Financial Research Journal

Print ISSN: 1024-8153
Online ISSN: 2423-5377

Stable and Cost-Efficient Tracking of the Tehran Stock Exchange Index through Robust Optimization and a Heuristic Algorithm

Danial Ramezani
MSc. Student, Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran. E-mail: danialramezani@khu.ac.ir

Mostafa Abouei Ardakan*
*Corresponding Author, Associate Prof., Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran. E-mail: abouei@khu.ac.ir

Mohamadreza Dehghani Ahmadabad
Assistant Prof., Department of Financial Management, Faculty of Financial Sciences, Kharazmi University, Tehran, Iran. E-mail: mr.dehghani@khu.ac.ir

Abstract

Objective

The rising popularity of passive management in recent years is largely due to its advantages, such as lower management fees and reduced transaction costs. A key component of passive management is index tracking, which aims to replicate the performance of a specific index using a smaller set of assets. This paper introduces a novel robust linear optimization approach for tracking indices that is not only more reliable than existing models but also exhibits superior performance on out-of-sample data, effectively tracking indices over extended periods with minimal deviation and without requiring frequent rebalancing. The proposed robust model is highly adaptable, allowing decision-makers to account for a wide range of potential future scenarios, including both the most certain outcomes and the worst-case possibilities. The importance of this model lies in its ability to adapt more effectively to market fluctuations and help investors get closer to their return objectives.

Citation: Ramezani, Danial; Abouei Ardakan, Mostafa & Dehghani Ahmadabad, Mohamadreza (2026). Stable and Cost-Efficient Tracking of the Tehran Stock Exchange Index through Robust Optimization and a Heuristic Algorithm. *Financial Research Journal*, 28(2), 494 – 526. https://doi.org/10.22059/FRJ.2025.379970.1007626 (*in Persian*)

Financial Research Journal, 2026, Vol. 28, No.2, pp. 494 - 526
Published by University of Tehran, College of Management
https://doi.org/10.22059/FRJ.2025.379970.1007626
Article Type: Research Paper
© Authors
Received: July 28, 2024
Revised: May 30, 2025
Accepted: July 07, 2025
Published online: June 20, 2026

Methods

Given the NP-hard nature of index tracking problems with cardinality constraints, solving them in polynomial time is challenging. To address this computational burden, a novel heuristic approach was designed. This heuristic integrates the exploration capabilities of genetic algorithms with the focused search capabilities of local search techniques and is designed to solve the problem efficiently. Daily data from the Tehran Stock Exchange and OR-library were utilized to validate the proposed model and heuristic. For comparison purposes, a commercial solver was employed to solve both the proposed robust model and the conventional linear model commonly used in the literature. This approach was adopted instead of the developed algorithm because, although the proposed algorithm demonstrates superior performance compared to the commercial solver, it is inherently stochastic in nature. Therefore, to eliminate the influence of random and probabilistic factors on the results, a deterministic and precise commercial solver was utilized.

Results

The results demonstrate that the proposed heuristic not only converges to optimal solutions for moderately sized problems but also produces portfolios that outperform those generated by commercial solvers in terms of both in-sample and out-of-sample data, all within a shorter time frame. Furthermore, given the structure of the Tehran Stock Exchange, the impact of including or excluding the ten largest stocks by market capitalization in the selected portfolio on the final results of the proposed tracking portfolio has been examined. This analysis indicates that the presence of these stocks in the portfolio can influence the performance of the tracking portfolio.

Conclusion

This study aimed to develop an effective approach for index tracking by proposing a robust linear mathematical model alongside a novel heuristic algorithm. The performance of the proposed approach was evaluated against a benchmark linear model using five quantitative indicators: correlation, mean absolute deviation (MAD), root mean squared error (RMSE), standard deviation, and beta. The proposed model outperformed the benchmark model across all indicators in the out-of-sample period. In addition to these indicators, the objective function value was used to assess the performance of the proposed heuristic. The algorithm outperformed the CPLEX solver in 77 out of 88 comparisons, 11 indicators under four configurations for the Tehran Stock Exchange and the Hang Seng Index. Furthermore, an alternative tracking portfolio with a novel weighting system was introduced, which effectively reflects the status of the largest industries represented in the Tehran Stock Exchange. The performance of this alternative portfolio was compared with that of the proposed portfolio in terms of return, risk, and tracking error. The tracking error of the proposed portfolio was found to be nearly three times lower than that of the alternative approach.

Keywords: Index tracking, Passive fund management, Heuristics, Mixed-integer linear programming, Portfolio management.

تحقیقات مالی
شاپا چاپی: ۱۰۲۴-۸۱۵۳
شاپا الکترونیکی: ۲۴۲۳-۵۳۷۷

ردیابی پایدار و اقتصادی شاخص بورس تهران با بهره‌گیری از بهینه‌سازی استوار و الگوریتم ابتکاری

دانیال رمضانی
دانشجوی کارشناسی ارشد، گروه مهندسی صنایع، دانشکده مهندسی، دانشگاه خوارزمی، تهران، ایران. رایانامه: danialramezani@khu.ac.ir

مصطفی ابویی اردکان*
* نویسنده مسئول، دانشیار، گروه مهندسی صنایع، دانشکده مهندسی، دانشگاه خوارزمی، تهران، ایران. رایانامه: abouei@khu.ac.ir

محمدرضا دهقانی احمدآباد
استادیار، گروه مدیریت مالی، دانشکده علوم مالی، دانشگاه خوارزمی، تهران، ایران. رایانامه: mr.dehghani@khu.ac.ir

چکیده

هدف: سرمایه‌گذاری منفعل در سال‌های اخیر، به‌دلیل داشتن هزینه‌های مدیریت و معاملات کمتر، به‌طور فزاینده‌ای در کانون توجه قرار گرفته است. ردیابی شاخص به‌عنوان یکی از انواع این سرمایه‌گذاری‌ها، به بازتولید عملکرد یک شاخص با استفاده از سبدی متشکل از تعداد کمی از سهام می‌پردازد. در این پژوهش، یک مدل ریاضی خطی استوار جدید، برای ردیابی شاخص معرفی شده است که نسبت به مدل‌های مشابه موجود، عملکرد بهتری دارد. این مدل قادر است شاخص را برای مدت طولانی‌تر، با خطای ردیابی کمتر و بدون نیاز به تعدیل سبد ردیابی کند. همچنین، مدل پیشنهادی انعطاف‌پذیری زیادی دارد و به تصمیم‌گیرنده این امکان را می‌دهد که تمامی سناریوهای ممکن، از حالت بدون‌عدم قطعیت تا بدترین حالت ممکن را در نظر بگیرد. اهمیت مدل ارائه شده در این است که می‌تواند به‌طور مؤثرتری با نوسان‌های بازار سازگار شود و به سرمایه‌گذاران کمک کند تا به اهداف خود در زمینه بازدهی نزدیک‌تر شوند.

روش: از آنجا که حل مسائل مرتبط با ردیابی شاخص با محدودیت کاردینالیتی در زمان چندجمله‌ای ممکن نیست، این پژوهش به توسعهٔ یک الگوریتم ابتکاری جدید مبتنی بر الگوریتم ژنتیک و انشعاب محلی پرداخته است. این الگوریتم ترکیبی از قابلیت‌های اکتشافی الگوریتم ژنتیک و ویژگی‌های جست‌وجوی محلی است و به‌گونه‌ای طراحی شده است که بتواند به‌طور کارآمدتر به حل مسئله بپردازد. در این تحقیق، داده‌های روزانهٔ بورس تهران و بازار هنگ‌سنگ از کتابخانه OR-library برای ارزیابی عملکرد مدل ریاضی و الگوریتم پیشنهادی استفاده شده است. برای بررسی عملکرد مدل خطی استوار پیشنهادی در مقایسه با مدل‌های خطی رایج در ادبیات موضوع، از یک حل‌کنندهٔ تجاری استفاده شده است. دلیل استفاده از این رویکرد و نه استفاده از الگوریتم توسعه‌یافته، این است که الگوریتم پیشنهادی با وجود برتری نسبت به حل‌کنندهٔ تجاری، ماهیت تصادفی دارد و به‌منظور حذف فاکتورهای تصادفی و احتمالاتی که ممکن است بر نتایج تأثیرگذار باشند، از حل‌کنندهٔ تجاری دقیق استفاده شده است.

استناد: رمضانی، دانیال؛ ابویی اردکان، مصطفی و دهقانی احمدآباد، محمدرضا (۱۴۰۵). ردیابی پایدار و اقتصادی شاخص بورس تهران با بهره‌گیری از بهینه‌سازی استوار و الگوریتم ابتکاری. *تحقیقات مالی*، ۲۸(۲)، ۴۹۴- ۵۲۶.

تحقیقات مالی، ۱۴۰۵، دوره ۲۸، شماره ۲، صص. ۴۹۴ - ۵۲۶
ناشر: دانشکدگان مدیریت دانشگاه تهران
نوع مقاله: علمی پژوهشی
© نویسندگان
تاریخ دریافت: ۱۴۰۳/۰۵/۰۷
تاریخ ویرایش: ۱۴۰۴/۰۳/۰۹
تاریخ پذیرش: ۱۴۰۴/۰۴/۱۶
تاریخ انتشار: ۱۴۰۵/۰۳/۳۰
doi: https://doi.org/10.22059/FRJ.2025.379970.1007626

یافته‌ها: نتایج به‌دست‌آمده از مقایسهٔ الگوریتم پیشنهادی با حل‌کنندهٔ تجاری روی مدل ارائه شده، نشان می‌دهد که رویکرد پیشنهادی نسبت به رویکردهای قبلی، عملکرد بهتری داشته و الگوریتم پیشنهادی قادر است جواب‌های بهتری در زمان کمتری بیابد. همچنین، با توجه به ساختار بورس تهران، تأثیر حضور یاعدم حضور ۱۰ سهم بزرگ بازار از نظر ارزش بازار در پرتفوی انتخابی، روی نتایج نهایی سبد ردیاب پیشنهادی بررسی شده است. این بررسی‌ها نشان می‌دهد که وجود این سهام در پرتفوی، می‌تواند بر عملکرد سبد ردیاب تأثیر داشته باشد.

نتیجه‌گیری: پژوهش حاضر با ارائه یک مدل ریاضی خطی استوار و الگوریتم ابتکاری جدید، رویکردی نوین و مؤثر برای ردیابی شاخص ارائه کرده است. عملکرد مدل پیشنهادی با ۵ شاخص کمی، مشتمل بر هم‌بستگی، میانگین انحراف مطلق، ریشه میانگین مربعات خطا، انحراف معیار و بتا، با مدل خطی معیار مقایسه شد. مدل پیشنهادی موفق به شکست مدل معیار در تمامی شاخص‌ها در بازهٔ خارج از نمونه شده و عملکرد بهتری ثبت کرده است. مقدار تابع هدف برای سنجش عملکرد الگوریتم پیشنهادی به شاخص‌های نام‌برده اضافه شده است. الگوریتم پیشنهادی موفق به شکست حل‌کنندهٔ CPLEX در ۷۷ مقایسه از ۸۸ مقایسهٔ انجام شده، ۱۱ معیار در ۴ پیکربندی مختلف برای دو بازار بورس تهران و هنگ‌سنگ شده است. همچنین، یک سبد ردیاب جایگزین با سیستم وزن‌دهی نوین معرفی شده است که به‌خوبی توانایی نمایش وضعیت بزرگ‌ترین صنایع حاضر در بورس تهران را دارد. عملکرد این سبد نیز از نظر بازده، ریسک و ریسک ردیابی با سبد پیشنهادی این پژوهش مورد قیاس قرار گرفته شده است. ریسک ردیابی سبد پیشنهادی این پژوهش تقریباً سه برابر از روش جایگزین کمتر است.

کلیدواژه‌ها: ردیابی شاخص، سرمایه‌گذاری منفعل، الگوریتم ابتکاری، برنامه‌ریزی عدد صحیح، مدیریت سبد سهام.

مقدمه

سبدگردانی یا مدیریت دارایی شامل تصمیم‌گیری از طرف مدیر صندوق جهت مدیریت سرمایه مشتریان با هدف کسب سود بر اساس سطح ریسک مورد انتظار مشتری است. به‌طورکلی، سبک مدیریت در این مؤسسه‌ها را می‌توان به دو گروه فعال و منفعل تقسیم‌بندی کرد. مدیریت فعال شامل بررسی و مدیریت دائم سبد سهام برای بررسی فرصت‌های موجود در بازار برای ساختن و تغییر دائم سبد سهام با هدف رسیدن به عملکردی بهتر و کسب سود بیشتر از بازار است. مدیران در این سبک با استفاده از تخصص خود در بررسی شرایط اقتصادی، بررسی صنایع، سیاست، زمان‌بندی بازار و مواردی از این‌دست، به خرید یا فروش سهام اقدام می‌کنند. حق‌الزحمه مدیریت در این صندوق‌ها در مقایسه با مدیریت منفعل بدون توجه به سودآور بودن یا زیان‌ده بودن مدیر صندوق بیشتر است. همچنین، گردش و هزینه معاملات سبد سهام بالاتر از صندوق‌های غیرفعال هستند. این هزینه‌های انباشته می‌توانند سود نهایی به‌دست‌آمده را به میزان زیادی تحت تأثیر قرار دهد.

دسته دوم از نحوه مدیریت صندوق‌ها مدیریت منفعل است. این دسته از مدیران معتقد هستند که شکست دادن بازار در یک دوره طولانی غیرممکن است. به عقیده این دسته از مدیران زمان ماندن در بازار بسیار مهم‌تر از زمان‌بندی در بازار است بنابراین هدف آن‌ها سرمایه‌گذاری در شاخص بورس است. ساده‌ترین راه‌حل، خریداری تمامی سهم‌های فهرست شده در بورس است که روی شاخص اثرگذاری دارند؛ اما خرید یا فروش بعضی از سهم‌ها ممکن است با توجه به نقدشوندگی پایین ممکن نباشد. علاوه‌براین، هزینه‌های معاملات، نظارت و تعدیل سبد می‌تواند بسیار زیاد باشد و این استراتژی را در واقعیت غیرعملی سازد. با توجه به موانع فوق، مدیران، زیرمجموعه‌ای از دارایی‌ها را در بازار انتخاب می‌کنند تا یک سبد برای ردیابی بازده بازار ایجاد کنند. با انجام این رویکرد، هزینه‌ها و مشکلات اشاره‌شده تا حدی مرتفع می‌شود. این استراتژی به‌عنوان ردیابی شاخص[1] شناخته می‌شود. ردیابی شاخص، به‌طور گسترده توسط مدیران صندوق‌های دارایی و پژوهشگران در دهه‌های اخیر مورد توجه قرار گرفته است. رویکرد مدل‌سازی ریاضی رایج‌ترین روشی است که محققان برای مدل‌سازی مسئله ردگیری شاخص از آن استفاده می‌کنند (سیلوا[2]، ۲۰۲۳). انتخاب یک زیرمجموعه از دارایی‌ها با هدف ردگیری یک شاخص خاص، به ایجاد یک مسئله بهینه‌سازی ترکیباتی منجر می‌شود که این مسئله ماهیت NP-hard دارد و با روش‌های معمول حل نمی‌شود؛ بنابراین، حجم وسیعی از تحقیقات به توسعه روش‌هایی برای حل این مسئله اختصاص داده شده است. نکته حائز اهمیت دیگر که باید مدنظر قرار گرفته شود نیز پیش‌فرض رویکرد بهینه‌سازی است؛ رویکرد بهینه‌سازی در ردیابی شاخص بر این فرض استوار است که یک «سبد سهام» که برای داده‌های داخل نمونه شاخص را به‌خوبی ردیابی کرده است، در آینده و برای داده‌های خارج از نمونه نیز عملکرد خوبی در ردیابی شاخص به جای خواهد گذاشت. این منطق لزوماً در دنیای واقعی صادق نیست و این مسئله، کار را برای یافتن یک سبد سهام ردیاب خوب در واقعیت سخت خواهد کرد.

