risk tolerances. However, while EvoFolio effectively incorporates user preferences as in the GTSPO framework (Chen et al., 2019), it does not explicitly model asset relationships.

Barroso et al. (2021), Lou (2023), and Bedoui et al. (2023) also further demonstrated the effectiveness of NSGAII in balancing the trade-offs between risk and return. For instance, Barroso et al. (2021) successfully applied it to the Brazilian stock market utilizing traditional metrics, expected return, and variance as bi-objectives, with the fitness evaluation based on the mean–variance framework proposed by Markowitz (1952).

Various fitness measure metrics have been explored in evolutionary based portfolio optimization. The portfolio optimization conducted by Bedoui et al. (2023) employed Conditional Value at Risk (CVaR) as the fitness measure. CVaR has also been employed by Kumar et al. (2023) together with data envelopment analysis under a credibilistic framework to handle the fuzziness of asset returns. Abbas and Raza (2023) optimized the Sharpe ratio while incorporating floor and ceiling constraints to manage asset allocations. Lim et al. (2020) enhanced the Sharpe ratio by incorporating a market valuation component based on the Capital Asset Pricing Model (CAPM) into the fitness function.

Different chromosome architectures have been utilized for portfolio optimization. In Wang et al.'s approach (Wang et al., 2023), the chromosome encoding consisted of two components: the weight and cardinality parts. Pal et al. (2021) used a clustering variable-length non-dominated sorting genetic algorithm to obtain the Pareto-optimal portfolios. They then employed a single-objective GA-based approach to optimize portfolio weights. While Pal et al. (2021) aimed to achieve clustering using genetic algorithms, it ultimately employed two algorithms: clustering and weight optimization, which may result in potentially longer processing times. Solares et al. (2022) implemented a genetic algorithm for portfolio optimization, incorporating an additional mechanism to eliminate undesirable stocks from the portfolio.

In addition to different architectures, hybrid approaches have also been used to bolster genetic algorithms' performance in portfolio optimization. A study by Yaman and Dalkılıç (2021) introduced a hybrid approach to cardinality constraint portfolio optimization by integrating a nonlinear neural network and GA. Lv et al. (2024) presented a hybrid approach to portfolio management that integrates deep learning-based stock selection with a multi-objective optimization framework. Their approach leveraged convolutional neural networks (CNNs) and bi-directional recurrent neural networks (Bi-RNNs) to predict stock trends and an improved NSGAIII to address the multi-objective constraints efficiently. Banerjee et al. (2024) combined the fuzzy-TODIM multi-criterion decision-making model with GA to optimize portfolio allocation. By integrating Markowitz's portfolio parameters into the fuzzy-TODIM framework, the study ranks stocks across three markets before dynamically generating portfolios using multiple weighting techniques and optimization constraints.

Other nature-inspired algorithms have also been applied to the portfolio optimization problem. Chou et al. (2022) introduced a quantum-inspired tabu search algorithm that incorporated novel performance metrics, including an emotion index and trend ratio, to enhance portfolio evaluation beyond the traditional Sharpe ratio. Similarly, Hasan et al. (2022) employed the whale optimization algorithm, optimizing both return and risk as objectives. While these methods introduced innovative techniques to improve portfolio performance, they, like others, did not explicitly address the interrelationships between assets,

Table 1
Evolutionary-based portfolio optimization studies.

Author	Type of Portfolio	Algorithm	Objective	Fitness evaluation	Input	Modularity
Abbas and Raza (2023)	Stock Portfolio	Genetic Algorithm & PSO	Single	Sharpe Ratio	Multiple Stock Series	No
Banerjee et al. (2024)	Stock Portfolio	Genetic Algorithm	Multiple	Return, Risk	Multiple Stock Series	No
Barroso et al. (2021)	Stock Portfolio	NSGAII	Single	CVaR	Multiple Stock Series	No
Bedoui et al. (2023)	Stock Portfolio	Genetic Algorithm	Multiple	CVaR	Multiple Stock Series	No
Chen et al. (2019, 2018)	Trading Strategy Portfolio	Grouping Genetic Algorithm	Single	Return, MDD	Single Stock Series	No
Chen et al. (2020, 2023)	Stock Portfolio	Grouping Genetic Algorithm	Single	Return, MDD	Multiple Stock Series	No
Chou et al. (2022)	Stock Portfolio	Quantum-inspired tabu search	Single	Trend Ratio, Emotion Index	Multiple Stock Series	No
Guarino et al. (2024)	Stock Portfolio	Whale Optimization Algorithm	Single	Return, Risk	Multiple Stock Series	No
Hasan et al. (2022)	Stock Portfolio	Whale Optimization Algorithm	Single	Return, Risk	Multiple Stock Series	No
Lim et al. (2020)	Stock Portfolio	Genetic Algorithm	Single	Return, Risk, SML	Multiple Stock Series	No
Lou (2023)	Stock Portfolio	NSGAII	Multiple	Returns	Multiple Stock Series	No
Lv et al. (2024)	Stock Portfolio	NSGAIII	Multiple	Return, Risk	Multiple Stock Series	No
Pal et al. (2021)	Stock Portfolio	NSGAII	Multiple	Returns, Risk	Multiple Stock Series	No
Solares et al. (2022)	Stock Portfolio	Genetic Algorithm	Single	Returns	Multiple Stock Series, Fundamental Indicators	No
Wang et al. (2023)	Stock Portfolio	Genetic Algorithm	Single	Returns	Multiple Stock Series	No
Yaman and Dalkılıç (2021)	Stock Portfolio	Genetic Algorithm	Single	Returns	Multiple Stock Series	No

a key factor in portfolio optimization that is addressed by the QGTSPO framework.

The reviewed literature highlights a wide range of approaches to portfolio optimization using evolutionary algorithms and nature-inspired techniques. While many studies, such as those by Barroso et al. (2021), Abbas and Raza (2023), and Wang et al. (2023), have focused on enhancing traditional metrics like return, risk, and the Sharpe ratio, they often overlook the structural relationships and inter-dependencies between assets. Approaches such as those by Chou et al. (2022), Lim et al. (2020), and Solares et al. (2022) have added unique components, such as market valuation and mechanisms for eliminating undesirable assets, to refine portfolio construction further. However, these approaches do not incorporate asset clustering techniques, which could provide a deeper understanding of asset dependencies and enhance diversification.

2.2.2. Graph-based portfolio optimization

To the best of our knowledge, no previous work has modeled trading strategy portfolios using a graph-based approach combined with evolutionary methods. Although not specifically applied to trading strategies, other works have explored machine learning and mathematical techniques to model portfolios as graphs for various asset classes.

Li et al. (2019) proposed an innovative approach to portfolio construction by leveraging network theory. They identified peripheral nodes within a network derived from the cross-correlation matrix of asset returns. Clemente et al. (2021) utilized the clustering coefficient to analyze the connectivity of assets within an asset correlation network. Ricca and Scozzari (2024) proposed an integer linear programming approach to generate "disassortative" portfolios by leveraging the "local degree" and "local strength" assortativity coefficients within a market graph. The weighted network was constructed from a correlation matrix, where edges represented the Pearson correlation coefficient.

Many network theory-based portfolio management studies have focused on identifying and modeling asset relationships, often using techniques such as clustering coefficients or correlation-based edges. However, these studies typically require an additional step to optimize weight allocation for actionable portfolio construction. This study addresses this gap by concurrently optimizing the identification of dissimilar nodes and the allocation of portfolio weights, streamlining the process and enhancing the effectiveness of portfolio management.

The QGTSPO framework offers a novel solution by addressing key gaps in previous approaches, optimizing portfolios based on asset relationships, and grouping similar strategies through modularity. This enables a more comprehensive optimization process, potentially leading to better-diversified and more resilient portfolios in complex market conditions. While prior research also employed GGA (Chen et al., 2019), they did not effectively address grouping similar assets. Other previous studies have made valuable contributions to portfolio optimization, and integrating asset inter-dependencies within QGTSPO represents a significant advancement in enhancing the effectiveness and robustness of portfolio construction strategies. Moreover, unlike traditional approaches that focus on portfolios of assets like stocks, the QGTSPO framework centers on portfolios of trading strategies.

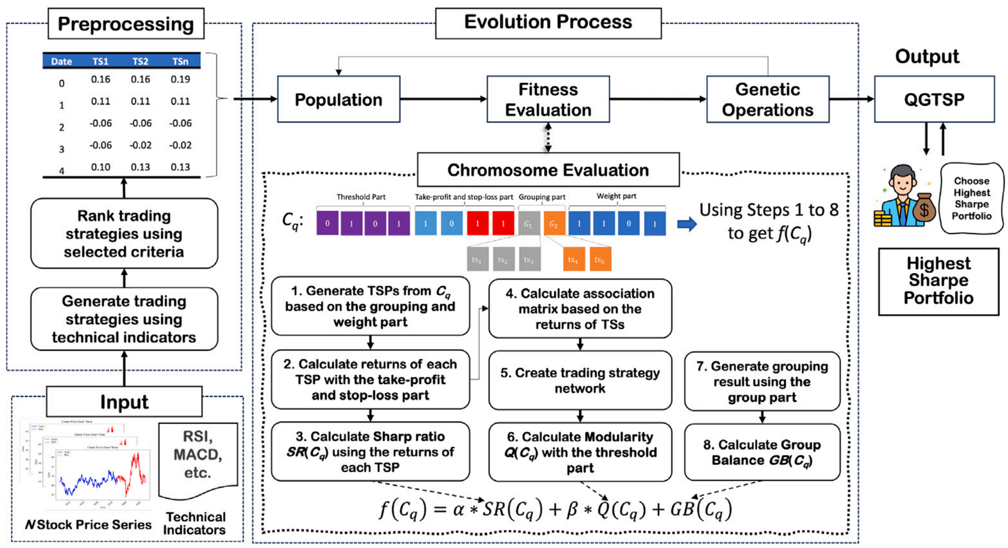

Fig. 1. Proposed QGTSPO framework.

3. Proposed approach

In this section, the details of the proposed approach are described. The framework of the proposed approach is given in Section 3.1. We then explore the details of the proposed approach in Section 3.2.

