



Σχολή Θετικών Επιστημών και Τεχνολογίας
Μεταπτυχιακή Ειδίκευση στα Πληροφοριακά Συστήματα

Διπλωματική Εργασία

Βελτιστοποίηση Χαρτοφυλακίου Αποτύπωσης Δείκτη με τη χρήση
Γενετικών Αλγορίθμων - Index-tracking Portfolio Optimization
using Genetic Algorithms

Απόστολος Κίκης

Επιβλέπων καθηγητής: Ευστράτιος Γεωργόπουλος

Πάτρα, 22^η Σεπτεμβρίου 2026

ΠΕΡΙΛΗΨΗ

Η βελτιστοποίηση χαρτοφυλακίου αποτελεί ένα από τα σημαντικότερα αντικείμενα μελέτης στον χώρο της χρηματοοικονομικής διοίκησης και της χρηματοοικονομικής μηχανικής. Μία ιδιαίτερα δημοφιλής μορφή παθητικής επενδυτικής στρατηγικής είναι η αποτύπωση δείκτη (index tracking), η οποία αποσκοπεί στην αναπαραγωγή της απόδοσης ενός χρηματιστηριακού δείκτη χωρίς την ανάγκη αγοράς όλων των μετοχών που τον απαρτίζουν. Στην πράξη, η πλήρης αναπαραγωγή ενός δείκτη είναι συχνά δύσκολη ή οικονομικά ασύμφορη, καθώς πολλοί δείκτες περιλαμβάνουν μεγάλο αριθμό μετοχών, ενώ παράλληλα προκύπτουν σημαντικά κόστη συναλλαγών κατά την αγορά και πώληση τίτλων. Για τον λόγο αυτό, αναπτύσσονται μέθοδοι επιλογής ενός περιορισμένου αριθμού μετοχών, οι οποίες μπορούν να αναπαράγουν όσο το δυνατόν πιστότερα τη συμπεριφορά του δείκτη.

Η παρούσα εργασία εξετάζει το πρόβλημα της βελτιστοποίησης χαρτοφυλακίου αποτύπωσης δείκτη μέσω της χρήσης γενετικών αλγορίθμων. Οι γενετικοί αλγόριθμοι αποτελούν μία κατηγορία εξελικτικών υπολογιστικών μεθόδων εμπνευσμένων από τις διαδικασίες της φυσικής επιλογής και της εξέλιξης. Οι αλγόριθμοι αυτοί επιτρέπουν την αναζήτηση βέλτιστων ή σχεδόν βέλτιστων λύσεων σε σύνθετα προβλήματα βελτιστοποίησης, όπως είναι το πρόβλημα της κατάλληλης επιλογής μετοχών για την αναπαραγωγή ενός δείκτη.

Στο πλαίσιο του προβλήματος αποτύπωσης δείκτη, στόχος είναι η δημιουργία ενός χαρτοφυλακίου με περιορισμένο αριθμό μετοχών, το οποίο να ελαχιστοποιεί την απόκλιση της απόδοσής του από την απόδοση του αντίστοιχου χρηματιστηριακού δείκτη. Η διατύπωση του προβλήματος λαμβάνει υπόψη σημαντικούς πρακτικούς περιορισμούς, όπως το συνολικό κόστος συναλλαγών που προκύπτει από τις αγοραπωλησίες μετοχών, καθώς και τον μέγιστο επιτρεπόμενο αριθμό τίτλων που μπορούν να περιλαμβάνονται στο χαρτοφυλάκιο. Οι περιορισμοί αυτοί καθιστούν το πρόβλημα ιδιαίτερα σύνθετο και υπολογιστικά απαιτητικό, γεγονός που καθιστά τις εξελικτικές ευρετικές μεθόδους ιδιαίτερα κατάλληλες για την επίλυσή του.

Η μεθοδολογική προσέγγιση της εργασίας βασίζεται στην εφαρμογή ενός εξελικτικού ευρετικού αλγορίθμου για την επιλογή του βέλτιστου υποσυνόλου μετοχών που αναπαράγει την πορεία του δείκτη με τη μικρότερη δυνατή απόκλιση. Μέσω μιας επαναληπτικής διαδικασίας αξιολόγησης και εξέλιξης πιθανών λύσεων, ο αλγόριθμος οδηγείται σταδιακά σε αποτελεσματικότερα χαρτοφυλάκια. Παράλληλα, εξετάζονται τεχνικές μείωσης και αξιολογούνται τα αποτελέσματα της μεθόδου σε δεδομένα από διαφορετικές χρηματιστηριακές αγορές.

Στόχος της εργασίας είναι να αναδείξει την αποτελεσματικότητα των γενετικών αλγορίθμων στην επίλυση προβλημάτων αποτύπωσης χρηματιστηριακών δεικτών με την ύπαρξη ρεαλιστικών περιορισμών και να αποδείξει ότι οι εξελικτικές προσεγγίσεις μπορούν να προσφέρουν πρακτικές και υπολογιστικά αποδοτικές λύσεις για τη διαχείριση παθητικών επενδυτικών χαρτοφυλακίων, συμβάλλοντας στη βελτίωση της απόδοσης και στη μείωση του κόστους διαχείρισης.

Λέξεις κλειδιά: αποτύπωση δείκτη, βελτιστοποίηση χαρτοφυλακίου, Γενετικοί Αλγόριθμοι, Εξελικτικοί Αλγόριθμοι, Ευρετικοί Αλγόριθμοι, κόστος συναλλαγών

ABSTRACT

Portfolio optimization is one of the most important research areas in financial management and financial engineering. A widely adopted form of passive investment strategy is index tracking, which aims to replicate the performance of a stock market index without requiring the purchase of all the securities included in the index. In practice, full index replication is often difficult or economically inefficient, since many indices consist of a large number of stocks, while substantial transaction costs arise from trading activities. Consequently, various approaches have been developed to select a limited subset of stocks capable of closely reproducing the behavior of the target index.

This thesis investigates the index tracking portfolio optimization problem through the application of genetic algorithms. Genetic algorithms belong to the broader class of evolutionary computation techniques inspired by the principles of natural selection and biological evolution. These algorithms provide an effective framework for identifying optimal or near-optimal solutions to complex optimization problems, such as selecting an appropriate subset of securities that can accurately replicate the performance of a market index.

Within the context of index tracking, the objective is to construct a portfolio composed of a limited number of stocks that minimizes the deviation between the portfolio returns and the returns of the benchmark index. The problem formulation incorporates important practical constraints, including the total transaction costs associated with buying and selling securities, as well as the maximum allowable number of assets that may be included in the portfolio. These constraints significantly increase the complexity and computational requirements of the problem, making evolutionary heuristic methods particularly suitable for addressing it.

The methodological approach of this study is based on the implementation of an evolutionary heuristic algorithm for selecting the optimal subset of stocks that replicates the index with the minimum possible tracking error. Through an iterative process of evaluation and evolution of candidate solutions, the algorithm progressively converges to more efficient portfolio configurations. Furthermore, dimensionality reduction techniques are explored and the performance of the proposed method is evaluated using data from different stock markets.

The aim of this thesis is to demonstrate the effectiveness of genetic algorithms in addressing index tracking problems under realistic constraints and to show that evolutionary approaches can provide practical and computationally efficient solutions for the management of passive investment portfolios, contributing to improved performance and reduced management costs.

Keywords: index tracking, portfolio optimization, genetic algorithms, evolutionary algorithms, transaction costs

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας, θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες προς τον επιβλέποντα καθηγητή μου, κύριο Ευστράτιο Γεωργόπουλο, για την πολύτιμη επιστημονική καθοδήγηση, τις εύστοχες παρατηρήσεις και τη συνεχή υποστήριξή του καθ' όλη τη διάρκεια της εκπόνησης της εργασίας. Η συμβολή του υπήρξε ιδιαίτερα σημαντική για τη διαμόρφωση και την ολοκλήρωση της παρούσας εργασίας.

Παράλληλα, θα ήθελα να εκφράσω τη βαθιά μου ευγνωμοσύνη προς την οικογένειά μου, για την αμέριστη συμπαράσταση, την υπομονή και τη διαρκή ηθική υποστήριξη που μου προσέφερε σε όλη τη διάρκεια αυτής της απαιτητικής προσπάθειας. Η εμπιστοσύνη και η στήριξή τους αποτέλεσαν ουσιαστικό παράγοντα για την επίτευξη του παρόντος στόχου μου.

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

[Πίνακας 5.2.1 Βασικές παράμετροι της τελικής πειραματικής εκτέλεσης](#)

[Πίνακας 5.3.1 Εκτός δείγματος MSE παρακολούθησης για τον δείκτη Hang Seng](#)

[Πίνακας 5.3.2 Εκτός δείγματος MSE παρακολούθησης για τον δείκτη FTSE 100](#)

[Πίνακας 5.3.3 Εκτός δείγματος MSE παρακολούθησης για τον δείκτη Nikkei 225](#)

[Πίνακας 5.4.1 Μεταβολή του καλύτερου MSE εκπαίδευσης κατά την εκτέλεση του Γενετικού Αλγορίθμου](#)

[Πίνακας 5.5.1 Συνολική κατάταξη των τεσσάρων μεθόδων στα εννέα σενάρια](#)

[Πίνακας 5.6.1 Έλεγχος ευαισθησίας βασικών παραμέτρων στο Hang Seng, \$K=10\$ \(5 runs ανά ρύθμιση\)](#)

[Πίνακας 5.7.1 Σταθερότητα του Γενετικού Αλγορίθμου σε 5 ανεξάρτητες επαναλήψεις ανά σενάριο](#)

[Πίνακας 5.8.1 Αρχικός πληθυσμός – γενιά 1: τρία καλύτερα χαρτοφυλάκια](#)

[Πίνακας 5.8.2 Ενδιάμεσος πληθυσμός – γενιά 75: τρία καλύτερα χαρτοφυλάκια](#)

[Πίνακας 5.8.3 Τελικός πληθυσμός – γενιά 150: τρία καλύτερα χαρτοφυλάκια](#)

[Πίνακας 5.9.1 Παραδείγματα τελικών χαρτοφυλακίων του ΓΑ για \$K=10\$](#)

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

[Σχήμα 5.1. Εκτός δείγματος MSE για τις τέσσερις μεθόδους στον Hang Seng](#)

[Σχήμα 5.2. Εκτός δείγματος MSE για τις τέσσερις μεθόδους στον FTSE 100](#)

[Σχήμα 5.3. Εκτός δείγματος MSE για τις τέσσερις μεθόδους στον Nikkei 225](#)

[Σχήμα 5.4. Σύγκλιση του Γενετικού Αλγορίθμου στον Hang Seng για \$K=5, 10\$ και \$15\$](#)

[Σχήμα 5.5. Σύγκλιση του Γενετικού Αλγορίθμου στον FTSE 100 για \$K=5, 10\$ και \$15\$](#)

[Σχήμα 5.6. Σύγκλιση του Γενετικού Αλγορίθμου στον Nikkei 225 για \$K=5, 10\$ και \$15\$](#)

[Σχήμα 5.7. Μέση κατάταξη των μεθόδων στα εννέα σενάρια](#)

[Σχήμα 5.8. Αριθμός νικών ανά μέθοδο στα εννέα σενάρια](#)

[Σχήμα 5.9. Μέσο σχετικό εκτός δείγματος MSE ως προς τον νικητή κάθε σεναρίου](#)

[Σχήμα 5.10. Μέσο out-of-sample MSE και τυπική απόκλιση για τις επτά ρυθμίσεις ευαισθησίας \(Hang Seng, \$K=10\$ \)](#)

[Σχήμα 5.11 Συντελεστής μεταβλητότητας του out-of-sample MSE του ΓΑ στις πέντε επαναλήψεις](#)

[Σχήμα 5.12. Out-of-sample MSE στις πέντε ανεξάρτητες εκτελέσεις για Hang Seng και \$K=10\$](#)

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ & ΑΚΡΩΝΥΜΙΑ

GA — Genetic Algorithm / Γενετικός Αλγόριθμος

EA — Evolutionary Algorithm / Εξελικτικός Αλγόριθμος

PSO — Particle Swarm Optimization / Βελτιστοποίηση Σμήνους Σωματιδίων

SA — Simulated Annealing / Προσομοιωμένη Ανόπτηση

TS — Tabu Search / Αναζήτηση Tabu

TA — Threshold Accepting / Μέθοδος Threshold Accepting

MIP — Mixed-Integer Programming / Μικτός Ακέραιος Προγραμματισμός

MV — Mean-Variance / Μέσος-Διασπορά

TE — Tracking Error / Σφάλμα Αποτύπωσης

MSE — Mean Squared Error / Μέσο Τετραγωνικό Σφάλμα

OOS — Out-of-Sample / Εκτός Δείγματος

IS — In-Sample / Εντός Δείγματος

Corr-EW — Correlation-based Equal Weighting / Ισοβαρής Στάθμιση βάσει Συσχέτισης

Corr-Opt — Correlation-based Optimized Portfolio / Βελτιστοποιημένο Χαρτοφυλάκιο βάσει Συσχέτισης

RS — Random Search / Τυχαία Αναζήτηση

TC — Transaction Cost / Κόστος Συναλλαγών

CSV — Comma-Separated Values / Αρχείο τιμών διαχωρισμένων με κόμμα

UCITS — Undertakings for Collective Investment in Transferable Securities / ΟΣΕΚΑ

OR-Library — Operations Research Library / Βιβλιοθήκη Δεδομένων Επιχειρησιακής Έρευνας

FTSE — Financial Times Stock Exchange

Nikkei — Nikkei Stock Average

HSI — Hang Seng Index

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ.....	2
ABSTRACT	3
ΕΥΧΑΡΙΣΤΙΕΣ	4
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	5
ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ	6
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ & ΑΚΡΩΝΥΜΙΑ	7
ΠΕΡΙΕΧΟΜΕΝΑ	8
ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ	11
1.1 Γενικό πλαίσιο της διαχείρισης επενδύσεων	11
1.2 Το πρόβλημα της βελτιστοποίησης χαρτοφυλακίου	12
1.3 Η έννοια της αποτύπωσης δείκτη (Index Tracking)	13
1.4 Οι Γενετικοί και Εξελικτικοί Αλγόριθμοι ως εργαλεία επίλυσης προβλημάτων βελτιστοποίησης 14	
1.5 Αντικείμενο της εργασίας	16
1.6 Σκοπός της εργασίας.....	17
1.7 Στόχοι της εργασίας	18
1.8 Ερευνητικά ερωτήματα.....	19
1.9 Μεθοδολογική προσέγγιση της εργασίας	21
1.10 Δομή της εργασίας.....	22
ΚΕΦΑΛΑΙΟ 2: ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΑΠΟΤΥΠΩΣΗΣ ΔΕΙΚΤΗ	24
2.1 Εισαγωγή.....	24
2.2 Χρηματοοικονομικό υπόβαθρο του προβλήματος	25
2.3 Το πρόβλημα της πλήρους και της μερικής αναπαραγωγής δείκτη	30
2.4 Βασικές συνιστώσες του προβλήματος της αποτύπωσης δείκτη	36
2.5 Κριτήρια και δείκτες αξιολόγησης της απόδοσης	46
2.6 Περιορισμοί στο πρόβλημα βελτιστοποίησης	52
2.7 Κλασικές προσεγγίσεις στη βελτιστοποίηση χαρτοφυλακίου	61
2.8 Κατηγορίες μεθόδων επίλυσης του προβλήματος	67
2.9 Εξελικτικές και μεταερευνητικές τεχνικές	72
2.10 Βιβλιογραφική ανασκόπηση	82
ΚΕΦΑΛΑΙΟ 3: ΠΡΟΤΕΙΝΟΜΕΝΗ ΜΕΘΟΔΟΣ ΕΠΙΛΥΣΗΣ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ	85
3.1 Εισαγωγή.....	85

3.2 Διατύπωση του προβλήματος ως προβλήματος βελτιστοποίησης.....	86
3.3 Ορισμός συμβολισμών, παραμέτρων και μεταβλητών.....	87
3.4 Αντικειμενική συνάρτηση.....	89
3.5 Περιορισμοί του προβλήματος	92
3.6 Υποθέσεις και απλοποιήσεις του μοντέλου	95
3.7 Σχεδιασμός του προτεινόμενου Γενετικού Αλγορίθμου	97
3.8 Αναπαράσταση χρωμοσώματος.....	99
3.9 Δημιουργία αρχικού πληθυσμού	101
3.10 Συνάρτηση καταλληλότητας	102
3.11 Τελεστής επιλογής.....	105
3.12 Τελεστής διασταύρωσης.....	107
3.13 Τελεστής μετάλλαξης	109
3.14 Μηχανισμός επιδιόρθωσης και διαχείρισης περιορισμών.....	112
3.15 Κριτήριο τερματισμού	115
3.16 Διάγραμμα ροής της προτεινόμενης μεθόδου.....	117
3.17 Υπολογιστική πολυπλοκότητα και σχολιασμός.....	119
ΚΕΦΑΛΑΙΟ 4: ΥΛΟΠΟΙΗΣΗ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΜΕΘΟΔΟΥ	122
4.1 Εισαγωγή.....	122
4.2 Αρχιτεκτονική της υπολογιστικής υλοποίησης.....	122
4.3 Εισαγωγή και προεπεξεργασία δεδομένων	123
4.4 Διαχείριση περιορισμών	124
4.5 Συνάρτηση αξιολόγησης και tracking MSE	125
4.6 Υλοποίηση της Corr-EW.....	126
4.7 Υλοποίηση της Corr-Opt	126
4.8 Υλοποίηση της Random Search	127
4.9 Υλοποίηση του Genetic Algorithm.....	127
4.10 Παραμετροποίηση του Genetic Algorithm.....	130
4.11 Ενιαία διαδικασία εκτέλεσης των πειραμάτων	131
4.12 Αξιολόγηση in-sample και out-of-sample.....	131
4.13 Αποθήκευση και οργάνωση αποτελεσμάτων	131
4.14 Γραφική παρακολούθηση της εξέλιξης.....	132
4.15 Αντιστοίχιση θεωρίας και υλοποίησης	132
ΚΕΦΑΛΑΙΟ 5: ΑΠΟΤΕΛΕΣΜΑΤΑ	133
5.1 Εισαγωγή.....	133
5.2 Πειραματικός Σχεδιασμός	135
5.2 Μέθοδοι και κριτήρια αξιολόγησης.....	137
5.3 Αποτελέσματα ανά χρηματιστηριακό δείκτη.....	137
5.3.1 Hang Seng	137

5.3.2 FTSE 100.....	138
5.3.3 Nikkei 225	139
5.4 Ανάλυση σύγκλισης του Γενετικού Αλγορίθμου	141
5.5 Συνολική σύγκριση των μεθόδων	143
5.6 Έλεγχος ευαισθησίας βασικών παραμέτρων.....	145
5.7 Ανεξάρτητες επαναλήψεις και έλεγχος στοχαστικής σταθερότητας.....	146
5.8 Παράδειγμα εξέλιξης του πληθυσμού του ΓΑ.....	148
5.9 Παραδείγματα τελικών χαρτοφυλακίων.....	149
5.10 Ψευδοκώδικας της υλοποίησης του Γενετικού Αλγορίθμου	150
5.11 Ερμηνεία των αποτελεσμάτων.....	151
5.12 Περιορισμοί της πειραματικής αξιολόγησης	151
5.13 Συμπεράσματα Κεφαλαίου	152
ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ	152
6.1 Συνοπτική ανακεφαλαίωση της εργασίας	152
6.2 Κύρια συμπεράσματα	153
6.3 Αξιολόγηση της προτεινόμενης μεθοδολογίας	154
6.4 Σύγκριση με άλλες μεθόδους.....	155
6.5 Περιορισμοί της εργασίας.....	156
6.6 Προτάσεις για μελλοντική έρευνα	157
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	158

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 Γενικό πλαίσιο της διαχείρισης επενδύσεων

Η διαχείριση επενδύσεων αποτελεί έναν από τους σημαντικότερους τομείς της σύγχρονης χρηματοοικονομικής επιστήμης, καθώς συνδέεται άμεσα με τη διαδικασία λήψης αποφάσεων σχετικά με την κατανομή διαθέσιμων κεφαλαίων σε εναλλακτικές επενδυτικές επιλογές (Markowitz, 1952; Merton, 1969). Στο πλαίσιο αυτό, ο επενδυτής ή ο διαχειριστής χαρτοφυλακίου καλείται να αποφασίσει ποια περιουσιακά στοιχεία θα επιλεγούν, σε ποιες ποσότητες θα συμμετέχουν στο χαρτοφυλάκιο και με ποιον τρόπο θα επιτυγχάνεται η επιθυμητή σχέση μεταξύ κινδύνου και απόδοσης (Markowitz, 1952).

Η ανάγκη για ορθολογική και συστηματική υποστήριξη των επενδυτικών αποφάσεων καθιστά τη βελτιστοποίηση χαρτοφυλακίου ένα ιδιαίτερα κρίσιμο ερευνητικό και πρακτικό πεδίο (Markowitz, 1952; Schaerf, 2002). Η έννοια του χαρτοφυλακίου αναφέρεται στη σύνθεση ενός συνόλου επενδυτικών στοιχείων, όπως μετοχές, ομόλογα και άλλα χρηματοοικονομικά μέσα, με στόχο την επίτευξη συγκεκριμένων επενδυτικών επιδιώξεων (Markowitz, 1952). Οι επιδιώξεις αυτές συνήθως σχετίζονται με τη μεγιστοποίηση της αναμενόμενης απόδοσης και την ταυτόχρονη ελαχιστοποίηση του αναλαμβανόμενου κινδύνου, αν και στην πράξη εμπλέκονται και άλλα κριτήρια, όπως η ρευστότητα, το κόστος συναλλαγών, η διαφοροποίηση και οι κανονιστικοί ή θεσμικοί περιορισμοί (Markowitz, 1952; Schaerf, 2002; Canakgoz & Beasley, 2009).

Η διαχείριση επενδύσεων δεν περιορίζεται μόνο στην αρχική επιλογή των στοιχείων ενός χαρτοφυλακίου, αλλά επεκτείνεται και στη συνεχή παρακολούθηση και αναπροσαρμογή του (Merton, 1969; Perold & Sharpe, 1988). Οι μεταβολές στις συνθήκες των αγορών, στις αποτιμήσεις των τίτλων, στις μακροοικονομικές μεταβλητές και στη συμπεριφορά των επενδυτών δημιουργούν ένα ιδιαίτερα δυναμικό περιβάλλον, μέσα στο οποίο οι αποφάσεις πρέπει να λαμβάνονται υπό συνθήκες αβεβαιότητας (Merton, 1969). Για τον λόγο αυτό, η ανάπτυξη μεθόδων και εργαλείων που επιτρέπουν την αποτελεσματική ανάλυση και βελτιστοποίηση των επενδυτικών επιλογών αποτελεί διαχρονικό ζητούμενο (Markowitz, 1952; Perold & Sharpe, 1988).

Στη χρηματοοικονομική θεωρία, η διαχείριση επενδύσεων έχει συνδεθεί στενά με το πρόβλημα της κατανομής κεφαλαίων μεταξύ εναλλακτικών τίτλων, ώστε να διαμορφωθεί ένα χαρτοφυλάκιο που να ανταποκρίνεται στις προτιμήσεις και στους περιορισμούς του επενδυτή (Markowitz, 1952; Merton, 1969). Η προσέγγιση αυτή αποκτά ιδιαίτερο ενδιαφέρον στις περιπτώσεις όπου ο επενδυτής δεν επιδιώκει απλώς υψηλή απόδοση, αλλά την επίτευξη απόδοσης που να προσεγγίζει ή να αναπαράγει τη συμπεριφορά ενός συγκεκριμένου δείκτη αναφοράς (Beasley et al., 2003; Maringer & Oyewumi, 2007). Στην περίπτωση αυτή, η επενδυτική διαδικασία συνδέεται άμεσα με την παθητική διαχείριση χαρτοφυλακίου και με το πρόβλημα της αποτύπωσης δείκτη (Beasley et al., 2003; Canakgoz & Beasley, 2009).

Η αυξανόμενη πολυπλοκότητα των χρηματοοικονομικών αγορών, σε συνδυασμό με την ύπαρξη πολλαπλών αντικρουόμενων στόχων και περιορισμών, έχει καταστήσει σαφές ότι τα προβλήματα διαχείρισης επενδύσεων δεν μπορούν πάντοτε να αντιμετωπιστούν αποτελεσματικά με απλές ή αμιγώς κλασικές μεθόδους (Schaerf, 2002; Li et al., 2014). Σε

πολλές περιπτώσεις, οι επενδυτικές αποφάσεις οδηγούν σε προβλήματα βελτιστοποίησης υψηλής πολυπλοκότητας, στα οποία οι αντικειμενικές συναρτήσεις μπορεί να είναι μη γραμμικές ή μη ομαλές, ενώ οι περιορισμοί μπορεί να είναι ασυνεχείς ή συνδυαστικοί (Schaerf, 2002; Beasley et al., 2003; Canakgoz & Beasley, 2009).

1.2 Το πρόβλημα της βελτιστοποίησης χαρτοφυλακίου

Η βελτιστοποίηση χαρτοφυλακίου αποτελεί ένα από τα κεντρικά προβλήματα της χρηματοοικονομικής θεωρίας και πρακτικής, καθώς σχετίζεται με την επιλογή της καταλληλότερης σύνθεσης επενδυτικών τίτλων υπό συγκεκριμένους στόχους και περιορισμούς (Markowitz, 1952; Schaerf, 2002). Στον πυρήνα του προβλήματος βρίσκεται η ανάγκη προσδιορισμού ενός χαρτοφυλακίου το οποίο να επιτυγχάνει την καλύτερη δυνατή ισορροπία μεταξύ ανταγωνιστικών κριτηρίων, όπως είναι η αναμενόμενη απόδοση και ο κίνδυνος (Markowitz, 1952). Η διαδικασία αυτή δεν είναι απλή, διότι οι επενδυτικές επιλογές λαμβάνονται σε περιβάλλον αβεβαιότητας και επηρεάζονται από μεγάλο αριθμό παραγόντων, τόσο ποσοτικών όσο και ποιοτικών (Merton, 1969).

Η κλασική θεωρητική αφετηρία του προβλήματος εντοπίζεται στη θεωρία χαρτοφυλακίου του Markowitz, σύμφωνα με την οποία ο επενδυτής επιδιώκει τη μεγιστοποίηση της αναμενόμενης απόδοσης και την ταυτόχρονη ελαχιστοποίηση της διασποράς των αποδόσεων, η οποία χρησιμοποιείται ως μέτρο κινδύνου (Markowitz, 1952). Η προσέγγιση αυτή υπήρξε θεμελιώδης για την εξέλιξη της χρηματοοικονομικής επιστήμης, καθώς εισήγαγε μια συστηματική και μαθηματικά διατυπωμένη διαδικασία επιλογής χαρτοφυλακίου (Markowitz, 1952). Παρά τη μεγάλη της σημασία, η συγκεκριμένη θεώρηση βασίζεται σε ορισμένες υποθέσεις που συχνά δυσκολεύουν την άμεση εφαρμογή της σε πραγματικά επενδυτικά προβλήματα (Schaerf, 2002).

Στην πράξη, το γενικότερο πρόβλημα της βελτιστοποίησης χαρτοφυλακίου είναι σημαντικά πιο σύνθετο από την «απλή» βελτιστοποίηση μέσου-διασποράς (Schaerf, 2002; Beasley et al., 2003). Αυτό συμβαίνει επειδή ο επενδυτής ή ο διαχειριστής χαρτοφυλακίου δεν καλείται μόνο να επιτύχει έναν ευνοϊκό συνδυασμό απόδοσης και κινδύνου, αλλά να λάβει υπόψη και μια σειρά από πρόσθετους παράγοντες. Οι παράγοντες αυτοί περιλαμβάνουν, μεταξύ άλλων, το κόστος συναλλαγών, τους περιορισμούς στον αριθμό των διαθέσιμων τίτλων, τα ανώτατα και κατώτατα όρια συμμετοχής κάθε τίτλου στο χαρτοφυλάκιο, ζητήματα ρευστότητας, ανάγκες αναπροσαρμογής του χαρτοφυλακίου, καθώς και ειδικές απαιτήσεις που επιβάλλονται από θεσμικά ή κανονιστικά πλαίσια (Schaerf, 2002; Beasley et al., 2003).

Ένα επιπλέον χαρακτηριστικό του προβλήματος είναι η πολυκριτηριακή του φύση. Σε πολλές περιπτώσεις, δεν υπάρχει ένας και μοναδικός στόχος, αλλά ένα σύνολο στόχων που είναι εν μέρει αντικρουόμενοι (Markowitz, 1952; Steuer et al., 2005). Για παράδειγμα, η επιδίωξη χαμηλότερου κινδύνου μπορεί να οδηγήσει σε χαμηλότερη απόδοση, ενώ η επιθυμία για ακριβέστερη αποτύπωση ενός δείκτη μπορεί να συνεπάγεται υψηλότερο κόστος συναλλαγών ή μεγαλύτερο αριθμό τίτλων στο χαρτοφυλάκιο (Beasley et al., 2003; Canakgoz & Beasley, 2009). Επομένως, η βελτιστοποίηση χαρτοφυλακίου δεν συνίσταται απλώς στην εύρεση μιας «καλύτερης» λύσης, αλλά συχνά στην αναζήτηση λύσεων που επιτυγχάνουν ικανοποιητικούς συμβιβασμούς μεταξύ πολλών επιμέρους κριτηρίων (Steuer et al., 2005).

Η πολυπλοκότητα του προβλήματος ενισχύεται και από τη μαθηματική του δομή. Οι αντικειμενικές συναρτήσεις που χρησιμοποιούνται μπορεί να είναι μη γραμμικές, μη κυρτές ή ακόμη και μη ομαλές, ενώ αρκετοί από τους περιορισμούς που τίθενται στην πράξη είναι ασυνεχείς ή συνδυαστικοί (Schaerf, 2002; Beasley et al., 2003). Ιδιαίτερα όταν εισάγονται περιορισμοί πληθικότητας, δηλαδή όταν ζητείται το χαρτοφυλάκιο να περιλαμβάνει ακριβώς ή το πολύ έναν συγκεκριμένο αριθμό τίτλων, το πρόβλημα αποκτά έντονα συνδυαστικό χαρακτήρα (Schaerf, 2002). Αυτό σημαίνει ότι ο αριθμός των πιθανών λύσεων αυξάνεται ραγδαία με την αύξηση του πλήθους των διαθέσιμων τίτλων, καθιστώντας εξαιρετικά δύσκολη την εξαντλητική αναζήτηση της βέλτιστης λύσης (Schaerf, 2002; Beasley et al., 2003).

1.3 Η έννοια της αποτύπωσης δείκτη (Index Tracking)

Η αποτύπωση δείκτη αποτελεί μια ειδική περίπτωση του προβλήματος βελτιστοποίησης χαρτοφυλακίου και συνδέεται άμεσα με τη φιλοσοφία της παθητικής διαχείρισης επενδύσεων (Beasley et al., 2003; Maringer & Oyewumi, 2007). Βασικός στόχος της είναι η κατασκευή ενός χαρτοφυλακίου το οποίο να αναπαράγει όσο το δυνατόν πιστότερα τη συμπεριφορά ενός συγκεκριμένου χρηματιστηριακού δείκτη αναφοράς (Beasley et al., 2003; Meade & Salkin, 1989). Με άλλα λόγια, ο διαχειριστής δεν επιδιώκει να υπεραποδώσει σε σχέση με την αγορά μέσω ενεργητικών επιλογών, αλλά να διαμορφώσει ένα χαρτοφυλάκιο που να ακολουθεί την πορεία του επιλεγμένου δείκτη με τη μικρότερη δυνατή απόκλιση (Beasley et al., 2003; Derigs & Nickel, 2003).

Η λογική της αποτύπωσης δείκτη βασίζεται στην παραδοχή ότι, σε πολλές περιπτώσεις, η συστηματική παρακολούθηση ενός ευρέως αναγνωρισμένου δείκτη μπορεί να αποτελέσει αποτελεσματική επενδυτική στρατηγική (Meade & Salkin, 1989; Maringer & Oyewumi, 2007). Οι χρηματιστηριακοί δείκτες χρησιμοποιούνται ως δείκτες αναφοράς της συνολικής πορείας μιας αγοράς, ενός κλάδου ή μιας συγκεκριμένης κατηγορίας τίτλων, και επομένως η αναπαραγωγή της συμπεριφοράς τους επιτρέπει στον επενδυτή να εκτεθεί στη συνολική τάση της αγοράς χωρίς να απαιτείται συνεχής ενεργητική διαχείριση (Meade & Salkin, 1989; Beasley et al., 2003). Για τον λόγο αυτό, η αποτύπωση δείκτη συνδέεται στενά με τη λειτουργία των index funds και γενικότερα με τις στρατηγικές παθητικής επένδυσης (Meade & Salkin, 1989; Maringer & Oyewumi, 2007).

Θεωρητικά, ο απλούστερος τρόπος αποτύπωσης ενός δείκτη είναι η πλήρης αναπαραγωγή του, δηλαδή η κατοχή όλων των τίτλων που συμμετέχουν στον δείκτη με τα ίδια ή αναλογικά βάρη (Beasley et al., 2003). Μια τέτοια προσέγγιση οδηγεί σε πολύ ακριβή παρακολούθηση της πορείας του δείκτη, καθώς το χαρτοφυλάκιο αντανακλά άμεσα τη σύνθεσή του (Meade & Salkin, 1989; Beasley et al., 2003). Ωστόσο, στην πράξη η πλήρης αναπαραγωγή δεν είναι πάντοτε εφικτή ή επιθυμητή. Σε δείκτες που περιλαμβάνουν μεγάλο αριθμό τίτλων, η διατήρηση όλων των μετοχών μπορεί να συνεπάγεται υψηλό διαχειριστικό κόστος, δυσκολίες αναπροσαρμογής, μικρές θέσεις σε πολλούς τίτλους και αυξημένα κόστη συναλλαγών, ιδιαίτερα όταν η σύνθεση του δείκτη μεταβάλλεται συχνά (Beasley et al., 2003; Derigs & Nickel, 2003; Maringer & Oyewumi, 2007).

Για τον λόγο αυτό, στην πράξη συναντάται συχνά η μερική αναπαραγωγή δείκτη, όπου το χαρτοφυλάκιο περιλαμβάνει μόνο ένα υποσύνολο των διαθέσιμων τίτλων (Beasley et al., 2003; Maringer & Oyewumi, 2007). Στην περίπτωση αυτή, το βασικό πρόβλημα είναι να επιλεγούν οι κατάλληλοι τίτλοι και τα αντίστοιχα βάρη τους, έτσι ώστε η συνολική

συμπεριφορά του χαρτοφυλακίου να προσεγγίζει όσο το δυνατόν καλύτερα εκείνη του δείκτη (Beasley et al., 2003; Derigs & Nickel, 2003). Η διαδικασία αυτή μετατρέπει την αποτύπωση δείκτη σε ένα απαιτητικό πρόβλημα βελτιστοποίησης, στο οποίο ο διαχειριστής πρέπει να ισορροπήσει ανάμεσα στην ακρίβεια παρακολούθησης και στους πρακτικούς περιορισμούς που επιβάλλονται από την αγορά και την επενδυτική στρατηγική (Beasley et al., 2003; Maringer & Oyewumi, 2007; Derigs & Nickel, 2003).

Κεντρική έννοια στο πρόβλημα της αποτύπωσης δείκτη είναι το σφάλμα αποτύπωσης, δηλαδή το μέτρο της απόκλισης μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη αναφοράς (Derigs & Nickel, 2003; Beasley et al., 2003). Όσο μικρότερη είναι αυτή η απόκλιση, τόσο πιο επιτυχημένη θεωρείται η διαδικασία αποτύπωσης (Derigs & Nickel, 2003). Στην πράξη, όμως, η ελαχιστοποίηση του σφάλματος αποτύπωσης δεν είναι ο μόνος στόχος. Συχνά πρέπει να ληφθούν υπόψη επιπλέον παράγοντες, όπως το κόστος συναλλαγών, ο περιορισμός στον αριθμό των τίτλων που επιτρέπεται να περιλαμβάνει το χαρτοφυλάκιο, τα ανώτατα και κατώτατα όρια βαρών κάθε τίτλου, καθώς και η ανάγκη περιοδικής αναπροσαρμογής του χαρτοφυλακίου. Έτσι, η αποτύπωση δείκτη αποκτά έναν έντονα ρεαλιστικό και πολυπαραγοντικό χαρακτήρα (Beasley et al., 2003; Derigs & Nickel, 2003).

1.4 Οι Γενετικοί και Εξελικτικοί Αλγόριθμοι ως εργαλεία επίλυσης προβλημάτων βελτιστοποίησης

Η επίλυση σύνθετων προβλημάτων βελτιστοποίησης έχει απασχολήσει εκτενώς τη διεθνή επιστημονική κοινότητα, ιδιαίτερα σε πεδία όπου οι κλασικές αναλυτικές μέθοδοι αδυνατούν να προσφέρουν αποδοτικές και πρακτικά εφαρμόσιμες λύσεις (Whitley, 1994; Schaerf, 2002). Η δυσκολία αυτή είναι εντονότερη σε προβλήματα που χαρακτηρίζονται από μη γραμμικότητα, ασυνέχεια, μεγάλο αριθμό μεταβλητών, πολλαπλούς περιορισμούς και πληθώρα τοπικών βέλτιστων λύσεων (Whitley, 1994; Chang et al., 2000). Σε αυτό το πλαίσιο, ιδιαίτερη σημασία έχουν αποκτήσει οι εξελικτικοί αλγόριθμοι, δηλαδή μια ευρύτερη κατηγορία υπολογιστικών μεθόδων που εμπνέονται από τις αρχές της φυσικής εξέλιξης και της φυσικής επιλογής (Beasley et al., 1993; Whitley, 1994).

Οι εξελικτικοί αλγόριθμοι βασίζονται στη σταδιακή βελτίωση ενός πληθυσμού υποψήφιων λύσεων μέσω διαδικασιών που προσομοιώνουν βιολογικούς μηχανισμούς, όπως η επιλογή, η αναπαραγωγή, η διασταύρωση και η μετάλλαξη (Beasley et al., 1993; Whitley, 1994). Αντί να επιχειρούν την άμεση κατασκευή μιας βέλτιστης λύσης με ντετερμινιστικό τρόπο, εξερευνούν επαναληπτικά τον χώρο αναζήτησης, διατηρώντας και συνδυάζοντας τα πιο υποσχόμενα χαρακτηριστικά των λύσεων που εμφανίζονται σε κάθε στάδιο (Whitley, 1994). Με αυτόν τον τρόπο επιδιώκουν τη σταδιακή προσέγγιση λύσεων υψηλής ποιότητας, ακόμη και όταν ο χώρος λύσεων είναι ιδιαίτερα πολύπλοκος ή δεν είναι εφικτή η αναλυτική αντιμετώπιση του προβλήματος (Beasley et al., 1993; Whitley, 1994).

Μία από τις σημαντικότερες και περισσότερο διαδεδομένες κατηγορίες εξελικτικών μεθόδων είναι οι Γενετικοί Αλγόριθμοι. Οι αλγόριθμοι αυτοί αναπαριστούν κάθε υποψήφια λύση ως ένα χρωμόσωμα, δηλαδή ως μια κωδικοποιημένη μορφή των χαρακτηριστικών της λύσης, και στη συνέχεια εφαρμόζουν επαναληπτικά τελεστές εξέλιξης πάνω σε έναν πληθυσμό τέτοιων χρωμοσωμάτων (Whitley, 1994; Beasley et al., 1993). Η ποιότητα κάθε λύσης αξιολογείται μέσω μιας συνάρτησης καταλληλότητας, η οποία αποτυπώνει το πόσο

καλά η συγκεκριμένη λύση ικανοποιεί τον στόχο του προβλήματος (Whitley, 1994). Οι καλύτερες λύσεις έχουν μεγαλύτερη πιθανότητα να επιλεγούν για την παραγωγή απογόνων, ενώ μέσω της διασταύρωσης και της μετάλλαξης δημιουργούνται νέες λύσεις που ενδέχεται να παρουσιάζουν βελτιωμένα χαρακτηριστικά (Beasley et al., 1993; Whitley, 1994).

Η ιδιαίτερη αξία των Γενετικών Αλγορίθμων έγκειται στην ικανότητά τους να αντιμετωπίζουν προβλήματα βελτιστοποίησης στα οποία δεν είναι εύκολα διαθέσιμες πληροφορίες όπως παράγωγοι, κυρτότητα ή αναλυτική μορφή της βέλτιστης λύσης (Whitley, 1994). Επιπλέον, λόγω του πληθυσμιακού χαρακτήρα τους, δεν εξετάζουν μια μόνο λύση κάθε φορά, αλλά πολλές ταυτόχρονα, γεγονός που τους επιτρέπει να διερευνούν αποτελεσματικότερα διαφορετικές περιοχές του χώρου λύσεων (Whitley, 1994; Beasley et al., 1993). Αυτό αυξάνει την πιθανότητα αποφυγής πρόωρης σύγκλισης σε κατώτερες τοπικές λύσεις και ενισχύει την ευελιξία τους σε προβλήματα όπου συνυπάρχουν συνεχείς και διακριτές μεταβλητές (Whitley, 1994; Chang et al., 2000).

Τα χαρακτηριστικά αυτά καθιστούν τους Γενετικούς Αλγορίθμους ιδιαίτερα κατάλληλους για προβλήματα χρηματοοικονομικής βελτιστοποίησης και ειδικότερα για προβλήματα κατασκευής χαρτοφυλακίων (Chang et al., 2000; Chang et al., 2009). Στο πεδίο αυτό, η ανάγκη ταυτόχρονης επιλογής τίτλων και προσδιορισμού των αντίστοιχων βαρών τους, υπό περιορισμούς όπως ο διαθέσιμος προϋπολογισμός, το κόστος συναλλαγών, τα όρια συμμετοχής κάθε τίτλου και ο μέγιστος επιτρεπτός αριθμός τίτλων, οδηγεί σε πολύπλοκα προβλήματα με έντονο συνδυαστικό χαρακτήρα (Chang et al., 2000; Lin & Liu, 2008; Anagnostopoulos & Mamanis, 2011). Τα προβλήματα αυτά δύσκολα αντιμετωπίζονται με κλασικές τεχνικές, ιδιαίτερα όταν απαιτείται ρεαλιστική μοντελοποίηση των συνθηκών της αγοράς (Schaerf, 2002; Chang et al., 2000).

Στην περίπτωση του προβλήματος της αποτύπωσης δείκτη, η χρήση Γενετικών Αλγορίθμων αποκτά ακόμη μεγαλύτερο ενδιαφέρον. Το ζητούμενο εδώ δεν είναι απλώς η επιλογή ενός αποδοτικού χαρτοφυλακίου με όρους κινδύνου και απόδοσης, αλλά η κατασκευή ενός χαρτοφυλακίου που να ακολουθεί όσο το δυνατόν πιστότερα την πορεία ενός δείκτη αναφοράς (Beasley et al., 2003; Oh et al., 2005). Η απαίτηση αυτή συνδυάζεται συνήθως με πρόσθετους περιορισμούς, όπως ο περιορισμός πληθικότητας, τα όρια στα βάρη των τίτλων και το κόστος αναπροσαρμογής του χαρτοφυλακίου (Beasley et al., 2003; Derigs & Nickel, 2003; Oh et al., 2005). Οι Γενετικοί Αλγόριθμοι προσφέρουν ένα ευέλικτο πλαίσιο μέσα στο οποίο είναι δυνατόν να ενσωματωθούν όλα τα παραπάνω στοιχεία, τόσο στη μορφή της αναπαράστασης των λύσεων όσο και στη διαδικασία αξιολόγησης και επιλογής τους (Oh et al., 2005; Beasley et al., 2003).

Ένα ακόμη σημαντικό πλεονέκτημα των Γενετικών Αλγορίθμων είναι ότι επιτρέπουν την εύκολη προσαρμογή τους στις ιδιαιτερότητες του υπό εξέταση προβλήματος. Ο τρόπος κωδικοποίησης των χρωμοσωμάτων, η μορφή της συνάρτησης καταλληλότητας, οι τελεστές επιλογής, διασταύρωσης και μετάλλαξης, καθώς και οι μηχανισμοί επιδιόρθωσης λύσεων που παραβιάζουν περιορισμούς, μπορούν να σχεδιαστούν έτσι ώστε να ανταποκρίνονται στις ανάγκες του συγκεκριμένου προβλήματος αποτύπωσης δείκτη (Whitley, 1994; Oh et al., 2005). Με αυτόν τον τρόπο, ο αλγόριθμος δεν αποτελεί μια γενική μέθοδο αναζήτησης, αλλά μια στοχευμένη υπολογιστική διαδικασία προσαρμοσμένη στα ιδιαίτερα χαρακτηριστικά της εφαρμογής (Oh et al., 2005; Anagnostopoulos & Mamanis, 2011).

1.5 Αντικείμενο της εργασίας

Αντικείμενο της παρούσας διπλωματικής εργασίας αποτελεί η μελέτη και αντιμετώπιση του προβλήματος της βελτιστοποίησης χαρτοφυλακίου αποτύπωσης δείκτη με τη χρήση Γενετικών Αλγορίθμων. Ειδικότερα, εξετάζεται το πρόβλημα κατασκευής ενός χαρτοφυλακίου το οποίο, από ένα διαθέσιμο σύνολο μετοχών, επιλέγει ένα κατάλληλο υποσύνολο τίτλων και καθορίζει τα αντίστοιχα βάρη συμμετοχής τους, έτσι ώστε να αναπαράγει όσο το δυνατόν πιστότερα την πορεία ενός χρηματιστηριακού δείκτη αναφοράς.

Η εργασία εντάσσεται στο ευρύτερο πεδίο της χρηματοοικονομικής μηχανικής και της υπολογιστικής βελτιστοποίησης, καθώς συνδυάζει τη θεωρητική διάσταση της διαχείρισης χαρτοφυλακίου με την ανάπτυξη αλγοριθμικών τεχνικών για την επίλυση ενός σύνθετου πρακτικού προβλήματος. Σε αντίθεση με πιο απλουστευμένες διατυπώσεις της βελτιστοποίησης χαρτοφυλακίου, εδώ το ζητούμενο δεν είναι γενικά η επίτευξη της βέλτιστης σχέσης απόδοσης και κινδύνου, αλλά η ελαχιστοποίηση της απόκλισης μεταξύ του χαρτοφυλακίου και του δείκτη, υπό την ταυτόχρονη ικανοποίηση συγκεκριμένων περιορισμών.

Συγκεκριμένα, το πρόβλημα που εξετάζεται περιλαμβάνει δύο αλληλένδετες αποφάσεις. Η πρώτη αφορά την επιλογή των τίτλων που θα συμμετέχουν στο χαρτοφυλάκιο, δηλαδή τον καθορισμό ενός υποσυνόλου μετοχών από το αρχικό σύνολο των διαθέσιμων τίτλων. Η δεύτερη αφορά τον προσδιορισμό των βαρών των επιλεγμένων τίτλων, ώστε η τελική σύνθεση του χαρτοφυλακίου να ακολουθεί κατά το δυνατόν τη συμπεριφορά του δείκτη αναφοράς. Η ταυτόχρονη αντιμετώπιση των δύο αυτών αποφάσεων καθιστά το πρόβλημα ιδιαίτερα απαιτητικό, καθώς συνδυάζει διακριτές και συνεχείς πτυχές βελτιστοποίησης. Στο πλαίσιο της εργασίας, το πρόβλημα προσεγγίζεται υπό ρεαλιστικές συνθήκες, οι οποίες αντανακλούν τις απαιτήσεις της πραγματικής επενδυτικής πρακτικής. Για τον λόγο αυτό, λαμβάνονται υπόψη περιορισμοί όπως ο διαθέσιμος προϋπολογισμός, ο περιορισμός στον αριθμό των τίτλων που μπορούν να περιλαμβάνονται στο χαρτοφυλάκιο, τα κατώτατα και ανώτατα όρια συμμετοχής κάθε τίτλου, καθώς και το κόστος συναλλαγών. Η ύπαρξη αυτών των περιορισμών προσδίδει στο πρόβλημα πιο ρεαλιστικό χαρακτήρα, αλλά ταυτόχρονα αυξάνει σημαντικά τη δυσκολία της επίλυσής του.

Για την αντιμετώπιση του προβλήματος προτείνεται η ανάπτυξη ενός Γενετικού Αλγορίθμου ειδικά προσαρμοσμένου στις απαιτήσεις της αποτύπωσης δείκτη. Η επιλογή της συγκεκριμένης μεθοδολογίας βασίζεται στην ικανότητα των Γενετικών Αλγορίθμων να αντιμετωπίζουν προβλήματα υψηλής πολυπλοκότητας, με μη γραμμικούς ή ασυνεχείς περιορισμούς και μεγάλο χώρο αναζήτησης πιθανών λύσεων. Στο πλαίσιο αυτό, η εργασία δεν περιορίζεται στη θεωρητική μελέτη του προβλήματος, αλλά επεκτείνεται και στον σχεδιασμό, την υλοποίηση και την πειραματική αξιολόγηση του προτεινόμενου αλγορίθμου.

1.6 Σκοπός της εργασίας

Σκοπός της παρούσας διπλωματικής εργασίας είναι η διερεύνηση της εφαρμογής Γενετικών Αλγορίθμων στο πρόβλημα της βελτιστοποίησης χαρτοφυλακίου αποτύπωσης δείκτη, με στόχο την ανάπτυξη μιας αποτελεσματικής και πρακτικά εφαρμόσιμης μεθοδολογίας για την κατασκευή χαρτοφυλακίων που αναπαράγουν όσο το δυνατόν πιστότερα την πορεία ενός χρηματιστηριακού δείκτη αναφοράς. Η διερεύνηση αυτή εντάσσεται στο ευρύτερο πλαίσιο της χρηματοοικονομικής μηχανικής, όπου η λήψη επενδυτικών αποφάσεων απαιτεί συνδυασμό θεωρητικής ανάλυσης, μαθηματικής μοντελοποίησης και υπολογιστικών τεχνικών βελτιστοποίησης.

Ειδικότερα, η εργασία αποσκοπεί στη μελέτη του προβλήματος της αποτύπωσης δείκτη ως σύνθετου προβλήματος βελτιστοποίησης χαρτοφυλακίου, στο οποίο πρέπει να επιλεγεί ένα κατάλληλο υποσύνολο μετοχών και να προσδιοριστούν τα αντίστοιχα βάρη συμμετοχής τους, έτσι ώστε να ελαχιστοποιείται η απόκλιση από τη συμπεριφορά του δείκτη. Ο σκοπός αυτός δεν περιορίζεται στην απλή θεωρητική περιγραφή του προβλήματος, αλλά επεκτείνεται στην ανάπτυξη μιας μεθοδολογίας που να μπορεί να λειτουργήσει υπό ρεαλιστικούς περιορισμούς, όπως ο περιορισμένος αριθμός τίτλων στο χαρτοφυλάκιο, τα όρια στα βάρη των μετοχών και το κόστος συναλλαγών.

Παράλληλα, σκοπός της εργασίας είναι να αναδείξει τη χρησιμότητα των Γενετικών Αλγορίθμων ως κατάλληλων εργαλείων επίλυσης για προβλήματα υψηλής πολυπλοκότητας, στα οποία οι κλασικές τεχνικές βελτιστοποίησης δυσκολεύονται να εφαρμοστούν αποτελεσματικά. Η μελέτη αυτή αποκτά ιδιαίτερη σημασία, καθώς το πρόβλημα της αποτύπωσης δείκτη περιλαμβάνει τόσο συνδυαστικά στοιχεία, όπως η επιλογή τίτλων, όσο και συνεχή στοιχεία, όπως ο καθορισμός των βαρών του χαρτοφυλακίου. Επομένως, η εργασία επιδιώκει να δείξει κατά πόσο ένας κατάλληλα σχεδιασμένος Γενετικός Αλγόριθμος μπορεί να αντιμετωπίσει αποτελεσματικά αυτή τη μικτή φύση του προβλήματος.

Επιπλέον, η εργασία στοχεύει στον σχεδιασμό και την ανάπτυξη ενός προσαρμοσμένου αλγορίθμου, ο οποίος δεν θα αποτελεί απλώς μια γενική εφαρμογή της εξελικτικής υπολογιστικής, αλλά μια ειδικά διαμορφωμένη μέθοδο για το συγκεκριμένο πρόβλημα. Αυτό σημαίνει ότι ο τρόπος αναπαράστασης των λύσεων, η διαμόρφωση της συνάρτησης καταλληλότητας, η επιλογή των γενετικών τελεστών και ο τρόπος διαχείρισης των περιορισμών θα σχεδιαστούν με γνώμονα τις απαιτήσεις της αποτύπωσης δείκτη. Κατά συνέπεια, ο σκοπός της εργασίας περιλαμβάνει και την αλγοριθμική προσαρμογή της μεθόδου στις ιδιαιτερότητες του υπό εξέταση χρηματοοικονομικού προβλήματος.

Ένας ακόμη βασικός στόχος που ενσωματώνεται στον συνολικό σκοπό της εργασίας είναι η πειραματική αξιολόγηση της προτεινόμενης μεθοδολογίας σε πραγματικά χρηματιστηριακά δεδομένα. Η εμπειρική αυτή διάσταση είναι ιδιαίτερα σημαντική, διότι επιτρέπει να εξεταστεί όχι μόνο η θεωρητική καταλληλότητα της προτεινόμενης προσέγγισης, αλλά και η πρακτική της αποτελεσματικότητα. Μέσα από την εφαρμογή σε πραγματικά δεδομένα και τη σύγκριση με αποτελέσματα άλλων μεθόδων, επιδιώκεται η εξαγωγή συμπερασμάτων σχετικά με την ποιότητα των λύσεων που παράγει ο αλγόριθμος, την ευαισθησία του στις παραμέτρους του και τη δυνατότητά του να αποτελέσει μια αξιόπιστη υπολογιστική προσέγγιση για το πρόβλημα.

1.7 Στόχοι της εργασίας

Οι επιμέρους στόχοι της παρούσας διπλωματικής εργασίας απορρέουν από τον γενικό σκοπό της και αφορούν τόσο τη θεωρητική μελέτη του προβλήματος της αποτύπωσης δείκτη όσο και την αλγοριθμική και πειραματική διερεύνηση της προτεινόμενης μεθοδολογίας. Η επίτευξη των στόχων αυτών είναι απαραίτητη ώστε η εργασία να αποκτήσει πληρότητα, συνέπεια και ερευνητική τεκμηρίωση, καλύπτοντας όλα τα βασικά στάδια από τη βιβλιογραφική θεμελίωση έως την αξιολόγηση των αποτελεσμάτων.

Πρώτος στόχος της εργασίας είναι η μελέτη της σχετικής βιβλιογραφίας που αφορά τη βελτιστοποίηση χαρτοφυλακίου και ειδικότερα το πρόβλημα της αποτύπωσης δείκτη. Στο πλαίσιο αυτό επιδιώκεται η κατανόηση των βασικών εννοιών, της χρηματοοικονομικής σημασίας του προβλήματος, των διαφορετικών μαθηματικών διατυπώσεων που έχουν προταθεί, καθώς και των κυριότερων μεθόδων που έχουν χρησιμοποιηθεί για την αντιμετώπισή του. Ιδιαίτερη έμφαση δίνεται στις προσεγγίσεις που αξιοποιούν ευρετικές, μεταερευτικές και εξελικτικές τεχνικές, ώστε να αποσαφηνιστεί το θεωρητικό υπόβαθρο πάνω στο οποίο θα στηριχθεί η προτεινόμενη μεθοδολογία.

Δεύτερος στόχος είναι η ανάλυση του προβλήματος ως προβλήματος βελτιστοποίησης χαρτοφυλακίου υπό περιορισμούς. Η εργασία επιδιώκει να παρουσιάσει με σαφήνεια τη μαθηματική δομή του προβλήματος, ορίζοντας την αντικειμενική συνάρτηση, τις μεταβλητές απόφασης, τις βασικές παραμέτρους και τους περιορισμούς που το χαρακτηρίζουν. Η διατύπωση αυτή είναι ιδιαίτερα σημαντική, διότι επιτρέπει τη μετάβαση από τη θεωρητική περιγραφή του προβλήματος στην ανάπτυξη μιας συστηματικής μεθόδου επίλυσης και καθιστά δυνατή την ακριβή αποτύπωση των συνθηκών μέσα στις οποίες θα λειτουργεί ο αλγόριθμος.

Τρίτος στόχος της εργασίας είναι ο σχεδιασμός ενός Γενετικού Αλγορίθμου κατάλληλα προσαρμοσμένου στο πρόβλημα της αποτύπωσης δείκτη. Ο στόχος αυτός περιλαμβάνει την επιλογή κατάλληλης αναπαράστασης των λύσεων, τον ορισμό συνάρτησης καταλληλότητας που να αντανakλά σωστά το σφάλμα αποτύπωσης και τους περιορισμούς του προβλήματος, καθώς και τη διαμόρφωση των βασικών γενετικών τελεστών, όπως η επιλογή, η διασταύρωση και η μετάλλαξη. Παράλληλα, επιδιώκεται η ανάπτυξη μηχανισμών που θα εξασφαλίζουν, στο μέτρο του δυνατού, τη δημιουργία και διατήρηση εφικτών λύσεων κατά τη διάρκεια της εξελικτικής διαδικασίας.

Τέταρτος στόχος είναι η υλοποίηση της προτεινόμενης μεθόδου σε περιβάλλον Python. Η υλοποίηση αυτή δεν αφορά μόνο τη μετατροπή της θεωρητικής μεθοδολογίας σε λειτουργικό κώδικα, αλλά και τη δημιουργία ενός ολοκληρωμένου πειραματικού πλαισίου που να επιτρέπει την εκτέλεση δοκιμών, τη μεταβολή παραμέτρων, την αποθήκευση αποτελεσμάτων και τη συγκριτική αξιολόγηση διαφορετικών ρυθμίσεων του αλγορίθμου. Η επιλογή της Python εξυπηρετεί τόσο τη σαφήνεια και ευελιξία της υλοποίησης όσο και την αξιοποίηση των υπολογιστικών εργαλείων που διαθέτει για ανάλυση δεδομένων και πειραματική διερεύνηση.

Πέμπτος στόχος της εργασίας είναι η αξιολόγηση της απόδοσης του προτεινόμενου αλγορίθμου με χρήση πραγματικών χρηματιστηριακών δεδομένων. Η επίτευξη αυτού του στόχου απαιτεί την επιλογή κατάλληλων δεδομένων, την περιγραφή των βασικών χαρακτηριστικών τους και τον σχεδιασμό ενός αξιόπιστου πειραματικού πλαισίου. Στο πλαίσιο αυτό, επιδιώκεται να εξεταστεί κατά πόσο ο προτεινόμενος Γενετικός Αλγόριθμος

μπορεί να κατασκευάσει χαρτοφυλάκια με μικρό σφάλμα αποτύπωσης, να ικανοποιεί τους τεθέντες περιορισμούς και να εμφανίζει σταθερή συμπεριφορά σε διαφορετικές συνθήκες και παραμετροποιήσεις.

Έκτος στόχος είναι η διερεύνηση της επίδρασης των παραμέτρων και των γενετικών τελεστών στην απόδοση του αλγορίθμου. Στο πλαίσιο αυτό, η εργασία επιδιώκει να αναλύσει πώς μεταβάλλονται τα αποτελέσματα όταν αλλάζουν στοιχεία όπως το μέγεθος του πληθυσμού, η πιθανότητα διασταύρωσης, η πιθανότητα μετάλλαξης, ο αριθμός γενεών και η επιλογή συγκεκριμένων τελεστών. Η διερεύνηση αυτή είναι κρίσιμη, καθώς επιτρέπει την καλύτερη κατανόηση της συμπεριφοράς της μεθόδου και τη διαμόρφωση τεκμηριωμένων συμπερασμάτων για τις πιο αποδοτικές επιλογές παραμετροποίησης.

Έβδομος στόχος είναι η σύγκριση της προτεινόμενης μεθόδου με άλλες τεχνικές επίλυσης του προβλήματος που έχουν παρουσιαστεί στη σχετική βιβλιογραφία ή με κατάλληλες βασικές προσεγγίσεις αναφοράς. Η σύγκριση αυτή είναι απαραίτητη ώστε να αξιολογηθεί η θέση της προτεινόμενης μεθοδολογίας σε σχέση με εναλλακτικές λύσεις και να εντοπιστούν τα σημεία στα οποία υπερτερεί ή υστερεί. Μέσα από αυτή τη διαδικασία επιδιώκεται η εξαγωγή πιο ουσιαστικών συμπερασμάτων σχετικά με τη χρησιμότητα των Γενετικών Αλγορίθμων στο συγκεκριμένο πρόβλημα.

Τέλος, στόχος της εργασίας είναι η διατύπωση τεκμηριωμένων συμπερασμάτων σχετικά με την αποτελεσματικότητα, τα πλεονεκτήματα, τους περιορισμούς και τις δυνατότητες περαιτέρω βελτίωσης της προτεινόμενης προσέγγισης. Η εργασία δεν αποσκοπεί μόνο στην παρουσίαση ενός λειτουργικού αλγορίθμου, αλλά και στην κριτική αποτίμηση της μεθόδου, ώστε να αναδειχθούν τόσο η πρακτική της αξία όσο και οι κατευθύνσεις στις οποίες θα μπορούσε να επεκταθεί ή να βελτιωθεί σε μελλοντικές ερευνητικές προσπάθειες.

1.8 Ερευνητικά ερωτήματα

Η παρούσα διπλωματική εργασία αναπτύσσεται γύρω από ένα σύνολο ερευνητικών ερωτημάτων που απορρέουν από τη φύση του προβλήματος της αποτύπωσης δείκτη και από την επιλογή των Γενετικών Αλγορίθμων ως βασικού εργαλείου επίλυσής του. Τα ερωτήματα αυτά λειτουργούν ως άξονες καθοδήγησης της έρευνας, καθώς συνδέουν τη θεωρητική μελέτη του προβλήματος με τη σχεδίαση της μεθόδου, την υλοποίηση του αλγορίθμου και την αξιολόγηση των αποτελεσμάτων.