1. Index tracking
2. Silva

پیشینهٔ پژوهش

در این قسمت به بررسی ادبیات موضوع مسئله ردیابی شاخص می‌پردازیم. رول[1] (۱۹۹۲) روی یافتن مرز کارا برای مسئله ردیابی شاخص تمرکز کرد که در آن هدف به حداقل رساندن واریانس خطای ردیابی و حداکثر کردن میزان بازده کسب شده بود. وی نشان داد که اضافه کردن بتا در محدودیت باعث بهبود جواب در بعضی از سناریوها می‌شود. یانسن و ون دایک[2] (۲۰۰۲) تعداد سهام انتخابی در سبد را به‌عنوان یکی از اهداف مسئله تعریف کردند. اوریون[3] (۲۰۰۳) بیان کرد که در نظر گرفتن خطای ردیابی به‌تنهایی مفید نیست و به جواب‌های ناکارآمد منجر می‌شود. برای حل این مشکل محدود کردن واریانس سبد سهام ردیاب در قالب یک محدودیت در مدل را پیشنهاد کرد، به‌نحوی‌ک ه واریانس سبد کمتر از واریانس شاخص باشد.

بیزلی، مید و چانگ[4] (۲۰۰۳) یک مدل ریاضی برای ردیابی شاخص با در نظر گرفتن هزینه‌های تراکنش و محدودیت تعداد سهام پیشنهاد کردند. همچنین الگوریتمی فراابتکاری برای حل مدل پیشنهادی ارائه کردند. آن‌ها رویکرد خود را بر روی پنج شاخص بازارهای بزرگ جهانی آزمایش کردند و زمان حل را به ازای پارامترهای مختلف مسئله در مدل خود گزارش کردند. گایورونسکی، کریلو و وان‌دروایست[5] (۲۰۰۵) تأثیر به‌کارگیری توابع ریسک و توابع هدف مختلف را روی مسئله ردیابی شاخص مورد مطالعه قرار دادند و مدل‌های ریاضی مربوط به آنان را ارائه کردند. همچنین آزمایش‌هایی روی داده‌های بورس اوراق بهادار اسلو انجام انجام دادند.

الکساندر و دیمیتریو[6] (۲۰۰۵) مدل‌های مبتنی بر همگرایی و مدل‌های ریاضی موجود در ادبیات موضوع ردیابی شاخص را مقایسه کردند و نشان دادند که استفاده از هر دو مدل، نتایج کمابیش مشابهی به همراه دارد. محققان زیادی الگوریتم‌های فراابتکاری یا ابتکاری را برای حل این مسئله توسعه دادند. اوه، کیم و مین[7] (۲۰۰۵) از الگوریتم ژنتیک[8] برای ردیابی شاخص در بازار سهام کره در رژیم‌های مختلف بازار استفاده کردند و خطای ردیابی را در سناریوهای مختلف بررسی کردند.

مارینگر و اویومی[9] (۲۰۰۷) الگوریتم تکامل دیفرانسیل[10] را ارائه و آن را روی شاخص میانگین صنعتی داو جونز آزمایش کردند. آن‌ها همچنین پیشنهاد کردند که استفاده از اطلاعات سهام در گذشته بسیار دور، می‌تواند در فرایند بهینه‌سازی نامناسب باشد. به اعتقاد آن‌ها، سبدهای بزرگ‌تر خطای ردیابی کمتری دارند.

1. Roll
2. Jansen & Van Dijk
3. Jorion
4. Beasley, Meade & Chang
5. Gaivoronski, Krylov & Van der Wijst
6. Alexander & Dimitriu
7. Oh, Kim & Min
8. Genetic Algorithm
9. Maringer & Oyewumi
10. Differential Evolution algorithm

رویز توروبیانو و سوارز[1] (۲۰۰۹) یک الگوریتم ژنتیک ترکیبی برای ردیابی شاخص ارائه کردند که در آن برنامه‌ریزی برنامه‌ریزی غیرخطی درجه دوم به‌عنوان تابع هدف برای تعیین وزن دارایی‌ها استفاده می‌شود.

کرینک، میتنیک و پترلینی[2] (۲۰۰۹) یک الگوریتم تکامل دیفرانسیل ترکیبی را پیشنهاد کرند که از عملگر جست‌وجوی ترکیبی استفاده می‌کرد. آن‌ها نتایج خود را روی یک شاخص مصنوعی که با در نظر گرفتن ارزش‌های بازار نیکی ژاپن ساخته شده است آزمایش کردند.

چن و کوون[3] (۲۰۱۲) یک مدل برنامه‌ریزی استوار برای ردیابی شاخص ارائه کردند و با استفاده از داده‌های شاخص شاخص S&P 100 آن را آزمایش کردند. وانگ، خو، خو و شوئه[4] (۲۰۱۲) محدودیت CVaR را به مسئله اضافه کردند. تجزیه‌وتحلیل عددی نشان داد که افزودن CVaR می‌تواند کاهش ارزش ناگهانی سبد را محدود کند.

وو، کوون و کاستا[5] (۲۰۱۷) مدل‌های ریاضی با محدودیت‌های گردش مالی، کارمزد معاملات و تنوع صنعت سبد را را توسعه دادند. همچنین از الگوریتم ابتکاری برای ایجاد جواب اولیه و از روش‌های لاگرانژی و نیمه لاگرانژی برای دستیابی به راه‌حل نهایی استفاده کردند. کائوسیچ، بربینی و کامروتا وردو[6] (۲۰۲۰) انحراف منفی از شاخص، سود بالقوه صعودی و نسبت شارپ اصلاح‌شده را به‌عنوان توابع هدف در نظر گرفته و با کمک برنامه‌ریزی آرمانی آن‌ها را فرموله کردند. همچنین با استفاده از الگوریتم ذرات[7] بهبود داده شده مسئله را حل کردند.

سنت‌آنا، الیویرا، فیلومنا و کالدریا[8] (۲۰۲۰) یک مدل برنامه‌ریزی غیرخطی عدد صحیح مختلط را برای ردیابی شاخص با استفاده از روش همگرایی فرموله کرده و با استفاده از الگوریتم شاخه و برش[9] آن را حل کردند. نتایج نشان داد داد که مدل آن‌ها سبدهایی با گردش مالی و هزینه تراکنش کمتر ارائه می‌دهد.

انیس، کاستا و کوون[10] (۲۰۲۳) مدل جدیدی را برای تقسیم ریسک سبد سهام ردیاب بین صنایع مختلف ارائه کردند. سیلوا، سیلوا و دی‌آلمیدا فیلهو[11] (۲۰۲۴) روش جست‌وجوی تصادفی تطبیقی حریصانه[12] را پیشنهاد و ادعا کردند که الگوریتم آن‌ها می‌تواند تقریباً به‌اندازه حل‌کننده‌های تجاری جواب خوبی برای مسئله ارائه دهد.

سنت‌آنا، ریگی، مولر و گودس[13] (۲۰۲۲) تأثیر اضافه کردن هفت معیار ریسک مختلف بر مدل برنامه‌ریزی ردیابی شاخص را مطالعه کردند و پس از معرفی معیارهای مناسب برای اندازه‌گیری عملکرد هر کدام در داده‌های خارج از نمونه

1. Ruiz-Torrubiano & Suárez
2. Krink, Mittnik & Paterlini
3. Chen & Kwon
4. Wang, Xu, Xu & Xue
5. Wu, Kwon & Costa
6. Kaucic, Barbini & Camerota Verdù
7. Particle Swarm Optimization algorithm
8. Sant'Anna, de Oliveira, Filomena & Caldeira
9. Branch and Cut algorithm
10. Anis, Costa & Kwon
11. Silva, Silva & de Almeida Filho
12. GRASP
13. Sant'Anna, Righi, Müller & Guedes

نتایج هر معیار را ارائه کردند. توری، جاکومتی و پاترلینی[1] (۲۰۲۴) از جریمه کردن برای کنترل انحراف وزن هر سهم در سبد ردیاب از وزن آن سهم در شاخص استفاده کردند. نسبت به روش‌های سنتی، روش آن‌ها اهمیت بیشتری روی ریسک و گردش مالی خود سبد تا ریسک نسبت به شاخص دارد.

سیلوا و دی‌آلمیدا فیلهو[2] (۲۰۲۳) الگوریتمی تکاملی فراابتکاری توسعه دادند و عملکرد آن را با استفاده از داده‌های تاریخی بازار و داده‌های شبیه‌سازی شده تست کردند.

ویرا، فیلومنا، سنت‌آنا و لژون[3] (۲۰۲۳) یک مطالعه جامع و کامل روی نقدینگی و تأثیر آن در ردیابی شاخص انجام دادند. همچنین برای مدل‌سازی نقدینگی دو روش مختلف پیشنهاد کردند. آنان نشان دادند که در روش اول که به سادگی جمع وزنی نقدینگی سهم‌های سبد بود، محدود کردن بیشتر باعث کاهش تعداد سهام سبد ردیاب می‌شود؛ اما مدل فرموله شده روش دوم، چنین تأثیری از خود نشان نمی‌دهد.

عیوضلو، شفیع‌زاده و قهرمانی (۱۳۹۶) از دو رویکرد هم‌بستگی و هم‌انباشتگی در بورس اوراق بهادار تهران برای ردیابی شاخص کل استفاده کردند. همچنین نشان دادند که افزایش داده‌های درون نمونه باعث بهبود عملکرد مدل‌ها شده و به‌طورکلی روش هم‌انباشتگی[4] عملکرد بهتری نسبت به هم‌بستگی داشته است.

آزادی و نجفی (۱۳۹۹) مدل برنامه‌ریزی برای ردیابی شاخص بورس تهران ارائه کردند که هدف آن کمینه‌سازی انحرافات منفی از شاخص و بیشینه‌سازی انحرافات مثبت از شاخص بوده و برای حل مدل ارائه شده از الگوریتم ژنتیک استفاده کردند. عیوضلو، فلاح‌پور و دهقانی اشکذری (۱۴۰۰) مدل برنامه‌ریزی خطی با استفاده از ارزش در معرض ریسک شرطی ترکیبی دو دنباله‌ای ارائه و آن را با روش پنجره غلطان روی داده‌های هفتگی بورس آزمایش کردند و نشان دادند که مدل موردنظر قادر به بازسازی عملکرد شاخص است.

بت‌شکن، بحرالعلوم، ارضا و تقی‌خان تجریشی (۱۳۹۸) معیار جدیدی برای بررسی کیفیت خطای ردیابی ارائه کردند و نشان دادند که این معیار نسبت به روش های سنتی دیگر از پایداری و ثبات بیشتری برخوردار بوده و جهت استفاده مناسب‌تر است. محمدی، راعی و تندنویس (۱۴۰۰) با تکنیک خوشه‌بندی سری‌های زمانی سبد ردیاب تشکیل داده و سپس اوزان سهام را به‌گونه‌ای محاسبه کردند که بازده مازاد سبد بیشینه شود.

تندنویس و حکیمیان (۱۳۹۹) تأثیر معیارهای متنوع سنجش شباهت سری‌های زمانی را بر روی مسئله ردیابی شاخص بررسی کردند و نشان دادند که خطای ردیابی با ضریب هم‌بستگی پیرسون به‌طور معناداری از سایر معیارها کمتر است. شهاب لواسانی، رضوانی و سماوی (۱۴۰۲) ارتباط بین روند شاخص کل و احساسات سرمایه‌گذاران را بررسی کرده و به این نتیجه رسیدند که ماندگاری در بازار سرمایه در رژیم رکود بیشتر است.

رستمی نوروزآباد، گل‌بابائی پاسندی، شهرازی و اسفندیاری (۱۴۰۳) جهت سنجش تحرکات سهام، شاخص ترس و طمع به‌صورت روزانه برای بورس اوراق بهادار تهران معرفی کردند. گل‌ارضی و ابوالفضلی (۱۴۰۳) اثرهای اهرمی،

1. Torri, Giacometti & Paterlini
2. Silva & de Almeida Filho
3. Vieira, Filomena, Sant'anna & Lejeune
4. Cointegration

هم‌بستگی شرطی پویا و سرایت‌پذیری تلاطم را در میان ۱۰ شاخص برتر بورس بررسی کرده و پیشنهاد کردند که جهت پوشش ریسک از یافته‌های پژوهش استفاده شود.

اصولیان، نیک‌مرام و کریمی (۱۴۰۴) یک شبکه عصبی هیبریدی با تمرکز بر استخراج ویژگی در مقیاس‌های زمانی چندگانه برای پیش‌بینی شاخص بازار سهام ارائه کردند. برخلاف مطالعات پیشین که برتری این نوع مدل‌ها را نشان می‌دادند، نتایج آن‌ها حاکی از آن بود که مدل‌های ساده‌تر در پیش‌بینی شاخص عملکرد بهتری دارند.

هدف از این پژوهش در وهلهٔ اول، ارائهٔ یک مدل برنامه‌ریزی خطی عدد صحیح مختلط استوار است که قادر باشد هم نسبت به مدل‌های موجود، جواب‌های بهتر خارج از نمونه ارائه دهد و هم بتواند شاخص را برای مدت طولانی‌تر، بدون نیاز به تعادل مجدد ردیابی کند. در وهلهٔ دوم، از آنجایی که این دسته از مدل‌ها NP-hard هستند، برای حل مسئله پیشنهادی، یک الگوریتم ابتکاری قدرتمند پیشنهاد خواهد شد که قادر است به جواب بهینه در اندازه‌های کوچک و متوسط برسد و همچنین، قابلیت دارد نسبت به حل‌کننده‌های تجاری جواب‌های بهتری را در زمانی کمتری پیدا کند. برای دستیابی به چنین هدفی، قدرت اکتشافی الگوریتم ژنتیک با ویژگی جست‌وجو محلی الگوریتم انشعاب محلی[1] ادغام شده است و الگوریتم ژنتیک با انشعاب محلی (GALB)[2] پیشنهاد شده است. GALB کاملاً قادر به پیشی گرفتن از حل‌کننده‌های تجاری یا همگرا شدن به جواب بهینه است.

ادامه مقاله به‌صورت زیر سازمان‌دهی شده است. در بخش سوم، مدل‌های ریاضی مرسوم در مرور ادبیات موضوع ردیابی شاخص مورد بررسی و تجزیه و تحلیل قرار گرفته و مدل برنامه‌ریزی ریاضی پیشنهادی ارائه خواهد شد. در بخش چهارم، الگوریتم ابتکاری پیشنهادی ارائه شده است. بخش پنجم به بررسی عملکرد مدل و الگوریتم پیشنهادی پرداخته و در بخش آخر نیز نتیجه‌گیری و یافته‌های مقاله گزارش شده است.

روش‌شناسی پژوهش

مدل‌های ریاضی سنتی

در این بخش مدل‌های ریاضی موجود در ادبیات موضوع ردیابی شاخص مورد بررسی قرار گرفته است. فرض کنید بازاری دارای N سهم است و ρ_{ij} معیار شباهت سهم i با j باشد. یک مدل ریاضی مرسوم برای ردیابی شاخص می‌تواند به شکل زیر نوشته شود. شایان ذکر است که در مرور ادبیات موضوع ردیابی شاخص، هم‌بستگی سهم‌ها معیار پذیرفته‌شده و مرسومی برای میزان شباهت بین آن‌هاست.