3.1. Framework

The proposed QGTSPO framework for the proposed approach is illustrated in Fig. 1, which provides an overview of the methodology. This diagram outlines the sequential and interconnected steps involved, beginning with the data pre-processing phase, followed by the core of the algorithmic components, and finally, to obtain a QGTSP for providing trading strategy portfolios to investors.

As shown in Fig. 1, a collection of time series data for N stock series data is gathered alongside a diverse set of technical indicators. Subsequently, N trading strategies are devised based on these indicators, and each is matched to a specific asset. Those exhibiting favorable returns, liquidity, and risk profiles are selected. These chosen strategies are then used to form a chromosome, the structure of which is elaborated in Section 3.2.2, including the threshold, the take-profit and stop-loss, the grouping, and the weight parts. The trading strategy portfolios are first generated according to the grouping and weight parts in the fitness evaluation of every chromosome. Then, their returns are calculated using the take-profit and stop-loss part to get the Sharpe ratio. Next, the returns of trading strategies are used to generate an association matrix to create the trading strategy network. The modularity of a chromosome is evaluated by utilizing the threshold part and the trading strategy network. Lastly, the three factors, Sharpe ratio, modularity, and group balance, are used to calculate fitness value. Genetic operators are executed to generate new offspring. After the evolution process, the QGTSP with the highest fitness value is given as output to provide Trading Strategy Portfolios (TSPs) to investors.

3.2. Components of framework

3.2.1. Data pre-processing

To enhance the diversity among trading strategies, data pre -processing employed the following process to generate strategies for portfolio optimization. Initially, it creates trading strategies utilizing various technical indicators, which simply match buying indicators with selling indicators to generate trading strategies. Subsequently, to decouple the underlying risk of each strategy from the other strategies, each strategy with a unique stock series dataset is paired. These strategies are then evaluated based on criteria such as trading frequency, MDD, and profit. Finally, the top q performing strategies are chosen for inclusion in the evolution process. Table 2 below shows an example of 25 strategies that are selected in the pre-processing stage.

Algorithm 1 outlines the data processing steps. As highlighted before, buying and selling technical indicators are taken as inputs together with multiple stock series data. In lines 1 to 7, trading strategies are generated by matching a buying indicator with a selling indicator. Following this, the performance of the generated strategies is calculated based on return, Sharpe ratio, and trading frequency (lines 8 to 13). From these metrics, the strategies are ranked, and the top N strategies are selected as the output (lines 14 to 15).

Algorithm 1 Data Pre-processing.

Input: M stock series sp, $\sqrt{M}$ Buying Indicators bi, $\sqrt{M}$ Selling Indicators si
Output: top N performing trading strategies $topTSs$.

1: $TSs \leftarrow \emptyset$
2: **for** b in bi **do**
3: **for** s in si **do**
4: strategy ← generateStrategy(b, s);
5: $TSs \leftarrow TSs \cup strategy$
6: **end for**
7: **end for**
8: **for** t in TSs **do**
9: **for** p in sp **do**
10: $strategyPerformance$ ← calculatePerformance(t, p);
11: t ← updatePerformance($strategyPerformance$);
12: **end for**
13: **end for**
14: $topTSs$ ← rankStrategies(TSs);
15: return $topTSs$;

3.2.2. Chromosome encoding

The chromosome encoding consists of four parts: the threshold part, the take-profit and stop-loss part, the grouping part, and the weight part, as depicted in Fig. 2.

Firstly, there is the node connection threshold, which consists of h bits. It sets the threshold value determining whether an edge should be formed between two trading strategies. The range of values in the threshold section depends on the association matrix parameter

Table 2
Trading strategies.

ID	Buying Indicator	Selling indicator	ID	Buying indicator	Selling indicator
1	5EMA > 20EMA	5EMA < 20EMA	14	WILLR < 80	MOM ≤ 0
2	5EMA > 20EMA	RSI < 70	15	WILLR < 80	CCI ≤ 100
3	5EMA > 20EMA	WILLR > 20	16	MOM > 0	5EMA < 20EMA
4	5EMA > 20EMA	MOM ≤ 0	17	MOM > 0	RSI < 70
5	5EMA > 20EMA	CCI ≤ 100	18	MOM > 0	WILLR > 20
6	RSI > 30	5EMA < 20EMA	19	MOM > 0	MOM ≤ 0
7	RSI > 30	RSI < 70	20	MOM > 0	CCI ≤ 100
8	RSI > 30	WILLR > 20	21	CCI > 100	5EMA < 20EMA
9	RSI > 30	MOM ≤ 0	22	CCI > 100	RSI < 70
10	RSI > 30	CCI ≤ 100	23	CCI > 100	WILLR > 20
11	WILLR < 80	5EMA < 20EMA	24	CCI > 100	MOM ≤ 0
12	WILLR < 80	RSI < 70	25	CCI > 100	CCI ≤ 100
13	WILLR < 80	WILLR > 20			

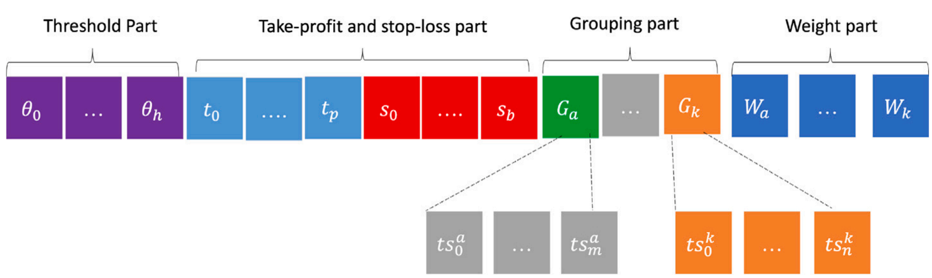

Fig. 2. Chromosome encoding schema.

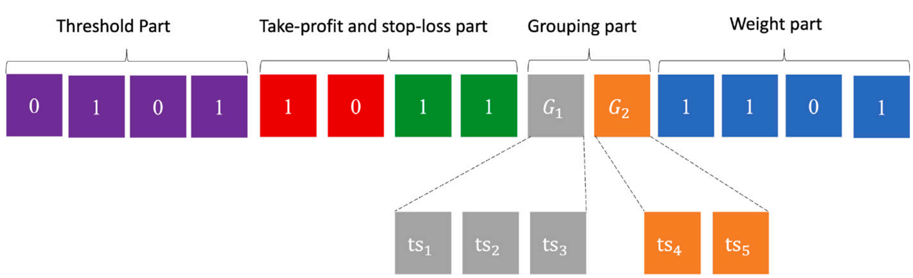

Fig. 3. An example of a chromosome.

provided to GGA. If the association matrix is correlation, the value range is from −1 to 1, and if it is mutual information, the value range is 0 to infinity.

Secondly, there is the take-profit and stop-loss segment, with the number of bits for take-profit being p and for stop-loss being b. These segments govern the exit points for trading strategies, specifying the return level at which a strategy terminates a buying position.

Thirdly, there is the grouping section, which categorizes trading strategies into specific groups. Lastly, the chromosome includes the weight component, which is influenced by the number of groups within the chromosome. This weight component facilitates the distribution of capital among the various groups. An example of a chromosome is shown in Fig. 3.

In Fig. 3, the threshold part $[0, 1, 0, 1]$ represents 0.028 if the association matrix is mutual information and 0.3125 if the association matrix is correlation. The take-profit and stop-loss part has two bits for each part. The first two bits [1,0] represent a take-profit threshold of 6.7%. The bits [1,1] represent a stop-loss threshold of 10%. The chromosome depicts two groups G_1 and G_2. Lastly, the weight part depicts 25% capital allocation to group G_1 and 75% to group G_2. There are three and two trading strategies in groups G_1 and G_2.

Following evolution, several portfolios are crafted by way of combinations of an element (trading strategy) from each group. The portfolio with the highest Sharpe ratio is then selected as the final output. An example is shown in Table 3.

In Table 3, there are two groups, which signifies portfolios consisting of two assets. Group G_1 has three assets, and G_2 has two assets;

Table 3
An example of QGTSPO output.

	Portfolio	Profit	MDD	StdDev	Sharpe ratio
1	ts_1, ts_4	0.72	−0.15	0.079	7.12
2	ts_1, ts_5	0.28	−0.17	0.065	4.01
3	ts_2, ts_4	0.34	−0.18	0.070	4.42
4	ts_2, ts_5	0.36	−0.20	0.072	4.68
5	ts_3, ts_4	0.29	−0.19	0.069	4.01
6	ts_3, ts_5	0.31	−0.21	0.071	4.91

hence, the total number of portfolios is 6 (= 3 x 2). Portfolio 1 has the highest Sharpe ratio of 7.12; hence, it was selected as the final output.

3.2.3. Fitness function

The QGTSPO aims to provide a well-diversified final portfolio with risk-adjusted returns. To get the desired outcome, we formulated an objective function that includes the Sharpe ratio, modularity, and group balance. The Sharpe ratio measures risk-adjusted returns, computed as the excess return per unit of risk, often represented by the portfolio's average return minus the risk-free rate divided by its standard deviation. Modularity, as defined by Newman (2006), is a measure of how effectively a network is partitioned into modules or communities, where nodes within the same module exhibit stronger connections among themselves compared to nodes in different modules. The last part is the grouping balance, which ensures that trading strategies are not allocated to a single group, leaving other groups with very few strategies. Sharpe ratio and modularity complement each other, so they are incorporated into a single metric, α, for Sharpe ratio and β for modularity, respectively. Eq. (2) provides the formula for evaluating the fitness value of each chromosome:

$$f\left(C_q\right) = \alpha * SR\left(C_q\right) + \beta * Q\left(C_q\right) + GB\left(C_q\right), \tag{2}$$

where $SR\left(C_q\right)$ is the Sharpe ratio of a GTSPO in chromosome C_q. $Q\left(C_q\right)$ is the modularity, and $GB\left(C_q\right)$ is the group balance. The formula for $SR\left(C_q\right)$ is given below in Eq. (3):

$$SR\left(C_q\right) = \frac{\sum_{j=1}^{nTSP} SR_t\left(TSP_j\right)}{nTSP}, \tag{3}$$

where $nTSP$ is the number of trading strategy portfolios generated and $SR_t\left(TSP_j\right)$ is given by equation below (4):

$$SR_t\left(TSP_j\right) = \frac{PReturn\left(TSP_j\right) - RFR}{PStdDev\left(TSP_j\right)}, \tag{4}$$

where $PReturn(TSP_j)$ is portfolio return, $PStdDev(TSP_j)$ is portfolio standard deviation and RFR is the risk free rate.