Κεντρικό ερευνητικό ερώτημα της εργασίας αποτελεί το κατά πόσο ένας Γενετικός Αλγόριθμος μπορεί να χρησιμοποιηθεί αποτελεσματικά για την επίλυση του προβλήματος της βελτιστοποίησης χαρτοφυλακίου αποτύπωσης δείκτη. Το ερώτημα αυτό είναι θεμελιώδες, διότι αφορά την ίδια τη δυνατότητα της εξελικτικής προσέγγισης να αντιμετωπίσει ένα πρόβλημα με συνδυαστικό χαρακτήρα, πολλαπλούς περιορισμούς και αυξημένη υπολογιστική δυσκολία. Η απάντησή του δεν είναι μόνο θεωρητικού ενδιαφέροντος, αλλά συνδέεται και με την πρακτική χρησιμότητα της προτεινόμενης μεθοδολογίας.

Σε άμεση συνάφεια με το προηγούμενο, τίθεται το ερώτημα κατά πόσο η προτεινόμενη προσέγγιση μπορεί να κατασκευάζει χαρτοφυλάκια που αναπαράγουν με ικανοποιητική ακρίβεια την πορεία ενός δείκτη αναφοράς. Εδώ το ενδιαφέρον εστιάζεται στη δυνατότητα του αλγορίθμου να επιτυγχάνει μικρό σφάλμα αποτύπωσης, διατηρώντας παράλληλα

εφικτές και ρεαλιστικές λύσεις. Το ερώτημα αυτό εξετάζει ουσιαστικά την ποιότητα των παραγόμενων χαρτοφυλακίων και τη χρηματοοικονομική επάρκεια των λύσεων που προκύπτουν.

Ένα τρίτο βασικό ερευνητικό ερώτημα αφορά την επίδραση των περιορισμών του προβλήματος στην απόδοση της μεθόδου. Πιο συγκεκριμένα, διερευνάται σε ποιο βαθμό περιορισμοί όπως ο περιορισμός πληθικότητας, τα κατώτατα και ανώτατα όρια συμμετοχής των τίτλων, ο διαθέσιμος προϋπολογισμός και το κόστος συναλλαγών επηρεάζουν την ικανότητα του αλγορίθμου να βρίσκει αποτελεσματικές λύσεις. Το ερώτημα αυτό είναι ιδιαίτερα σημαντικό, διότι οι περιορισμοί αυτοί προσδίδουν ρεαλισμό στο μοντέλο, αλλά ταυτόχρονα αυξάνουν την πολυπλοκότητα της αναζήτησης.

Επιπλέον, η εργασία διερευνά το κατά πόσο η επιλογή της αναπαράστασης των λύσεων και των επιμέρους μηχανισμών του Γενετικού Αλγορίθμου επηρεάζει την αποτελεσματικότητα της μεθόδου. Συγκεκριμένα, τίθεται το ερώτημα αν ο τρόπος κωδικοποίησης του χρωμοσώματος, η μορφή της συνάρτησης καταλληλότητας, ο τρόπος χειρισμού των περιορισμών και η επιλογή των γενετικών τελεστών μπορούν να βελτιώσουν ή να επιβαρύνουν την τελική απόδοση του αλγορίθμου. Η διερεύνηση αυτού του ερωτήματος είναι ουσιαστική, καθώς σε προβλήματα όπως το index tracking η απόδοση μιας εξελικτικής μεθόδου εξαρτάται σε μεγάλο βαθμό από τον ειδικό σχεδιασμό της.

Ένα ακόμη ερευνητικό ερώτημα αφορά την επίδραση των παραμέτρων του αλγορίθμου στα αποτελέσματά του. Η εργασία εξετάζει αν και σε ποιο βαθμό παράμετροι όπως το μέγεθος του πληθυσμού, η πιθανότητα διασταύρωσης, η πιθανότητα μετάλλαξης, ο αριθμός γενεών και η στρατηγική ελιτισμού επηρεάζουν την ποιότητα των λύσεων, τη σταθερότητα της συμπεριφοράς του αλγορίθμου και τον απαιτούμενο υπολογιστικό χρόνο.

Παράλληλα, τίθεται το ερώτημα κατά πόσο η προτεινόμενη μεθοδολογία μπορεί να σταθεί ανταγωνιστικά σε σύγκριση με άλλες προσεγγίσεις που έχουν χρησιμοποιηθεί στη σχετική βιβλιογραφία. Η σύγκριση αυτή μπορεί να αφορά είτε άλλες ευρετικές και μεταευρετικές τεχνικές είτε κλασικότερες μεθόδους επίλυσης, ανάλογα με τα διαθέσιμα συγκριτικά αποτελέσματα και τα δεδομένα που θα χρησιμοποιηθούν. Το συγκεκριμένο ερώτημα είναι ιδιαίτερα σημαντικό, διότι επιτρέπει να αξιολογηθεί όχι μόνο αν η προτεινόμενη μέθοδος λειτουργεί, αλλά και ποια είναι η θέση της σε σχέση με τις ήδη γνωστές εναλλακτικές.

Τέλος, η εργασία εξετάζει κατά πόσο η χρήση Γενετικών Αλγορίθμων αποτελεί μια γενικότερα κατάλληλη και υποσχόμενη κατεύθυνση για την αντιμετώπιση προβλημάτων χρηματοοικονομικής βελτιστοποίησης με σύνθετη δομή. Το ερώτημα αυτό υπερβαίνει εν μέρει τα στενά όρια του συγκεκριμένου προβλήματος και συνδέει τα ευρήματα της εργασίας με το ευρύτερο πεδίο της υπολογιστικής χρηματοοικονομικής. Η απάντησή του αναμένεται να συμβάλει στην αποτίμηση της εξελικτικής υπολογιστικής ως εργαλείου για ρεαλιστικά προβλήματα λήψης επενδυτικών αποφάσεων.

1.9 Μεθοδολογική προσέγγιση της εργασίας

Η μεθοδολογική προσέγγιση της παρούσας διπλωματικής εργασίας βασίζεται στον συνδυασμό θεωρητικής ανάλυσης, μαθηματικής μοντελοποίησης, αλγοριθμικού σχεδιασμού, υπολογιστικής υλοποίησης και πειραματικής αξιολόγησης. Η επιλογή αυτής της πολυεπίπεδης προσέγγισης κρίνεται αναγκαία, καθώς το πρόβλημα της αποτύπωσης δείκτη δεν αποτελεί μόνο ένα θεωρητικό ζήτημα της χρηματοοικονομικής, αλλά και ένα σύνθετο εφαρμοσμένο πρόβλημα βελτιστοποίησης, το οποίο απαιτεί τόσο εννοιολογική αποσαφήνιση όσο και πρακτική διερεύνηση.

Σε πρώτο στάδιο, η εργασία στηρίζεται σε συστηματική μελέτη της σχετικής βιβλιογραφίας. Στο πλαίσιο αυτό εξετάζονται η θεωρία της βελτιστοποίησης χαρτοφυλακίου, οι βασικές έννοιες της παθητικής διαχείρισης και της αποτύπωσης δείκτη, καθώς και οι κυριότερες μεθοδολογικές προσεγγίσεις που έχουν χρησιμοποιηθεί για την αντιμετώπιση του προβλήματος. Ιδιαίτερη έμφαση δίνεται στις ευρετικές, μεταερευνητικές και εξελικτικές τεχνικές, καθώς αυτές συνδέονται άμεσα με τη βασική μεθοδολογική κατεύθυνση της παρούσας εργασίας. Η βιβλιογραφική αυτή διερεύνηση δεν λειτουργεί μόνο ως θεωρητικό υπόβαθρο, αλλά και ως μέσο εντοπισμού των βασικών στοιχείων που πρέπει να ενσωματωθούν στο προτεινόμενο μοντέλο.

Σε δεύτερο στάδιο, το πρόβλημα διατυπώνεται αυστηρά ως πρόβλημα βελτιστοποίησης. Η διαδικασία αυτή περιλαμβάνει τον σαφή ορισμό της αντικειμενικής συνάρτησης, των μεταβλητών απόφασης, των παραμέτρων και των περιορισμών που χαρακτηρίζουν το πρόβλημα της αποτύπωσης δείκτη. Στο σημείο αυτό η εργασία ακολουθεί μια προσέγγιση συμβατή με τη σχετική επιστημονική βιβλιογραφία, ώστε η μαθηματική διατύπωση να είναι σαφής, συνεπής και κατάλληλη για αλγοριθμική επεξεργασία. Η φάση αυτή είναι καθοριστική, διότι μετατρέπει το χρηματοοικονομικό πρόβλημα σε δομημένο υπολογιστικό πρόβλημα, πάνω στο οποίο μπορεί να αναπτυχθεί η προτεινόμενη μέθοδος επίλυσης.

Στη συνέχεια, η εργασία προχωρά στον σχεδιασμό ενός Γενετικού Αλγορίθμου προσαρμοσμένου στο συγκεκριμένο πρόβλημα. Ο σχεδιασμός αυτός δεν βασίζεται σε μια γενική ή αφηρημένη εφαρμογή της μεθόδου, αλλά σε στοχευμένες επιλογές που λαμβάνουν υπόψη τις ιδιαιτερότητες του index tracking. Πιο συγκεκριμένα, καθορίζεται ο τρόπος αναπαράστασης κάθε υποψήφιας λύσης, η μορφή της συνάρτησης καταλληλότητας, οι τελεστές επιλογής, διασταύρωσης και μετάλλαξης, καθώς και οι μηχανισμοί διαχείρισης των περιορισμών. Στόχος αυτής της φάσης είναι η ανάπτυξη ενός αλγορίθμου που να ανταποκρίνεται τόσο στις απαιτήσεις της θεωρητικής μοντελοποίησης όσο και στις πρακτικές ανάγκες της υπολογιστικής επίλυσης.

Ακολουθεί η φάση της υλοποίησης της προτεινόμενης μεθόδου σε περιβάλλον Python. Στο στάδιο αυτό η θεωρητική και αλγοριθμική περιγραφή μετατρέπεται σε λειτουργικό υπολογιστικό σύστημα, το οποίο επιτρέπει την εφαρμογή της μεθόδου σε πραγματικά δεδομένα. Η υλοποίηση περιλαμβάνει τόσο την ανάπτυξη των βασικών συναρτήσεων του αλγορίθμου όσο και την οργάνωση της συνολικής ροής εκτέλεσης, από την ανάγνωση και προεπεξεργασία των δεδομένων έως την παραγωγή και αξιολόγηση των λύσεων. Η επιλογή της Python στηρίζεται στην ευχρηστία της, στη σαφήνεια του κώδικα και στις δυνατότητες που προσφέρει για επεξεργασία δεδομένων, αριθμητικούς υπολογισμούς και πειραματική ανάλυση.

Ιδιαίτερα κρίσιμο στοιχείο της μεθοδολογικής προσέγγισης αποτελεί η πειραματική αξιολόγηση της προτεινόμενης μεθόδου σε πραγματικά χρηματιστηριακά δεδομένα. Η αξιολόγηση αυτή θα πραγματοποιηθεί με χρήση δεδομένων που αντιστοιχούν σε έναν ή περισσότερους χρηματιστηριακούς δείκτες, με στόχο να εκτιμηθεί κατά πόσο ο αλγόριθμος μπορεί να παράγει χαρτοφυλάκια που ακολουθούν αποτελεσματικά τον δείκτη αναφοράς.

Στο πλαίσιο αυτό θα εξεταστούν διαφορετικές τιμές παραμέτρων, εναλλακτικοί γενετικοί τελεστές και διαφορετικά πειραματικά σενάρια, ώστε να αποτιμηθεί όχι μόνο η τελική ποιότητα των λύσεων, αλλά και η συμπεριφορά της μεθόδου υπό διαφορετικές συνθήκες.

Παράλληλα, η μεθοδολογική προσέγγιση της εργασίας περιλαμβάνει και συγκριτική αξιολόγηση της προτεινόμενης λύσης. Η σύγκριση αυτή θα γίνει, στον βαθμό που το επιτρέπουν τα διαθέσιμα δεδομένα και η σχετική βιβλιογραφία, με αποτελέσματα άλλων προσεγγίσεων που έχουν εφαρμοστεί στο ίδιο ή σε συναφές πρόβλημα.

1.10 Δομή της εργασίας

Η παρούσα διπλωματική εργασία οργανώνεται σε έξι κύρια κεφάλαια, τα οποία ακολουθούν μια λογική αλληλουχία που ξεκινά από τη θεωρητική θεμελίωση του προβλήματος και καταλήγει στη σχεδίαση, υλοποίηση και αξιολόγηση της προτεινόμενης μεθοδολογίας. Η διάρθρωση αυτή έχει επιλεγεί έτσι ώστε να εξασφαλίζεται η σαφής παρουσίαση του αντικειμένου, η ομαλή μετάβαση από τη θεωρία στην εφαρμογή και η συστηματική ανάπτυξη των επιμέρους πτυχών της εργασίας.

Στο πρώτο κεφάλαιο παρουσιάζεται το γενικό πλαίσιο της διαχείρισης επενδύσεων και της βελτιστοποίησης χαρτοφυλακίου, ενώ εισάγεται η έννοια της αποτύπωσης δείκτη ως ειδική περίπτωση του γενικότερου προβλήματος. Παράλληλα, γίνεται συνοπτική αναφορά στους Γενετικούς και Εξελικτικούς Αλγορίθμους ως κατάλληλα εργαλεία για την αντιμετώπιση σύνθετων προβλημάτων βελτιστοποίησης. Στο ίδιο κεφάλαιο προσδιορίζονται το αντικείμενο, ο σκοπός, οι στόχοι, τα ερευνητικά ερωτήματα και η μεθοδολογική προσέγγιση της εργασίας, ώστε να διαμορφωθεί από την αρχή ένα σαφές και συνεκτικό πλαίσιο αναφοράς.

Στο δεύτερο κεφάλαιο αναλύεται εκτενώς το πρόβλημα της αποτύπωσης δείκτη. Αρχικά παρουσιάζεται το σχετικό χρηματοοικονομικό υπόβαθρο, με έμφαση στη διάκριση μεταξύ ενεργητικής και παθητικής διαχείρισης χαρτοφυλακίου, στον ρόλο των χρηματιστηριακών δεικτών και στη σημασία της παρακολούθησης δεικτών αναφοράς. Στη συνέχεια εξετάζονται οι βασικές συνιστώσες του προβλήματος, όπως η επιλογή τίτλων, ο προσδιορισμός βαρών, το σφάλμα αποτύπωσης, το κόστος συναλλαγών και οι περιορισμοί πληθικότητας. Το κεφάλαιο αυτό περιλαμβάνει επίσης τη βιβλιογραφική ανασκόπηση, στην οποία παρουσιάζονται και συγκρίνονται σχετικές επιστημονικές εργασίες και διαφορετικές κατηγορίες μεθόδων επίλυσης, με ιδιαίτερη έμφαση σε ευρετικές, μεταευρετικές και εξελικτικές προσεγγίσεις.

Στο τρίτο κεφάλαιο παρουσιάζεται η προτεινόμενη μέθοδος επίλυσης του προβλήματος. Αρχικά διατυπώνεται το πρόβλημα της αποτύπωσης δείκτη σε αυστηρή μαθηματική μορφή, ως πρόβλημα βελτιστοποίησης, με σαφή προσδιορισμό της αντικειμενικής συνάρτησης,

των μεταβλητών απόφασης, των παραμέτρων και των περιορισμών. Στη συνέχεια αναλύεται ο σχεδιασμός του προτεινόμενου Γενετικού Αλγορίθμου, συμπεριλαμβανομένης της αναπαράστασης των χρωμοσωμάτων, της συνάρτησης καταλληλότητας, των γενετικών τελεστών και των μηχανισμών διαχείρισης περιορισμών. Το κεφάλαιο ολοκληρώνεται με λεκτική περιγραφή της μεθόδου, ψευδοκώδικα και διάγραμμα ροής, ώστε να αποδίδεται με σαφήνεια η συνολική λειτουργία του αλγορίθμου.

Το τέταρτο κεφάλαιο αφορά την υλοποίηση της προτεινόμενης μεθόδου. Σε αυτό παρουσιάζεται το τεχνολογικό περιβάλλον ανάπτυξης, η γενική αρχιτεκτονική του συστήματος και ο τρόπος με τον οποίο μετατρέπεται ο θεωρητικός σχεδιασμός σε λειτουργική υπολογιστική εφαρμογή σε Python. Περιγράφονται η προεπεξεργασία των δεδομένων, η υλοποίηση της αναπαράστασης των λύσεων, η υλοποίηση της συνάρτησης καταλληλότητας και των γενετικών τελεστών, καθώς και οι βασικές συναρτήσεις και μέθοδοι του αλγορίθμου. Παράλληλα, εξετάζονται ζητήματα αποδοτικότητας, αξιοπιστίας και οργάνωσης του πειραματικού πλαισίου, ενώ παρουσιάζονται ενδεικτικά αποσπάσματα κώδικα, με τον πλήρη κώδικα να παρατίθεται στο παράρτημα.

Στο πέμπτο κεφάλαιο παρουσιάζονται τα αποτελέσματα της εργασίας. Αρχικά γίνεται αναλυτική περιγραφή των δεδομένων που χρησιμοποιούνται, της προέλευσής τους και των βασικών χαρακτηριστικών τους, καθώς και του πειραματικού σχεδιασμού που ακολουθείται. Στη συνέχεια παρουσιάζονται οι βασικές τιμές των παραμέτρων του αλγορίθμου, τα διαφορετικά πειραματικά σενάρια και τα κριτήρια αξιολόγησης που χρησιμοποιούνται για την αποτίμηση της απόδοσής του. Ακολουθεί η παρουσίαση και ανάλυση των αποτελεσμάτων του προτεινόμενου αλγορίθμου, η διερεύνηση της επίδρασης των παραμέτρων και των τελεστών, καθώς και η σύγκριση με άλλες τεχνικές επίλυσης, ώστε να εξαχθούν τεκμηριωμένα συμπεράσματα σχετικά με την αποτελεσματικότητα της προτεινόμενης προσέγγισης.

Το έκτο και τελευταίο κεφάλαιο περιλαμβάνει τα συμπεράσματα και τη συζήτηση της εργασίας. Σε αυτό γίνεται συνοπτική ανακεφαλαίωση του αντικειμένου, της μεθοδολογίας και των βασικών ευρημάτων, ενώ αξιολογείται η προτεινόμενη μέθοδος ως προς τα πλεονεκτήματα, τις αδυναμίες και τους περιορισμούς της. Παράλληλα, συζητείται η θέση της σε σχέση με άλλες προσεγγίσεις που έχουν προταθεί στη σχετική βιβλιογραφία και διατυπώνονται προτάσεις για μελλοντική έρευνα και πιθανές επεκτάσεις της μεθοδολογίας.

Τέλος, στο τέλος της εργασίας παρατίθεται η βιβλιογραφία, όπου συγκεντρώνονται όλες οι επιστημονικές αναφορές που χρησιμοποιήθηκαν κατά τη συγγραφή, καθώς και τα παραρτήματα, στα οποία περιλαμβάνονται συμπληρωματικά στοιχεία, όπως η πλήρης μαθηματική διατύπωση, ο ψευδοκώδικας, τα διαγράμματα ροής, ο πλήρης κώδικας της υλοποίησης και πρόσθετοι πίνακες ή γραφήματα. Με τη διάρθρωση αυτή επιδιώκεται η συνολική εργασία να παρουσιάζει ενότητα, σαφήνεια και επιστημονική πληρότητα, καλύπτοντας τόσο τη θεωρητική όσο και την εφαρμοσμένη διάσταση του προβλήματος.

ΚΕΦΑΛΑΙΟ 2: ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΑΠΟΤΥΠΩΣΗΣ ΔΕΙΚΤΗ

2.1 Εισαγωγή

Το πρόβλημα της αποτύπωσης δείκτη αποτελεί μία από τις σημαντικότερες εφαρμογές της παθητικής διαχείρισης χαρτοφυλακίου και έχει αποκτήσει ιδιαίτερη σημασία τόσο στη θεωρία όσο και στην πρακτική της σύγχρονης χρηματοοικονομικής (Beasley et al., 2003; Maringer & Oyewumi, 2007). Η βασική του ιδέα είναι σχετικά απλή: ζητείται η κατασκευή ενός χαρτοφυλακίου το οποίο να αναπαράγει όσο το δυνατόν πιστότερα την πορεία ενός χρηματιστηριακού δείκτη αναφοράς (Beasley et al., 2003). Παρά τη φαινομενική απλότητά του, το πρόβλημα αυτό παρουσιάζει σημαντικές θεωρητικές και υπολογιστικές δυσκολίες, ιδίως όταν επιχειρείται η αντιμετώπισή του υπό ρεαλιστικές συνθήκες αγοράς (Derigs & Nickel, 2003; Maringer & Oyewumi, 2007).

Οι χρηματιστηριακοί δείκτες αποτελούν συνοπτικές εκφράσεις της συμπεριφοράς ενός συνόλου τίτλων και χρησιμοποιούνται ευρέως ως σημεία αναφοράς για την αποτίμηση της απόδοσης επενδυτικών χαρτοφυλακίων (Meade & Salkin, 1989; Beasley et al., 2003). Η δυνατότητα κατασκευής χαρτοφυλακίων που να ακολουθούν με ακρίβεια έναν δείκτη συνδέεται άμεσα με τη λειτουργία των index funds και με την ευρύτερη λογική της παθητικής επένδυσης (Meade & Salkin, 1989; Maringer & Oyewumi, 2007). Στο πλαίσιο αυτό, η αποτύπωση δείκτη δεν αφορά μόνο ένα θεωρητικό πρόβλημα βελτιστοποίησης, αλλά και μια πρακτική επενδυτική στρατηγική με ευρεία εφαρμογή στις διεθνείς χρηματοοικονομικές αγορές (Beasley et al., 2003; Derigs & Nickel, 2003).

Σε θεωρητικό επίπεδο, η πιο άμεση προσέγγιση για την παρακολούθηση ενός δείκτη είναι η πλήρης αναπαραγωγή της σύνθεσής του, δηλαδή η κατοχή όλων των τίτλων που συμμετέχουν στον δείκτη με αντίστοιχα βάρη (Beasley et al., 2003). Ωστόσο, στην πράξη, μια τέτοια προσέγγιση δεν είναι πάντοτε η πιο αποδοτική ή εφικτή (Maringer & Oyewumi, 2007; Derigs & Nickel, 2003). Η ύπαρξη μεγάλου αριθμού τίτλων, το κόστος συναλλαγών, οι ανάγκες συνεχούς αναπροσαρμογής, η περιορισμένη ρευστότητα ορισμένων μετοχών και η επιθυμία διατήρησης πιο «συμπαγών» χαρτοφυλακίων καθιστούν συχνά προτιμότερη την επιλογή ενός υποσυνόλου τίτλων που να προσεγγίζει επαρκώς τη συμπεριφορά του δείκτη (Beasley et al., 2003; Derigs & Nickel, 2003; Maringer & Oyewumi, 2007). Στο σημείο αυτό ακριβώς αναδεικνύεται ο πυρήνας του προβλήματος της αποτύπωσης δείκτη ως προβλήματος βελτιστοποίησης (Beasley et al., 2003).

Η δυσκολία του προβλήματος έγκειται στο ότι πρέπει να ληφθούν ταυτόχρονα δύο ειδών αποφάσεις. Από τη μία πλευρά, απαιτείται η επιλογή των τίτλων που θα περιληφθούν στο χαρτοφυλάκιο και, από την άλλη, ο καθορισμός των αντίστοιχων βαρών συμμετοχής τους (Beasley et al., 2003; Maringer & Oyewumi, 2007). Οι αποφάσεις αυτές πρέπει να οδηγούν σε μικρή απόκλιση από τον δείκτη αναφοράς, ενώ συγχρόνως να ικανοποιούν μια σειρά από περιορισμούς, όπως ο διαθέσιμος προϋπολογισμός, τα κατώτατα και ανώτατα όρια βαρών, ο μέγιστος αριθμός τίτλων και το κόστος συναλλαγών (Beasley et al., 2003; Canakgoz & Beasley, 2009). Επομένως, το πρόβλημα συνδυάζει διακριτές και συνεχείς μεταβλητές, γεγονός που ενισχύει τον υπολογιστικό του χαρακτήρα (Derigs & Nickel, 2003; Canakgoz & Beasley, 2009).

Η σημασία του προβλήματος της αποτύπωσης δείκτη δεν περιορίζεται στη διαχείριση παθητικών χαρτοφυλακίων, αλλά επεκτείνεται και στη γενικότερη μελέτη σύνθετων προβλημάτων χρηματοοικονομικής βελτιστοποίησης (Beasley et al., 2003; Derigs & Nickel, 2003). Πράγματι, το index tracking αποτελεί χαρακτηριστικό παράδειγμα προβλήματος όπου η θεωρητική οικονομική λογική συναντά τη μαθηματική μοντελοποίηση και την υπολογιστική επίλυση (Maringer & Oyewumi, 2007; Derigs & Nickel, 2003). Για τον λόγο αυτό, στη σχετική βιβλιογραφία έχουν προταθεί διαφορετικές μέθοδοι αντιμετώπισής του, που εκτείνονται από κλασικές τεχνικές μαθηματικού προγραμματισμού έως ευρετικές, μεταευρετικές και εξελικτικές προσεγγίσεις (Beasley et al., 2003; Derigs & Nickel, 2003; Canakgoz & Beasley, 2009).

2.2 Χρηματοοικονομικό υπόβαθρο του προβλήματος

2.2.1 Ενεργητική και παθητική διαχείριση χαρτοφυλακίου

Η διαχείριση χαρτοφυλακίου διακρίνεται συνήθως σε δύο βασικές κατηγορίες: την ενεργητική και την παθητική διαχείριση. Η διάκριση αυτή είναι ιδιαίτερα σημαντική για την κατανόηση του προβλήματος της αποτύπωσης δείκτη, καθώς το τελευταίο εντάσσεται άμεσα στη φιλοσοφία της παθητικής επενδυτικής στρατηγικής (Sharpe, 1991; Beasley et al., 2003). Οι δύο αυτές προσεγγίσεις διαφέρουν ουσιαστικά ως προς τους στόχους, τις μεθόδους λήψης αποφάσεων και τον τρόπο με τον οποίο αντιλαμβάνονται τη σχέση του χαρτοφυλακίου με την αγορά (Sharpe, 1991).

Η ενεργητική διαχείριση χαρτοφυλακίου στηρίζεται στην επιδίωξη υπεραπόδοσης σε σχέση με την αγορά ή με έναν δείκτη αναφοράς (Sharpe, 1991; Goodwin, 1998). Στην περίπτωση αυτή, ο διαχειριστής επιλέγει τίτλους με βάση την εκτίμηση ότι ορισμένες επενδυτικές ευκαιρίες είναι δυνατόν να αποφέρουν υψηλότερες αποδόσεις από εκείνες που αντανάκλα η συνολική αγορά (Sharpe, 1991). Η διαδικασία αυτή περιλαμβάνει συνήθως συνεχή ανάλυση των οικονομικών δεδομένων, αξιολόγηση των επιχειρήσεων, εκτίμηση μακροοικονομικών συνθηκών, πρόβλεψη μεταβολών στις τιμές και συχνές αναπροσαρμογές του χαρτοφυλακίου. Επομένως, η ενεργητική διαχείριση χαρακτηρίζεται από υψηλό βαθμό παρέμβασης του διαχειριστή και από συνεχή προσπάθεια εκμετάλλευσης αποκλίσεων της αγοράς ή λανθασμένης αποτίμησης τίτλων (Sharpe, 1991).

Βασική παραδοχή της ενεργητικής προσέγγισης είναι ότι ο διαχειριστής μπορεί, μέσω ανάλυσης, εμπειρίας ή ειδικών πληροφοριών, να επιλέξει ένα χαρτοφυλάκιο που θα αποδώσει καλύτερα από την αγορά. Ωστόσο, η στρατηγική αυτή συνοδεύεται από αυξημένο κόστος διαχείρισης και συναλλαγών, καθώς η συχνή αναθεώρηση των επενδυτικών επιλογών προϋποθέτει συχνές αγοραπωλησίες τίτλων (Sharpe, 1991). Επιπλέον, η επιτυχία της ενεργητικής διαχείρισης δεν είναι διασφαλισμένη, αφού εξαρτάται από την ικανότητα του διαχειριστή να εντοπίζει και να αξιοποιεί αποτελεσματικά τις ευκαιρίες της αγοράς (Sharpe, 1991; Pastor et al., 2015). Για τον λόγο αυτό, η ενεργητική διαχείριση συνδέεται με υψηλότερο βαθμό αβεβαιότητας, αλλά και με τη δυνατότητα επίτευξης ανώτερων αποδόσεων σε περίπτωση επιτυχούς στρατηγικής (Sharpe, 1991).

Αντίθετα, η παθητική διαχείριση χαρτοφυλακίου δεν επιδιώκει την υπεραπόδοση έναντι της αγοράς, αλλά τη συστηματική παρακολούθηση της συμπεριφοράς ενός δείκτη ή μιας συγκεκριμένης αγοράς (Sharpe, 1991; Meade & Salkin, 1989). Στην προσέγγιση αυτή, ο διαχειριστής δεν προσπαθεί να «νικήσει» την αγορά, αλλά να την αναπαράγει όσο το

δυνατόν πιστότερα (Meade & Salkin, 1989; Beasley et al., 2003). Η λογική αυτή βασίζεται στην άποψη ότι, σε πολλές περιπτώσεις, η επίτευξη σταθερής υπεραπόδοσης μέσω ενεργητικής διαχείρισης είναι εξαιρετικά δύσκολη, ιδιαίτερα μετά την αφαίρεση των εξόδων διαχείρισης και των συναλλακτικών δαπανών (Sharpe, 1991). Ως εκ τούτου, μπορεί να είναι προτιμότερο να κατασκευαστεί ένα χαρτοφυλάκιο που να ακολουθεί στενά έναν δείκτη αναφοράς, εξασφαλίζοντας χαμηλότερο κόστος και μεγαλύτερη διαφάνεια στη στρατηγική (Meade & Salkin, 1989; Beasley et al., 2003).

Η παθητική διαχείριση συνδέεται στενά με τα χαρτοφυλάκια αποτύπωσης δείκτη και με τα επενδυτικά κεφάλαια που έχουν ως στόχο την παρακολούθηση συγκεκριμένων χρηματιστηριακών δεικτών (Meade & Salkin, 1989; Beasley et al., 2003). Στην πράξη, η στρατηγική αυτή μπορεί να υλοποιηθεί είτε μέσω πλήρους αναπαραγωγής του δείκτη, είτε μέσω μερικής αναπαραγωγής, όπου επιλέγεται ένα αντιπροσωπευτικό υποσύνολο τίτλων (Beasley et al., 2003; Maringer & Oyewumi, 2007). Και στις δύο περιπτώσεις, το ζητούμενο είναι η ελαχιστοποίηση της απόκλισης μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη (Derigs & Nickel, 2003; Beasley et al., 2003). Για τον λόγο αυτό, η παθητική διαχείριση δεν συνεπάγεται απουσία ανάλυσης ή βελτιστοποίησης· αντίθετα, απαιτεί προσεκτικό σχεδιασμό, ιδίως όταν υπάρχουν περιορισμοί στον αριθμό των τίτλων, στο διαθέσιμο κεφάλαιο και στο κόστος συναλλαγών (Derigs & Nickel, 2003; Canakgoz & Beasley, 2009).

Μια βασική διαφορά μεταξύ ενεργητικής και παθητικής διαχείρισης αφορά συνεπώς το επενδυτικό σημείο αναφοράς. Στην ενεργητική διαχείριση, ο δείκτης χρησιμοποιείται ως benchmark έναντι του οποίου αξιολογείται η επίδοση του διαχειριστή (Goodwin, 1998; Sharpe, 1991). Στην παθητική διαχείριση, αντιθέτως, ο ίδιος ο δείκτης αποτελεί τον στόχο που πρέπει να προσεγγιστεί (Meade & Salkin, 1989; Beasley et al., 2003). Η διαφορά αυτή έχει ουσιαστικές συνέπειες τόσο στον τρόπο κατασκευής του χαρτοφυλακίου όσο και στα κριτήρια αξιολόγησης της επιτυχίας της επενδυτικής στρατηγικής. Στην πρώτη περίπτωση, η επιτυχία μετράται με όρους υπεραπόδοσης· στη δεύτερη, με όρους ακρίβειας παρακολούθησης και περιορισμού του σφάλματος αποτύπωσης (Goodwin, 1998; Derigs & Nickel, 2003).

Παρά τη σαφή εννοιολογική διάκριση, οι δύο στρατηγικές δεν είναι πάντοτε πλήρως απομονωμένες στην πράξη. Υπάρχουν περιπτώσεις μικτών στρατηγικών, στις οποίες ένα μέρος του χαρτοφυλακίου επενδύεται παθητικά και ένα άλλο μέρος διαχειρίζεται ενεργητικά (Pástor et al., 2015). Ωστόσο, για τους σκοπούς της παρούσας εργασίας, το ενδιαφέρον εστιάζεται στην παθητική διαχείριση και ειδικότερα στην κατασκευή χαρτοφυλακίων που παρακολουθούν χρηματιστηριακούς δείκτες (Beasley et al., 2003; Maringer & Oyewumi, 2007). Η κατανόηση της διάκρισης αυτής είναι κρίσιμη, διότι επιτρέπει να τοποθετηθεί σωστά το πρόβλημα της αποτύπωσης δείκτη μέσα στο ευρύτερο πεδίο της διαχείρισης επενδύσεων (Sharpe, 1991; Beasley et al., 2003).

2.2.2 Χρηματιστηριακοί δείκτες και ο ρόλος τους στις επενδυτικές αποφάσεις

Οι χρηματιστηριακοί δείκτες αποτελούν ένα από τα σημαντικότερα εργαλεία παρακολούθησης, ανάλυσης και αξιολόγησης της πορείας των κεφαλαιαγορών. Πρόκειται για σύνθετα μεγέθη που αποτυπώνουν τη συνολική συμπεριφορά ενός συνόλου τίτλων, οι οποίοι επιλέγονται με βάση συγκεκριμένα κριτήρια, όπως η κεφαλαιοποίηση, η ρευστότητα, ο κλάδος δραστηριότητας ή η γεωγραφική αγορά στην οποία

διαπραγματεύονται. Μέσω αυτής της συνθετικής αποτύπωσης, οι δείκτες παρέχουν μια συνοπτική αλλά ιδιαίτερα χρήσιμη εικόνα της πορείας μιας αγοράς ή ενός επιμέρους τμήματός της (Meade & Salkin, 1989; Beasley et al., 2003).

Η σημασία των χρηματιστηριακών δεικτών είναι πολυδιάστατη. Σε πρώτο επίπεδο, λειτουργούν ως μέτρα αναφοράς της γενικής κατεύθυνσης της αγοράς. Όταν, για παράδειγμα, ένας ευρύς δείκτης παρουσιάζει άνοδο ή πτώση, η μεταβολή αυτή ερμηνεύεται συνήθως ως ένδειξη της συνολικής τάσης που επικρατεί στην αντίστοιχη αγορά. Με αυτόν τον τρόπο, οι δείκτες επιτρέπουν στους επενδυτές, στους αναλυτές και στους διαχειριστές χαρτοφυλακίων να παρακολουθούν εύκολα τις εξελίξεις στις αγορές χωρίς να χρειάζεται να εξετάζουν ξεχωριστά την πορεία κάθε επιμέρους τίτλου (Meade & Salkin, 1989).

Σε δεύτερο επίπεδο, οι χρηματιστηριακοί δείκτες χρησιμοποιούνται ευρέως ως δείκτες αναφοράς για την αξιολόγηση της απόδοσης επενδυτικών χαρτοφυλακίων. Ένα χαρτοφυλάκιο δεν εξετάζεται συνήθως μόνο με απόλυτους όρους απόδοσης, αλλά και σε σχέση με κάποιο κατάλληλο benchmark, δηλαδή έναν δείκτη που θεωρείται αντιπροσωπευτικός της αγοράς ή της κατηγορίας επένδυσης στην οποία ανήκει το χαρτοφυλάκιο. Έτσι, οι δείκτες αποκτούν κεντρικό ρόλο στη διαδικασία αξιολόγησης, καθώς επιτρέπουν να εκτιμηθεί αν ένα χαρτοφυλάκιο υπεραπέδωσε, υποαπέδωσε ή απλώς ακολούθησε τη γενική πορεία της αγοράς (Sharpe, 1991; Goodwin, 1998).

Η χρήση των δεικτών ως benchmarks είναι ιδιαίτερα σημαντική τόσο στην ενεργητική όσο και στην παθητική διαχείριση. Στην ενεργητική διαχείριση, ο δείκτης χρησιμοποιείται ως σημείο σύγκρισης, έναντι του οποίου αξιολογείται η ικανότητα του διαχειριστή να επιτύχει υψηλότερη απόδοση. Στην παθητική διαχείριση, αντιθέτως, ο δείκτης δεν αποτελεί απλώς μέτρο αναφοράς, αλλά τον ίδιο τον στόχο της επενδυτικής στρατηγικής. Σε αυτή την περίπτωση, το ζητούμενο είναι η όσο το δυνατόν πιστότερη αναπαραγωγή της συμπεριφοράς του δείκτη, γεγονός που καθιστά τους χρηματιστηριακούς δείκτες κεντρικό στοιχείο στη διαμόρφωση και αξιολόγηση χαρτοφυλακίων αποτύπωσης δείκτη (Sharpe, 1991; Beasley et al., 2003; Derigs & Nickel, 2003).

Ένας ακόμη σημαντικός ρόλος των χρηματιστηριακών δεικτών συνδέεται με τη δυνατότητα που προσφέρουν για κατηγοριοποίηση και τμηματοποίηση της αγοράς. Υπάρχουν δείκτες που αναφέρονται στο σύνολο μιας εθνικής αγοράς, δείκτες που αποτυπώνουν συγκεκριμένους κλάδους, δείκτες που εστιάζουν σε μετοχές μεγάλης ή μικρής κεφαλαιοποίησης, καθώς και δείκτες που αφορούν διεθνείς ή περιφερειακές αγορές. Η ύπαρξη αυτής της ποικιλίας επιτρέπει στους επενδυτές να επιλέγουν το κατάλληλο σημείο αναφοράς ανάλογα με τη στρατηγική, το επίπεδο κινδύνου και το πεδίο επένδυσης που τους ενδιαφέρει. Έτσι, οι δείκτες δεν αποτυπώνουν απλώς την αγορά συνολικά, αλλά και επιμέρους επενδυτικά σύμπαντα με διαφορετικά χαρακτηριστικά (Meade & Salkin, 1989; Walkshäusl & Lobe, 2010).

Η μεθοδολογία κατασκευής ενός δείκτη επηρεάζει σημαντικά τη χρησιμότητά του στις επενδυτικές αποφάσεις. Οι δείκτες μπορεί να βασίζονται σε διαφορετικούς μηχανισμούς στάθμισης, όπως στάθμιση με βάση την κεφαλαιοποίηση, στάθμιση με βάση την τιμή ή ίση στάθμιση μεταξύ των συμμετεχόντων τίτλων. Ο τρόπος αυτός επηρεάζει τη σχετική βαρύτητα κάθε τίτλου στον δείκτη και, κατά συνέπεια, τον τρόπο με τον οποίο οι μεταβολές συγκεκριμένων μετοχών επηρεάζουν τη συνολική του πορεία. Για τον επενδυτή ή τον διαχειριστή χαρτοφυλακίου, η κατανόηση της μεθοδολογίας κατασκευής ενός δείκτη είναι απαραίτητη, καθώς επηρεάζει τον τρόπο με τον οποίο πρέπει να σχεδιαστεί ένα

χαρτοφυλάκιο που θα τον ακολουθεί αποτελεσματικά (Arnott et al., 2005; Walkshäusl & Lobe, 2010; Perold, 2007).

Οι δείκτες έχουν επίσης καθοριστική σημασία για τη δημιουργία επενδυτικών προϊόντων και στρατηγικών. Πολλά σύγχρονα επενδυτικά προϊόντα, όπως αμοιβαία κεφάλαια παθητικής διαχείρισης και διαπραγματεύσιμα αμοιβαία κεφάλαια, βασίζονται ακριβώς στη λογική παρακολούθησης ενός δείκτη. Σε αυτές τις περιπτώσεις, ο δείκτης λειτουργεί ως το θεμέλιο πάνω στο οποίο χτίζεται ολόκληρη η επενδυτική πολιτική του προϊόντος. Επομένως, η κατανόηση του ρόλου των δεικτών δεν είναι μόνο θεωρητικής σημασίας, αλλά συνδέεται άμεσα με τον τρόπο με τον οποίο οργανώνονται στην πράξη σημαντικά τμήματα της επενδυτικής βιομηχανίας (Meade & Salkin, 1989; Brogaard et al., 2026).

Στο πλαίσιο του προβλήματος της αποτύπωσης δείκτη, οι χρηματιστηριακοί δείκτες αποκτούν ακόμη πιο εξειδικευμένο ρόλο, καθώς αποτελούν το μέγεθος-στόχο που επιχειρεί να προσεγγίσει το υπό κατασκευή χαρτοφυλάκιο. Αυτό σημαίνει ότι η επιλογή του κατάλληλου δείκτη αναφοράς, η κατανόηση της σύνθεσής του, των αναθεωρήσεών του και της μεθοδολογίας υπολογισμού του είναι κρίσιμα στοιχεία για τη διατύπωση και επίλυση του προβλήματος. Δεν αρκεί δηλαδή να θεωρηθεί ο δείκτης ως μια αφηρημένη αναφορά της αγοράς· απαιτείται να αντιμετωπιστεί ως συγκεκριμένο χρηματοοικονομικό αντικείμενο, με σαφώς καθορισμένα χαρακτηριστικά που επηρεάζουν άμεσα τη συμπεριφορά του χαρτοφυλακίου αποτύπωσης (Beasley et al., 2003; Derigs & Nickel, 2003; Canakgoz & Beasley, 2009).

2.2.3 Χαρτοφυλάκια και δείκτες αναφοράς

Η έννοια του δείκτη αναφοράς κατέχει κεντρική θέση στη θεωρία και στην πρακτική της διαχείρισης χαρτοφυλακίου, καθώς παρέχει το σημείο σύγκρισης με βάση το οποίο αξιολογείται η επενδυτική επίδοση (Sharpe, 1991; Daniel et al., 1997). Ένα χαρτοφυλάκιο δεν ξεετάζεται συνήθως απομονωμένα, αλλά σε σχέση με κάποιο πρότυπο μέγεθος που αποτυπώνει την αγορά ή το επενδυτικό σύμπαν μέσα στο οποίο δραστηριοποιείται. Το πρότυπο αυτό είναι ο δείκτης αναφοράς και η επιλογή του επηρεάζει ουσιαστικά τόσο τη διαδικασία διαχείρισης όσο και την ερμηνεία των αποτελεσμάτων (Daniel et al., 1997; Goodwin, 1998).

Ένας δείκτης αναφοράς μπορεί να οριστεί ως το εξωτερικό μέτρο έναντι του οποίου συγκρίνεται η απόδοση, ο κίνδυνος και γενικότερα η συμπεριφορά ενός χαρτοφυλακίου (Daniel et al., 1997; Chen & Knez, 1996). Στην πράξη, ο δείκτης αυτός επιλέγεται έτσι ώστε να είναι αντιπροσωπευτικός της αγοράς, του κλάδου ή της κατηγορίας επενδύσεων στην οποία ανήκει το χαρτοφυλάκιο (Daniel et al., 1997). Για παράδειγμα, ένα χαρτοφυλάκιο που επενδύει σε μετοχές μεγάλης κεφαλαιοποίησης μιας συγκεκριμένης αγοράς είναι εύλογο να αξιολογείται σε σχέση με έναν δείκτη που αποτυπώνει ακριβώς αυτή την κατηγορία τίτλων. Η αντιπροσωπευτικότητα του δείκτη αναφοράς είναι ιδιαίτερα σημαντική, διότι μόνο τότε η σύγκριση με το χαρτοφυλάκιο αποκτά ουσιαστικό επενδυτικό περιεχόμενο (Daniel et al., 1997).

Η χρήση δείκτη αναφοράς συνδέεται πρωτίστως με την ανάγκη αξιολόγησης της επενδυτικής απόδοσης. Σε ένα περιβάλλον όπου οι αποδόσεις των αγορών μεταβάλλονται διαρκώς, η αποτίμηση ενός χαρτοφυλακίου μόνο με βάση την απόλυτη απόδοσή του δεν είναι πάντοτε επαρκής. Ένα χαρτοφυλάκιο μπορεί να έχει θετική απόδοση, αλλά να υπολείπεται σημαντικά της αγοράς, ή αντίστροφα να έχει αρνητική απόδοση σε περίοδο

γενικής πτώσης, αλλά να εμφανίζει σχετική υπεροχή έναντι του δείκτη αναφοράς. Επομένως, ο δείκτης αναφοράς επιτρέπει τη σχετική αξιολόγηση της επίδοσης και καθιστά πιο ακριβή την αποτίμηση της αποτελεσματικότητας της επενδυτικής στρατηγικής (Goodwin, 1998; Daniel et al., 1997; Chen & Knez, 1996).

Η σημασία του δείκτη αναφοράς διαφοροποιείται ανάλογα με το είδος της διαχείρισης που εφαρμόζεται. Στην ενεργητική διαχείριση, ο δείκτης χρησιμοποιείται ως μέτρο σύγκρισης, δηλαδή ως το επίπεδο απόδοσης που επιδιώκει να υπερβεί ο διαχειριστής μέσω των επενδυτικών του επιλογών (Sharpe, 1991; Goodwin, 1998). Η αξιολόγηση επικεντρώνεται στην ικανότητα του χαρτοφυλακίου να παρουσιάζει καλύτερη απόδοση ή ευνοϊκότερη σχέση απόδοσης και κινδύνου σε σχέση με τον επιλεγμένο benchmark (Sharpe, 1991; Goodwin, 1998). Αντίθετα, στην παθητική διαχείριση, ο δείκτης αναφοράς δεν αποτελεί απλώς σημείο σύγκρισης αλλά τον ίδιο τον στόχο της επενδυτικής στρατηγικής. Στην περίπτωση αυτή, το χαρτοφυλάκιο σχεδιάζεται με σκοπό να ακολουθεί όσο το δυνατόν πιστότερα την πορεία του δείκτη (Meade & Salkin, 1989; Beasley et al., 2003).

Η σχέση μεταξύ χαρτοφυλακίου και δείκτη αναφοράς αποκτά ιδιαίτερο ενδιαφέρον στο πρόβλημα της αποτύπωσης δείκτη, όπου ο βασικός στόχος είναι η ελαχιστοποίηση της απόκλισης μεταξύ των δύο (Beasley et al., 2003; Derigs & Nickel, 2003). Εδώ, το χαρτοφυλάκιο δεν κρίνεται ως προς το αν υπεραποδίδει ή υποαποδίδει στρατηγικά, αλλά ως προς το πόσο ακριβώς αναπαράγει τη συμπεριφορά του benchmark (Derigs & Nickel, 2003). Το στοιχείο αυτό διαφοροποιεί ουσιαστικά την έννοια του benchmark στην παθητική διαχείριση από τη χρήση του στην ενεργητική, καθώς μετατοπίζει το ενδιαφέρον από την επίδοση σε όρους υπεραπόδοσης στην επίδοση σε όρους πιστότητας παρακολούθησης (Sharpe, 1991; Derigs & Nickel, 2003).

Η καταλληλότητα ενός δείκτη αναφοράς εξαρτάται από διάφορους παράγοντες. Ένας κατάλληλος benchmark πρέπει να είναι σαφώς προσδιορισμένος, παρατηρήσιμος, αντιπροσωπευτικός του αντίστοιχου επενδυτικού πεδίου και επαρκώς σταθερός ως προς τη μεθοδολογία του (Daniel et al., 1997; Chen & Knez, 1996). Παράλληλα, πρέπει να αντανακλά τις επενδυτικές δυνατότητες που είναι διαθέσιμες στον διαχειριστή και να επιτρέπει δίκαιη σύγκριση με το χαρτοφυλάκιο (Daniel et al., 1997). Αν ο δείκτης αναφοράς δεν είναι συμβατός με το επενδυτικό σύμπαν ή με τους περιορισμούς του χαρτοφυλακίου, τότε η σύγκριση μπορεί να οδηγήσει σε παραπλανητικά συμπεράσματα σχετικά με την επίδοση (Daniel et al., 1997; Goodwin, 1998).

Στο πλαίσιο της αποτύπωσης δείκτη, το benchmark αποκτά και έναν ακόμη ρόλο: λειτουργεί ως πηγή πληροφοριών για τη σύνθεση του χαρτοφυλακίου. Η σύνθεση του δείκτη, τα βάρη των επιμέρους τίτλων, η συχνότητα αναθεώρησής του και η μεθοδολογία υπολογισμού του αποτελούν κρίσιμα δεδομένα για την κατασκευή του χαρτοφυλακίου αποτύπωσης (Beasley et al., 2003; Canakgoz & Beasley, 2009). Με άλλα λόγια, ο δείκτης αναφοράς δεν χρησιμοποιείται μόνο για τελική αξιολόγηση, αλλά και ως βασικό στοιχείο εισόδου στη διαδικασία βελτιστοποίησης (Beasley et al., 2003; Derigs & Nickel, 2003). Αυτό είναι ιδιαίτερα εμφανές όταν η προσέγγιση που ακολουθείται δεν είναι πλήρης αναπαραγωγή, αλλά επιλογή υποσυνόλου τίτλων που πρέπει να αποδώσει με τρόπο όσο το δυνατόν πιο κοντινό στον δείκτη (Beasley et al., 2003; Canakgoz & Beasley, 2009).

Παράλληλα, η έννοια του χαρτοφυλακίου αναφοράς μπορεί να χρησιμοποιηθεί και με ευρύτερη έννοια, ως η υποθετική ή πραγματική σύνθεση επενδύσεων που αντιστοιχεί στον επιλεγμένο benchmark (Daniel et al., 1997; Chen & Knez, 1996). Στην περίπτωση της

πλήρους αναπαραγωγής, το χαρτοφυλάκιο αναφοράς ταυτίζεται πρακτικά με τη σύνθεση του δείκτη. Στην περίπτωση της μερικής αναπαραγωγής, το χαρτοφυλάκιο αναφοράς εξακολουθεί να λειτουργεί ως το πρότυπο που επιχειρεί να προσεγγίσει το υπό κατασκευή χαρτοφυλάκιο, χωρίς όμως να απαιτείται πλήρης ταύτιση της σύνθεσης. Αυτή η διάκριση είναι κρίσιμη, διότι δείχνει ότι στο πρόβλημα του index tracking η επιτυχία δεν εξαρτάται κατ' ανάγκη από την ακριβή αντιγραφή όλων των τίτλων, αλλά από την αποτελεσματική προσέγγιση της συνολικής συμπεριφοράς του δείκτη (Beasley et al., 2003; Maringer & Oyewumi, 2007).

2.3 Το πρόβλημα της πλήρους και της μερικής αναπαραγωγής δείκτη

2.3.1 Πλήρης αναπαραγωγή δείκτη

Η πλήρης αναπαραγωγή δείκτη αποτελεί την πιο άμεση και θεωρητικά ακριβή προσέγγιση για την κατασκευή ενός χαρτοφυλακίου αποτύπωσης δείκτη (Jiang et al., 2021; Beasley et al., 2003). Στην προσέγγιση αυτή, το χαρτοφυλάκιο περιλαμβάνει όλους τους τίτλους που συμμετέχουν στον δείκτη αναφοράς και κάθε τίτλος κατέχει βάρος ανάλογο προς εκείνο που έχει στον ίδιο τον δείκτη (Jiang et al., 2021; Zhang et al., 2024). Με τον τρόπο αυτό, η σύνθεση του χαρτοφυλακίου ταυτίζεται, τουλάχιστον κατά το δυνατόν, με τη σύνθεση του δείκτη, γεγονός που επιτρέπει την πολύ στενή παρακολούθηση της πορείας του (Jiang et al., 2021; Beasley et al., 2003).

Η λογική της πλήρους αναπαραγωγής είναι ιδιαίτερα απλή από εννοιολογική άποψη. Εφόσον ο δείκτης αποτελείται από ένα συγκεκριμένο σύνολο τίτλων με γνωστά βάρη, η πιο φυσική στρατηγική για την παρακολούθησή του είναι η δημιουργία ενός χαρτοφυλακίου που να αναπαράγει ακριβώς αυτή τη σύνθεση (Jiang et al., 2021). Αν η διαδικασία αυτή μπορούσε να εφαρμοστεί χωρίς κανέναν περιορισμό και χωρίς κόστος, τότε το χαρτοφυλάκιο θα ακολουθούσε σχεδόν απόλυτα τον δείκτη, με αποτέλεσμα το σφάλμα αποτύπωσης να είναι πρακτικά μηδενικό ή εξαιρετικά μικρό (Beasley et al., 2003; Derigs & Nickel, 2003). Για τον λόγο αυτό, η πλήρης αναπαραγωγή θεωρείται συχνά το θεωρητικό σημείο αναφοράς στο πρόβλημα της αποτύπωσης δείκτη (Jiang et al., 2021; Beasley et al., 2003).

Ένα από τα βασικά πλεονεκτήματα της πλήρους αναπαραγωγής είναι ακριβώς η υψηλή ακρίβεια παρακολούθησης που μπορεί να προσφέρει (Beasley et al., 2003; Meade & Salkin, 1989). Εφόσον το χαρτοφυλάκιο κατανέμει τα κεφάλαια σύμφωνα με τη σύνθεση του δείκτη, οι μεταβολές στις τιμές των τίτλων αντικατοπτρίζονται στο χαρτοφυλάκιο με τρόπο παρόμοιο με εκείνον που επηρεάζουν τον δείκτη. Αυτό μειώνει σημαντικά την απόκλιση μεταξύ των δύο και καθιστά τη συγκεκριμένη προσέγγιση ιδιαίτερα ελκυστική όταν ο πρωταρχικός στόχος είναι η μέγιστη δυνατή πιστότητα αποτύπωσης (Beasley et al., 2003; Derigs & Nickel, 2003).

Ένα δεύτερο πλεονέκτημα της πλήρους αναπαραγωγής είναι η διαφάνεια και η απλότητα της στρατηγικής. Η επενδυτική πολιτική είναι σαφής, εύκολα ερμηνεύσιμη και βασίζεται σε έναν αντικειμενικό κανόνα, δηλαδή στην ακριβή παρακολούθηση της σύνθεσης του δείκτη (Meade & Salkin, 1989; Jiang et al., 2021). Δεν απαιτείται πολύπλοκη διαδικασία επιλογής τίτλων, καθώς όλοι οι τίτλοι του δείκτη περιλαμβάνονται υποχρεωτικά στο χαρτοφυλάκιο (Jiang et al., 2021). Επιπλέον, η λογική της στρατηγικής είναι εύκολα

κατανοητή από επενδυτές και διαχειριστές, γεγονός που την καθιστά ιδιαίτερα κατάλληλη για επενδυτικά προϊόντα που επιδιώκουν διαφάνεια και σαφή προσανατολισμό (Meade & Salkin, 1989).

Παρά τα πλεονεκτήματα αυτά, η πλήρης αναπαραγωγή παρουσιάζει σημαντικούς πρακτικούς περιορισμούς. Ο σημαντικότερος από αυτούς σχετίζεται με το πλήθος των τίτλων που περιλαμβάνει ένας δείκτης. Σε περιπτώσεις όπου ο δείκτης αποτελείται από μεγάλο αριθμό μετοχών, η κατοχή όλων των τίτλων μπορεί να είναι διοικητικά και λειτουργικά δύσκολη (Beasley et al., 2003; Ni & Yin, 2013). Το πρόβλημα γίνεται ακόμη πιο έντονο όταν ορισμένοι τίτλοι έχουν πολύ μικρό βάρος, καθώς η διατήρηση εξαιρετικά μικρών θέσεων μπορεί να μην είναι αποδοτική από πλευράς διαχείρισης και να αυξάνει δυσανάλογα την πολυπλοκότητα του χαρτοφυλακίου (Beasley et al., 2003; Ni & Yin, 2013).

Ένας ακόμη σοβαρός περιορισμός αφορά το κόστος συναλλαγών. Η πλήρης αναπαραγωγή προϋποθέτει την αγορά όλων των τίτλων του δείκτη και, σε πολλές περιπτώσεις, τη συχνή αναπροσαρμογή του χαρτοφυλακίου ώστε να αντικατοπτρίζει μεταβολές στη σύνθεση ή στα βάρη του δείκτη (Beasley et al., 2003; Wu et al., 2019). Κάθε τέτοια μεταβολή συνεπάγεται συναλλαγές, οι οποίες επιβαρύνουν το χαρτοφυλάκιο με κόστος. Όταν οι αναθεωρήσεις του δείκτη είναι συχνές ή όταν το πλήθος των τίτλων είναι μεγάλο, το συνολικό κόστος μπορεί να γίνει σημαντικό και να μειώσει την πρακτική ελκυστικότητα της προσέγγισης (Wu et al., 2019; Ni & Yin, 2013).

Επιπλέον, η πλήρης αναπαραγωγή μπορεί να προσκρούει σε περιορισμούς ρευστότητας και διαθεσιμότητας τίτλων. Δεν είναι πάντοτε εύκολο να αποκτηθούν όλοι οι τίτλοι στις ακριβείς αναλογίες που απαιτεί ο δείκτης, ιδιαίτερα όταν ορισμένοι από αυτούς παρουσιάζουν περιορισμένη εμπορευσιμότητα ή όταν οι θέσεις που αντιστοιχούν σε αυτούς είναι πολύ μικρές σε σχέση με το συνολικό χαρτοφυλάκιο (Ni & Yin, 2013). Σε τέτοιες περιπτώσεις, η προσπάθεια ακριβούς αναπαραγωγής μπορεί να οδηγήσει σε αυξημένες δυσκολίες εκτέλεσης και σε αποκλίσεις από την επιθυμητή σύνθεση (Ni & Yin, 2013; Derigs & Nickel, 2003).

Ένα ακόμη ζήτημα συνδέεται με τη λειτουργική ευελιξία του χαρτοφυλακίου. Η πλήρης αναπαραγωγή επιβάλλει ουσιαστικά την πιστή τήρηση της σύνθεσης του δείκτη, αφήνοντας πολύ μικρά περιθώρια προσαρμογής της στρατηγικής σε πρακτικούς περιορισμούς ή ειδικές επενδυτικές ανάγκες (Beasley et al., 2003; Canakgoz & Beasley, 2009). Αντίθετα, σε πολλές περιπτώσεις είναι προτιμότερο να υπάρχει η δυνατότητα επιλογής ενός πιο περιορισμένου και διαχειρίσιμου συνόλου τίτλων, το οποίο να επιτυγχάνει ικανοποιητική αποτύπωση χωρίς να απαιτεί την ακριβή κατοχή κάθε στοιχείου του δείκτη (Maringer & Oyewumi, 2007; Jiang et al., 2021).

Για τους λόγους αυτούς, η πλήρης αναπαραγωγή, παρότι αποτελεί την πιο άμεση και θεωρητικά ορθή προσέγγιση, δεν είναι πάντοτε η βέλτιστη λύση υπό πραγματικές συνθήκες (Jiang et al., 2021; Ni & Yin, 2013). Αποτελεί μεν ένα φυσικό σημείο εκκίνησης και ένα χρήσιμο πρότυπο αναφοράς, αλλά η πρακτική εφαρμογή της συχνά περιορίζεται από το μέγεθος του δείκτη, το κόστος συναλλαγών, τη ρευστότητα των τίτλων και τις λειτουργικές απαιτήσεις της διαχείρισης (Beasley et al., 2003; Wu et al., 2019; Ni & Yin, 2013). Για τον λόγο αυτό, σε πολλές εφαρμογές προτιμάται η μερική αναπαραγωγή δείκτη, η οποία επιχειρεί να ισορροπήσει ανάμεσα στην ακρίβεια παρακολούθησης και στη διαχειριστική αποδοτικότητα (Maringer & Oyewumi, 2007; Jiang et al., 2021).

2.3.2 Μερική αναπαραγωγή δείκτη

Η μερική αναπαραγωγή δείκτη αποτελεί την εναλλακτική προσέγγιση απέναντι στην πλήρη αναπαραγωγή και, σε πολλές περιπτώσεις, την πιο ρεαλιστική στρατηγική για την αποτύπωση της πορείας ενός χρηματιστηριακού δείκτη (Maringer & Oyewumi, 2007; Beasley et al., 2003; Sant'Anna et al., 2020). Στην προσέγγιση αυτή, το χαρτοφυλάκιο δεν περιλαμβάνει το σύνολο των τίτλων που συμμετέχουν στον δείκτη αναφοράς, αλλά μόνο ένα επιλεγμένο υποσύνολο αυτών. Παράλληλα, καθορίζονται κατάλληλα βάρη για τους επιλεγμένους τίτλους, έτσι ώστε η συνολική συμπεριφορά του χαρτοφυλακίου να προσεγγίζει όσο το δυνατόν καλύτερα τη συμπεριφορά του δείκτη (Maringer & Oyewumi, 2007; Beasley et al., 2003).

Η λογική της μερικής αναπαραγωγής βασίζεται στην παραδοχή ότι δεν είναι πάντοτε αναγκαίο να περιλαμβάνονται όλοι οι τίτλοι του δείκτη προκειμένου να επιτευχθεί ικανοποιητική παρακολούθηση της πορείας του. Εφόσον ένας δείκτης αποτυπώνει τη συνολική κίνηση μιας αγοράς ή ενός επενδυτικού συνόλου, είναι δυνατόν σε αρκετές περιπτώσεις ένα μικρότερο και προσεκτικά σχεδιασμένο χαρτοφυλάκιο να αναπαράγει σε ικανοποιητικό βαθμό τη συνολική του συμπεριφορά (Beasley et al., 2003; Coleman et al., 2006; Sant'Anna et al., 2020). Η βασική πρόκληση, επομένως, δεν είναι η απλή κατοχή όλων των τίτλων, αλλά η επιλογή εκείνων που συνεισφέρουν περισσότερο στην παρακολούθηση του δείκτη και ο κατάλληλος προσδιορισμός των βαρών τους (Maringer & Oyewumi, 2007; Coleman et al., 2006).