مدل ۱

$$Z = \max \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} x_{ij} \qquad \text{رابطهٔ ۱)}$$

1. Local Branching
2. Genetic Algorithm with Local Branching

Subject to:

$$\sum_{j=1}^{N} y_j = q \qquad \text{رابطهٔ ۲)}$$

$$\sum_{j=1}^{N} x_{ij} = 1 \qquad \forall\, i = 1, \dots, N \qquad \text{رابطهٔ ۳)}$$

$$x_{ij} \leq y_j \qquad \forall\, i, j = 1, \dots, N \qquad \text{رابطهٔ ۴)}$$

$$x_{ij}, y_j = 0,1 \qquad \forall\, i, j = 1, \dots, N \qquad \text{رابطهٔ ۵)}$$

در مدل فوق، y_j یک متغیر صفر و یک است و اگر دارایی j در سبد ردیابی باشد، مقدار ۱ به خود می‌گیرد و در غیر این صورت صفر است. x_{ij} نیز یک متغیر صفر و یک است که اگر دارایی j نماینده دارایی i باشد برابر با ۱ است و در غیر این صورت مقدار صفر می‌گیرد. رابطهٔ ۱ تابع هدف است و میزان شباهت[1] سبد انتخابی با باقی سهم‌ها را بیشینه می‌کند. رابطهٔ ۲، تعداد سهم‌های انتخابی در سبد ردیاب را محدود می‌کند. رابطهٔ ۳، از اینکه سهم i فقط یک نماینده در پورتفو داشته باشد، اطمینان حاصل می‌کند. بر اساس رابطهٔ ۴، سهم i تنها زمانی می‌تواند نمایندهٔ سهم j باشد که سهم j در سبد ردیاب باشد. آخرین معادله نیز دامنه متغیرهای تصمیم را معین می‌کند. این مدل تنها قادر به ارائه ترکیب سهم‌های بهینه است و توانایی مشخص کردن وزن آن‌ها را ندارد. برای تعیین وزن سهام سبد ردیاب می‌توان از نسبت ارزش بازار سهم‌های انتخاب‌شده استفاده کرد.

این رویکرد دارای چندین اشکال جدی است. اولین موردی که می‌تواند کاملاً واضح باشد، استفاده از هم‌بستگی بین دارایی‌ها به‌عنوان معیار اندازه‌گیری شباهت است که می‌تواند باعث ایجاد مشکل شود؛ زیرا هم‌بستگی بین دارایی‌ها می‌تواند اختلال[2] داشته باشد و در شرایط مختلف بازار تغییر کند. انجام فرایند بهینه‌سازی با استفاده از داده‌های گذشته، برای تخمین این دسته از پارامترها و سپس آزمایش آن روی داده‌های جدید، می‌تواند نتایج بسیار نامطلوبی را به همراه داشته باشد. برای حل این مشکل چن و کوون (۲۰۱۲) ساختار استوار شده مدل ۱ را به‌صورت زیر پیشنهاد کردند:

مدل ۲

$$\mathrm{Z} = \max \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\rho_{ij} x_{ij} - v_{ij}\right) - \Gamma\theta \qquad \text{رابطهٔ ۶)}$$

Subject to:
(۲)(۳)(۴)(۵)

$$v_{ij} + \theta \geq d_{ij} x_{ij} \qquad \forall\, i, j = 1, \dots, N \qquad \text{رابطهٔ ۷)}$$

1. Similarity
2. Noise

رابطهٔ ۸) $$\theta, v_{ij} \geq 0 \qquad \forall\, i,j = 1, \dots, N$$

در مدل فوق، ρ_{ij} مقدار اسمی هم‌بستگی جفت سهم (ارزش انتظاری) و d_{ij} مقدار انحراف معیار مثبت آن از مقدار اسمی است. Γ بیانگر تعداد پارامترهایی است که تحت تأثیر عدم قطعیت قرار می‌گیرند. مدل ۲، مدل استوار مسئله برنامه‌ریزی خطی مدل ۱ است. مدل ۲ قادر است تا حدی نقض‌های مدل ۱ را پوشش دهد؛ اما در نظر نگرفتن اوزان سهام انتخابی باعث ناتوانی در درنظرگرفتن محدودیت‌های دنیای واقعی و درنتیجه مدل‌سازی نه‌چندان دقیق می‌شود. لذا دستهٔ دوم مدل‌های حاضر در مرور ادبیات موضوع به‌صورت زیر است:

فرض کنید که برای T دوره داده وجود دارد و شما می‌خواهید یک سبدی از سهام پیدا کنید که کمترین اختلاف تاریخی را در این دوره با شاخص داشته باشد و محدودیت‌های موجود را نیز لحاظ کند. به‌دلیل اینکه این مدل وزن سهم‌ها را نیز در نظر می‌گیرد اضافه کردن محدودیت‌های دنیای واقعی مانند محدودیت حداقل و حداکثر مقدار نگهداری، کارمزد تراکنش‌ها، معیارهای ریسک پورتفوی از قبیل واریانس، ارزش در معرض ریسک و غیره روی این مدل راحت است.

مدل ۳

رابطهٔ ۹) $$Z = \min \frac{1}{T} \sum_{t=1}^{T} \left| I_t - \sum_{i=1}^{N} r_{it} w_i \right|$$

رابطهٔ ۱۰) $$\sum_{i=1}^{N} w_i = 1$$

رابطهٔ ۱۱) $$\sum_{i=1}^{N} z_i = K$$

رابطهٔ ۱۲) $$\epsilon_i z_i \leq w_i \leq \delta_i z_i \qquad \forall\, i,j = 1, \dots, N$$

رابطهٔ ۱۳) $$w_i \geq 0, z_i = 0,1 \qquad \forall\, i,j = 1, \dots, N$$

در مدل ۳، w_i وزن سهم i؛ z_i متغیر تصمیم‌گیری صفر و یک است که در صورتی مقدار می‌گیرد که سهم i در سبد ردیاب باشد و در غیر این صورت صفر است. I_t بازده شاخص و r_{it} بازده سهم i در زمان t است. رابطهٔ ۹، میانگین قدر مطلق انحرافات سبد را با شاخص کمینه می‌کند. رابطهٔ ۱۰ از اینکه تمام بودجه صرف سرمایه‌گذاری شود اطمینان حاصل می‌کند. رابطهٔ ۱۱، تعداد سهم‌ها انتخابی در سبد ردیاب را محدود می‌کند و رابطهٔ ۱۲، حداقل و حداکثر مقدار مجاز نگهداری هر سهم را مشخص می‌کند. آخرین معادله نیز دامنه متغیر تصمیم را تعیین می‌کند.

شایان ذکر است که میانگین قدر مطلق انحراف‌ها، معیاری معروف در ادبیات موضوع است؛ اما استفاده از باقی معیارهای فاصله مانند میانگین مربعات انحراف‌ها یا میانگین مجذور مربعات خطا نیز، معیارهای محبوبی در ادبیات ردیابی

شاخص هستند. ویژگی برتر میانگین قدر مطلق انحرافات نسبت به باقی معیارها خصلت خطی بودن آن است. ساختار خطی شده مدل ۳ را می‌تواند به شکل زیر در نظر گرفت.

$$Z = \min \frac{1}{T}\sum_{t=1}^{T}(u_t + d_t)$$ رابطهٔ ۱۴)

Subject to:
(۱۰)(۱۱)(۱۲)(۱۳)

$$u_t - d_t = I_t - \sum_{i=1}^{N} r_{it} w_i \qquad \forall\, t = 1, \dots, T$$ رابطهٔ ۱۵)

$$u_t, d_t \geq 0$$ رابطهٔ ۱۶)

مدل برنامه‌ریزی استوار خطی پیشنهادی

مدل ۴ مدل ریاضی استوار خطی پیشنهادی این مقاله برای ردیابی شاخص است.

مدل ٤

$$Z = min \frac{\xi}{T}\sum_{t=1}^{T}(u_{t+}d_t) - \frac{\sum_{i=1}^{N}\sum_{i=1}^{N}(\rho_{ij}x_{ij} - v_{ij}) - \Gamma\theta}{\psi}$$ رابطهٔ ۱۷)

Subject to:
(۳)(۵)(۷)(۸)(۱۰)(۱۱)(۱۲)(۱۳)(۱۵)(۱۶)

$$x_{ij} \leq z_j \qquad \forall\, i, j = 1, \dots, N$$ رابطهٔ ۱۸)

که در آن ψ مقدار تابع هدف مدل ۱ و ξ یک مقدار ثابت است. این دو پارامتر برای تنظیم کردن مقیاس دو تابع هدف طراحی‌شده است تا هر دو عبارت بین ۰ و ۱ باشند. برخلاف باقی مدل‌های استوار در ادبیات ردیابی شاخص که وزن‌دهی سبد با توجه به ارزش بازار سهم‌های انتخابی صورت می‌گیرد مدل پیشنهادی قادر به پیدا کردن سبد K سهمی با در نظر گرفتن عدم قطعیت و محدودیت‌های پیچیده‌تر دنیای واقعی به همراه اوزان بهینه آن‌ها است. از آنجایی که میزان هم‌بستگی سهام با یکدیگر به شرایط مختلف بازار بستگی دارد، استفاده از داده‌های درون نمونه نمی‌تواند لزوماً به تخمین مناسبی از هم‌بستگی سهام بینجامد. به‌عنوان نمونه دو سهمی که هم‌بستگی زیادی در شرایط عادی بازار دارند می‌توانند هم‌بستگی بسیار کمتری در شرایط غیرعادی از خود نشان دهند یا برعکس. قدم بعدی تعیین میزان انحراف از مقدار اسمی مقادیر هم‌بستگی سهام است. برای تحقق این هدف بازه‌های داخل‌نمونه به بازه‌های مساوی تقسیم شده و میزان هم‌بستگی سهام در داخل این بازه‌ها باهم سنجیده شده است. در مرحله بعد میزان انحراف معیار برای هم‌بستگی هر سهم به دست آورده شده است.

الگوریتم‌ها

در این قسمت رویکرد پیشنهادی ابتکاری برای حل مدل جدید ارائه‌شده مورد بحث قرار می‌گیرد. همان‌طور که قبلاً اشاره شد، مسائل ردیابی شاخص در گروه مسائل NP-hard طبقه‌بندی می‌شوند. بنابراین دستیابی به راه‌حل بهینه در یک بازه زمانی قابل‌قبول غیرممکن می‌شود. برای رفع این مشکل، در این پژوهش یک الگوریتم ابتکاری جدید بر اساس الگوریتم‌های تکاملی و انشعاب محلی توسعه داده شده است که به‌طور کامل در بخش‌های بعدی مورد بحث قرار خواهد گرفت. الگوریتم ژنتیک که یک الگوریتم تکاملی است به‌عنوان الگوریتم جست‌وجوی اصلی در رویکرد پیشنهادی این پژوهش عمل می‌کند. بعد از جست‌وجو چندین تکرار در صورت رعایت برخی شرایط مشخص‌شده، الگوریتم انشعاب محلی روی بهترین جواب پیداشده فعال می‌شود. در این مقاله برای بررسی عملکرد الگوریتم ابتکاری GALB در مقابل حل‌کننده‌های دقیق تجاری از داده‌های بازار بورس تهران و داده‌های بازار هنگ‌سنگ[1] از کتابخانه OR-library (بیزلی، ۱۹۹۰) استفاده شده است.

الگوریتم ژنتیک

الگوریتم ژنتیک یک الگوریتم حریصانه[2] بهینه‌سازی الهام گرفته‌شده از طبیعت است که از مفهوم انتخاب طبیعی برای پیدا کردن بهترین جواب استفاده می‌کند. ایده اصلی الگوریتم ایجاد جمعیتی اولیه از جواب‌های کاندید و ایجاد جواب‌های جدید با استفاده از عملگرهای[3] مختلف و سپس حفظ بهترین دسته از جواب‌ها و حذف جواب‌های بدتر و تکرار این فرایند است. از آنجایی‌که GA یک تکنیک بهینه‌سازی تقریبی است، با احتمال بالا می‌توان گفت که جواب نهایی یافته شده توسط الگوریتم ژنتیک بهینه محلی بوده و قادر به همگرایی به جواب بهینه مسئله را ندارد. الگوریتم ۱ یک شبه کد برای GA ساده را نشان می‌دهد:

الگوریتم۱: الگوریتم ژنتیک

یک جمعیت اولیه از جواب‌های مسئله به اندازهٔ p ایجاد کن؛
تا زمانی که شرط توقف محقق نشد:
  با استفاده از اپراتور تقاطع تعدادی جواب جدید ایجاد کن؛
  با استفاده از اپراتور جهش تعدادی جواب جدید ایجاد کن؛
  جواب‌های ایجادشده را به سبد جواب‌های اولیه اضافه کن؛
  میزان تابع هدف را برای تمام جواب‌ها را محاسبه کن؛
  p عدد از بهترین جواب‌ها را حفظ و مابقی را حذف کن؛
بهترین جواب را به‌عنوان جواب مسئله انتخاب کن؛

1. Hang Seng Index
2. Greedy Algorithm
3. Operators

اصلی‌ترین چالش برای پیاده‌سازی الگوریتم ژنتیک در این دسته از مسائل محاسبه میزان تابع هدف برای آن‌هاست. یک راه استفاده از حل‌کننده‌های تجاری در قسمت میزان تعیین تابع الگوریتم است؛ اما این روش یک عیب بسیار واضح دارد و آن، زمان و نیاز محاسباتی مضاعفی است که این حل‌کننده‌ها به الگوریتم اضافه می‌کنند. گزینه دیگر روی میز، محاسبه مستقیم تابع هدف و تخصیص پنالتی برای نقض کردن محدودیت‌ها است. نکته منفی این روش این است که تعداد تکرارها یا زمان بیشتری باید به الگوریتم اختصاص داده شود تا بتواند یک جواب خوب یا حتی یک جواب موجه پیدا کند. موارد گفته‌شده به تضعیف عملکرد الگوریتم منجر می‌شود. علاوه‌بر موارد گفته‌شده در طراحی یک روش برای سنجش میزان تابع هدف برای الگوریتم، باید در نظر داشت که برتری نسبی جواب‌ها حفظ شود دقیقاً به‌مانند زمانی که یک حل‌کنندهٔ تجاری میزان تابع هدف جواب‌ها را به‌دست می‌آورد. فرض کنید تابع هدف سبدهای الف و ب توسط یک حل‌کنندهٔ تجاری تعیین‌شده است. اگر تابع هدف الف از ب بهتر باشد روش طراحی‌شده نیز باید این برتری را نشان دهد. هرچند شانس زیادی وجود دارد که در روش دوم با این مشکل مواجه شویم. اگر بازاری با N سهم در نظر بگیرید، هر جواب مسئله در این مقاله با یک بردار N عضوی نمایش داده می‌شود که هر عضو نشان‌دهنده وزن هر سهم در سبد است. الگوریتم ۲ رویکرد این پژوهش در طراحی روشی برای تعیین میزان تابع هدف مسئله که به‌اندازه حل‌کننده‌های تجاری زمان‌بر نبود و با مشکلات برتری مواجه نشود را توضیح می‌دهد.

الگوریتم۲: تابع هدف

P جمعیت اولیه جواب‌ها است؛

W_{ij} برابر با $\rho_{ij} - d_{ij}$ است؛

N سهم‌های بازار است؛

$D_i = \{d_i | d_i = argmax\ P_i, |d_i| = K\} \forall i \in P$ // مرحله دکدینگ؛

$E = \{N\} - D$؛

برای تمام i های در D:

$R = 0$؛

$temp = \{\}$؛

برای تمامی $e \in E$:

$M_x = \max(W_{eD_i})$ // بیشنیه مقدار e و عضوهای در D_i؛

$R = R + M_x$؛

اگر: $\Gamma = N$

ادامه بده؛

در غیر این صورت:

$y = \{y \mid W_{ey} = M_x\}$؛

d_{ey} را به $temp$ اضافه کن؛

اگر: $\Gamma \neq N$

اعضای $temp$ را به‌صورت تصاعدی مرتب کن و به $temp_S$ اضافه کن؛

$$R = R + \sum_{i=1}^{|temp|+K-\Gamma} temp_{S_i}$$

$$F_i = \frac{\xi}{T}\sum_{t=1}^{T}\left|I_t - \sum_{i=1}^{N} r_{it}w_i\right| - \frac{R}{\psi}$$

در مرحله اول شکل نمایش کروموزوم‌ها به صورت شماره هر سهم تغییر کرده تا سهم‌های انتخابی هر جواب مشخص شود. برای محاسبه عبارت دوم تابع هدف در ابتدا بدترین مقادیر ممکن در نظر گرفته می‌شود. بیشتر مقدار W_{ij} برای دارایی که در سبد ردیاب نیست و تمامی دارایی‌هایی که در سبد سهام هستند محاسبه می شود. در گام بعدی، اگر قرار بر این است که تمامی پارامترها بدترین مقدار خود را اختیار نکنند، انحراف معیاری که کمترین مقدار d_{ij} را دارند به مقادیر اسمی اضافه می‌شوند. با این روش پیشنهادی و مکانیزم تعمیر جواب‌های ناموجه، ابزاری قدرتمند و کارا جهت مقایسه جواب‌ها برای الگوریتم خواهیم داشت.