The modularity of the grouping structure in chromosome C_q is denoted as $Q\left(C_q\right)$, following Newman's formulation (2006) (Newman, 2006), represented by Eq. (5).

$$Q\left(C_q\right) = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right)\delta(G_i, G_j), \tag{5}$$

where A_{ij} is the adjacency matrix denoting the connections between trading strategies within the network. Each trading strategy corresponds to a node. There are two ways that can be used to construct the adjacency matrix. In the first approach, an edge exists between two nodes if their correlation surpasses a predefined threshold θ. In the second approach, it utilizes mutual information to model the dependencies. The degrees of nodes i and j are denoted as k_i and k_j, respectively, while m signifies the total number of edges in the network. The Kronecker delta function $\delta(G_i, G_j)$ is utilized to determine whether nodes i and j belong to the same community ($G_i = G_j$). When G_i and G_j are equal, the function returns 1; otherwise, it returns 0. Essentially, it signifies whether ts_i and ts_j are similar.

The proposed approach treats each trading strategy as a node within a network. An edge linking two distinct nodes ts_1 and ts_2 is established when the value in the association matrix exceeds or equals a specified threshold θ. The threshold θ is, in turn, determined by the threshold part encoded in the chromosome. The association matrix in this paper is created using two approaches: correlation and mutual information. This framework yields an undirected network with a modularity value ranging from −0.5 to 1. Consequently, each chromosome represents a potential solution for grouping the network of trading strategies. Subsequently, the modularity metric is employed to evaluate the effectiveness of each chromosome's grouping in organizing similar trading strategies into distinct groups.

The group balance $GB\left(C_q\right)$ tries to steer the solution towards a chromosome with an even number of trading strategies across the groups and is given by the formula below in Eq. (6):

$$GB(C_q) = \sum_{i=1}^{K} -\frac{|G_i|}{N} log\frac{|G_i|}{N}, \tag{6}$$

where K is number of groups, $|G_i|$ is number of trading strategies in group G_i, and N is number of trading strategies.

3.2.4. Genetic operations

The genetic operations are tailored to steer the evolution process towards an optimal solution. As described in Section 3.2.2, the chromosome is encoded into four parts. Each part optimizes a certain part of the QGTSPO. The first part is the node threshold part, which optimizes the value at which an edge is created between two nodes. This part is coded in binary. A one-point crossover is applied to this part. The mutation operation works by flipping the bit. Next, the take-profit and stop-loss part is also coded in binary. Thirdly, the grouping part, which is described in Section 3.2.2, puts similar trading strategies into the same groups. The mutation operation for this part works by moving a trading strategy from one group to another. Finally, the elite selection strategy chooses the top M chromosomes from the population for the subsequent generation.

3.3. Pseudocode of the proposed approach

In this section, we outline the pseudocode of the proposed approach. Firstly, N stock series data and technical indicators are taken as input. Algorithm 2 outlines the steps for the optimization process. N trading strategies are devised, matching with a single stock series as shown in line 1. pop_size chromosomes are formed using the trading strategies in line 2. The chromosome encoding is shown in Fig. 2. Lines 3–10 indicate the evolution process, with the optimized QGTSPO given by lines 11 and 12.

Algorithm 2 Pseudocode of the Proposed Algorithm (QGTSPO).

Input: N stock series data $spdata$, and technical indicators TI.
Parameters: Population size pop_size, Number of groups K, number of stop loss bits n, number of take profit bits b, take profit level $take\_profit$, stop loss level $stop\_loss$, number of iterations num_iter, fitness weights α, β, cross over rate r_c, mutation rate r_m, and inversion rate r_i.
Output: Optimized QGTSP $globalBest$.

1: $candidateTSs \leftarrow$ generateTS($spdata, TI$);
2: $population \leftarrow$ init_population($pop\_size, K, candidateTSs, n, b, tp, sl$);
3: $population \leftarrow$ fitness_function($population, \alpha, \beta$);
4: **for** i in range(num_iter) **do**
5:     $population \leftarrow$ selection($population, pop\_size$);
6:     $population \leftarrow$ crossover($population, r_c$);
7:     $population \leftarrow$ mutation($population, r_m$);
8:     $population \leftarrow$ inversion($population, r_i$);
9:     $population \leftarrow$ fitness_function($population, \alpha, \beta$);
10: **end for**
11: $globalBest \leftarrow$ get_global_best($population$);
12: return $globalBest$;

In line 1, the function $generateTS$ adheres to the specifications outlined in Algorithm 1. The fitness evaluation, detailed in Algorithm 3, is presented in more detail. This algorithm assesses the quality of a population of trading strategy chromosomes, representing potential solutions to the Group Trading Strategy Portfolio optimization problem. The algorithm takes the population of chromosomes as input.

For each chromosome in the population, the algorithm extracts the trading strategies and the threshold value encoded within it. Subsequently, it computes the association matrix among these strategies and constructs an adjacency matrix A based on a threshold determined by the threshold part in the chromosome. In this paper, two methods are explored for determining the association matrix. The first method employs correlation, while the second method utilizes mutual information. The adjacency matrix is composed of ones and zeros, where a value of one indicates that the strategies are associated above the threshold value.

Next, the algorithm identifies groups within the chromosome and calculates the modularity Q of these groups using the adjacency matrix A. Additionally, it determines the group balance GB within the chromosome and computes the Sharpe ratio SR of the trading strategies. These metrics are then combined to calculate the fitness of the chromosome. After evaluating all chromosomes in the population, the algorithm updates each chromosome with its computed fitness value. Finally, the algorithm returns the evaluated population.

Algorithm 3 Fitness Evaluation.

Input: chromosome population $population_a$.
Output: Evaluated population $population_b$.

for $chromosome$ in $population$ **do**
  $trading\_strategies \leftarrow$ get_ts($chromosome$)
  $assoc\_matrix \leftarrow$ calc_association($trading\_strategies$);
  $threshold \leftarrow$ get_threshold($chromosome$)
  $A \leftarrow$ get_adj($assoc\_matrix, threshold$);
  $groups \leftarrow$ get_groups($chromosome$)
  $Q \leftarrow$ get_modularity($groups, A$);
  $GB \leftarrow$ get_group_balance($groups$);
  $SR \leftarrow$ get_sharpe($trading\_strategies$);
  $fitness \leftarrow SR + Q + GB$;
  $chromosome \leftarrow$ update($chromosome, fitness$);
end for
return $population$;

4. Experimental results

This section analyzes of the proposed framework and is organized as follows: Firstly, Section 4.1 describes the utilized datasets, and Section 4.2 elaborates on the choice of parameter settings used in the proposed algorithm. Sections 4.3 and 4.4 then give and discuss the portfolio performance results and grouping results, respectively. Finally, Section 4.5 assesses the impact of different components of the QGTSPO framework.

4.1. Data description

Data was collected from two different capital markets, namely the New York Stock Exchange (NYSE) and the Taiwan Stock Exchange (TWSE). In this case, we were able to investigate the performance of our proposed approach on a matured market like the New York Stock Exchange (NYSE) and an emerging market like the TWSE. The period 2007–2008 has been one of the latest periods in recent times to be characterized by uncertainty in the financial markets. As such, the period for validating the proposed approach was 2005 to 2010. In each market, stock price series were collected via Yahoo Finance. Furthermore, in each market, the data was processed as detailed in Section 3.2.1. The GTSPO only utilizes a single stock, so the Microsoft Corporation and the TSMC stock were selected for the NYSE and TWSE markets, respectively. Section 4.3 highlights the portfolio performance of the proposed approach in these markets. Table 4 presents the stocks chosen from each market.

Table 4
List of stock ticker symbols from NYSE and TWSE.

NYSE	TWSE
Microsoft Corporation	TSMC
JPMorgan Chase	Hon Hai Precision Industry Co.
Bank of America Corporation	Formosa Plastics
Citigroup Inc.	Formosa Chemicals & Fibre
Wells Fargo	Chunghwa Telecom
Goldman Sachs	Cathay Financial Holding
Morgan Stanley	Fubon Financial Holding
American International	Uni-President Enterprises
U.S. Bancorp	Taiwan Cement
BNY	Delta Electronics
PNC Financial Services	MediaTek
Johnson & Johnson	ASE Technology
Procter & Gamble Co.	China Development Financial
The Coca-Cola Company	Yuanta Financial Holding
Walmart Inc.	CTBC Financial Holding
McDonald's Corporation	Mega Financial Holding
Exxon Mobil Corporation	Evergreen Marine
General Electric	EVA Airways
IBM	Yang Ming Marine Transport
AT&T Inc.	Yulon Motor
United Technologies Corporation	Far Eastern New Century
Starbucks Corporation	United Microelectronics
Nike, Inc.	Asustek Computer
Lowe's Companies	Largan Precision
The TJX Companies	HTC Corporation
Marriott International	
Yum! Brands	

Table 5
Parameter values in the QGTSPO algorithm for crossover and mutation analysis.

Parameter	Value
Population Size	30
Number of generations	30
Selection Rate	0.5
Inversion Rate	0.2
Number of take profit bits	8
Number of stop loss bits	8
Association Matrix	Mutual Information

As formulated in Section 3.2.1, trading strategies are traded on each stock. This generates a time series of returns for each trading strategy.

4.2. Parameter settings

Setting the parameters for a genetic algorithm has been shown to affect performance (Deb et al., 1999; Deb and Deb, 2014), (Vlasov et al., 2021). The values for each parameter are usually dependent on the optimization problem. In the portfolio optimization problem tackled in this study, an analysis was conducted to determine the values for the key parameters for crossover, mutation, and the number of groups in GGA.

4.2.1. Crossover and mutation

The crossover and mutation operations within a genetic algorithm are pivotal parameters influencing its performance, with the former focusing on exploitation and the latter exploration. The objective was to ascertain the optimal values for these operations to attain superior results. Each combination of parameter values was run a total of 10 times. In our evaluation of this analysis, the settings of other crucial parameters as detailed in Table 5.

Fig. 4 shows the results collected and averaged for each combination of parameter values.