Η προσέγγιση αυτή αποκτά ιδιαίτερη σημασία όταν ο δείκτης περιλαμβάνει μεγάλο αριθμό τίτλων ή όταν η πλήρης αναπαραγωγή συνεπάγεται αυξημένα διοικητικά και συναλλακτικά βάρη (Beasley et al., 2003; Maringer & Oyewumi, 2007; Ni & Yin, 2013). Με τη μερική αναπαραγωγή, ο διαχειριστής επιδιώκει να περιορίσει τον αριθμό των τίτλων που κατέχει, να μειώσει το κόστος συναλλαγών και να διατηρήσει ένα πιο διαχειρίσιμο χαρτοφυλάκιο, χωρίς όμως να θυσιάσει υπερβολικά την ακρίβεια παρακολούθησης (Canakgoz & Beasley, 2009; Ni & Yin, 2013; Sant'Anna et al., 2020). Έτσι, η στρατηγική αυτή συνδέεται στενά με την ανάγκη εξισορρόπησης δύο αντικρουόμενων επιδιώξεων: αφενός της μικρής απόκλισης από τον δείκτη και αφετέρου της λειτουργικής και οικονομικής αποδοτικότητας της διαχείρισης (Maringer & Oyewumi, 2007; Canakgoz & Beasley, 2009).

Ένα από τα σημαντικότερα πλεονεκτήματα της μερικής αναπαραγωγής είναι ακριβώς η μείωση του αριθμού των τίτλων που περιλαμβάνει το χαρτοφυλάκιο. Η μείωση αυτή απλοποιεί τη διαδικασία διαχείρισης, μειώνει την πολυπλοκότητα της παρακολούθησης και περιορίζει τις ανάγκες για συχνές και πολυάριθμες συναλλαγές (Coleman et al., 2006; Ni & Yin, 2013). Παράλληλα, η αποφυγή διατήρησης πολύ μικρών θέσεων σε τίτλους με χαμηλή συμμετοχή στον δείκτη μπορεί να βελτιώσει τη λειτουργική αποδοτικότητα του χαρτοφυλακίου (Ni & Yin, 2013; Sant'Anna et al., 2020). Ιδίως σε μεγάλους δείκτες, η δυνατότητα συγκρότησης ενός μικρότερου αλλά αντιπροσωπευτικού χαρτοφυλακίου αποτελεί σημαντικό πρακτικό πλεονέκτημα (Maringer & Oyewumi, 2007; Sant'Anna et al., 2020).

Ένα δεύτερο πλεονέκτημα της μερικής αναπαραγωγής είναι ότι επιτρέπει μεγαλύτερη ευελιξία στην ενσωμάτωση ρεαλιστικών περιορισμών. Σε αντίθεση με την πλήρη αναπαραγωγή, όπου η σύνθεση του δείκτη καθορίζει σχεδόν απόλυτα τη σύνθεση του χαρτοφυλακίου, η μερική αναπαραγωγή επιτρέπει την προσαρμογή της λύσης σε

παράγοντες όπως ο διαθέσιμος προϋπολογισμός, τα όρια συμμετοχής των τίτλων, η ρευστότητα της αγοράς και το κόστος συναλλαγών (Canakgoz & Beasley, 2009; Ni & Yin, 2013; Yuen et al., 2022). Με άλλα λόγια, η στρατηγική αυτή δεν στοχεύει μόνο στην πιστή παρακολούθηση του δείκτη, αλλά στην επίτευξή της υπό συνθήκες που είναι περισσότερο συμβατές με τις ανάγκες της πραγματικής διαχείρισης χαρτοφυλακίου (Canakgoz & Beasley, 2009; Yuen et al., 2022).

Ωστόσο, η μερική αναπαραγωγή συνοδεύεται και από σημαντικές δυσκολίες. Σε αντίθεση με την πλήρη αναπαραγωγή, όπου η σύνθεση του χαρτοφυλακίου καθορίζεται άμεσα από τη σύνθεση του δείκτη, εδώ πρέπει να ληφθούν αποφάσεις τόσο για την επιλογή των τίτλων όσο και για τον καθορισμό των βαρών τους. Αυτό μετατρέπει τη διαδικασία σε ένα σύνθετο πρόβλημα βελτιστοποίησης, στο οποίο συνυπάρχουν διακριτές αποφάσεις επιλογής και συνεχείς αποφάσεις κατανομής κεφαλαίου (Beasley et al., 2003; Canakgoz & Beasley, 2009; Coleman et al., 2006). Η πολυπλοκότητα αυξάνεται ακόμη περισσότερο όταν εισάγονται περιορισμοί στον αριθμό των τίτλων, στο κόστος συναλλαγών ή στα κατώτατα και ανώτατα όρια συμμετοχής κάθε μετοχής (Canakgoz & Beasley, 2009; Yuen et al., 2022).

Η βασική δυσκολία της μερικής αναπαραγωγής έγκειται στο ότι το χαρτοφυλάκιο πρέπει να είναι όσο το δυνατόν πιο αντιπροσωπευτικό της συνολικής συμπεριφοράς του δείκτη, παρά το γεγονός ότι περιλαμβάνει μόνο μέρος των τίτλων του. Αυτό σημαίνει ότι οι επιλεγμένοι τίτλοι πρέπει να έχουν την ικανότητα να «συμπυκνώνουν» την πληροφορία και τη δυναμική του συνόλου (Maringer & Oyewumi, 2007; Sant'Anna et al., 2020). Η επιλογή αυτή δεν είναι προφανής, καθώς δεν αρκεί να προτιμηθούν οι τίτλοι με τα μεγαλύτερα βάρη ή οι πλέον γνωστές μετοχές· απαιτείται μια πιο σύνθετη αξιολόγηση της συμβολής κάθε τίτλου στη συνολική κίνηση του δείκτη και της αλληλεπίδρασής του με τους υπόλοιπους τίτλους (Maringer & Oyewumi, 2007; Coleman et al., 2006).

Για τον λόγο αυτό, η μερική αναπαραγωγή βρίσκεται στον πυρήνα του προβλήματος της αποτύπωσης δείκτη ως προβλήματος βελτιστοποίησης. Στην πραγματικότητα, το ενδιαφέρον της σχετικής βιβλιογραφίας εστιάζεται σε μεγάλο βαθμό ακριβώς σε αυτή τη μορφή του προβλήματος, καθώς εκεί εμφανίζονται οι ουσιαστικές δυσκολίες επιλογής, κατανομής και ελέγχου των περιορισμών (Beasley et al., 2003; Maringer & Oyewumi, 2007; Canakgoz & Beasley, 2009). Η ανάγκη εξεύρεσης χαρτοφυλακίων που να επιτυγχάνουν μικρό σφάλμα αποτύπωσης με περιορισμένο αριθμό τίτλων και ελεγχόμενο κόστος συναλλαγών καθιστά τη μερική αναπαραγωγή ιδιαίτερα πρόσφορο πεδίο εφαρμογής προηγμένων μεθόδων βελτιστοποίησης (Canakgoz & Beasley, 2009; Coleman et al., 2006; Sant'Anna et al., 2020).

Επιπλέον, η μερική αναπαραγωγή προσφέρεται ιδιαίτερα για την εφαρμογή ευρετικών και μεταευρετικών τεχνικών, όπως οι Γενετικοί Αλγόριθμοι. Αυτό συμβαίνει επειδή ο χώρος των δυνατών λύσεων είναι πολύ μεγάλος και η αναζήτηση του βέλτιστου συνδυασμού τίτλων και βαρών δεν είναι εύκολο να επιτευχθεί με κλασικές ακριβείς μεθόδους, ιδίως όταν το μοντέλο περιλαμβάνει ρεαλιστικούς περιορισμούς (Beasley et al., 2003; Ni & Yin, 2013; Yuen et al., 2022). Έτσι, η μερική αναπαραγωγή δεν αποτελεί μόνο μια πρακτική εναλλακτική της πλήρους αναπαραγωγής, αλλά και το βασικό πεδίο μέσα στο οποίο αναπτύσσονται σύγχρονες αλγοριθμικές προσεγγίσεις για το index tracking (Beasley et al., 2003; Ni & Yin, 2013).

Συνολικά, η μερική αναπαραγωγή δείκτη συνιστά μια πιο ευέλικτη και πρακτική προσέγγιση σε σχέση με την πλήρη αναπαραγωγή, καθώς επιτρέπει τη δημιουργία πιο συμπαγών και διαχειρίσιμων χαρτοφυλακίων, με χαμηλότερο κόστος και μεγαλύτερη προσαρμοστικότητα στους περιορισμούς της αγοράς (Maringer & Oyewumi, 2007; Canakgoz & Beasley, 2009; Sant'Anna et al., 2020). Ταυτόχρονα, όμως, εισάγει ένα πιο απαιτητικό πρόβλημα βελτιστοποίησης, στο οποίο η επιτυχία εξαρτάται από την ορθότητα της επιλογής των τίτλων και την αποτελεσματικότητα της μεθόδου προσδιορισμού των βαρών τους (Beasley et al., 2003; Coleman et al., 2006).

2.3.3 Πλεονεκτήματα και μειονεκτήματα των επιμέρους προσεγγίσεων

Η πλήρης και η μερική αναπαραγωγή δείκτη αποτελούν τις δύο βασικές στρατηγικές αντιμετώπισης του προβλήματος της αποτύπωσης δείκτη (Beasley et al., 2003; Maringer & Oyewumi, 2007). Καθεμία από αυτές παρουσιάζει συγκεκριμένα πλεονεκτήματα και περιορισμούς, οι οποίοι συνδέονται τόσο με την ακρίβεια παρακολούθησης του δείκτη όσο και με τη λειτουργική και οικονομική διάσταση της διαχείρισης του χαρτοφυλακίου (Beasley et al., 2003; Canakgoz & Beasley, 2009). Η συγκριτική εξέταση των δύο προσεγγίσεων είναι ιδιαίτερα σημαντική, διότι επιτρέπει να αναδειχθούν οι βασικοί συμβιβασμοί που χαρακτηρίζουν το πρόβλημα και να αιτιολογηθεί η στροφή της βιβλιογραφίας προς πιο σύνθετες μεθόδους βελτιστοποίησης (Beasley et al., 2003; Canakgoz & Beasley, 2009; Maringer & Oyewumi, 2007).

Το σημαντικότερο πλεονέκτημα της πλήρους αναπαραγωγής είναι η υψηλή ακρίβεια αποτύπωσης που μπορεί να προσφέρει. Εφόσον το χαρτοφυλάκιο περιλαμβάνει όλους τους τίτλους του δείκτη με τα αντίστοιχα βάρη τους, η συμπεριφορά του είναι, σε θεωρητικό επίπεδο, σχεδόν ταυτόσημη με εκείνη του δείκτη αναφοράς (Jiang et al., 2021; Beasley et al., 2003). Αυτό έχει ως αποτέλεσμα την ελαχιστοποίηση του σφάλματος αποτύπωσης και καθιστά την πλήρη αναπαραγωγή την πιο φυσική επιλογή όταν ο πρωταρχικός στόχος είναι η μέγιστη δυνατή πιστότητα παρακολούθησης (Jiang et al., 2021; Derigs & Nickel, 2003). Επιπλέον, η προσέγγιση αυτή διακρίνεται για τη διαφάνεια και την απλότητά της, αφού η σύνθεση του χαρτοφυλακίου καθορίζεται άμεσα από τη σύνθεση του δείκτη, χωρίς να απαιτείται διαδικασία επιλογής τίτλων (Jiang et al., 2021).

Παράλληλα, η πλήρης αναπαραγωγή διευκολύνει την ερμηνεία και την αιτιολόγηση της στρατηγικής, καθώς το χαρτοφυλάκιο βασίζεται σε σαφή και αντικειμενικά δεδομένα (Jiang et al., 2021). Από την άποψη αυτή, είναι ιδιαίτερα κατάλληλη για επενδυτικά προϊόντα στα οποία η σαφήνεια της επενδυτικής πολιτικής και η πιστή σύνδεση με έναν δείκτη αναφοράς αποτελούν βασικές απαιτήσεις (Meade & Salkin, 1989; Jiang et al., 2021). Επίσης, σε περιπτώσεις όπου ο δείκτης περιλαμβάνει σχετικά περιορισμένο αριθμό τίτλων και οι συνθήκες συναλλαγών είναι ευνοϊκές, η πλήρης αναπαραγωγή μπορεί να είναι μια απολύτως λειτουργική και αποδοτική λύση (Jiang et al., 2021).

Ωστόσο, τα πλεονεκτήματα αυτά συνοδεύονται από σημαντικά μειονεκτήματα. Το βασικότερο είναι ότι η πλήρης αναπαραγωγή μπορεί να είναι πρακτικά δύσκολη ή και ασύμφορη όταν ο δείκτης περιλαμβάνει μεγάλο αριθμό τίτλων (Beasley et al., 2003; Jiang et al., 2021). Η ανάγκη διατήρησης όλων των μετοχών, ακόμη και εκείνων με πολύ μικρό βάρος, αυξάνει σημαντικά την πολυπλοκότητα της διαχείρισης και επιβαρύνει το χαρτοφυλάκιο με λειτουργικό κόστος (Beasley et al., 2003). Επιπλέον, οι συχνές αναθεωρήσεις στη σύνθεση ή στα βάρη του δείκτη απαιτούν συνεχή προσαρμογή του

χαρτοφυλακίου, γεγονός που οδηγεί σε αυξημένα κόστη συναλλαγών και σε πιθανές δυσκολίες εκτέλεσης, ιδίως όταν ορισμένοι τίτλοι εμφανίζουν περιορισμένη ρευστότητα (Canakgoz & Beasley, 2009; Jiang et al., 2021).

Αντίθετα, η μερική αναπαραγωγή προσφέρει μεγαλύτερη ευελιξία και διαχειριστική αποδοτικότητα. Το βασικό της πλεονέκτημα είναι ότι επιτρέπει τη δημιουργία χαρτοφυλακίων με μικρότερο αριθμό τίτλων, γεγονός που μειώνει την πολυπλοκότητα της διαχείρισης, διευκολύνει την παρακολούθηση των θέσεων και περιορίζει το κόστος συναλλαγών (Maringer & Oyewumi, 2007; Canakgoz & Beasley, 2009). Με τον τρόπο αυτό, η προσέγγιση αυτή είναι πιο συμβατή με τις συνθήκες της πραγματικής αγοράς, ιδίως όταν οι επενδυτικοί πόροι είναι περιορισμένοι ή όταν ο δείκτης περιλαμβάνει πολύ μεγάλο πλήθος τίτλων (Beasley et al., 2003; Maringer & Oyewumi, 2007).

Ένα ακόμη σημαντικό πλεονέκτημα της μερικής αναπαραγωγής είναι ότι επιτρέπει την ενσωμάτωση πρόσθετων περιορισμών και πρακτικών απαιτήσεων στη διαδικασία κατασκευής του χαρτοφυλακίου. Η επιλογή ενός υποσυνόλου τίτλων επιτρέπει στον διαχειριστή να λάβει υπόψη ζητήματα όπως η ρευστότητα, τα κατώτατα και ανώτατα όρια βαρών, η ανάγκη περιορισμού του πλήθους των θέσεων και η μείωση του κόστους αναπροσαρμογής (Canakgoz & Beasley, 2009; Maringer & Oyewumi, 2007). Έτσι, η μερική αναπαραγωγή αποτελεί πιο ευέλικτη και ρεαλιστική στρατηγική, ιδιαίτερα σε περιβάλλοντα όπου η διαχείριση πρέπει να είναι ταυτόχρονα αποτελεσματική και λειτουργικά βιώσιμη (Canakgoz & Beasley, 2009; Beasley et al., 2003).

Το βασικό μειονέκτημα της μερικής αναπαραγωγής είναι ότι δεν μπορεί, κατά κανόνα, να επιτύχει την ίδια ακρίβεια παρακολούθησης με την πλήρη αναπαραγωγή. Από τη στιγμή που το χαρτοφυλάκιο περιλαμβάνει μόνο ένα μέρος των τίτλων του δείκτη, είναι αναπόφευκτο να προκύπτει κάποιος βαθμός απόκλισης από τη συμπεριφορά του (Beasley et al., 2003; Maringer & Oyewumi, 2007). Το σφάλμα αποτύπωσης εξαρτάται από την ποιότητα της επιλογής των τίτλων, από τον προσδιορισμό των βαρών τους και από το πόσο αντιπροσωπευτικό είναι το τελικό χαρτοφυλάκιο σε σχέση με τον δείκτη (Maringer & Oyewumi, 2007; Jiang et al., 2021). Επομένως, η επιτυχία της προσέγγισης δεν είναι αυτονόητη, αλλά εξαρτάται άμεσα από την αποτελεσματικότητα της μεθόδου βελτιστοποίησης που χρησιμοποιείται (Beasley et al., 2003; Maringer & Oyewumi, 2007).

Ένα δεύτερο μειονέκτημα της μερικής αναπαραγωγής είναι η αυξημένη υπολογιστική δυσκολία που συνεπάγεται. Σε αντίθεση με την πλήρη αναπαραγωγή, όπου η σύνθεση του χαρτοφυλακίου είναι σχεδόν προκαθορισμένη, εδώ πρέπει να ληφθούν αποφάσεις τόσο για την επιλογή των τίτλων όσο και για τον καθορισμό των βαρών τους (Beasley et al., 2003; Canakgoz & Beasley, 2009). Το πρόβλημα αυτό είναι σύνθετο, συνδυαστικό και συχνά μη γραμμικό, ιδιαίτερα όταν συνυπάρχουν περιορισμοί πληθικότητας, κόστους συναλλαγών και ορίων συμμετοχής (Canakgoz & Beasley, 2009; Maringer & Oyewumi, 2007). Για τον λόγο αυτό, η μερική αναπαραγωγή απαιτεί συνήθως τη χρήση εξελιγμένων μεθόδων βελτιστοποίησης, όπως οι ευρετικές, μεταερευτικές και εξελικτικές τεχνικές (Beasley et al., 2003; Maringer & Oyewumi, 2007).

Η σύγκριση των δύο προσεγγίσεων αναδεικνύει έναν σαφή συμβιβασμό μεταξύ ακρίβειας και διαχειριστικής αποδοτικότητας. Η πλήρης αναπαραγωγή προσφέρει υψηλότερη πιστότητα παρακολούθησης, αλλά συνοδεύεται από αυξημένο κόστος, μειωμένη ευελιξία και μεγαλύτερη λειτουργική επιβάρυνση (Jiang et al., 2021; Beasley et al., 2003). Αντίθετα, η μερική αναπαραγωγή μειώνει την πολυπλοκότητα και το κόστος, αλλά απαιτεί πιο

σύνθετη διαδικασία σχεδιασμού και συνήθως οδηγεί σε μεγαλύτερο, αν και ενδεχομένως αποδεκτό, σφάλμα αποτύπωσης (Maringer & Oyewumi, 2007; Canakgoz & Beasley, 2009). Συνεπώς, η επιλογή μεταξύ των δύο προσεγγίσεων δεν είναι απόλυτη, αλλά εξαρτάται από τα χαρακτηριστικά του δείκτη, τους περιορισμούς της αγοράς, τους διαθέσιμους πόρους και τους ειδικούς στόχους της διαχείρισης (Jiang et al., 2021; Canakgoz & Beasley, 2009).

Στο πλαίσιο της παρούσας εργασίας, το ενδιαφέρον επικεντρώνεται κυρίως στη μερική αναπαραγωγή δείκτη, ακριβώς επειδή αυτή συνιστά την πιο απαιτητική αλλά και πιο ρεαλιστική μορφή του προβλήματος αποτύπωσης δείκτη (Beasley et al., 2003; Maringer & Oyewumi, 2007). Η ανάγκη επιλογής υποσυνόλου τίτλων και προσδιορισμού κατάλληλων βαρών, υπό περιορισμούς και με στόχο τη μικρότερη δυνατή απόκλιση από τον δείκτη, δημιουργεί ένα σύνθετο πρόβλημα βελτιστοποίησης που προσφέρεται ιδιαίτερα για την εφαρμογή Γενετικών Αλγορίθμων (Beasley et al., 2003).

2.4 Βασικές συνιστώσες του προβλήματος της αποτύπωσης δείκτη

2.4.1 Επιλογή τίτλων

Η επιλογή τίτλων αποτελεί μία από τις θεμελιώδεις συνιστώσες του προβλήματος της αποτύπωσης δείκτη και, ταυτόχρονα, ένα από τα δυσκολότερα στάδιά του από πλευράς βελτιστοποίησης (Beasley et al., 2003; Canakgoz & Beasley, 2009; Maringer & Oyewumi, 2007). Στο πλαίσιο της πλήρους αναπαραγωγής, το ζήτημα αυτό ουσιαστικά δεν υφίσταται, καθώς όλοι οι τίτλοι του δείκτη περιλαμβάνονται υποχρεωτικά στο χαρτοφυλάκιο. Αντίθετα, στη μερική αναπαραγωγή, η επιλογή τίτλων καθίσταται κεντρικό πρόβλημα, αφού το χαρτοφυλάκιο πρέπει να συγκροτηθεί από ένα περιορισμένο υποσύνολο μετοχών, το οποίο να είναι ικανό να αναπαράγει όσο το δυνατόν πιστότερα τη συνολική συμπεριφορά του δείκτη (Beasley et al., 2003; Maringer & Oyewumi, 2007).

Η σημασία της επιλογής τίτλων είναι ιδιαίτερα μεγάλη, διότι από αυτήν εξαρτάται σε σημαντικό βαθμό η ποιότητα της τελικής λύσης. Ένα χαρτοφυλάκιο που περιλαμβάνει κατάλληλα επιλεγμένους τίτλους μπορεί να επιτύχει μικρό σφάλμα αποτύπωσης ακόμη και αν ο αριθμός των θέσεων είναι περιορισμένος. Αντίθετα, μια ακατάλληλη επιλογή μπορεί να οδηγήσει σε σημαντική απόκλιση από τον δείκτη, ακόμη και αν τα επιλεγμένα βάρη βελτιστοποιηθούν με τον καλύτερο δυνατό τρόπο. Επομένως, η διαδικασία επιλογής τίτλων δεν αποτελεί απλώς ένα προκαταρκτικό βήμα, αλλά έναν κρίσιμο μηχανισμό διαμόρφωσης της επενδυτικής δομής του χαρτοφυλακίου (Coleman et al., 2006; Beasley et al., 2003; Strub & Baumann, 2018).

Το πρόβλημα αυτό είναι σύνθετο, επειδή η καταλληλότητα ενός τίτλου δεν εξαρτάται μόνο από το ατομικό του βάρος στον δείκτη ή από την απόδοσή του μεμονωμένα, αλλά από τη συμβολή του στη συνολική δυνατότητα του χαρτοφυλακίου να παρακολουθεί τον δείκτη. Με άλλα λόγια, ένας τίτλος πρέπει να αξιολογείται όχι απομονωμένα, αλλά σε συνδυασμό με τους υπόλοιπους τίτλους που επιλέγονται. Αυτό σημαίνει ότι η διαδικασία επιλογής έχει έντονα συνδυαστικό χαρακτήρα, καθώς το ενδιαφέρον δεν περιορίζεται στο «ποιοι τίτλοι είναι σημαντικοί», αλλά στο «ποιος συνδυασμός τίτλων είναι καταλληλότερος» για την αποτελεσματική αποτύπωση του δείκτη (Beasley et al., 2003; Canakgoz & Beasley, 2009; Coleman et al., 2006).

Στην πράξη, διάφοροι παράγοντες μπορούν να επηρεάσουν την επιλογή των τίτλων. Ένας πρώτος παράγοντας είναι η συμμετοχή κάθε τίτλου στον δείκτη αναφοράς. Τίτλοι με υψηλότερα βάρη ασκούν συνήθως μεγαλύτερη επίδραση στη συνολική κίνηση του δείκτη και, συνεπώς, η παρουσία τους στο χαρτοφυλάκιο μπορεί να συμβάλλει περισσότερο στη μείωση του σφάλματος αποτύπωσης. Ωστόσο, η επιλογή μόνο με βάση το βάρος δεν είναι πάντοτε επαρκής, διότι ενδέχεται τίτλοι με μικρότερη στάθμιση να παίζουν ουσιαστικό ρόλο στην αναπαραγωγή της δυναμικής συγκεκριμένων τμημάτων του δείκτη ή να εμφανίζουν ισχυρή στατιστική σχέση με τη συνολική του πορεία (Beasley et al., 2003; Maringer & Oyewumi, 2007; Strub & Baumann, 2018).

Ένας δεύτερος σημαντικός παράγοντας είναι η συσχέτιση των τίτλων μεταξύ τους και με τον ίδιο τον δείκτη. Η επιλογή τίτλων που κινούνται με τρόπο παρόμοιο προς τον δείκτη μπορεί να βελτιώσει την ποιότητα της αποτύπωσης, αλλά εξίσου σημαντικό είναι να αποφεύγεται η επιλογή πολλών τίτλων που είναι υπερβολικά όμοιοι μεταξύ τους και προσφέρουν ελάχιστη πρόσθετη πληροφορία στο χαρτοφυλάκιο. Επομένως, η αποτελεσματική επιλογή τίτλων απαιτεί συνήθως έναν βαθμό διαφοροποίησης, ώστε το χαρτοφυλάκιο να μπορεί να αποτυπώνει διαφορετικές πηγές μεταβολής που συνθέτουν την κίνηση του δείκτη (Maringer & Oyewumi, 2007; Jiang et al., 2021; Strub & Baumann, 2018).

Η επιλογή τίτλων επηρεάζεται επίσης από πρακτικούς και λειτουργικούς περιορισμούς. Στο πλαίσιο της πραγματικής διαχείρισης χαρτοφυλακίου, δεν αρκεί οι τίτλοι να είναι θεωρητικά κατάλληλοι για την παρακολούθηση του δείκτη· πρέπει επίσης να είναι επαρκώς ρευστοί, διαπραγματεύσιμοι και συμβατοί με το διαθέσιμο επίπεδο επενδυτικών πόρων. Τίτλοι με χαμηλή ρευστότητα, υψηλό κόστος συναλλαγών ή ιδιαίτερα μικρές επιθυμητές θέσεις μπορεί να δημιουργούν δυσανάλογες δυσκολίες στη διαχείριση και, κατά συνέπεια, να μην είναι επιθυμητοί, ακόμη και αν η καθαρά στατιστική τους συμβολή στη μείωση του tracking error είναι θετική (Canakgoz & Beasley, 2009; Ni & Yin, 2013; Strub & Baumann, 2018).

Επιπλέον, η επιλογή τίτλων συνδέεται στενά με τον περιορισμό πληθικότητας του χαρτοφυλακίου. Όταν ο διαχειριστής ή το μοντέλο επιβάλλει ότι το χαρτοφυλάκιο μπορεί να περιλαμβάνει το πολύ ή ακριβώς έναν συγκεκριμένο αριθμό τίτλων, τότε η διαδικασία επιλογής αποκτά ακόμη μεγαλύτερη βαρύτητα. Στην περίπτωση αυτή, κάθε επιλεγμένος τίτλος «καταναλώνει» μία από τις διαθέσιμες θέσεις του χαρτοφυλακίου και, συνεπώς, η απόφασή του πρέπει να είναι ιδιαίτερα προσεκτική. Η ύπαρξη του περιορισμού πληθικότητας μετατρέπει το πρόβλημα σε καθαρά συνδυαστικό, καθώς πρέπει να αναζητηθεί ο βέλτιστος ή τουλάχιστον ένας ιδιαίτερα αποδοτικός συνδυασμός τίτλων μέσα από έναν πολύ μεγάλο αριθμό εναλλακτικών υποσυνόλων (Coleman et al., 2006; Beasley et al., 2003; Zheng et al., 2020).

Η δυσκολία αυτή εξηγεί γιατί η επιλογή τίτλων αποτελεί ένα από τα βασικά σημεία στα οποία οι κλασικές μέθοδοι βελτιστοποίησης συχνά δυσκολεύονται να αποδώσουν. Καθώς το πλήθος των διαθέσιμων τίτλων αυξάνεται, ο αριθμός των δυνατών υποσυνόλων αυξάνεται εκθετικά, γεγονός που καθιστά ανέφικτη την εξαντλητική αναζήτηση. Για τον λόγο αυτό, η διαδικασία επιλογής τίτλων είναι ιδιαίτερα πρόσφορη για την εφαρμογή ευρετικών και μεταευρετικών τεχνικών, οι οποίες επιτρέπουν την εξερεύνηση μεγάλων χώρων λύσεων και την αναζήτηση ποιοτικών συνδυασμών χωρίς να απαιτείται πλήρης απαρίθμηση όλων των δυνατών επιλογών (Beasley et al., 2003; Maringer & Oyewumi, 2007; Zheng et al., 2020).

2.4.2 Προσδιορισμός βαρών χαρτοφυλακίου

Ο προσδιορισμός των βαρών των τίτλων αποτελεί τη δεύτερη βασική συνιστώσα του προβλήματος της αποτύπωσης δείκτη και είναι άρρηκτα συνδεδεμένος με την επιλογή των τίτλων που θα συμμετέχουν στο χαρτοφυλάκιο (Beasley et al., 2003; de Paulo et al., 2016). Αφού επιλεγεί το υποσύνολο των μετοχών που θα χρησιμοποιηθεί για την αναπαραγωγή της πορείας του δείκτη, το επόμενο κρίσιμο βήμα είναι να καθοριστεί το ποσοστό συμμετοχής κάθε τίτλου στο συνολικό χαρτοφυλάκιο. Η διαδικασία αυτή είναι καθοριστικής σημασίας, καθώς ακόμη και μια ιδιαίτερα καλή επιλογή τίτλων μπορεί να οδηγήσει σε χαμηλής ποιότητας αποτέλεσμα αν τα βάρη δεν προσδιοριστούν με τρόπο συμβατό προς τον στόχο της αποτύπωσης (de Paulo et al., 2016; Yuen et al., 2022).

Σε γενικές γραμμές, τα βάρη του χαρτοφυλακίου εκφράζουν τον τρόπο με τον οποίο κατανέμεται το διαθέσιμο επενδυτικό κεφάλαιο μεταξύ των επιλεγμένων τίτλων. Η κατανομή αυτή δεν είναι ουδέτερη, αφού επηρεάζει άμεσα την τελική συμπεριφορά του χαρτοφυλακίου και, συνεπώς, την απόκλιση του από τον δείκτη αναφοράς. Για τον λόγο αυτό, ο προσδιορισμός των βαρών δεν μπορεί να βασίζεται αποκλειστικά σε απλούς εμπειρικούς κανόνες, αλλά πρέπει να αποτελεί μέρος της συνολικής διαδικασίας βελτιστοποίησης. Στο πλαίσιο του index tracking, τα βάρη πρέπει να επιλέγονται έτσι ώστε η συνδυασμένη κίνηση των τίτλων να αναπαράγει όσο το δυνατόν πιστότερα τη συνολική δυναμική του δείκτη (Chavez-Bedoya & Alonso-Becerra, 2014; Yuen et al., 2022).

Η πιο προφανής λογική θα ήταν τα βάρη του χαρτοφυλακίου να αντιστοιχούν, κατά το δυνατόν, στα βάρη των ίδιων τίτλων μέσα στον δείκτη. Στην περίπτωση της πλήρους αναπαραγωγής, αυτή ακριβώς είναι η συνήθης πρακτική, αφού το χαρτοφυλάκιο επιδιώκει να ταυτιστεί με τη σύνθεση του δείκτη. Όταν όμως εξετάζεται η μερική αναπαραγωγή, η κατάσταση γίνεται πιο σύνθετη. Εφόσον δεν περιλαμβάνονται όλοι οι τίτλοι, τα βάρη των επιλεγμένων μετοχών δεν μπορούν απλώς να αντιγραφούν από τον δείκτη, αλλά πρέπει να αναπροσαρμοστούν έτσι ώστε το τελικό χαρτοφυλάκιο να αντισταθμίζει την απουσία των μη επιλεγμένων τίτλων και να επιτυγχάνει τη βέλτιστη δυνατή προσέγγιση της συνολικής του συμπεριφοράς (Beasley et al., 2003; de Paulo et al., 2016; Yuen et al., 2022).

Ο προσδιορισμός των βαρών είναι ουσιαστικά μια διαδικασία συνεχούς βελτιστοποίησης, καθώς τα βάρη αποτελούν συνήθως συνεχείς μεταβλητές απόφασης. Η δυσκολία του προβλήματος έγκειται στο ότι η μεταβολή του βάρους ενός τίτλου δεν επηρεάζει μόνο την ατομική συμβολή του στο χαρτοφυλάκιο, αλλά και τη συνολική αλληλεπίδραση με τους υπόλοιπους τίτλους. Επομένως, το κατάλληλο βάρος για κάθε μετοχή δεν μπορεί να καθοριστεί ανεξάρτητα, αλλά μόνο σε συνάρτηση με τα βάρη των άλλων επιλεγμένων τίτλων και με τον συνολικό στόχο ελαχιστοποίησης του σφάλματος αποτύπωσης (Chavez-Bedoya & Alonso-Becerra, 2014; de Paulo et al., 2016).

Η διαδικασία αυτή επηρεάζεται από πολλούς παράγοντες. Ένας πρώτος παράγοντας είναι η ιστορική συμπεριφορά των αποδόσεων των τίτλων και η σχέση τους με τις αποδόσεις του δείκτη. Τίτλοι που εμφανίζουν υψηλή συσχέτιση με τον δείκτη ή που αντανάκλουν ισχυρά επιμέρους τμήματα της συνολικής του κίνησης ενδέχεται να λάβουν υψηλότερα βάρη, εφόσον συμβάλλουν ουσιαστικά στη μείωση της απόκλισης. Ωστόσο, η κατανομή βαρών δεν εξαρτάται μόνο από τη μεμονωμένη συμπεριφορά κάθε τίτλου, αλλά και από το πώς ο συνδυασμός των τίτλων αλληλεπιδρά στο επίπεδο του χαρτοφυλακίου (Chavez-Bedoya & Alonso-Becerra, 2014; de Paulo et al., 2016).

Ένας δεύτερος παράγοντας είναι οι περιορισμοί που τίθενται στη σύνθεση του χαρτοφυλακίου. Στην πράξη, τα βάρη δεν είναι δυνατόν να λάβουν οποιεσδήποτε τιμές. Συνήθως ισχύει περιορισμός προϋπολογισμού, σύμφωνα με τον οποίο το άθροισμα των βαρών πρέπει να ισούται με τη μονάδα, εφόσον θεωρείται ότι όλο το διαθέσιμο κεφάλαιο επενδύεται στο χαρτοφυλάκιο. Παράλληλα, μπορεί να επιβάλλονται κατώτατα και ανώτατα όρια συμμετοχής για κάθε τίτλο, ώστε να αποφεύγονται υπερβολικά μικρές ή υπερβολικά μεγάλες θέσεις. Οι περιορισμοί αυτοί είναι ιδιαίτερα σημαντικοί, καθώς εξασφαλίζουν ότι το χαρτοφυλάκιο παραμένει ρεαλιστικό και συμβατό με τις ανάγκες διαχείρισης (Canakgoz & Beasley, 2009; Yuen et al., 2022).

Επιπλέον, ο προσδιορισμός των βαρών συνδέεται στενά με το κόστος συναλλαγών και με την ανάγκη αναπροσαρμογής του χαρτοφυλακίου. Σε περιπτώσεις όπου το χαρτοφυλάκιο προκύπτει από αναθεώρηση μιας ήδη υπάρχουσας επενδυτικής σύνθεσης, η μεταβολή των βαρών συνεπάγεται αγορές και πωλήσεις τίτλων, άρα και κόστος. Συνεπώς, η επιλογή των βαρών δεν μπορεί να εξετάζεται μόνο με βάση τη μείωση του tracking error, αλλά και με βάση το αν οι αναγκαίες μεταβολές είναι οικονομικά αποδεκτές. Το στοιχείο αυτό καθιστά ακόμη πιο έντονο τον πολυκριτηριακό χαρακτήρα του προβλήματος, καθώς η βελτίωση της αποτύπωσης μπορεί να συγκρούεται με την ανάγκη περιορισμού του κόστους (Beasley et al., 2003; Canakgoz & Beasley, 2009; Yuen et al., 2022).

Στο υπολογιστικό επίπεδο, ο προσδιορισμός των βαρών είναι ένα από τα σημεία στα οποία φαίνεται η μικτή φύση του προβλήματος της αποτύπωσης δείκτη. Από τη μία πλευρά υπάρχει η διακριτή επιλογή των τίτλων, και από την άλλη ο συνεχής προσδιορισμός των βαρών τους. Η συνδυασμένη αντιμετώπιση αυτών των δύο αποφάσεων είναι ιδιαίτερα απαιτητική, καθώς η καταλληλότητα ενός συνόλου βαρών εξαρτάται από το ποιοι τίτλοι έχουν επιλεγεί, ενώ η καταλληλότητα των επιλεγμένων τίτλων εξαρτάται από το αν μπορούν να τους αποδοθούν βάρη που οδηγούν σε μικρό σφάλμα αποτύπωσης. Αυτή η αλληλεξάρτηση καθιστά το πρόβλημα ιδιαίτερα κατάλληλο για μεθόδους όπως οι Γενετικοί Αλγόριθμοι, οι οποίοι έχουν τη δυνατότητα να διαχειρίζονται ταυτόχρονα πολλαπλές και ετερογενείς μεταβλητές απόφασης (Beasley et al., 2003; Yuen et al., 2022).

2.4.3 Σφάλμα αποτύπωσης (Tracking Error)

Το σφάλμα αποτύπωσης αποτελεί την κεντρική έννοια στο πρόβλημα της αποτύπωσης δείκτη, καθώς εκφράζει τον βαθμό στον οποίο ένα χαρτοφυλάκιο κατορθώνει να αναπαράγει τη συμπεριφορά του δείκτη αναφοράς. Με απλά λόγια, το tracking error μετρά την απόκλιση μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη. Όσο μικρότερη είναι αυτή η απόκλιση, τόσο επιτυχέστερη θεωρείται η διαδικασία αποτύπωσης και τόσο πιο αποτελεσματικό κρίνεται το χαρτοφυλάκιο ως μέσο παθητικής παρακολούθησης της αγοράς (Roll, 1992; Rudolf et al., 1999; Guastaroba et al., 2020).

Η σημασία του tracking error είναι θεμελιώδης, διότι στο πλαίσιο της παθητικής διαχείρισης η επιτυχία του χαρτοφυλακίου δεν αποτιμάται κυρίως από την απόλυτη απόδοσή του, αλλά από το πόσο κοντά βρίσκεται η πορεία του στην πορεία του δείκτη. Ένα χαρτοφυλάκιο μπορεί να έχει υψηλή απόδοση σε απόλυτους όρους, αλλά να θεωρείται ανεπιτυχές ως χαρτοφυλάκιο αποτύπωσης αν αποκλίνει σημαντικά από τον δείκτη. Αντίστοιχα, ακόμη και σε περιόδους χαμηλών ή αρνητικών αποδόσεων της αγοράς, ένα χαρτοφυλάκιο μπορεί να θεωρηθεί επιτυχημένο εφόσον ακολουθεί με ακρίβεια την αντίστοιχη πτωτική πορεία του

benchmark. Επομένως, το σφάλμα αποτύπωσης αποτελεί το βασικό κριτήριο αξιολόγησης στην περίπτωση του index tracking (Roll, 1992; Beasley et al., 2003; Guastaroba et al., 2020).

Από μαθηματική άποψη, το tracking error ορίζεται ως κάποιο μέτρο της διαφοράς μεταξύ των αποδόσεων του χαρτοφυλακίου και των αποδόσεων του δείκτη σε μια δεδομένη χρονική περίοδο. Η έννοια αυτή μπορεί να αποτυπωθεί με διάφορους τρόπους, ανάλογα με το ποιο ακριβώς χαρακτηριστικό της απόκλισης επιδιώκεται να μετρηθεί. Σε ορισμένες περιπτώσεις, ενδιαφέρει το μέσο επίπεδο απόκλισης, ενώ σε άλλες η έμφαση δίνεται στη διακύμανση των διαφορών ή στη βαρύτερη τιμωρία μεγαλύτερων αποκλίσεων. Για τον λόγο αυτό, στη σχετική βιβλιογραφία έχουν προταθεί διαφορετικοί ορισμοί και μέτρα tracking error, τα οποία εξυπηρετούν διαφορετικές ανάγκες μοντελοποίησης και αξιολόγησης (Roll, 1992; Rudolf et al., 1999; Barro & Canestrelli, 2009).

Μια βασική διάκριση αφορά τη διαφορά ανάμεσα στο tracking error ως μέτρο μέσης απόκλισης και στο tracking error ως μέτρο μεταβλητότητας της απόκλισης. Στην πρώτη περίπτωση, ενδιαφέρει το πόσο κοντά είναι κατά μέσο όρο η απόδοση του χαρτοφυλακίου στην απόδοση του δείκτη. Στη δεύτερη, εξετάζεται το πόσο σταθερή ή ασταθής είναι η διαφορά μεταξύ των δύο σειρών αποδόσεων μέσα στον χρόνο. Η διάκριση αυτή είναι σημαντική, διότι δύο χαρτοφυλάκια μπορεί να έχουν παρόμοια μέση απόκλιση αλλά να διαφέρουν σημαντικά ως προς τη μεταβλητότητα των αποκλίσεών τους, γεγονός που επηρεάζει διαφορετικά την αξιολόγηση της ποιότητάς τους (Roll, 1992; Barro & Canestrelli, 2009).

Στο πλαίσιο της αποτύπωσης δείκτη, συνηθίζεται η χρήση μέτρων που βασίζονται στις διαφορές μεταξύ των αποδόσεων του χαρτοφυλακίου και των αποδόσεων του δείκτη σε διαδοχικές χρονικές στιγμές. Οι διαφορές αυτές μπορούν να αθροίζονται σε απόλυτες τιμές ή σε τετραγωνική μορφή, έτσι ώστε να δίνεται μεγαλύτερη βαρύτητα στις μεγαλύτερες αποκλίσεις. Η επιλογή του κατάλληλου μέτρου εξαρτάται από το πώς ο ερευνητής ή ο διαχειριστής επιθυμεί να ποσοτικοποιήσει την έννοια της «καλής παρακολούθησης». Αν, για παράδειγμα, θεωρείται ότι οι μεγάλες αποκλίσεις είναι ιδιαίτερα ανεπιθύμητες, τότε ένα μέτρο που τις τιμωρεί δυσανάλογα περισσότερο είναι πιο κατάλληλο (Rudolf et al., 1999; Barro & Canestrelli, 2009).

Ιδιαίτερο ενδιαφέρον παρουσιάζει και η περίπτωση στην οποία η έμφαση δίνεται μόνο στις αρνητικές αποκλίσεις, δηλαδή στις περιόδους όπου η απόδοση του χαρτοφυλακίου υπολείπεται της απόδοσης του δείκτη. Η λογική πίσω από αυτή την προσέγγιση είναι ότι, σε ορισμένα επενδυτικά πλαίσια, η υπεραπόδοση του χαρτοφυλακίου έναντι του δείκτη δεν θεωρείται προβληματική, ενώ η υποαπόδοση κρίνεται ανεπιθύμητη. Έτσι, έχουν προταθεί και measures που εστιάζουν στη downside διάσταση του tracking error, δίνοντας έμφαση μόνο στις αρνητικές αποκλίσεις. Αυτή η οπτική είναι σημαντική, επειδή αναδεικνύει ότι το σφάλμα αποτύπωσης δεν είναι αποκλειστικά τεχνικό μέγεθος, αλλά συνδέεται και με το πώς ερμηνεύεται επενδυτικά η απόκλιση από τον benchmark (Rudolf et al., 1999; Huang et al., 2018; Guastaroba et al., 2020).

Το tracking error επηρεάζεται άμεσα από τις βασικές επιλογές κατασκευής του χαρτοφυλακίου. Η επιλογή των τίτλων, ο προσδιορισμός των βαρών τους, η ύπαρξη περιορισμών πληθικότητας, η επιβολή ανώτατων ή κατώτατων ορίων συμμετοχής και το κόστος συναλλαγών διαμορφώνουν από κοινού τη δυνατότητα του χαρτοφυλακίου να ακολουθεί αποτελεσματικά τον δείκτη. Για παράδειγμα, ένας αυστηρός περιορισμός στον

αριθμό των τίτλων είναι πιθανό να αυξήσει το tracking error, καθώς περιορίζει τη δυνατότητα του χαρτοφυλακίου να συλλάβει όλο το εύρος της συμπεριφοράς του δείκτη. Αντίστοιχα, η ανάγκη περιορισμού του κόστους συναλλαγών μπορεί να εμποδίσει την επίτευξη της θεωρητικά καλύτερης αναπροσαρμογής των βαρών (Beasley et al., 2003; Canakgoz & Beasley, 2009; Barro & Canestrelli, 2009).

Από υπολογιστική άποψη, το tracking error αποτελεί συνήθως τον βασικό πυρήνα της αντικειμενικής συνάρτησης του προβλήματος βελτιστοποίησης. Δηλαδή, οι περισσότερες προσεγγίσεις στο index tracking οργανώνονται γύρω από την ελαχιστοποίηση κάποιου κατάλληλου μέτρου απόκλισης από τον δείκτη, υπό ένα σύνολο περιορισμών. Αυτό σημαίνει ότι η επιλογή του ορισμού του tracking error δεν είναι δευτερεύον θέμα, αλλά επηρεάζει άμεσα τη μορφή του μαθηματικού προβλήματος και τη συμπεριφορά της μεθόδου επίλυσης. Διαφορετικοί ορισμοί μπορούν να οδηγήσουν σε διαφορετικά χαρτοφυλάκια, ακόμη και όταν τα διαθέσιμα δεδομένα και οι περιορισμοί παραμένουν οι ίδιοι (Rudolf et al., 1999; Barro & Canestrelli, 2009; Guastaroba et al., 2020).

2.4.4 Κόστος συναλλαγών

Το κόστος συναλλαγών αποτελεί μία από τις σημαντικότερες πρακτικές παραμέτρους στο πρόβλημα της αποτύπωσης δείκτη, καθώς επηρεάζει άμεσα τόσο τη διαμόρφωση όσο και την αναπροσαρμογή του χαρτοφυλακίου (Canakgoz & Beasley, 2009; Mezali & Beasley, 2014; Strub & Baumann, 2018). Σε θεωρητικά υποδείγματα βελτιστοποίησης είναι συχνά δελεαστικό να υποτεθεί ότι οι αγορές και οι πωλήσεις τίτλων πραγματοποιούνται χωρίς επιβάρυνση. Ωστόσο, μια τέτοια παραδοχή απομακρύνει σημαντικά το μοντέλο από τις πραγματικές συνθήκες των χρηματοοικονομικών αγορών. Στην πράξη, κάθε μεταβολή στη σύνθεση ή στα βάρη ενός χαρτοφυλακίου συνεπάγεται συγκεκριμένο κόστος, το οποίο πρέπει να λαμβάνεται υπόψη αν επιδιώκεται μια ρεαλιστική και εφαρμόσιμη λύση (Lobo et al., 2007; Mezali & Beasley, 2014).

Το κόστος συναλλαγών συνδέεται με όλες τις ενέργειες αγοράς και πώλησης τίτλων που απαιτούνται για τη δημιουργία ή την αναθεώρηση ενός χαρτοφυλακίου. Στην περίπτωση της αρχικής κατασκευής ενός χαρτοφυλακίου αποτύπωσης δείκτη, το κόστος αυτό προκύπτει από την απόκτηση των επιλεγμένων μετοχών στις επιθυμητές ποσότητες. Στην περίπτωση της αναπροσαρμογής ενός ήδη υφιστάμενου χαρτοφυλακίου, προκύπτει από τις αναγκαίες αγοραπωλησίες που απαιτούνται ώστε το χαρτοφυλάκιο να προσαρμοστεί είτε σε νέες συνθήκες αγοράς είτε σε αλλαγές στη σύνθεση και στα βάρη του δείκτη αναφοράς. Συνεπώς, το κόστος συναλλαγών δεν αφορά μόνο μια εφάπαξ αρχική δαπάνη, αλλά μπορεί να αποτελεί διαρκή επιβάρυνση της επενδυτικής στρατηγικής (Strub & Baumann, 2018; Wu et al., 2019).

Η σημασία του κόστους συναλλαγών είναι ιδιαίτερα έντονη στο πρόβλημα της αποτύπωσης δείκτη, επειδή ο στόχος της ακριβούς παρακολούθησης του benchmark οδηγεί συχνά στην ανάγκη συχνών αναπροσαρμογών. Όσο περισσότερο επιδιώκεται η ελαχιστοποίηση του σφάλματος αποτύπωσης, τόσο πιθανότερο είναι να απαιτούνται μεταβολές στα βάρη ή ακόμη και στη σύνθεση του χαρτοφυλακίου. Αυτές οι παρεμβάσεις, αν και μπορεί να βελτιώνουν την ακρίβεια παρακολούθησης, συνοδεύονται από επιπλέον κόστος. Επομένως, ανακύπτει ένας ουσιαστικός συμβιβασμός μεταξύ της επιδίωξης πολύ μικρού tracking error και της ανάγκης περιορισμού των δαπανών συναλλαγών (Mezali & Beasley, 2014; Strub & Baumann, 2018; Yuen et al., 2022).

Το κόστος συναλλαγών μπορεί να λάβει διαφορετικές μορφές. Σε ένα γενικό επίπεδο, περιλαμβάνει άμεσα κόστη, όπως προμήθειες μεσιτείας, τέλη αγοράς και λοιπές χρεώσεις που επιβάλλονται κατά την εκτέλεση των συναλλαγών. Πέρα όμως από τα άμεσα αυτά στοιχεία, υπάρχουν και έμμεσες μορφές κόστους, όπως η διαφορά μεταξύ τιμών αγοράς και πώλησης, η επίδραση μεγάλων εντολών στην τιμή του τίτλου και η δυσκολία εκτέλεσης πράξεων σε τίτλους περιορισμένης ρευστότητας. Αν και τα έμμεσα κόστη μπορεί να είναι δυσκολότερο να αποτυπωθούν με ακρίβεια, αποτελούν ουσιαστικό μέρος του πραγματικού οικονομικού βάρους που συνεπάγεται η διαχείριση ενός χαρτοφυλακίου (Lobo et al., 2007; Wu et al., 2019).

Στη μαθηματική μοντελοποίηση του προβλήματος, το κόστος συναλλαγών μπορεί να ενσωματωθεί με διαφορετικούς τρόπους. Μια πρώτη προσέγγιση είναι να θεωρηθεί ως μέρος της αντικειμενικής συνάρτησης, έτσι ώστε η διαδικασία βελτιστοποίησης να αναζητά λύσεις που εξισορροπούν την ελαχιστοποίηση του tracking error με τον περιορισμό των δαπανών. Μια δεύτερη προσέγγιση είναι να επιβληθεί ως ρητός περιορισμός, δηλαδή να ορίζεται ένα ανώτατο επιτρεπτό επίπεδο συνολικού κόστους συναλλαγών που το χαρτοφυλάκιο δεν μπορεί να υπερβεί. Η δεύτερη αυτή λογική είναι ιδιαίτερα συμβατή με περιβάλλοντα στα οποία η διαχείριση διέπεται από συγκεκριμένα λειτουργικά ή θεσμικά όρια ως προς το ύψος των συναλλακτικών δαπανών (Canakgoz & Beasley, 2009; Mezali & Beasley, 2014; Yuen et al., 2022).

Η ενσωμάτωση του κόστους συναλλαγών στο μοντέλο προσδίδει μεγαλύτερο ρεαλισμό, αλλά ταυτόχρονα αυξάνει σημαντικά την πολυπλοκότητα του προβλήματος. Αυτό συμβαίνει επειδή η επίδρασή του δεν είναι ανεξάρτητη από τις υπόλοιπες αποφάσεις του χαρτοφυλακίου. Η επιλογή των τίτλων, ο προσδιορισμός των βαρών, η ανάγκη διατήρησης ή μεταβολής θέσεων και η συχνότητα αναπροσαρμογής αλληλεπιδρούν με το κόστος συναλλαγών και επηρεάζουν από κοινού την ποιότητα της τελικής λύσης. Για παράδειγμα, ένα χαρτοφυλάκιο που θεωρητικά επιτυγχάνει πολύ μικρό σφάλμα αποτύπωσης μπορεί να αποδειχθεί πρακτικά μη αποδοτικό αν απαιτεί υπερβολικά μεγάλο αριθμό συναλλαγών (Strub & Baumann, 2018; Wu et al., 2019).

Το κόστος συναλλαγών συνδέεται επίσης άμεσα με τον αριθμό των τίτλων που περιλαμβάνει το χαρτοφυλάκιο. Ένα χαρτοφυλάκιο με μεγάλο πλήθος θέσεων είναι πιθανό να απαιτεί περισσότερες πράξεις για την αρχική του συγκρότηση και για κάθε μεταγενέστερη αναπροσαρμογή. Αντίθετα, ένα πιο συμπαγές χαρτοφυλάκιο μπορεί να είναι ευκολότερο και φθηνότερο στη διαχείριση, ακόμη και αν συνοδεύεται από κάπως μεγαλύτερο σφάλμα αποτύπωσης. Αυτή η σχέση εξηγεί σε μεγάλο βαθμό γιατί το πρόβλημα της αποτύπωσης δείκτη συνδέεται τόσο στενά με τον περιορισμό πληθικότητας και γιατί η μερική αναπαραγωγή αποτελεί συχνά πιο ρεαλιστική επιλογή σε σχέση με την πλήρη αναπαραγωγή (Beasley et al., 2003; Mezali & Beasley, 2014; Yuen et al., 2022).

Ιδιαίτερο ενδιαφέρον παρουσιάζει το κόστος συναλλαγών και στη διαδικασία επανασταθμίσεων του χαρτοφυλακίου. Οι χρηματιστηριακοί δείκτες δεν παραμένουν σταθεροί στο χρόνο, αλλά μεταβάλλουν τη σύνθεση και τα βάρη τους λόγω αλλαγών στις τιμές των μετοχών, εισόδου ή εξόδου τίτλων και αναθεωρήσεων της μεθοδολογίας τους. Αν ένα χαρτοφυλάκιο επιδιώκει να παρακολουθεί στενά αυτές τις μεταβολές, τότε πρέπει να αναπροσαρμόζεται ανάλογα. Η συνεχής αυτή προσαρμογή μπορεί να αυξήσει σημαντικά το συνολικό κόστος διαχείρισης, γεγονός που καθιστά απαραίτητη την εξισορρόπηση μεταξύ συχνότητας αναθεώρησης και οικονομικής αποδοτικότητας (Strub & Baumann, 2018; Sammon, 2026).

Από την πλευρά της βελτιστοποίησης, το κόστος συναλλαγών αποτελεί έναν ακόμη λόγο για τον οποίο το πρόβλημα της αποτύπωσης δείκτη δεν μπορεί να αντιμετωπιστεί επαρκώς μόνο με απλοποιημένα κλασικά μοντέλα. Η παρουσία του εισάγει μη γραμμικότητα, ενδεχόμενη ασυνέχεια και ισχυρές αλληλεξαρτήσεις μεταξύ των μεταβλητών απόφασης, καθιστώντας το πρόβλημα πιο απαιτητικό. Για τον λόγο αυτό, η σχετική βιβλιογραφία έχει στραφεί σε μεγάλο βαθμό προς ευρετικές και εξελικτικές μεθόδους, οι οποίες μπορούν να διαχειριστούν πιο αποτελεσματικά τέτοιου είδους ρεαλιστικούς περιορισμούς (Mezali & Beasley, 2014; Yuen et al., 2022; Rubio-García et al., 2024).

2.4.5 Περιορισμός πλήθους τίτλων

Ο περιορισμός πλήθους τίτλων αποτελεί μία από τις πιο χαρακτηριστικές και ουσιαστικές διαστάσεις του προβλήματος της αποτύπωσης δείκτη. Η έννοιά του αναφέρεται στον περιορισμό του αριθμού των μετοχών που επιτρέπεται να συμμετέχουν στο χαρτοφυλάκιο, ανεξάρτητα από το πόσοι τίτλοι περιλαμβάνονται συνολικά στον δείκτη αναφοράς. Στην πράξη, ο περιορισμός αυτός εισάγεται ώστε το χαρτοφυλάκιο να παραμένει διαχειρίσιμο, λειτουργικά αποδοτικό και οικονομικά βιώσιμο, ιδιαίτερα όταν ο δείκτης περιλαμβάνει πολύ μεγάλο αριθμό τίτλων (Beasley et al., 2003; Chang et al., 2000; Yuen et al., 2022).

Η ανάγκη επιβολής περιορισμού στο πλήθος των τίτλων προκύπτει από το γεγονός ότι η κατοχή μεγάλου αριθμού μετοχών δεν είναι πάντοτε επιθυμητή ή εφικτή. Ένα χαρτοφυλάκιο που περιλαμβάνει υπερβολικά πολλές θέσεις μπορεί να είναι δύσκολο στην παρακολούθηση, να απαιτεί αυξημένο διοικητικό έλεγχο και να συνεπάγεται σημαντικό κόστος τόσο κατά την αρχική συγκρότησή του όσο και κατά την επαναστάθμιση. Επιπλέον, η διατήρηση πολύ μικρών θέσεων σε μεγάλο αριθμό τίτλων συχνά δεν είναι πρακτικά αποδοτική, ιδιαίτερα όταν τα σχετικά κόστη συναλλαγών και διαχείρισης είναι δυσανάλογα σε σχέση με τη συνεισφορά αυτών των θέσεων στην ακρίβεια αποτύπωσης του δείκτη (Chang et al., 2000; Strub & Baumann, 2018; Leung et al., 2022).

Στο πλαίσιο της πλήρους αναπαραγωγής, ο περιορισμός πλήθους τίτλων ουσιαστικά δεν εφαρμόζεται, αφού όλοι οι τίτλοι του δείκτη περιλαμβάνονται στο χαρτοφυλάκιο. Αντίθετα, στη μερική αναπαραγωγή, ο περιορισμός αυτός βρίσκεται στον πυρήνα του προβλήματος, καθώς το ζητούμενο είναι να επιλεγεί ένα σχετικά μικρό αλλά κατάλληλο υποσύνολο τίτλων που να επιτρέπει την αποτελεσματική παρακολούθηση του δείκτη. Για τον λόγο αυτό, ο περιορισμός πλήθους δεν είναι απλώς μια πρακτική επιλογή του διαχειριστή, αλλά βασικός παράγοντας που καθορίζει τη μορφή και τη δυσκολία του προβλήματος βελτιστοποίησης (Beasley et al., 2003; Maringer & Oyewumi, 2007; Zheng et al., 2020).

Η εισαγωγή περιορισμού στον αριθμό των τίτλων μετασχηματίζει ουσιαστικά το πρόβλημα από ένα καθαρά συνεχές πρόβλημα κατανομής βαρών σε ένα μικτό πρόβλημα επιλογής και κατανομής. Πλέον δεν αρκεί να καθοριστούν μόνο τα κατάλληλα βάρη των τίτλων, αλλά πρέπει προηγουμένως ή ταυτόχρονα να αποφασιστεί ποιοι τίτλοι θα περιληφθούν στο χαρτοφυλάκιο και ποιοι θα αποκλειστούν. Αυτό προσδίδει στο πρόβλημα έντονα συνδυαστικό χαρακτήρα, καθώς ο αριθμός των δυνατών υποσυνόλων τίτλων αυξάνεται εκθετικά με το πλήθος των διαθέσιμων μετοχών. Έτσι, ο περιορισμός πλήθους είναι ένας από τους βασικούς λόγους για τους οποίους το πρόβλημα της αποτύπωσης δείκτη καθίσταται υπολογιστικά δύσκολο (Chang et al., 2000; Ruiz-Torrubiano & Suárez, 2009; Zheng et al., 2020).

Από χρηματοοικονομική άποψη, ο περιορισμός πλήθους συνδέεται με την προσπάθεια εξισορρόπησης ανάμεσα στην ακρίβεια παρακολούθησης και στη λειτουργική αποδοτικότητα. Όσο μεγαλύτερος είναι ο επιτρεπτός αριθμός τίτλων, τόσο περισσότερες δυνατότητες έχει το χαρτοφυλάκιο να προσεγγίσει τη συμπεριφορά του δείκτη, καθώς μπορεί να ενσωματώσει μεγαλύτερο μέρος της πληροφορίας που περιέχεται στη σύνθεσή του. Αντίθετα, όσο μικρότερος είναι ο αριθμός των επιτρεπτών θέσεων, τόσο πιο δύσκολο γίνεται να αναπαραχθεί πιστά η πορεία του δείκτη, αλλά τόσο πιο απλή και οικονομική γίνεται η διαχείριση του χαρτοφυλακίου. Συνεπώς, ο περιορισμός αυτός εκφράζει έναν σαφή συμβιβασμό μεταξύ ποιότητας αποτύπωσης και διαχειριστικής απλότητας (Coleman et al., 2006; Canakgoz & Beasley, 2009; García et al., 2017).