انشعاب محلی

به‌صورت ساده، ایده انشعاب محلی جست‌وجو در نزدیکی یک جواب (که در این مقاله بهترین جواب پیدا شده تا تکرار mام است)، جهت یافتن جوابی بهتر است. میزان سنجش فاصله دو جواب در این روش با استفاده از فاصله همینگ[1] است. فاصله همینگ بین دو بردار از یک اندازه (صفر و یک) تعداد عضوهایی هستند که دو جواب با هم تفاوت دارند. برای مثال، دو جواب $B = [۱,۱,۰,۰]$ و $A = [۱,۰,۰,۱]$ را در نظر بگیرید. فاصله بین این دو جواب یا همان تعداد تعویض‌های[2] لازم برای ساختن یک جواب از جوابی دیگر است. در این مثال عضو اول و سوم یکی هستند پس به هیچ تعویضی نیاز نیست ولی برای ساختن B از Aعضو دوم باید به یک و آخرین عضو باید به صفر تغییر کند. چون دو تغییر نیاز است فاصله بین این دو جواب ۲ است. همین منطق ساده نیز در جست‌وجو محلی استفاده می‌شود. این موضوع را می‌توان به صورت ریاضی در مدل ۵ توضیح داد.

مدل۵

$$Z = \min \frac{\xi}{T}\sum_{t=1}^{T}(u_t + d_t) - \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\{\rho_{ij}x_{ij} - v_{ij}\} - \Gamma\theta}{\psi}$$ رابطهٔ ۱۹)

Subject to:
(۳)(۵)(۷)(۸)(۱۰)(۱۱)(۱۲)(۱۳)(۱۵)(۱۶)(۱۸)

1. Hamming distance
2. flips

$$a \le \sum_{j \in J} (1 - z_j) + \sum_{i \in I/J} z_i \le b$$ رابطهٔ ۲۰)

$$\sum_{i \in N/I} z_i = 0$$ رابطهٔ ۲۱)

مدل ۵ معادل مدل ۴ است؛ ولی با دو محدودیت بیشتر که برای انشعاب محلی استفاده می‌شود. رابطهٔ ۲۰، تعداد تغییرهای لازم در بهترین جواب یافته شده J به همراه در نظر گرفتن سهم‌های I را می‌شمارد. رابطهٔ ۲۱، از صفر بودن متغیر تصمیم‌گیری سهم‌هایی که در مجموعه I هستند اطمینان حاصل می‌کند. برای ایجاد مجموعه برای انشعاب محلی (S_s) مدل ۴ ریلکس‌شده[1] حل می‌شود. در گام بعدی، درصد مشخصی از اندازهٔ بازار برای انتخاب مقداری که در ادامه توضیح داده خواهد شد تعیین می‌شود. تعداد L عدد از بیشترین دارایی‌هایی که بیشترین مقدار z و w دارند انتخاب می‌شوند. ماتریس بدترین مقادیر هم‌بستگی بین سهم‌ها را در نظر بگیرید. L عدد از بزرگ‌ترین مقادیر میانگین هر سطر ماتریس محاسبه‌شده و نگهداری می‌شود. بنابراین، یک مجموعه‌ای با پتانسیل بالا ایجاد می‌شود که در فرایند انشعاب محلی می‌تواند نقش مهمی ایفا کند. نکته حائز اهمیت دیگر این است که اگر قرار بر این است که سبد ردیاب اندازه مشخص داشته و به یک عدد مانند ۱۰ محدود شود مقادیر a, b باید همواره زوج باشند تا جواب‌های ناموجه تولید نشود. همچنین حداقل ۲ تغییر لازم است تا سبد ردیاب موجه تولید شود. علاوه‌براین حداکثر تعداد تغییرات مجاز برای جست‌وجو محلی جهت آن که جوابی کاملاً بی‌ربط به جواب اصلی پیدا نشود، برابر با $2(K - 1)$ است. فرض کنید اندازهٔ یک سبد ردیاب ۳ باشد. تغییر ۴ بیت (برابر قرار دادن دو ۱ از سبد اصلی به مقدار صفر و دو دارایی جدید به ۱) بیشینه فاصله‌ای است که ما حق جست‌وجو در آن را داریم؛ چرا که ۵ تغییر سبد را ناموجه می‌کند و برای ۶ تغییر باید تمامی یک‌های سبد به صفر تبدیل شود. در این تحقیق، به‌جای داشتن مجموعه‌های بزرگ و جست‌وجو در فواصل خیلی کم، مجموعه‌های قوی‌تر و با پتانسیل بیشتر را انتخاب می‌کنیم و در فواصل اندکی بیشتر به جست‌وجو می‌پردازیم. الگوریتم ۳، شبه کد الگوریتم ابتکاری انشعاب محلی پیشنهادی این پژوهش است و الگوریتم ۴ شبه کد الگوریتم پیشنهادی GALB را نمایش می‌دهد.

الگوریتم۳: الگوریتم ابتکاری انشعاب محلی

rep تعداد تکرارهایی است که بهبودی در جواب فرایند انشعاب محلی ایجاد نشده است؛

$c_1, c_2 \in \mathbb{R}, (0 \le c_1, c_2 \le 1), c_2 \ge c_1, c_1 \le 0.5, c_2 \ge 0.5$ ؛

$best$ بهترین جواب یافت شده در نسل؛

S_s مجموعه کوچک‌تر انتخاب‌شده جهت انجام انشعاب محلی؛

$\alpha, \beta \in \mathbb{Z}^+, \beta \ge \alpha$؛

$D = \{d \mid d = argmax\ P_i, |d| = K, i \in best\}$؛

1. Relaxation of model 4

$S_N = \{N\} - D$؛

$c = c_1$؛

اگر: $rep \geq \alpha$

$c = c_2$؛

اگر: $rep \geq \beta$

$r = max(r + r, K - 1)$ ؛

اگر: $random(0,1) \leq c$ یا $|S_s - D| \leq r$

$I = D \cup \{j \mid j \in N - D, |j| = r\}$؛

در غیر این صورت:

$I = D \cup \{j \mid j \in S_s - D, |j| = r\}$؛

مدل ۵ را با محدودیت زمانی مشخص حل کن $(I = I, J = D)$

الگوریتم۴: الگوریتم ژنتیک با انشعاب محلی

LB تعداد تکرارهایی است که باید از الگوریتم ژنتیک بگذرد تا یک‌بار الگوریتم انشعاب محلی اجرا شود.

$G = 0$

B یک عدد بسیار بزرگ است؛

تا زمانی که شرط توقف محقق نشد:

$G = G + 1$؛

فرزندان تولیدشده از عملگر تقاطع $= C$؛

فرزندان تولیدشده از عملگر جهش $= M$؛

$A = P \cup \{C \cup M\}$؛

$\forall i \in P$؛ P_iرا مرتب کن و $N - K$ عدد از اوزان را برابر صفر قرار بده و اوزان را نرمال کن؛

اگر : $G = LB$

بهترین جواب را پیدا کرده و الگوریتم ۳ را اجرا کن؛

اگر تابع هدف جواب یافته شده از B کمتر بود:

جواب را به A اضافه کن؛

$G = 0$؛

بهترین جواب‌های پیدا شده به اندازهٔ P که از مجموعه A انتخاب‌شده است ; $= P$؛

بهترین جواب مجموعه $b = P$؛

اگر: $b \leq B$

$B = b$؛

بهترین جواب را نشان بده؛

استفاده از الگوریتم انشعاب محلی در الگوریتم ژنتیک باعث استفاده از قدرت هر دو الگوریتم می‌شود. از یک طرف ژنتیک به قدرت جست‌وجو جهانی[1] معروف است و قادر به یافتن تقریبی باکیفیت در زمان کوتاه است. از سوی دیگر انشعاب محلی یک روش جست‌وجوی محلی[2] بوده که می‌تواند فرایند جست‌وجو را تقویت کند. در این پژوهش با اعمال جست‌وجو محلی در فواصل مشخص بعد از تکرارهای الگوریتم ژنتیک بین جست‌وجو محلی و جهانی تعادل ایجاد شده است.

یافته‌های پژوهش

در این بخش نتایج مقایسه عملکرد الگوریتم ابتکاری ارائه‌شده را با حل‌کننده‌های تجاری مقایسه کرده و به بررسی سبدهای انتخابی می‌پردازیم. برای انجام این کار از حل‌کنندهٔ CPLEX به‌عنوان یک معیار برای مقایسه استفاده شده است. الگوریتم‌ها بر روی داده‌های روزانه قیمت سهام شرکت‌های موجود در شاخص کل بورس اوراق بهادار تهران از تاریخ ۲۹ آذر ۱۴۰۰ تا ۸ فروردین ۱۴۰۳ و بازار هنگ‌سنگ از معیار پایه[3] OR-library اجرا شده است. شکل ۱ نمای داده‌های درون‌نمونه و خارج از نمونه هر دو بازار را نشان می‌دهد. در این بخش ابتدا عملکرد مدل ریاضی پیشنهادی این مقاله با مدل‌های ریاضی موجود (مدل ۳) مقایسه می‌شود. در قسمت بعدی عملکرد GALB بررسی شده، سپس به بررسی سبدهای سهام پیشنهادشده پرداخته شده است و در بخش آخر نیز این رویکرد با یک سبد معیار مقایسه شده است. اگرچه در ادامه نشان خواهیم داد که GALB قابلیت همگرایی به جواب بهینه را دارد اما برای ایجاد مقایسه‌ای منصفانه و کنار گذاشتن احتمالات، مدل پیشنهادی و مدل ۳ با حل‌کنندهٔ CPLEX حل می‌شوند. در تمامی آزمایش‌های این مقاله تعداد دارایی‌های در سبدها ۱۰، کف و سقف خرید سهام به ترتیب ۱ و ۹۰ درصد و ξ برابر با ۱۰۰ تعریف‌شده است. برای مقایسه ۵ معیار مختلف در نظر گرفته شده است. که به ترتیب توضیح داده خواهد شد.

ρ: میزان هم‌بستگی بازده سبد ردیاب با بازده شاخص

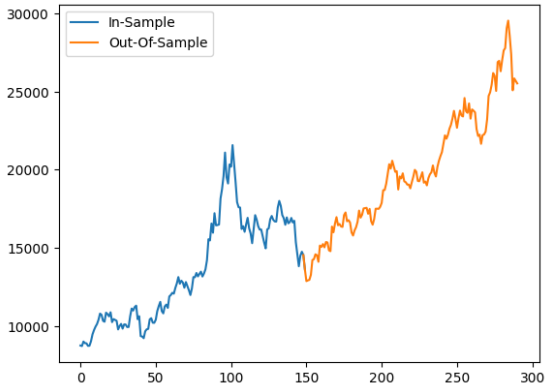
$$Mean\ Absolute\ Deviation\ (MAD) = \frac{1}{T}\sum_{t=1}^{T}\left|I_t - \sum_{i=1}^{N} r_{it}w_i\right| \quad \text{رابطهٔ ۲۲)}$$

$$Root\ Mean\ Squared\ Error\ (RMSE) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(I_t - \sum_{i=1}^{N} r_{it}w_i)^2} \quad \text{رابطهٔ ۲۳)}$$

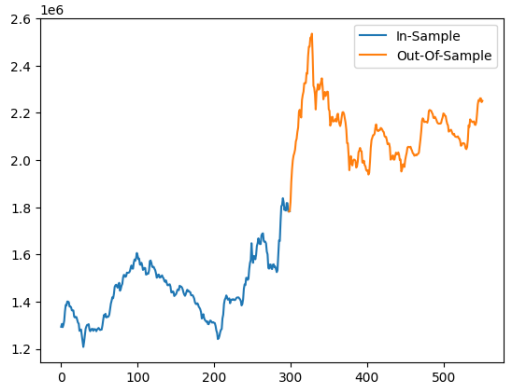
STD : انحراف معیار خطای ردیابی

$Beta$: بتا سبد ردیاب با شاخص

1. Global Search (Exploration)
2. Local Search
3. Benchmark



هنگ سنگ



بورس تهران

شکل ۱. نمای داده‌های داخل و خارج نمونه بازارهای استفاده شده در این تحقیق

اگرچه مقادیر کمتر $MAD, RMSE, STD$ و مقادیر بیشتر ρ جذاب‌تر هستند، مقادیر نزدیک‌تر به یک برای $Beta$ مورد توجه‌تر است. نتایج برای درون و خارج نمونه نمایش داده شده است که به ترتیب در جداول با (I) و (O) قبل از هر معیار مشخص شده است. به تمامی مدل‌ها به مدت ۱۸۰۰ ثانیه فرصت داده شده است تا جواب بهینه را پیدا کنند. جدول ۱ نتایج سبد پیشنهادی CPLEX برای مدل۳ و مدل ریاضی پیشنهادی این مقاله روی معیارهای معرفی‌شده را نشان می‌دهد.

جدول ۱. مقایسه مدل پیشنهادی با مدل۳ بر روی داده‌های بورس اوراق بهادار تهران

معیار مقایسه	مدل پیشنهادی	مدل ۳
(I)ρ	۰/۹۶۲۶	۰/۹۷۹۹
(I)MAD	۰/۰۰۲۲۳۰	۰/۰۰۱۶۹۶
(I)RMSE	۰/۰۰۳۲۰۹	۰/۰۰۲۳۰۸
(I)STD	۰/۰۰۳۲۱۳	۰/۰۰۲۳۱۵
(I)BETA	۰/۹۸۸۱	۰/۹۸۲۸
(O)ρ	۰/۹۴۳۲	۰/۹۰۷۹
(O) MAD	۰/۰۰۲۸۶۹	۰/۰۰۳۲۱۲
(O)RMSE	۰/۰۰۴۰۱۲	۰/۰۰۴۸۱۳
(O)STD	۰/۰۰۳۹۹۲	۰/۰۰۴۸۳۲
(O)BETA	۰/۹۸۲۸	۰/۸۲۴۵

همان‌طور که از جدول بالا مشخص است، سبد ردیاب هر دو مدل نتایجی رقابتی و نزدیک به همدیگر روی داده‌های درون نمونه دارند؛ ولی به غیر از بتا، نتایج مدل۳ از مدل پیشنهادی در داده‌های درون نمونه بهتر است؛ اما در

داده‌های خارج از نمونه، مدل ریاضی پیشنهادی با اختلاف بیشتری عملکرد بهتر از مدل ۳ بـه نمـایش مـی‌گـذارد. سـبد ردیاب مدل ۳ نتایج بسیار ضعیف‌تری نسبت به مدل پیشنهادی ما در تمامی معیارها در داده‌های خارج از نمونـه داشـته و نتایج به‌طور کامل نشان‌دهنده برتری مدل پیشنهادی این مقاله بر روی داده‌های خارج از نمونه است. باید در نظر داشـت که برای این مقایسه در مدل پیشنهادی عدم قطعیت پارامترها در این حالت در نظر گرفته نشده است. شایان ذکـر اسـت که هر دو مدل یک سبد ردیاب برای بازار هنگ‌سنگ پیدا کرده و قبل از تمام شدن محـدودیت زمـانی بـه فراینـد حـل خاتمه داده‌اند، لذا از نمایش نتایج برای این بازار جهت مقایسه عملکرد مدل‌ها خودداری شده است.