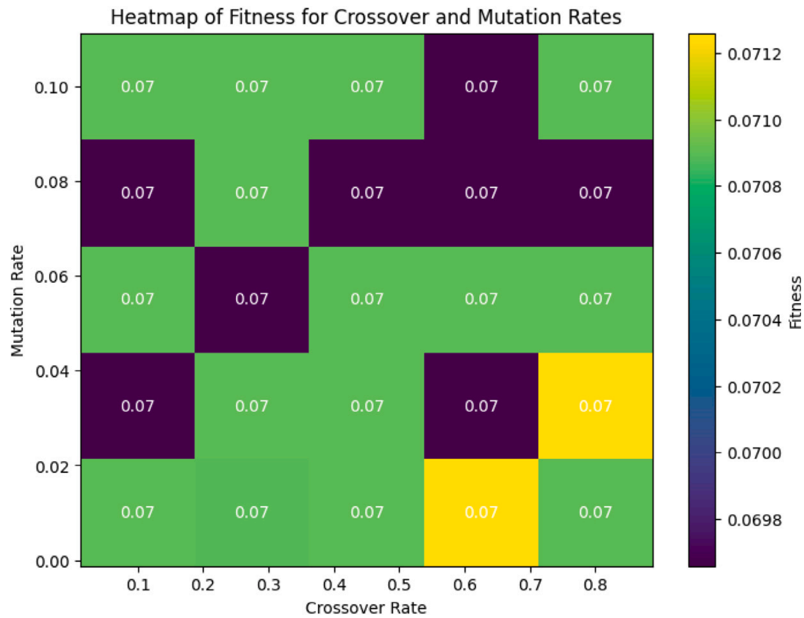

Fig. 4. Crossover and mutation.

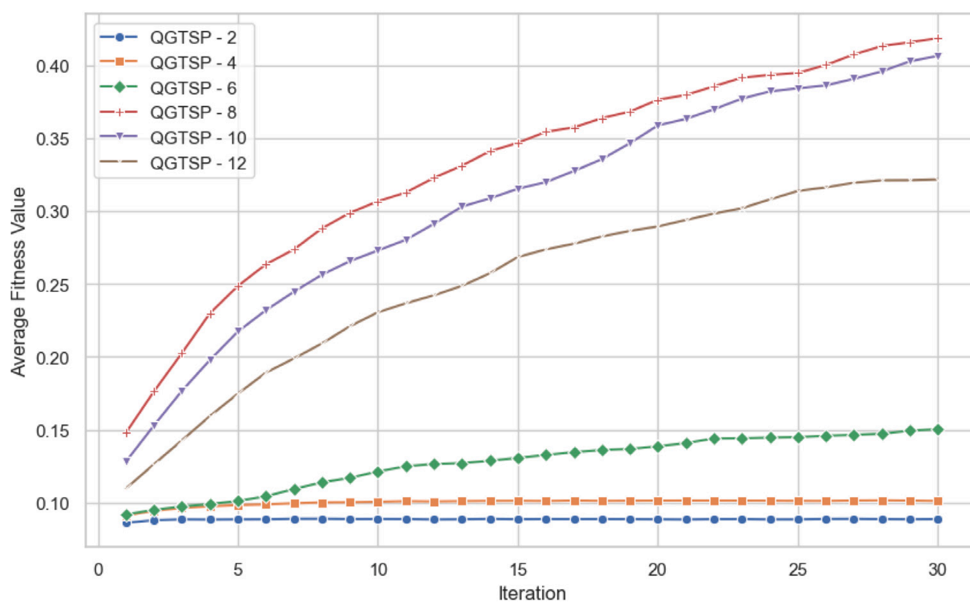

Fig. 5. Average fitness value of QGTSPO on different number of groups.

As depicted in Fig. 4, the investigation reveals that higher crossover values coupled with lower mutation values correspond to elevated fitness levels within the proposed QGTSPO algorithm. In our further analyses, we utilized values of 0.6 and 0.02 for crossover and mutation, respectively.

4.2.2. *Number of groups*

The number of groups is another crucial parameter in GGA. In the QGTSPO algorithm, various settings were explored for the number of groups, denoted by K. Each initialization was labeled QGTSPO-k to indicate the specific number of groups, K. The experiments were conducted by averaging the results of 10 runs for each initialization, and the results are depicted in Fig. 5.

In particular, as illustrated in Fig. 5, an initialization with eight groups clearly results in a higher fitness function for the dataset under consideration. In subsequent experiments, the QGTSPO algorithm implemented utilized this value.

4.2.3. *Fitness weight combinations*

Our proposed approach comprises three components in the fitness function. Section 3.2.3 delves into the direct link between two of these components, the Sharpe ratio and modularity, and their impact on portfolio performance. Consequently, these two components are combined into a single metric, weighted by α and β, respectively. Subsequently, an analysis was conducted to determine the optimal weighting combination for these two components with respect to the Sharpe ratio. This analysis was conducted under the specified parameters for each run, as outlined in Table 6. The average results obtained from 10 runs are presented in Fig. 6.

The analysis in Fig. 6 indicates that by allocating more weight to the Sharpe ratio, the portfolio's performance improves. Notably, when the Sharpe component's weight exceeds 50%, the portfolio's performance continues to improve. Furthermore, the optimal ratio of Sharpe = 0.8 and the optimal modularity of 0.2 were discovered.

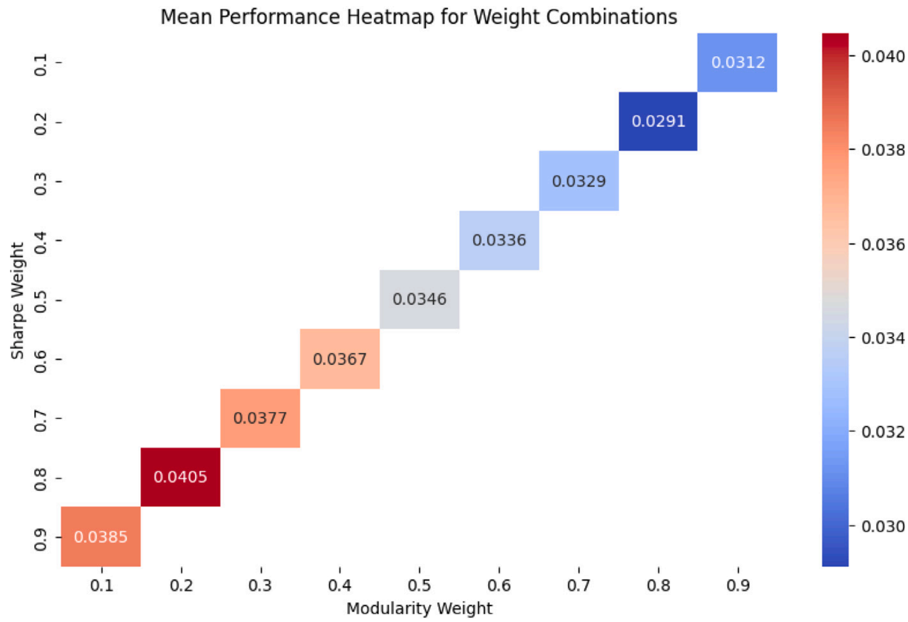

Fig. 6. Average sharpe ratio for weight combinations.

Table 6
Parameter values in the QGTSPO algorithm for weight combination analysis.

Parameter	Value
Population Size	30
Number of generations	30
Selection Rate	0.5
Inversion Rate	0.2
Number of take profit bits	8
Number of stop loss bits	8
Association Matrix	Correlation
Data	NYSE

4.3. Portfolio performance analysis

This study aimed to enhance the diversification of a portfolio of trading strategies. A series of experiments were conducted to compare the proposed QGTSPO approach to the existing approaches and evaluate the efficacy of the proposed approach. The QGTSPO framework has two different association matrix configurations: correlation-based and mutual information-based association. Hence, two QGTSPO models were presented, QGTSPO-SMC and QGTSPO-SMI, with the former being correlation-based and the latter mutual information-based.

In the comparative analysis of portfolio performance, six metrics were considered, including the Sharpe ratio, Sortino ratio, Calmar ratio, volatility (standard deviation), MDD, and cumulative return. The Sharpe ratio, a risk-adjusted return metric, calculates the excess return per unit of risk as the portfolio's average return minus the risk-free rate divided by its standard deviation. Unlike the Sharpe ratio, which considers total volatility, the Sortino ratio focuses solely on downside risk. It divides the returns below the investor's target returns by the volatility of those returns. Similarly, the Calmar ratio also considers downside risk by using the MDD as the denominator. Cumulative return provides a straightforward measure of overall portfolio performance by reflecting the total growth or decline in the portfolio value over a specified period. Volatility measures the extent of fluctuation in portfolio returns, indicating the level of risk or uncertainty. Lower volatility typically signifies a more stable and less risky portfolio, while higher volatility suggests heightened uncertainty and potential for larger price swings. MDD quantifies the maximum loss from a peak to a trough during a specific period, offering insights into the worst-case scenario for an investor. In the evolutionary process of QGTSPO, the portfolio with the highest Sharpe ratio is taken as the terminal portfolio for comparisons.

As mentioned earlier, in the previous approach (Chen et al., 2019), GTSPO, all trading strategies (portfolio constituents) traded the same stock. However, in the proposed method, QGTSPO, each trading strategy trades a unique stock, inadvertently introducing diversity. To ensure a fair comparison between GTSPO and QGTSPO, the trading strategy generation process in GTSPO was modified to guarantee that each trading strategy trades a distinct stock in the same way as the trading strategy process in QGTSPO. Hence, a modified model from the previous approach, GTSPO-M, was also given. This modification allows us to highlight the contributions of introducing unique stocks for each trading strategy. Additionally, because GTSPO-M and QGTSPO use the same data processing mechanism, the diversity introduced by the aforementioned mechanism is eliminated, leaving the analysis to focus on the effects of modularity on portfolio performance.

The GGA chromosome structure enables the GTSPO framework, which generates multiple portfolio suggestions for users. In contrast, other evolutionary methods are not inherently suited to the GTSPO problem. However, to broaden the analysis, we introduced Particle Swarm Optimization (PSO) despite a key limitation: PSO produces only a single portfolio, making direct comparisons with the GTSPO-based models (QGTSPO, GTSPO) challenging. Nevertheless, a comparison between PSO and the GTSPO-based models is conducted. In the PSO framework, trading strategies are developed the same way as the QGTPO outlined in Section 3.2.1, and the PSO model is used specifically for portfolio weight assignment. Thus, the comparison primarily evaluates the weight assignment methods of PSO versus the GTSPO-based models.