Επιπλέον, ο περιορισμός πλήθους τίτλων συνδέεται στενά με το πρόβλημα της διαφοροποίησης. Ένα χαρτοφυλάκιο με πολύ μικρό αριθμό τίτλων ενδέχεται να μην αποτυπώνει επαρκώς τη συνολική δομή και τη διασπορά του δείκτη, γεγονός που μπορεί να αυξήσει το σφάλμα αποτύπωσης και να καταστήσει το χαρτοφυλάκιο εύάλωτο σε ιδιοσυγκρασιακές μεταβολές συγκεκριμένων μετοχών. Από την άλλη πλευρά, η αύξηση του αριθμού των τίτλων μπορεί να ενισχύσει την αντιπροσωπευτικότητα του χαρτοφυλακίου, αλλά με κόστος μεγαλύτερης πολυπλοκότητας και υψηλότερων συναλλακτικών απαιτήσεων. Επομένως, ο καθορισμός του κατάλληλου επιπέδου πληθικότητας είναι στρατηγικής σημασίας για την κατασκευή χαρτοφυλακίων index tracking (Maringer & Oyewumi, 2007; Coleman et al., 2006; García et al., 2017).

Στη μαθηματική διατύπωση του προβλήματος, ο περιορισμός πλήθους εισάγεται συνήθως μέσω δυαδικών μεταβλητών που υποδηλώνουν αν ένας τίτλος συμμετέχει ή όχι στο χαρτοφυλάκιο. Με αυτόν τον τρόπο, επιβάλλεται ότι μόνο ένας συγκεκριμένος αριθμός τίτλων μπορεί να λάβει θετικό βάρος. Η παρουσία τέτοιων δυαδικών μεταβλητών ενισχύει τη μη γραμμικότητα και τον διακριτό χαρακτήρα του προβλήματος, καθιστώντας δύσκολη την αντιμετώπισή του με κλασικές τεχνικές συνεχούς βελτιστοποίησης. Για τον λόγο αυτό, οι περιορισμοί πληθικότητας έχουν αποτελέσει έναν από τους βασικούς λόγους ανάπτυξης και εφαρμογής ευρετικών και μεταευρετικών μεθόδων στο πρόβλημα της βελτιστοποίησης χαρτοφυλακίου (Chang et al., 2000; Canakgoz & Beasley, 2009; Ruiz-Torrubiano & Suárez, 2009).

Η πρακτική σημασία του περιορισμού πλήθους είναι επίσης μεγάλη σε περιβάλλοντα θεσμικής διαχείρισης. Πολλοί επενδυτικοί οργανισμοί ή διαχειριστές κεφαλαίων θέτουν εκ των προτέρων όρια στον αριθμό των τίτλων που μπορούν να περιλαμβάνονται σε ένα χαρτοφυλάκιο, είτε για λόγους διαχειριστικής πολιτικής είτε για λόγους ελέγχου του λειτουργικού κόστους. Σε αυτές τις περιπτώσεις, ο περιορισμός πληθικότητας δεν είναι απλώς τεχνικό χαρακτηριστικό του μοντέλου, αλλά πραγματική επιχειρησιακή απαίτηση. Έτσι, η ενσωμάτωσή του στο πρόβλημα βελτιστοποίησης είναι αναγκαία ώστε οι λύσεις που προκύπτουν να είναι εφαρμόσιμες και όχι μόνο θεωρητικά αποδοτικές (Chang et al., 2000; Leung et al., 2022).

Από αλγοριθμική άποψη, ο περιορισμός πλήθους είναι ιδιαίτερα συμβατός με την εφαρμογή Γενετικών Αλγορίθμων, καθώς η επιλογή των τίτλων μπορεί να αναπαρασταθεί άμεσα μέσα στη δομή του χρωμοσώματος. Ένας Γενετικός Αλγόριθμος μπορεί να εξερευνά εναλλακτικούς συνδυασμούς τίτλων, διατηρώντας ή επιβάλλοντας συγκεκριμένο αριθμό ενεργών θέσεων, και ταυτόχρονα να αναζητά τα κατάλληλα βάρη για τους επιλεγμένους τίτλους. Με αυτόν τον τρόπο, ο περιορισμός πλήθους δεν αντιμετωπίζεται ως εξωτερική

μεταγενέστερη διόρθωση, αλλά ενσωματώνεται οργανικά στη διαδικασία αναζήτησης της λύσης (Beasley et al., 2003; García et al., 2017; Chang et al., 2000).

2.4.6 Αναπροσαρμογή χαρτοφυλακίου

Η αναπροσαρμογή χαρτοφυλακίου αποτελεί αναπόσπαστο στοιχείο του προβλήματος της αποτύπωσης δείκτη, καθώς η παρακολούθηση ενός χρηματιστηριακού δείκτη δεν είναι μια στατική διαδικασία, αλλά μια δυναμική επενδυτική πρακτική που εξελίσσεται μέσα στον χρόνο (Strub & Baumann, 2018; Chiam et al., 2013). Ένα χαρτοφυλάκιο το οποίο σε μια δεδομένη χρονική στιγμή αποτυπώνει ικανοποιητικά τη συμπεριφορά ενός δείκτη δεν είναι βέβαιο ότι θα συνεχίσει να το πράττει με την ίδια αποτελεσματικότητα και στο μέλλον. Οι μεταβολές στις τιμές των τίτλων, οι αλλαγές στη σύνθεση του δείκτη, οι μετατοπίσεις στα σχετικά βάρη των μετοχών και οι γενικότερες συνθήκες της αγοράς καθιστούν αναγκαία τη διαδικασία περιοδικής αναθεώρησης και προσαρμογής του χαρτοφυλακίου (Strub & Baumann, 2018; Sammon, 2026).

Η βασική λογική της αναπροσαρμογής είναι ότι, καθώς οι αγορές μεταβάλλονται, το χαρτοφυλάκιο αποκλίνει σταδιακά από τη σύνθεση ή τη συμπεριφορά που απαιτείται για την αποτελεσματική παρακολούθηση του δείκτη. Ακόμη και αν στην αρχική χρονική στιγμή οι τίτλοι και τα βάρη του χαρτοφυλακίου έχουν επιλεγεί με βέλτιστο τρόπο, οι μεταγενέστερες αλλαγές στις αποδόσεις των τίτλων μπορούν να τροποποιήσουν ουσιαστικά τη συμβολή κάθε θέσης στο συνολικό αποτέλεσμα. Επομένως, για να διατηρηθεί η ποιότητα της αποτύπωσης, απαιτείται επανεξέταση της σύνθεσης του χαρτοφυλακίου και, όπου είναι αναγκαίο, προσαρμογή των βαρών ή ακόμη και αντικατάσταση επιμέρους τίτλων (Strub & Baumann, 2018; Mezali & Beasley, 2014).

Η ανάγκη αναπροσαρμογής είναι ιδιαίτερα έντονη στην περίπτωση δεικτών που αναθεωρούνται συχνά. Οι χρηματιστηριακοί δείκτες δεν παραμένουν σταθεροί ως προς τη σύνθεσή τους, καθώς νέοι τίτλοι μπορεί να εισάγονται και άλλοι να απομακρύνονται, ενώ τα βάρη των επιμέρους μετοχών μεταβάλλονται λόγω αλλαγών στις τιμές ή στη μεθοδολογία στάθμισης. Ένα χαρτοφυλάκιο που επιδιώκει να παρακολουθεί στενά έναν τέτοιο δείκτη πρέπει να είναι σε θέση να ανταποκρίνεται στις αλλαγές αυτές, διαφορετικά η απόκλισή του από τον benchmark είναι πιθανό να αυξηθεί σημαντικά. Έτσι, η αναπροσαρμογή δεν είναι απλώς τεχνική δυνατότητα, αλλά ουσιαστική προϋπόθεση για τη διαχρονική διατήρηση της αποτελεσματικότητας της στρατηγικής (Strub & Baumann, 2018; Sammon, 2026).

Από χρηματοοικονομική άποψη, η αναπροσαρμογή εκφράζει τη δυναμική διάσταση της διαχείρισης χαρτοφυλακίου. Το χαρτοφυλάκιο δεν αντιμετωπίζεται ως μια σταθερή κατασκευή που δημιουργείται μία φορά και παραμένει αμετάβλητη, αλλά ως ένα σύστημα που απαιτεί συνεχή ή περιοδική προσαρμογή στις εξελίξεις της αγοράς. Η διαχείριση αυτή μπορεί να είναι περισσότερο ή λιγότερο συχνή, ανάλογα με τη στρατηγική που ακολουθείται. Σε ορισμένες περιπτώσεις επιλέγονται τακτικές αναπροσαρμογές σε προκαθορισμένα χρονικά διαστήματα, ενώ σε άλλες η αναθεώρηση γίνεται όταν η απόκλιση από τον δείκτη ή από την επιθυμητή σύνθεση υπερβεί κάποιο αποδεκτό όριο (Strub & Baumann, 2018; El Bernoussi & Rockinger, 2023).

Παρά τη σημασία της, η αναπροσαρμογή χαρτοφυλακίου δεν είναι χωρίς κόστος. Κάθε μεταβολή στα βάρη ή στη σύνθεση συνεπάγεται αγορές και πωλήσεις τίτλων, άρα και κόστος συναλλαγών. Συνεπώς, η συχνή αναπροσαρμογή μπορεί πράγματι να μειώνει το

σφάλμα αποτύπωσης, αλλά ταυτόχρονα να αυξάνει σημαντικά το συνολικό κόστος διαχείρισης. Εδώ ανακύπτει ένας ακόμη βασικός συμβιβασμός του προβλήματος: η επιδίωξη μεγαλύτερης ακρίβειας παρακολούθησης μέσω συχνών προσαρμογών συγκρούεται με την ανάγκη περιορισμού των συναλλακτικών δαπανών. Για τον λόγο αυτό, η αναπροσαρμογή δεν μπορεί να αξιολογείται μόνο με βάση τη βελτίωση του tracking error, αλλά πρέπει να εξετάζεται συνολικά σε σχέση με το οικονομικό της αποτέλεσμα (Strub & Baumann, 2018; Mezali & Beasley, 2014; Chiam et al., 2013).

Επιπλέον, η αναπροσαρμογή συνδέεται και με τον κίνδυνο υπερβολικής ευαισθησίας στις βραχυχρόνιες διακυμάνσεις των δεδομένων. Αν ένα χαρτοφυλάκιο αναθεωρείται υπερβολικά συχνά με βάση βραχυπρόθεσμες μεταβολές, υπάρχει ο κίνδυνος η διαδικασία να ανταποκρίνεται περισσότερο σε προσωρινό «θόρυβο» της αγοράς παρά σε ουσιαστικές διαρθρωτικές αλλαγές. Σε μια τέτοια περίπτωση, η αναπροσαρμογή μπορεί να καταστεί αντιπαραγωγική, αυξάνοντας το κόστος χωρίς να προσφέρει ανάλογη βελτίωση στην ποιότητα αποτύπωσης. Συνεπώς, η συχνότητα και ο τρόπος αναπροσαρμογής αποτελούν κρίσιμες στρατηγικές αποφάσεις (Strub & Baumann, 2018; Frino et al., 2005).

Στη μαθηματική και αλγοριθμική διατύπωση του προβλήματος, η αναπροσαρμογή μπορεί να αντιμετωπιστεί με διαφορετικούς τρόπους. Μια απλούστερη προσέγγιση είναι να θεωρηθεί ότι το πρόβλημα επιλύεται σε μία μόνο χρονική στιγμή, δηλαδή να εξετάζεται μόνο η αρχική κατασκευή του χαρτοφυλακίου. Μια πιο ρεαλιστική προσέγγιση είναι να ληφθεί υπόψη και η ύπαρξη ενός τρέχοντος χαρτοφυλακίου, από το οποίο πρέπει να μεταβούμε σε ένα νέο χαρτοφυλάκιο μέσω συναλλαγών. Στην περίπτωση αυτή, η αναπροσαρμογή συνδέεται άμεσα με το κόστος μετάβασης από την υπάρχουσα στη νέα σύνθεση, προσδίδοντας στο πρόβλημα σαφώς πιο πρακτικό χαρακτήρα (Strub & Baumann, 2018; Mezali & Beasley, 2014).

Η ύπαρξη αναπροσαρμογής αυξάνει περαιτέρω την πολυπλοκότητα του προβλήματος της αποτύπωσης δείκτη. Πλέον δεν αρκεί να βρεθεί απλώς ένα «καλό» χαρτοφυλάκιο ως προς τον δείκτη, αλλά πρέπει να βρεθεί ένα χαρτοφυλάκιο το οποίο να είναι καλύτερο από το υφιστάμενο, λαμβάνοντας υπόψη και το κόστος της μετάβασης προς αυτό. Με άλλα λόγια, η διαδικασία βελτιστοποίησης αποκτά διαχρονική διάσταση, καθώς η σημερινή απόφαση εξαρτάται όχι μόνο από το πώς συμπεριφέρεται ο δείκτης, αλλά και από την τρέχουσα κατάσταση του χαρτοφυλακίου. Αυτό το στοιχείο καθιστά το πρόβλημα πιο ρεαλιστικό, αλλά και πιο απαιτητικό υπολογιστικά (Strub & Baumann, 2018; Chiam et al., 2013; Mezali & Beasley, 2014).

2.5 Κριτήρια και δείκτες αξιολόγησης της απόδοσης

2.5.1 Μέτρα σφάλματος αποτύπωσης

Η αξιολόγηση ενός χαρτοφυλακίου αποτύπωσης δείκτη βασίζεται πρωτίστως στη μέτρηση του βαθμού απόκλισής του από τον δείκτη αναφοράς. Η απόκλιση αυτή αποτυπώνεται μέσω των μέτρων σφάλματος αποτύπωσης, τα οποία αποτελούν τον βασικό μηχανισμό ποσοτικοποίησης της επιτυχίας ή μη της στρατηγικής παρακολούθησης. Η επιλογή του κατάλληλου μέτρου είναι ιδιαίτερα σημαντική, διότι καθορίζει όχι μόνο τον τρόπο αξιολόγησης των αποτελεσμάτων, αλλά και τη μορφή της αντικειμενικής συνάρτησης όταν το πρόβλημα διατυπώνεται ως πρόβλημα βελτιστοποίησης (Roll, 1992; Rudolf et al., 1999; Barro & Canestrelli, 2009).

Σε γενικές γραμμές, τα μέτρα σφάλματος αποτύπωσης βασίζονται στη σύγκριση των αποδόσεων του χαρτοφυλακίου με τις αντίστοιχες αποδόσεις του δείκτη σε μια σειρά χρονικών περιόδων. Αν συμβολιστεί με r_{trt} η απόδοση του χαρτοφυλακίου και με R_{trt} η απόδοση του δείκτη κατά την περίοδο tt , τότε η βασική πληροφορία πάνω στην οποία στηρίζεται κάθε *measure tracking error* είναι η διαφορά $r_t - R_{trt} - R_t$. Όσο πιο κοντά βρίσκεται αυτή η διαφορά στο μηδέν διαχρονικά, τόσο πιο επιτυχής θεωρείται η αποτύπωση του δείκτη (Roll, 1992; Rudolf et al., 1999).

Ένα από τα πιο απλά και διαισθητικά μέτρα είναι η μέση απόλυτη απόκλιση των αποδόσεων. Στην περίπτωση αυτή, το σφάλμα αποτύπωσης προκύπτει από τον μέσο όρο των απόλυτων διαφορών μεταξύ αποδόσεων χαρτοφυλακίου και δείκτη. Το συγκεκριμένο *measure* είναι εύκολα ερμηνεύσιμο, καθώς δείχνει πόσο αποκλίνει κατά μέσο όρο το χαρτοφυλάκιο από τον δείκτη σε κάθε περίοδο. Πλεονέκτημά του είναι ότι αντιμετωπίζει συμμετρικά θετικές και αρνητικές αποκλίσεις και δεν υπερτιμά δυσανάλογα τις πολύ μεγάλες αποκλίσεις. Ωστόσο, ακριβώς επειδή δεν δίνει αυξημένο βάρος στις ακραίες αποκλίσεις, μπορεί να μην είναι το καταλληλότερο όταν οι μεγάλες αποκλίσεις θεωρούνται ιδιαίτερα ανεπιθύμητες (Rudolf et al., 1999).

Για τον λόγο αυτό, στη σχετική βιβλιογραφία χρησιμοποιείται πολύ συχνά το μέσο τετραγωνικό σφάλμα ή η τετραγωνική μορφή της απόκλισης. Στην περίπτωση αυτή, οι διαφορές μεταξύ αποδόσεων χαρτοφυλακίου και δείκτη υψώνονται στο τετράγωνο πριν αθροιστούν ή ληφθεί ο μέσος όρος τους. Η προσέγγιση αυτή έχει το πλεονέκτημα ότι τιμωρεί περισσότερο τις μεγαλύτερες αποκλίσεις, γεγονός που την καθιστά ιδιαίτερα χρήσιμη όταν η επιδίωξη είναι η αποφυγή έντονων απομακρύνσεων από τον *benchmark*. Επιπλέον, η τετραγωνική μορφή είναι ιδιαίτερα βολική από μαθηματική άποψη και χρησιμοποιείται συχνά σε μοντέλα βελτιστοποίησης, καθώς οδηγεί σε ομαλότερες και πιο επεξεργάσιμες συναρτήσεις στόχου (Roll, 1992; Barro & Canestrelli, 2009).

Συγγενές *measure* με το παραπάνω είναι η ρίζα του μέσου τετραγωνικού σφάλματος, η οποία εκφράζει το *tracking error* σε μονάδες συγκρίσιμες με εκείνες των αποδόσεων. Το πλεονέκτημα αυτής της μορφής είναι ότι παρέχει ένα μέγεθος πιο άμεσα ερμηνεύσιμο, καθώς δεν βρίσκεται πλέον σε τετραγωνικές μονάδες. Παρ' όλα αυτά, από πλευράς βελτιστοποίησης χρησιμοποιείται συχνά η τετραγωνική μορφή χωρίς ρίζα, επειδή είναι υπολογιστικά πιο εύχρηστη και οδηγεί στην ίδια ουσιαστικά κατάταξη των λύσεων (Roll, 1992; Barro & Canestrelli, 2009).

Μια άλλη κατηγορία μέτρων εστιάζει στη διακύμανση ή την τυπική απόκλιση της διαφοράς μεταξύ των αποδόσεων χαρτοφυλακίου και δείκτη. Εδώ το ενδιαφέρον δεν είναι μόνο αν το χαρτοφυλάκιο αποκλίνει από τον δείκτη, αλλά και το πόσο σταθερή ή ασταθής είναι αυτή η απόκλιση στον χρόνο. Ένα χαρτοφυλάκιο μπορεί να παρουσιάζει μικρή μέση απόκλιση, αλλά έντονες διακυμάνσεις γύρω από τον δείκτη, γεγονός που μπορεί να θεωρηθεί ανεπιθύμητο σε ένα πλαίσιο παθητικής διαχείρισης. Για τον λόγο αυτό, *measures* που αποτυπώνουν τη μεταβλητότητα της απόκλισης χρησιμοποιούνται συχνά όταν δίνεται έμφαση στη σταθερότητα της παρακολούθησης και όχι μόνο στη μέση εγγύτητα προς τον *benchmark* (Roll, 1992; Ammann & Zimmermann, 2001).

Ιδιαίτερο ενδιαφέρον παρουσιάζουν και τα *measures* που επικεντρώνονται αποκλειστικά στις αρνητικές αποκλίσεις, δηλαδή στις περιόδους όπου η απόδοση του χαρτοφυλακίου υπολείπεται της απόδοσης του δείκτη. Η λογική πίσω από την προσέγγιση αυτή είναι ότι, σε ορισμένα επενδυτικά περιβάλλοντα, η υπεραπόδοση έναντι του δείκτη δεν θεωρείται

πρόβλημα, ενώ η υποαπόδοση είναι σαφώς ανεπιθύμητη. Έτσι, μπορούν να χρησιμοποιηθούν measures downside tracking error, τα οποία λαμβάνουν υπόψη μόνο τις περιπτώσεις στις οποίες το χαρτοφυλάκιο αποδίδει χειρότερα από τον benchmark. Η προσέγγιση αυτή είναι χρήσιμη όταν το επενδυτικό ενδιαφέρον επικεντρώνεται περισσότερο στην αποφυγή της υστέρησης έναντι του δείκτη παρά στην πλήρη συμμετρική ταύτιση με αυτόν (Rudolf et al., 1999; Huang et al., 2018; Guastaroba et al., 2020).

Πέρα από τις παραπάνω βασικές κατηγορίες, έχουν προταθεί και πιο εξειδικευμένα measures, τα οποία ενσωματώνουν διαφορετικούς βαθμούς ποινής στις αποκλίσεις ή χρησιμοποιούν εναλλακτικές συναρτησιακές μορφές. Σε ορισμένα μοντέλα, για παράδειγμα, η διαφορά μεταξύ χαρτοφυλακίου και δείκτη μπορεί να υψώνεται σε δύναμη διαφορετική του δύο, ώστε να ελέγχεται ο βαθμός ευαισθησίας του measure απέναντι σε μεγαλύτερες αποκλίσεις. Με αυτόν τον τρόπο, το measure tracking error μπορεί να προσαρμόζεται στις ανάγκες της εκάστοτε εφαρμογής και να αντανakλά καλύτερα το πώς ο ερευνητής ή ο διαχειριστής αντιλαμβάνεται τη σοβαρότητα των αποκλίσεων (Guastaroba et al., 2020; Barro & Canestrelli, 2009).

Η επιλογή του κατάλληλου μέτρου σφάλματος αποτύπωσης δεν είναι ουδέτερη, καθώς επηρεάζει ουσιαστικά τα αποτελέσματα της βελτιστοποίησης. Ένα measure που δίνει μεγαλύτερη βαρύτητα στις ακραίες αποκλίσεις μπορεί να οδηγήσει σε πιο «συντηρητικά» χαρτοφυλάκια, τα οποία θυσιάζουν ίσως τη μέση επίδοση για χάρη της αποφυγής έντονων σφαλμάτων. Αντίθετα, ένα measure που εστιάζει περισσότερο στη μέση απόκλιση μπορεί να επιτρέπει μεγαλύτερη ανοχή σε ορισμένες ακραίες διαφορές, αρκεί το χαρτοφυλάκιο να παραμένει γενικά κοντά στον δείκτη. Κατά συνέπεια, η επιλογή measure συνδέεται άμεσα με τη φιλοσοφία της διαχείρισης και με τους ειδικούς στόχους της εφαρμογής (Roll, 1992; Rudolf et al., 1999; Guastaroba et al., 2020).

2.5.2 Απόδοση χαρτοφυλακίου σε σχέση με τον δείκτη

Η αξιολόγηση ενός χαρτοφυλακίου αποτύπωσης δείκτη δεν εξαντλείται μόνο στη μέτρηση του σφάλματος αποτύπωσης, αλλά συνδέεται και με τη συνολική απόδοσή του σε σχέση με τον δείκτη αναφοράς. Παρότι στην παθητική διαχείριση ο κύριος στόχος είναι η όσο το δυνατόν πιστότερη παρακολούθηση του benchmark, η ανάλυση της απόδοσης παραμένει ουσιώδης, διότι επιτρέπει να αποτιμηθεί το αν και σε ποιο βαθμό το χαρτοφυλάκιο ακολουθεί τη γενική κατεύθυνση της αγοράς, αν υπολείπεται συστηματικά του δείκτη ή αν εμφανίζει αποκλίσεις που έχουν ιδιαίτερη επενδυτική σημασία (Roll, 1992; Frino et al., 2005; Ammann & Zimmermann, 2001).

Η απόδοση του χαρτοφυλακίου σε σχέση με τον δείκτη εκφράζει ουσιαστικά τη συγκριτική επενδυτική συμπεριφορά των δύο μεγεθών μέσα στον χρόνο. Σε κάθε χρονική περίοδο, μπορεί να εξεταστεί αν το χαρτοφυλάκιο παρήγαγε απόδοση ίση, χαμηλότερη ή υψηλότερη από εκείνη του δείκτη. Η πληροφορία αυτή είναι ιδιαίτερα χρήσιμη, καθώς επιτρέπει να εντοπιστεί όχι μόνο η μέση απόκλιση, αλλά και η κατεύθυνση της απόκλισης. Ένα χαρτοφυλάκιο μπορεί, για παράδειγμα, να εμφανίζει μικρό tracking error κατά μέσο όρο, αλλά να υπολείπεται συστηματικά του δείκτη σε κρίσιμες χρονικές περιόδους, γεγονός που έχει διαφορετική ερμηνεία από ένα χαρτοφυλάκιο που αποκλίνει συμμετρικά γύρω από την πορεία του benchmark (Roll, 1992; Rudolf et al., 1999).

Η σχετική απόδοση του χαρτοφυλακίου έναντι του δείκτη έχει ιδιαίτερη σημασία επειδή συνδέεται με την οικονομική ερμηνεία της επιτυχίας της στρατηγικής. Σε ένα χαρτοφυλάκιο αποτύπωσης δείκτη δεν είναι απαραίτητο το ζητούμενο να είναι η υπεραπόδοση σε σχέση με τον benchmark, όπως συμβαίνει στην ενεργητική διαχείριση. Αντίθετα, η επιτυχία συνδέεται κυρίως με την εγγύτητα των αποδόσεων. Παρ' όλα αυτά, η ύπαρξη συστηματικής υποαπόδοσης μπορεί να θεωρηθεί ένδειξη ότι η μέθοδος επιλογής τίτλων, η διαμόρφωση των βαρών ή οι επιβαρύνσεις από το κόστος συναλλαγών επηρεάζουν αρνητικά τη δυνατότητα του χαρτοφυλακίου να ανταποκρίνεται στον βασικό του στόχο (Roll, 1992; Frino et al., 2005).

Η σύγκριση της απόδοσης χαρτοφυλακίου και δείκτη μπορεί να εξεταστεί τόσο σε επίπεδο επιμέρους χρονικών περιόδων όσο και σε σωρευτική βάση. Σε επίπεδο επιμέρους περιόδων, η σύγκριση επιτρέπει την παρακολούθηση της βραχυχρόνιας συμπεριφοράς του χαρτοφυλακίου και την ανίχνευση χρονικών διαστημάτων στα οποία η απόκλιση είναι εντονότερη. Σε σωρευτική βάση, εξετάζεται η συνολική πορεία των αποδόσεων στο εξεταζόμενο χρονικό διάστημα, γεγονός που παρέχει πιο ολοκληρωμένη εικόνα της διαχρονικής σχέσης μεταξύ χαρτοφυλακίου και δείκτη. Η σωρευτική θεώρηση είναι ιδιαίτερα σημαντική, διότι μικρές περιοδικές αποκλίσεις μπορεί να συσσωρεύονται με την πάροδο του χρόνου και να οδηγούν σε ουσιαστικά διαφορετική τελική επίδοση (Frino et al., 2005; Roll, 1992).

Ένα χαρτοφυλάκιο αποτύπωσης δείκτη κρίνεται συνήθως επιτυχημένο όταν όχι μόνο παρουσιάζει μικρό tracking error, αλλά και όταν η σωρευτική του απόδοση παραμένει πολύ κοντά στη σωρευτική απόδοση του δείκτη. Αυτό συμβαίνει επειδή ο επενδυτής που επιλέγει στρατηγική παθητικής διαχείρισης προσδοκά ουσιαστικά να συμμετέχει στη γενική πορεία της αγοράς όπως αυτή εκφράζεται από τον benchmark. Αν το χαρτοφυλάκιο υστερεί σημαντικά σε σωρευτικούς όρους, τότε ακόμη και μικρές περιοδικές αποκλίσεις αποκτούν μεγαλύτερη βαρύτητα, καθώς μεταφράζονται σε πραγματική απώλεια απόδοσης σε σχέση με τον στόχο (Frino et al., 2005; Roll, 1992).

Η ανάλυση της σχετικής απόδοσης είναι επίσης χρήσιμη για την κατανόηση της επίδρασης παραγόντων όπως το κόστος συναλλαγών, η συχνότητα αναπροσαρμογής και οι περιορισμοί πληθικότητας. Για παράδειγμα, ένα χαρτοφυλάκιο μπορεί θεωρητικά να προσεγγίζει καλά τον δείκτη σε όρους δομής, αλλά η επιβάρυνση από συχνές συναλλαγές να οδηγεί τελικά σε ελαφρώς χαμηλότερη καθαρή απόδοση. Αντίστοιχα, ένας αυστηρός περιορισμός στον αριθμό των τίτλων μπορεί να μειώνει τη δυνατότητα πλήρους αναπαραγωγής της πορείας του δείκτη και να οδηγεί σε ελαφρά αλλά συστηματική υστέρηση. Επομένως, η εξέταση της απόδοσης σε σχέση με τον δείκτη λειτουργεί συμπληρωματικά προς το tracking error και επιτρέπει ουσιαστικότερη ερμηνεία των αποτελεσμάτων (Frino et al., 2005; Ammann & Zimmermann, 2001).

Επιπλέον, η σχετική απόδοση μπορεί να εξεταστεί όχι μόνο ως διαφορά αποδόσεων, αλλά και ως λόγος ή ποσοστό συμμετοχής στη συνολική κίνηση του δείκτη. Μια τέτοια προσέγγιση είναι χρήσιμη όταν ενδιαφέρει να αξιολογηθεί ο βαθμός στον οποίο το χαρτοφυλάκιο «συλλαμβάνει» την ανοδική ή πτωτική πορεία της αγοράς. Παρότι σε μια καθαρά παθητική στρατηγική η ιδανική κατάσταση είναι η όσο το δυνατόν πλησιέστερη ταύτιση, η ανάλυση αυτή μπορεί να δώσει πρόσθετη πληροφόρηση για το αν το χαρτοφυλάκιο εμφανίζει μεγαλύτερη ευαισθησία σε συγκεκριμένες φάσεις της αγοράς ή αν η συμπεριφορά του αποκλίνει συστηματικά σε περιόδους αυξημένης μεταβλητότητας (Ammann & Zimmermann, 2001; Roll, 1992).

Από μεθοδολογική άποψη, η εξέταση της απόδοσης του χαρτοφυλακίου έναντι του δείκτη είναι σημαντική και επειδή επιτρέπει τη συμπληρωματική αξιολόγηση της λύσης πέρα από τη συνάρτηση στόχου. Σε πολλά προβλήματα βελτιστοποίησης index tracking, η αντικειμενική συνάρτηση διαμορφώνεται αποκλειστικά με βάση κάποιο μέτρο tracking error. Ωστόσο, κατά την εμπειρική αξιολόγηση των αποτελεσμάτων, είναι συχνά αναγκαίο να εξετάζονται και πρόσθετοι δείκτες, ώστε να διασφαλίζεται ότι η μείωση του tracking error δεν συνοδεύεται από ανεπιθύμητη συμπεριφορά σε επίπεδο σωρευτικής ή μέσης απόδοσης. Με άλλα λόγια, η σχετική απόδοση λειτουργεί ως ουσιαστικό συμπληρωματικό κριτήριο ποιότητας των παραγόμενων χαρτοφυλακίων (Rudolf et al., 1999; Roll, 1992; Frino et al., 2005).

2.5.3 Αξιολόγηση εντός και εκτός δείγματος

Η αξιολόγηση ενός χαρτοφυλακίου αποτύπωσης δείκτη δεν είναι επαρκές να βασίζεται αποκλειστικά στην επίδοσή του πάνω στα δεδομένα που χρησιμοποιήθηκαν για την κατασκευή του. Αντιθέτως, για να εξαχθούν ουσιαστικά και αξιόπιστα συμπεράσματα σχετικά με την αποτελεσματικότητα μιας μεθόδου, είναι αναγκαίο να διακρίνεται η αξιολόγηση σε δύο βασικές κατηγορίες: την αξιολόγηση εντός δείγματος και την αξιολόγηση εκτός δείγματος. Η διάκριση αυτή αποτελεί θεμελιώδη μεθοδολογική αρχή τόσο στη χρηματοοικονομική μοντελοποίηση όσο και στη γενικότερη υπολογιστική βελτιστοποίηση, διότι επιτρέπει να εξεταστεί όχι μόνο το κατά πόσο ένα μοντέλο προσαρμόζεται καλά στα διαθέσιμα ιστορικά δεδομένα, αλλά και το κατά πόσο μπορεί να διατηρήσει την αποτελεσματικότητά του σε νέα, μη χρησιμοποιημένα δεδομένα (Takeda et al., 2013; Yuen et al., 2022; Bailey et al., 2014).

Η αξιολόγηση εντός δείγματος αναφέρεται στην αποτίμηση της επίδοσης του χαρτοφυλακίου πάνω στο ίδιο σύνολο δεδομένων που χρησιμοποιήθηκε για την εκτίμηση των παραμέτρων του μοντέλου ή για τη βελτιστοποίηση της λύσης. Στο πλαίσιο του προβλήματος της αποτύπωσης δείκτη, αυτό σημαίνει ότι οι αποδόσεις των τίτλων και του δείκτη σε μια συγκεκριμένη ιστορική περίοδο χρησιμοποιούνται τόσο για την κατασκευή του χαρτοφυλακίου όσο και για τη μέτρηση του tracking error ή άλλων δεικτών απόδοσης. Η αξιολόγηση αυτή είναι ιδιαίτερα χρήσιμη, διότι δείχνει κατά πόσο η προτεινόμενη μέθοδος μπορεί να εντοπίσει, μέσα στα διαθέσιμα δεδομένα, μια λύση που να προσεγγίζει αποτελεσματικά τη συμπεριφορά του δείκτη (Chavez-Bedoya & Birge, 2014; Yuen et al., 2022).

Παρά τη χρησιμότητά της, η εντός δείγματος αξιολόγηση δεν αρκεί από μόνη της για να κριθεί η πραγματική αξία μιας μεθόδου. Ο βασικός λόγος είναι ότι ένα μοντέλο ή ένας αλγόριθμος μπορεί να προσαρμοστεί υπερβολικά καλά στα συγκεκριμένα ιστορικά δεδομένα που χρησιμοποιήθηκαν κατά την εκπαίδευση ή τη βελτιστοποίηση, χωρίς όμως να διατηρεί την ίδια αποτελεσματικότητα σε νέα δεδομένα. Το φαινόμενο αυτό είναι γνωστό ως υπερπροσαρμογή και αποτελεί σημαντικό κίνδυνο σε προβλήματα χρηματοοικονομικής βελτιστοποίησης, ιδίως όταν χρησιμοποιούνται ισχυρές και ευέλικτες υπολογιστικές τεχνικές, όπως οι εξελικτικοί αλγόριθμοι. Ένα χαρτοφυλάκιο που εμφανίζει εξαιρετική εντός δείγματος επίδοση μπορεί στην πραγματικότητα να «μαθαίνει» ιδιαιτερότητες ή θόρυβο της συγκεκριμένης περιόδου και όχι σταθερά χαρακτηριστικά της σχέσης μεταξύ τίτλων και δείκτη (Bailey et al., 2014; Takeda et al., 2013).

Για τον λόγο αυτό, είναι απαραίτητη και η αξιολόγηση εκτός δείγματος. Η εκτός δείγματος αξιολόγηση αναφέρεται στην αποτίμηση του χαρτοφυλακίου σε μια χρονική περίοδο διαφορετική από εκείνη που χρησιμοποιήθηκε για την κατασκευή του. Συνήθως, ένα πρώτο τμήμα των διαθέσιμων δεδομένων χρησιμοποιείται ως περίοδος εκτίμησης ή εκπαίδευσης, ενώ ένα μεταγενέστερο τμήμα διατηρείται για έλεγχο της επίδοσης του προτεινόμενου χαρτοφυλακίου. Με αυτόν τον τρόπο εξετάζεται κατά πόσο το χαρτοφυλάκιο που προέκυψε από ιστορικά δεδομένα είναι ικανό να διατηρήσει μικρό σφάλμα αποτύπωσης και ικανοποιητική σχετική απόδοση και σε μελλοντικές συνθήκες αγοράς (Takeda et al., 2013; Chavez-Bedoya & Birge, 2014; Yuen et al., 2022).

Η εκτός δείγματος αξιολόγηση είναι ιδιαίτερα σημαντική στο index tracking, επειδή το πραγματικό επενδυτικό ενδιαφέρον αφορά τελικά τη μελλοντική και όχι την παρελθούσα συμπεριφορά του χαρτοφυλακίου. Η χρήση ιστορικών δεδομένων στη βελτιστοποίηση βασίζεται στην παραδοχή ότι το παρελθόν περιέχει πληροφορία χρήσιμη για τη διαμόρφωση μελλοντικών αποφάσεων. Ωστόσο, η παραδοχή αυτή δεν είναι πάντοτε απόλυτα έγκυρη, καθώς οι σχέσεις μεταξύ τίτλων και δείκτη μπορεί να μεταβάλλονται με την πάροδο του χρόνου. Επομένως, η εκτός δείγματος αξιολόγηση λειτουργεί ως βασικός έλεγχος της σταθερότητας και της γενικευσιμότητας της προτεινόμενης μεθόδου (Takeda et al., 2013; Yuen et al., 2022).

Η διάκριση μεταξύ εντός και εκτός δείγματος αξιολόγησης επιτρέπει επίσης να εντοπιστεί η διαφορά ανάμεσα στην ικανότητα προσαρμογής και στην ικανότητα γενίκευσης ενός μοντέλου. Η εντός δείγματος επίδοση δείχνει πόσο καλά μπορεί η μέθοδος να «εξηγήσει» ή να προσεγγίσει τα γνωστά δεδομένα. Η εκτός δείγματος επίδοση, αντίθετα, δείχνει πόσο καλά η ίδια μέθοδος μπορεί να λειτουργήσει σε συνθήκες που δεν έχουν χρησιμοποιηθεί κατά τη φάση βελτιστοποίησης. Σε ένα επιτυχημένο μοντέλο, οι δύο αυτές διαστάσεις πρέπει να βρίσκονται σε σχετική ισορροπία: δεν αρκεί δηλαδή το χαρτοφυλάκιο να επιτυγχάνει εξαιρετικά μικρό tracking error εντός δείγματος, αν αυτό συνοδεύεται από αισθητή επιδείνωση εκτός δείγματος (Bailey et al., 2014; Takeda et al., 2013).

Στην πράξη, η αξιολόγηση εκτός δείγματος μπορεί να οργανωθεί με διαφορετικούς τρόπους. Η απλούστερη προσέγγιση είναι ο διαχωρισμός του διαθέσιμου χρονικού διαστήματος σε δύο τμήματα: ένα τμήμα εκπαίδευσης και ένα τμήμα ελέγχου. Πιο σύνθετες προσεγγίσεις περιλαμβάνουν διαδοχικές ή κυλιόμενες επανεκτιμήσεις, όπου το χαρτοφυλάκιο επαναπροσδιορίζεται σε τακτά χρονικά διαστήματα και κάθε φορά αξιολογείται σε επόμενες περιόδους. Οι κυλιόμενες αυτές διαδικασίες είναι συχνά πιο ρεαλιστικές από επενδυτική άποψη, καθώς αντανακλούν καλύτερα τη διαχρονική πρακτική διαχείρισης χαρτοφυλακίου και τη σταδιακή ενσωμάτωση νέων πληροφοριών (Yuen et al., 2022; Chavez-Bedoya & Birge, 2014).

Η μεθοδολογική σημασία της εντός και εκτός δείγματος αξιολόγησης είναι ακόμη μεγαλύτερη όταν πρόκειται να συγκριθούν διαφορετικές μέθοδοι ή παραμετροποιήσεις. Αν μια μέθοδος αξιολογηθεί μόνο εντός δείγματος, υπάρχει ο κίνδυνος να φαίνεται τεχνητά ανώτερη επειδή έχει απλώς προσαρμοστεί καλύτερα στα συγκεκριμένα ιστορικά δεδομένα. Αντίθετα, η εκτός δείγματος αξιολόγηση προσφέρει ισχυρότερο κριτήριο σύγκρισης, καθώς εξετάζει ποια μέθοδος διατηρεί καλύτερα την αποτελεσματικότητά της σε νέα δεδομένα. Επομένως, για μια αξιόπιστη συγκριτική αποτίμηση είναι απαραίτητο να λαμβάνονται υπόψη και οι δύο μορφές αξιολόγησης (Takeda et al., 2013; Yuen et al., 2022; Chavez-Bedoya & Birge, 2014).

2.6 Περιορισμοί στο πρόβλημα βελτιστοποίησης

2.6.1 Περιορισμός προϋπολογισμού

Ο περιορισμός προϋπολογισμού αποτελεί έναν από τους πιο βασικούς και αναγκαίους περιορισμούς σε κάθε πρόβλημα βελτιστοποίησης χαρτοφυλακίου και, κατ' επέκταση, και στο πρόβλημα της αποτύπωσης δείκτη (Markowitz, 1952; Beasley et al., 2003; Canakgoz & Beasley, 2009). Η ουσία του περιορισμού αυτού είναι ότι το χαρτοφυλάκιο πρέπει να διαμορφωθεί μέσα στα όρια των διαθέσιμων επενδυτικών πόρων. Με άλλα λόγια, η κατανομή κεφαλαίου στους επιλεγμένους τίτλους δεν μπορεί να γίνεται αυθαίρετα, αλλά οφείλει να σέβεται το συνολικό ποσό που είναι διαθέσιμο για επένδυση (Markowitz, 1952; Canakgoz & Beasley, 2009).

Σε θεωρητικό επίπεδο, ο περιορισμός προϋπολογισμού εκφράζει την αρχή ότι το άθροισμα των επενδυμένων ποσών ή, ισοδύναμα, το άθροισμα των βαρών των τίτλων στο χαρτοφυλάκιο πρέπει να ισούται με το διαθέσιμο κεφάλαιο (Markowitz, 1952). Στις περισσότερες διατυπώσεις του προβλήματος, όταν τα βάρη εκφράζονται ως ποσοστά του συνολικού χαρτοφυλακίου, ο περιορισμός αυτός λαμβάνει τη μορφή ότι το άθροισμα των βαρών πρέπει να είναι ίσο με τη μονάδα (Markowitz, 1952; Beasley et al., 2003). Η διατύπωση αυτή σημαίνει ότι όλο το διαθέσιμο κεφάλαιο κατανέμεται μεταξύ των επιλεγμένων τίτλων και ότι δεν παραμένει αχρησιμοποίητο υπόλοιπο, εκτός αν αυτό επιτρέπεται ρητά από το μοντέλο (Canakgoz & Beasley, 2009).

Η σημασία του περιορισμού προϋπολογισμού είναι θεμελιώδης, διότι εξασφαλίζει τη χρηματοοικονομική συνοχή του χαρτοφυλακίου (Markowitz, 1952; Schaerf, 2002). Χωρίς την ύπαρξή του, το μοντέλο θα μπορούσε να οδηγηθεί σε μη ρεαλιστικές λύσεις, όπου τα συνολικά επενδυμένα κεφάλαια υπερβαίνουν το διαθέσιμο ποσό ή, αντίθετα, όπου μέρος του κεφαλαίου παραμένει αδιάθετο χωρίς αυτό να είναι συμβατό με τη λογική της στρατηγικής (Schaerf, 2002; Canakgoz & Beasley, 2009). Επομένως, ο περιορισμός αυτός δεν αποτελεί απλώς μια τεχνική μαθηματική συνθήκη, αλλά τη βασική έκφραση της οικονομικής πραγματικότητας μέσα στην οποία λειτουργεί κάθε επενδυτική απόφαση (Markowitz, 1952).

Στο πρόβλημα της αποτύπωσης δείκτη, ο περιορισμός προϋπολογισμού αποκτά ακόμη μεγαλύτερη σημασία, επειδή το ζητούμενο δεν είναι μόνο να επιλεγεί ένα κατάλληλο σύνολο τίτλων, αλλά και να κατανεμηθεί σωστά το διαθέσιμο κεφάλαιο μεταξύ τους, έτσι ώστε το χαρτοφυλάκιο να παρακολουθεί αποτελεσματικά τον δείκτη (Beasley et al., 2003; Canakgoz & Beasley, 2009). Η ύπαρξη του περιορισμού σημαίνει ότι οποιαδήποτε αύξηση του βάρους ενός τίτλου πρέπει αναγκαστικά να συνοδεύεται από αντίστοιχη μείωση του βάρους κάποιου άλλου τίτλου, εφόσον το συνολικό κεφάλαιο παραμένει σταθερό. Κατά συνέπεια, ο προσδιορισμός των βαρών είναι κατ' ουσίαν μια διαδικασία ανακατανομής περιορισμένων πόρων (Markowitz, 1952; Beasley et al., 2003).

Από μαθηματική άποψη, ο περιορισμός προϋπολογισμού δημιουργεί μια βασική σχέση αλληλεξάρτησης μεταξύ όλων των μεταβλητών βάρους του χαρτοφυλακίου. Τα βάρη δεν μπορούν να επιλεγούν ανεξάρτητα το ένα από το άλλο, αλλά πρέπει να προσδιορίζονται ταυτόχρονα έτσι ώστε να ικανοποιείται η συνολική συνθήκη ισότητας (Markowitz, 1952; Canakgoz & Beasley, 2009). Αυτό το χαρακτηριστικό ενισχύει τη δομική συνοχή του προβλήματος και καθιστά σαφές ότι η κατασκευή του χαρτοφυλακίου είναι ζήτημα

συνδυασμένης απόφασης και όχι μεμονωμένων επιλογών για κάθε τίτλο ξεχωριστά (Markowitz, 1952).

Σε πολλές εφαρμογές, ο περιορισμός προϋπολογισμού συνδέεται και με την υπόθεση ότι δεν επιτρέπονται αρνητικά βάρη, δηλαδή δεν επιτρέπονται ανοικτές πωλήσεις. Στην περίπτωση αυτή, όλα τα βάρη του χαρτοφυλακίου είναι μη αρνητικά και αθροίζονται στη μονάδα (Markowitz, 1952; Beasley et al., 2003). Η συνδυασμένη παρουσία των δύο αυτών συνθηκών εξασφαλίζει ότι το χαρτοφυλάκιο αντιστοιχεί σε πραγματική κατανομή του διαθέσιμου κεφαλαίου μεταξύ τίτλων που έχουν πράγματι αποκτηθεί. Η παραδοχή αυτή είναι ιδιαίτερα συμβατή με τη λογική της παθητικής διαχείρισης και της αποτύπωσης δείκτη, όπου συνήθως το ενδιαφέρον εστιάζεται σε long-only χαρτοφυλάκια (Beasley et al., 2003; Canakgoz & Beasley, 2009).

Ωστόσο, σε πιο γενικές ή θεωρητικές διατυπώσεις του προβλήματος, είναι δυνατόν να εξεταστούν και περιπτώσεις στις οποίες μέρος του κεφαλαίου παραμένει σε ρευστή μορφή ή επιτρέπονται ειδικές επενδυτικές θέσεις (Markowitz, 1952; Schaerf, 2002). Παρά ταύτα, για τους σκοπούς μιας ρεαλιστικής στρατηγικής αποτύπωσης δείκτη, η τυπική προσέγγιση είναι ότι το σύνολο του διαθέσιμου προϋπολογισμού κατανέμεται στους τίτλους του χαρτοφυλακίου (Beasley et al., 2003; Canakgoz & Beasley, 2009). Αυτό καθιστά τον περιορισμό προϋπολογισμού έναν από τους πιο σταθερούς και αναλλοίωτους περιορισμούς του προβλήματος (Canakgoz & Beasley, 2009).

Η πρακτική σημασία του περιορισμού αυτού συνδέεται επίσης με τη δυνατότητα ελέγχου της έκθεσης του χαρτοφυλακίου στην αγορά. Μέσω του περιορισμού προϋπολογισμού, διασφαλίζεται ότι το χαρτοφυλάκιο παραμένει πλήρως επενδεδυμένο σύμφωνα με τις διαθέσιμες δυνατότητες του επενδυτή και ότι δεν αναλαμβάνει υπέρμετρη ή μη χρηματοδοτούμενη έκθεση. Έτσι, ο περιορισμός αυτός λειτουργεί ως βασικός μηχανισμός πειθαρχίας της διαδικασίας βελτιστοποίησης και αποκλείει λύσεις που μπορεί να είναι μαθηματικά ελκυστικές αλλά οικονομικά ανεφάρμοστες (Markowitz, 1952; Schaerf, 2002; Canakgoz & Beasley, 2009).

2.6.2 Περιορισμός πληθικότητας (cardinality constraints)

Ο περιορισμός πληθικότητας αποτελεί έναν από τους σημαντικότερους και πιο χαρακτηριστικούς περιορισμούς στο πρόβλημα της αποτύπωσης δείκτη, καθώς καθορίζει πόσοι τίτλοι επιτρέπεται να συμμετέχουν στο χαρτοφυλάκιο. Αν και η έννοια του περιορισμού αυτού έχει ήδη αναφερθεί στο πλαίσιο της συζήτησης για τον περιορισμό πλήθους τίτλων, εδώ εξετάζεται ειδικότερα ως μαθηματικός και υπολογιστικός περιορισμός του προβλήματος βελτιστοποίησης. Η σημασία του είναι ιδιαίτερα μεγάλη, διότι αποτελεί το βασικό στοιχείο που μετατρέπει ένα κατά τα άλλα συνεχές πρόβλημα κατανομής βαρών σε ένα μικτό πρόβλημα διακριτής και συνεχούς βελτιστοποίησης (Chang et al., 2000; Ruiz-Torrubiano & Suárez, 2009; Beasley et al., 2003).

Ο περιορισμός πληθικότητας εκφράζει την απαίτηση το χαρτοφυλάκιο να περιλαμβάνει είτε ακριβώς είτε το πολύ έναν προκαθορισμένο αριθμό τίτλων. Στην πράξη, αυτό σημαίνει ότι, από το σύνολο των διαθέσιμων μετοχών που συμμετέχουν στον δείκτη ή στο επενδυτικό σύμπαν, μόνο ένα περιορισμένο υποσύνολο μπορεί να λάβει θετικό βάρος στο χαρτοφυλάκιο. Ο περιορισμός αυτός είναι ιδιαίτερα ρεαλιστικός, καθώς οι διαχειριστές κεφαλαίων σπάνια επιλέγουν να διατηρούν πολύ μεγάλο αριθμό θέσεων, ιδίως όταν

επιδιώκουν αποτελεσματική και οικονομικά αποδοτική διαχείριση (Chang et al., 2000; Beasley et al., 2003).

Η χρηματοοικονομική αιτιολόγηση του περιορισμού πληθικότητας συνδέεται με πολλούς παράγοντες. Πρώτον, η διατήρηση περιορισμένου αριθμού τίτλων μειώνει τη διοικητική και λειτουργική πολυπλοκότητα του χαρτοφυλακίου. Δεύτερον, επιτρέπει καλύτερο έλεγχο των θέσεων και ευκολότερη παρακολούθηση της αγοράς. Τρίτον, μειώνει το κόστος συναλλαγών και τις ανάγκες συνεχούς αναπροσαρμογής, ιδίως όταν οι μικρές θέσεις σε πολλούς τίτλους δεν προσφέρουν ουσιαστικό όφελος στην ποιότητα της αποτύπωσης. Επομένως, ο περιορισμός πληθικότητας δεν αποτελεί αυθαίρετο μαθηματικό περιορισμό, αλλά αντανακλά πραγματικές ανάγκες της επενδυτικής πρακτικής (Chang et al., 2000; Beasley et al., 2003).

Από την άλλη πλευρά, ο περιορισμός αυτός επιβάλλει έναν σημαντικό συμβιβασμό. Όσο μικρότερος είναι ο επιτρεπτός αριθμός τίτλων, τόσο πιο συμπαγές και διαχειρίσιμο γίνεται το χαρτοφυλάκιο, αλλά τόσο πιο δύσκολο είναι να προσεγγιστεί με ακρίβεια η συμπεριφορά του δείκτη. Αντίθετα, η αύξηση του αριθμού των επιτρεπτών τίτλων διευρύνει τις δυνατότητες αποτύπωσης, αλλά ταυτόχρονα αυξάνει την πολυπλοκότητα, το κόστος και τις ανάγκες διαχείρισης. Έτσι, ο περιορισμός πληθικότητας αποτελεί σαφή έκφραση της ισορροπίας ανάμεσα στην ποιότητα παρακολούθησης και στην πρακτική βιωσιμότητα του χαρτοφυλακίου (Beasley et al., 2003; Ruiz-Torrubiano & Suárez, 2009).

Η μαθηματική σημασία του περιορισμού πληθικότητας είναι ακόμη μεγαλύτερη. Σε ένα πρόβλημα όπου ζητείται απλώς ο προσδιορισμός βαρών για ένα προκαθορισμένο σύνολο τίτλων, οι μεταβλητές απόφασης είναι συνεχείς και το πρόβλημα, τουλάχιστον σε ορισμένες περιπτώσεις, μπορεί να αντιμετωπιστεί με τεχνικές συνεχούς βελτιστοποίησης. Όταν όμως εισάγεται ο περιορισμός πληθικότητας, απαιτείται να αποφασιστεί ταυτόχρονα ποιοι τίτλοι θα έχουν μη μηδενικό βάρος και ποιοι θα αποκλειστούν. Η απόφαση αυτή έχει διακριτό χαρακτήρα και συνήθως μοντελοποιείται με δυαδικές μεταβλητές, οι οποίες λαμβάνουν την τιμή 1 όταν ένας τίτλος συμμετέχει στο χαρτοφυλάκιο και 0 όταν αποκλείεται (Chang et al., 2000; Beasley et al., 2003; Ruiz-Torrubiano & Suárez, 2009).

Η εισαγωγή δυαδικών μεταβλητών αυξάνει δραστικά τη δυσκολία του προβλήματος. Ο αριθμός των δυνατών υποσυνόλων τίτλων αυξάνεται εκθετικά με το πλήθος των διαθέσιμων μετοχών, γεγονός που καθιστά αδύνατη στην πράξη την εξαντλητική αναζήτηση της βέλτιστης λύσης για μεγάλα προβλήματα. Αυτό είναι ιδιαίτερα κρίσιμο στο index tracking, όπου συχνά οι διαθέσιμοι τίτλοι μπορεί να είναι δεκάδες ή και εκατοντάδες. Συνεπώς, ο περιορισμός πληθικότητας είναι ένας από τους βασικούς λόγους για τους οποίους το πρόβλημα θεωρείται υπολογιστικά δύσκολο και απαιτεί τη χρήση ειδικών τεχνικών αντιμετώπισης (Ruiz-Torrubiano & Suárez, 2009; García et al., 2018).

Η επίδραση του περιορισμού πληθικότητας δεν είναι μόνο ποσοτική αλλά και ποιοτική. Δεν καθορίζει απλώς πόσοι τίτλοι θα περιλαμβάνονται στο χαρτοφυλάκιο, αλλά επηρεάζει έμμεσα και το είδος των λύσεων που θεωρούνται εφικτές. Όταν, για παράδειγμα, το μέγιστο επιτρεπτό πλήθος τίτλων είναι πολύ μικρό, το μοντέλο αναγκάζεται να επιλέξει πιο «αντιπροσωπευτικούς» τίτλους και να τους αποδώσει βάρη που θα καλύψουν όσο το δυνατόν μεγαλύτερο μέρος της συνολικής δυναμικής του δείκτη. Αντίθετα, όταν το επιτρεπτό πλήθος είναι μεγαλύτερο, υπάρχει μεγαλύτερη ευχέρεια λεπτομερέστερης αναπαράστασης του benchmark. Επομένως, ο περιορισμός πληθικότητας επηρεάζει άμεσα τη μορφή και τη λογική της βέλτιστης λύσης (Beasley et al., 2003; García et al., 2018).

Στη βιβλιογραφία της βελτιστοποίησης χαρτοφυλακίου, ο περιορισμός πληθικότητας έχει αναγνωριστεί ως ένας από τους σημαντικότερους παράγοντες που καθιστούν τα κλασικά μοντέλα δύσκολα ή και ανεπαρκή για ρεαλιστικές εφαρμογές. Ενώ η θεωρία μέσου-διασποράς του Markowitz σε απλή μορφή μπορεί να επιλυθεί σχετικά εύκολα όταν δεν υπάρχουν τέτοιοι περιορισμοί, η εισαγωγή περιορισμού πληθικότητας μετατρέπει το πρόβλημα σε πολύ πιο σύνθετο. Για τον λόγο αυτό, έχουν προταθεί πολυάριθμες ευρετικές, μεταευρετικές και εξελικτικές προσεγγίσεις, ακριβώς με σκοπό να αντιμετωπιστεί η δυσκολία που δημιουργεί ο συνδυασμός δυαδικών επιλογών και συνεχών βαρών (Chang et al., 2000; Ruiz-Torrubiano & Suárez, 2009; García et al., 2018).

Στο πρόβλημα της αποτύπωσης δείκτη, ο περιορισμός πληθικότητας είναι ιδιαίτερα συμβατός με την εφαρμογή Γενετικών Αλγορίθμων. Οι αλγόριθμοι αυτοί μπορούν να αναπαραστήσουν φυσικά την επιλογή ή μη των τίτλων μέσα από δομές χρωμοσώματος και να εξερευνούν διαφορετικούς συνδυασμούς συμμετοχής τίτλων, ενώ παράλληλα ρυθμίζουν τα βάρη τους. Με αυτόν τον τρόπο, ο περιορισμός πληθικότητας δεν αντιμετωπίζεται ως απλή μεταγενέστερη διόρθωση, αλλά ως οργανικό μέρος της εξελικτικής διαδικασίας αναζήτησης. Αυτός είναι και ένας από τους βασικούς λόγους για τους οποίους οι Γενετικοί Αλγόριθμοι θεωρούνται ιδιαίτερα κατάλληλοι για το συγκεκριμένο πρόβλημα (Chang et al., 2000; Beasley et al., 2003; García et al., 2018).

2.6.3 Κατώτατα και ανώτατα όρια συμμετοχής τίτλων

Τα κατώτατα και ανώτατα όρια συμμετοχής τίτλων αποτελούν έναν ακόμη ουσιώδη περιορισμό στο πρόβλημα της αποτύπωσης δείκτη, καθώς ρυθμίζουν το εύρος τιμών που μπορούν να λάβουν τα βάρη των επιλεγμένων τίτλων στο χαρτοφυλάκιο. Ο περιορισμός αυτός είναι ιδιαίτερα σημαντικός, επειδή δεν αρκεί ένα χαρτοφυλάκιο να είναι απλώς εφικτό ως προς τον συνολικό προϋπολογισμό και το πλήθος των τίτλων, αλλά πρέπει επιπλέον να διαθέτει εσωτερική δομή που να είναι χρηματοοικονομικά λογική και πρακτικά διαχειρίσιμη. Με άλλα λόγια, ακόμη και όταν έχει αποφασιστεί ποιοι τίτλοι θα συμμετάσχουν στο χαρτοφυλάκιο, είναι αναγκαίο να ελεγχθεί και το πόσο μεγάλη ή πόσο μικρή μπορεί να είναι η συμμετοχή καθενός από αυτούς (Chang et al., 2000; Beasley et al., 2003; Canakgoz & Beasley, 2009).

Τα κατώτατα όρια συμμετοχής καθορίζουν τη μικρότερη επιτρεπτή θέση που μπορεί να λάβει ένας τίτλος, εφόσον περιλαμβάνεται στο χαρτοφυλάκιο. Η ύπαρξή τους είναι ιδιαίτερα χρήσιμη, διότι αποτρέπει την εμφάνιση πολύ μικρών θέσεων που, αν και μαθηματικά μπορεί να συμβάλλουν οριακά στη βελτιστοποίηση, στην πράξη συχνά δεν είναι λειτουργικά ή οικονομικά αποδοτικές. Μια πολύ μικρή επενδυτική θέση μπορεί να επιβαρύνεται δυσανάλογα από κόστη συναλλαγών, να αυξάνει την πολυπλοκότητα της διαχείρισης και να προσφέρει ελάχιστη πραγματική συμβολή στην παρακολούθηση του δείκτη. Επομένως, τα κατώτατα όρια λειτουργούν ως μηχανισμός αποκλεισμού λύσεων που βασίζονται σε υπερβολικά κατακερματισμένες κατανομές κεφαλαίου (Chang et al., 2000; Canakgoz & Beasley, 2009).

Αντίστοιχα, τα ανώτατα όρια συμμετοχής ορίζουν τη μέγιστη επιτρεπτή στάθμιση κάθε τίτλου στο χαρτοφυλάκιο. Η χρηματοοικονομική σημασία τους συνδέεται με την ανάγκη αποφυγής υπερβολικής συγκέντρωσης κεφαλαίου σε έναν μικρό αριθμό μετοχών. Ακόμη και αν ένα συγκεκριμένο χαρτί εμφανίζεται ιδιαίτερα «ελκυστικό» από την άποψη της

ελαχιστοποίησης του tracking error, η υπέρμετρη στάθμισή του μπορεί να οδηγήσει σε χαρτοφυλάκιο με χαμηλή διαφοροποίηση και αυξημένη εξάρτηση από τη συμπεριφορά ενός ή λίγων τίτλων. Έτσι, τα ανώτατα όρια λειτουργούν ως προστατευτικός μηχανισμός έναντι της υπερσυγκέντρωσης και συμβάλλουν στη διατήρηση μιας πιο ισορροπημένης κατανομής (Chang et al., 2000; Ruiz-Torrubiano & Suárez, 2009).

Από θεωρητική άποψη, τα κατώτατα και ανώτατα όρια συμμετοχής μπορούν να ιδωθούν ως περιορισμοί που διαμορφώνουν το επιτρεπτό πεδίο τιμών των μεταβλητών βάρους. Δεν αρκεί δηλαδή ένα βάρος να είναι απλώς μη αρνητικό· πρέπει επιπλέον να βρίσκεται μέσα σε ένα συγκεκριμένο διάστημα, το οποίο ορίζεται από την κατώτερη και την ανώτερη επιτρεπτή τιμή. Η ύπαρξη τέτοιων ορίων προσδίδει μεγαλύτερο ρεαλισμό στο μοντέλο, καθώς αντικατοπτρίζει τις πραγματικές επενδυτικές πρακτικές, όπου οι διαχειριστές κεφαλαίων σπάνια επιτρέπουν είτε αμελητέες είτε υπερβολικά μεγάλες θέσεις (Chang et al., 2000; Canakgoz & Beasley, 2009).