در قسمت دوم نتایج، هدف بررسی عملکرد الگوریتم ابتکاری پیشنهادی در مقابل حل‌کنندهٔ تجاری CPLEX است. برای مقایسه تنظیمات مختلف بدون حالت استوار و حالت استوار ۵۰، ۷۵ و۹۰ درصد در نظر گرفته شده اسـت. همچنـین برای پیدا کردن انحراف معیار هم‌بستگی‌ها از بازه‌های ۱۰تایی استفاده شده است. برای ایجاد یک مقایسه منصفانه میزان زمان لازم برای حل مدل برای الگوریتم ابتکاری برای بازار هنگ سنگ ۱۲۰ ثانیه و بـرای بـورس تهـران ۲۰۰ ثانیـه و برای حل‌کنندهٔ CPLEX این مقدار ۱۸۰۰ ثانیه در نظر گرفته شده است. در این قسمت علاوه‌بر معیارهای معرفی‌شده دو معیار جدید بهترین مقدار تابع هدف (OFV) و انحراف معیار نتایج ثبت شده برای OFV یا ($\mathrm{OFV}(\sigma)$) که برای ۵ اجـرا مختلف محاسبه شده فقط برای داده‌های درون نمونه نیز استفاده می‌شود. از آنجایی که مسئله مورد مطالعه کمینه‌سـازی است، مقادیر کمتر تابع هدف و انحراف معیار تابع هدف جذاب‌ترند. برای داده‌های درون نمونه انتظار مـی‌رود کـه مقـدار تابع هدف هم‌بستگی کامل با معیارهای اندازه‌گیری عملکرد در داده‌های درون نمونه داشته باشد؛ به این معنـی کـه تـابع هدف کمتر مقادیر مطلوب‌تر از نظر معیارهای سنجش خطای ردیابی در داده‌های درون نمونه داشته باشد.

اما همان‌طور که بیزلی و همکاران (۲۰۰۳) اشاره کردند، نتایج درون و خارج نمونه با همدیگر هم‌بستگی دارنـد؛ امـا نتایج خوب درون نمونه همواره به نتایج خوب خارج از نمونه ختم نمی‌شود. با وجود این، انتظـار ایـن را داریـم کـه سـبد ردیابی که در داخل نمونه بهترین عملکرد را دارد، در خارج نمونه نیز این عملکرد را حفظ کند و معیـار مـا بـرای انتخـاب سبد برای ردیابی شاخص در آینده، بهترین نتیجه به‌دست آمده در گذشته است. جدول ۲ نتایج مقایسه عملکرد الگـوریتم پیشنهادی GALB را برای داده‌های درون و خارج نمونه روی داده‌های بورس تهران نشان می‌دهد.

همان‌طور که از جدول ۲ نیز مشخص است، الگوریتم ابتکاری ارائه شده در این مقاله موفق به دستیابی به نتـایجی بهتری نسبت به CPLEX در تنظیمات مختلف شده است. در حالتی که عدم قطعیت در نظر گرفته نشـده باشـد، اگرچـه نتایج درون نمونه نتایجی رقابتی و نزدیک است ولی نتایج سبد ردیاب الگوریتم پیشـنهادی در داده‌هـای خـارج از نمونـه کاملاً بهتر و اختلاف بیشتری نسبت به CPLEX نشان داده است. در دو حالت عدم قطعیت ۵۰ و ۷۵ درصد نیز الگوریتم پیشنهادی هم در داده‌های داخل و هم خارج نمونه موفق به شکست CPLEX با اختلاف ملموس شده اسـت. در حـالتی که ۹۰ درصد پارامترها بدترین مقدار خود را گرفته‌اند، اگرچه CPLEX نتایجی بهتر ولی رقـابتی ماننـد حالـت اول ارائـه کرده، اما این عملکرد در داده‌های خارج از نمونه ادامه پیدا نکرده و سبد ردیاب GALB نتایج بهتری در تمامی معیارهای خارج نمونه به غیر از بتا کسب کرده است.

جدول ۲. نتایج مقایسه GALB و CPLEX با داده‌های بورس تهران

CPLEX Γ = 90%	GALB Γ = 90%	CPLEX Γ = 75%	GALB Γ = 75%	CPELX Γ = 50%	GALB Γ = 50%	CPELX Γ = 0	GALB Γ = 0	measure
	−۰/۷۴۸۷۰	−۰/۷۵۶۸۰	−۰/۷۶۹۱۰	−۰/۷۵۶۹۰	−۰/۷۶۱۴۰	−۰/۷۵۶۹۰	−۰/۷۶۶۱۰	(I)OFV
	۰/۰۰۶۳۱۹	–	۰/۰۱۳۲۴	–	۰/۰۱۷۳۵	–	۰/۰۱۲۷۷	(I)OFV(σ)
	۰/۹۶۴۹۰	۰/۹۵۶۶۰	۰/۹۶۳۵۰	۰/۹۶۲۶۰	۰/۹۶۸۶۰	۰/۹۶۲۶۰	۰/۹۷۰۸	(I)ρ
	۰/۰۰۲۲۵	۰/۰۰۲۳۵	۰/۰۰۲۲۲	۰/۰۰۲۲۳	۰/۰۰۲۱۴	۰/۰۰۲۲۳	۰/۰۰۲۲۲	(I)MAD
	۰/۰۰۳۰۶	۰/۰۰۳۴۲	۰/۰۰۳۱۳	۰/۰۰۳۲۰	۰/۰۰۲۹۵	۰/۰۰۳۲۰	۰/۰۰۳۰۱	(I)RMSE
	۰/۰۰۳۰۷	۰/۰۰۳۴۲	۰/۰۰۳۱۲	۰/۰۰۳۲۱	۰/۰۰۲۹۶	۰/۰۰۳۲۱	۰/۰۰۳۰۲	(I)STD
	۰/۹۶۷۹۰	۰/۹۶۹۰۰	۰/۹۷۰۷۰	۰/۹۸۸۱۰	۰/۱۰۰۰۰	۰/۹۸۸۱۰	۱/۰۴۵۸۰	(I)BETA
	۰/۹۴۸۱۰	۰/۹۲۳۳۰	۰/۹۴۵۳۰	۰/۹۴۳۲۰	۰/۹۶۰۹۰	۰/۹۴۳۲۰	۰/۹۵۷۸۰	(O) ρ
	۰/۰۰۲۵۹	۰/۰۰۳۴۶	۰/۰۰۲۵۲	۰/۰۰۲۸۶	۰/۰۰۲۳۶	۰/۰۰۲۸۶	۰/۰۰۲۲۷	(O) MAD
	۰/۰۰۳۷۳	۰/۰۰۴۹۱	۰/۰۰۳۷۷	۰/۰۰۴۰۱	۰/۰۰۳۳۱	۰/۰۰۴۰۱	۰/۰۰۳۳۱	(O)RMSE
	۰/۰۰۳۷۵	۰/۰۰۴۸۹	۰/۰۰۳۷۸	۰/۰۰۳۹۹	۰/۰۰۳۳۲	۰/۰۰۳۹۹	۰/۰۰۳۳۲	(O)STD
	۰/۹۶۶۱۰	۱/۰۱۸۸۵	۰/۹۳۲۷۰	۱/۰۰۲۸۰	۱/۰۰۲۸۰	۱/۰۰۲۸۰	۰/۹۴۱۴۰	(O)BETA

جدول ۳. نتایج مقایسه GALB و CPLEX با داده‌های هنگ سنگ OR-library

CPLEX Γ = 90%	GALB Γ = 90%	CPLEX Γ = 75%	GALB Γ = 75%	CPELX Γ = 50%	GALB Γ = 50%	CPELX Γ = 0	GALB Γ = 0	Measure
−۰/۵۰۳۰۰	−۰/۵۰۹۶۰	−۰/۵۰۳۰۰	−۰/۵۰۶۸۰	−۰/۵۰۳۰۰	−۰/۵۳۰۶۰	−۰/۶۵۹۷۰	−۰/۶۵۹۷۰	(I)OFV
–	۰/۰۰۹۷۰	–	۰/۰۰۴۴۰	–	۰/۰۰۹۹۰	–	۰	(I)OFV(σ)
۰/۹۹۴۱۰	۰/۹۹۵۱۰	۰/۹۹۴۱۰	۰/۹۹۴۸۰	۰/۹۹۴۱۰	۰/۹۹۴۹۰	۰/۹۹۵۱۰	۰/۹۹۵۱۰	(I)ρ
۰/۰۰۳۰۹	۰/۰۰۲۸۱	۰/۰۰۳۰۹	۰/۰۰۲۸۶	۰/۰۰۳۰۹	۰/۰۰۲۸۷	۰/۰۰۲۸۱	۰/۰۰۲۸۱	(I)MAD
۰/۰۰۴۰۶	۰/۰۰۳۷۳	۰/۰۰۴۰۶	۰/۰۰۳۸۴	۰/۰۰۴۰۶	۰/۰۰۳۸۳	۰/۰۰۳۷۳	۰/۰۰۳۷۳	(I)RMSE
۰/۰۰۴۰۷	۰/۰۰۳۷۳	۰/۰۰۴۰۷	۰/۰۰۳۸۴	۰/۰۰۴۰۷	۰/۰۰۳۸۲	۰/۰۰۳۷۳	۰/۰۰۳۷۳	(I)STD
۰/۹۹۱۸۰	۰/۹۹۵۵۰	۰/۹۹۱۸۰	۰/۹۸۹۴۰	۰/۹۹۱۵۰	۰/۹۷۸۳۰	۰/۹۹۵۵۰	۰/۹۹۵۵۰	(I)BETA
۰/۹۸۱۹۰	۰/۹۸۶۲۰	۰/۹۸۱۸۰	۰/۹۸۶۶۰	۰/۹۸۱۸۰	۰/۹۸۵۵۰	۰/۹۸۶۲۰	۰/۹۸۶۲۰	(O) ρ
۰/۰۰۳۸۹	۰/۰۰۳۵۲	۰/۰۰۳۸۹	۰/۰۰۳۳۷	۰/۰۰۳۸۹	۰/۰۰۳۴۰	۰/۰۰۳۵۲	۰/۰۰۳۵۲	(O) MAD
۰/۰۰۵۲۶	۰/۰۰۴۶۸	۰/۰۰۵۲۶	۰/۰۰۴۴۹	۰/۰۰۵۲۶	۰/۰۰۴۶۴	۰/۰۰۴۶۸	۰/۰۰۴۶۸	(O)RMSE
۰/۰۰۵۳۰	۰/۰۰۴۷۱	۰/۰۰۵۳۰	۰/۰۰۴۵۲	۰/۰۰۵۳۰	۰/۰۰۴۶۵	۰/۰۰۴۷۱	۰/۰۰۴۷۱	(O)STD
۱/۰۰۱۶۰	۱/۰۲۰۲۰	۱/۰۰۱۶۰	۰/۹۹۹۵۰	۱/۰۰۱۶۰	۰/۹۸۴۵۰	۱/۰۲۰۲۰	۱/۰۲۰۲۰	(O)BETA

جدول ۳ همین نتایج را بر روی بازار هنگ‌سنگ OR-library نشان می‌دهد. شایان ذکر است که هدف اصلی از آزمایش الگوریتم پیشنهادی و CPLEX بر روی این بازار نشان دادن قدرت الگوریتم (همگرایی به جواب بهینه) در بازارهای کوچک‌تر است. در حالت عدم قطعیت، CPLEX به‌دلیل بزرگ نبودن اندازهٔ بازار، حل را پس از چند دقیقهٔ اول متوقف کرده و به جواب بهینه رسیده است. همان‌طور که مشخص است GALB نیز توانسته در هر ۵ تکرار به جواب بهینه همگرا شود و دقیقا همان سبد ردیابی که CPLEX گزارش کرده را پیدا کند. همچنین در تمامی تنظیمات عدم قطعیت در نظر گرفته الگوریتم ابتکاری پیشنهادی هم در داده‌های داخل و خارج نمونه توانسته نتایجی بهتر از CPLEX ارائه کند. آزمون ناپارامتری ویلکوکسون[1] بر روی تمامی شاخص‌های خارج‌نمونه اعمال شد. نتایج این آزمون نشان داد که تفاوت بین دو روش در سطح اطمینان ۹۰ درصد برای تمامی شاخص‌ها معنادار است. همچنین در سطح اطمینان ۹۵ درصد نیز تفاوت معناداری برای همه شاخص‌ها به جز بتا مشاهده شد.

در این قسمت از نتایج با در نظر گرفتن داده‌های بورس اوراق بهادار تهران، به بررسی سبدهای سهام پیشنهادشده توسط GALB و حل‌کنندهٔ تجاری CPLEX به همراه نمایش گرافیکی عملکرد آن‌ها بر روی داده‌های خارج از نمونه پرداخته شده است. این قسمت از نتایج به ۴ قسمت جداگانه تقسیم می‌شود. اولین قسمت سبدهای پیشنهادشده در حالت عادی و بدون عدم قطعیت و دومین قسمت، حالت قبلی است ولی با در نظر گرفتن عدم قطعیت. در دو قسمت بعدی ما ۱۰ سهم بزرگ بازار از نظر ارزش بازار را برای انتخاب شدن در سبد پیشنهادی ممنوع کرده و مسئله را با سایر سهام موجود در بازار حل کرده‌ایم. نمادهای کنار گذاشته‌شده از محاسبات در حالت‌های سوم و چهارم در جدول ۴ نشان داده شده است. در جدول‌های مربوط به سبدهای ارائه شده سهم‌ها به ترتیب درصد وزنی مرتب شده‌اند، ستون CPLEX مربوط به سبد پیشنهادی این حل‌کنندهٔ و بقیه ستون‌ها مربوط به اجراهای مختلف GALB و جواب‌های پیشنهادی آن است. باید دوباره خاطرنشان کرد رویکرد این مقاله یک رویکردی تصادفی و فضای حل موجه آن بسیار بزرگ است. برای انتخاب سبدی ۱۰ سهمی از میان ۵۰۰ سهم بازار بورس تهران تعداد $\binom{۵۰۰}{۱۰}$ جایگشت که به معنای وجود $e^{۲۰}$ ۲/۴۵ حالت مختلف بدون در نظر گرفتن اوزان سبد وجود دارد. الگوریتم پیشنهادی از بین این جواب‌ها سعی در انتخاب بهترین ترکیب به همراه تخصیص بهترین اوزان ممکن برای هر سهم سبد را دارد که لزوماً منحصربه‌فرد هم نیست. به همین دلیل، سبدهای مختلفی را می‌توان یافت که بتوانند شاخص کل را به‌صورت مناسبی پایش کنند. همچنین پس از جدول هر حالت، تفاوت گرافیکی چند نمونه از سبدها و CPLEX از نظر بازده و بازده تجمعی به نمایش گذاشته شده است. حالت عدم قطعیت تمامی حالت‌ها نیز حالت ۷۵ درصد در نظر گرفته شده است.