Upon determining the optimal parameters for K and the other genetic operators from the analysis done in Section 4.2, the QGTSPO framework was compared against the previous GTSPO approach (with the trading strategy generation process modification). A sliding window methodology involved ten runs for each data partition. After removing the best and worst results according to the evaluation metrics, the eight results were aggregated to ensure robustness. Each window had a duration of 2 years (777 days) for training, and the test period lasted for 328 days. A consistent random seed was established across all experiments to ensure a fair comparison in the stochastic processes (e.g., crossover, mutation rates) integral to genetic algorithms.

Table 7
Parameter values for portfolio performance analysis.

Parameter	Value
Population Size	30
Number of Generations	30
Selection Rate	0.5
Crossover Rate	0.6
Mutation Rate	0.02
Inversion Rate	0.2
Number of Take-Profit Bits	8
Number of Stop-Loss Bits	8
Sharpe Component Weighting	0.8
Modularity Component Weighting	0.2

Other parameter configurations for portfolio experiments are detailed in Table 7. Results comparing the proposed approach with the previous approaches are analyzed in Sections 4.3.1 and 4.3.2.

4.3.1. Portfolio performance: NYSE dataset

The initial dataset employed is sourced from the New York Stock Exchange (NYSE) spanning the years 2005 to 2010, a period encompassing both stable market conditions and the 2008 financial crisis. The randomly selected companies are presented in Table 4. As explained in the preceding sections, the experiments were conducted using a sliding window approach, with a training window of 777 days and a testing window of 328 days. This configuration resulted in four testing periods, whose outcomes are presented in Table 8.

In the year 2007, market conditions were highly volatile. The QGTSPO-SMC model achieved a positive Sharpe ratio (0.0058) and relatively stable returns (0.4137). While the modularity the modified previous approach model GTSPO-M (18.8985%) and PSO (2.3568%) posted better raw returns, their high volatility and larger MDD highlight the risk involved. The QGTSPO-SMC model's lower MDD (−5.4508%) and improved stability suggest the strength of modularity consideration in reducing downside risk. The better performance of multi-stock driven GTSPO-M over the single stock GTSPO model also highlight the advantage of introducing unique stocks for each trading strategy in the portfolio.

The following (2008) period was marked by the global financial crisis, leading to sharp declines across models. Despite all models experiencing negative returns, QGTSPO-SMI (−14.08%) limited its losses more effectively than GTSPO-M (−16.2396%) and PSO (−14.59%). Moreover, QGTSPO models exhibited lower drawdowns, with QGTSPO-SMI at −10.25%, compared to GTSPO at −12.93%, and PSO at −11.88%. This again suggests that QGTSPO models with modularity were better at mitigating extreme losses in market downturns. In the previous approach, GTSPO posted a better performance than all models, and this can be attributed to the trading strategies in the portfolio being linked to the same stock (Microsoft Inc.), which happened to be one of the stocks that did not suffer considerable losses in 2008.

Since market conditions in 2009 showed some recovery, the risk-adjusted performance became a key differentiator. The QGTSPO-SMI model (0.0373) and QGTSPO-SMC (0.0214) provided competitive Sharpe ratios compared to PSO (O.0012), and GTSPO (0.0134). GSTPO-M posted a high Sharpe ratio (0.619), however, the proposed framework QGTSPO exhibited better risk management through having the lowest volatility (QGTSPO-SMC model: 11.93%, QGTSPO-SMI: 12.80%) and MDD (QGTSPO-SMC model: −3.70%, QGTSPO-SMI: −3.47%) compared to the volatility given by previous approach (GTSPO: 57.66%, GTSPO-M: 17.66%). In addition, QGTSPO had high Sortino ratios (QGTSPO-SMI: 0.9132, QGTSPO-SMC: 0.5270), further indicating superior downside risk management.

By 2010, the QGTSPO model consistently demonstrated advantages in risk-adjusted returns. The QGTSPO-SMC model achieved a Sharpe ratio of 0.0248, outperforming GTSPO-M (0.0104), PSO (−0.0041), and the GTSPO model (−0.0041). Additionally, QGTSPO-SMC (0.5516) and QGTSPO-SMI (0.5205) exhibited strong Sortino ratios, reinforcing their ability to deliver better returns with lower risk. The Calmar ratio analysis further supported this, with QGTSPO-SMC (0.0412) and QGTSPO-SMI (0.0411) showing superior performance over all the other models presented, indicating greater capital efficiency and stability.

The QGTSPO models, QGTSPO-SCI and QGTSPO-SMC, show clear advantages in risk management while maintaining competitive returns. Their superior volatility, Calmar ratios, and controlled MDDs make them strong alternatives to the previous model, especially for risk-sensitive investment strategies. Although their raw returns might not always be the highest, their improved risk-adjusted performance positions them as excellent candidates for stability-focused portfolio management. The previous GSTPO model's performance exhibited very high volatility levels and even suffered losses in recovery periods, highlighting the better performance of the proposed approach. On average, the performance of QGTSPO is significantly better, as the results in Table 9.

From Table 9, while GTSPO-M achieved the highest return (7.15%) it also exhibited the highest standard deviation (21.41%), indicating greater exposure to risk. The GTSPO model on average suffered losses for the period 2007–2010 and also had very high volatility levels 20.22%during that period. In contrast, QGTSPO-SMI had a more stable performance, as reflected in its relatively low standard deviation (12.76%), suggesting better downside risk control compared to PSO (16.21%), GTSPO (21.41%), and QGTSPO-S (14.04%). Looking at MDD, QGTSPO-SMI had the smallest drawdown (−6.53%), reinforcing its stability compared to PSO (−9.30%), GTSPO (−8.69%), and GTSPO-M (−7.58%). Overall, GTSPO suffered the highest loss, and while the modified version, GTSPO-M, delivered the best performance in terms of absolute returns, QGTSPO proved to be the most stable model with the lowest volatility and drawdown, making them suitable choices for investors prioritizing risk minimization over high returns.

4.3.2. Portfolio performance: TWSE dataset

Emerging markets, such as the Taiwan Stock Exchange, may exhibit different characteristics compared to mature markets like the New York Stock Exchange. Therefore, an analysis was conducted to evaluate the performance of the proposed model in such a market. The results are presented in Table 10.

In 2007, the QGTSPO model demonstrated competitive risk-adjusted returns compared to other models. QGTSPO-SMI achieved a Sharpe ratio (0.0574) and Sortino ratio (1.3092), indicating strong returns with a better downside risk profile. Compared to GTSPO, GTSPO-M, and PSO, the QGTSPO exhibited a better balance between returns and volatility, showcasing their robustness in trading strategy portfolio optimization.

The market downturn during 2008 saw negative returns across most models presented. However, QGTSPO models with modularity mitigated the losses better than GTSPO and GTSPO-M models. In this period, the PSO has positive returns but still showed high volatility and MDD levels. QGTSPO managed to garner a better Sharpe ratio (QGTSPO-SMC: −0.0029, QGTSPO-SMI: −0.0041,GTSPO: −0.0876, GTSPO-M: −0.0325) which signifies superior risk management. Despite slightly negative Sharpe ratios, the QGTSPO models demonstrated more stability compared to the larger drawdowns experienced by other models.

During the 2009 financial market recovery phase, all the models presented achieved competitive results. In particular, the QGTSPO demonstrated its effectiveness in capturing good results both in a downtrend (2008) and an uptrend (2009), whereas the other models (PSO, GTSPO, GTSPO-M) exhibited varied performance behavior in these two distinct periods.

In 2010, the proposed model again displayed superior performance. QGTSPO-SMI recorded one of the highest Sharpe ratios (0.0726) and Sortino ratios (1.7224), indicating consistent gains with controlled risk exposure. The Calmar ratios of QGTSPO models, QGTSPO-SMC and

Table 8
Performance metrics for various models over different time windows for the NYSE dataset.

Period	Model	Sharpe ratio	Cumulative return (%)	Volatility (%)	MDD (%)	Sortino ratio	Calmar ratio
2007	PSO	0.0162	2.3568	13.5725	−7.6970	0.3831	0.0295
	GTSPO	−0.1228	−24.3801	15.5540	−12.5468	−2.4747	−0.1504
	GTSPO-M	**0.0418**	**18.8985**	27.8183	−7.9395	**0.9988**	**0.1723**
	QGTSPO-SMC	0.0058	0.4137	11.0269	−5.4508	0.1399	0.0179
	QGTSPO-SMI	−0.0049	−1.2338	**10.7136**	**−5.3272**	−0.1097	−0.0099
2008	PSO	−0.0470	−14.5909	18.5200	−11.8789	−1.0853	−0.0731
	GTSPO	**−0.0038**	**−4.7974**	25.1827	**−8.2198**	**−0.0921**	**−0.0112**
	GTSPO-M	−0.0363	−16.2396	30.2684	−12.9316	−0.7772	−0.0767
	QGTSPO-SMC	−0.0606	−14.9306	15.2776	−10.6778	−1.3631	−0.0864
	QGTSPO-SMI	−0.0579	−14.0843	**15.0993**	−10.2545	−1.3089	−0.0848
2009	PSO	0.0012	−0.6730	14.3508	−6.6787	0.0395	0.0044
	GTSPO	0.0134	−5.5194	57.6657	−14.9244	0.3349	0.0516
	GTSPO-M	**0.0619**	**14.3212**	17.6575	−4.5282	**1.5384**	**0.2410**
	QGTSPO-SMC	0.0214	3.1504	**11.9285**	−3.7045	0.5270	0.0724
	QGTSPO-SMI	0.0373	5.9628	12.7971	**−3.4724**	0.9132	0.1372
2010	PSO	−0.0041	−2.3103	20.3809	−12.6113	−0.0808	−0.0018
	GTSPO	−0.0041	−2.1328	13.5540	**−6.4029**	−0.0847	−0.0086
	GTSPO-M	0.0104	1.1345	**13.6853**	−7.1029	0.2480	0.0229
	QGTSPO-SMC	**0.0248**	**6.0852**	16.2660	−9.9616	0.5516	**0.0412**
	QGTSPO-SMI	0.0232	5.7605	14.6148	−8.6648	0.5205	0.0411

Table 9
Average values of performance metrics from 2007 to 2010 on NYSE Dataset.