Στο πρόβλημα της αποτύπωσης δείκτη, τα όρια αυτά έχουν ιδιαίτερη σημασία επειδή αλληλεπιδρούν άμεσα με τον περιορισμό πληθικότητας. Όταν, για παράδειγμα, ο αριθμός των επιλεγμένων τίτλων είναι περιορισμένος, η κατανομή των βαρών γίνεται πιο «σφιχτή», καθώς το συνολικό κεφάλαιο πρέπει να διανεμηθεί μεταξύ λιγότερων θέσεων. Σε αυτή την περίπτωση, τα κατώτατα όρια διασφαλίζουν ότι κάθε επιλεγμένος τίτλος έχει ουσιαστική παρουσία στο χαρτοφυλάκιο, ενώ τα ανώτατα όρια αποτρέπουν την υπερβολική συγκέντρωση σε λίγες μετοχές. Επομένως, οι δύο αυτοί τύποι περιορισμών συνδυάζονται με τον cardinality constraint και επηρεάζουν καθοριστικά τη μορφή των εφικτών λύσεων (Chang et al., 2000; Ruiz-Torrubiano & Suárez, 2009; Beasley et al., 2003).

Η πρακτική αξία των κατώτατων ορίων συνδέεται επίσης με τη σταθερότητα της διαχείρισης. Σε ένα χαρτοφυλάκιο χωρίς τέτοιους περιορισμούς, είναι δυνατόν να προκύπτουν λύσεις όπου πολλοί τίτλοι συμμετέχουν με εξαιρετικά μικρά βάρη μόνο και μόνο επειδή βελτιώνουν οριακά μια μαθηματική συνάρτηση στόχου. Όμως, τέτοιες λύσεις είναι συχνά ανεπιθύμητες, γιατί δημιουργούν χαρτοφυλάκια φαινομενικά διαφοροποιημένα αλλά ουσιαστικά δύσχρηστα. Με την επιβολή ενός ελάχιστου επιπέδου συμμετοχής, το μοντέλο «υποχρεώνεται» να επιλέγει τίτλους που έχουν ουσιαστικό και όχι απλώς οριακό ρόλο στη δομή του χαρτοφυλακίου (Chang et al., 2000; Canakgoz & Beasley, 2009).

Αντίστοιχα, τα ανώτατα όρια έχουν έντονη σημασία σε περιπτώσεις όπου η διαδικασία βελτιστοποίησης τείνει να συγκεντρώνει το μεγαλύτερο μέρος του κεφαλαίου σε λίγους τίτλους που φαίνονται ιδιαίτερα αποτελεσματικοί για την παρακολούθηση του δείκτη. Αν και μια τέτοια λύση μπορεί βραχυπρόθεσμα να μειώνει το tracking error, συχνά συνεπάγεται αυξημένο ειδικό κίνδυνο και μειωμένη αντιπροσωπευτικότητα του χαρτοφυλακίου σε σχέση με τον δείκτη. Με άλλα λόγια, ένα χαρτοφυλάκιο που στηρίζεται υπερβολικά σε λίγες μετοχές μπορεί να είναι μαθηματικά αποδεκτό, αλλά να απέχει από τη χρηματοοικονομική λογική της παθητικής παρακολούθησης ενός ευρύτερου benchmark. Τα ανώτατα όρια αποτρέπουν ακριβώς αυτό το ενδεχόμενο (Ruiz-Torrubiano & Suárez, 2009; Chang et al., 2000).

Από υπολογιστική άποψη, η ύπαρξη κατώτατων και ανώτατων ορίων αυξάνει περαιτέρω τη δυσκολία του προβλήματος, ιδίως όταν συνδυάζεται με δυαδικές μεταβλητές συμμετοχής τίτλων. Στην πράξη, όταν ένας τίτλος επιλέγεται να συμμετάσχει στο χαρτοφυλάκιο, το βάρος του πρέπει να βρίσκεται μέσα σε ένα συγκεκριμένο διάστημα, ενώ όταν δεν επιλέγεται, το βάρος του πρέπει να είναι μηδενικό. Αυτή η συνθήκη δημιουργεί

μια ισχυρή σύνδεση ανάμεσα στις διακριτές και τις συνεχείς μεταβλητές του προβλήματος και ενισχύει τη μη γραμμική και μικτή φύση της βελτιστοποίησης. Για τον λόγο αυτό, οι περιορισμοί αυτοί συχνά δυσχεραίνουν την εφαρμογή κλασικών μεθόδων και ενισχύουν την ανάγκη για πιο ευέλικτες αλγοριθμικές προσεγγίσεις (Chang et al., 2000; Canakgoz & Beasley, 2009; Rubio-García et al., 2024).

Οι Γενετικοί Αλγόριθμοι προσφέρουν σημαντικά πλεονεκτήματα στην αντιμετώπιση τέτοιων περιορισμών, καθώς επιτρέπουν την ενσωμάτωσή τους είτε άμεσα στη σχεδίαση του χρωμοσώματος είτε μέσω μηχανισμών επιδιόρθωσης και κανονικοποίησης. Για παράδειγμα, μετά από μια πράξη διασταύρωσης ή μετάλλαξης, μπορεί να απαιτείται αναπροσαρμογή των βαρών ώστε κάθε ενεργός τίτλος να παραμένει μέσα στα επιτρεπτά του όρια και το σύνολο των βαρών να συνεχίζει να ικανοποιεί τον περιορισμό προϋπολογισμού. Έτσι, τα κατώτατα και ανώτατα όρια συμμετοχής δεν αντιμετωπίζονται απλώς ως παθητικές μαθηματικές συνθήκες, αλλά ως ενεργά στοιχεία που επηρεάζουν άμεσα τη συμπεριφορά του αλγορίθμου (Chang et al., 2000; Beasley et al., 2003; García et al., 2018).

2.6.4 Περιορισμοί κόστους συναλλαγών

Οι περιορισμοί κόστους συναλλαγών αποτελούν μία από τις σημαντικότερες προϋποθέσεις για τη ρεαλιστική διατύπωση του προβλήματος της αποτύπωσης δείκτη. Αν και σε πολλές απλουστευμένες θεωρητικές προσεγγίσεις υποτίθεται ότι οι αγορές και οι πωλήσεις τίτλων πραγματοποιούνται χωρίς επιβάρυνση, στην πραγματικότητα κάθε μεταβολή στη σύνθεση ή στα βάρη ενός χαρτοφυλακίου συνεπάγεται συγκεκριμένο οικονομικό κόστος. Για τον λόγο αυτό, δεν αρκεί να λαμβάνεται υπόψη μόνο η ύπαρξη κόστους συναλλαγών ως γενική έννοια, αλλά είναι συχνά αναγκαίο να επιβάλλονται και ρητοί περιορισμοί στο μέγιστο αποδεκτό επίπεδό του (Canakgoz & Beasley, 2009; Mezali & Beasley, 2014; Lobo et al., 2007).

Η λογική των περιορισμών κόστους συναλλαγών είναι ότι το χαρτοφυλάκιο δεν πρέπει να αναδιαμορφώνεται με τρόπο που, παρότι ενδεχομένως βελτιώνει το tracking error, συνεπάγεται υπερβολικά μεγάλη οικονομική επιβάρυνση. Ένα χαρτοφυλάκιο μπορεί θεωρητικά να επιτυγχάνει πολύ μικρό σφάλμα αποτύπωσης αν αναπροσαρμόζεται διαρκώς και επιθετικά, όμως η πρακτική αξία μιας τέτοιας στρατηγικής μειώνεται σημαντικά όταν το κόστος των απαιτούμενων συναλλαγών απορροφά μεγάλο μέρος του διαθέσιμου κεφαλαίου ή της επενδυτικής απόδοσης. Επομένως, οι περιορισμοί αυτοί εκφράζουν την ανάγκη ισορροπίας ανάμεσα στη βελτίωση της ακρίβειας παρακολούθησης και στη διατήρηση της οικονομικής αποδοτικότητας της στρατηγικής (Mezali & Beasley, 2014; Strub & Baumann, 2018).

Στο πρόβλημα της αποτύπωσης δείκτη, οι περιορισμοί κόστους συναλλαγών αποκτούν ιδιαίτερη σημασία τόσο κατά την αρχική συγκρότηση όσο και κατά την αναπροσαρμογή του χαρτοφυλακίου. Κατά την αρχική κατασκευή, το κόστος συνδέεται με τις αγορές των τίτλων που απαιτούνται για τη δημιουργία της επενδυτικής θέσης. Κατά την αναθεώρηση ενός ήδη υφιστάμενου χαρτοφυλακίου, το κόστος συνδέεται με τις αγορές και τις πωλήσεις που απαιτούνται για τη μετάβαση από την τρέχουσα στη νέα επιθυμητή σύνθεση. Στη δεύτερη περίπτωση μάλιστα, η σημασία του περιορισμού είναι ακόμη μεγαλύτερη, διότι η διαδικασία βελτιστοποίησης δεν εξετάζει ένα χαρτοφυλάκιο «από το μηδέν», αλλά μια μετάβαση από μία υπάρχουσα κατάσταση σε μία νέα, με αντίστοιχο οικονομικό τίμημα (Strub & Baumann, 2018; Mezali & Beasley, 2014).

Οι περιορισμοί κόστους συναλλαγών μπορούν να λάβουν διαφορετικές μορφές στη μαθηματική διατύπωση του προβλήματος. Η πιο άμεση προσέγγιση είναι να οριστεί ένα ανώτατο αποδεκτό επίπεδο συνολικού κόστους, το οποίο το χαρτοφυλάκιο δεν επιτρέπεται να υπερβεί. Η τιμή αυτή μπορεί να εκφράζεται είτε ως απόλυτο ποσό είτε ως ποσοστό του διαθέσιμου προϋπολογισμού ή της αξίας του χαρτοφυλακίου. Η μορφή αυτή είναι ιδιαίτερα χρήσιμη όταν ο διαχειριστής επιθυμεί να θέσει εκ των προτέρων ένα σαφές λειτουργικό ή οικονομικό όριο, μέσα στο οποίο πρέπει να κινηθεί η διαδικασία αναπροσαρμογής (Canakgoz & Beasley, 2009; Mezali & Beasley, 2014).

Η ύπαρξη ενός τέτοιου περιορισμού έχει ουσιαστικές συνέπειες στη διαμόρφωση των λύσεων. Από τη στιγμή που το συνολικό κόστος συναλλαγών δεν μπορεί να υπερβεί ένα δεδομένο όριο, η διαδικασία βελτιστοποίησης αναγκάζεται να αναζητήσει λύσεις που είναι όχι μόνο αποτελεσματικές ως προς το tracking error, αλλά και «συγκρατημένες» ως προς το μέγεθος των απαιτούμενων μεταβολών. Αυτό σημαίνει ότι, σε ορισμένες περιπτώσεις, το μοντέλο μπορεί να προτιμήσει μια λύση με ελαφρώς μεγαλύτερη απόκλιση από τον δείκτη, εφόσον η λύση αυτή είναι πολύ φθηνότερη στην εφαρμογή της. Ο περιορισμός κόστους συναλλαγών, επομένως, λειτουργεί ως μηχανισμός οικονομικής πειθαρχίας της διαδικασίας βελτιστοποίησης (Strub & Baumann, 2018; Mezali & Beasley, 2014).

Από χρηματοοικονομική άποψη, οι περιορισμοί αυτοί συνδέονται άμεσα με τη διατήρηση της καθαρής αποδοτικότητας της επενδυτικής στρατηγικής. Ένα χαρτοφυλάκιο αποτύπωσης δείκτη δεν έχει νόημα να παρακολουθεί εξαιρετικά καλά τον benchmark αν το τίμημα αυτής της παρακολούθησης είναι τόσο υψηλό ώστε να μειώνει σημαντικά την καθαρή του απόδοση. Αυτό είναι ιδιαίτερα σημαντικό σε περιβάλλοντα παθητικής διαχείρισης, όπου ένα από τα βασικά πλεονεκτήματα της στρατηγικής είναι ακριβώς το χαμηλότερο κόστος σε σχέση με την ενεργητική διαχείριση. Αν το κόστος συναλλαγών ξεφύγει από λογικά επίπεδα, τότε αναιρείται σε σημαντικό βαθμό ένα από τα κύρια πλεονεκτήματα της παθητικής παρακολούθησης δείκτη (Mezali & Beasley, 2014; Lobo et al., 2007).

Επιπλέον, οι περιορισμοί κόστους συναλλαγών συνδέονται στενά με τη συχνότητα αναπροσαρμογής του χαρτοφυλακίου. Όσο συχνότερα αναθεωρείται το χαρτοφυλάκιο, τόσο μεγαλύτερη είναι η πιθανότητα να αυξάνονται οι συναλλακτικές δαπάνες. Συνεπώς, η επιβολή ανώτατου ορίου στο κόστος λειτουργεί έμμεσα και ως περιορισμός στην ένταση της διαχείρισης. Το μοντέλο εξαναγκάζεται να επιλέγει προσεκτικά ποιες αλλαγές αξίζει πράγματι να πραγματοποιηθούν και ποιες μεταβολές του δείκτη ή των αποδόσεων δεν δικαιολογούν το αντίστοιχο οικονομικό βάρος. Με αυτόν τον τρόπο, ο περιορισμός συμβάλλει και στην αποφυγή υπερβολικά «νευρικών» στρατηγικών αναπροσαρμογής (Strub & Baumann, 2018; Mezali & Beasley, 2014).

Από μαθηματική και υπολογιστική σκοπιά, οι περιορισμοί κόστους συναλλαγών αυξάνουν σημαντικά τη δυσκολία του προβλήματος. Αυτό συμβαίνει επειδή το συνολικό κόστος εξαρτάται από τη διαφορά μεταξύ της τρέχουσας και της νέας σύνθεσης του χαρτοφυλακίου, άρα συνδέεται με τη δυναμική διάσταση του προβλήματος. Δεν αρκεί πλέον να εξεταστούν μόνο τα τελικά βάρη των τίτλων, αλλά απαιτείται να ληφθεί υπόψη και η μετάβαση προς αυτά. Έτσι, οι περιορισμοί αυτοί εισάγουν πρόσθετη αλληλεξάρτηση μεταξύ των μεταβλητών απόφασης και καθιστούν πιο δύσκολη την εύρεση λύσεων που είναι ταυτόχρονα εφικτές και ποιοτικές (Strub & Baumann, 2018; Lobo et al., 2007).

Η ύπαρξη τέτοιων περιορισμών είναι ένας ακόμη λόγος για τον οποίο οι κλασικές μέθοδοι βελτιστοποίησης συχνά δυσκολεύονται να αντιμετωπίσουν ρεαλιστικά προβλήματα αποτύπωσης δείκτη. Αντίθετα, οι ευρετικές και εξελικτικές προσεγγίσεις, όπως οι Γενετικοί Αλγόριθμοι, μπορούν να ενσωματώσουν πιο ευέλικτα τέτοιες απαιτήσεις, είτε μέσω ρητών ελέγχων εφικτότητας είτε μέσω μηχανισμών ποινής και επιδιόρθωσης λύσεων. Στην περίπτωση αυτή, οι περιορισμοί κόστους δεν αποτελούν μόνο εξωτερικές συνθήκες του προβλήματος, αλλά κρίσιμα στοιχεία που πρέπει να επηρεάζουν ενεργά τη διαδικασία αξιολόγησης και επιλογής των υποψηφίων λύσεων (Mezali & Beasley, 2014; Guastaroba et al., 2012).

2.6.5 Περιορισμοί αναπροσαρμογής

Οι περιορισμοί αναπροσαρμογής αποτελούν μια ιδιαίτερα σημαντική κατηγορία περιορισμών στο πρόβλημα της αποτύπωσης δείκτη, καθώς συνδέονται άμεσα με τη δυναμική φύση της διαχείρισης χαρτοφυλακίου. Σε αντίθεση με μια καθαρά στατική θεώρηση, όπου το ενδιαφέρον περιορίζεται στην αρχική κατασκευή ενός χαρτοφυλακίου, στην πράξη η διατήρηση της αποτελεσματικής παρακολούθησης ενός δείκτη απαιτεί συχνά τη σταδιακή προσαρμογή του χαρτοφυλακίου μέσα στον χρόνο. Οι περιορισμοί αναπροσαρμογής αφορούν ακριβώς τα όρια και τις συνθήκες υπό τις οποίες μπορούν να πραγματοποιούνται αυτές οι μεταβολές (Strub & Baumann, 2018; Chiam et al., 2013).

Η ανάγκη εισαγωγής τέτοιων περιορισμών προκύπτει από το γεγονός ότι ένα χαρτοφυλάκιο δεν αναπροσαρμόζεται σε ένα κενό πλαίσιο, αλλά υπό πραγματικούς λειτουργικούς, οικονομικούς και στρατηγικούς περιορισμούς. Θεωρητικά, ένα χαρτοφυλάκιο θα μπορούσε να αναθεωρείται κάθε φορά που προκύπτει έστω και μικρή βελτίωση στο tracking error. Ωστόσο, μια τέτοια προσέγγιση δεν είναι ρεαλιστική, καθώς συνεπάγεται υπερβολικά μεγάλη συχνότητα παρεμβάσεων, αυξημένο κόστος συναλλαγών και πιθανή αστάθεια στη συνολική επενδυτική στρατηγική. Επομένως, είναι συχνά αναγκαίο να τίθενται ρητοί ή έμμεσοι περιορισμοί στον βαθμό, στη συχνότητα ή στο μέγεθος των επιτρεπόμενων αλλαγών (Strub & Baumann, 2018; Mezali & Beasley, 2014).

Ένας βασικός τύπος περιορισμού αναπροσαρμογής αφορά το μέγεθος της μεταβολής των θέσεων του χαρτοφυλακίου. Στην περίπτωση αυτή, δεν επιτρέπεται οι νέες σταθμίσεις των τίτλων να αποκλίνουν υπερβολικά από τις υπάρχουσες, ακόμη και αν μια τέτοια μεταβολή θα βελτίωνε θεωρητικά την αποτύπωση του δείκτη. Η λογική πίσω από αυτή την απαίτηση είναι ότι μεγάλες αλλαγές στη σύνθεση ή στα βάρη του χαρτοφυλακίου συνεπάγονται όχι μόνο υψηλό κόστος συναλλαγών, αλλά και μεγαλύτερη έκθεση σε λειτουργικό κίνδυνο, δυσκολίες εκτέλεσης και απώλεια σταθερότητας της στρατηγικής. Έτσι, ο περιορισμός αυτός λειτουργεί ως μηχανισμός ελέγχου της έντασης της αναπροσαρμογής (Strub & Baumann, 2018).

Ένας δεύτερος τύπος περιορισμού σχετίζεται με τη συχνότητα των αναθεωρήσεων. Στην πράξη, πολλά χαρτοφυλάκια δεν αναπροσαρμόζονται συνεχώς, αλλά μόνο σε προκαθορισμένα χρονικά σημεία ή όταν συμβούν συγκεκριμένα γεγονότα, όπως αναθεώρηση της σύνθεσης του δείκτη ή υπέρβαση κάποιου ορίου απόκλισης. Η επιβολή τέτοιου είδους περιορισμού αποσκοπεί στη διατήρηση μιας πιο πειθαρχημένης στρατηγικής διαχείρισης, αποφεύγοντας υπερβολικά συχνές παρεμβάσεις που μπορεί να βασίζονται σε βραχυχρόνιες και ενδεχομένως μη ουσιαστικές μεταβολές της αγοράς (Frino et al., 2005; Strub & Baumann, 2018).

Οι περιορισμοί αναπροσαρμογής μπορούν επίσης να συνδέονται με τον αριθμό των τίτλων που επιτρέπεται να αλλάξουν μεταξύ δύο διαδοχικών περιόδων. Για παράδειγμα, μπορεί να θεωρείται επιθυμητό ότι μόνο περιορισμένος αριθμός τίτλων μπορεί να εισέλθει ή να εξέλθει από το χαρτοφυλάκιο σε κάθε κύκλο αναθεώρησης. Η προσέγγιση αυτή διατηρεί σε μεγαλύτερο βαθμό τη συνέχεια της επενδυτικής πολιτικής και μειώνει την πιθανότητα να μεταβάλλεται ριζικά η σύνθεση του χαρτοφυλακίου από περίοδο σε περίοδο. Με αυτόν τον τρόπο, το χαρτοφυλάκιο παραμένει πιο σταθερό και πιο εύκολα διαχειρίσιμο (Strub & Baumann, 2018; Chiam et al., 2013).

Η χρηματοοικονομική σημασία των περιορισμών αναπροσαρμογής είναι ιδιαίτερα μεγάλη, διότι σχετίζεται με τη σταθερότητα και τη συνοχή της επενδυτικής στρατηγικής. Ένα χαρτοφυλάκιο που μεταβάλλεται υπερβολικά συχνά ή υπερβολικά έντονα μπορεί να εμφανίζει θεωρητικά βελτιωμένη βραχυπρόθεσμη αποτύπωση, αλλά συχνά χάνει σε αξιοπιστία, προβλεψιμότητα και οικονομική αποδοτικότητα. Αντίθετα, η ύπαρξη περιορισμών επιτρέπει τη διατήρηση μιας πιο συνεκτικής και ελεγχόμενης διαδικασίας προσαρμογής, η οποία είναι περισσότερο συμβατή με τις πρακτικές απαιτήσεις της επαγγελματικής διαχείρισης χαρτοφυλακίου (Frino et al., 2005; Strub & Baumann, 2018).

Παράλληλα, οι περιορισμοί αυτοί αλληλεπιδρούν στενά με το κόστος συναλλαγών. Στην πραγματικότητα, πολλές φορές οι περιορισμοί αναπροσαρμογής λειτουργούν ως έμμεσος τρόπος περιορισμού του κόστους, καθώς ελέγχουν το μέγεθος των αγοραπωλησιών που είναι δυνατόν να πραγματοποιηθούν. Όσο πιο περιορισμένη είναι η δυνατότητα αλλαγής των θέσεων, τόσο μικρότερο είναι συνήθως και το συνολικό συναλλακτικό βάρος. Επομένως, οι περιορισμοί αναπροσαρμογής δεν πρέπει να θεωρούνται ανεξάρτητοι από τους περιορισμούς κόστους, αλλά μάλλον ως συμπληρωματικός μηχανισμός που συμβάλλει στη διατήρηση της οικονομικής βιωσιμότητας της στρατηγικής (Strub & Baumann, 2018; Chiam et al., 2013).

Από μαθηματική άποψη, οι περιορισμοί αναπροσαρμογής εισάγουν μια επιπλέον διαχρονική διάσταση στο πρόβλημα. Δεν αρκεί πλέον να προσδιοριστεί απλώς η βέλτιστη σύνθεση ενός χαρτοφυλακίου σε μια χρονική στιγμή, αλλά πρέπει να ληφθεί υπόψη και η σχέση της νέας λύσης με την προηγούμενη κατάσταση. Αυτό σημαίνει ότι το πρόβλημα γίνεται πιο σύνθετο, καθώς κάθε νέα λύση δεν αξιολογείται μόνο ως προς το tracking error και τους στατικούς περιορισμούς, αλλά και ως προς το κατά πόσο μπορεί να επιτευχθεί από την υφιστάμενη σύνθεση χωρίς να παραβιάζονται τα επιτρεπτά όρια αναπροσαρμογής (Strub & Baumann, 2018).

Η υπολογιστική δυσκολία αυξάνεται ακόμη περισσότερο όταν οι περιορισμοί αυτοί συνδυάζονται με cardinality constraints, όρια βαρών και περιορισμούς κόστους συναλλαγών. Σε αυτή την περίπτωση, η διαδικασία βελτιστοποίησης πρέπει να αναζητήσει όχι απλώς ένα «καλό» χαρτοφυλάκιο, αλλά ένα χαρτοφυλάκιο που να είναι καλύτερο από το τρέχον, να παραμένει κοντά στον δείκτη και ταυτόχρονα να μην απαιτεί υπερβολικά μεγάλη μεταβολή στη σύνθεσή του. Η πολυπλοκότητα αυτή αποτελεί έναν από τους βασικούς λόγους για τους οποίους ευέλικτες μεθοδολογίες, όπως οι Γενετικοί Αλγόριθμοι, είναι ιδιαίτερα κατάλληλες για τέτοια προβλήματα (Chiam et al., 2013; Strub & Baumann, 2018).

2.7 Κλασικές προσεγγίσεις στη βελτιστοποίηση χαρτοφυλακίου

2.7.1 Η θεωρία χαρτοφυλακίου κατά Markowitz

Η θεωρία χαρτοφυλακίου κατά Markowitz αποτελεί το θεμέλιο της σύγχρονης θεωρητικής προσέγγισης στη βελτιστοποίηση χαρτοφυλακίου και έχει διαμορφώσει σε μεγάλο βαθμό τον τρόπο με τον οποίο αντιμετωπίζεται η σχέση μεταξύ απόδοσης και κινδύνου στις επενδυτικές αποφάσεις (Markowitz, 1952). Η βασική συμβολή της θεωρίας αυτής έγκειται στο ότι εισήγαγε ένα συστηματικό και μαθηματικά διατυπωμένο πλαίσιο για την επιλογή χαρτοφυλακίων, μετατοπίζοντας την έμφαση από την αξιολόγηση μεμονωμένων τίτλων στην αξιολόγηση του χαρτοφυλακίου ως συνόλου.

Η κεντρική ιδέα της θεωρίας είναι ότι ο επενδυτής δεν πρέπει να εξετάζει κάθε τίτλο ανεξάρτητα, αλλά να λαμβάνει υπόψη τον τρόπο με τον οποίο οι αποδόσεις των διαφορετικών τίτλων αλληλεπιδρούν μέσα στο χαρτοφυλάκιο. Έτσι, η συνολική συμπεριφορά του χαρτοφυλακίου δεν εξαρτάται μόνο από τις αναμενόμενες αποδόσεις των επιμέρους στοιχείων του, αλλά και από τη μεταξύ τους συνδιακύμανση. Με τον τρόπο αυτό, η θεωρία του Markowitz ανέδειξε τη διαφοροποίηση ως βασικό μηχανισμό μείωσης του κινδύνου, δείχνοντας ότι η κατάλληλη συνδυαστική επιλογή τίτλων μπορεί να οδηγήσει σε χαρτοφυλάκια με ευνοϊκότερο προφίλ κινδύνου–απόδοσης από εκείνο που θα προέκυπτε από την εξέταση κάθε τίτλου ξεχωριστά (Markowitz, 1952).

Στο πλαίσιο αυτό, η απόδοση ενός χαρτοφυλακίου εκφράζεται ως ο σταθμισμένος μέσος των αναμενόμενων αποδόσεων των τίτλων που το απαρτίζουν, ενώ ο κίνδυνος αποτυπώνεται μέσω της διασποράς ή της τυπικής απόκλισης των αποδόσεών του. Το κρίσιμο σημείο είναι ότι η διασπορά του χαρτοφυλακίου δεν εξαρτάται μόνο από τις διασπορές των επιμέρους τίτλων, αλλά και από τις συνδιακυμάνσεις τους. Αυτό σημαίνει ότι ένας τίτλος με υψηλή ατομική μεταβλητότητα δεν είναι κατ' ανάγκη ανεπιθύμητος, αν η συμμετοχή του συμβάλλει στη μείωση της συνολικής μεταβλητότητας του χαρτοφυλακίου μέσω κατάλληλης αλληλεπίδρασης με τους υπόλοιπους τίτλους (Markowitz, 1952).

Ένα από τα πιο γνωστά αποτελέσματα της θεωρίας του Markowitz είναι η έννοια του αποδοτικού συνόρου. Το αποδοτικό σύνορο περιγράφει το σύνολο των χαρτοφυλακίων που, για κάθε δεδομένο επίπεδο κινδύνου, προσφέρουν τη μέγιστη δυνατή αναμενόμενη απόδοση ή, αντίστροφα, για κάθε δεδομένο επίπεδο αναμενόμενης απόδοσης συνεπάγονται τον ελάχιστο δυνατό κίνδυνο. Τα χαρτοφυλάκια που δεν ανήκουν σε αυτό το σύνορο θεωρούνται μη αποδοτικά, καθώς υπάρχει τουλάχιστον ένα άλλο χαρτοφυλάκιο που τα υπερτερεί είτε σε όρους χαμηλότερου κινδύνου είτε σε όρους υψηλότερης απόδοσης. Η έννοια αυτή υπήρξε εξαιρετικά σημαντική, επειδή μετέτρεψε το πρόβλημα της επιλογής χαρτοφυλακίου σε πρόβλημα συστηματικής αναζήτησης βέλτιστων συμβιβασμών (Markowitz, 1952).

Η θεωρία του Markowitz στηρίζεται σε ορισμένες βασικές παραδοχές, οι οποίες επιτρέπουν τη μαθηματική της διατύπωση. Τυπικά υποτίθεται ότι οι επενδυτές είναι ορθολογικοί και επιδιώκουν τη μεγιστοποίηση της χρησιμότητάς τους, προτιμώντας υψηλότερες αποδόσεις και χαμηλότερο κίνδυνο. Επιπλέον, θεωρείται ότι η απόδοση και ο κίνδυνος μπορούν να συνοψιστούν επαρκώς από τη μέση τιμή και τη διασπορά των αποδόσεων αντίστοιχα. Παρότι οι παραδοχές αυτές έχουν δεχθεί κριτική ως προς τον βαθμό ρεαλισμού τους, η

θεωρία εξακολουθεί να αποτελεί το πιο επιδραστικό σημείο εκκίνησης για τη μελέτη της βελτιστοποίησης χαρτοφυλακίου (Markowitz, 1952).

Η σημασία της θεωρίας του Markowitz για τη χρηματοοικονομική επιστήμη είναι ιδιαίτερα μεγάλη, διότι καθιέρωσε ένα ενιαίο αναλυτικό πλαίσιο για την αντιμετώπιση των επενδυτικών αποφάσεων. Για πρώτη φορά, η διαδικασία επιλογής χαρτοφυλακίου δεν βασιζόταν αποκλειστικά σε εμπειρικές κρίσεις ή απλές προτιμήσεις, αλλά σε ρητά ορισμένο μαθηματικό πρόβλημα. Η συνεισφορά αυτή υπήρξε καθοριστική όχι μόνο για την ανάπτυξη της θεωρίας χαρτοφυλακίου, αλλά και για την εξέλιξη μεταγενέστερων προσεγγίσεων, όπως τα υποδείγματα τιμολόγησης περιουσιακών στοιχείων και οι σύγχρονες μέθοδοι χρηματοοικονομικής βελτιστοποίησης (Markowitz, 1952).

Παρά την ιστορική και θεωρητική της σημασία, η θεωρία του Markowitz δεν σχεδιάστηκε ειδικά για το πρόβλημα της αποτύπωσης δείκτη. Ο αρχικός της στόχος ήταν η εύρεση χαρτοφυλακίων που επιτυγχάνουν βέλτιστη σχέση απόδοσης και κινδύνου, όχι η κατασκευή χαρτοφυλακίων που να παρακολουθούν πιστά τη συμπεριφορά ενός δείκτη αναφοράς. Παρ' όλα αυτά, η λογική της μαθηματικής βελτιστοποίησης που εισήγαγε, καθώς και η έμφαση στη δομή του χαρτοφυλακίου ως συνδυασμού τίτλων και βαρών, επηρέασαν έντονα και τη μετέπειτα βιβλιογραφία του index tracking. Πολλές από τις μεθόδους που αναπτύχθηκαν για την αποτύπωση δείκτη μπορούν να θεωρηθούν, σε έναν βαθμό, ως εξειδικευμένες ή τροποποιημένες επεκτάσεις του γενικότερου πλαισίου που θεμελίωσε ο Markowitz (Beasley et al., 2003).

Στο πρόβλημα της αποτύπωσης δείκτη, η κληρονομιά της θεωρίας του Markowitz είναι ορατή κυρίως στην ιδέα ότι το χαρτοφυλάκιο πρέπει να προσδιορίζεται μέσω ρητής βελτιστοποίησης των βαρών των τίτλων υπό συγκεκριμένους στόχους και περιορισμούς. Αντί όμως ο στόχος να είναι η ελαχιστοποίηση της διασποράς για δεδομένη απόδοση ή η μεγιστοποίηση της απόδοσης για δεδομένο κίνδυνο, στο index tracking το ενδιαφέρον μετατοπίζεται στην ελαχιστοποίηση της απόκλισης από τον δείκτη. Έτσι, ενώ ο βασικός στόχος αλλάζει, η λογική του μαθηματικού προσδιορισμού του χαρτοφυλακίου παραμένει στενά συνδεδεμένη με την παράδοση που εγκαινίασε η θεωρία χαρτοφυλακίου κατά Markowitz (Beasley et al., 2003).

Ταυτόχρονα, η εφαρμογή της θεωρίας αυτής σε πραγματικά προβλήματα ανέδειξε και τους περιορισμούς της. Η ανάγκη εισαγωγής πρακτικών περιορισμών, όπως cardinality constraints, όρια συμμετοχής τίτλων, κόστη συναλλαγών και δυναμικές αναπροσαρμογές, οδήγησε σε προβλήματα πολύ πιο σύνθετα από την αρχική mean–variance διατύπωση. Αυτό ακριβώς είναι και το σημείο στο οποίο αρχίζουν να εμφανίζονται οι δυσκολίες των κλασικών προσεγγίσεων και να αναδεικνύεται η ανάγκη για πιο ευέλικτες τεχνικές επίλυσης (Schaerf, 2002).

Συνολικά, η θεωρία χαρτοφυλακίου κατά Markowitz αποτελεί το θεμελιώδες θεωρητικό υπόβαθρο πάνω στο οποίο αναπτύχθηκε η σύγχρονη επιστημονική αντιμετώπιση της βελτιστοποίησης χαρτοφυλακίου. Παρότι δεν σχεδιάστηκε ειδικά για την αποτύπωση δείκτη, εισήγαγε τις βασικές έννοιες της αναμενόμενης απόδοσης, του κινδύνου, της διαφοροποίησης και της μαθηματικής επιλογής χαρτοφυλακίων, οι οποίες εξακολουθούν να διαδραματίζουν καθοριστικό ρόλο και στο index tracking. Στην επόμενη υποενότητα εξετάζεται ειδικότερα η προσέγγιση μέσου–διασποράς, η οποία αποτελεί την πιο χαρακτηριστική μαθηματική έκφραση της θεωρίας του Markowitz (Markowitz, 1952; Beasley et al., 2003).

2.7.2 Η προσέγγιση μέσου–διασποράς

Η προσέγγιση μέσου–διασποράς αποτελεί την πιο χαρακτηριστική και ευρέως γνωστή μαθηματική έκφραση της θεωρίας χαρτοφυλακίου κατά Markowitz (Markowitz, 1952). Στον πυρήνα της βρίσκεται η ιδέα ότι η επενδυτική απόφαση μπορεί να περιγραφεί μέσα από δύο βασικά μεγέθη: την αναμενόμενη απόδοση και τον κίνδυνο του χαρτοφυλακίου. Η αναμενόμενη απόδοση αποτυπώνει το προσδοκώμενο οικονομικό όφελος από την επένδυση, ενώ η διασπορά ή η τυπική απόκλιση των αποδόσεων χρησιμοποιείται ως μέτρο του κινδύνου, δηλαδή της αβεβαιότητας που συνοδεύει την πραγματοποίηση αυτής της απόδοσης (Markowitz, 1952).

Η βασική λογική της προσέγγισης είναι ότι ο επενδυτής επιδιώκει είτε να μεγιστοποιήσει την αναμενόμενη απόδοση για δεδομένο επίπεδο κινδύνου είτε να ελαχιστοποιήσει τον κίνδυνο για δεδομένο επίπεδο αναμενόμενης απόδοσης (Markowitz, 1952). Η διατύπωση αυτή αντανάκλα τον θεμελιώδη συμβιβασμό μεταξύ απόδοσης και κινδύνου, ο οποίος βρίσκεται στον πυρήνα σχεδόν κάθε προβλήματος χρηματοοικονομικής λήψης αποφάσεων. Έτσι, η προσέγγιση μέσου–διασποράς προσφέρει ένα σαφές και συστηματικό πλαίσιο μέσα στο οποίο μπορούν να αναζητηθούν χαρτοφυλάκια που θεωρούνται αποδοτικά με βάση αυτόν τον διπλό στόχο (Markowitz, 1952).

Η αναμενόμενη απόδοση του χαρτοφυλακίου εκφράζεται ως ο σταθμισμένος μέσος των αναμενόμενων αποδόσεων των επιμέρους τίτλων. Δηλαδή, κάθε τίτλος συμβάλλει στη συνολική αναμενόμενη απόδοση ανάλογα με το βάρος του στο χαρτοφυλάκιο. Η διασπορά, αντίθετα, έχει πιο σύνθετη μορφή, καθώς δεν εξαρτάται μόνο από τις διασπορές των μεμονωμένων τίτλων, αλλά και από τις συνδιακυμάνσεις μεταξύ των αποδόσεών τους. Αυτό σημαίνει ότι ο κίνδυνος του χαρτοφυλακίου δεν είναι απλό άθροισμα των ατομικών κινδύνων, αλλά αποτέλεσμα της αλληλεπίδρασης μεταξύ των τίτλων που το συγκροτούν (Markowitz, 1952).

Το χαρακτηριστικό αυτό είναι ιδιαίτερα σημαντικό, επειδή αναδεικνύει τον ρόλο της διαφοροποίησης. Η προσέγγιση μέσου–διασποράς δείχνει ότι η κατάλληλη συνδυαστική επιλογή τίτλων μπορεί να μειώσει τη συνολική μεταβλητότητα του χαρτοφυλακίου χωρίς απαραίτητα να μειώσει δραστικά την αναμενόμενη απόδοση. Με τον τρόπο αυτό, η διαφοροποίηση δεν αντιμετωπίζεται απλώς ως εμπειρικός κανόνας επενδυτικής πρακτικής, αλλά ως αποτέλεσμα συγκεκριμένης μαθηματικής δομής. Αυτή ακριβώς η συμβολή υπήρξε ένας από τους κύριους λόγους για τους οποίους η προσέγγιση μέσου–διασποράς επηρέασε βαθιά τη χρηματοοικονομική θεωρία (Markowitz, 1952).

Στη μαθηματική της μορφή, η προσέγγιση μέσου–διασποράς οδηγεί συνήθως σε ένα πρόβλημα τετραγωνικής βελτιστοποίησης. Η αντικειμενική συνάρτηση περιλαμβάνει όρο που σχετίζεται με τη διασπορά του χαρτοφυλακίου, ενώ η αναμενόμενη απόδοση εμφανίζεται είτε ως στόχος είτε ως περιορισμός. Παράλληλα, επιβάλλονται συνήθως περιορισμοί όπως ο περιορισμός προϋπολογισμού και, σε πολλές πρακτικές εφαρμογές, η μη αρνητικότητα των βαρών (Markowitz, 1952; Best & Grauer, 1991). Με αυτή τη μορφή, το πρόβλημα γίνεται σαφές, μαθηματικά δομημένο και επιδεκτικό ανάλυσης, γεγονός που συνέβαλε καθοριστικά στην ευρεία χρήση του στη θεωρία και στην πράξη (Markowitz, 1952).

Η προσέγγιση αυτή υπήρξε εξαιρετικά επιδραστική επειδή προσέφερε έναν σαφή τρόπο ποσοτικοποίησης της επενδυτικής επιλογής. Ο επενδυτής δεν χρειάζεται πλέον να βασίζεται

μόνο σε ποιοτικές κρίσεις ή γενικές προτιμήσεις, αλλά μπορεί να διατυπώσει το πρόβλημα ως αναζήτηση χαρτοφυλακίου που ικανοποιεί ρητά κριτήρια απόδοσης και κινδύνου. Η λογική αυτή εισήγαγε μια νέα εποχή στη χρηματοοικονομική επιστήμη, όπου οι επενδυτικές αποφάσεις μπορούσαν να αντιμετωπιστούν με εργαλεία μαθηματικής ανάλυσης και βελτιστοποίησης (Markowitz, 1952).

Παρά τα πλεονεκτήματά της, η προσέγγιση μέσου–διασποράς βασίζεται σε ορισμένες παραδοχές που περιορίζουν την εφαρμογή της σε πραγματικά προβλήματα. Πρώτον, προϋποθέτει ότι η αναμενόμενη απόδοση και η διασπορά αρκούν για να περιγράψουν ικανοποιητικά την επενδυτική επιλογή, κάτι που δεν είναι πάντοτε επαρκές όταν οι αποδόσεις εμφανίζουν ασυμμετρία ή ακραία γεγονότα (Konno & Yamazaki, 1991). Δεύτερον, στηρίζεται σε εκτιμήσεις των αναμενόμενων αποδόσεων και των συνδιακυμάνσεων, οι οποίες είναι συχνά ασταθείς και ευαίσθητες στα διαθέσιμα δεδομένα (Best & Grauer, 1991; Michaud, 1989). Τρίτον, στην απλή της μορφή δεν ενσωματώνει άμεσα πολλούς από τους πρακτικούς περιορισμούς που εμφανίζονται σε πραγματικά προβλήματα διαχείρισης χαρτοφυλακίου (Schaerf, 2002).

Οι περιορισμοί αυτοί γίνονται ακόμη πιο εμφανείς όταν η προσέγγιση επιχειρεί να εφαρμοστεί στο πρόβλημα της αποτύπωσης δείκτη. Στην περίπτωση αυτή, ο στόχος δεν είναι γενικά η επίτευξη ενός βέλτιστου συνδυασμού αναμενόμενης απόδοσης και κινδύνου, αλλά η όσο το δυνατόν ακριβέστερη παρακολούθηση ενός δείκτη αναφοράς. Αυτό σημαίνει ότι η έννοια του κινδύνου μεταβάλλεται: δεν ενδιαφέρει μόνο η συνολική διασπορά του χαρτοφυλακίου, αλλά κυρίως η απόκλισή του από τον δείκτη. Επομένως, παρότι η προσέγγιση μέσου–διασποράς παρέχει σημαντικό θεωρητικό υπόβαθρο, δεν αρκεί από μόνη της για να περιγράψει πλήρως τις απαιτήσεις του index tracking (Beasley et al., 2003; Rudolf et al., 1999).

Επιπλέον, στην πράξη το πρόβλημα της αποτύπωσης δείκτη συνοδεύεται συχνά από περιορισμούς όπως cardinality constraints, κατώτατα και ανώτατα όρια συμμετοχής τίτλων, κόστος συναλλαγών και περιορισμούς αναπροσαρμογής. Οι περιορισμοί αυτοί υπερβαίνουν το πλαίσιο της απλής mean–variance διατύπωσης και καθιστούν το πρόβλημα σαφώς πιο πολύπλοκο. Έτσι, ενώ η προσέγγιση μέσου–διασποράς παραμένει βασικό θεωρητικό σημείο αναφοράς, οι ρεαλιστικές εκδοχές του προβλήματος απαιτούν σημαντικές επεκτάσεις ή και διαφορετικές μεθοδολογικές επιλογές (Schaerf, 2002; Beasley et al., 2003).

Παρά τις αδυναμίες της, η προσέγγιση μέσου–διασποράς εξακολουθεί να κατέχει κεντρική θέση στη χρηματοοικονομική ανάλυση. Η σημασία της δεν περιορίζεται μόνο στην αρχική της εφαρμογή, αλλά και στο ότι αποτέλεσε τη βάση πάνω στην οποία αναπτύχθηκαν πλήθος μεταγενέστερων μοντέλων και τεχνικών. Ακόμη και όταν ένα πρόβλημα, όπως το index tracking, απαιτεί τροποποίηση του στόχου ή εισαγωγή πρόσθετων περιορισμών, η λογική της ρητής μαθηματικής βελτιστοποίησης του χαρτοφυλακίου παραμένει άμεσα συνδεδεμένη με την παράδοση της mean–variance προσέγγισης (Markowitz, 1952; Schaerf, 2002).

2.7.3 Περιορισμοί των κλασικών προσεγγίσεων στο πρόβλημα της αποτύπωσης δείκτη

Παρά τη θεμελιώδη συμβολή της θεωρίας χαρτοφυλακίου κατά Markowitz και της προσέγγισης μέσου–διασποράς στη διαμόρφωση της σύγχρονης χρηματοοικονομικής ανάλυσης, οι κλασικές αυτές προσεγγίσεις παρουσιάζουν σημαντικούς περιορισμούς όταν επιχειρείται η εφαρμογή τους στο πρόβλημα της αποτύπωσης δείκτη. Ο βασικός λόγος είναι ότι τα κλασικά υποδείγματα αναπτύχθηκαν πρωτίστως για την αντιμετώπιση του γενικού προβλήματος επιλογής χαρτοφυλακίου υπό το πρίσμα της σχέσης απόδοσης–κινδύνου και όχι για την ειδική περίπτωση όπου το ζητούμενο είναι η αναπαραγωγή της πορείας ενός συγκεκριμένου δείκτη αναφοράς (Markowitz, 1952; Beasley et al., 2003).

Ο πρώτος και σημαντικότερος περιορισμός αφορά την ίδια τη στόχευση της βελτιστοποίησης. Στην κλασική θεωρία, ο επενδυτής επιδιώκει τη μεγιστοποίηση της αναμενόμενης απόδοσης για δεδομένο επίπεδο κινδύνου ή την ελαχιστοποίηση της διασποράς για δεδομένη απόδοση. Αντίθετα, στο πρόβλημα της αποτύπωσης δείκτη, το ζητούμενο δεν είναι απλώς η επίτευξη ενός αποδοτικού συνδυασμού απόδοσης και κινδύνου, αλλά η ελαχιστοποίηση της απόκλισης μεταξύ του χαρτοφυλακίου και του δείκτη αναφοράς. Η μετατόπιση αυτή του στόχου σημαίνει ότι ο κίνδυνος δεν νοείται πια μόνο ως συνολική μεταβλητότητα του χαρτοφυλακίου, αλλά κυρίως ως σχετική απόκλιση από ένα benchmark. Επομένως, η κλασική mean–variance διατύπωση δεν επαρκεί για να αποτυπώσει με ακρίβεια τη λογική του index tracking (Markowitz, 1952; Rudolf et al., 1999; Beasley et al., 2003).

Ένας δεύτερος περιορισμός σχετίζεται με τη μορφή των περιορισμών που εμφανίζονται στα ρεαλιστικά προβλήματα αποτύπωσης δείκτη. Τα κλασικά μοντέλα συνήθως εξετάζουν σχετικά απλές διατυπώσεις, όπου οι κύριοι περιορισμοί αφορούν τον συνολικό προϋπολογισμό και, ενδεχομένως, τη μη αρνητικότητα των βαρών. Στην πράξη όμως, τα χαρτοφυλάκια αποτύπωσης δείκτη συνοδεύονται συχνά από πολύ πιο σύνθετους περιορισμούς, όπως περιορισμούς πληθικότητας, κατώτατα και ανώτατα όρια συμμετοχής τίτλων, περιορισμούς κόστους συναλλαγών και περιορισμούς αναπροσαρμογής. Η ύπαρξη αυτών των στοιχείων μεταβάλλει ουσιαστικά τη φύση του προβλήματος και καθιστά ανεπαρκείς τις απλές κλασικές διατυπώσεις (Schaerf, 2002; Beasley et al., 2003).

Ιδιαίτερη σημασία έχει ο περιορισμός πληθικότητας, ο οποίος εισάγει διακριτές μεταβλητές και μετατρέπει το πρόβλημα από συνεχές σε μικτό διακριτό–συνεχές. Η κλασική προσέγγιση του Markowitz δεν έχει σχεδιαστεί για να χειρίζεται άμεσα τέτοιου είδους συνδυαστικούς περιορισμούς. Όταν απαιτείται να επιλεγεί ένας συγκεκριμένος αριθμός τίτλων από ένα μεγαλύτερο σύνολο, το πρόβλημα παύει να είναι απλώς ζήτημα κατανομής βαρών και μετατρέπεται σε πρόβλημα ταυτόχρονης επιλογής και στάθμισης τίτλων. Το χαρακτηριστικό αυτό αυξάνει ραγδαία την υπολογιστική δυσκολία και απομακρύνει το πρόβλημα από το πλαίσιο στο οποίο οι κλασικές μέθοδοι είναι πιο αποτελεσματικές (Schaerf, 2002; Beasley et al., 2003).

Ένας ακόμη περιορισμός των κλασικών προσεγγίσεων αφορά την αδυναμία τους να ενσωματώσουν εύκολα το κόστος συναλλαγών. Στα περισσότερα βασικά θεωρητικά υποδείγματα υποτίθεται σιωπηρά ότι οι συναλλαγές πραγματοποιούνται χωρίς κόστος ή ότι το κόστος αυτό είναι αμελητέο. Ωστόσο, στο index tracking το κόστος συναλλαγών αποτελεί ουσιώδες στοιχείο, ιδίως όταν το χαρτοφυλάκιο πρέπει να αναπροσαρμόζεται συχνά ή όταν ο αριθμός των τίτλων είναι μεγάλος. Η παράλειψη αυτού του παράγοντα

μπορεί να οδηγήσει σε λύσεις θεωρητικά αποδοτικές αλλά πρακτικά μη εφαρμόσιμες ή οικονομικά μη συμφέρουσες (Schaerf, 2002; Beasley et al., 2003).

Επιπλέον, οι κλασικές προσεγγίσεις εμφανίζουν περιορισμούς και ως προς τη δυναμική διάσταση του προβλήματος. Η αρχική θεωρία χαρτοφυλακίου είναι κατεξοχήν στατική, με την έννοια ότι εξετάζει την επιλογή ενός χαρτοφυλακίου για μια δεδομένη στιγμή ή περίοδο χωρίς να ενσωματώνει άμεσα τη συνεχή διαδικασία αναπροσαρμογής. Αντίθετα, στην αποτύπωση δείκτη η ανάγκη περιοδικής επαναστάθμισης είναι κεντρική, καθώς τόσο οι αποδόσεις των τίτλων όσο και η σύνθεση του ίδιου του δείκτη μεταβάλλονται μέσα στον χρόνο. Η αδυναμία των κλασικών μοντέλων να ενσωματώσουν αποτελεσματικά αυτή τη χρονική δυναμική περιορίζει σημαντικά την καταλληλότητά τους για το συγκεκριμένο πρόβλημα (Markowitz, 1952; Beasley et al., 2003).

Σημαντικός είναι επίσης ο περιορισμός που σχετίζεται με την ευαισθησία των κλασικών προσεγγίσεων στις εκτιμήσεις των παραμέτρων. Η mean–variance βελτιστοποίηση βασίζεται σε εκτιμήσεις αναμενόμενων αποδόσεων, διασπορών και συνδιακυμάνσεων, οι οποίες στην πράξη μπορεί να είναι ασταθείς και ευαίσθητες στα δεδομένα που χρησιμοποιούνται. Μικρές μεταβολές στις εκτιμήσεις αυτές είναι δυνατόν να οδηγήσουν σε σημαντικά διαφορετικές λύσεις, γεγονός που δημιουργεί πρόβλημα αξιοπιστίας και σταθερότητας. Το ζήτημα αυτό γίνεται ακόμη πιο έντονο όταν το χαρτοφυλάκιο πρέπει να σχεδιαστεί για να παρακολουθεί έναν δείκτη σε πραγματικές συνθήκες αγοράς, όπου οι στατιστικές σχέσεις μεταξύ τίτλων μεταβάλλονται με την πάροδο του χρόνου (Michaud, 1989; Lai et al., 2011).

Από υπολογιστική άποψη, οι περιορισμοί αυτοί έχουν ως αποτέλεσμα ότι οι κλασικές τεχνικές συχνά δεν μπορούν να αντιμετωπίσουν αποτελεσματικά ρεαλιστικά προβλήματα αποτύπωσης δείκτη, ιδίως όταν το μέγεθος του προβλήματος είναι μεγάλο. Η ύπαρξη πολλών τίτλων, μικτών μεταβλητών, μη γραμμικών περιορισμών και συναλλακτικών δαπανών δημιουργεί ένα χώρο λύσεων εξαιρετικά πολύπλοκο, στον οποίο οι ακριβείς μέθοδοι είτε γίνονται υπολογιστικά δυσβάσταχτες είτε απαιτούν απλουστευτικές παραδοχές που μειώνουν τον ρεαλισμό του μοντέλου. Έτσι, η πρακτική εφαρμογή των κλασικών προσεγγίσεων περιορίζεται συχνά σε εξιδανικευμένες ή σχετικά μικρής κλίμακας περιπτώσεις (Schaerf, 2002; Beasley et al., 2003).

Πέρα όμως από τα τεχνικά και υπολογιστικά ζητήματα, υπάρχει και ένας ευρύτερος εννοιολογικός περιορισμός: οι κλασικές προσεγγίσεις αντανακλούν μια εποχή και ένα θεωρητικό πλαίσιο όπου η έμφαση δινόταν κυρίως στην αποδοτική επιλογή χαρτοφυλακίου υπό γενικά κριτήρια απόδοσης και κινδύνου. Το πρόβλημα της αποτύπωσης δείκτη, αντίθετα, εντάσσεται σε ένα πιο εξειδικευμένο πλαίσιο παθητικής διαχείρισης, στο οποίο το ζητούμενο είναι η συστηματική παρακολούθηση ενός benchmark υπό ρεαλιστικούς επενδυτικούς περιορισμούς. Αυτή η μεταβολή της οπτικής απαιτεί μεθοδολογίες πιο ευέλικτες και πιο κοντά στις πρακτικές απαιτήσεις της σύγχρονης επενδυτικής διαχείρισης (Markowitz, 1952; Beasley et al., 2003).

Για όλους τους παραπάνω λόγους, η σχετική βιβλιογραφία έχει σταδιακά στραφεί προς πιο εξειδικευμένες και ευέλικτες προσεγγίσεις για το πρόβλημα της αποτύπωσης δείκτη. Στις προσεγγίσεις αυτές περιλαμβάνονται τόσο μέθοδοι μαθηματικού προγραμματισμού με πιο σύνθετες διατυπώσεις όσο και ευρετικές, μεταευρετικές και εξελικτικές τεχνικές, οι οποίες έχουν τη δυνατότητα να χειρίζονται πιο αποτελεσματικά περιορισμούς όπως ο cardinality constraint, τα όρια συμμετοχής, τα κόστη συναλλαγών και η δυναμική αναπροσαρμογής.

2.8 Κατηγορίες μεθόδων επίλυσης του προβλήματος

2.8.1 Μέθοδοι μαθηματικού προγραμματισμού

Οι μέθοδοι μαθηματικού προγραμματισμού αποτελούν μία από τις βασικές κατηγορίες προσεγγίσεων που έχουν χρησιμοποιηθεί για την αντιμετώπιση προβλημάτων βελτιστοποίησης χαρτοφυλακίου και, κατ' επέκταση, και για το πρόβλημα της αποτύπωσης δείκτη. Η βασική τους λογική στηρίζεται στη διατύπωση του προβλήματος με αυστηρά μαθηματικό τρόπο, μέσω μιας αντικειμενικής συνάρτησης και ενός συνόλου περιορισμών, έτσι ώστε η λύση να προκύπτει από τη συστηματική αναζήτηση του βέλτιστου σημείου μέσα σε ένα καθορισμένο πεδίο εφικτών λύσεων. Πρόκειται για προσεγγίσεις που χαρακτηρίζονται από υψηλό βαθμό τυπικότητας και σαφήνειας, καθώς επιχειρούν να μεταφράσουν το επενδυτικό πρόβλημα σε αυστηρά ορισμένη μαθηματική δομή.

Στο γενικότερο πεδίο της βελτιστοποίησης χαρτοφυλακίου, οι μέθοδοι μαθηματικού προγραμματισμού συνδέονται ιστορικά με τη δυνατότητα προσδιορισμού χαρτοφυλακίων που ικανοποιούν συγκεκριμένους στόχους απόδοσης και κινδύνου. Η βασική διαφορά τους από πιο εμπειρικές ή ευρετικές προσεγγίσεις είναι ότι δεν βασίζονται κυρίως σε κανόνες αναζήτησης ή σε προοδευτική βελτίωση υποψήφιων λύσεων, αλλά σε αυστηρή ανάλυση του ίδιου του μαθηματικού προβλήματος. Εφόσον η μορφή της αντικειμενικής συνάρτησης και των περιορισμών το επιτρέπει, οι μέθοδοι αυτές μπορούν να οδηγήσουν σε πολύ υψηλής ποιότητας λύσεις ή ακόμη και σε αποδεδειγμένα βέλτιστες λύσεις.

Στην περίπτωση της αποτύπωσης δείκτη, η εφαρμογή μαθηματικού προγραμματισμού βασίζεται στην ιδέα ότι το χαρτοφυλάκιο πρέπει να προσδιοριστεί έτσι ώστε να ελαχιστοποιείται ένα κατάλληλο μέτρο απόκλισης από τον δείκτη αναφοράς, υπό ένα σύνολο επενδυτικών περιορισμών. Η αντικειμενική συνάρτηση μπορεί, για παράδειγμα, να εκφράζει κάποιο μέτρο tracking error, ενώ οι περιορισμοί μπορεί να αφορούν τον προϋπολογισμό, τα όρια συμμετοχής των τίτλων, το κόστος συναλλαγών ή το πλήθος των επιλεγμένων μετοχών. Έτσι, το πρόβλημα αποκτά μια μορφή η οποία, τουλάχιστον σε ορισμένες εκδοχές του, είναι δυνατόν να αντιμετωπιστεί με εργαλεία μαθηματικού προγραμματισμού.

Μια σημαντική αρετή των μεθόδων αυτών είναι ότι προσφέρουν σαφή ερμηνεία των μεταβλητών, των περιορισμών και του στόχου. Αυτό είναι ιδιαίτερα χρήσιμο σε ένα πρόβλημα όπως το index tracking, όπου απαιτείται όχι μόνο η εύρεση μιας καλής λύσης, αλλά και η σαφής κατανόηση του πώς και γιατί αυτή η λύση προκύπτει. Η μαθηματική διατύπωση επιτρέπει στον ερευνητή να αποσαφηνίσει τις υποθέσεις του μοντέλου, να διαχωρίσει τις αποφάσεις επιλογής τίτλων από τις αποφάσεις κατανομής βαρών και να εξετάσει συστηματικά τον ρόλο κάθε περιορισμού στη διαμόρφωση της τελικής λύσης.

Παράλληλα, οι μέθοδοι μαθηματικού προγραμματισμού είναι ιδιαίτερα χρήσιμες όταν το πρόβλημα διατηρεί σχετικά ευνοϊκή δομή. Για παράδειγμα, όταν η αντικειμενική συνάρτηση και οι περιορισμοί έχουν γραμμική ή κυρτή μορφή, είναι συχνά δυνατόν να χρησιμοποιηθούν τεχνικές που οδηγούν σε ακριβείς και υπολογιστικά αποδοτικές λύσεις. Σε τέτοιες περιπτώσεις, οι μέθοδοι αυτές αποτελούν ένα ιδιαίτερα ισχυρό εργαλείο, καθώς δεν περιορίζονται στην αναζήτηση απλώς ικανοποιητικών λύσεων, αλλά μπορούν να προσφέρουν ισχυρότερα συμπεράσματα ως προς τη βέλτιστη επιλογή.

Ωστόσο, η αποτελεσματικότητα των μεθόδων μαθηματικού προγραμματισμού εξαρτάται σε μεγάλο βαθμό από τη μορφή του προβλήματος. Στο index tracking, η εισαγωγή ρεαλιστικών περιορισμών, όπως ο περιορισμός πληθικότητας, τα κατώτατα και ανώτατα όρια συμμετοχής, οι περιορισμοί κόστους συναλλαγών και η δυναμική αναπροσαρμογής, συχνά μεταβάλλει τη δομή του προβλήματος με τρόπο που καθιστά δυσκολότερη την άμεση εφαρμογή κλασικών τεχνικών. Η παρουσία δυαδικών μεταβλητών, μη γραμμικών σχέσεων και αλληλεξαρτήσεων μεταξύ μεταβλητών αυξάνει σημαντικά την πολυπλοκότητα και σε αρκετές περιπτώσεις περιορίζει την πρακτική αποτελεσματικότητα των αμιγώς ακριβών μαθηματικών μεθόδων.

Επιπλέον, οι μέθοδοι μαθηματικού προγραμματισμού συχνά απαιτούν σχετικά καθαρή και αυστηρή μορφοποίηση του προβλήματος, κάτι που δεν είναι πάντοτε εύκολο όταν επιχειρείται ρεαλιστική αναπαράσταση των συνθηκών της αγοράς. Πολλές φορές, για να καταστεί ένα πρόβλημα επίλυσιμο μέσω τέτοιων τεχνικών, απαιτούνται απλουστευτικές παραδοχές ή παραλείψεις στοιχείων που είναι ουσιαστικά από επενδυτική άποψη. Για παράδειγμα, μπορεί να παραλείπονται μη γραμμικά κόστη, λεπτομερείς περιορισμοί αναπροσαρμογής ή πιο σύνθετες μορφές tracking error, προκειμένου να διατηρηθεί το μοντέλο σε μορφή που να είναι διαχειρίσιμη υπολογιστικά.

Παρά τους περιορισμούς αυτούς, οι μέθοδοι μαθηματικού προγραμματισμού διατηρούν κεντρική θέση στη μελέτη του προβλήματος της αποτύπωσης δείκτη, διότι προσφέρουν το βασικό τυπικό πλαίσιο μέσα στο οποίο διατυπώνεται το πρόβλημα. Ακόμη και όταν δεν επαρκούν από μόνες τους για την πλήρη επίλυσή του, λειτουργούν ως αφετηρία για την ανάπτυξη πιο σύνθετων ή υβριδικών προσεγγίσεων. Συχνά, άλλωστε, οι ευρετικές και μεταευρετικές τεχνικές δεν αντικαθιστούν πλήρως τον μαθηματικό προγραμματισμό, αλλά βασίζονται σε μαθηματική διατύπωση του προβλήματος και στη συνέχεια χρησιμοποιούν πιο ευέλικτες στρατηγικές αναζήτησης για την εύρεση αποδοτικών λύσεων.