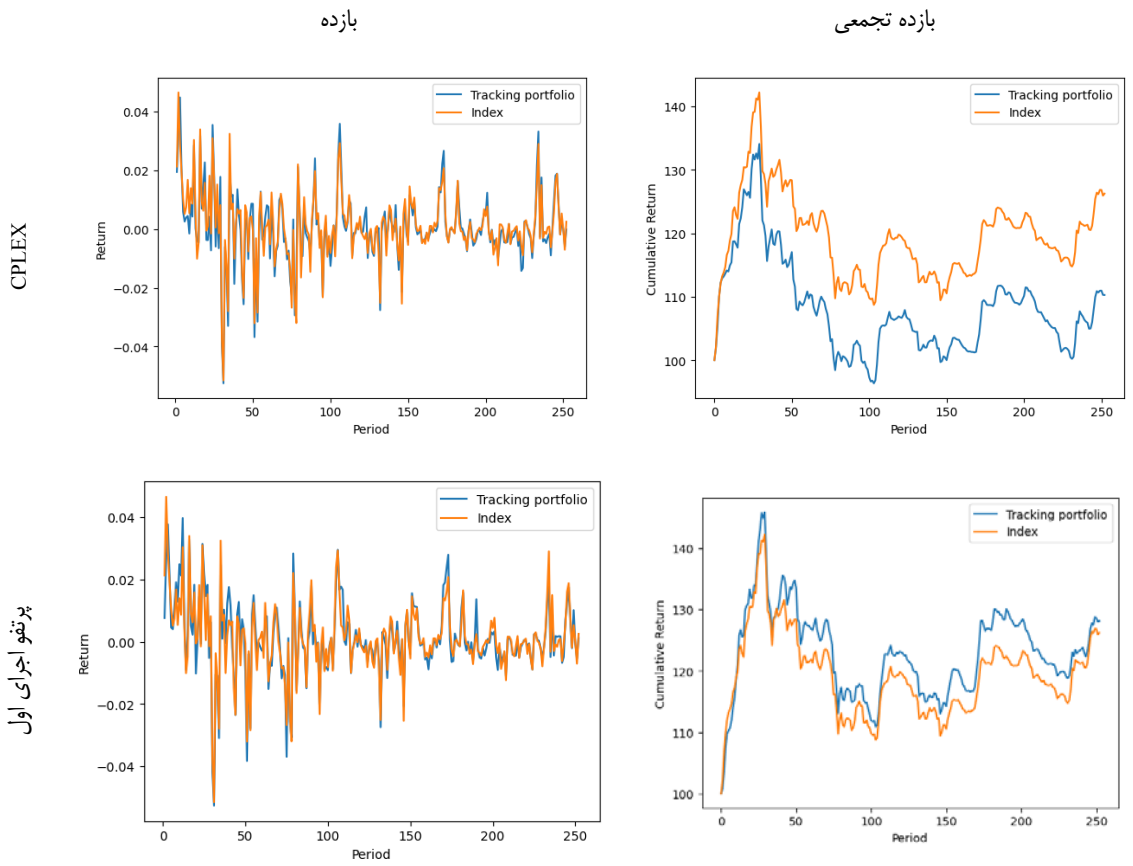
جدول ٤. سهم‌های خارج‌شده از محاسبات در حالت سوم و چهارم این بخش

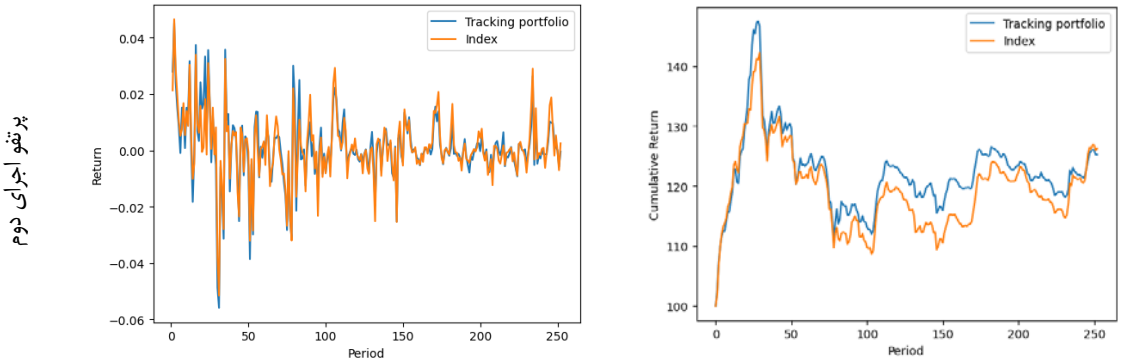
۱	۲	۳	٤	٥	٦	۷	۸	۹	۱۰
فملی	فارس	فولاد	وبملت	شپنا	شستا	وغدیر	کگل	تاپیکو	وامید

1. Wilcoxon signed-rank test

جدول ۵. مقایسه سبد پیشنهادی CPLEX با سبد پیشنهادی رویکرد این پژوهش در حالت عادی بدون عدم قطعیت

CPLEX		اجرای چهارم		اجرای سوم		اجرای دوم		اجرای اول	
٪۱۹/۰	امین	٪۲۱/۷	وغدیر	٪۲۷/۳	فارس	٪۲۰/۹	امین	٪۲۴/۰	اپال
٪۱۶/۰	وغدیر	٪۱۶/۰	فولاد	٪۱۳/۸	سیتا	٪۱۸/۱	وغدیر	٪۱۵/۵	فولاد
٪۱۴/۷	فولاد	٪۱۵/۲	امین	٪۱۳/۳	فولاد	٪۱۶/۳	فولاد	٪۱۸/۷	وغدیر
٪۱۱/۱	پارس	٪۱۴/۵	لوتوس	٪۱۲/۷	وغدیر	٪۹/۵	میدکو	٪۱۰/۶	امین
٪۱۰/۶	وپاسار	٪۱۱/۴	شستا	٪۸/۶	دارو	٪۷/۷	شپنا	٪۹/۰	تیپیکو
٪۸/۱	شبندر	٪۶/۲	شبندر	٪۷/۵	شبندر	٪۶/۷	تیپیکو	٪۷/۳	شبندر
٪۷/۳	تیپیکو	٪۵/۷	وبشهر	٪۷/۰	وبشهر	٪۶/۴	وپاسار	٪۴/۹	وتجارت
٪۵/۴	وخارزم	٪۴/۵	وتجارت	٪۵/۰	وتجارت	٪۵/۰	سفارس	٪۴/۸	چدن
٪۴/۶	خساپا	٪۲/۵	دارو	٪۳/۲	خگستر	٪۴/۶	خساپا	٪۲/۹	وخارزم
٪۳/۲	پکرمان	٪۲/۲	وخارزم	٪۱/۶	وخارزم	٪۴/۴	وخارزم	٪۲/۳	خبهمن



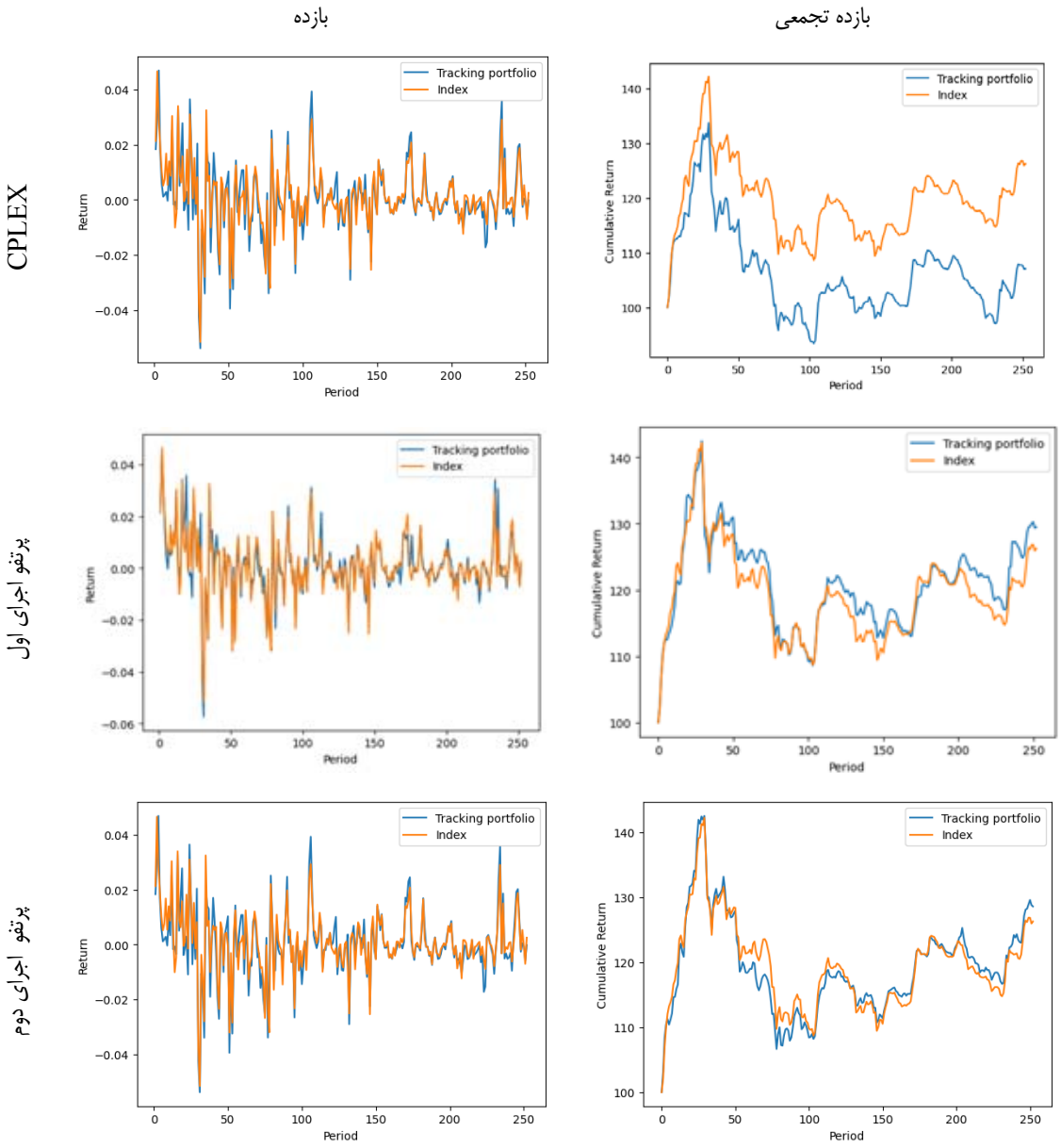


شکل ۲. مقایسه عملکرد گرافیکی سبد رویکرد پیشنهادی با CPLEX در حالت عادی بدون عدم قطعیت

عملکرد دو پرتفوی پیشنهادی اول در مقایسه با پرتفوی پیشنهادی CPLEX به‌صورت گرافیکی در شکل ۲ نمایش داده شده است. از تصویر فوق مشخص است که عملکرد حل‌کننده CPELX از نظر ردگیری شاخص نسبت به جواب‌های به‌دست‌آمده اول و دوم توسط الگوریتم پیشنهادی ضعیف‌تر است. دو پرتفو دیگر نیز عملکرد به مراتب بهتری نسبت به حل کننده تجاری داشته اند. از نظر بازده تجمعی نیز سبد ردیاب CPLEX اختلافی بیشتر از شاخص داشته و در تمامی دوره‌های آزمودن پایین‌تر شاخص است، درحالی‌که سبد GALB اختلاف ناچیزی از شاخص داشته و به‌خوبی آن را ردیابی کرده است. نکته اصلی در این مسئله اختلاف کم پرتفوی پیشنهادی از شاخص تحت بررسی هست و اختلاف بالا یا پایین نشان از ضعف الگوریتم در حرکت کردن هماهنگ با شاخص مورد نظر است.

جدول ٦. مقایسه سبد پیشنهادی CPLEX با سبد پیشنهادی رویکرد این پژوهش در حالت عادی با در نظر گرفتن عدم قطعیت ۷۵ درصد

CPLEX		اجرای چهارم		اجرای سوم		اجرای دوم		اجرای اول	
٪۲۴/۰	امین	٪۲۱/۷	جم	٪۲۰/۲	فملی	٪۲۹/۱	اپال	٪۱۷/۰	وغدیر
٪۲۱/۰	وغدیر	٪۲۰/۱	فملی	٪۱۶/۸	تاپیکو	٪۱۹/۰	وغدیر	٪۱۶/۸	امین
٪۱۶/۲	فولاد	٪۱۷/۸	وغدیر	٪۱۴/۸	امین	٪۱۷/۵	فولاد	٪۱۱/۰	کچاد
٪۱۳/۳	وپاسار	٪۱۳/۰	میدکو	٪۱۴/۵	میدکو	٪۱۰/۳	تیپیکو	٪۱۰/۰	فخوز
٪۷/۶	شبندر	٪۶/۰	خساپا	٪۸/۳	وخارزم	٪۸/۰	شپنا	٪۹/۴	شپنا
٪۶/۰	تیپیکو	٪۵/۲	تیپیکو	٪۸/۱	شتران	٪۴/۰	تنوین	٪۹/۰	سیستم
٪۳/۹	وخارزم	٪۴/۹	ومدیر	٪۶/۱	تنوین	٪۳/۷	وبصادر	٪۸/۴	سیدکو
٪۳/۴	خساپا	٪۴/۶	وبصادر	٪۵/۷	سفارس	٪۳/۶	خبهمن	٪۷/۷	وسکاب
٪۳/۲	سفارس	٪۴/۰	اخابر	٪۳/۳	سیمرغ	٪۳/۰	وخارزم	٪۶/۵	وبصادر
٪۱/۶	انرژی	٪۲/۶	اتکام	٪۲/۰	دلر	٪۱/۸	دجابر	٪۴/۱	وخارزم



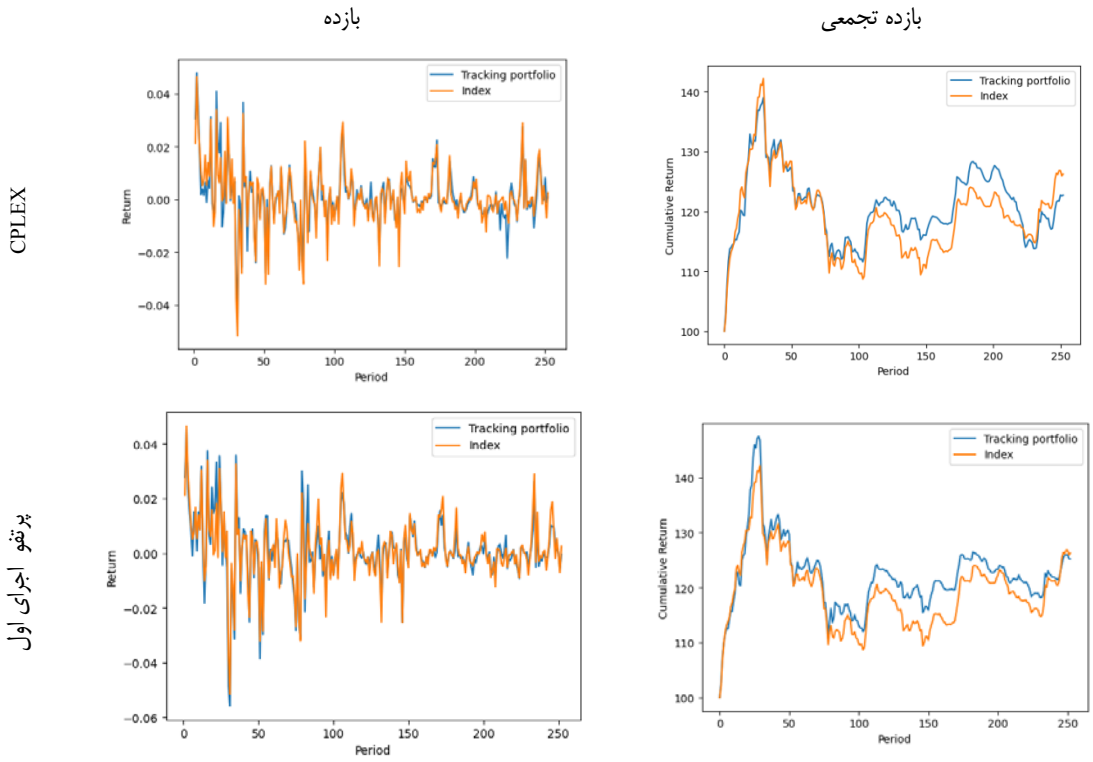
شکل ۳. مقایسه عملکرد گرافیکی سبد الگوریتم پیشنهادی با سبد CPLEX در حالت عادی با در نظر گرفتن عدم قطعیت ۷۵ درصد

با در نظر گرفتن عدم قطعیت مشخص است که عملکرد سبدهای ردیاب الگوریتم ابتکاری نسبت به حالت قبل بهبود چشمگیری پیدا کرده و اختلاف سبدها با شاخص بسیار ناچیز شده است.