Model	Sharpe ratio	Cumulative return (%)	Volatility (%)	MDD (%)	Sortino ratio	Calmar ratio
PSO	−0.0021	−2.1096	16.2113	−9.2983	−0.0390	−0.0000
GTSPO	−0.0268	−7.6842	20.2297	−8.6937	−0.5443	−0.0349
GTSPO-M	**0.0257**	**7.1542**	21.4057	−7.5845	**0.6421**	**0.1073**
QGTSPO-SMC	0.0004	−0.6899	13.3672	−7.2746	0.0196	0.0138
QGTSPO-SMI	0.0026	0.0265	**12.7634**	**−6.5297**	0.0726	0.0210

Table 10
Performance metrics of different models over various time windows for TWSE dataset.

Period	Model	Sharpe ratio	Cumulative return (%)	Volatility (%)	MDD (%)	Sortino ratio	Calmar ratio
2007	PSO	0.0453	10.6345	20.1208	−6.6016	1.0413	0.1363
	GTSPO	−0.0732	−13.1336	**13.6008**	−7.1834	−1.4679	−0.1346
	GTSPO-M	0.0007	−0.6851	18.6525	−8.9323	0.0373	0.0042
	QGTSPO-SMC	0.0502	11.3441	17.9851	−6.6746	1.1361	0.1363
	QGTSPO-SMI	**0.0574**	**13.2108**	18.4764	**−6.3597**	**1.3092**	**0.1664**
2008	PSO	**0.0312**	**7.7804**	19.6064	−13.5491	**0.8018**	**0.0461**
	GTSPO	−0.0876	−23.1097	20.1442	−15.0968	−1.7076	−0.1132
	GTSPO-M	−0.0325	−8.2423	14.9897	**−9.7752**	−0.7205	−0.0495
	QGTSPO-SMC	−0.0029	−1.8179	14.6131	−11.8830	−0.0656	−0.0036
	QGTSPO-SMI	−0.0041	−1.9584	**14.3836**	−11.7493	−0.0896	−0.0046
2009	PSO	0.0318	11.2984	29.7518	−12.8172	0.7162	0.0736
	GTSPO	0.0598	21.9744	28.0649	**−6.5397**	1.4848	0.2560
	GTSPO-M	**0.0657**	**25.1148**	**24.4980**	−7.4454	**1.5088**	**0.2417**
	QGTSPO-SMC	0.0618	23.9225	25.7080	−8.6999	1.4186	0.1832
	QGTSPO-SMI	0.0647	22.8095	25.2678	−9.4097	1.4788	0.1741
2010	PSO	0.0465	11.6884	17.4458	−8.0359	1.1653	0.1064
	GTSPO	−0.0049	−0.9212	**8.0943**	**−3.6459**	−0.0996	−0.0108
	GTSPO-M	0.0552	13.4214	19.2443	−6.2576	1.3135	**0.1776**
	QGTSPO-SMC	0.0703	14.7753	14.0751	−6.6798	1.6653	0.1477
	QGTSPO-SMI	**0.0726**	**15.1709**	14.3246	−7.0540	**1.7224**	0.1481

QGTSPO-SMI, were significantly higher than GTSPO, GTSPO-M, and PSO, proving their efficacy in reducing drawdowns while maintaining healthy returns.

Across the multiple periods as presented in Table 11, the QGTSPO model consistently demonstrated superior risk-adjusted performance over the GTSPO-M model, the previous GTSPO model, and the PSO model. Particularly in the year 2007, QGTSPO showed better downside risk management. In 2008, during the global financial crisis, we saw the QGTSPO providing a better return and risk trade-off. It provided better downside risk management during downturns and outperformed in recovery phases. These results suggest that the QGTSPO model is well-suited for optimizing portfolios in volatile market conditions, particularly in an emerging market like the Taiwanese Stock Exchange.

From Table 11, we can see that the average performance from 2007 to 2010 shows that the QGTSPO-SMI exhibited superior risk-adjusted performance compared to the other models. The Sharpe ratio, which measures the risk-adjusted return, was the highest, with QGTSPO-SMI achieving the highest Sharpe ratio of 0.0584. MDD for the previous model GTSPO (−6.54%) was slightly lower, but its failure to generate positive returns and worse risk performance (Sharpe ratio: −0.0532, Sortino ratio: −1.0974, Calmar: −0.1074) compared to the proposed

Table 11
Average values of performance metrics from 2007 to 2010 on TWSE Dataset.

Model	Sharpe ratio	Cumulative return (%)	Volatility (%)	MDD (%)	Sortino ratio	Calmar ratio
PSO	0.0350	10.0245	19.3843	−8.5764	0.8460	0.0760
GTSPO	−0.0531	−11.7660	**13.3254**	**−6.5438**	−1.0974	−0.1074
GTSPO-M	0.0120	0.5618	17.5558	−8.1880	0.2660	0.0287
QGTSPO-SMC	0.0534	**14.9140**	16.6826	−7.0498	1.2250	0.1450
QGTSPO-SMI	**0.0584**	13.1226	17.4361	−6.9310	**1.3549**	**0.1588**

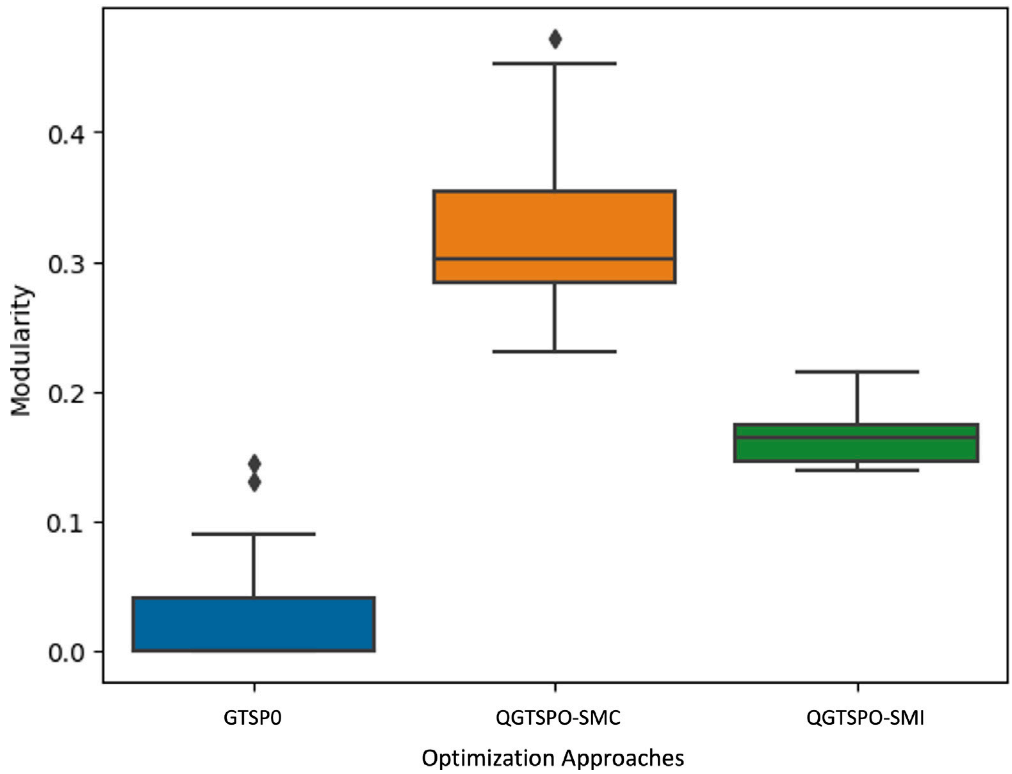

Fig. 7. Modularity values across eight runs.

approach indicate that the proposed approach has better risk management, in particular QGTSPO-SMI (Sharpe ratio: 0.0584, Sortino ratio: 1.3549, Calmar: 1.588) posted the best performance among the two QGTSPO implementations. In conclusion, the proposed model demonstrates a well-balanced approach to optimizing returns while minimizing risk and managing drawdowns more effectively compared to the other models. In particular, QGTSPO-SMI exhibited balanced risk-adjusted returns, highlighting how mutual information is an effective tool in modeling trading strategy relationships in a volatile market (TWSE) and a stable market (NYSE).

4.4. Similarity analysis

Aligned with the objectives set out in Section 1, the efficacy of the proposed algorithm was examined in clustering related trading strategies within the same group. The experiment employs modularity as a metric to gauge the robustness of these groupings. A higher modularity score signifies a superior division of strategies into related groups. Runs for each algorithm were conducted 8 times with the other parameter values set similarly to the portfolio performance experiment in the previous section (see Table 7). The results are presented in Fig. 7.

Fig. 7 illustrates that both implementations of the proposed approach, QGTSPO-SMC and QGTSPO-SMI, surpass the modularity value of the previous approach, GTSPO. QGTSPO-SMC utilizes correlation as the association matrix, while QGTSPO-SMI employs mutual information. For comparative analysis, the threshold value for QGTSPO-SMC was fixed at 0.3, and for mutual information, it was set to 0.05 for all eight runs. An edge is established between two nodes or trading strategies if the correlation in QGTSPO-SMC or mutual information value in QGTSPO-SMI is greater than or equal to the threshold. These values were determined by examining the correlation matrix (see Fig. 8) and mutual information matrix of the trading strategies (see Fig. 9).

Although correlation appears to yield higher modularity than mutual information, the portfolio performance does not exhibit significant differences. Notably, the previous method, GTSPO, fails to match the performance of QGTSPO, suggesting that QGTSPO effectively groups trading strategies into similar groups, thereby promoting diversification in portfolio results. As emphasized in Section 3, a portfolio is constructed by selecting a strategy from each group. Therefore, grouping similar strategies results in a final portfolio comprising diverse trading strategies.