2.8.2 Ακριβείς μέθοδοι επίλυσης

Οι ακριβείς μέθοδοι επίλυσης αποτελούν ειδική κατηγορία μέσα στο ευρύτερο πεδίο του μαθηματικού προγραμματισμού και χαρακτηρίζονται από το γεγονός ότι επιδιώκουν την εύρεση της βέλτιστης λύσης ενός προβλήματος με αυστηρό και αποδεδειγμένο τρόπο. Σε αντίθεση με τις ευρετικές και μεταευρετικές προσεγγίσεις, οι οποίες στοχεύουν συνήθως στην αναζήτηση πολύ καλών ή σχεδόν βέλτιστων λύσεων μέσα σε σύνθετους χώρους αναζήτησης, οι ακριβείς μέθοδοι αποσκοπούν στη μαθηματικά βέβαιη ταυτοποίηση της άριστης λύσης, εφόσον φυσικά το πρόβλημα είναι διατυπωμένο σε μορφή που καθιστά δυνατή μια τέτοια διαδικασία (Canakgoz & Beasley, 2009; Scozzari et al., 2013).

Η βασική αξία των ακριβών μεθόδων έγκειται στη θεωρητική τους ισχύ. Όταν ένα πρόβλημα μπορεί να επιλυθεί με ακρίβεια, ο ερευνητής ή ο διαχειριστής έχει τη δυνατότητα να γνωρίζει ότι η λύση που προέκυψε δεν είναι απλώς ικανοποιητική ή καλή, αλλά η καλύτερη δυνατή υπό τις δεδομένες παραδοχές, τους στόχους και τους περιορισμούς του μοντέλου. Η δυνατότητα αυτή είναι ιδιαίτερα σημαντική σε θεωρητικές αναλύσεις, σε εφαρμογές μικρής ή μέσης κλίμακας, καθώς και σε περιπτώσεις όπου η ποιότητα της λύσης πρέπει να τεκμηριώνεται με αυστηρό τρόπο (Scozzari et al., 2013).

Στο πεδίο της βελτιστοποίησης χαρτοφυλακίου, οι ακριβείς μέθοδοι έχουν χρησιμοποιηθεί κυρίως σε περιπτώσεις όπου η μορφή του προβλήματος επιτρέπει την αξιοποίηση καθιερωμένων τεχνικών, όπως ο γραμμικός, ο τετραγωνικός ή ο μικτός αθέριος

προγραμματισμός. Όταν, για παράδειγμα, η αντικειμενική συνάρτηση και οι περιορισμοί έχουν κατάλληλη δομή, είναι δυνατόν να εφαρμοστούν αλγόριθμοι που εγγυώνται την εύρεση της βέλτιστης λύσης ή, τουλάχιστον, την εύρεση λύσης με σαφώς προσδιορισμένο επίπεδο βελτιστότητας. Αυτή η δυνατότητα υπήρξε ιδιαίτερα σημαντική για την εξέλιξη της θεωρίας, καθώς προσέφερε σημείο αναφοράς για τη σύγκριση εναλλακτικών μεθόδων και για την κατανόηση της δομής του προβλήματος (Canakgoz & Beasley, 2009; Bonami & Lejeune, 2009).

Στην περίπτωση της αποτύπωσης δείκτη, οι ακριβείς μέθοδοι είναι περισσότερο εφαρμόσιμες όταν το πρόβλημα διατυπώνεται σε σχετικά απλή μορφή. Αν, για παράδειγμα, ο στόχος εκφράζεται μέσω μιας καλά συμπεριφερόμενης αντικειμενικής συνάρτησης και οι περιορισμοί παραμένουν περιορισμένοι σε βασικές συνθήκες, όπως ο προϋπολογισμός και η μη αρνητικότητα των βαρών, τότε η επίλυση μπορεί να είναι εφικτή μέσω κλασικών τεχνικών βελτιστοποίησης. Σε τέτοιες περιπτώσεις, οι ακριβείς μέθοδοι προσφέρουν σημαντικό πλεονέκτημα, καθώς επιτρέπουν την αυστηρή αναζήτηση της λύσης που ελαχιστοποιεί το tracking error ή ικανοποιεί με τον καλύτερο δυνατό τρόπο τον επιλεγμένο στόχο (Canakgoz & Beasley, 2009).

Ωστόσο, η πραγματική δυσκολία αναδύεται όταν επιχειρείται η ενσωμάτωση πιο ρεαλιστικών χαρακτηριστικών του προβλήματος. Η εισαγωγή περιορισμών πληθικότητας, κατώτατων και ανώτατων ορίων συμμετοχής τίτλων, κόστους συναλλαγών ή περιορισμών αναπροσαρμογής μεταβάλλει ουσιαστικά τη μαθηματική δομή του προβλήματος. Σε αυτές τις περιπτώσεις, το πρόβλημα γίνεται πολύ πιο δύσκολο, καθώς συνδυάζει συνεχείς και διακριτές μεταβλητές, ενώ συχνά αποκτά μη γραμμικό ή μη κυρτό χαρακτήρα. Παρότι οι ακριβείς μέθοδοι μπορούν θεωρητικά να εφαρμοστούν και σε ορισμένες από αυτές τις διατυπώσεις, το υπολογιστικό κόστος αυξάνεται δραματικά και η επίλυση καθίσταται συχνά δυσχερής ή πρακτικά ανέφικτη για προβλήματα μεγάλης κλίμακας (Canakgoz & Beasley, 2009; Scozzari et al., 2013; Bonami & Lejeune, 2009).

Το βασικό μειονέκτημα των ακριβών μεθόδων στο index tracking είναι ακριβώς αυτή η ευαισθησία τους στην πολυπλοκότητα του προβλήματος. Όσο αυξάνεται ο αριθμός των διαθέσιμων τίτλων και όσο εμπλουτίζεται το μοντέλο με πιο ρεαλιστικούς περιορισμούς, τόσο δυσκολότερο γίνεται να βρεθεί η βέλτιστη λύση με αυστηρά υπολογιστικά μέσα. Ο χώρος των πιθανών λύσεων διευρύνεται ραγδαία, ενώ η συνδυαστική φύση του προβλήματος, ιδίως όταν υπάρχουν cardinality constraints, καθιστά την εξαντλητική ή αυστηρά βελτιστοποιητική αναζήτηση εξαιρετικά απαιτητική. Σε πολλές περιπτώσεις, ο απαιτούμενος χρόνος ή οι υπολογιστικοί πόροι ξεπερνούν τα όρια της πρακτικής εφαρμογής (Canakgoz & Beasley, 2009; Fu et al., 2023).

Ένα ακόμη σημαντικό σημείο είναι ότι οι ακριβείς μέθοδοι απαιτούν συνήθως πολύ σαφή και σχετικά «καθαρή» διατύπωση του προβλήματος. Όταν το πρόβλημα περιλαμβάνει στοιχεία που είναι δύσκολο να αποτυπωθούν με γραμμικό ή ομαλό τρόπο, όπως σύνθετα κόστη συναλλαγών, εναλλακτικούς ορισμούς του tracking error ή πιο ευέλικτες επενδυτικές απαιτήσεις, η εφαρμογή ακριβών τεχνικών γίνεται ακόμη δυσκολότερη. Συχνά, για να είναι εφικτή η χρήση τέτοιων μεθόδων, απαιτείται απλοποίηση του μοντέλου, γεγονός που μειώνει τον ρεαλισμό της προσέγγισης και περιορίζει την πρακτική αξία των αποτελεσμάτων (Canakgoz & Beasley, 2009; Bonami & Lejeune, 2009).

Παρά τις δυσκολίες αυτές, οι ακριβείς μέθοδοι διατηρούν ιδιαίτερη σημασία στη μελέτη του προβλήματος της αποτύπωσης δείκτη. Πρώτον, προσφέρουν ένα χρήσιμο θεωρητικό

σημείο αναφοράς για την κατανόηση του προβλήματος και τη διαμόρφωση πιο εξελιγμένων μοντέλων. Δεύτερον, σε μικρότερης κλίμακας ή απλουστευμένες περιπτώσεις, μπορούν να χρησιμοποιηθούν για την παραγωγή βέλτιστων λύσεων, οι οποίες λειτουργούν ως benchmark για τη σύγκριση με ευρετικές και μεταευρετικές τεχνικές. Τρίτον, η λογική και η δομή τους επηρεάζουν ακόμη και τις προσεγγίσεις που δεν είναι ακριβείς, καθώς πολλές ευρετικές μέθοδοι βασίζονται σε μαθηματική διατύπωση του προβλήματος και στη συνέχεια αναζητούν πιο ευέλικτο τρόπο διερεύνησης του χώρου λύσεων (Scozzari et al., 2013; Canakgoz & Beasley, 2009).

2.8.3 Ευρετικές και μεταευρετικές μέθοδοι

Οι ευρετικές και μεταευρετικές μέθοδοι έχουν αποκτήσει ιδιαίτερη σημασία στο πρόβλημα της αποτύπωσης δείκτη, κυρίως λόγω των περιορισμών που εμφανίζουν οι κλασικές και οι ακριβείς προσεγγίσεις όταν το πρόβλημα αποκτά ρεαλιστική και υπολογιστικά σύνθετη μορφή. Όπως έχει ήδη επισημανθεί, η εισαγωγή περιορισμών πληθικότητας, ορίων συμμετοχής τίτλων, κόστους συναλλαγών και δυναμικών αναπροσαρμογών καθιστά το πρόβλημα ιδιαίτερα δύσκολο, τόσο ως προς τη μαθηματική του διατύπωση όσο και ως προς την υπολογιστική του επίλυση. Σε αυτό το πλαίσιο, οι ευρετικές και μεταευρετικές τεχνικές εμφανίζονται ως μια εναλλακτική κατηγορία προσεγγίσεων που δεν επιδιώκει απαραίτητα την αυστηρά βέλτιστη λύση, αλλά την εύρεση λύσεων υψηλής ποιότητας μέσα σε εύλογο υπολογιστικό χρόνο (Beasley et al., 2003; Maringer & Oyewumi, 2007; Guastaroba et al., 2012).

Η βασική λογική των ευρετικών μεθόδων στηρίζεται στην αξιοποίηση απλών αλλά στοχευμένων κανόνων αναζήτησης, οι οποίοι επιτρέπουν τη σταδιακή προσέγγιση καλών λύσεων χωρίς να απαιτείται εξαντλητική διερεύνηση όλων των δυνατοτήτων. Οι ευρετικές μέθοδοι είναι συνήθως σχεδιασμένες για ένα συγκεκριμένο πρόβλημα ή για μια συγκεκριμένη κατηγορία προβλημάτων και εκμεταλλεύονται στοιχεία της δομής τους για να περιορίσουν το πεδίο αναζήτησης. Αυτό σημαίνει ότι, παρότι δεν εγγυώνται τη βέλτιστη λύση, συχνά επιτυγχάνουν πολύ ικανοποιητικά αποτελέσματα, ιδιαίτερα όταν η πλήρης βελτιστοποίηση είναι υπολογιστικά απαγορευτική (Beasley et al., 2003; Guastaroba et al., 2012).

Οι μεταευρετικές μέθοδοι αποτελούν πιο γενικευμένες στρατηγικές αναζήτησης, οι οποίες μπορούν να εφαρμοστούν σε ευρύ φάσμα προβλημάτων βελτιστοποίησης. Σε αντίθεση με τις απλές ευρετικές, οι μεταευρετικές δεν βασίζονται αποκλειστικά σε κανόνες ειδικά σχεδιασμένους για μία μόνο εφαρμογή, αλλά σε ευρύτερους μηχανισμούς εξερεύνησης και εκμετάλλευσης του χώρου λύσεων. Οι μέθοδοι αυτές έχουν σχεδιαστεί έτσι ώστε να μπορούν να διαφεύγουν από τοπικά βέλτιστα, να εξερευνούν διαφορετικές περιοχές του πεδίου λύσεων και να διαχειρίζονται αποτελεσματικά προβλήματα υψηλής πολυπλοκότητας, ακόμη και όταν η μαθηματική τους δομή δεν είναι ιδιαίτερα ευνοϊκή για ακριβείς τεχνικές (Maringer & Oyewumi, 2007; Chiam et al., 2013).

Η ανάγκη χρήσης ευρετικών και μεταευρετικών μεθόδων στο πρόβλημα της αποτύπωσης δείκτη προκύπτει ακριβώς από τη φύση του προβλήματος. Όταν το χαρτοφυλάκιο πρέπει να επιλεγεί από μεγάλο αριθμό διαθέσιμων τίτλων, υπό περιορισμό πληθικότητας, με όρια στα βάρη και με έλεγχο κόστους συναλλαγών, ο χώρος λύσεων γίνεται τεράστιος. Κάθε συνδυασμός τίτλων και βαρών συνιστά μια πιθανή λύση, και ο αριθμός αυτών των λύσεων αυξάνεται ταχύτατα όσο αυξάνεται το μέγεθος του προβλήματος. Σε τέτοιες συνθήκες, η

εξαντλητική ή αυστηρά ακριβής αναζήτηση δεν είναι συνήθως πρακτικά εφαρμόσιμη, ενώ οι ευρετικές και μεταευρετικές μέθοδοι προσφέρουν μια πιο ρεαλιστική διέξοδο (Beasley et al., 2003; Maringer & Oyewumi, 2007; Guastaroba et al., 2012).

Ένα από τα κύρια πλεονεκτήματα των μεθόδων αυτών είναι η ευελιξία τους. Μπορούν να ενσωματώσουν σχετικά εύκολα πολύπλοκους και ετερογενείς περιορισμούς, χωρίς να απαιτούν το πρόβλημα να λάβει απόλυτα «καθαρή» ή απλοποιημένη μαθηματική μορφή. Για παράδειγμα, μπορούν να χειριστούν ταυτόχρονα διακριτές αποφάσεις, όπως η επιλογή τίτλων, και συνεχείς αποφάσεις, όπως ο καθορισμός βαρών. Μπορούν επίσης να προσαρμοστούν σε διαφορετικούς ορισμούς του tracking error και να διαχειριστούν περιορισμούς κόστους ή αναπροσαρμογής μέσω μηχανισμών ποινής, επιδιόρθωσης ή ελέγχου εφικτότητας. Αυτή η ικανότητα προσαρμογής είναι εξαιρετικά χρήσιμη σε ένα πρόβλημα όπως το index tracking, όπου ο ρεαλισμός του μοντέλου είναι κρίσιμος (Beasley et al., 2003; Chiam et al., 2013).

Ένα δεύτερο σημαντικό πλεονέκτημα είναι ότι οι μεταευρετικές τεχνικές προσφέρουν καλύτερη ισορροπία ανάμεσα στην ποιότητα της λύσης και στον υπολογιστικό χρόνο. Παρότι δεν παρέχουν εγγύηση βέλτιστης λύσης, μπορούν συχνά να εντοπίσουν λύσεις πολύ υψηλής ποιότητας μέσα σε αποδεκτό χρονικό διάστημα, ακόμη και για προβλήματα μεγάλης κλίμακας. Αυτό είναι ιδιαίτερα σημαντικό σε εφαρμογές χρηματοοικονομικής βελτιστοποίησης, όπου η πρακτική χρησιμότητα μιας μεθόδου εξαρτάται όχι μόνο από τη θεωρητική της ισχύ, αλλά και από το κατά πόσο μπορεί να χρησιμοποιηθεί σε πραγματικά δεδομένα και σε ρεαλιστικά χρονικά πλαίσια (Guastaroba et al., 2012; Chiam et al., 2013).

Ωστόσο, οι ευρετικές και μεταευρετικές μέθοδοι έχουν και ορισμένους περιορισμούς. Ο βασικότερος είναι ότι η ποιότητα της λύσης εξαρτάται συχνά σε σημαντικό βαθμό από τον σχεδιασμό της μεθόδου, την επιλογή των παραμέτρων της και τον τρόπο με τον οποίο ενσωματώνονται οι περιορισμοί του προβλήματος. Δεν υπάρχει δηλαδή ενιαία εγγύηση ότι μια μεταευρετική θα αποδώσει εξίσου καλά σε όλες τις μορφές του προβλήματος. Επιπλέον, η απόδοσή τους μπορεί να είναι ευαίσθητη σε παράγοντες όπως η αρχικοποίηση, το μέγεθος του πληθυσμού ή οι κανόνες μετάβασης από μία λύση σε άλλη. Αυτό σημαίνει ότι η εφαρμογή τους απαιτεί προσεκτικό σχεδιασμό και πειραματική διερεύνηση, ώστε να διασφαλιστεί ότι η επιλεγμένη παραμετροποίηση είναι πράγματι κατάλληλη (Maringer & Oyewumi, 2007; Chiam et al., 2013).

Παρά τους περιορισμούς αυτούς, οι ευρετικές και μεταευρετικές μέθοδοι έχουν κυριαρχήσει σε μεγάλο βαθμό στη βιβλιογραφία των σύνθετων προβλημάτων χαρτοφυλακίου και ειδικότερα της αποτύπωσης δείκτη. Ο λόγος είναι ότι προσφέρουν ένα επίπεδο ευελιξίας και πρακτικής αποδοτικότητας που δύσκολα επιτυγχάνεται με τις ακριβείς μεθόδους όταν το πρόβλημα περιλαμβάνει σύνθετους και ρεαλιστικούς περιορισμούς. Ειδικά για το index tracking, όπου το ενδιαφέρον εστιάζεται συχνά σε μεγάλης κλίμακας προβλήματα και σε ανάγκη ισορροπίας ανάμεσα στην πιστή παρακολούθηση και στο λειτουργικό κόστος, οι μέθοδοι αυτές έχουν αποδειχθεί ιδιαίτερα χρήσιμες (Beasley et al., 2003; Guastaroba et al., 2012; Chiam et al., 2013).

2.9 Εξελικτικές και μεταευρετικές τεχνικές

2.9.1 Γενετικοί Αλγόριθμοι

Οι Γενετικοί Αλγόριθμοι αποτελούν μία από τις πιο γνωστές και διαδεδομένες κατηγορίες εξελικτικών αλγορίθμων και έχουν χρησιμοποιηθεί ευρέως σε προβλήματα βελτιστοποίησης υψηλής πολυπλοκότητας. Η βασική τους φιλοσοφία είναι εμπνευσμένη από τους μηχανισμούς της φυσικής εξέλιξης, όπως η επιλογή, η αναπαραγωγή, η διασταύρωση και η μετάλλαξη. Μέσω αυτών των μηχανισμών, ένας πληθυσμός υποψήφιας λύσεων εξελίσσεται σταδιακά από γενιά σε γενιά, με στόχο την παραγωγή λύσεων ολοένα και καλύτερης ποιότητας ως προς ένα προκαθορισμένο κριτήριο αξιολόγησης (De Jong, 1988; McCall, 2005).

Η κεντρική ιδέα των Γενετικών Αλγορίθμων είναι ότι κάθε υποψήφια λύση αναπαρίσταται με μια δομή που συχνά αποκαλείται χρωμόσωμα. Το χρωμόσωμα αυτό περιέχει την κωδικοποιημένη πληροφορία της λύσης, δηλαδή τα στοιχεία που καθορίζουν τις αποφάσεις του προβλήματος. Στην περίπτωση της βελτιστοποίησης χαρτοφυλακίου, τα στοιχεία αυτά μπορεί να αφορούν την επιλογή των τίτλων που περιλαμβάνονται στο χαρτοφυλάκιο, τα βάρη τους ή και τα δύο ταυτόχρονα. Έτσι, ο Γενετικός Αλγόριθμος δεν λειτουργεί απευθείας πάνω στις οικονομικές έννοιες, αλλά πάνω σε μια αναπαράσταση των πιθανών λύσεων, την οποία τροποποιεί και εξελίσσει συστηματικά (Beasley et al., 2003; Lin & Liu, 2008).

Η ποιότητα κάθε λύσης εκτιμάται μέσω μιας συνάρτησης καταλληλότητας. Η συνάρτηση αυτή παίζει κεντρικό ρόλο στη λειτουργία του αλγορίθμου, καθώς εκφράζει το πόσο καλή θεωρείται μια υποψήφια λύση ως προς τον στόχο του προβλήματος. Στην αποτύπωση δείκτη, η συνάρτηση καταλληλότητας μπορεί να συνδέεται άμεσα με το tracking error, δηλαδή με την απόκλιση μεταξύ χαρτοφυλακίου και δείκτη, και να περιλαμβάνει επιπλέον όρους που λαμβάνουν υπόψη παραβιάσεις περιορισμών ή άλλες πρακτικές απαιτήσεις. Όσο καλύτερη είναι μια λύση σύμφωνα με αυτή τη συνάρτηση, τόσο μεγαλύτερη είναι η πιθανότητα να επιλεγεί για τη δημιουργία απογόνων στις επόμενες γενιές (Beasley et al., 2003; Ni & Yin, 2013).

Η διαδικασία εξέλιξης του πληθυσμού βασίζεται σε τρεις κύριους μηχανισμούς. Ο πρώτος είναι η επιλογή, μέσω της οποίας κάποιες από τις τρέχουσες λύσεις επιλέγονται ως γονείς για τη δημιουργία νέων λύσεων. Συνήθως προτιμώνται οι πιο κατάλληλες λύσεις, χωρίς όμως να αποκλείονται πλήρως και λιγότερο αποδοτικές, ώστε να διατηρείται ένα επίπεδο ποικιλίας στον πληθυσμό. Ο δεύτερος είναι η διασταύρωση, κατά την οποία δύο ή περισσότερες γονικές λύσεις συνδυάζονται για να παραχθούν νέες λύσεις που κληρονομούν χαρακτηριστικά από τους γονείς τους. Ο τρίτος είναι η μετάλλαξη, μέσω της οποίας εισάγονται μικρές τυχαίες τροποποιήσεις στα χρωμοσώματα, με στόχο τη διατήρηση της ποικιλίας και την αποφυγή πρόωρης σύγκλισης σε κατώτερες τοπικές λύσεις (De Jong, 1988; McCall, 2005).

Ένα από τα σημαντικότερα πλεονεκτήματα των Γενετικών Αλγορίθμων είναι ο πληθυσμιακός χαρακτήρας τους. Σε αντίθεση με πολλές κλασικές ή ντετερμινιστικές τεχνικές, οι οποίες εξετάζουν κάθε φορά μία λύση και κινούνται τοπικά γύρω από αυτήν, οι Γενετικοί Αλγόριθμοι διαχειρίζονται ταυτόχρονα έναν ολόκληρο πληθυσμό λύσεων.

Αυτό τους επιτρέπει να εξερευνούν πολλαπλές περιοχές του χώρου αναζήτησης την ίδια χρονική στιγμή, γεγονός που αυξάνει την πιθανότητα εντοπισμού ποιοτικών λύσεων και μειώνει τον κίνδυνο εγκλωβισμού σε τοπικά βέλτιστα. Το χαρακτηριστικό αυτό είναι ιδιαίτερα σημαντικό για σύνθετα προβλήματα, όπως η αποτύπωση δείκτη, όπου ο χώρος λύσεων είναι μεγάλος, μικτός και συχνά μη ομαλός (De Jong, 1988; McCall, 2005).

Ένα ακόμη ουσιαστικό πλεονέκτημα είναι η ευελιξία τους ως προς τη μορφή του προβλήματος. Οι Γενετικοί Αλγόριθμοι δεν απαιτούν αυστηρές παραδοχές για κυρτότητα, διαφορισιμότητα ή απλή γραμμική δομή της αντικειμενικής συνάρτησης. Μπορούν να εφαρμοστούν σε προβλήματα με διακριτές και συνεχείς μεταβλητές, σύνθετους περιορισμούς και εναλλακτικούς ορισμούς στόχου. Αυτό εξηγεί γιατί έχουν βρει ιδιαίτερα πρόσφορο πεδίο εφαρμογής σε προβλήματα χαρτοφυλακίου, όπου η επιλογή τίτλων, ο καθορισμός βαρών, το κόστος συναλλαγών και οι περιορισμοί πληθικότητας συνυπάρχουν σε ένα ενιαίο πλαίσιο βελτιστοποίησης (Lin & Liu, 2008; Ni & Yin, 2013).

Στην αποτύπωση δείκτη, η καταλληλότητα των Γενετικών Αλγορίθμων συνδέεται ακριβώς με τη μικτή φύση του προβλήματος. Το index tracking απαιτεί ταυτόχρονα επιλογή υποσυνόλου τίτλων και προσδιορισμό βαρών, υπό περιορισμούς όπως ο προϋπολογισμός, τα όρια συμμετοχής, το κόστος συναλλαγών και ο cardinality constraint. Η δυνατότητα των Γενετικών Αλγορίθμων να αναπαριστούν φυσικά τέτοιου είδους αποφάσεις μέσα στη δομή των χρωμοσωμάτων και να εξελίσσουν σταδιακά τις λύσεις τους καθιστά ιδιαίτερα κατάλληλους για το συγκεκριμένο πρόβλημα. Ιδίως όταν το μέγεθος του προβλήματος αυξάνεται, η χρήση μιας τέτοιας εξελικτικής μεθόδου καθίσταται πολύ ελκυστική σε σχέση με αμιγώς ακριβείς τεχνικές (Beasley et al., 2003; Ni & Yin, 2013; Chiam et al., 2013).

Παρά τα σημαντικά τους πλεονεκτήματα, οι Γενετικοί Αλγόριθμοι δεν στερούνται περιορισμών. Η απόδοσή τους εξαρτάται σε σημαντικό βαθμό από τον τρόπο με τον οποίο έχει σχεδιαστεί η αναπαράσταση των λύσεων, η συνάρτηση καταλληλότητας και οι γενετικοί τελεστές. Μια ακατάλληλη κωδικοποίηση ή μια μη καλά σχεδιασμένη fitness function μπορεί να οδηγήσει τον αλγόριθμο σε αδύναμα αποτελέσματα, ακόμη και αν η βασική ιδέα της μεθόδου είναι θεωρητικά κατάλληλη. Επιπλέον, η παραμετροποίησή τους, όπως το μέγεθος του πληθυσμού, η πιθανότητα διασταύρωσης, η πιθανότητα μετάλλαξης και ο αριθμός γενεών, επηρεάζει άμεσα την ποιότητα των λύσεων και τον απαιτούμενο υπολογιστικό χρόνο (De Jong, 1988; McCall, 2005; Maringer & Oyewumi, 2007).

Ακριβώς για τον λόγο αυτό, η εφαρμογή ενός Γενετικού Αλγορίθμου σε ένα πρόβλημα όπως το index tracking δεν είναι απλή υπόθεση χρήσης μιας γενικής μεθόδου, αλλά απαιτεί προσεκτικό σχεδιασμό και προσαρμογή στα ιδιαίτερα χαρακτηριστικά του προβλήματος. Η αναπαράσταση του χαρτοφυλακίου, η διαχείριση των εφικτών λύσεων, ο τρόπος ελέγχου των περιορισμών και η μορφή της συνάρτησης στόχου πρέπει να σχεδιάζονται έτσι ώστε να εξυπηρετούν την οικονομική λογική του προβλήματος και όχι απλώς τη γενική λειτουργία του αλγορίθμου (Beasley et al., 2003; Ni & Yin, 2013; Chiam et al., 2013).

2.9.2 Εξελικτικές στρατηγικές

Οι εξελικτικές στρατηγικές αποτελούν μία ακόμη σημαντική κατηγορία εξελικτικών μεθόδων βελτιστοποίησης και ανήκουν στο ίδιο ευρύτερο πεδίο με τους Γενετικούς Αλγορίθμους. Παρότι μοιράζονται με αυτούς την έμπνευση από τις αρχές της φυσικής

εξέλιξης, διαφοροποιούνται ως προς τον τρόπο αναπαράστασης των λύσεων, τους μηχανισμούς παραγωγής νέων υποψήφίων λύσεων και τη γενικότερη φιλοσοφία αναζήτησης. Οι εξελικτικές στρατηγικές αναπτύχθηκαν αρχικά με έμφαση σε προβλήματα συνεχούς βελτιστοποίησης και έχουν συνδεθεί ιδιαίτερα με την προσαρμοστική τροποποίηση αριθμητικών παραμέτρων μέσα από διαδικασίες επαναληπτικής εξέλιξης (Beyer & Schwefel, 2002; Li et al., 2020).

Η βασική ιδέα των εξελικτικών στρατηγικών είναι ότι ένας πληθυσμός λύσεων εξελίσσεται σταδιακά μέσω μηχανισμών που δίνουν μεγαλύτερη έμφαση στη μετάλλαξη και λιγότερη στη διασταύρωση, σε σύγκριση με ό,τι συναντάται συνήθως στους Γενετικούς Αλγορίθμους. Κάθε λύση αναπαρίσταται συνήθως ως διάνυσμα πραγματικών τιμών, γεγονός που καθιστά τη μέθοδο ιδιαίτερα κατάλληλη για προβλήματα όπου οι μεταβλητές απόφασης είναι συνεχείς. Η παραγωγή νέων λύσεων πραγματοποιείται μέσω μικρών μεταβολών στις τιμές των παραμέτρων, ενώ οι πιο κατάλληλες λύσεις επιλέγονται για να αποτελέσουν τη βάση της επόμενης γενιάς (Beyer & Schwefel, 2002; Hansen & Ostermeier, 2001).

Ένα βασικό χαρακτηριστικό των εξελικτικών στρατηγικών είναι η προσαρμοστικότητα των ίδιων των μηχανισμών αναζήτησης. Σε πολλές παραλλαγές της μεθόδου, δεν εξελίσσονται μόνο οι τιμές των μεταβλητών απόφασης, αλλά και οι παράμετροι που ελέγχουν το μέγεθος και την ένταση των μεταβολών τους. Με άλλα λόγια, η στρατηγική αναζήτησης προσαρμόζεται δυναμικά κατά τη διάρκεια της εξέλιξης. Το χαρακτηριστικό αυτό δίνει στις εξελικτικές στρατηγικές μεγάλη ευελιξία και τις καθιστά ιδιαίτερα ισχυρές σε προβλήματα όπου απαιτείται λεπτή ρύθμιση συνεχών παραμέτρων (Hansen & Ostermeier, 2001; Beyer & Schwefel, 2002).

Σε σχέση με το πρόβλημα της αποτύπωσης δείκτη, οι εξελικτικές στρατηγικές παρουσιάζουν ενδιαφέρον κυρίως στις περιπτώσεις όπου το βάρος δίνεται στον προσδιορισμό των βαρών των τίτλων και γενικότερα σε συνεχή τμήματα του προβλήματος. Όταν, για παράδειγμα, έχει ήδη καθοριστεί ένα συγκεκριμένο σύνολο τίτλων ή όταν η έμφαση βρίσκεται περισσότερο στη βελτιστοποίηση αριθμητικών παραμέτρων του χαρτοφυλακίου, οι εξελικτικές στρατηγικές μπορούν να λειτουργήσουν αποτελεσματικά. Η ικανότητά τους να εξερευνούν συνεχείς χώρους λύσεων με προσαρμοστικό τρόπο τις καθιστά ιδιαίτερα χρήσιμες σε τέτοιου είδους εφαρμογές (Beyer et al., 2014; Li et al., 2020).

Ωστόσο, η καταλληλότητά τους στο πλήρες πρόβλημα του index tracking εξαρτάται από τη μορφή της διατύπωσης. Όταν το πρόβλημα περιλαμβάνει έντονα διακριτές αποφάσεις, όπως η επιλογή υποσυνόλου τίτλων υπό cardinality constraint, οι εξελικτικές στρατηγικές δεν είναι πάντοτε τόσο φυσικές ή άμεσες στην εφαρμογή τους όσο οι Γενετικοί Αλγόριθμοι. Αυτό συμβαίνει επειδή η κλασική τους μορφή είναι περισσότερο προσανατολισμένη σε συνεχείς μεταβλητές παρά σε μικτές ή αμιγώς συνδυαστικές δομές. Παρ' όλα αυτά, με κατάλληλες τροποποιήσεις ή σε υβριδικά σχήματα, μπορούν να χρησιμοποιηθούν και σε πιο σύνθετες μορφές του προβλήματος (Beyer & Schwefel, 2002; Beasley et al., 2003).

Ένα σημαντικό πλεονέκτημα των εξελικτικών στρατηγικών είναι η ικανότητά τους να επιτυγχάνουν καλή ισορροπία ανάμεσα στη διερεύνηση και στην εκμετάλλευση του χώρου λύσεων. Η προσαρμοστική φύση των μεταβολών επιτρέπει στη μέθοδο να κινείται πιο «τολμηρά» όταν βρίσκεται μακριά από καλές λύσεις και πιο «λεπτομερώς» όταν πλησιάζει σε υποσχόμενες περιοχές του χώρου αναζήτησης. Αυτό είναι ιδιαίτερα χρήσιμο σε

προβλήματα όπου το πεδίο λύσεων είναι πολύπλοκο και η ποιότητα της λύσης εξαρτάται από ακριβή ρύθμιση συνεχών παραμέτρων (Hansen & Ostermeier, 2001; Li et al., 2020).

Από την άλλη πλευρά, οι εξελικτικές στρατηγικές παρουσιάζουν και ορισμένους περιορισμούς. Η αποτελεσματικότητά τους εξαρτάται από τον σωστό σχεδιασμό της αναπαράστασης, των κανόνων μετάλλαξης και των μηχανισμών επιλογής. Επιπλέον, όταν το πρόβλημα έχει ισχυρό συνδυαστικό χαρακτήρα ή περιλαμβάνει πολλούς δύσκολους διακριτούς περιορισμούς, μπορεί να απαιτείται ειδική προσαρμογή της μεθόδου, ώστε να παραμένει αποδοτική και να μην παράγει μεγάλο αριθμό μη εφικτών λύσεων. Συνεπώς, παρότι είναι ισχυρό εργαλείο, η εφαρμογή τους δεν είναι πάντοτε εξίσου φυσική σε κάθε είδος προβλήματος χαρτοφυλακίου (Beyer & Schwefel, 2002; Li et al., 2020; Beyer et al., 2014).

2.9.3 Προσομοιωμένη ανόπτηση

Η προσομοιωμένη ανόπτηση αποτελεί μία από τις πιο γνωστές και ευρέως χρησιμοποιούμενες μεταερευτικές τεχνικές βελτιστοποίησης και έχει εφαρμοστεί σε πλήθος σύνθετων προβλημάτων όπου η αναζήτηση της βέλτιστης λύσης είναι υπολογιστικά δύσκολη. Η λογική της μεθόδου είναι εμπνευσμένη από τη φυσική διαδικασία της ανόπτησης στα μέταλλα, κατά την οποία ένα υλικό θερμαίνεται και στη συνέχεια ψύχεται σταδιακά, ώστε να αποκτήσει πιο σταθερή και οργανωμένη εσωτερική δομή. Στο πλαίσιο της βελτιστοποίησης, η ιδέα αυτή μεταφέρεται σε μια διαδικασία αναζήτησης λύσεων, όπου ο αλγόριθμος εξερευνά τον χώρο λύσεων επιτρέποντας προσωρινά και τη μετάβαση σε χειρότερες λύσεις, με σκοπό να αποφύγει τον εγκλωβισμό σε τοπικά βέλτιστα.

Η βασική διαφορά της προσομοιωμένης ανόπτησης από πιο απλές τοπικές μεθόδους αναζήτησης είναι ακριβώς αυτή η δυνατότητα αποδοχής μη βελτιωτικών μετακινήσεων. Σε μια κλασική τοπική αναζήτηση, ο αλγόριθμος μετακινείται συνήθως μόνο όταν εντοπίζει καλύτερη λύση από την τρέχουσα. Έτσι όμως υπάρχει αυξημένος κίνδυνος να παγιδευτεί γρήγορα σε μια τοπικά καλή αλλά συνολικά κατώτερη περιοχή του χώρου λύσεων. Αντίθετα, η προσομοιωμένη ανόπτηση επιτρέπει, με ορισμένη πιθανότητα, τη μετάβαση και σε χειρότερες λύσεις, ιδιαίτερα στα αρχικά στάδια της διαδικασίας, όταν η «θερμοκρασία» του συστήματος είναι υψηλή. Με τον τρόπο αυτό, ο αλγόριθμος διατηρεί μεγαλύτερη ελευθερία εξερεύνησης και μπορεί να διαφύγει από τοπικές παγίδες.

Η λειτουργία της μεθόδου βασίζεται σε τρία βασικά στοιχεία: την τρέχουσα λύση, τον μηχανισμό παραγωγής γειτονικών λύσεων και το σχήμα ψύξης. Αρχικά επιλέγεται μια αρχική λύση και καθορίζεται μια αρχική θερμοκρασία. Σε κάθε επανάληψη, παράγεται μια νέα υποψήφια λύση μέσω κάποιας μικρής τροποποίησης της τρέχουσας. Αν η νέα λύση είναι καλύτερη, γίνεται αποδεκτή. Αν είναι χειρότερη, μπορεί και πάλι να γίνει αποδεκτή με πιθανότητα που εξαρτάται από το πόσο χειρότερη είναι και από το τρέχον επίπεδο θερμοκρασίας. Όσο η διαδικασία προχωρά και η θερμοκρασία μειώνεται, η πιθανότητα αποδοχής χειρότερων λύσεων μειώνεται επίσης, με αποτέλεσμα ο αλγόριθμος να γίνεται σταδιακά πιο «συντηρητικός» και να συγκεντρώνεται σε πιο τοπική βελτίωση.

Η προσομοιωμένη ανόπτηση είναι ιδιαίτερα ελκυστική για προβλήματα όπου ο χώρος λύσεων παρουσιάζει μεγάλο αριθμό τοπικών βέλτιστων. Το χαρακτηριστικό αυτό απαντάται πολύ συχνά σε προβλήματα χρηματοοικονομικής βελτιστοποίησης και ειδικότερα στην αποτύπωση δείκτη, ιδίως όταν συνυπάρχουν περιορισμοί επιλογής τίτλων, ορίων βαρών και κόστους συναλλαγών. Σε τέτοιες περιπτώσεις, η ικανότητα του

αλγορίθμου να αποφεύγει πρόωρο εγκλωβισμό αποτελεί σημαντικό πλεονέκτημα, καθώς επιτρέπει την πιο εκτεταμένη διερεύνηση του χώρου λύσεων σε σχέση με απλούστερες τοπικές προσεγγίσεις.

Στο πρόβλημα της αποτύπωσης δείκτη, η προσομοιωμένη ανόπτηση μπορεί να εφαρμοστεί είτε στην επιλογή τίτλων είτε στον προσδιορισμό των βαρών είτε και στα δύο ταυτόχρονα, ανάλογα με τον τρόπο που ορίζεται η έννοια της γειτονικής λύσης. Για παράδειγμα, μια γειτονική λύση μπορεί να προκύπτει από την αντικατάσταση ενός τίτλου με κάποιον άλλο, από μικρή τροποποίηση των βαρών ή από συνδυασμό των δύο. Η ποιότητα κάθε υποψήφιας λύσης αποτιμάται με βάση κάποιο μέτρο tracking error και, ενδεχομένως, με τη συνυπολογισμένη επίδραση περιορισμών και ποινών. Έτσι, η μέθοδος μπορεί να διαμορφωθεί με τρόπο που να προσαρμόζεται στις ιδιαιτερότητες του προβλήματος.

Ένα από τα σημαντικά πλεονεκτήματα της προσομοιωμένης ανόπτησης είναι η σχετική απλότητα της βασικής της δομής. Σε σύγκριση με πληθυσμιακές μεθόδους, όπως οι Γενετικοί Αλγόριθμοι, η προσομοιωμένη ανόπτηση λειτουργεί συνήθως πάνω σε μία λύση κάθε φορά, γεγονός που απλοποιεί την υλοποίηση και τον έλεγχο της διαδικασίας. Παράλληλα, η δυνατότητα αποδοχής χειρότερων λύσεων της δίνει μια μορφή «ευφύους» τυχαιότητας, η οποία την καθιστά πιο ισχυρή από τις καθαρά άπληστες τοπικές τεχνικές. Για προβλήματα μεσαίας κλίμακας ή για καλά σχεδιασμένους μηχανισμούς γειτονικότητας, η μέθοδος μπορεί να αποδώσει ιδιαίτερα καλά.

Ωστόσο, η αποτελεσματικότητα της προσομοιωμένης ανόπτησης εξαρτάται σε μεγάλο βαθμό από τον σωστό σχεδιασμό του σχήματος ψύξης και του μηχανισμού μετάβασης. Η επιλογή αρχικής θερμοκρασίας, ο ρυθμός μείωσής της, ο αριθμός επαναλήψεων σε κάθε επίπεδο θερμοκρασίας και ο τρόπος δημιουργίας γειτονικών λύσεων επηρεάζουν άμεσα την ποιότητα των αποτελεσμάτων. Αν η ψύξη είναι υπερβολικά γρήγορη, ο αλγόριθμος μπορεί να χάσει το πλεονέκτημα της διερεύνησης και να συμπεριφερθεί σαν απλή τοπική αναζήτηση. Αντίθετα, αν είναι υπερβολικά αργή, το υπολογιστικό κόστος μπορεί να γίνει μεγάλο. Συνεπώς, η πρακτική επιτυχία της μεθόδου εξαρτάται σε σημαντικό βαθμό από την παραμετροποίηση.

Ένας ακόμη περιορισμός της προσομοιωμένης ανόπτησης είναι ότι, ως μονολυτική μέθοδος αναζήτησης, δεν έχει τον πληθυσμιακό χαρακτήρα των Γενετικών Αλγορίθμων. Αυτό σημαίνει ότι σε κάθε χρονική στιγμή η διαδικασία εξετάζει ουσιαστικά μία τρέχουσα λύση και τις τοπικές μεταβολές της, αντί να διατηρεί και να εξελίσσει ταυτόχρονα πολλές διαφορετικές λύσεις. Αν και η πιθανότητα αποδοχής χειρότερων λύσεων ενισχύει την εξερευνητική της ικανότητα, η μέθοδος μπορεί σε ορισμένες περιπτώσεις να προσφέρει μικρότερη ποικιλία αναζήτησης σε σύγκριση με πληθυσμιακές εξελικτικές τεχνικές, ιδίως όταν το πρόβλημα περιλαμβάνει ισχυρά μικτές ή πολύπλοκες δομές απόφασης.

Παρά τα παραπάνω, η προσομοιωμένη ανόπτηση έχει χρησιμοποιηθεί με επιτυχία σε αρκετές εφαρμογές χαρτοφυλακίου και αποτύπωσης δείκτη, ιδίως σε περιπτώσεις όπου το πρόβλημα είναι αρκετά σύνθετο ώστε να δυσκολεύει τις ακριβείς μεθόδους, αλλά επιτρέπει τη διαμόρφωση αποτελεσματικής έννοιας γειτονικότητας.

2.9.4 Αναζήτηση Tabu

Η αναζήτηση Tabu αποτελεί μία από τις σημαντικότερες μεταερευνητικές τεχνικές βελτιστοποίησης και έχει χρησιμοποιηθεί εκτενώς σε προβλήματα συνδυαστικής και μικτής

βελτιστοποίησης, όπου η απλή τοπική αναζήτηση εμφανίζει σαφείς περιορισμούς. Η βασική ιδέα της μεθόδου είναι ότι η αναζήτηση δεν πρέπει να περιορίζεται μόνο στην άμεση βελτίωση της τρέχουσας λύσης, αλλά να διαθέτει μηχανισμούς που της επιτρέπουν να αποφεύγει την επαναλαμβανόμενη επιστροφή σε ήδη εξερευνημένες περιοχές του χώρου λύσεων. Με αυτόν τον τρόπο, η αναζήτηση Tabu επιχειρεί να υπερβεί ένα από τα βασικά προβλήματα των κλασικών τοπικών μεθόδων, δηλαδή τον εγκλωβισμό σε τοπικά βέλτιστα ή την κυκλική περιπλάνηση ανάμεσα σε παρόμοιες λύσεις (Glover, 1989; Glover, 1990).

Η μέθοδος βασίζεται στη διατήρηση μιας ειδικής μνήμης αναζήτησης, γνωστής ως *tabu list*. Η λίστα αυτή περιέχει πληροφορία για πρόσφατες κινήσεις ή χαρακτηριστικά λύσεων που έχουν ήδη εξεταστεί και οι οποίες, για ορισμένο χρονικό διάστημα, χαρακτηρίζονται ως «απαγορευμένες». Έτσι, όταν ο αλγόριθμος εξετάζει νέες υποψήφιες λύσεις, αποκλείει προσωρινά εκείνες που θα τον οδηγούσαν πίσω σε περιοχές που επισκέφθηκε πολύ πρόσφατα. Η ιδέα αυτή δεν σημαίνει ότι οι λύσεις αυτές είναι οριστικά ανεπιθύμητες, αλλά ότι η προσωρινή απαγόρευση της επιστροφής τους δίνει στην αναζήτηση μεγαλύτερη δυνατότητα διερεύνησης νέων περιοχών (Glover, 1989; Glover, 1990).

Σε αντίθεση με πιο απλές τοπικές τεχνικές, η αναζήτηση Tabu επιτρέπει συχνά τη μετάβαση ακόμη και σε λύσεις που δεν είναι καλύτερες από την τρέχουσα, εφόσον αυτό κρίνεται χρήσιμο για τη συνολική πορεία της αναζήτησης. Η μέθοδος δεν λειτουργεί αποκλειστικά με άπληστο τρόπο, αλλά επιχειρεί να ισορροπήσει ανάμεσα στη βραχυπρόθεσμη βελτίωση και στη στρατηγική αποφυγή πρόωρης σύγκλισης. Με αυτόν τον τρόπο, ο αλγόριθμος αποκτά τη δυνατότητα να εγκαταλείπει περιοχές όπου έχει παγιδευτεί και να κατευθύνεται σε νέες υποψήφιες περιοχές του χώρου λύσεων που μπορεί να αποδειχθούν πιο ελπιδοφόρες (Glover, 1989; Glover, 1990).

Η μνήμη αποτελεί το κατεξοχήν διακριτικό χαρακτηριστικό της αναζήτησης Tabu. Ενώ πολλές άλλες μεταερευνητικές τεχνικές στηρίζονται κυρίως σε στοχαστικούς μηχανισμούς ή σε πληθυσμιακή εξερεύνηση, η αναζήτηση Tabu εισάγει τη δυνατότητα συστηματικής αξιοποίησης της ιστορίας της διαδικασίας. Η πληροφορία για προηγούμενες κινήσεις δεν χρησιμοποιείται μόνο για αποφυγή επανάληψης, αλλά μπορεί να συμβάλει και στη στρατηγική ενίσχυση υποσχόμενων περιοχών ή στην εστίαση της αναζήτησης σε χαρακτηριστικά που έχουν αποδειχθεί γόνιμα. Έτσι, η μέθοδος αποκτά πιο «στρατηγικό» χαρακτήρα σε σύγκριση με καθαρά τοπικές ή αμιγώς τυχαίες μεταερευνητικές διαδικασίες (Glover, 1989; Glover, 1990).

Στο πρόβλημα της αποτύπωσης δείκτη, η αναζήτηση Tabu παρουσιάζει ιδιαίτερο ενδιαφέρον επειδή η φύση του προβλήματος είναι έντονα συνδυαστική, ιδίως όταν περιλαμβάνονται περιορισμοί πληθικότητας και επιλογής τίτλων. Σε ένα τέτοιο πλαίσιο, η τοπική αναζήτηση μπορεί να οριστεί με βάση μικρές μεταβολές στη σύνθεση του χαρτοφυλακίου, όπως η αντικατάσταση ενός τίτλου με κάποιον άλλο, η προσαρμογή επιμέρους βαρών ή ο συνδυασμός των δύο. Η αναζήτηση Tabu μπορεί να διαχειριστεί αποτελεσματικά τέτοιες κινήσεις, αποφεύγοντας την επαναλαμβανόμενη επιστροφή στα ίδια υποσύνολα τίτλων ή στις ίδιες μορφές κατανομής, και έτσι να διερευνήσει πιο αποδοτικά τον χώρο των πιθανών χαρτοφυλακίων (García et al., 2018; Chang et al., 2000).

Ένα βασικό πλεονέκτημα της μεθόδου στο index tracking είναι ότι μπορεί να προσαρμοστεί εύκολα σε προβλήματα όπου η έννοια της γειτονικότητας είναι σαφής. Αν μπορεί να οριστεί με πειστικό τρόπο τι σημαίνει «μικρή μεταβολή» ενός χαρτοφυλακίου, τότε η αναζήτηση Tabu μπορεί να κινηθεί αποτελεσματικά από λύση σε λύση, αξιοποιώντας τη μνήμη της για

να αποφεύγει αδιέξοδες περιοχές. Αυτό είναι ιδιαίτερα χρήσιμο σε προβλήματα όπου οι διακριτές αποφάσεις επιλογής τίτλων είναι καθοριστικές και όπου η τοπική βελτίωση από μόνη της δεν αρκεί για την επίτευξη λύσεων υψηλής ποιότητας (García et al., 2018).

Η αναζήτηση Tabu έχει επίσης το πλεονέκτημα ότι μπορεί να ενσωματώσει πρόσθετα κριτήρια και περιορισμούς με σχετικά ευέλικτο τρόπο. Η ποιότητα κάθε υποψήφιας λύσης μπορεί να αξιολογείται με βάση το *tracking error*, ενώ παράλληλα μπορούν να ληφθούν υπόψη περιορισμοί κόστους συναλλαγών, ορίων βαρών ή αναπροσαρμογής. Επιπλέον, μπορούν να χρησιμοποιηθούν κριτήρια εξαίρεσης από τον κανόνα *tabu*, γνωστά ως *aspiration criteria*, τα οποία επιτρέπουν την αποδοχή μιας κανονικά απαγορευμένης κίνησης εφόσον αυτή οδηγεί σε λύση εξαιρετικά υψηλής ποιότητας. Με τον τρόπο αυτό, η μέθοδος δεν γίνεται άκαμπτη, αλλά διατηρεί ισορροπία ανάμεσα στον έλεγχο της αναζήτησης και στην ευελιξία (Glover, 1989; Glover, 1990).

Παρά τα πλεονεκτήματά της, η αναζήτηση Tabu παρουσιάζει και ορισμένους περιορισμούς. Η απόδοσή της εξαρτάται σημαντικά από τον τρόπο ορισμού της γειτονιάς, από τη δομή της *tabu list* και από τη διάρκεια για την οποία μια κίνηση παραμένει απαγορευμένη. Αν η διάρκεια αυτή είναι πολύ μικρή, ο μηχανισμός μνήμης μπορεί να μην είναι αρκετός για να αποτρέψει την κυκλική συμπεριφορά. Αν είναι πολύ μεγάλη, ο αλγόριθμος μπορεί να γίνει υπερβολικά περιοριστικός και να αποκλείει χρήσιμες διαδρομές αναζήτησης. Επομένως, η επιτυχία της μεθόδου εξαρτάται σε σημαντικό βαθμό από τον προσεκτικό σχεδιασμό των κανόνων λειτουργίας της (Glover, 1989; Glover, 1990).

Ένας ακόμη περιορισμός είναι ότι η αναζήτηση Tabu, όπως και άλλες μονολυτικές μεταερευτικές τεχνικές, δεν διαθέτει τον πληθυσμιακό χαρακτήρα των Γενετικών Αλγορίθμων. Δηλαδή, εξελίσσει κάθε φορά ουσιαστικά μία κεντρική πορεία αναζήτησης, αντί να διατηρεί ταυτόχρονα πολλαπλές εναλλακτικές λύσεις. Παρότι η μνήμη της της δίνει σημαντική στρατηγική ισχύ, η έλλειψη πληθυσμιακής ποικιλίας μπορεί σε ορισμένα προβλήματα να την καταστήσει λιγότερο ευέλικτη από εξελικτικές μεθόδους, ιδιαίτερα όταν ο χώρος λύσεων είναι εξαιρετικά μεγάλος και η ταυτόχρονη διερεύνηση πολλών περιοχών κρίνεται σημαντική (Glover, 1989; García et al., 2018).

2.9.5 Threshold Accepting

Η μέθοδος Threshold Accepting αποτελεί μία μεταερευτική τεχνική βελτιστοποίησης που αναπτύχθηκε ως εναλλακτική προσέγγιση σε στοχαστικές μεθόδους όπως η προσομοιωμένη απόπτηση (Dueck & Scheuer, 1990). Η βασική της φιλοσοφία στηρίζεται στην ιδέα ότι, κατά τη διάρκεια της αναζήτησης, δεν είναι πάντοτε σκόπιμο να γίνονται αποδεκτές μόνο οι λύσεις που βελτιώνουν άμεσα την αντικειμενική συνάρτηση. Αντίθετα, μπορεί να είναι χρήσιμο να επιτρέπεται και η αποδοχή ελαφρώς χειρότερων λύσεων, αρκεί η επιδείνωση που προκαλούν να μην υπερβαίνει ένα προκαθορισμένο όριο, το οποίο αποκαλείται *threshold*. Με τον τρόπο αυτό, η μέθοδος επιδιώκει να αποφεύγει τον πρόωρο εγκλωβισμό σε τοπικά βέλτιστα και να διατηρεί μεγαλύτερη ελευθερία διερεύνησης του χώρου λύσεων (Dueck & Scheuer, 1990).

Η λογική της Threshold Accepting παρουσιάζει ομοιότητες με την προσομοιωμένη απόπτηση, καθώς και οι δύο μέθοδοι επιτρέπουν τη μετάβαση σε χειρότερες λύσεις. Η βασική τους διαφορά έγκειται στον τρόπο με τον οποίο λαμβάνεται η απόφαση αποδοχής.

Στην προσομοιωμένη ανόπτηση η αποδοχή χειρότερων λύσεων βασίζεται σε στοχαστικό κανόνα, δηλαδή σε πιθανότητα που εξαρτάται από το μέγεθος της επιδείνωσης και από τη θερμοκρασία του συστήματος. Αντίθετα, στο Threshold Accepting η απόφαση είναι ντετερμινιστική: μια χειρότερη λύση γίνεται αποδεκτή μόνο αν η υποβάθμιση της ποιότητάς της δεν υπερβαίνει το τρέχον όριο αποδοχής. Αυτή η διαφοροποίηση δίνει στη μέθοδο έναν πιο άμεσο και ελεγχόμενο χαρακτήρα (Dueck & Scheuer, 1990; Eglese, 1990).

Η λειτουργία της μεθόδου είναι σχετικά απλή στη βασική της μορφή. Ξεκινώντας από μια αρχική λύση, ο αλγόριθμος παράγει διαδοχικά γειτονικές υποψήφιες λύσεις. Αν η νέα λύση είναι καλύτερη από την τρέχουσα, γίνεται αποδεκτή. Αν είναι χειρότερη, ελέγχεται αν η διαφορά στην ποιότητά της βρίσκεται εντός του προκαθορισμένου *threshold*. Αν ναι, η λύση αποδέχεται προσωρινά, παρά τη χειρότερη αντικειμενική της τιμή. Καθώς η διαδικασία εξελίσσεται, το *threshold* μειώνεται σταδιακά, με αποτέλεσμα ο αλγόριθμος να γίνεται όλο και πιο αυστηρός. Έτσι, στα αρχικά στάδια επιτρέπεται ευρύτερη διερεύνηση, ενώ στα τελικά στάδια η αναζήτηση γίνεται πιο τοπική και πιο προσανατολισμένη στη βελτίωση (Dueck & Scheuer, 1990).

Ένα βασικό πλεονέκτημα της μεθόδου είναι η απλότητά της σε σύγκριση με πιο σύνθετες στοχαστικές τεχνικές. Η απουσία πιθανοτικού κανόνα αποδοχής καθιστά τη συμπεριφορά της πιο εύκολα ελέγξιμη και συχνά πιο απλή στην υλοποίηση. Ο αλγόριθμος δεν χρειάζεται να υπολογίζει πιθανότητες αποδοχής ούτε να βασίζεται σε γεννήτριες τυχαίων αποφάσεων για τη διαχείριση χειρότερων λύσεων. Η διαδικασία βασίζεται σε έναν καθαρό και διαφανή κανόνα, ο οποίος μπορεί να προσαρμόζεται σταδιακά μέσα από τη μείωση του *threshold* (Dueck & Scheuer, 1990; Eglese, 1990).

Η μέθοδος αυτή είναι ιδιαίτερα κατάλληλη για προβλήματα στα οποία ο χώρος λύσεων είναι πολύπλοκος και περιλαμβάνει πολλά τοπικά βέλτιστα, αλλά όπου ταυτόχρονα είναι εφικτός ο ορισμός μιας σαφούς έννοιας γειτονικότητας. Στο πρόβλημα της αποτύπωσης δείκτη, μια τέτοια έννοια μπορεί να βασίζεται σε μικρές μεταβολές της σύνθεσης του χαρτοφυλακίου ή των βαρών των επιλεγμένων τίτλων. Για παράδειγμα, η αντικατάσταση ενός τίτλου, η τροποποίηση των βαρών ή ένας συνδυασμός τέτοιων αλλαγών μπορούν να θεωρηθούν γειτονικές κινήσεις. Μέσα σε αυτό το πλαίσιο, το Threshold Accepting μπορεί να κινηθεί αποτελεσματικά από λύση σε λύση, επιτρέποντας ελεγχόμενη απομάκρυνση από την άμεση τοπική βελτίωση (Dueck & Scheuer, 1990; Erwin et al., 2023).

Στην αποτύπωση δείκτη, η σημασία της μεθόδου συνδέεται κυρίως με την ικανότητά της να διαχειρίζεται συνδυαστικά προβλήματα όπου οι τοπικά καλύτερες κινήσεις δεν αρκούν για την εύρεση συνολικά ισχυρών λύσεων. Όταν υπάρχουν περιορισμοί πληθικότητας, όρια συμμετοχής και κόστη συναλλαγών, ο χώρος λύσεων γίνεται ιδιαίτερα τραχύς και διακεκομμένος. Σε τέτοιες συνθήκες, μια μέθοδος που επιτρέπει την προσωρινή αποδοχή ελαφρώς χειρότερων λύσεων μπορεί να είναι πολύ χρήσιμη, διότι προσφέρει έναν απλό αλλά αποτελεσματικό τρόπο υπέρβασης τοπικών εμποδίων (Dueck & Scheuer, 1990; Erwin et al., 2023).

Ένα επιπλέον πλεονέκτημα του Threshold Accepting είναι ότι μπορεί να προσαρμοστεί με ευκολία σε διαφορετικούς τύπους αντικειμενικών συναρτήσεων και περιορισμών. Η αξιολόγηση των λύσεων μπορεί να βασίζεται σε μέτρα *tracking error*, ενώ οι περιορισμοί μπορούν να ενσωματωθούν είτε άμεσα είτε μέσω ποινών στην τιμή της συνάρτησης αξιολόγησης. Παράλληλα, το ίδιο το *threshold* μπορεί να σχεδιαστεί έτσι ώστε να μειώνεται με διαφορετικούς ρυθμούς, ανάλογα με τα χαρακτηριστικά του προβλήματος και

τη συμπεριφορά της αναζήτησης. Έτσι, η μέθοδος αποκτά έναν βαθμό ευελιξίας που την καθιστά χρήσιμη σε πρακτικές εφαρμογές (Dueck & Scheuer, 1990; Gilli et al., 2010).

Παρόλα αυτά, η αποτελεσματικότητα της μεθόδου εξαρτάται σε σημαντικό βαθμό από τον ορθό καθορισμό του αρχικού *threshold* και του τρόπου μείωσής του. Αν το *threshold* είναι υπερβολικά μικρό από την αρχή, η μέθοδος θα συμπεριφέρεται περίπου σαν κλασική τοπική αναζήτηση και θα χάνει το βασικό της πλεονέκτημα. Αντίθετα, αν είναι υπερβολικά μεγάλο, υπάρχει κίνδυνος να αποδέχεται πάρα πολλές κατώτερες λύσεις και να περιπλανιέται άσκοπα στον χώρο αναζήτησης. Επομένως, όπως συμβαίνει και με άλλες μεταερευτικές προσεγγίσεις, η παραμετροποίηση αποτελεί κρίσιμο στοιχείο για την επιτυχημένη εφαρμογή της (Dueck & Scheuer, 1990; Eglese, 1990).

Επιπλέον, η μέθοδος διατηρεί τον μονολυτικό χαρακτήρα των τεχνικών τοπικής αναζήτησης, καθώς σε κάθε χρονική στιγμή διαχειρίζεται ουσιαστικά μία πορεία λύσης και όχι έναν πληθυσμό λύσεων. Αυτό σημαίνει ότι, παρότι μπορεί να υπερβεί τοπικά βέλτιστα μέσω ελεγχόμενης αποδοχής χειρότερων λύσεων, δεν διαθέτει τη συλλογική διερευνητική δύναμη πληθυσμιακών μεθόδων όπως οι Γενετικοί Αλγόριθμοι. Σε προβλήματα όπου η ταυτόχρονη εξέταση πολλών εναλλακτικών περιοχών του χώρου λύσεων είναι καθοριστική, αυτό μπορεί να αποτελεί σχετικό μειονέκτημα (Dueck & Scheuer, 1990; Erwin et al., 2023).

2.9.6 Βελτιστοποίηση σμήνους σωματιδίων

Η βελτιστοποίηση σμήνους σωματιδίων αποτελεί μία ακόμη σημαντική μεταερευτική τεχνική, η οποία έχει χρησιμοποιηθεί σε πλήθος προβλημάτων συνεχούς και μικτής βελτιστοποίησης. Η μέθοδος αυτή ανήκει στις τεχνικές εμπνευσμένες από τη φύση και βασίζεται στην παρατήρηση συλλογικών συμπεριφορών που εμφανίζονται σε βιολογικά συστήματα, όπως η κίνηση σμηνών πτηνών ή κοπαδιών ψαριών. Η βασική ιδέα είναι ότι μια ομάδα απλών μονάδων, οι οποίες αλληλεπιδρούν μεταξύ τους και με το περιβάλλον τους, μπορεί να οδηγηθεί συλλογικά σε ιδιαίτερα αποτελεσματική εξερεύνηση ενός χώρου αναζήτησης (Kennedy & Eberhart, 1995; Cheng & Jin, 2015).