با توجه به اینکه شاخص بورس اوراق بهادار تهران، نسب به سهام بزرگ موجود در بورس هم‌بستگی زیادی دارد، در این بخش فرض شده است که سبد پیشنهادی برای ردیابی شاخص بورس نمی تواند از ۱۰ سهم با بزرگ‌ترین ارزش بازار (جدول ۴) باشد و الگوریتم باید با سایر سهام موجود در بازار بتواند شاخص بورس تهران را به خوبی ردیابی کند. نتایج این بخش به صورت زیر است.

جدول ۷. مقایسه سبد پیشنهادی CPLEX با سبد پیشنهادی در حالت حذف ۱۰ سهم بزرگ بدون عدم قطعیت

CPLEX		اجرای سوم		اجرای دوم		اجرای اول	
٪۱۸/۴	کیمیاتک	٪۲۶/۹	اپال	٪۱۸/۵	جم	٪۲۰/۸	جم
٪۱۳/۰	شبریز	٪۱۸/۴	تاصیکو	٪۱۷/۹	اپال	٪۱۳/۴	شبندر
٪۱۰/۹	وپارس	٪۱۱/۲	شبندر	٪۱۲/۳	شبندر	٪۱۲/۴	امین
٪۱۰/۲	شفن	٪۱۰/۵	کرماشا	٪۱۲/۰	کرماشا	٪۱۱/۸	مبین
٪۹/۶	ثفارس	٪۸/۰	ومدیر	٪۱۰/۰	تیپیکو	٪۹/۳	تیپیکو
٪۹/۵	چدن	٪۶/۷	شیران	٪۷/۶	وخارزم	٪۸/۲	ومدیر
٪۹/۱	حفارس	٪۶/۵	نوری	٪۷/۵	سیتا	٪۷/۹	کروی
٪۷/۳	فخاس	٪۵/۴	خساپا	٪۶/۶	خساپا	٪۷/۲	خودرو
٪۷/۲	دسبحا	٪۳/۴	کروی	٪۶/۵	کروی	٪۵/۱	اخابر
٪۴/۷	فاما	٪۲/۹	وتجارت	٪۱/۲	فخاس	٪۳/۶	کاما

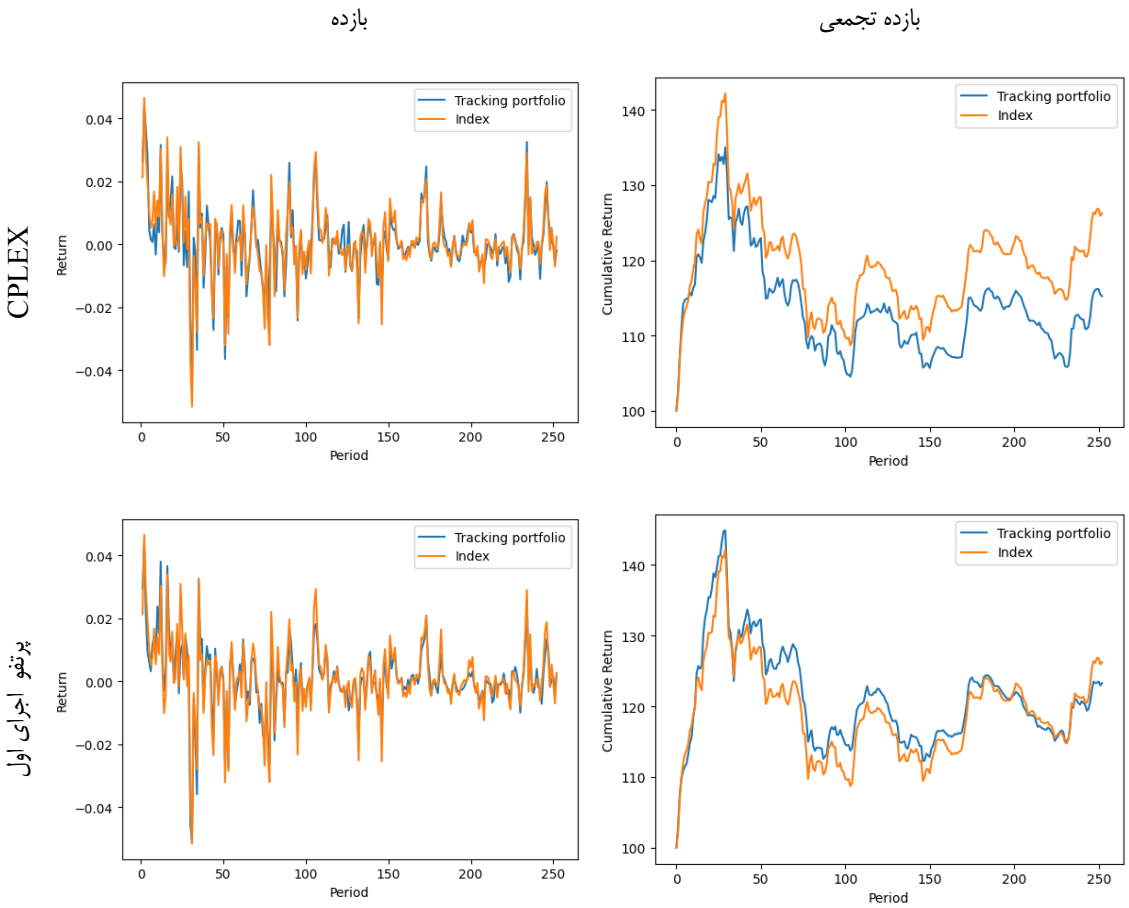


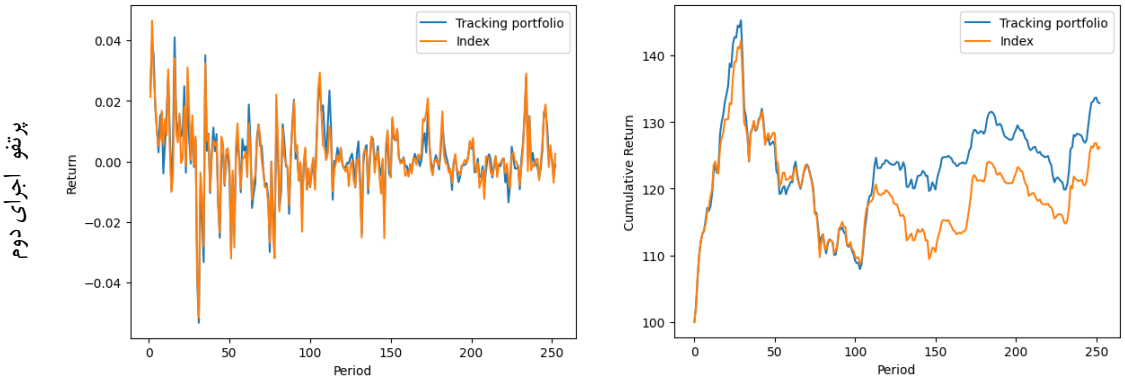
شکل ٤. مقایسه عملکرد گرافیکی سبد رویکرد پیشنهادی با CPLEX در حالت حذف ۱۰ سهم بزرگ بدون عدم قطعیت

در این حالت عملکرد CPLEX از نظر بازده تجمعی بهبود پیدا کرده، اما در بعضی از دوره‌ها مانند دوره بین ۲۰۰ تا ۲۵۰ قادر به حفظ کردن شکل و بازتولید مناسب شاخص نبوده و در ردیابی اختلاف داشته. سبد ردیاب الگوریتم ابتکاری پیشنهادی اما با حالت قبلی که تمامی سهم‌ها در نظر گرفته شده بودند، اختلاف چشمگیری ندارد.

جدول ۸. مقایسه سبد پیشنهادی CPLEX با سبد پیشنهادی در حالت حذف ۱۰ سهم بزرگ با در نظر گرفتن عدم قطعیت ۷۵ درصد

CPLEX		اجرای چهارم		اجرای سوم		اجرای دوم		اجرای اول	
٪۱۷/۳	شبریز	٪۲۸/۱	اپال	٪۲۸/۶	اپال	٪۳۱/۳	اپال	٪۲۲/۴	اپال
٪۱۵/۳	وتوسم	٪۱۸/۴	وسپه	٪۱۵/۴	تاصیکو	٪۱۵/۳	تاصیکو	٪۱۴/۵	وسپه
٪۱۳/۰	دپارس	٪۱۲/۱	شبندر	٪۱۰/۸	کرماشا	٪۱۲/۹	شبندر	٪۱۱/۱	شبندر
٪۱۱/۵	وپارس	٪۸/۴	تیپیکو	٪۱۲/۲	شبندر	٪۱۱/۴	وسپه	٪۹/۰	ومدیر
٪۱۱/۴	شفن	٪۷/۷	امین	٪۹/۹	وصندوق	٪۷/۳	شپدیس	٪۸/۶	مبین
٪۹/۴	دسبحا	٪۷/۱	کروی	٪۶/۵	وخارزم	٪۵/۱	وتجارت	٪۷/۷	تیپیکو
٪۹/۲	فخاس	٪۵/۴	خودرو	٪۵/۳	خساپا	٪۴/۹	خساپا	٪۷/۷	نوری
٪۵/۸	حفارس	٪۵/۳	وخارزم	٪۴/۱	فاسمین	٪۴/۵	وتوکا	٪۷/۲	خودرو
٪۴/۱	سفانو	٪۴/۱	سفارس	٪۳/۶	سفارس	٪۳/۹	شیران	٪۶/۸	فاسمین
٪۲/۸	فاما	٪۳/۳	پترول	٪۳/۳	سیمرغ	٪۳/۲	سفارس	٪۴/۸	اخابر





شکل ۵. مقایسه عملکرد گرافیکی سبد رویکرد پیشنهادی با CPLEX در حالت بدون ۱۰ سهم بزرگ با در نظر گرفتن عدم قطعیت ۷۵ درصد

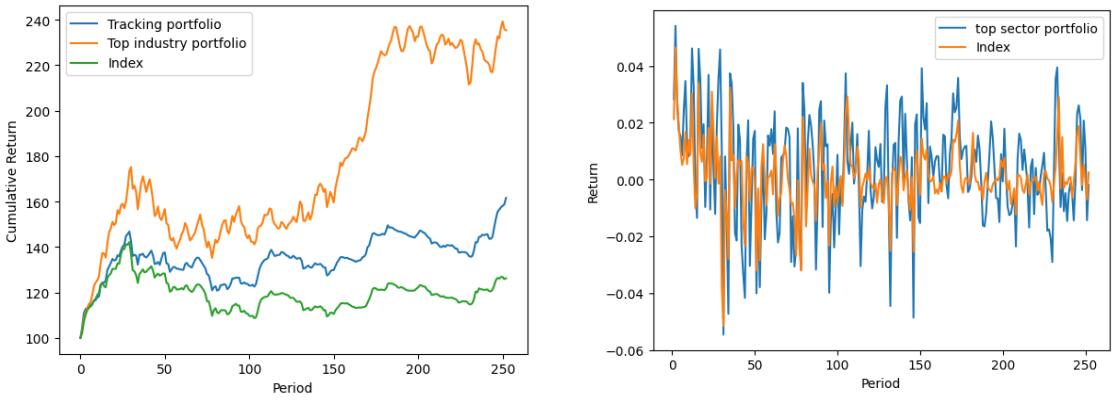
در این حالت با در نظر گرفتن عدم قطعیت ۷۵ درصد عملکرد CPLEX نسبت به بازتولید شاخص بهتر شده اما از نظر بازده تجمعی با شاخص اختلاف نامطلوبی دارد اما الگوریتم ابتکاری پیشنهادی توانسته عملکردی بسیار مناسب با اختلاف کم در دوره تست کسب کند. در تمامی چهار حالت مشخص است که سبد ردیاب عملکرد بهتری در حالت‌هایی که عدم قطعیت در نظر گرفته شده نسبت به حالت‌های بدون عدم قطعیت دارد.

در بخش آخر یک «سبد سهام ردیاب معیار جایگزین» جهت مقایسه با سبد ردیاب پیشنهادی این پژوهش انتخاب شده تا از نظر عملکردی بین دو سبد مقایسه انجام شود. سبد انتخاب شده باید سبدی باشد که بتواند تا حد زیادی رفتار شاخص را از خود نشان دهد. از آنجاییکه شاخص کل بورس تهران بیشترین تأثیر را از سهم‌های با بیشترین ارزش بازار می‌پذیرد سبدی که برای مقایسه انتخاب شده ۱۰ سهم برتر از ۱۰ صنعت بزرگ بورس تهران از نظر ارزش بازار از تاریخ ۲۹ آذر ۱۴۰۰ تا ۸ فروردین ۱۴۰۳ است. وزن هر دارایی در این سبد متناسب با درصد وزنی ارزش بازار آن صنعت است. این سبد به دو دلیل می‌تواند معیار خوبی برای مقایسه باشد. دلیل اول دربرداشتن صنایع بزرگ و ارزش بازار آن سهم‌ها است که بسیار بر شاخص کل تأثیرگذار است. دلیل دوم تنوع[1] سبد است که باعث کاهش ریسک سرمایه‌گذاری می‌شود. شکل ۶ عملکرد بازده و بازده تجمعی سبد معیار، سبد پیشنهادی و شاخص را در یک قاب نشان داده و جدول ۹ نیز سهام سبد ردیاب جایگزین به همراه اوزان هر یک را نشان می‌دهد.

جدول ۹. سبد معیار جایگزین به همراه اوزان هر سهم

سهم	فارس	فملی	کگل	شستا	شپنا	وبملت	خودرو	دعبید	سیتا	مبین
وزن	٪۲۹/۶	٪۲۳/۵۴	٪۹/۶	٪۹/۲	٪۸/۴	٪۷/۰	٪۴/۴	٪۳/۱	٪۳/۰	٪۲/۱

1. Diversification



شکل ٦. مقایسه سبد سهام معیار با سبد سهام پیشنهادی این پژوهش با هدف ردیابی شاخص

از نظر بازده تجمعی بازده سبد معیار که در سمت چپ با رنگ نارنجی نشان داده شده است از سبد ردیاب و شاخص کل بیشتر بوده ولی از شکل سمت راست واضح است که اصلاً عملکردی متناسب با شاخص نداشته است نتوانسته بازده شاخص را بازتولید کند. لازم به تأکید مجدد است که هدف این پژوهش تولید بازده شاخص و نه به دست آوردن بازده‌ای بیشتر یا کمتر از آن است. با در نظر گرفتن این نکته مشخص است که سبد معیار نتوانسته بازده نزدیک به شاخص داشته باشد. از طرف دیگر میدانیم که دستیابی به بازده بیشتر همواره با ریسک بیشتر برای سرمایه‌گذار همراه است لذا در قسمت بعدی به بررسی ریسک این دو سبد می‌پردازیم. جدول ۱۰ ریسک (انحراف معیار بازده) هر دو سبد را نسبت به شاخص نشان می‌دهد.