4.5. Ablation study

To assess the impact of different components of the QGTSPO framework, a comprehensive evaluation was done to compare QGTSPO's performance under different settings of association matrices and the influence of incorporating modularity. While preserving the fitness function component related to group balance, modifications were made to the inclusion or exclusion of the Sharpe ratio or modularity and the type of association matrix. As part of this evaluation, the QGTSPO-S implementation was introduced, focusing solely on the Sharpe ratio

Table 12
Summary of the QGTSPO algorithm implementations: association matrix and fitness evaluation criteria, where the fitness function weights for the Sharpe and modularity components are α and β, respectively.

Algorithm	Fitness function combination weights	Association Matrix
QGTSPO-S	$\alpha = 1,\ \beta = 0$	–
QGTSPO-MI	$\alpha = 0,\ \beta = 1$	Mutual Information
QGTSPO-CI	$\alpha = 0,\ \beta = 1$	Correlation
QGTSPO-SMC	$\alpha = 0.8,\ \beta = 0.2$	Correlation
QGTSPO-SMI	$\alpha = 0.8,\ \beta = 0.2$	Mutual Information

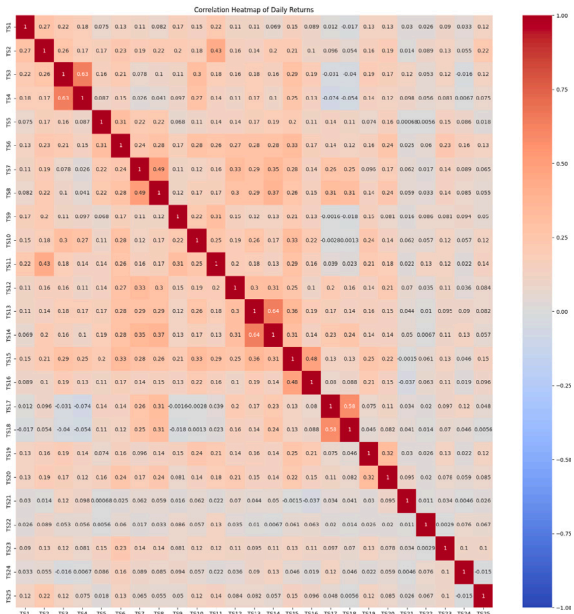

Fig. 8. Correlation between trading strategies.

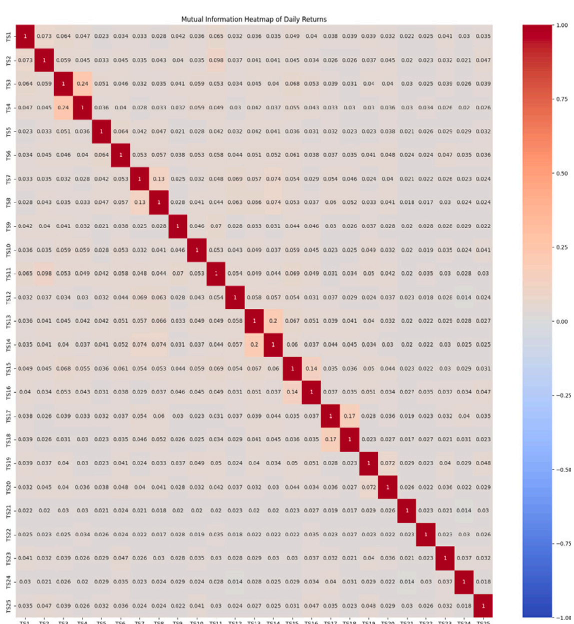

Fig. 9. Mutual information between trading strategies.

component. This was done to determine the effects of modularity on the optimization process accurately. Table 12 provides a summary of the QGTSPO implementations.

The performance of the QGTSPO models varies across the NYSE and TWSE markets, reflecting differences in market structure and the effectiveness of different configurations on the QGTSPO framework. In general, models incorporating mutual information (QGTSPO-MI, QGTSPO-SMI) tend to perform more consistently than those relying on correlation (QGTSPO-CI, QGTSPO-SMC). Furthermore, models that balance the Sharpe ratio and modularity ($\alpha = 0.8, \beta = 0.2$) demonstrate more stable performance compared to those optimizing purely for modularity ($\alpha = 0, \beta = 1$). These differences highlight the impact of the fitness function and association matrix on asset selection across markets. Average results for the period 2007–2010 are presented in Table 13. The distribution results from the data presented in Tables 13 are illustrated in Fig. 10.

When examining risk-adjusted performance in the TWSE market, QGTSPO-SMC and QGTSPO-SMI demonstrate particularly consistent results. While in terms of returns, the QGTSP-S model led with a slightly higher average return, the modularity models provided a more balanced approach by offering competitive returns while keeping risk lower. On comparing modularity and Sharpe ratio, the modularity-only models (QGTSPO-MI, QGTSPO-CI) offer the best risk mitigation while the Sharpe ratio model QGTSPO-S offers the best returns. This finding gives more clues on how the two components in the QGTSPO framework interact. The Sharpe ratio component aims to achieve high returns, and the modularity component aims for better risk management through better diversification. The models which incorporate modularity and Sharpe ratio (QGTSPO-SMI and QGTSPO-SMC) maintain favorable Sharpe, Sortino, and Calmar ratios with minimal variability, indicating reliable risk-adjusted returns in an unpredictable market condition. These two models deliver a more balanced performance, sacrificing only marginal returns to achieve significantly improved risk characteristics—a favorable tradeoff for investors prioritizing risk-adjusted returns over raw performance. This strongly indicates that incorporating modularity and Sharpe ratio into the QGTSPO, as in the SMC and SMI implementations, yields meaningful performance enhancements relative to the QGTSPO-S model which only looks at the Sharpe ratio or the QGTSPO-MI and QGTSPO-CI models which only look at modularity.

Analyzing the analysis presented in Table 13, Correlation-based models, particularly QGTSPO-CI (Sharpe ratio: 0.0488, Sortino ratio: 1.0861, Calmar ratio: 0.1366) and QGTSPO-SMC (Sharpe ratio: 0.0534, Sortino ratio: 1.2250, Calmar ratio: 0.1450), offer worse risk management indicating their vulnerability to extreme market conditions. In contrast, QGTSPO-MI (Sharpe ratio: 0.0612, Sortino ratio: 1.4062, Calmar ratio: 0.1488) and QGTSPO-SMI (Sharpe ratio: 0.0584, Sortino ratio: 1.3549, Calmar ratio: 0.1588) achieve better risk-adjusted performance. The performance of QGTSPO-SMI and QGTSPO-MI suggests that mutual information models the relationships of the trading strategies better than correlation to minimize risk. Mutual information provides a superior diversification framework in the TWSE market during periods of financial stress.

Overall, the results indicate that when the QGTSPO framework leverages mutual information, it outperforms those relying on correlation, particularly in volatile and recovering market conditions. Additionally, QGTSPO models that balance Sharpe ratio and modularity ($\alpha = 0.8, \beta = 0.2$) demonstrate greater stability in the TWSE markets as evidenced by results in Tables 13. These findings highlight the importance of selecting an appropriate fitness function and association matrix when applying QGTSPO-based models across different financial markets.

Table 13
Ablation study results: TWSE dataset (2007–2010).

Model	Sharpe ratio	Cumulative return (%)	Volatility (%)	MDD (%)	Sortino ratio	Calmar ratio
QGTSPO-S	0.0569	**15.2249**	18.9005	−8.2580	1.3022	0.1604
QGTSPO-CI	0.0488	9.1521	14.7026	**−5.6126**	1.0861	0.1366
QGTSPO-MI	0.0612	12.2485	15.8612	−6.6558	**1.4062**	0.1488
QGTSPO-SMC	0.0534	14.9140	16.6826	−7.0498	1.2250	0.1450
QGTSPO-SMI	**0.0584**	13.1226	17.4361	−6.9310	1.3549	**0.1588**

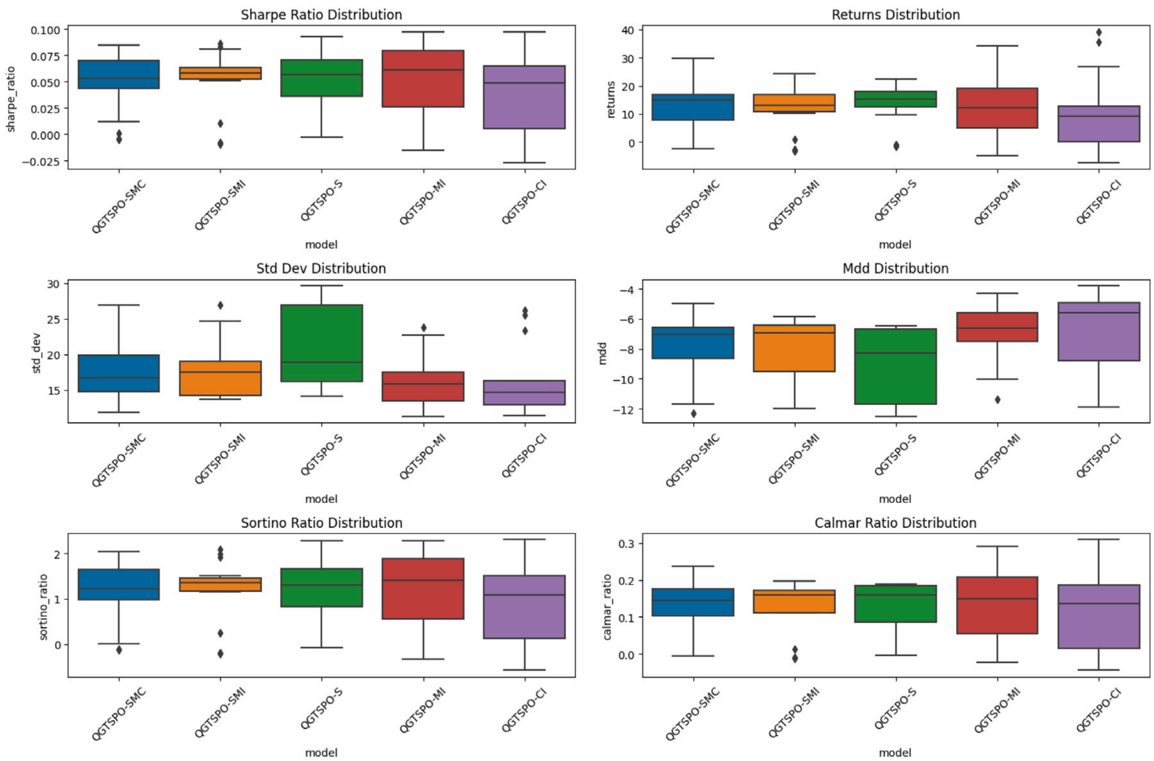

Fig. 10. Portfolio performance metric distributions of GTSPO, QGTSPO-S, QGSTPO-SMC, and QGTSPO-SMI in the TWSE market in the period 2007 to 2010.