Στο πλαίσιο της μεθόδου, κάθε υποψήφια λύση αναπαρίσταται ως σωματίδιο που κινείται μέσα στον χώρο λύσεων. Η θέση του σωματιδίου αντιστοιχεί σε μια συγκεκριμένη λύση του προβλήματος, ενώ η κίνησή του καθορίζεται από έναν μηχανισμό ενημέρωσης που λαμβάνει υπόψη τόσο την ατομική του εμπειρία όσο και τη συλλογική εμπειρία του πληθυσμού. Κάθε σωματίδιο «θυμάται» την καλύτερη θέση που έχει επισκεφθεί μέχρι εκείνη τη στιγμή, ενώ παράλληλα επηρεάζεται και από την καλύτερη θέση που έχει εντοπίσει ολόκληρο το σμήνος ή μια τοπική γειτονιά του. Μέσω αυτής της διπλής επιρροής, η μέθοδος συνδυάζει ατομική αναζήτηση και κοινωνική μάθηση (Kennedy & Eberhart, 1995; Cheng & Jin, 2015).

Ένα από τα πιο χαρακτηριστικά στοιχεία της βελτιστοποίησης σμήνους σωματιδίων είναι ότι η διαδικασία αναζήτησης δεν βασίζεται ούτε σε τελεστές αναπαραγωγής, όπως στους Γενετικούς Αλγόριθμους, ούτε σε αποδοχή χειρότερων λύσεων, όπως στην προσομοιωμένη απόκτηση και στο Threshold Accepting. Αντίθετα, η εξέλιξη του πληθυσμού προκύπτει από τη διαρκή αναπροσαρμογή των θέσεων και των ταχυτήτων των σωματιδίων. Κάθε νέα θέση επηρεάζεται από την προηγούμενη πορεία του σωματιδίου, από την καλύτερη προσωπική του επίδοση και από τη συλλογικά καλύτερη γνωστή λύση. Έτσι, η μέθοδος συνδυάζει τη

διερεύνηση του χώρου λύσεων με τη σταδιακή σύγκλιση προς περιοχές που φαίνονται ιδιαίτερα υποσχόμενες (Kennedy & Eberhart, 1995; Cheng & Jin, 2015).

Η απλότητα της βασικής της δομής αποτελεί ένα από τα βασικά πλεονεκτήματα της μεθόδου. Η βελτιστοποίηση σμήνους σωματιδίων απαιτεί σχετικά λίγες παραμέτρους σε σύγκριση με άλλες μεταερευτικές τεχνικές και η λογική της είναι αρκετά διαφανής. Παράλληλα, η πληθυσμιακή της φύση επιτρέπει την ταυτόχρονη εξερεύνηση πολλών διαφορετικών περιοχών του χώρου λύσεων, γεγονός που μειώνει τον κίνδυνο πρόωρης σύγκλισης σε τοπικά βέλτιστα. Το χαρακτηριστικό αυτό την καθιστά ελκυστική για προβλήματα όπου η ποιότητα της λύσης εξαρτάται από την ευρύτερη και όχι μόνο τοπική αναζήτηση (Cheng & Jin, 2015).

Στη βελτιστοποίηση χαρτοφυλακίου, η μέθοδος έχει χρησιμοποιηθεί κυρίως σε προβλήματα όπου οι μεταβλητές απόφασης έχουν συνεχή μορφή, όπως στον προσδιορισμό των βαρών των τίτλων. Σε τέτοιες περιπτώσεις, κάθε σωματίδιο μπορεί να αναπαριστά ένα διάνυμα βαρών και η εξέλιξή του να αποσκοπεί στη σταδιακή βελτίωση κάποιου μέτρου απόδοσης, κινδύνου ή απόκλισης από benchmark. Η ομαλή αυτή αντιστοίχιση ανάμεσα στις θέσεις των σωματιδίων και στις αριθμητικές τιμές των βαρών καθιστά τη μέθοδο ιδιαίτερα φυσική σε προβλήματα συνεχούς βελτιστοποίησης (Cura, 2009; Zhu et al., 2011).

Στο πρόβλημα της αποτύπωσης δείκτη, η βελτιστοποίηση σμήνους σωματιδίων παρουσιάζει ενδιαφέρον κυρίως όταν το βάρος δίνεται στον προσδιορισμό των κατανομών κεφαλαίου και λιγότερο στην καθαρά διακριτή επιλογή τίτλων. Όταν το πρόβλημα διατυπώνεται με δεδομένο ή σχετικά σταθερό σύνολο τίτλων, η μέθοδος μπορεί να αναζητήσει αποτελεσματικά βάρη που ελαχιστοποιούν το tracking error. Αντίθετα, όταν το πρόβλημα περιλαμβάνει έντονα συνδυαστικό χαρακτήρα, όπως αυστηρούς cardinality constraints και αποφάσεις ένταξης ή αποκλεισμού τίτλων, η κλασική μορφή της μεθόδου δεν είναι πάντοτε η πιο φυσική επιλογή. Σε αυτές τις περιπτώσεις, απαιτούνται προσαρμογές ή υβριδικές μορφές, ώστε να μπορεί να χειριστεί αποτελεσματικά και διακριτές μεταβλητές (Zhu et al., 2010; Corazza et al., 2013).

Ένα από τα σημαντικά πλεονεκτήματα της μεθόδου είναι ότι συχνά εμφανίζει γρήγορη σύγκλιση σε προβλήματα συνεχούς βελτιστοποίησης. Η συλλογική έλξη των σωματιδίων προς τις καλύτερες γνωστές περιοχές του χώρου λύσεων μπορεί να οδηγήσει σε αποδοτική βελτίωση της ποιότητας των λύσεων μέσα σε σχετικά λίγες επαναλήψεις. Αυτό αποτελεί σημαντικό στοιχείο σε εφαρμογές όπου ο χρόνος υπολογισμού είναι κρίσιμος ή όπου απαιτείται επαναλαμβανόμενη εφαρμογή της μεθόδου σε πολλά πειραματικά σενάρια (Cheng & Jin, 2015; Zhu et al., 2011).

Ωστόσο, η βελτιστοποίηση σμήνους σωματιδίων παρουσιάζει και ορισμένους περιορισμούς. Ένας βασικός περιορισμός είναι ότι η μέθοδος μπορεί να γίνει υπερβολικά συγκεντρωτική γύρω από ορισμένες περιοχές του χώρου λύσεων, ιδίως όταν οι παράμετροι κίνησης των σωματιδίων δεν έχουν ρυθμιστεί κατάλληλα. Σε τέτοιες περιπτώσεις, υπάρχει κίνδυνος πρόωρης σύγκλισης, δηλαδή γρήγορης συγκεντρώσης γύρω από λύσεις που είναι καλές αλλά όχι κατ' ανάγκη οι καλύτερες δυνατές. Επιπλέον, σε προβλήματα με έντονα διακριτούς ή μικτούς περιορισμούς, η μεταφορά της φυσικής λογικής των συνεχών θέσεων και ταχυτήτων δεν είναι πάντοτε άμεση ούτε απλή (Cheng & Jin, 2015; Corazza et al., 2013).

Ένας ακόμη περιορισμός σχετίζεται με τη διαχείριση εφικτότητας. Σε προβλήματα χαρτοφυλακίου, τα βάρη πρέπει συνήθως να ικανοποιούν περιορισμούς όπως το άθροισμα ίσο με τη μονάδα, μη αρνητικότητα, κατώτατα και ανώτατα όρια συμμετοχής, καθώς και ενδεχομένως περιορισμούς πληθικότητας. Η μέθοδος μπορεί να εφαρμοστεί σε τέτοια πλαίσια, αλλά συχνά απαιτεί ειδικούς μηχανισμούς κανονικοποίησης, επιδιόρθωσης ή ποινής, ώστε οι νέες θέσεις των σωματιδίων να αντιστοιχούν σε εφικτές λύσεις. Αυτό σημαίνει ότι, όπως και σε άλλες μεταερευτικές τεχνικές, η επιτυχία της εξαρτάται σε σημαντικό βαθμό από τον τρόπο προσαρμογής της στις ιδιαιτερότητες του προβλήματος (Zhu et al., 2011; Corazza et al., 2013).

2.10 Βιβλιογραφική ανασκόπηση

Η βιβλιογραφία για το πρόβλημα της αποτύπωσης δείκτη έχει εξελιχθεί σημαντικά τις τελευταίες δεκαετίες, ακολουθώντας τη γενικότερη εξέλιξη της θεωρίας χαρτοφυλακίου και της υπολογιστικής βελτιστοποίησης. Στον πυρήνα του προβλήματος βρίσκεται η ανάγκη κατασκευής ενός χαρτοφυλακίου που να αναπαράγει όσο το δυνατόν πιστότερα τη συμπεριφορά ενός δείκτη αναφοράς, χωρίς να απαιτείται κατ' ανάγκη η κατοχή όλων των τίτλων που τον απαρτίζουν. Η σχετική έρευνα ξεκίνησε από πιο κλασικές προσεγγίσεις βασισμένες σε μαθηματικά μοντέλα επιλογής χαρτοφυλακίου και σταδιακά επεκτάθηκε σε ακριβείς, ευρετικές, μεταερευτικές και πιο πρόσφατα υβριδικές ή εναλλακτικές υπολογιστικές τεχνικές.

Οι πρώτες προσεγγίσεις στο index tracking συνδέθηκαν στενά με την κλασική θεωρία χαρτοφυλακίου και με την προσπάθεια κατασκευής παθητικών χαρτοφυλακίων που να παρακολουθούν έναν δείκτη με όσο το δυνατόν μικρότερη απόκλιση. Ο Rudd (1980) υπήρξε μία από τις πρώτες κεντρικές αναφορές στο πεδίο, εξετάζοντας τη λογική της παθητικής διαχείρισης και της επιλογής χαρτοφυλακίων που να μιμούνται τη συμπεριφορά της αγοράς. Λίγα χρόνια αργότερα, οι Meade και Salkin (1990) διατύπωσαν το πρόβλημα σε πλαίσιο τετραγωνικού προγραμματισμού και το συνέδεσαν με την πρακτική ανάπτυξη και συντήρηση ενός equity index fund, αναδεικνύοντας ήδη από νωρίς ότι η διαχείριση ενός χαρτοφυλακίου αποτύπωσης δείκτη είναι πρόβλημα όχι μόνο αρχικής επιλογής αλλά και διαρκούς συντήρησης.

Σημαντικό ορόσημο στην εξέλιξη της βιβλιογραφίας αποτέλεσε η εργασία των Beasley, Meade και Chang (2003), η οποία θεωρείται μία από τις πιο επιδραστικές μελέτες στο πεδίο. Οι συγγραφείς διατύπωσαν το index tracking ως πρόβλημα βελτιστοποίησης που περιλαμβάνει ρητά περιορισμούς πληθικότητας και κόστους συναλλαγών και πρότειναν εξελικτική ευρετική μέθοδο για την αντιμετώπισή του. Η συμβολή της συγκεκριμένης εργασίας ήταν διπλή: αφενός συστηματοποίησε μεγάλο μέρος της μέχρι τότε βιβλιογραφίας, αφετέρου ανέδειξε με καθαρό τρόπο τη δυσκολία των ρεαλιστικών εκδοχών του προβλήματος και τη σημασία των εξελικτικών τεχνικών στην επίλυσή του.

Η βιβλιογραφία δείχνει επίσης ότι, καθώς το πρόβλημα έγινε πιο ρεαλιστικό, η έρευνα μετακινήθηκε από απλές continuous formulations προς διατυπώσεις που ενσωματώνουν σύνθετους περιορισμούς. Οι Canakgoz και Beasley (2009) πρότειναν προσεγγίσεις μικτού αέριου προγραμματισμού για index tracking και enhanced indexation, αναδεικνύοντας ότι το πρόβλημα μπορεί να αντιμετωπιστεί με ακριβείς τεχνικές όταν η διατύπωση είναι

επαρκώς ελεγχόμενη, αλλά και ότι η υπολογιστική δυσκολία αυξάνεται σημαντικά όταν ενσωματώνονται πρακτικοί περιορισμοί. Στο ίδιο πνεύμα, οι Scozzari, Tardella, Paterlini και Krink (2013) εξέτασαν το πρόβλημα υπό περιορισμούς UCITS, συνδυάζοντας ακριβείς και ευριστικές προσεγγίσεις και δείχνοντας ότι η ανάγκη συμμόρφωσης με κανονιστικά και επενδυτικά όρια καθιστά το πρόβλημα ακόμη πιο απαιτητικό.

Παράλληλα με τις ακριβείς μεθόδους, αναπτύχθηκε έντονα και η βιβλιογραφία των μεταερευνητικών τεχνικών. Οι Gilli και Këllezi (2002) διερεύνησαν την απόδοση της threshold accepting heuristic στο index tracking, προσφέροντας μία από τις σημαντικές πρώιμες μεταερευνητικές εφαρμογές στο πεδίο. Η εργασία αυτή είναι χαρακτηριστική της στροφής της βιβλιογραφίας προς τεχνικές που μπορούν να διαχειριστούν καλύτερα το μεγάλο μέγεθος του χώρου λύσεων και την ύπαρξη πολλών τοπικών βέλτιστων, ιδιαίτερα όταν ο αριθμός των διαθέσιμων τίτλων είναι υψηλός και το πρόβλημα δεν είναι πρακτικά εφικτό να επιλυθεί με εξαντλητικές τεχνικές.

Ιδιαίτερα ισχυρή είναι και η γραμμή της βιβλιογραφίας που εστιάζει ειδικά στους Γενετικούς Αλγόριθμους. Οι Oh, Kim και Min (2005) πρότειναν χρήση Genetic Algorithm για υποστήριξη της βελτιστοποίησης index funds, αναδεικνύοντας τη δυνατότητα των GAs να ενσωματώνουν ταυτόχρονα επιλογή τίτλων και προσδιορισμό βαρών. Αργότερα, οι Ni και Wang (2013) παρουσίασαν Pareto-efficient genetic algorithm για stock index tracking, συνδέοντας πιο άμεσα τη γενετική αναζήτηση με πολυκριτηριακή λογική και με την ανάγκη ισορροπίας ανάμεσα σε tracking performance και πρακτικούς περιορισμούς. Οι εργασίες αυτές είναι ιδιαίτερα σημαντικές για την παρούσα μελέτη, καθώς επιβεβαιώνουν ότι οι Γενετικοί Αλγόριθμοι αποτελούν μία καθιερωμένη και ερευνητικά ώριμη επιλογή για το συγκεκριμένο πρόβλημα.

Η βιβλιογραφία των τελευταίων ετών δείχνει επίσης ότι το ερευνητικό ενδιαφέρον έχει στραφεί σε όλο και πιο ολοκληρωμένες διατυπώσεις του προβλήματος. Ο Moeini (2022) πρότεινε προσέγγιση συνεχής βελτιστοποίησης για το index tracking, επιχειρώντας να βελτιώσει την αποτελεσματικότητα μαθηματικών τεχνικών σε πιο σύνθετες διατυπώσεις. Την ίδια περίοδο, οι Yuen, Ng, Leung και Che (2022) παρουσίασαν μεταερευνητικό περιβάλλον για index tracking με πρακτικούς περιορισμούς, ενσωματώνοντας ταυτόχρονα αραιότητα, όρια βαρών, τέλη συναλλαγών, διαφοροποίηση και άλλες ρεαλιστικές συνθήκες. Η κατεύθυνση αυτή είναι ενδεικτική της ωρίμανσης του πεδίου: το index tracking δεν μελετάται πλέον ως αφηρημένο πρόβλημα αντιγραφής δείκτη, αλλά ως ολοκληρωμένο πρόβλημα επενδυτικής απόφασης με πρακτικές, κανονιστικές και υπολογιστικές διαστάσεις.

Ενδιαφέρουσα εξέλιξη αποτελεί και η εμφάνιση εναλλακτικών υπολογιστικών προσεγγίσεων πέρα από τις κλασικές μεταερευνητικές. Οι Zheng, Zhu και Li (2020) πρότειναν στοχαστική νευρωνική προσέγγιση για index tracking με περιορισμούς πληθικότητας, αντιμετωπίζοντας το πρόβλημα μέσω αναπαραμετροποίησης και συνδετικής διατύπωσης. Η εργασία αυτή δείχνει ότι η σύγχρονη βιβλιογραφία διευρύνεται σταδιακά και προς πιο σύγχρονα υπολογιστικά εργαλεία, χωρίς όμως να αναιρεί τη σημασία των μεταερευνητικών μεθόδων. Αντίθετα, επιβεβαιώνει ότι η δυσκολία του προβλήματος είναι τέτοια, ώστε συνεχίζει να κινητοποιεί την ανάπτυξη νέων πλαισίων επίλυσης.

Από τη συνολική επισκόπηση της βιβλιογραφίας προκύπτουν ορισμένα σαφή συμπεράσματα. Πρώτον, το index tracking έχει εξελιχθεί από ένα σχετικά απλό πρόβλημα παθητικής διαχείρισης σε ένα σύνθετο πρόβλημα βελτιστοποίησης, στο οποίο συνυπάρχουν

συνεχείς και διακριτές μεταβλητές, καθώς και ποικίλοι ρεαλιστικοί περιορισμοί. Δεύτερον, οι ακριβείς μέθοδοι διατηρούν θεωρητική σημασία, αλλά η αποτελεσματικότητά τους περιορίζεται όσο αυξάνεται η ρεαλιστικότητα της διατύπωσης. Τρίτον, οι μεταερευνητικές τεχνικές και ειδικότερα οι Γενετικοί Αλγόριθμοι έχουν αποκτήσει κεντρικό ρόλο, καθώς προσφέρουν την απαιτούμενη ευελιξία για τη διαχείριση περιορισμών πληθικότητας, κόστη συναλλαγών και σύνθετων περιορισμών βαρών. Τέταρτον, η πιο πρόσφατη βιβλιογραφία κινείται προς ολοκληρωμένα, πρακτικά και συχνά υβριδικά πλαίσια, γεγονός που ενισχύει τη σημασία της ανάπτυξης προσαρμοσμένων αλγοριθμικών λύσεων για πραγματικά δεδομένα.

Με βάση τα παραπάνω, η παρούσα εργασία τοποθετείται στο σημείο τομής ανάμεσα στην εξελικτική υπολογιστική και στη ρεαλιστική χρηματοοικονομική μοντελοποίηση. Η επιλογή των Γενετικών Αλγορίθμων ως βασικής μεθοδολογίας δεν αποτελεί απλώς τεχνική προτίμηση, αλλά ευθυγραμμίζεται με μία ισχυρή και διαρκώς εξελισσόμενη γραμμή της διεθνούς βιβλιογραφίας, η οποία έχει δείξει ότι τα προβλήματα αποτύπωσης δείκτη με περιορισμούς πληθικότητας, ορίων βαρών και κόστους συναλλαγών είναι ιδιαίτερα κατάλληλα για εξελικτικές και μεταερευνητικές προσεγγίσεις.

ΚΕΦΑΛΑΙΟ 3: ΠΡΟΤΕΙΝΟΜΕΝΗ ΜΕΘΟΔΟΣ ΕΠΙΛΥΣΗΣ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

3.1 Εισαγωγή

Στο παρόν κεφάλαιο η εργασία μεταβαίνει από το επίπεδο της θεωρητικής και βιβλιογραφικής ανάλυσης στο επίπεδο της διαμόρφωσης της προτεινόμενης μεθοδολογίας. Ειδικότερα, σκοπός του κεφαλαίου είναι η αυστηρή διατύπωση του προβλήματος της αποτύπωσης δείκτη ως προβλήματος βελτιστοποίησης, καθώς και η αναλυτική παρουσίαση του προτεινόμενου Γενετικού Αλγορίθμου που θα χρησιμοποιηθεί για την επίλυσή του. Με άλλα λόγια, το κεφάλαιο αυτό αποτελεί τον πυρήνα της μεθοδολογικής συμβολής της εργασίας, αφού εδώ ορίζονται με σαφήνεια οι μεταβλητές, οι παράμετροι, η αντικειμενική συνάρτηση, οι περιορισμοί και ο τρόπος με τον οποίο όλα αυτά ενσωματώνονται σε ένα ενιαίο αλγοριθμικό πλαίσιο.

Η μετάβαση αυτή είναι ιδιαίτερα σημαντική, διότι το πρόβλημα της αποτύπωσης δείκτη δεν μπορεί να αντιμετωπιστεί μόνο με αφηρημένους θεωρητικούς όρους. Για να καταστεί επιλύσιμο, πρέπει να διατυπωθεί σε σαφή μαθηματική μορφή, ώστε να είναι δυνατόν να εφαρμοστεί πάνω του μια οργανωμένη διαδικασία αναζήτησης λύσεων. Η διατύπωση αυτή πρέπει να αποτυπώνει τόσο τον βασικό στόχο του προβλήματος, δηλαδή την ελαχιστοποίηση της απόκλισης μεταξύ χαρτοφυλακίου και δείκτη, όσο και τους ρεαλιστικούς περιορισμούς που χαρακτηρίζουν την επενδυτική πρακτική. Κατά συνέπεια, το παρόν κεφάλαιο δεν περιορίζεται σε μια απλή περιγραφή του αλγορίθμου, αλλά αναπτύσσει πρώτα το αναγκαίο μαθηματικό υπόβαθρο πάνω στο οποίο ο αλγόριθμος θα μπορέσει να λειτουργήσει αποτελεσματικά.

Στην αρχή του κεφαλαίου θα παρουσιαστεί η μαθηματική διατύπωση του προβλήματος. Θα οριστούν το σύνολο των διαθέσιμων τίτλων, η χρονική περίοδος ανάλυσης, οι αποδόσεις των μετοχών και του δείκτη, οι μεταβλητές απόφασης και οι βασικές παράμετροι του μοντέλου. Στη συνέχεια θα διατυπωθεί η αντικειμενική συνάρτηση και θα αναλυθούν οι βασικοί περιορισμοί, όπως ο περιορισμός προϋπολογισμού, ο περιορισμός πληθικότητας, τα κατώτατα και ανώτατα όρια συμμετοχής και, όπου απαιτείται, οι περιορισμοί που συνδέονται με το κόστος συναλλαγών. Η διατύπωση αυτή είναι απαραίτητη ώστε το πρόβλημα να αποκτήσει σαφή και ελέγξιμη μορφή, κατάλληλη για την εφαρμογή του προτεινόμενου αλγορίθμου.

Στη συνέχεια, το ενδιαφέρον θα στραφεί στον σχεδιασμό του Γενετικού Αλγορίθμου. Σε αυτό το σημείο θα αναλυθεί ο τρόπος αναπαράστασης των λύσεων, δηλαδή η μορφή του χρωμοσώματος, ο τρόπος δημιουργίας του αρχικού πληθυσμού, η συνάρτηση καταλληλότητας και οι βασικοί γενετικοί τελεστές, όπως η επιλογή, η διασταύρωση και η μετάλλαξη. Ιδιαίτερη έμφαση θα δοθεί και στον τρόπο διαχείρισης των περιορισμών, καθώς η αποτελεσματική αντιμετώπισή τους αποτελεί κρίσιμο στοιχείο για την ποιότητα και τη ρεαλιστικότητα των παραγόμενων λύσεων. Η παρουσίαση αυτή θα ολοκληρωθεί με ψευδοκώδικα και διάγραμμα ροής, ώστε να αποδοθεί με σαφήνεια η συνολική λογική της προτεινόμενης μεθόδου.

Είναι σημαντικό να τονιστεί ότι ο προτεινόμενος αλγόριθμος δεν αντιμετωπίζεται ως γενικής χρήσης εργαλείο που εφαρμόζεται μηχανιστικά στο πρόβλημα, αλλά ως μέθοδος ειδικά προσαρμοσμένη στις απαιτήσεις της αποτύπωσης δείκτη.

3.2 Διατύπωση του προβλήματος ως προβλήματος βελτιστοποίησης

Το πρόβλημα της αποτύπωσης δείκτη μπορεί να διατυπωθεί αυστηρά ως πρόβλημα βελτιστοποίησης, στο οποίο στόχος είναι η κατασκευή ενός χαρτοφυλακίου που να αναπαράγει όσο το δυνατόν πιστότερα την πορεία ενός χρηματιστηριακού δείκτη αναφοράς. Η διατύπωση αυτή είναι απαραίτητη, διότι επιτρέπει τη συστηματική σύνδεση του χρηματοοικονομικού προβλήματος με τις τεχνικές υπολογιστικής επίλυσης που θα χρησιμοποιηθούν στη συνέχεια. Μέσω της μαθηματικής του μορφής, το πρόβλημα καθίσταται σαφές ως προς τις μεταβλητές απόφασης, τον στόχο που επιδιώκεται και τους περιορισμούς που πρέπει να ικανοποιούνται.

Σε γενικούς όρους, θεωρείται ότι υπάρχει ένα σύνολο διαθέσιμων μετοχών, από το οποίο πρέπει να επιλεγεί ένα υποσύνολο τίτλων για τη συγκρότηση του χαρτοφυλακίου. Για κάθε επιλεγμένο τίτλο πρέπει να προσδιοριστεί το αντίστοιχο βάρος συμμετοχής του, έτσι ώστε η συνολική απόδοση του χαρτοφυλακίου να προσεγγίζει όσο το δυνατόν καλύτερα την απόδοση του δείκτη. Η προσέγγιση αυτή αποτυπώνει τη βασική λογική της μερικής αναπαραγωγής δείκτη, όπου δεν απαιτείται η πλήρης κατοχή όλων των τίτλων του benchmark, αλλά η επιλογή μιας σύνθεσης που να επιτυγχάνει υψηλή πιστότητα παρακολούθησης με περιορισμένο αριθμό θέσεων.

Η μαθηματική φύση του προβλήματος είναι μικτή. Από τη μία πλευρά, υπάρχουν διακριτές αποφάσεις που σχετίζονται με το αν ένας τίτλος θα συμπεριληφθεί ή όχι στο χαρτοφυλάκιο. Από την άλλη πλευρά, υπάρχουν συνεχείς αποφάσεις που αφορούν το ποσοστό του συνολικού κεφαλαίου που θα επενδυθεί σε κάθε επιλεγμένο τίτλο. Η συνύπαρξη αυτών των δύο ειδών μεταβλητών είναι ένας από τους βασικούς λόγους που το πρόβλημα καθίσταται υπολογιστικά απαιτητικό. Δεν πρόκειται δηλαδή μόνο για πρόβλημα κατανομής βαρών, ούτε μόνο για πρόβλημα επιλογής συνόλου τίτλων, αλλά για συνδυασμό των δύο αυτών διαστάσεων σε ένα ενιαίο πλαίσιο.

Κεντρικό στοιχείο της διατύπωσης είναι η αντικειμενική συνάρτηση. Στο πρόβλημα της αποτύπωσης δείκτη, η συνάρτηση αυτή πρέπει να εκφράζει την απόκλιση μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη αναφοράς. Η απόκλιση αυτή μετράται μέσω ενός κατάλληλου μέτρου tracking error, το οποίο λειτουργεί ως βασικό κριτήριο αξιολόγησης της ποιότητας μιας λύσης. Συνεπώς, το πρόβλημα διατυπώνεται ως πρόβλημα ελαχιστοποίησης: ζητείται το χαρτοφυλάκιο εκείνο που, υπό τους δεδομένους περιορισμούς, ελαχιστοποιεί το σφάλμα αποτύπωσης.

Πέρα από τον βασικό αυτό στόχο, η διατύπωση του προβλήματος πρέπει να περιλαμβάνει και ένα σύνολο περιορισμών που αντανakλούν τις πραγματικές συνθήκες της επενδυτικής απόφασης. Πρώτος και θεμελιώδης περιορισμός είναι ο περιορισμός προϋπολογισμού, σύμφωνα με τον οποίο το συνολικό επενδυμένο κεφάλαιο δεν μπορεί να υπερβαίνει το διαθέσιμο ποσό. Στην τυπική του μορφή, αυτό σημαίνει ότι το άθροισμα των βαρών των τίτλων του χαρτοφυλακίου πρέπει να είναι ίσο με τη μονάδα. Παράλληλα, επιβάλλεται συνήθως ότι τα βάρη είναι μη αρνητικά, ώστε το χαρτοφυλάκιο να αντιστοιχεί σε long-only επενδυτική στρατηγική.

Ένας δεύτερος κρίσιμος περιορισμός είναι ο περιορισμός πληθικότητας. Ο περιορισμός αυτός επιβάλλει ότι το χαρτοφυλάκιο μπορεί να περιλαμβάνει το πολύ ή ακριβώς έναν συγκεκριμένο αριθμό τίτλων. Η παρουσία του έχει μεγάλη σημασία, διότι αντικατοπτρίζει την πρακτική ανάγκη διατήρησης ενός διαχειρίσιμου χαρτοφυλακίου και ταυτόχρονα καθιστά το πρόβλημα σημαντικά πιο δύσκολο. Πλέον, η μέθοδος επίλυσης δεν αρκεί να βρει μόνο κατάλληλα βάρη, αλλά πρέπει και να επιλέξει ποιοι τίτλοι θα έχουν μη μηδενική συμμετοχή.

Πέρα από τους δύο αυτούς βασικούς περιορισμούς, η διατύπωση περιλαμβάνει και όρια συμμετοχής των τίτλων. Συγκεκριμένα, μπορεί να απαιτείται κάθε επιλεγμένος τίτλος να κατέχει βάρος όχι μικρότερο από ένα ελάχιστο επιτρεπτό επίπεδο και όχι μεγαλύτερο από ένα μέγιστο. Τα όρια αυτά αποτρέπουν τόσο τη δημιουργία αμελητέων θέσεων όσο και την υπερβολική συγκέντρωση κεφαλαίου σε λίγους τίτλους. Με τον τρόπο αυτό, το χαρτοφυλάκιο παραμένει πιο ρεαλιστικό και καλύτερα ευθυγραμμισμένο με τις πρακτικές ανάγκες της διαχείρισης.

Σε πιο ρεαλιστικές διατυπώσεις του προβλήματος, μπορεί να προστεθεί και περιορισμός κόστους συναλλαγών. Ο περιορισμός αυτός είναι ιδιαίτερα σημαντικός όταν το χαρτοφυλάκιο δεν δημιουργείται από μηδενική βάση, αλλά προκύπτει μέσω αναπροσαρμογής μιας ήδη υπάρχουσας επενδυτικής σύνθεσης. Σε αυτή την περίπτωση, η μετάβαση από το παλαιό στο νέο χαρτοφυλάκιο συνεπάγεται αγορές και πωλήσεις τίτλων, άρα και κόστος. Η ενσωμάτωση αυτού του στοιχείου καθιστά το πρόβλημα πιο ρεαλιστικό, αφού η βέλτιστη λύση δεν είναι εκείνη που απλώς ελαχιστοποιεί το tracking error, αλλά εκείνη που το επιτυγχάνει υπό αποδεκτό επίπεδο συναλλακτικής επιβάρυνσης.

Η χρονική διάσταση του προβλήματος είναι επίσης σημαντική. Για την εκτίμηση της ποιότητας ενός χαρτοφυλακίου απαιτείται η παρατήρηση των αποδόσεων τόσο των τίτλων όσο και του δείκτη σε ένα συγκεκριμένο ιστορικό διάστημα. Η περίοδος αυτή λειτουργεί ως βάση πάνω στην οποία υπολογίζεται το measure tracking error και διαμορφώνεται η συνολική συνάρτηση στόχου. Συνεπώς, το πρόβλημα βελτιστοποίησης δεν είναι ανεξάρτητο από τα δεδομένα, αλλά προσδιορίζεται άμεσα από το εύρος, τη συχνότητα και τη δομή των χρηματοοικονομικών παρατηρήσεων που χρησιμοποιούνται.

Από μεθοδολογική άποψη, η παραπάνω διατύπωση δείχνει ότι το πρόβλημα της αποτύπωσης δείκτη είναι πρόβλημα περιορισμένης βελτιστοποίησης σε μη απλό χώρο λύσεων. Η ταυτόχρονη παρουσία στόχου ελαχιστοποίησης απόκλισης, διακριτών επιλογών, συνεχών βαρών και πολλαπλών επενδυτικών περιορισμών καθιστά τη συνολική δομή ιδιαίτερα απαιτητική. Αυτή ακριβώς η φύση του προβλήματος δικαιολογεί την επιλογή μιας εξελικτικής μεθόδου, όπως ο Γενετικός Αλγόριθμος, η οποία μπορεί να εξερευνήσει αποτελεσματικά τον σύνθετο αυτό χώρο αναζήτησης χωρίς να απαιτεί αυστηρές παραδοχές απλότητας ή κυρτότητας.

3.3 Ορισμός συμβολισμών, παραμέτρων και μεταβλητών

Για την αυστηρή μαθηματική διατύπωση του προβλήματος της αποτύπωσης δείκτη είναι απαραίτητος ο σαφής ορισμός των συμβολισμών, των παραμέτρων και των μεταβλητών απόφασης που θα χρησιμοποιηθούν. Η ύπαρξη ενός συνεκτικού συστήματος συμβολισμών επιτρέπει την ακριβή περιγραφή του προβλήματος, διευκολύνει τη διατύπωση της αντικειμενικής συνάρτησης και των περιορισμών και παρέχει το απαραίτητο υπόβαθρο για τον σχεδιασμό του προτεινόμενου Γενετικού Αλγορίθμου. Στο πλαίσιο αυτό, η παρούσα

ενότητα οργανώνει τα στοιχεία του μοντέλου σε τρεις βασικές κατηγορίες: τα σύνολα και τους δείκτες, τις παραμέτρους του προβλήματος και τις μεταβλητές απόφασης

3.3.1 Σύνολα και δείκτες

Θεωρείται ότι υπάρχει ένα σύνολο διαθέσιμων τίτλων, το οποίο αντιστοιχεί στις μετοχές που μπορούν να χρησιμοποιηθούν για την κατασκευή του χαρτοφυλακίου αποτύπωσης δείκτη. Το σύνολο αυτό συμβολίζεται με $I = \{1, 2, \dots, n\}$, όπου n είναι το συνολικό πλήθος των διαθέσιμων μετοχών. Κάθε στοιχείο $i \in I$ αντιστοιχεί σε έναν συγκεκριμένο τίτλο του επενδυτικού συνόλου.

Παράλληλα, θεωρείται ένα σύνολο χρονικών περιόδων, το οποίο χρησιμοποιείται για την παρατήρηση των αποδόσεων των τίτλων και του δείκτη. Το σύνολο αυτό συμβολίζεται με $T = \{1, 2, \dots, m\}$, όπου m είναι το πλήθος των χρονικών παρατηρήσεων. Κάθε στοιχείο $t \in T$ αντιστοιχεί σε μία χρονική περίοδο, όπως ημέρα, εβδομάδα ή μήνα, ανάλογα με τη συχνότητα των δεδομένων που θα χρησιμοποιηθούν.

Ο δείκτης i χρησιμοποιείται επομένως για να δηλώσει τίτλους, ενώ ο δείκτης t χρησιμοποιείται για να δηλώσει χρονικές περιόδους. Η διάκριση αυτή είναι ουσιώδης, διότι επιτρέπει τη σαφή σύνδεση της διαχρονικής συμπεριφοράς των αποδόσεων με τη δομή του χαρτοφυλακίου.

3.3.2 Παράμετροι του προβλήματος

Οι παράμετροι του προβλήματος είναι μεγέθη γνωστά εκ των προτέρων, τα οποία προκύπτουν από τα διαθέσιμα δεδομένα ή από τις παραδοχές και τους περιορισμούς του μοντέλου.

Έστω r_{it} η απόδοση του τίτλου i κατά τη χρονική περίοδο t . Το μέγεθος αυτό αποτελεί βασική εισροή του μοντέλου, καθώς αποτυπώνει τη διαχρονική συμπεριφορά κάθε μετοχής. Αντίστοιχα, έστω R_t η απόδοση του δείκτη αναφοράς κατά τη χρονική περίοδο t . Το μέγεθος αυτό εκφράζει τη συμπεριφορά του benchmark που το χαρτοφυλάκιο επιδιώκει να αναπαράγει.

Έστω ακόμη K ο μέγιστος επιτρεπτός αριθμός τίτλων που μπορούν να συμμετέχουν στο χαρτοφυλάκιο. Η παράμετρος αυτή συνδέεται με τον περιορισμό πληθικότητας και καθορίζει τη μέγιστη επιτρεπτή πολυπλοκότητα της τελικής λύσης.

Αν l_i και u_i δηλώνουν αντίστοιχα το κατώτατο και το ανώτατο επιτρεπτό βάρος για τον τίτλο i , τότε οι παράμετροι αυτές χρησιμοποιούνται για να διασφαλιστεί ότι οι επιλεγμένοι τίτλοι δεν θα έχουν ούτε αμελητέα ούτε υπερβολικά μεγάλη συμμετοχή στο χαρτοφυλάκιο. Σε απλούστερες διατυπώσεις, τα όρια αυτά μπορεί να θεωρούνται κοινά για όλους τους τίτλους, δηλαδή $l_i = l$ και $u_i = u$ για κάθε $i \in I$.

Σε περιπτώσεις όπου λαμβάνεται υπόψη το κόστος συναλλαγών, μπορεί να οριστεί επιπλέον μια παράμετρος c_i , η οποία εκφράζει το κόστος που συνδέεται με την επένδυση ή την αναπροσαρμογή της θέσης στον τίτλο i . Ανάλογα με τη διατύπωση του προβλήματος, το κόστος αυτό μπορεί να είναι γραμμικό, σταθερό ή μικτό. Εναλλακτικά, μπορεί να

χρησιμοποιηθεί μια συνολική παράμετρος $C_{\max}$, η οποία να εκφράζει το μέγιστο αποδεκτό επίπεδο κόστους συναλλαγών.

Εφόσον η διατύπωση περιλαμβάνει αναπροσαρμογή ενός ήδη υπάρχοντος χαρτοφυλακίου, μπορεί επίσης να οριστεί η παράμετρος w_i^0 , η οποία εκφράζει το αρχικό ή τρέχον βάρος του τίτλου i πριν από την εφαρμογή της νέας βελτιστοποίησης. Η παράμετρος αυτή είναι χρήσιμη όταν η μετάβαση από την υφιστάμενη στη νέα σύνθεση του χαρτοφυλακίου επηρεάζει το κόστος συναλλαγών ή τους περιορισμούς αναπροσαρμογής.

3.3.3 Μεταβλητές απόφασης

Οι μεταβλητές απόφασης είναι τα μεγέθη που πρέπει να προσδιοριστούν από τη διαδικασία βελτιστοποίησης. Στο πρόβλημα της αποτύπωσης δείκτη, οι βασικές μεταβλητές είναι δύο ειδών: συνεχείς μεταβλητές που εκφράζουν τα βάρη των τίτλων και δυαδικές μεταβλητές που εκφράζουν την επιλογή ή μη των τίτλων.

Έστω w_i το βάρος του τίτλου i στο χαρτοφυλάκιο. Η μεταβλητή αυτή είναι συνεχής και εκφράζει το ποσοστό του συνολικού διαθέσιμου κεφαλαίου που επενδύεται στη συγκεκριμένη μετοχή. Αν $w_i = 0$, τότε ο τίτλος δεν συμμετέχει ουσιαστικά στο χαρτοφυλάκιο. Αν $w_i > 0$, τότε ο τίτλος περιλαμβάνεται με συμμετοχή ανάλογη προς το μέγεθος αυτού του βάρους.

Για την αυστηρή μοντελοποίηση του περιορισμού πληθικότητας, ορίζεται επίσης η δυαδική μεταβλητή z_i , όπου:

$$z_i = \begin{cases} 1, & \text{αν ο τίτλος } i \text{ επιλεγεί στο χαρτοφυλάκιο} \\ 0, & \text{διαφορετικά} \end{cases}$$

Η μεταβλητή αυτή επιτρέπει να διατυπωθεί ρητά ο περιορισμός ότι μόνο περιορισμένος αριθμός τίτλων μπορεί να έχει θετική συμμετοχή στο χαρτοφυλάκιο. Παράλληλα, συνδέει τις αποφάσεις επιλογής τίτλων με τα αντίστοιχα βάρη τους, αφού συνήθως απαιτείται το βάρος w_i να είναι θετικό μόνο όταν $z_i = 1$.

Αν το πρόβλημα επεκτείνεται σε πλαίσιο αναπροσαρμογής, μπορούν να οριστούν και πρόσθετες μεταβλητές, όπως οι μεταβλητές αγοράς και πώλησης τίτλων. Για παράδειγμα, οι μεταβλητές b_i και s_i μπορούν να εκφράζουν αντίστοιχα την αύξηση και τη μείωση της θέσης στον τίτλο i . Οι μεταβλητές αυτές είναι χρήσιμες όταν το κόστος συναλλαγών εξαρτάται ρητά από τη μεταβολή των θέσεων σε σχέση με ένα υφιστάμενο χαρτοφυλάκιο.

3.4 Αντικειμενική συνάρτηση

Η αντικειμενική συνάρτηση αποτελεί το βασικό στοιχείο κάθε προβλήματος βελτιστοποίησης, καθώς αποτυπώνει με ακριβή μαθηματικό τρόπο τον στόχο που επιδιώκεται. Στο πρόβλημα της αποτύπωσης δείκτη, ο στόχος αυτός δεν είναι η γενική μεγιστοποίηση της απόδοσης ούτε η ελαχιστοποίηση του συνολικού κινδύνου με την κλασική έννοια της διασποράς, αλλά η ελαχιστοποίηση της απόκλισης μεταξύ της

απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη αναφοράς. Κατά συνέπεια, η αντικειμενική συνάρτηση πρέπει να εκφράζει κατάλληλα το μέγεθος του σφάλματος αποτύπωσης και να λειτουργεί ως το κριτήριο βάσει του οποίου θα συγκρίνονται οι εναλλακτικές λύσεις.

Για κάθε χρονική περίοδο $t \in T$, η απόδοση του χαρτοφυλακίου μπορεί να εκφραστεί ως το σταθμισμένο άθροισμα των αποδόσεων των επιλεγμένων τίτλων:

$$P_t = \sum_{i=1}^n w_i r_{it} \quad (3.1)$$

όπου P_t είναι η απόδοση του χαρτοφυλακίου κατά την περίοδο t , w_i είναι το βάρος του τίτλου i στο χαρτοφυλάκιο και r_{it} είναι η απόδοση του τίτλου i κατά την ίδια περίοδο. Η σχέση αυτή εκφράζει με απλό και σαφή τρόπο τη συνολική συμπεριφορά του χαρτοφυλακίου ως γραμμικό συνδυασμό των αποδόσεων των τίτλων που το απαρτίζουν.

Ο δείκτης αναφοράς έχει απόδοση R_t κατά την περίοδο t . Επομένως, η απόκλιση μεταξύ χαρτοφυλακίου και δείκτη στην περίοδο αυτή μπορεί να εκφραστεί ως:

$$e_t = P_t - R_t = \sum_{i=1}^n w_i r_{it} - R_t \quad (3.2)$$

όπου e_t είναι το σφάλμα αποτύπωσης κατά την περίοδο t . Το συνολικό πρόβλημα συνίσταται στην κατασκευή ενός χαρτοφυλακίου για το οποίο οι αποκλίσεις e_t παραμένουν όσο το δυνατόν μικρότερες σε όλη τη διάρκεια της εξεταζόμενης περιόδου.

3.4.1 Ορισμός του στόχου βελτιστοποίησης

Ο φυσικός στόχος του προβλήματος είναι η ελαχιστοποίηση ενός κατάλληλου συνολικού μέτρου των αποκλίσεων e_t . Στην παρούσα εργασία, ως βασική μορφή της αντικειμενικής συνάρτησης υιοθετείται η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος αποτύπωσης, δηλαδή:

$$\min TE = \frac{1}{m} \sum_{t=1}^m (\sum_{i=1}^n w_i r_{it} - R_t)^2 \quad (3.3)$$

όπου TE δηλώνει το συνολικό tracking error και m είναι το πλήθος των χρονικών περιόδων. Η μορφή αυτή είναι ιδιαίτερα διαδεδομένη στη σχετική βιβλιογραφία, καθώς τιμωρεί περισσότερο τις μεγαλύτερες αποκλίσεις και παρέχει σαφή και ομαλή μαθηματική διατύπωση του στόχου.

Η επιλογή της τετραγωνικής μορφής είναι σκόπιμη, διότι δεν ενδιαφέρει μόνο η μέση κατεύθυνση της απόκλισης, αλλά και το μέγεθός της. Αν χρησιμοποιούνταν απλό άθροισμα διαφορών, οι θετικές και αρνητικές αποκλίσεις θα μπορούσαν να αλληλοαναιρεθούν, οδηγώντας σε παραπλανητικά μικρή συνολική τιμή. Με την τετραγωνική μορφή, κάθε

απόκλιση συμβάλλει θετικά στο συνολικό σφάλμα, ενώ οι μεγαλύτερες αποκλίσεις επιβαρύνουν δυσανάλογα περισσότερο την αντικειμενική συνάρτηση.

3.4.2 Μέτρο σφάλματος αποτύπωσης

Η παραπάνω αντικειμενική συνάρτηση βασίζεται στο τετραγωνικό μέτρο tracking error, το οποίο είναι συμβατό τόσο με τη θεωρητική λογική του προβλήματος όσο και με την ανάγκη αριθμητικά σταθερής και ερμηνεύσιμης βελτιστοποίησης. Ειδικότερα, η χρήση του μέσου τετραγωνικού σφάλματος επιτρέπει την αποτίμηση του βαθμού στον οποίο το χαρτοφυλάκιο ακολουθεί τον δείκτη σε όλη τη διάρκεια της περιόδου ανάλυσης, χωρίς να εστιάζει μόνο σε μεμονωμένες χρονικές στιγμές.

Σε ορισμένες περιπτώσεις, η ίδια λογική μπορεί να αποδοθεί και χωρίς τον παράγοντα $\frac{1}{m}$, δηλαδή ως:

$$\min \sum_{t=1}^m (\sum_{i=1}^n w_i r_{it} - R_t)^2 \quad (3.4)$$

Οι δύο μορφές είναι ισοδύναμες ως προς τη βέλτιστη λύση, αφού ο παράγοντας $\frac{1}{m}$ είναι σταθερός και δεν επηρεάζει τη θέση του ελάχιστου. Η χρήση του μέσου όρου, ωστόσο, είναι συχνά προτιμότερη, επειδή επιτρέπει πιο εύκολη ερμηνεία της τιμής του tracking error και καλύτερη συγκρισιμότητα μεταξύ διαφορετικών συνόλων δεδομένων.

Αν ζητείται η έκφραση του σφάλματος αποτύπωσης σε μορφή πιο κοντά στην αρχική κλίμακα των αποδόσεων, μπορεί να χρησιμοποιηθεί και η τετραγωνική ρίζα του παραπάνω μέτρου:

$$\min \sqrt{\frac{1}{m} \sum_{t=1}^m (\sum_{i=1}^n w_i r_{it} - R_t)^2} \quad (3.5)$$

Η μορφή αυτή είναι ερμηνευτικά πιο άμεση, αλλά για σκοπούς βελτιστοποίησης η μη ριζική τετραγωνική διατύπωση παραμένει συνήθως προτιμότερη, επειδή είναι απλούστερη και οδηγεί στην ίδια κατάταξη λύσεων.

3.4.3 Αιτιολόγηση της επιλεγμένης αντικειμενικής συνάρτησης

Η επιλογή της συγκεκριμένης αντικειμενικής συνάρτησης βασίζεται σε τρεις κύριους λόγους. Πρώτον, είναι άμεσα συμβατή με τη βασική χρηματοοικονομική λογική του προβλήματος, καθώς μετρά ακριβώς αυτό που επιδιώκει να ελαχιστοποιήσει ένα χαρτοφυλάκιο αποτύπωσης δείκτη: την απόκλιση από τον benchmark. Δεύτερον, η τετραγωνική μορφή τιμωρεί αποτελεσματικά τις μεγάλες αποκλίσεις, κάτι που είναι ιδιαίτερα επιθυμητό σε εφαρμογές παθητικής διαχείρισης, όπου η σταθερή και πιστή παρακολούθηση του δείκτη έχει μεγαλύτερη σημασία από την περιστασιακή εγγύτητα σε

μέσους όρους. Τρίτον, η μαθηματική της μορφή είναι κατάλληλη για ενσωμάτωση σε αλγοριθμικά σχήματα βελτιστοποίησης και επιτρέπει σαφή αξιολόγηση των υποψήφιων λύσεων μέσα στο πλαίσιο του Γενετικού Αλγορίθμου.

Πρέπει επίσης να σημειωθεί ότι η αντικειμενική συνάρτηση εκφράζει μόνο τον βασικό στόχο του προβλήματος και όχι ολόκληρη τη δυσκολία του. Η πραγματική πολυπλοκότητα προκύπτει από τον συνδυασμό της με τους επενδυτικούς περιορισμούς που συνοδεύουν το χαρτοφυλάκιο. Ένα χαρτοφυλάκιο με πολύ μικρό tracking error δεν είναι κατ' ανάγκη αποδεκτό, αν παραβιάζει τον περιορισμό προϋπολογισμού, αν υπερβαίνει το επιτρεπτό πλήθος τίτλων, αν αποδίδει υπερβολικό βάρος σε λίγες μετοχές ή αν συνεπάγεται μη αποδεκτό κόστος συναλλαγών. Για τον λόγο αυτό, η αντικειμενική συνάρτηση πρέπει πάντοτε να εξετάζεται σε συνδυασμό με τους περιορισμούς του μοντέλου.

Σε πιο εκτεταμένες ή εναλλακτικές διατυπώσεις του προβλήματος, θα μπορούσε να εξεταστεί και η ενσωμάτωση πρόσθετων όρων στην αντικειμενική συνάρτηση, όπως ποινές για παραβίαση περιορισμών ή όροι που σχετίζονται με το κόστος συναλλαγών. Στην παρούσα εργασία, όμως, η βασική λογική είναι να διατηρηθεί η αντικειμενική συνάρτηση προσανατολισμένη στον πυρήνα του προβλήματος, δηλαδή στο tracking error, ενώ οι λοιπές απαιτήσεις να αποτυπώνονται μέσω ρητών περιορισμών ή μέσω κατάλληλων μηχανισμών ελέγχου εφικτότητας στο αλγοριθμικό επίπεδο.

3.5 Περιορισμοί του προβλήματος

Η αντικειμενική συνάρτηση που παρουσιάστηκε στην προηγούμενη ενότητα αποτυπώνει τον βασικό στόχο του προβλήματος της αποτύπωσης δείκτη, δηλαδή την ελαχιστοποίηση της απόκλισης μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη αναφοράς. Ωστόσο, η αναζήτηση μιας λύσης δεν μπορεί να βασίζεται αποκλειστικά στην ελαχιστοποίηση του tracking error, διότι στην πράξη η κατασκευή ενός χαρτοφυλακίου διέπεται από συγκεκριμένους επενδυτικούς και λειτουργικούς περιορισμούς. Οι περιορισμοί αυτοί είναι απαραίτητοι ώστε η τελική λύση να είναι όχι μόνο μαθηματικά αποδοτική, αλλά και χρηματοοικονομικά εφικτή, ρεαλιστική και εφαρμόσιμη.

Στο πλαίσιο της παρούσας εργασίας, το πρόβλημα διατυπώνεται με ένα σύνολο βασικών περιορισμών που αντιστοιχούν στις πιο ουσιώδεις απαιτήσεις της αποτύπωσης δείκτη υπό ρεαλιστικές συνθήκες. Οι περιορισμοί αυτοί αφορούν τον συνολικό προϋπολογισμό, τον μέγιστο αριθμό τίτλων που μπορούν να συμμετέχουν στο χαρτοφυλάκιο, τα κατώτατα και ανώτατα όρια συμμετοχής των τίτλων, καθώς και το κόστος συναλλαγών, όπου αυτό κρίνεται αναγκαίο. Η συνολική μορφή του προβλήματος προκύπτει επομένως από τον συνδυασμό της αντικειμενικής συνάρτησης με το παρακάτω σύνολο περιορισμών.

3.5.1 Περιορισμός προϋπολογισμού

Ο πρώτος και θεμελιώδης περιορισμός είναι ο περιορισμός προϋπολογισμού. Ο περιορισμός αυτός επιβάλλει ότι το σύνολο του διαθέσιμου κεφαλαίου κατανέμεται μεταξύ των τίτλων του χαρτοφυλακίου και ότι το άθροισμα των βαρών τους πρέπει να είναι ίσο με τη μονάδα:

$$\sum_{i=1}^n w_i = 1 \quad (3.6)$$

Η σχέση αυτή διασφαλίζει ότι το χαρτοφυλάκιο είναι πλήρως επενδεδυμένο και ότι δεν υπερβαίνει το διαθέσιμο κεφάλαιο. Παράλληλα, εφόσον θεωρείται long-only στρατηγική, ισχύει και ο περιορισμός μη αρνητικότητας:

$$w_i \geq 0, \forall i \in I \quad (3.7)$$

Με τον τρόπο αυτό, αποκλείονται αρνητικές θέσεις και ανοικτές πωλήσεις, γεγονός που καθιστά τη διατύπωση συμβατή με την κλασική λογική παθητικής διαχείρισης χαρτοφυλακίου.

3.5.2 Περιορισμός πληθικότητας

Ο δεύτερος βασικός περιορισμός αφορά τον αριθμό των τίτλων που μπορούν να περιλαμβάνονται στο χαρτοφυλάκιο. Αν z_i είναι η δυαδική μεταβλητή που δηλώνει αν ο τίτλος i επιλέγεται ή όχι, τότε ο περιορισμός πληθικότητας διατυπώνεται ως εξής:

$$\sum_{i=1}^n z_i \leq K \quad (3.8)$$

όπου K είναι ο μέγιστος επιτρεπτός αριθμός τίτλων στο χαρτοφυλάκιο. Σε ορισμένες διατυπώσεις, ο περιορισμός μπορεί να λαμβάνει και τη μορφή ακριβούς ισότητας,

$$\sum_{i=1}^n z_i = K \quad (3.9)$$

όταν απαιτείται το χαρτοφυλάκιο να περιλαμβάνει ακριβώς K τίτλους. Η γενική διατύπωση με ανώτατο όριο είναι πιο ευέλικτη, καθώς επιτρέπει στον αλγόριθμο να λειτουργεί με μεγαλύτερη ελευθερία μέσα στα επιτρεπτά όρια πολυπλοκότητας του χαρτοφυλακίου.

Ο περιορισμός αυτός είναι καθοριστικής σημασίας, διότι διασφαλίζει ότι το χαρτοφυλάκιο παραμένει διαχειρίσιμο και πρακτικά εφαρμόσιμο, ενώ ταυτόχρονα μετατρέπει το πρόβλημα σε μικτό πρόβλημα διακριτής και συνεχούς βελτιστοποίησης.

3.5.3 Κατώτατα και ανώτατα όρια συμμετοχής

Για να διασφαλιστεί ότι κάθε επιλεγμένος τίτλος έχει ουσιαστική αλλά όχι υπερβολική συμμετοχή στο χαρτοφυλάκιο, επιβάλλονται κατώτατα και ανώτατα όρια στα βάρη των τίτλων. Η σύνδεση ανάμεσα στις δυαδικές μεταβλητές επιλογής και στα βάρη διατυπώνεται ως:

$$l_i z_i \leq w_i \leq u_i z_i, \forall i \in I \quad (3.10)$$

όπου l_i είναι το κατώτατο επιτρεπτό βάρος και u_i το ανώτατο επιτρεπτό βάρος για τον τίτλο i .

Η παραπάνω σχέση έχει διπλή λειτουργία. Αν $z_i = 0$, τότε αναγκαστικά $w_i = 0$, άρα ο τίτλος δεν συμμετέχει στο χαρτοφυλάκιο. Αν $z_i = 1$, τότε το βάρος του τίτλου πρέπει να βρίσκεται μέσα στο επιτρεπτό διάστημα $[l_i, u_i]$. Με τον τρόπο αυτό αποφεύγονται τόσο οι αμελητέες συμμετοχές όσο και η υπερσυγκέντρωση κεφαλαίου σε λίγες μετοχές.

Σε απλουστευμένη μορφή, όταν χρησιμοποιούνται κοινά όρια για όλους τους τίτλους, η σχέση μπορεί να γραφτεί ως:

$$l z_i \leq w_i \leq u z_i, \forall i \in I \quad (3.11)$$

όπου l και u είναι σταθερές για όλο το επενδυτικό σύνολο.

3.5.4 Περιορισμός κόστους συναλλαγών

Εφόσον το πρόβλημα εξετάζεται σε ρεαλιστικό πλαίσιο, μπορεί να ενσωματωθεί και περιορισμός κόστους συναλλαγών. Αν θεωρηθεί ότι υπάρχει υφιστάμενο χαρτοφυλάκιο με αρχικά βάρη w_i^0 , τότε το κόστος μετάβασης προς το νέο χαρτοφυλάκιο μπορεί να εκφραστεί με βάση τη διαφορά μεταξύ των νέων και των αρχικών θέσεων.

Μια γενική μορφή του συνολικού κόστους συναλλαγών είναι:

$$\sum_{i=1}^n c_i |w_i - w_i^0| \quad (3.12)$$

όπου c_i είναι ο συντελεστής κόστους για τον τίτλο i . Αν υπάρχει ανώτατο αποδεκτό επίπεδο συνολικού κόστους, τότε ο αντίστοιχος περιορισμός γράφεται:

$$\sum_{i=1}^n c_i |w_i - w_i^0| \leq C_{\max} \quad (3.13)$$

όπου $C_{\max}$ είναι το μέγιστο αποδεκτό κόστος συναλλαγών.

Η παρουσία του απόλυτου όρου καθιστά τη διατύπωση πιο ρεαλιστική, καθώς λαμβάνει υπόψη τόσο αυξήσεις όσο και μειώσεις θέσεων. Παράλληλα, ο περιορισμός αυτός επιτρέπει στον αλγόριθμο να αποφεύγει λύσεις που, παρότι μπορεί να μειώνουν το tracking error, απαιτούν υπερβολικά μεγάλες ανακατανομές κεφαλαίου.

3.5.5 Οικονομική ερμηνεία των περιορισμών

Οι παραπάνω περιορισμοί δεν αποτελούν απλές μαθηματικές συνθήκες, αλλά έχουν σαφή οικονομική σημασία. Ο περιορισμός προϋπολογισμού αποτυπώνει το γεγονός ότι ο

επενδυτής λειτουργεί μέσα σε συγκεκριμένα χρηματοδοτικά όρια και ότι το σύνολο του διαθέσιμου κεφαλαίου πρέπει να κατανεμηθεί συνεκτικά. Ο περιορισμός πληθικότητας αντανακλά την ανάγκη για χαρτοφυλάκια περιορισμένης πολυπλοκότητας, τα οποία να είναι διαχειρίσιμα και λειτουργικά αποδοτικά. Τα κατώτατα και ανώτατα όρια συμμετοχής διασφαλίζουν ότι οι επιμέρους θέσεις έχουν ρεαλιστικό μέγεθος και ότι το χαρτοφυλάκιο δεν γίνεται ούτε υπερβολικά κατακερματισμένο ούτε υπερβολικά συγκεντρωμένο. Τέλος, ο περιορισμός κόστους συναλλαγών συνδέει τη διαδικασία βελτιστοποίησης με τη χρηματοοικονομική πραγματικότητα των αγορών, όπου κάθε αναπροσαρμογή συνεπάγεται οικονομικό κόστος.

Συνδυαστικά, οι περιορισμοί αυτοί διαμορφώνουν το επιτρεπτό σύνολο λύσεων μέσα στο οποίο θα κινηθεί ο Γενετικός Αλγόριθμος. Η ποιότητα μιας λύσης δεν εξαρτάται επομένως μόνο από το πόσο μικρό tracking error επιτυγχάνει, αλλά και από το κατά πόσο παραμένει συμβατή με όλες τις παραπάνω απαιτήσεις. Για τον λόγο αυτό, η επιτυχία της προτεινόμενης μεθόδου δεν θα εξαρτηθεί μόνο από τον ορισμό της αντικειμενικής συνάρτησης, αλλά και από τον αποτελεσματικό χειρισμό αυτών των περιορισμών κατά τη διαδικασία αναζήτησης.

3.6 Υποθέσεις και απλοποιήσεις του μοντέλου

Η διατύπωση ενός προβλήματος βελτιστοποίησης χαρτοφυλακίου με στόχο την αποτύπωση δείκτη προϋποθέτει την υιοθέτηση ορισμένων υποθέσεων και απλοποιήσεων, οι οποίες είναι αναγκαίες ώστε το πρόβλημα να αποκτήσει σαφή, διαχειρίσιμη και υπολογιστικά εφαρμόσιμη μορφή. Οι υποθέσεις αυτές δεν αναιρούν τη ρεαλιστική στόχευση της εργασίας, αλλά λειτουργούν ως πλαίσιο μέσα στο οποίο το πρόβλημα μπορεί να μελετηθεί συστηματικά και να επιλυθεί με μια οργανωμένη αλγοριθμική διαδικασία. Παράλληλα, η ρητή διατύπωσή τους είναι σημαντική, διότι επιτρέπει την κατανόηση των ορίων του μοντέλου και συμβάλλει στην ορθή ερμηνεία των αποτελεσμάτων.

Μία πρώτη βασική υπόθεση αφορά το σύνολο των διαθέσιμων τίτλων. Θεωρείται ότι το επενδυτικό σύμπαν είναι δεδομένο και προκαθορισμένο για κάθε εξεταζόμενο πρόβλημα, δηλαδή ότι οι τίτλοι από τους οποίους θα επιλεγεί το χαρτοφυλάκιο είναι γνωστοί εκ των προτέρων. Η υπόθεση αυτή είναι εύλογη στο πλαίσιο της αποτύπωσης δείκτη, καθώς το χαρτοφυλάκιο κατασκευάζεται με αναφορά σε συγκεκριμένο benchmark και επομένως το σύνολο των υποψηφίων τίτλων προκύπτει από τη σύνθεση του δείκτη ή από ένα συγκεκριμένο υποσύνολο σχετικό με αυτόν. Με τον τρόπο αυτό, το πρόβλημα επικεντρώνεται στην επιλογή και στάθμιση των τίτλων και όχι στον καθορισμό του ίδιου του επενδυτικού σύμπαντος.