جدول ۱۰. مقایسه ریسک سبد ردیاب معیار با سبد ردیاب پیشنهادی

	ریسک	ریسک فاصله از شاخص
شاخص	۰/۰۱۱۵۲	–
سبد ردیاب پیشنهادی	۰/۰۱۲۰۹	۰/۰۰۴۵۶۵
سبد ردیاب معیار	۰/۰۱۸۹۹	۰/۰۱۲۹۸۴

اگر شاخص کل را به‌عنوان یک سبد سهام فرضی تشکیل‌شده از تمامی سهم‌های تأثیرگذار بر شاخص در نظر بگیریم پس از محاسبه ریسک این سبد و مقایسه آن با دو سبد مورد بحث دیگر مشخص است که سبد ردیاب پیشنهادی، ریسکی متناسب با شاخص ولی فقط با تعداد ۱۰ سهم دارد و این در حالی است که ریسک سبد معیار بسیار بیشتر از شاخص و سبد پیشنهادی ما است. یک معیار ریسک دیگر در ادبیات ردیابی شاخص فاصله از شاخص است. در ستون ریسک فاصله از شاخص، انحراف معیار اختلاف میان بازده شاخص و سبد موردنظر محاسبه شده است. همان طور که مشخص است ریسک سبد ردیاب معیار با این معیار حتی از حالت قبلی نیز بیشتر است. ریسک سبد معیار در حالت آخر ۲/۸۴ برابر سبد پیشنهادی ما بوده، در حالی‌که در حالت قبلی، این نسبت تقریباً ۱/۶ برابر است.

نتیجه‌گیری و پیشنهادها

در این پژوهش یک مدل جدید استوار برنامه‌ریزی خطی عدد صحیح مختلط، برای مسئله ردیابی شاخص ارائه شد که قادر است شاخص بازار را برای مدت طولانی، بدون نیاز به تعادل مجدد ردیابی کند. مدل پیشنهادی با مدل‌های خطی موجود با استفاده از حل‌کنندهٔ تجاری CPLEX مقایسه شد و نتایج برتری مدل ریاضی ارائه شده در داده‌های خارج نمونه، در تمامی شاخص‌های ارزیابی از قبیل ضریب هم‌بستگی، بتا، انحراف معیار و سایر معیارهای مبتنی بر فاصله را نشان داد. مدل توسعه‌یافتهٔ این پژوهش، به‌طور میانگین موجب بهبود ۳۲ درصد در دقت ردیابی شاخص بر داده‌های خارج نمونه شده است. همچنین، در پژوهش حاضر، الگوریتم ابتکاری جدیدی به‌نام GALB، برای حل این مسئله ارائه و عملکرد آن را با حل‌کنندهٔ تجاری موجود مقایسه شد. نتایج مقایسه نشان داد که الگوریتم پیشنهادی قادر است با جواب بهینه در مسائل با اندازه‌های متوسط و کوچک همگرا شود و در مسائل با اندازهٔ بزرگ‌تر، کاراتر و برتر از حل‌کنندهٔ تجاری موجود است. الگوریتم پیشنهادی این پژوهش در ۹۰ درصد مقایسه‌های رودررو، روی داده‌های خارج از نمونه، عملکرد بهتری نسبت به حل‌کنندهٔ CPLEX داشته است. بر اساس آزمون ناپارامتری ویلکوکسون، تفاوت عملکرد در تمامی شاخص‌ها با سطح اطمینان ۹۰ درصد و در همه شاخص‌ها به‌جز بتا با سطح اطمینان ۹۵ درصد، از نظر آماری معنادار گزارش شده است. همچنین، این الگوریتم با بهره‌گیری از شاخص‌های مختلف در مجموع ۸۷/۵ درصد برتری نسبت به حل‌کنندهٔ مرجع در کل مجموعه داده‌ها از خود نشان داده است. علاوه‌براین، سهم‌های انتخابی به همراه درصد وزنی سبدهای پیشنهادی و عملکرد آن‌ها در مقایسه با حل‌کنندهٔ تجاری در داده‌های خارج نمونه برای حالت‌های عادی و حالتی که ۱۰ سهم بزرگ بازار در محاسبات دخالت داده نشده، بررسی شد. همچنین سبد معیاری متشکل از ۱۰ صنعت بزرگ بورس برای مقایسه با سبد ردیاب پیشنهادی این پژوهش ارائه شد تا از نظر قابلیت ردیابی، بازتولید عملکرد شاخص و ریسک مقایسه صورت گیرد. نتایج نشان می‌دهد که سبد ردیاب پیشنهادی، در بازتولید عملکرد شاخص با ریسکی متناسب با شاخص عملکرد بهتری دارد، در حالی‌که ریسک سبد معیار بسیار بیشتر بود و قابلیت ردیابی شاخص را نداشت.

با وجود نتایج رضایت‌بخش این مطالعه، همچنان فرصت‌هایی برای گسترش و بهبود وجود دارد. در گام نخست، پیشنهاد می‌شود که عملکرد الگوریتم پیشنهادی در سایر مسائل مالی، از جمله بهینه‌سازی سبد سهام، بررسی شود. با توجه به ماهیت کاربردی و پتانسیل تجاری رویکرد ارائه‌شده، انجام مطالعاتی در خصوص پیاده‌سازی عملی آن در فرایند سبدگردانی و مقایسه عملکرد آن با روش‌های سنتی و رایج، می‌تواند افق‌های جدیدی را در کاربردهای واقعی این رویکرد فراهم سازد.

منابع

آزادی، امیر و نجفی، امیرعباس (۱۳۹۹). مدل بهبودیافته ردیابی شاخص با درنظر گرفتن هزینه‌های معاملاتی. *مهندسی مالی و مدیریت اوراق بهادار (مدیریت پرتفوی)*، ۱۱(۴۲)، ۲۳-۴۳.

اصولیان، محمد؛ نیک‌مرام، علی و کریمی، مهدی (۱۴۰۴). پیش‌بینی روند شاخص کل با استفاده از شبکه‌های عصبی هیبریدی با تمرکز بر استخراج ویژگی مقیاس زمانی چندگانه در بورس اوراق بهادار تهران. *تحقیقات مالی*، ۲۷(۱)، ۸۵-۱۱۳.

بت‌شکن، محمدهاشم؛ بحرالعلوم، محمدمهدی؛ ارضا، امیرحسین و تقی خان تجریشی، امیر (۱۳۹۸). توسعهٔ معیار پایدار ردیابی شاخص کل بورس اوراق بهادار تهران. *چشم‌انداز مدیریت مالی*، ۹(۲۷)، ۵۱-۷۹.

تندنویس، فرید و حکیمیان، حسن (۱۳۹۹). کاربرد الگوریتم پیچش زمانی پویا و ضرایب هم‌بستگی در خوشه‌بندی سری‌های زمانی به‌منظور تشکیل پرتفوی مبتنی بر شاخص. *دانش سرمایه‌گذاری*، ۹(۳۵)، ۱۸۹-۲۰۵.

رستمی نوروزآباد، مجتبی؛ گل‌بابائی پاسندی، علی؛ شهرازی، میلاد و اسفندیاری، سمیه (۱۴۰۳). اندازه‌گیری شاخص ترس و طمع در بازار سهام: شواهدی از بورس اوراق بهادار تهران. *تحقیقات مالی*، ۲۶(۲)، ۳۹۷-۴۱۴.

عیوضلو، رضا؛ فلاح‌پور، سعید و دهقانی اشکذری، مهدی (۱۴۰۰). ردیابی شاخص با استفاده از معیار ارزش در معرض ریسک شرطی ترکیبی دو دنباله‌ای در بورس اوراق بهادار تهران. *تحقیقات مالی*، ۲۳(۴)، ۵۴۵-۵۶۳.

عیوضلو، رضا؛ شفیع‌زاده، مجتبی و قهرمانی، علی (۱۳۹۶). ردیابی شاخص و شاخص بهبودیافته با استفاده از رویکردهای هم‌انباشتگی و هم‌بستگی. *تحقیقات مالی*، ۱۹(۳)، ۴۵۷-۴۷۴.

گل‌ارضی، غلامحسین و ابوالفضلی، سید رامین (۱۴۰۳). بررسی اثرهای اهرمی، هم‌بستگی شرطی پویا و سرایت‌پذیری تلاطم میان شاخص‌های صنایع بورسی با استفاده از مدل ARMA-DCC-GJR-GARCH. *تحقیقات مالی*، ۲۶(۱)، ۵۸-۸۶.

شهاب لواسانی، کیوان؛ شعبانی رضوانی، لیلا و سماوی، محمد ابراهیم (۱۴۰۲). بررسی ارتباط نامتقارن احساسات سرمایه‌گذاران و نوسان‌های شاخص کل به‌روش مارکوف سوئیچینگ. *تحقیقات مالی*، ۲۵(۴)، ۶۶۱-۶۸۷.

محمدی، شاپور؛ راعی، رضا و تندنویس، فرید (۱۴۰۰). کاربرد ضرایب هم‌بستگی مبتنی بر کاپولا و اطلاعات متقابل در خوشه‌بندی سری‌های زمانی و تشکیل پرتفوی شاخصی ارتقایافته با استفاده از رویکرد بهینه‌سازی استوار. *تحقیقات مالی*، ۲۳(۴)، ۴۹۷-۵۲۲.

References

Alexander, C. & Dimitriu, A. (2005). Indexing and statistical arbitrage. *The Journal of Portfolio Management*, *31*(2), 50-63.

Anis, H. T., Costa, G. & Kwon, R. H. (2023). Risk-allocation-based index tracking. *Computers & operations research*, *154*, 106219.

Azadi, A. & Najafi, A.A. (2020). An enhanced model for the index tracking problem with transaction costs. *Financial Engineering and Securities Management (Portfolio Management),* 11(42), 23-43. SID. https://sid.ir/paper/367271/en (*in Persian*)

Beasley, J. E. (1990). OR-Library: distributing test problems by electronic mail. *Journal of the operational research society*, *41*(11), 1069-1072.

Beasley, J. E., Meade, N. & Chang, T.J. (2003). An evolutionary heuristic for the index tracking problem. *European Journal of Operational Research*, *148*(3), 621-643.

Botshekan, M.H., bahrololoum, M.M., Erza, A.H. & Taghikhan Tajrishi, A. (2019). Development of a Stable Tracking Measure for Tehran Stock Exchange. *Financial Management Perspective*, *9*(27), 51-79. doi: 10.52547/jfmp.9.27.51.

Chen, C. & Kwon, R. H. (2012). Robust portfolio selection for index tracking. *Computers & operations research*, *39*(4), 829-837.

Eyvazloo, R., Fallahpour, S. & Dehghani Ashkezari, M. (2022). Index tracking using Two-tail Mixed Conditional Value-at-risk in Tehran Stock Exchange. *Financial Research Journal*, *23*(4), 545-563. doi: 10.22059/frj.2020.289344.1006927 (*in Persian*)

Eyvazloo, R., Shafizadeh, M. & Ghahramani, A. (2017). Index Tracking and Enhanced Indexing Using Co-integration and Correlation Approaches. *Financial Research Journal*, *19*(3), 457-474. doi: 10.22059/jfr.2018.245816.1006551 (*in Persian*)

Gaivoronski, A. A., Krylov, S. & Van der Wijst, N. (2005). Optimal portfolio selection and dynamic benchmark tracking. *European Journal of Operational Research*, *163*(1), 115-131.

Golarzi, G. & Abolfazli, S. R. (2024). Examining the Leverage Effect, Dynamic Conditional Correlation, and Volatility Spillover Among Selected Indices of the Tehran Stock Exchange: Evidence from the ARMA-DCC-GJR-GARCH Model. *Financial Research Journal*, *26*(1), 58-86. doi: 10.22059/frj.2023.361936.1007487 (*in Persian*)

Jansen, R. & Van Dijk, R. (2002). Optimal benchmark tracking with small portfolios. *Journal of portfolio management*, *28*(2), 33.

Jorion, P. (2003). Portfolio optimization with tracking-error constraints. *Financial Analysts Journal*, *59*(5), 70-82.

Kaucic, M., Barbini, F. & Camerota Verdù, F. J. (2020). Polynomial goal programming and particle swarm optimization for enhanced indexation. *Soft Computing*, *24*(12), 8535-8551.

Krink, T., Mittnik, S. & Paterlini, S. (2009). Differential evolution and combinatorial search for constrained index-tracking. *Annals of Operations Research*, *172*, 153-176.

Maringer, D. & Oyewumi, O. (2007). Index tracking with constrained portfolios. *Intelligent Systems in Accounting, Finance & Management: International Journal*, *15*(1-2), 57-71.

Mohammadi, S., Raei, R. & Tondnevis, F. (2022). Application of Copula Based Correlations and Mutual Information in Time Series Clustering and Enhanced Indexing by Adopting the Robust Optimization Approach. *Financial Research Journal*, *23*(4), 497-522. doi: 10.22059/frj.2020.286126.1006904 (*in Persian*)

Oh, K. J., Kim, T. Y. & Min, S. (2005). Using genetic algorithm to support portfolio optimization for index fund management. *Expert Systems with applications*, *28*(2), 371-379.

Osoolian, M., Nikmaram, A. & Karimi, M. (2025). Predicting Index Trend Using Hybrid Neural Networks with a Focus on Multi-Scale Temporal Feature Extraction in the Tehran Stock

Exchange. *Financial Research Journal*, *27*(1), 85-113. doi: 10.22059/frj.2024.377816.1007611 (*in Persian*)

Roll, R. (1992). A mean/variance analysis of tracking error. *Journal of portfolio management*, *18*(4), 13.

Rostami Noroozabad, M., Golbabaei Pasandi, A., Shahrazi, M. & Esfandyari, S. (2024). Measuring Fear and Greed Index in Stock Market: Evidence from the Tehran Stock Exchange. *Financial Research Journal*, *26*(2), 397-414. doi: 10.22059/frj.2023.365536.1007511 (*in Persian*)

Ruiz-Torrubiano, R. & Suárez, A. (2009). A hybrid optimization approach to index tracking. *Annals of Operations Research*, *166*, 57-71.

Sant'Anna, L. R., Righi, M. B., Müller, F. M. & Guedes, P. C. (2022). Risk measure index tracking model. *International Review of Economics & Finance*, *80*, 361-383.

Sant'Anna, L. R., de Oliveira, A. D., Filomena, T. P. & Caldeira, J. F. (2020). Solving the index tracking problem based on a convex reformulation for cointegration. *Finance Research Letters*, *37*, 101356.

Shahab Lavasani, K., Shabani Rezvani, L. & Samavi, M. E. (2023). Investigating the Asymmetric Relationship between Investor Sentiments and Fluctuations in the Overall Index via the Markov Switching Method. *Financial Research Journal*, *25*(4), 661-687. doi: 10.22059/frj.2023.356277.1007444 (*in Persian*)

Silva, J. C. S. & de Almeida Filho, A. T. (2023). Using GAN-generated market simulations to guide genetic algorithms in index tracking optimization. *Applied Soft Computing*, *145*, 110587.

Silva, J. C. S. (2023). A systematic literature review on solution approaches for the index tracking problem in the last decade. *arXiv preprint arXiv:2306.01660*.

Silva, J. C. S., Silva, D. F. d. L. & de Almeida Filho, A. T. (2024). An enhanced grasp approach for the index tracking problem. *International Transactions in Operational Research*, *31*(3), 1828-1858.

Tondnevis, F. & Hakimian, H. (2020). The application of Dynamic Time Wrapping Algorithm and correlation coefficients in Time Sires Clustering for Index Tracking. *Journal of Investment Knowledge*, *9*(35), 189-205. (*in Persian*)

Torri, G., Giacometti, R. & Paterlini, S. (2024). Penalized enhanced portfolio replication with asymmetric deviation measures. *Annals of Operations Research*, *332*(1), 481-531.

Vieira, E. B. F., Filomena, T. P., Sant'anna, L. R. & Lejeune, M. A. (2023). Liquidity-constrained index tracking optimization models. *Annals of Operations Research*, *330*(1), 73-118.

Wang, M., Xu, C., Xu, F. & Xue, H. (2012). A mixed 0–1 LP for index tracking problem with CVaR risk constraints. *Annals of Operations Research*, *196*, 591-609.

Wu, D., Kwon, R. H. & Costa, G. (2017). A constrained cluster-based approach for tracking the S&P 500 index. *International Journal of Production Economics*, *193*, 222-243.