5. Conclusions & future work

To increase the diversification of the suggested trading strategy portfolio, this paper proposes the Q-Grouping Trading Strategy Portfolio Optimization (QGTSPO) approach for finding QGTSP. QGTSPO enhances the previous work, GTSPO, by incorporating modularity into the fitness function, leveraging unique stock to each trading strategy, and modeling the relationship among the trading strategies using correlation and mutual information. In addition, introducing threshold value optimization in chromosome encoding has enabled more effective construction of trading strategy networks, resulting in more reliable portfolio performance across diverse market conditions. Experimental results demonstrate that the proposed QGTSPO approach consistently outperforms the previous methods in terms of risk-adjusted returns, with particularly notable improvements during crisis periods. Most significantly, our empirical evaluation demonstrates that the QGTSPO approach provides superior drawdown protection, more stable recovery patterns, and reduced volatility during market stress periods such as the 2008 financial crisis. This enhanced performance was observed across multiple markets, including the New York Stock Exchange and Taiwan Stock Exchange, validating the robustness of our methodology. The consistent volatility patterns exhibited by QGTSPO across various market conditions make it particularly valuable for investors with strict risk controls and capital preservation requirements. By simultaneously addressing the critical challenges of diversification, risk aversion, and network optimization, QGTSPO represents a significant advancement in constructing resilient trading strategy portfolio strategies capable of delivering favorable risk-adjusted returns even during periods of extreme market turbulence. Looking ahead, we will enhance our work in the following ways: (1) we will try to explore the utilization of weighted graphs for constructing the association matrix representing the trading strategies graph. Works such as that by Clemente et al. (2021) have examined the use of weighted correlation networks utilizing the Pearson correlation coefficient as weights; (2) Additionally, we plan to investigate alternative population initialization methods to broaden the search space for potential solutions; (3) Then, other evolutionary algorithms will also be used to find QGTSP, e.g., the Grey Wolf Optimizer (GWO) and Whale Optimization Algorithm (WOA); (4) At last, we will try to transform the GTSPO problem into multi-objective optimization problem and use multi-objective optimization algorithms, like NSGAII, to find the Pareto solutions.

CRediT authorship contribution statement

Kudakwashe Chideme: Visualization, Methodology, Data curation, Writing – original draft, Software, Investigation, Conceptualization. **Chun-Hao Chen:** Resources, Supervision, Methodology, Writing – review & editing, Project administration.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Chun Hao Chen reports financial support was provided by National Science and Technology Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the National Science and Technology Council of the Republic of China under grant NSTC 111-2221-E-992-088-MY2.

Data availability

Data will be made available on request.

References

Abbas, A., Raza, K., 2023. Sharpe index based portfolio optimization using computational intelligence. In: IEEE International Conference on Artificial Intelli- gence in Engineering and Technology. pp. 218–223.

Asawa, Y.S., 2022. Modern machine learning solutions for portfolio selection. IEEE Eng. Manag. Rev. 50, 94–112.

Banerjee, A.K., Pradhan, H., Sensoy, A., Fabozzi, F., Mahapatra, B., 2024. Robust portfolio optimization with fuzzy TODIM, genetic algorithm and multi-criteria constraints. Ann. Oper. Res. 337 (1), 1–22.

Barroso, B.C., Cardoso, R.T., Melo, M.K., 2021. Performance analysis of the integration between portfolio optimization and technical analysis strategies in the Brazilian stock market. Expert Syst. Appl. 186.

Bedoui, R., Benkraiem, R., Guesmi, K., Kedidi, I., 2023. Portfolio optimization through hybrid deep learning and genetic algorithms vine copula-GARCH-EVT-cvar model. Technol. Forecast. Soc. Change 197, 122887.

Chen, C.H., Chen, Y.H., Lin, J.C.W., Wu, M.E., 2019. An effective approach for obtaining a group trading strategy portfolio using grouping genetic algorithm. IEEE Access 7, 7313–7325.

Chen, C.H., Chen, Y.H., Wu, M.E., Hong, T.P., 2018. A sophisticated optimization algorithm for obtaining a group trading strategy portfolio and its stop-loss and take-profit points. In: IEEE International Conference on Systems, Man, and Cybernetics. IEEE, pp. 3417–3420.

Chen, C., Cheng Yu, L., Rui-Dong, C., Gautam, S., Jerry, C.W.L., 2023. An evolutionary-based approach for optimising diverse group stock portfolio with active and inactive stocks. Enterp. Inf. Syst. 17 (8), 2180328.

Chen, C.H., Lu, C.Y., Lin, C.B., 2020. An intelligence approach for group stock portfolio optimization with a trading mechanism. Knowl. Inf. Syst. 62, 287–316.

Chou, Y.H., Jiang, Y.C., Hsu, Y.R., Kuo, S.-Y., Kuo, S.Y., 2022. A weighted portfolio optimization model based on the trend ratio, emotion index, and ANGQTS. IEEE Trans. Emerg. Top. Comput. Intell. 6 (4), 867–882.

Clemente, G.P., Grassi, R., Hitaj, A., 2021. Asset allocation: New evidence through network approaches. Ann. Oper. Res. 299, 61–80.

Deb, K., Agrawal, S., et al., 1999. Understanding interactions among genetic algorithm parameters. Found. Genet. Algorithms 5 (5), 265–286.

Deb, K., Deb, D., 2014. Analysing mutation schemes for real-parameter genetic algorithms. Int. J. Artif. Intell. Soft Comput. 4 (1), 1–28.

Falkenauer, E., 1998. Genetic Algorithms and Grouping Problems. John Wiley & Sons, Inc..

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486 (3–5), 75–174.

Giudici, P., Pagnottoni, P., Polinesi, G., 2020. Network models to enhance automated cryptocurrency portfolio management. Front. Artif. Intell. 3.

Guarino, A., Santoro, D., Grilli, L., Zaccagnino, R., Balbi, M., 2024. EvoFolio: a portfolio optimization method based on multi-objective evolutionary algorithms. Neural Comput. Appl. 36 (13), 7221–7243.

Gunjan, A., Bhattacharyya, S., 2023. A brief review of portfolio optimization techniques. Artif. Intell. Rev. 56 (5), 3847–3886.

Hasan, F., Ahmad, F., Shahid, M., Khan, A., Ahmad, G., 2022. Solving portfolio selection problem using whale optimization algorithm. 3rd Int. Conf. Comput. Autom. Knowledge Manag. 1–5.

Ioannidis, E., Sarikeisoglou, I., Angelidis, G., 2023. Portfolio construction: A network approach. Mathematics 11 (22).

Kumar, A., Yadav, S., Gupta, P., Mehlawat, M.K., 2023. A credibilistic multiobjective multiperiod efficient portfolio selection approach using data envelopment analysis. IEEE Trans. Eng. Manage. 70 (6), 2334–2348.

Li, Y., Jiang, X.F., Tian, Y., Li, S.P., Zheng, B., 2019. Portfolio optimization based on network topology. Phys. A 515.

Lim, S., Kim, M.J., Ahn, C.W., 2020. A genetic algorithm approach to the portfolio design based on market movements and asset valuations. IEEE Access 8, 140234–140249.

Lopez de Prado, M., 2016. Building diversified portfolios that outperform out-of-sample. J. Portf. Manag..

Lou, K., 2023. Optimizing portfolios with modified NSGA-II solutions. In: IEEE 3rd International Conference on Data Science and Computer Application. pp. 375–380.

Lou, H., Li, S., Zhao, Y., 2013. Detecting community structure using label propagation with weighted coherent neighborhood propinquity. Phys. A. 392, 3095–3105.

Lv, M., Wang, J., Wang, S., Gao, J., Guo, H., 2024. Developing a hybrid system for stock selection and portfolio optimization with many-objective optimization based on deep learning and improved NSGA-III. Inform. Sci. 670, 120549.

Markowitz, H., 1952. Portfolio selection. J. Financ. 7 (1), 77–91.

Newman, M., 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. 103 (23), 8577–8582.

Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. Phys. Rev. E. 69, 026113.

Pal, R., Chaudhuri, T.D., Mukhopadhyay, S., 2021. Portfolio formation and optimization with continuous realignment: a suggested method for choosing the best portfolio of stocks using variable length NSGA-II. Expert Syst. Appl. 186, 115732.

Raghavan, U., Albert, R., Kumara, S., 2007. Near linear time algorithm to detect community structures in large-scale networks. Phys. Rev E. 76, 036106.

Ricca, F., Scozzari, A., 2024. Portfolio optimization through a network approach: Network assortative mixing and portfolio diversification. European J. Oper. Res. 312 (2).

Solares, E., Salas, F.G., De-Leon-Gomez, V., Diaz, R., 2022. A comprehensive soft computing-based approach to portfolio management by discarding undesirable stocks. IEEE Access 10, 40467–40481.

Šubelj, L., Bajec, M., 2012. Ubiquitousness of link-density and link-pattern communities in real-world networks. Eur. Phys. J. B. 85, 1–11.

Vehlow, C., Reinhardt, T., Weiskopf, D., 2013. Visualizing fuzzy overlapping communities in networks. IEEE Trans. Vis. Comput. Graphics 19, 2486–2495.

Vlasov, A., Khomchenko, A., Faizliev, A., Mironov, S., Grigoriev, A., 2021. Parameter tuning of a genetic algorithm for finding central vertices in graphs. J. Phys.: Conf. Ser. 1784, 012009.

Wang, F., Huang, Z., Wang, S., 2023. I $\epsilon+$ LGEA a learning-guided evolutionary algorithm based on I $\epsilon+$ indicator for portfolio optimization. Complex Syst. Model. Simul. 3 (3), 191–201.

Yaman, I., Dalkılıç, T.E., 2021. A hybrid approach to cardinality constraint portfolio selection problem based on nonlinear neural network and genetic algorithm. Expert Syst. Appl. 169, 114517.