Μία δεύτερη σημαντική υπόθεση αφορά τα δεδομένα αποδόσεων. Θεωρείται ότι οι ιστορικές αποδόσεις των τίτλων και του δείκτη είναι διαθέσιμες, αξιόπιστες και κατάλληλες για χρήση στη διαδικασία βελτιστοποίησης. Η υπόθεση αυτή δεν σημαίνει ότι το παρελθόν αναπαράγει πλήρως το μέλλον, αλλά ότι η ιστορική πληροφορία παρέχει χρήσιμο υπόβαθρο για την εκτίμηση της σχέσης μεταξύ των τίτλων και του δείκτη. Στο πλαίσιο της παρούσας εργασίας, οι αποδόσεις θεωρούνται ως το βασικό εμπειρικό υλικό πάνω στο οποίο υπολογίζεται το tracking error και αξιολογείται η ποιότητα των λύσεων.

Μία τρίτη υπόθεση είναι ότι η μορφή του προβλήματος παραμένει στατική μέσα στην επιλεγμένη περίοδο βελτιστοποίησης. Δηλαδή, θεωρείται ότι οι μεταβλητές απόφασης, οι περιορισμοί και οι παράμετροι του προβλήματος ορίζονται σε ένα συγκεκριμένο χρονικό πλαίσιο και ότι η διαδικασία βελτιστοποίησης αναφέρεται στη συγκεκριμένη αυτή περίοδο ως ενιαίο σύνολο παρατηρήσεων. Αν και στην πραγματικότητα οι αγορές μεταβάλλονται συνεχώς, η υπόθεση της στατικής διατύπωσης είναι αναγκαία για να καταστεί το πρόβλημα επιλύσιμο σε μία δεδομένη φάση. Η δυναμική πλευρά της διαχείρισης εξετάζεται έμμεσα μέσω της αξιολόγησης και της πιθανής αναπροσαρμογής, αλλά όχι ως πλήρως δυναμικό στοχαστικό πρόβλημα.

Επιπλέον, θεωρείται ότι η επένδυση πραγματοποιείται σε long-only πλαίσιο. Αυτό σημαίνει ότι δεν επιτρέπονται αρνητικά βάρη και, κατά συνέπεια, δεν επιτρέπονται ανοικτές πωλήσεις. Η υπόθεση αυτή είναι ιδιαίτερα συμβατή με τη λογική της παθητικής διαχείρισης και της αποτύπωσης δείκτη, όπου το χαρτοφυλάκιο επιδιώκει να αναπαράγει τη συμπεριφορά ενός benchmark μέσω πραγματικής κατοχής τίτλων και όχι μέσω σύνθετων στρατηγικών leverage ή short selling. Η επιλογή αυτή απλοποιεί το πρόβλημα και ταυτόχρονα το διατηρεί κοντά στις συνήθεις πρακτικές των index funds και των σχετικών παθητικών επενδυτικών σχημάτων.

Μία ακόμη βασική υπόθεση αφορά τον περιορισμό προϋπολογισμού. Θεωρείται ότι το σύνολο του διαθέσιμου κεφαλαίου επενδύεται στο χαρτοφυλάκιο και ότι δεν διατηρείται μέρος του σε μετρητά ή άλλη μορφή μη επενδεδυμένου αποθέματος. Η συνθήκη αυτή εκφράζεται μέσω της απαίτησης το άθροισμα των βαρών να είναι ίσο με τη μονάδα. Πρόκειται για μια τυπική και εύλογη παραδοχή σε μοντέλα χαρτοφυλακίου, η οποία διευκολύνει τη μαθηματική διατύπωση και καθιστά τη σύγκριση των λύσεων πιο σαφή.

Ως προς τα όρια συμμετοχής των τίτλων, θεωρείται ότι αυτά είναι γνωστά και μπορούν να επιβληθούν με σαφή τρόπο στο μοντέλο. Η παραδοχή αυτή σημαίνει ότι το πρόβλημα δεν εξετάζει την ενδογενή διαμόρφωση των επιτρεπτών ορίων, αλλά δέχεται ότι αυτά αποτελούν δεδομένες συνθήκες της επενδυτικής πολιτικής. Με τον τρόπο αυτό, το μοντέλο εστιάζει στην εύρεση της καλύτερης δυνατής λύσης υπό τα συγκεκριμένα όρια και όχι στην επιλογή των ιδίων των κανόνων διαχείρισης.

Σε σχέση με το κόστος συναλλαγών, υιοθετείται μια απλοποιημένη αναπαράστασή του, ώστε να μπορεί να ενσωματωθεί αποτελεσματικά στη διατύπωση του προβλήματος. Στην πράξη, το κόστος συναλλαγών είναι πολυδιάστατο και μπορεί να περιλαμβάνει προμήθειες, bid-ask spreads, market impact και άλλες μορφές επιβάρυνσης. Ωστόσο, για σκοπούς μοντελοποίησης θεωρείται ότι το κόστος αυτό μπορεί να προσεγγιστεί μέσω απλούστερης συνάρτησης που εξαρτάται από τη μεταβολή των θέσεων ή από προκαθορισμένους συντελεστές κόστους ανά τίτλο. Η απλοποίηση αυτή είναι συνήθης στη σχετική βιβλιογραφία και επιτρέπει τη ρεαλιστική αλλά υπολογιστικά διαχειρίσιμη ενσωμάτωση του κόστους στο μοντέλο.

Ανάλογη απλοποίηση υιοθετείται και ως προς τη διαδικασία αναπροσαρμογής του χαρτοφυλακίου. Αν και στην πράξη η αναπροσαρμογή μπορεί να εξαρτάται από πλήθος παραγόντων, όπως οι ακριβείς χρονισμοί συναλλαγών, η ρευστότητα της αγοράς ή οι ενδοπερίοδοι μεταβολές των τιμών, στο παρόν μοντέλο θεωρείται ότι η αναπροσαρμογή μπορεί να περιγραφεί μέσω της μετάβασης από ένα σύνολο αρχικών βαρών σε ένα νέο σύνολο τελικών βαρών. Η υπόθεση αυτή καθιστά το πρόβλημα περισσότερο ελέγξιμο και

επιτρέπει την αποτύπωση του κόστους και της μεταβολής της δομής του χαρτοφυλακίου χωρίς να απαιτείται πλήρης δυναμική προσομοίωση της αγοράς.

Μία ακόμη ουσιώδης απλοποίηση αφορά το ίδιο το μέτρο tracking error. Στην παρούσα εργασία, το πρόβλημα διατυπώνεται με βάση ένα συγκεκριμένο μέτρο απόκλισης, το οποίο θεωρείται ότι αποτυπώνει επαρκώς την ποιότητα της παρακολούθησης του δείκτη. Αν και στη βιβλιογραφία υπάρχουν εναλλακτικοί τρόποι ορισμού του σφάλματος αποτύπωσης, η υιοθέτηση ενός βασικού και σαφούς measure είναι αναγκαία ώστε το πρόβλημα να παραμένει σαφώς ορισμένο και η αντικειμενική συνάρτηση να είναι συγκεκριμένη και λειτουργική.

Πρέπει επίσης να τονιστεί ότι το μοντέλο δεν αποσκοπεί στην πλήρη περιγραφή όλων των πραγματικών διαστάσεων των αγορών, αλλά στην ανάπτυξη ενός επαρκώς ρεαλιστικού και ταυτόχρονα υπολογιστικά επιλύσιμου πλαισίου. Αυτό σημαίνει ότι ορισμένες πλευρές της επενδυτικής πραγματικότητας, όπως φορολογικές επιβαρύνσεις, περιορισμοί ρευστότητας σε πολύ λεπτομερές επίπεδο, θεσμικές διαφοροποιήσεις μεταξύ αγορών ή στοχαστικές μεταβολές παραμέτρων σε πραγματικό χρόνο, δεν ενσωματώνονται πλήρως στο μοντέλο. Η επιλογή αυτή δεν συνιστά αδυναμία, αλλά αναγκαίο στοιχείο μεθοδολογικής εστίασης, ώστε η εργασία να παραμείνει συνεκτική και τεχνικά εφαρμόσιμη.

3.7 Σχεδιασμός του προτεινόμενου Γενετικού Αλγορίθμου

Η επίλυση του προβλήματος της αποτύπωσης δείκτη, όπως διατυπώθηκε στις προηγούμενες ενότητες, απαιτεί μια μεθοδολογία ικανή να διαχειριστεί ταυτόχρονα διακριτές και συνεχείς αποφάσεις, καθώς και ένα σύνολο ρεαλιστικών περιορισμών. Στο πλαίσιο αυτό, επιλέγεται η χρήση Γενετικού Αλγορίθμου, καθώς η συγκεκριμένη κατηγορία μεθόδων προσφέρει την απαραίτητη ευελιξία για την αντιμετώπιση σύνθετων προβλημάτων βελτιστοποίησης, όπου ο χώρος λύσεων είναι μεγάλος, μη ομαλός και έντονα συνδυαστικός. Ο σχεδιασμός του προτεινόμενου αλγορίθμου δεν αντιμετωπίζεται ως τυπική εφαρμογή μιας γενικής τεχνικής, αλλά ως στοχευμένη προσαρμογή των βασικών αρχών των Γενετικών Αλγορίθμων στις ιδιαιτερότητες του index tracking.

Η βασική ιδέα του προτεινόμενου Γενετικού Αλγορίθμου είναι ότι κάθε υποψήφια λύση αναπαριστά ένα πλήρες χαρτοφυλάκιο αποτύπωσης δείκτη. Αυτό σημαίνει ότι κάθε άτομο του πληθυσμού πρέπει να ενσωματώνει τόσο την πληροφορία για το ποιοι τίτλοι περιλαμβάνονται στο χαρτοφυλάκιο όσο και την πληροφορία για τα αντίστοιχα βάρη συμμετοχής τους. Με άλλα λόγια, η λύση πρέπει να αποτυπώνει ταυτόχρονα τη διακριτή διάσταση του προβλήματος, δηλαδή την επιλογή τίτλων, και τη συνεχή διάσταση, δηλαδή την κατανομή κεφαλαίου μεταξύ των επιλεγμένων τίτλων. Η επιλογή αυτού του τύπου αναπαράστασης είναι καθοριστική, διότι επιτρέπει στον αλγόριθμο να λειτουργεί άμεσα πάνω στην ίδια τη δομή του προβλήματος.

Ο προτεινόμενος αλγόριθμος θα λειτουργεί σε επίπεδο πληθυσμού. Σε κάθε χρονική στιγμή θα διατηρεί ένα σύνολο υποψήφιων χαρτοφυλακίων, τα οποία θα εξελίσσονται προοδευτικά μέσα από την εφαρμογή τελεστών επιλογής, διασταύρωσης και μετάλλαξης. Η λογική αυτής της πληθυσμιακής προσέγγισης είναι ότι επιτρέπει την ταυτόχρονη διερεύνηση πολλών διαφορετικών περιοχών του χώρου λύσεων και έτσι μειώνει τον κίνδυνο εγκλωβισμού σε τοπικά βέλτιστα. Το χαρακτηριστικό αυτό είναι ιδιαίτερα σημαντικό στο πρόβλημα της αποτύπωσης δείκτη, όπου ο συνδυασμός cardinality constraints, ορίων

βαρών και κόστους συναλλαγών δημιουργεί έναν πολύπλοκο και συχνά μη κυρτό χώρο εφικτών λύσεων.

Ο σχεδιασμός του αλγορίθμου βασίζεται στην αρχή ότι οι καλύτερες λύσεις, δηλαδή εκείνες που επιτυγχάνουν χαμηλότερο tracking error και ικανοποιούν τους περιορισμούς, πρέπει να έχουν μεγαλύτερη πιθανότητα να επηρεάσουν τις επόμενες γενιές. Για τον λόγο αυτό, ο πληθυσμός θα αξιολογείται σε κάθε επανάληψη μέσω κατάλληλης συνάρτησης καταλληλότητας, η οποία θα συνδέεται άμεσα με την αντικειμενική συνάρτηση του προβλήματος και, ενδεχομένως, με μηχανισμούς ποινής για παραβιάσεις περιορισμών. Έτσι, ο αλγόριθμος θα ευνοεί συστηματικά εκείνες τις λύσεις που είναι πιο κοντά στον επιθυμητό στόχο, χωρίς όμως να αποκλείει πλήρως την παρουσία ποικιλίας στον πληθυσμό.

Η επιλογή των γονέων αποτελεί κρίσιμο τμήμα του σχεδιασμού, καθώς καθορίζει ποια χαρτοφυλάκια θα χρησιμοποιηθούν για τη δημιουργία απογόνων. Η διαδικασία αυτή πρέπει να εξισορροπεί δύο αντίρροπες ανάγκες: από τη μία πλευρά, να δίνει πλεονέκτημα στις πιο αποδοτικές λύσεις και, από την άλλη, να διατηρεί επαρκή ποικιλία ώστε να μην οδηγείται ο πληθυσμός σε πρόωρη ομογενοποίηση. Για τον λόγο αυτό, η επιλογή δεν πρέπει να είναι υπερβολικά επιθετική, αλλά να επιτρέπει και σε λιγότερο ισχυρές λύσεις να συμμετέχουν περιστασιακά στη διαδικασία αναπαραγωγής, διατηρώντας έτσι μεγαλύτερη διερευνητική ικανότητα.

Η διασταύρωση θα λειτουργεί ως μηχανισμός συνδυασμού χαρακτηριστικών από δύο ή περισσότερες γονικές λύσεις. Στο πρόβλημα της αποτύπωσης δείκτη, αυτό σημαίνει ότι ένας απόγονος μπορεί να κληρονομεί μέρος της επιλογής τίτλων από έναν γονέα και μέρος της κατανομής βαρών από έναν άλλο. Η επιτυχία του τελεστή αυτού εξαρτάται από το κατά πόσο μπορεί να συνδυάζει χρήσιμα χαρακτηριστικά χωρίς να παράγει υπερβολικά μεγάλο αριθμό μη εφικτών λύσεων. Για τον λόγο αυτό, ο σχεδιασμός της διασταύρωσης πρέπει να λαμβάνει υπόψη όχι μόνο τη γενική λογική των Γενετικών Αλγορίθμων, αλλά και τους ειδικούς περιορισμούς του χαρτοφυλακίου, όπως ο περιορισμός πληθικότητας και τα όρια συμμετοχής.

Η μετάλλαξη, από την άλλη πλευρά, θα αποτελεί τον βασικό μηχανισμό εισαγωγής νέας πληροφορίας στον πληθυσμό. Μέσω μικρών αλλά στοχευμένων τροποποιήσεων σε μια λύση, ο αλγόριθμος μπορεί να διατηρεί την ποικιλία και να αποφεύγει τη στασιμότητα της αναζήτησης. Στο πλαίσιο του index tracking, η μετάλλαξη μπορεί να αφορά είτε την αντικατάσταση ενός τίτλου με κάποιον άλλο είτε τη μικρή ανακατανομή βαρών μεταξύ υφιστάμενων τίτλων είτε και τα δύο ταυτόχρονα. Ο ρόλος της είναι ιδιαίτερα σημαντικός σε μεταγενέστερα στάδια της εξέλιξης, όταν ο πληθυσμός έχει ήδη αρχίσει να συγκλίνει και απαιτείται νέος διερευνητικός μηχανισμός για την αποφυγή πρόωρης σύγκλισης.

Ιδιαίτερη βαρύτητα στον σχεδιασμό του αλγορίθμου έχει η διαχείριση της εφικτότητας των λύσεων. Επειδή το πρόβλημα περιλαμβάνει πολλούς περιορισμούς, είναι πολύ πιθανό οι τελεστές διασταύρωσης και μετάλλαξης να παράγουν προσωρινά λύσεις που παραβιάζουν τον προϋπολογισμό, το πλήθος τίτλων ή τα επιτρεπτά όρια βαρών. Για τον λόγο αυτό, ο προτεινόμενος αλγόριθμος θα πρέπει να περιλαμβάνει είτε μηχανισμούς επιδιόρθωσης είτε κατάλληλους κανόνες κανονικοποίησης και ελέγχου, ώστε οι παραγόμενοι απόγονοι να επανέρχονται εντός των ορίων του εφικτού χώρου. Η διαχείριση αυτή είναι καθοριστική, καθώς η ποιότητα του Γενετικού Αλγορίθμου δεν εξαρτάται μόνο από την ικανότητά του να παράγει καλές λύσεις, αλλά και από την ικανότητά του να διατηρεί την αναζήτηση μέσα σε οικονομικά και επενδυτικά αποδεκτά πλαίσια.

Ένα ακόμη σημαντικό στοιχείο του σχεδιασμού είναι ο καθορισμός του κριτηρίου τερματισμού. Ο αλγόριθμος πρέπει να σταματά όταν έχει επιτευχθεί ικανοποιητικός βαθμός σύγκλισης ή όταν η συνέχιση της διαδικασίας δεν αναμένεται να προσφέρει ουσιαστική βελτίωση. Συνήθεις επιλογές είναι ο μέγιστος αριθμός γενεών, η απουσία ουσιαστικής βελτίωσης στη βέλτιστη λύση για συγκεκριμένο αριθμό επαναλήψεων ή ο συνδυασμός των δύο. Η επιλογή του κριτηρίου αυτού επηρεάζει άμεσα τόσο τον χρόνο εκτέλεσης όσο και την ποιότητα του αποτελέσματος και, συνεπώς, πρέπει να είναι συμβατή με τον συνολικό πειραματικό σχεδιασμό της εργασίας.

3.8 Αναπαράσταση χρωμοσώματος

Η αναπαράσταση του χρωμοσώματος αποτελεί ένα από τα πιο κρίσιμα σημεία στον σχεδιασμό ενός Γενετικού Αλγορίθμου, καθώς καθορίζει τον τρόπο με τον οποίο κωδικοποιείται κάθε υποψήφια λύση και, κατ' επέκταση, επηρεάζει άμεσα την αποτελεσματικότητα των τελεστών επιλογής, διασταύρωσης και μετάλλαξης. Στο πρόβλημα της αποτύπωσης δείκτη, η αναπαράσταση πρέπει να είναι σε θέση να αποτυπώνει με σαφή και λειτουργικό τρόπο τόσο την επιλογή των τίτλων όσο και τα αντίστοιχα βάρη συμμετοχής τους στο χαρτοφυλάκιο. Επομένως, η μορφή του χρωμοσώματος οφείλει να ανταποκρίνεται στη μικτή φύση του προβλήματος, το οποίο περιλαμβάνει ταυτόχρονα διακριτές και συνεχείς αποφάσεις.

Στην παρούσα εργασία, κάθε χρωμόσωμα αναπαριστά ένα πλήρες χαρτοφυλάκιο αποτύπωσης δείκτη. Η δομή του περιλαμβάνει πληροφορία για το ποιο τίτλοι συμμετέχουν στο χαρτοφυλάκιο και για το ποιο είναι το βάρος καθενός από αυτούς. Με αυτή τη λογική, το χρωμόσωμα δεν αποτελεί απλώς μια αφηρημένη ακολουθία τιμών, αλλά μια πλήρη κωδικοποίηση μιας επενδυτικής λύσης. Η επιλογή αυτής της αναπαράστασης είναι ουσιώδης, διότι επιτρέπει στον αλγόριθμο να λειτουργεί απευθείας πάνω στα χαρακτηριστικά που έχουν πραγματικό οικονομικό νόημα για το πρόβλημα.

Η πιο κατάλληλη μορφή για το συγκεκριμένο πρόβλημα είναι μια σύνθετη αναπαράσταση, στην οποία κάθε χρωμόσωμα συνδέεται με δύο διαστάσεις: τη δυαδική διάσταση επιλογής τίτλων και τη συνεχή διάσταση προσδιορισμού βαρών. Πιο συγκεκριμένα, σε εννοιολογικό επίπεδο, ένα χρωμόσωμα μπορεί να θεωρηθεί ότι αποτελείται από δύο παράλληλα μέρη. Το πρώτο μέρος δηλώνει αν ένας τίτλος συμμετέχει ή όχι στο χαρτοφυλάκιο, ενώ το δεύτερο αποδίδει το αντίστοιχο βάρος του. Έτσι, για κάθε τίτλο i , το χρωμόσωμα περιλαμβάνει αφενός μια μεταβλητή συμμετοχής $z_i \in \{0,1\}$ και αφετέρου μια μεταβλητή βάρους $w_i \geq 0$.

Η αναπαράσταση αυτή μπορεί να αποδοθεί σχηματικά ως εξής:

$$\text{Χρωμόσωμα} = [(z_1, w_1), (z_2, w_2), \dots, (z_n, w_n)] \quad (3.14)$$

όπου κάθε ζεύγος (z_i, w_i) αντιστοιχεί σε έναν τίτλο του επενδυτικού συνόλου. Αν $z_i = 0$, τότε ο τίτλος δεν συμμετέχει στο χαρτοφυλάκιο και το αντίστοιχο βάρος του πρέπει να είναι $w_i = 0$. Αν $z_i = 1$, τότε ο τίτλος είναι ενεργός στο χαρτοφυλάκιο και το βάρος του πρέπει να ικανοποιεί τους περιορισμούς του μοντέλου.

Η σύνδεση ανάμεσα στα δύο μέρη της αναπαράστασης είναι εξαιρετικά σημαντική. Δεν αρκεί δηλαδή να υπάρχει απλώς μια δυαδική απόφαση επιλογής και μια αριθμητική τιμή

βάρους, αλλά οι δύο αυτές πληροφορίες πρέπει να είναι συνεπείς μεταξύ τους. Ο τίτλος που δεν έχει επιλεγεί δεν μπορεί να έχει θετικό βάρος, ενώ ο τίτλος που έχει επιλεγεί πρέπει να έχει βάρος μέσα στα επιτρεπτά όρια. Αυτή η συνέπεια θα διασφαλίζεται είτε άμεσα από τον τρόπο δημιουργίας και τροποποίησης των χρωμοσωμάτων είτε μέσω μηχανισμών επιδιόρθωσης που θα εφαρμόζονται μετά από γενετικούς τελεστές.

Η επιλογή αυτής της σύνθετης αναπαράστασης προσφέρει σημαντικά πλεονεκτήματα. Πρώτον, είναι απολύτως συμβατή με τη μαθηματική διατύπωση του προβλήματος, όπου επίσης υπάρχουν δυαδικές μεταβλητές συμμετοχής και συνεχείς μεταβλητές βαρών. Δεύτερον, επιτρέπει την άμεση ενσωμάτωση του περιορισμού πληθικότητας, αφού ο αριθμός των τίτλων που συμμετέχουν στο χαρτοφυλάκιο μπορεί να ελεγχθεί μέσω του πλήθους των γονιδίων για τα οποία ισχύει $z_i = 1$. Τρίτον, επιτρέπει την εύκολη προσαρμογή των γενετικών τελεστών, καθώς είναι δυνατόν να σχεδιαστούν διαφορετικοί μηχανισμοί παρέμβασης είτε στη διάσταση της επιλογής τίτλων είτε στη διάσταση των βαρών.

Παράλληλα, η μορφή αυτή επιτρέπει και μεγαλύτερη ευελιξία σε σύγκριση με πιο απλουστευμένες αναπαραστάσεις. Σε ορισμένες εφαρμογές της βιβλιογραφίας χρησιμοποιούνται καθαρά δυαδικά χρωμοσώματα, όπου ο αλγόριθμος επιλέγει μόνο τίτλους και τα βάρη προκύπτουν σε δεύτερο στάδιο από άλλη μέθοδο. Αν και αυτή η προσέγγιση μπορεί να είναι λειτουργική, διαχωρίζει τεχνητά δύο διαστάσεις του προβλήματος που στην πραγματικότητα είναι στενά αλληλεξαρτώμενες. Στην παρούσα εργασία προτιμάται η άμεση και ενιαία αναπαράσταση, ώστε η επιλογή τίτλων και ο προσδιορισμός βαρών να αποτελούν ταυτόχρονο αντικείμενο της εξελικτικής διαδικασίας.

Εναλλακτικά, σε πιο πρακτικό υπολογιστικό επίπεδο, το χρωμόσωμα μπορεί να αναπαρίσταται ως ένα διάνυσμα πραγματικών τιμών $w = (w_1, w_2, \dots, w_n)$, όπου η συμμετοχή των τίτλων προκύπτει έμμεσα από το αν το αντίστοιχο βάρος είναι μη μηδενικό. Σε αυτή την περίπτωση, η δυαδική πληροφορία δεν αποθηκεύεται ξεχωριστά, αλλά εξάγεται από τη δομή του διανύσματος βαρών. Η προσέγγιση αυτή είναι επίσης χρήσιμη, ιδίως όταν ο αλγόριθμος εφαρμόζει μηχανισμούς κανονικοποίησης ή επιδιόρθωσης που μετατρέπουν αυτόματα ένα γενικό διάνυσμα τιμών σε εφικτό χαρτοφυλάκιο με συγκεκριμένο πλήθος ενεργών τίτλων. Παρ' όλα αυτά, από εννοιολογική και μαθηματική άποψη, η δυαδική και η συνεχής διάσταση του χρωμοσώματος παραμένουν διακριτές.

Ένα ιδιαίτερα σημαντικό ζήτημα στην αναπαράσταση είναι η εφικτότητα του χρωμοσώματος. Κάθε χρωμόσωμα θα πρέπει, ιδανικά, να αντιστοιχεί σε αποδεκτό χαρτοφυλάκιο, δηλαδή σε λύση που σέβεται τους περιορισμούς του προβλήματος. Ωστόσο, επειδή η διαδικασία εξέλιξης περιλαμβάνει τελεστές όπως η διασταύρωση και η μετάλλαξη, είναι πιθανό να παράγονται προσωρινά χρωμοσώματα που δεν ικανοποιούν πλήρως τους περιορισμούς. Για τον λόγο αυτό, η αναπαράσταση πρέπει να είναι σχεδιασμένη με τέτοιο τρόπο ώστε να υποστηρίζει εύκολα την επιδιόρθωση ή την κανονικοποίηση των λύσεων. Αυτό σημαίνει ότι η δομή του χρωμοσώματος δεν πρέπει να είναι μόνο περιγραφικά επαρκής, αλλά και λειτουργικά κατάλληλη για τις ανάγκες της εξελικτικής διαδικασίας.

Επιπλέον, η επιλεγμένη αναπαράσταση επηρεάζει άμεσα την ποιότητα των απογόνων που δημιουργούνται από τους γενετικούς τελεστές. Αν η δομή του χρωμοσώματος είναι ασαφής ή πολύπλοκη, τότε η διασταύρωση και η μετάλλαξη μπορεί να οδηγούν συχνά σε μη ερμηνεύσιμες ή μη εφικτές λύσεις. Αντίθετα, μια αναπαράσταση που είναι ευθυγραμμισμένη με τη λογική του προβλήματος επιτρέπει στους τελεστές να λειτουργούν

πιο αποτελεσματικά και να παράγουν χαρτοφυλάκια με μεγαλύτερη πιθανότητα χρηματοοικονομικής και μαθηματικής συνοχής.

3.9 Δημιουργία αρχικού πληθυσμού

Η δημιουργία του αρχικού πληθυσμού αποτελεί κρίσιμο στάδιο στη λειτουργία του προτεινόμενου Γενετικού Αλγορίθμου, καθώς καθορίζει το σημείο εκκίνησης της εξελικτικής διαδικασίας και επηρεάζει σε σημαντικό βαθμό τόσο την ταχύτητα σύγκλισης όσο και τη συνολική ποιότητα των λύσεων που θα παραχθούν. Εφόσον ο Γενετικός Αλγόριθμος εξελίσσει έναν πληθυσμό υποψήφιων χαρτοφυλακίων, η αρχική σύνθεσή του δεν πρέπει να είναι τυχαία με ανεξέλεγκτο τρόπο, αλλά να διαμορφώνεται έτσι ώστε να παρέχει επαρκή ποικιλία λύσεων και ταυτόχρονα να κινείται όσο το δυνατόν περισσότερο μέσα στον εφικτό χώρο του προβλήματος.

Στο πρόβλημα της αποτύπωσης δείκτη, η ανάγκη προσεκτικού σχεδιασμού του αρχικού πληθυσμού είναι ακόμη μεγαλύτερη, διότι κάθε άτομο πρέπει να ανταποκρίνεται σε ένα σύνθετο σύνολο απαιτήσεων. Δεν αρκεί να αντιστοιχεί σε ένα οποιοδήποτε διάνυμα βαρών, αλλά πρέπει να αποτυπώνει μια λογική επιλογή τίτλων, να ικανοποιεί τον περιορισμό προϋπολογισμού, να σέβεται τον περιορισμό πληθικότητας και να βρίσκεται, όσο είναι δυνατόν, εντός των επιτρεπτών ορίων συμμετοχής. Επομένως, η αρχικοποίηση δεν είναι μια απλή τεχνική διαδικασία, αλλά ουσιαστικό στοιχείο της συνολικής αλγοριθμικής σχεδίασης.

Ο βασικός στόχος κατά τη δημιουργία του αρχικού πληθυσμού είναι διττός. Από τη μία πλευρά, πρέπει να εξασφαλίζεται ποικιλία, ώστε ο πληθυσμός να καλύπτει διαφορετικές περιοχές του χώρου λύσεων και να μην ξεκινά ήδη από μια στενή ή ομοιογενή περιοχή αναζήτησης. Από την άλλη πλευρά, πρέπει να αποφεύγεται η παραγωγή μεγάλου αριθμού προφανώς κακών ή εξ ολοκλήρου μη εφικτών λύσεων, καθώς κάτι τέτοιο θα καθιστούσε την αρχική φάση της εξέλιξης ιδιαίτερα αναποτελεσματική. Συνεπώς, απαιτείται ισορροπία ανάμεσα στη διαφοροποίηση και στη στοιχειώδη ποιότητα της αρχικοποίησης.

Στην παρούσα εργασία, ο αρχικός πληθυσμός θα δημιουργείται με ημι-τυχαίο τρόπο. Αυτό σημαίνει ότι η παραγωγή των αρχικών χρωμοσωμάτων θα βασίζεται μεν σε τυχαία επιλογή, αλλά μέσα σε ένα ελεγχόμενο πλαίσιο κανόνων που διασφαλίζει τη βασική συμβατότητα των λύσεων με το πρόβλημα. Συγκεκριμένα, για κάθε άτομο του αρχικού πληθυσμού θα επιλέγεται πρώτα ένα υποσύνολο τίτλων, του οποίου το πλήθος δεν θα υπερβαίνει το επιτρεπτό όριο K . Η επιλογή αυτή μπορεί να γίνεται τυχαία από το διαθέσιμο επενδυτικό σύνολο, με τρόπο που να επιτρέπει τη δημιουργία χαρτοφυλακίων διαφορετικής σύνθεσης. Αφού καθοριστεί το υποσύνολο των επιλεγμένων τίτλων, θα ακολουθεί η απόδοση αρχικών βαρών. Η διαδικασία αυτή πρέπει να είναι τέτοια ώστε το άθροισμα των βαρών να ισούται με τη μονάδα και, παράλληλα, κάθε ενεργός τίτλος να ικανοποιεί τα κατώτατα και ανώτατα όρια συμμετοχής. Ένας πρακτικός τρόπος είναι να παράγονται αρχικά τυχαίες θετικές τιμές για τους επιλεγμένους τίτλους και στη συνέχεια να εφαρμόζεται κατάλληλη κανονικοποίηση, έτσι ώστε τα βάρη να αθροίζονται στη μονάδα. Αν μετά την κανονικοποίηση κάποια βάρη παραβιάζουν τα επιτρεπτά όρια, θα εφαρμόζεται διαδικασία επιδιόρθωσης ώστε να επανέρχονται εντός των αποδεκτών ορίων.

Η ύπαρξη μηχανισμού επιδιόρθωσης ήδη από τη φάση της αρχικοποίησης είναι ιδιαίτερα σημαντική. Αν ο αρχικός πληθυσμός περιέχει μεγάλο ποσοστό λύσεων που παραβιάζουν βασικούς περιορισμούς, τότε ο αλγόριθμος αναγκάζεται να αφιερώνει σημαντικό μέρος της

εξελικτικής του προσπάθειας όχι στη βελτίωση του tracking error αλλά στην απλή αποκατάσταση της εφικτότητας. Αντίθετα, ένας αρχικός πληθυσμός που είναι κατά κανόνα εφικτός επιτρέπει στην εξελικτική διαδικασία να ξεκινήσει αμέσως από χαρτοφυλάκια με νόημα και να επικεντρωθεί γρηγορότερα στην ουσιαστική βελτιστοποίηση.

Παράλληλα, η δημιουργία του αρχικού πληθυσμού δεν πρέπει να είναι υπερβολικά «καθοδηγούμενη», διότι κάτι τέτοιο θα μπορούσε να περιορίσει τεχνητά την ποικιλία των λύσεων. Αν όλα τα άτομα αρχικοποιούνται πολύ κοντά σε παρόμοιες δομές, τότε ο πληθυσμός μπορεί να στερηθεί την αναγκαία εξερευνητική ικανότητα και να συγκλίνει πρόωρα σε περιορισμένο τμήμα του χώρου λύσεων. Για τον λόγο αυτό, ακόμη και αν εφαρμόζονται κανόνες εφικτότητας, η διαδικασία πρέπει να επιτρέπει σημαντικό βαθμό διαφοροποίησης τόσο ως προς το ποιοι τίτλοι επιλέγονται όσο και ως προς το πώς κατανέμονται τα βάρη τους.

Μια επιπλέον σχεδιαστική επιλογή είναι το αν ο αρχικός πληθυσμός θα περιέχει αποκλειστικά τυχαία παραγόμενες λύσεις ή αν θα ενισχύεται και με ορισμένες πιο «καθοδηγούμενες» αρχικές λύσεις. Στο πλαίσιο της παρούσας μεθοδολογίας, είναι δυνατόν να περιληφθούν, εκτός από τυχαία χαρτοφυλάκια, και ορισμένες λύσεις που βασίζονται σε απλούς κανόνες, όπως για παράδειγμα επιλογή τίτλων με μεγαλύτερη στάθμιση στον δείκτη ή αρχική κατανομή βαρών που προσεγγίζει τη δομή του benchmark. Η επιλογή αυτή μπορεί να ενισχύσει την ποιότητα του αρχικού πληθυσμού, εφόσον όμως δεν μειώνει υπερβολικά την ποικιλία του. Για τον λόγο αυτό, μια μικτή στρατηγική αρχικοποίησης θεωρείται συχνά πιο ισορροπημένη από την αποκλειστικά τυχαία ή την αποκλειστικά καθοδηγούμενη. Στην πράξη, κάθε χρωμόσωμα του αρχικού πληθυσμού θα κατασκευάζεται μέσω των εξής βασικών βημάτων: πρώτον, επιλογή επιτρεπτού αριθμού τίτλων από το διαθέσιμο σύνολο· δεύτερον, απόδοση αρχικών μη μηδενικών τιμών βαρών στους επιλεγμένους τίτλους· τρίτον, κανονικοποίηση των βαρών ώστε να ικανοποιείται ο περιορισμός προϋπολογισμού· τέταρτον, έλεγχος και, εφόσον απαιτείται, επιδιόρθωση ως προς τον περιορισμό πληθικότητας και τα επιτρεπτά όρια συμμετοχής. Το αποτέλεσμα αυτής της διαδικασίας θα είναι ένα σύνολο αρχικών χαρτοφυλακίων που θα μπορούν να αποτελέσουν ουσιαστική βάση για τη μετέπειτα εξέλιξη.

Η ποιότητα του αρχικού πληθυσμού επηρεάζει άμεσα και την ταχύτητα σύγκλισης του αλγορίθμου. Αν ο αρχικός πληθυσμός είναι πλήρως τυχαίος και περιλαμβάνει πολλές χαμηλής ποιότητας λύσεις, ο αλγόριθμος μπορεί να χρειαστεί περισσότερες γενιές για να προσεγγίσει αποδοτικές περιοχές του χώρου λύσεων. Αντίθετα, ένας πληθυσμός που συνδυάζει ποικιλία και στοιχειώδη εφικτότητα μπορεί να οδηγήσει σε πιο αποδοτική και σταθερή εξελικτική διαδικασία. Παρ' όλα αυτά, πρέπει να αποφεύγεται η υπερβολική «ωρίμανση» του αρχικού πληθυσμού, διότι κάτι τέτοιο μπορεί να μειώσει τον εξερευνητικό χαρακτήρα του αλγορίθμου.

3.10 Συνάρτηση καταλληλότητας

Η συνάρτηση καταλληλότητας αποτελεί τον πυρήνα της λειτουργίας του Γενετικού Αλγορίθμου, καθώς είναι ο μηχανισμός μέσω του οποίου αποτιμάται η ποιότητα κάθε υποψήφιας λύσης. Σε κάθε γενιά, όλα τα χρωμοσώματα του πληθυσμού αξιολογούνται βάσει της συνάρτησης αυτής και, ανάλογα με την τιμή που λαμβάνουν, αποκτούν μεγαλύτερη ή μικρότερη πιθανότητα να επιλεγούν για αναπαραγωγή. Κατά συνέπεια, η

συνάρτηση καταλληλότητας δεν έχει μόνο περιγραφικό ρόλο, αλλά κατευθύνει άμεσα ολόκληρη την εξελικτική διαδικασία. Στο πρόβλημα της αποτύπωσης δείκτη, η κατάλληλη διαμόρφωσή της είναι ιδιαίτερα σημαντική, διότι πρέπει να αντανακλά πιστά τον στόχο της βελτιστοποίησης και ταυτόχρονα να χειρίζεται αποτελεσματικά τους περιορισμούς του μοντέλου.

Ο βασικός στόχος του προβλήματος, όπως έχει ήδη διατυπωθεί, είναι η ελαχιστοποίηση του tracking error. Συνεπώς, η συνάρτηση καταλληλότητας πρέπει να σχετίζεται άμεσα με το μέτρο απόκλισης μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη. Αν θεωρηθεί ότι για ένα χρωμόσωμα η αντίστοιχη λύση παράγει βάρη $w_1, w_2, \dots, w_n$, τότε το μέτρο tracking error ορίζεται ως:

$$TE = \frac{1}{m} \sum_{t=1}^m (\sum_{i=1}^n w_i r_{it} - R_t)^2 \quad (3.15)$$

Η φυσική πρώτη επιλογή θα ήταν να χρησιμοποιηθεί η παραπάνω τιμή ως βάση αξιολόγησης της λύσης. Ωστόσο, στους Γενετικούς Αλγορίθμους η συνάρτηση καταλληλότητας πρέπει να σχεδιαστεί έτσι ώστε μεγαλύτερη τιμή να αντιστοιχεί σε καλύτερη λύση, ιδίως όταν η διαδικασία επιλογής στηρίζεται σε αναλογίες ή σε συγκρίσεις καταλληλότητας. Επειδή εδώ ο στόχος είναι η ελαχιστοποίηση, απαιτείται μια μετασχηματισμένη μορφή που να μετατρέπει μικρότερο tracking error σε μεγαλύτερη καταλληλότητα.

Μια απλή και αποτελεσματική μορφή συνάρτησης καταλληλότητας είναι η εξής:

$$F = \frac{1}{1 + TE} \quad (3.16)$$

όπου F είναι η τιμή καταλληλότητας του χρωμοσώματος. Η σχέση αυτή εξασφαλίζει ότι όσο μικρότερο είναι το tracking error, τόσο μεγαλύτερη είναι η τιμή της καταλληλότητας. Ταυτόχρονα, αποφεύγεται το πρόβλημα διαίρεσης με το μηδέν και η συνάρτηση παραμένει θετική για κάθε εφικτή λύση. Πρόκειται για μια ομαλή, εύκολα υπολογίσιμη και ερμηνεύσιμη μετατροπή, η οποία είναι κατάλληλη για χρήση σε εξελικτικά σχήματα.

Παρότι η παραπάνω μορφή είναι ικανοποιητική σε θεωρητικό επίπεδο, στην πράξη η συνάρτηση καταλληλότητας πρέπει να λαμβάνει υπόψη και τις πιθανές παραβιάσεις περιορισμών. Ο λόγος είναι ότι, κατά τη διαδικασία διασταύρωσης και μετάλλαξης, είναι δυνατόν να παράγονται χρωμοσώματα που εμφανίζουν χαμηλό tracking error αλλά δεν είναι πλήρως εφικτά ως προς τον περιορισμό προϋπολογισμού, το πλήθος τίτλων ή τα όρια συμμετοχής. Αν τέτοιες λύσεις αξιολογούνται μόνο βάσει του tracking error, τότε υπάρχει κίνδυνος να ευνοηθούν υπερβολικά από τη διαδικασία επιλογής, παρότι δεν είναι στην πραγματικότητα αποδεκτές.

Για τον λόγο αυτό, η συνάρτηση καταλληλότητας μπορεί να επεκταθεί ώστε να περιλαμβάνει όρους ποινής για τυχόν παραβιάσεις περιορισμών. Σε αυτή την περίπτωση, ορίζεται πρώτα μια γενικευμένη τιμή κόστους:

$$C = TE + \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3 + \lambda_4 P_4 \quad (3.17)$$

όπου:

- TE είναι το tracking error,
- P_1 είναι όρος ποινής για παραβίαση του περιορισμού προϋπολογισμού,
- P_2 είναι όρος ποινής για παραβίαση του περιορισμού πληθικότητας,
- P_3 είναι όρος ποινής για παραβίαση των ορίων βαρών,
- P_4 είναι όρος ποινής για παραβίαση περιορισμού κόστους συναλλαγών, εφόσον αυτός χρησιμοποιείται,
- και $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ είναι θετικοί συντελεστές βαρύτητας των ποινών.

Με βάση αυτή τη συνολική τιμή κόστους, η συνάρτηση καταλληλότητας μπορεί να γραφτεί ως:

$$F = \frac{1}{1 + C} \quad (3.18)$$

Η μορφή αυτή επιτρέπει στον αλγόριθμο να αξιολογεί τις λύσεις με ενιαίο τρόπο, συνδυάζοντας τον βασικό στόχο ελαχιστοποίησης του tracking error με την ανάγκη συμμόρφωσης προς τους περιορισμούς. Όσο καλύτερη και πιο εφικτή είναι μια λύση, τόσο μικρότερη θα είναι η τιμή C και τόσο μεγαλύτερη η καταλληλότητά της.

Οι όροι ποινής μπορούν να οριστούν με διάφορους τρόπους. Για παράδειγμα, για τον περιορισμό προϋπολογισμού μπορεί να χρησιμοποιηθεί η απόλυτη απόκλιση από τη μονάδα:

$$P_1 = \left| \sum_{i=1}^n w_i - 1 \right| \quad (3.19)$$

Για τον περιορισμό πληθικότητας, μια ενδεικτική μορφή είναι:

$$P_2 = \max \left(0, \sum_{i=1}^n z_i - K \right) \quad (3.20)$$

η οποία τιμωρεί μόνο την υπέρβαση του επιτρεπτού αριθμού τίτλων. Για τα όρια βαρών, ένας γενικός όρος ποινής μπορεί να βασίζεται στο σύνολο των παραβιάσεων κάτω ή πάνω από τα επιτρεπτά όρια:

$$P_3 = \sum_{i=1}^n [\max(0, l_i z_i - w_i) + \max(0, w_i - u_i z_i)] \quad (3.21)$$

Αν χρησιμοποιείται περιορισμός κόστους συναλλαγών, τότε ο αντίστοιχος όρος ποινής μπορεί να λάβει τη μορφή:

$$P_4 = \max \left(0, \sum_{i=1}^n c_i |w_i - w_i^0| - C_{max} \right) \quad (3.22)$$

Οι παραπάνω ποινές έχουν το πλεονέκτημα ότι είναι μηδενικές όταν ο περιορισμός ικανοποιείται και θετικές μόνο όταν υπάρχει παραβίαση. Έτσι, οι εφικτές λύσεις δεν

επιβαρύνονται τεχνητά, ενώ οι μη εφικτές λύσεις αποθαρρύνονται σε βαθμό που εξαρτάται από το μέγεθος της παραβίασης.

Η επιλογή των συντελεστών $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ είναι ιδιαίτερα σημαντική. Αν οι συντελεστές είναι υπερβολικά μικροί, τότε οι ποινές δεν θα έχουν επαρκή επίδραση και ο αλγόριθμος μπορεί να διατηρεί ή να προωθεί λύσεις που είναι μεν αποδοτικές ως προς το tracking error αλλά μη εφικτές. Αντίθετα, αν οι συντελεστές είναι υπερβολικά μεγάλοι, τότε η διαδικασία επιλογής μπορεί να γίνεται υπερβολικά αυστηρή και να αποκλείει λύσεις που θα μπορούσαν να γίνουν εφικτές μέσω μικρής επιδιόρθωσης. Συνεπώς, οι τιμές αυτές πρέπει να επιλέγονται με προσοχή και να ελέγχονται πειραματικά, ώστε να επιτυγχάνεται ισορροπία ανάμεσα στη βελτιστοποίηση του στόχου και στην τήρηση των περιορισμών.

Στην παρούσα εργασία, η βασική λογική είναι ότι η συνάρτηση καταλληλότητας πρέπει πρωτίστως να αντανακλά τον χρηματοοικονομικό στόχο της αποτύπωσης δείκτη, χωρίς όμως να αγνοεί την εφικτότητα των λύσεων. Για τον λόγο αυτό, η προτεινόμενη fitness function θα βασίζεται στον συνδυασμό tracking error και ποινών περιορισμών, ενώ παράλληλα θα υποστηρίζεται από μηχανισμούς επιδιόρθωσης των λύσεων, ώστε να μειώνεται η ανάγκη υπερβολικά μεγάλων ποινών. Με άλλα λόγια, η συνάρτηση καταλληλότητας δεν θα λειτουργεί απομονωμένα, αλλά ως μέρος ευρύτερου πλαισίου διαχείρισης της ποιότητας και εφικτότητας των χαρτοφυλακίων.

Ένα ακόμη σημαντικό στοιχείο είναι ότι η συνάρτηση καταλληλότητας πρέπει να είναι υπολογιστικά αποδοτική. Επειδή ο Γενετικός Αλγόριθμος αξιολογεί μεγάλο αριθμό χρωμοσωμάτων σε κάθε γενιά και η διαδικασία επαναλαμβάνεται πολλές φορές, ο υπολογισμός του F πρέπει να είναι σχετικά απλός και σταθερός. Η χρήση του tracking error σε κλειστή μορφή και απλών ποινικών όρων εξυπηρετεί ακριβώς αυτή την ανάγκη, επιτρέποντας στον αλγόριθμο να παραμένει πρακτικά εφαρμόσιμος ακόμη και για σχετικά μεγάλα σύνολα δεδομένων.

3.11 Τελεστής επιλογής

Ο τελεστής επιλογής αποτελεί ένα από τα βασικότερα συστατικά του Γενετικού Αλγορίθμου, καθώς καθορίζει ποια άτομα του τρέχοντος πληθυσμού θα έχουν τη δυνατότητα να συμμετάσχουν στη διαδικασία αναπαραγωγής και, συνεπώς, να μεταφέρουν τα χαρακτηριστικά τους στις επόμενες γενιές. Μέσω του τελεστή αυτού υλοποιείται στην πράξη η βασική εξελικτική αρχή σύμφωνα με την οποία οι καλύτερες λύσεις πρέπει να έχουν αυξημένη πιθανότητα επιβίωσης και αναπαραγωγής, χωρίς όμως να αποκλείεται πλήρως η παρουσία λιγότερο αποδοτικών λύσεων. Η ισορροπία αυτή είναι ιδιαίτερα σημαντική, διότι επιτρέπει αφενός την προοδευτική βελτίωση της ποιότητας του πληθυσμού και αφετέρου τη διατήρηση της ποικιλίας, η οποία είναι αναγκαία για την αποφυγή πρόωρης σύγκλισης.

Στο πρόβλημα της αποτύπωσης δείκτη, ο ρόλος του τελεστή επιλογής είναι ακόμη πιο κρίσιμος, επειδή κάθε χρωμόσωμα αντιστοιχεί σε ένα πλήρες χαρτοφυλάκιο που συνδυάζει επιλογή τίτλων, κατανομή βαρών και συμμόρφωση προς περιορισμούς. Η επιλογή, επομένως, δεν αφορά απλώς αφηρημένες ακολουθίες γονιδίων, αλλά εναλλακτικές χρηματοοικονομικές λύσεις διαφορετικής ποιότητας. Ο τελεστής επιλογής πρέπει να ευνοεί εκείνα τα χαρτοφυλάκια που επιτυγχάνουν χαμηλότερο tracking error και μεγαλύτερο βαθμό εφικτότητας, χωρίς ωστόσο να περιορίζει υπερβολικά την εξερεύνηση του χώρου λύσεων.

Γενικά, ένας τελεστής επιλογής πρέπει να ικανοποιεί δύο βασικές απαιτήσεις. Πρώτον, να είναι επιλεκτικός, δηλαδή να προσδίδει μεγαλύτερη πιθανότητα αναπαραγωγής στις λύσεις υψηλής καταλληλότητας. Δεύτερον, να μην είναι υπερβολικά επιθετικός, ώστε να μην οδηγεί πολύ γρήγορα τον πληθυσμό σε ομογενοποίηση. Αν μόνο οι απολύτως καλύτερες λύσεις επιλέγονται διαρκώς, τότε ο πληθυσμός χάνει σε ποικιλία και ο αλγόριθμος ενδέχεται να παγιδευτεί πρόωρα σε περιορισμένη περιοχή του χώρου αναζήτησης. Αντίθετα, αν η επιλογή είναι υπερβολικά χαλαρή και δεν διαφοροποιεί επαρκώς τις καλές από τις αδύναμες λύσεις, τότε η εξελικτική διαδικασία γίνεται αργή και λιγότερο αποτελεσματική.

Στην παρούσα εργασία, προτείνεται η χρήση επιλογής μέσω τουρνουά. Η μέθοδος αυτή θεωρείται ιδιαίτερα κατάλληλη για το συγκεκριμένο πρόβλημα, επειδή είναι απλή στην υλοποίηση, υπολογιστικά αποδοτική και ταυτόχρονα προσφέρει ικανοποιητικό έλεγχο της επιλεκτικής πίεσης. Η βασική της λογική είναι ότι από τον πληθυσμό επιλέγεται τυχαία ένα μικρό υποσύνολο ατόμων και, ανάμεσα σε αυτά, επιλέγεται εκείνο με την υψηλότερη τιμή καταλληλότητας ως γονέας. Η διαδικασία επαναλαμβάνεται όσες φορές απαιτείται για να σχηματιστεί το σύνολο των γονέων που θα χρησιμοποιηθούν στη διασταύρωση.

Αν ο αριθμός των ατόμων που συμμετέχουν σε κάθε τουρνουά είναι q , τότε κάθε διαδικασία επιλογής μπορεί να περιγραφεί ως εξής: επιλέγονται τυχαία q χρωμοσώματα από τον πληθυσμό και στη συνέχεια επιλέγεται το χρωμόσωμα με τη μέγιστη τιμή καταλληλότητας F . Η παράμετρος q επηρεάζει άμεσα την ένταση της επιλογής. Όσο μεγαλύτερο είναι το q , τόσο αυξάνεται η πιθανότητα να επιλέγονται οι ισχυρότερες λύσεις και άρα τόσο ισχυρότερη γίνεται η επιλεκτική πίεση. Αντίθετα, μικρότερες τιμές του q ευνοούν περισσότερο τη διατήρηση ποικιλίας. Στην πράξη, μικρές ή μεσαίες τιμές του q είναι συνήθως καταλληλότερες, διότι παρέχουν καλή ισορροπία ανάμεσα στην πρόοδο της βελτιστοποίησης και στη διατήρηση της ερευνητικής ικανότητας του πληθυσμού.

Η επιλογή μέσω τουρνουά παρουσιάζει αρκετά πλεονεκτήματα σε σχέση με άλλες κλασικές μεθόδους επιλογής. Πρώτον, δεν απαιτεί κανονικοποίηση των τιμών καταλληλότητας ούτε ειδικούς υπολογισμούς πιθανοτήτων. Δεύτερον, παραμένει σταθερή ακόμη και όταν υπάρχουν μεγάλες διαφορές κλίμακας μεταξύ των τιμών fitness, κάτι που είναι σημαντικό όταν η συνάρτηση καταλληλότητας περιλαμβάνει όρους ποινής για παραβίαση περιορισμών. Τρίτον, η μέθοδος είναι εύκολα παραμετροποιήσιμη μέσω του μεγέθους του τουρνουά, γεγονός που επιτρέπει καλύτερο έλεγχο της συμπεριφοράς του αλγορίθμου.

Στο πλαίσιο του προβλήματος της αποτύπωσης δείκτη, το πλεονέκτημα αυτό είναι ιδιαίτερα χρήσιμο. Επειδή οι λύσεις μπορεί να διαφέρουν όχι μόνο ως προς το tracking error αλλά και ως προς το αν είναι πλήρως εφικτές ή ελαφρώς παραβατικές, είναι σημαντικό ο τελεστής επιλογής να μπορεί να λειτουργεί αξιόπιστα ακόμη και όταν οι τιμές της καταλληλότητας δεν ακολουθούν ομαλή ή στενή κατανομή. Η επιλογή μέσω τουρνουά ανταποκρίνεται καλά σε αυτή την ανάγκη και γι' αυτό θεωρείται κατάλληλη για το προτεινόμενο σχήμα.

Η διαδικασία επιλογής στην παρούσα εργασία μπορεί να περιγραφεί ως εξής:

Έστω πληθυσμός $\mathcal{P}$ μεγέθους N , όπου κάθε χρωμόσωμα $X^{(j)}$, για $j = 1, 2, \dots, N$, έχει τιμή καταλληλότητας $F(X^{(j)})$. Για κάθε επιλογή γονέα:

1. Επιλέγονται τυχαία q άτομα από τον πληθυσμό.
2. Συγκρίνονται οι τιμές καταλληλότητάς τους.

3. Επιλέγεται ως γονέας το άτομο με τη μέγιστη τιμή F .

Η διαδικασία αυτή επαναλαμβάνεται μέχρι να προκύψει ο απαιτούμενος αριθμός γονέων για τη φάση της αναπαραγωγής.

Σημαντικό στοιχείο του τελεστή επιλογής είναι ότι δεν μεταφέρει ακόμη άμεσα γενετική πληροφορία από τη μία γενιά στην άλλη· απλώς καθορίζει ποιοι θα έχουν τη δυνατότητα να το κάνουν μέσω της διασταύρωσης και της μετάλλαξης. Παρ' όλα αυτά, η επίδρασή του είναι καθοριστική, καθώς μέσω αυτού εγκαθίσταται η λεγόμενη επιλεκτική πίεση. Με τον όρο αυτό εννοείται η τάση του πληθυσμού να στρέφεται προοδευτικά προς λύσεις ανώτερης ποιότητας. Αν η πίεση αυτή είναι επαρκώς ρυθμισμένη, ο αλγόριθμος βελτιώνεται σταδιακά χωρίς να χάνει την ικανότητά του να εξερευνά νέες περιοχές του χώρου λύσεων.

Σε ορισμένες περιπτώσεις, η επιλογή μπορεί να συνδυαστεί και με μηχανισμό ελιτισμού. Δηλαδή, ένα μικρό ποσοστό των καλύτερων λύσεων της τρέχουσας γενιάς μπορεί να μεταφέρεται αυτούσιο στην επόμενη, χωρίς να υπόκειται σε διασταύρωση ή μετάλλαξη. Ο μηχανισμός αυτός δεν υποκαθιστά τον τελεστή επιλογής, αλλά λειτουργεί συμπληρωματικά προς αυτόν, διασφαλίζοντας ότι οι καλύτερες μέχρι στιγμής λύσεις δεν θα χαθούν εξαιτίας στοχαστικών μεταβολών. Η χρήση ελιτισμού είναι ιδιαίτερα χρήσιμη σε προβλήματα όπως το index tracking, όπου η εύρεση μιας πολύ καλής και εφικτής λύσης έχει υψηλή αξία και πρέπει να προστατεύεται από τυχαία υποβάθμιση.

3.12 Τελεστής διασταύρωσης

Ο τελεστής διασταύρωσης αποτελεί τον βασικό μηχανισμό μέσω του οποίου ο Γενετικός Αλγόριθμος συνδυάζει πληροφορία από δύο γονικές λύσεις και παράγει νέους απογόνους. Ο τελεστής διασταύρωσης που χρησιμοποιήθηκε στην εργασία αποτελεί καθιερωμένο γενετικό τελεστή και έχει χρησιμοποιηθεί σε εφαρμογές Γενετικών Αλγορίθμων για προβλήματα χαρτοφυλακίου και index tracking (Oh et al., 2005). Στην παρούσα εργασία ο τελεστής προσαρμόζεται στην αναπαράσταση των χαρτοφυλακίων και στους περιορισμούς του προβλήματος. Η λειτουργία του είναι καθοριστική, διότι μέσω αυτής της διαδικασίας επιχειρείται η κληρονομία και ο συνδυασμός ευνοϊκών χαρακτηριστικών από διαφορετικά χρωμοσώματα, με στόχο τη δημιουργία χαρτοφυλακίων που να είναι δυναμικά καλύτερα από τους γονείς τους. Στο πρόβλημα της αποτύπωσης δείκτη, όπου κάθε χρωμόσωμα εκφράζει ταυτόχρονα επιλογή τίτλων και κατανομή βαρών, ο σχεδιασμός του τελεστή διασταύρωσης πρέπει να είναι ιδιαίτερα προσεκτικός, ώστε να συνδυάζει αποτελεσματικά τις δύο αυτές διαστάσεις χωρίς να οδηγεί συστηματικά σε μη εφικτές λύσεις.

Η γενική λογική της διασταύρωσης είναι ότι δύο γονικές λύσεις, έστω $X^{(1)}$ και $X^{(2)}$, χρησιμοποιούνται για την παραγωγή ενός ή περισσότερων απογόνων, οι οποίοι κληρονομούν τμήματα της δομής τους από τους γονείς. Στο πλαίσιο του προβλήματος που εξετάζεται, αυτό σημαίνει ότι ένας απόγονος μπορεί να κληρονομεί μέρος του συνόλου των επιλεγμένων τίτλων από τον πρώτο γονέα και μέρος από τον δεύτερο, ενώ παράλληλα μπορεί να κληρονομεί ή να συνδυάζει τα αντίστοιχα βάρη τους. Η ιδέα αυτή είναι ιδιαίτερα ελκυστική, επειδή επιτρέπει στον αλγόριθμο να αξιοποιεί υπάρχουσες καλές δομές χαρτοφυλακίου και να δημιουργεί νέες λύσεις που ενδεχομένως ενσωματώνουν τα πλεονεκτήματα και των δύο γονικών συνθέσεων.

Ωστόσο, στο πρόβλημα της αποτύπωσης δείκτη η διασταύρωση δεν μπορεί να εφαρμοστεί με εντελώς αφελή τρόπο. Αν ένας τελεστής διασταύρωσης σχεδιαστεί χωρίς να λαμβάνει υπόψη τους περιορισμούς του μοντέλου, είναι πολύ πιθανό να παράγει χαρτοφυλάκια που παραβιάζουν τον περιορισμό προϋπολογισμού, υπερβαίνουν το επιτρεπτό πλήθος τίτλων ή περιλαμβάνουν βάρη εκτός των επιτρεπτών ορίων. Για τον λόγο αυτό, ο προτεινόμενος τελεστής πρέπει να είναι δομημένος έτσι ώστε να συνδυάζει πληροφορία από τους γονείς με όσο το δυνατόν πιο συνεπή τρόπο και να συνοδεύεται, όπου χρειάζεται, από διαδικασίες επιδιόρθωσης ή κανονικοποίησης.

Στην παρούσα εργασία προτείνεται μια διασταύρωση δύο σταδίων, η οποία ανταποκρίνεται στη μικτή φύση του χρωμοσώματος. Στο πρώτο στάδιο, η διασταύρωση λειτουργεί στο επίπεδο της επιλογής τίτλων. Δηλαδή, από τα δύο γονικά χαρτοφυλάκια προσδιορίζεται ένα νέο υποσύνολο τίτλων, το οποίο προκύπτει από συνδυασμό των ενεργών τίτλων των δύο γονέων. Στο δεύτερο στάδιο, καθορίζονται τα βάρη των τίτλων του απογόνου, με τρόπο που λαμβάνει υπόψη τις αντίστοιχες βαρύτητες των γονικών λύσεων και στη συνέχεια εφαρμόζεται κανονικοποίηση ώστε να ικανοποιείται ο περιορισμός προϋπολογισμού.

Πιο συγκεκριμένα, αν $S^{(1)}$ και $S^{(2)}$ είναι τα σύνολα των ενεργών τίτλων των δύο γονέων, τότε το σύνολο υποψήφιων τίτλων του απογόνου μπορεί να οριστεί αρχικά ως το ενωσιακό σύνολο:

$$S^* = S^{(1)} \cup S^{(2)} \quad (3.23)$$

Το σύνολο αυτό περιλαμβάνει όλους τους τίτλους που εμφανίζονται τουλάχιστον σε έναν από τους δύο γονείς. Αν το πλήθος των τίτλων του S^* είναι μικρότερο ή ίσο του επιτρεπτού ορίου K , τότε το σύνολο αυτό μπορεί να αποτελέσει άμεσα τη βάση του απογόνου. Αν όμως το πλήθος του υπερβαίνει το K , τότε απαιτείται διαδικασία επιλογής ενός επιτρεπτού υποσυνόλου. Σε αυτή την περίπτωση, μπορούν να διατηρηθούν οι τίτλοι με ισχυρότερη παρουσία στους γονείς, δηλαδή εκείνοι που εμφανίζονται και στους δύο ή εκείνοι που συνοδεύονται από μεγαλύτερα γονικά βάρη, ενώ οι υπόλοιποι αποκόπτονται μέχρι να ικανοποιηθεί ο περιορισμός πληθικότητας.

Αφού καθοριστεί το σύνολο τίτλων του απογόνου, ακολουθεί η απόδοση αρχικών βαρών. Για κάθε τίτλο i που περιλαμβάνεται στον απόγονο, μπορεί να οριστεί ένα προσωρινό βάρος ως γραμμικός συνδυασμός των αντίστοιχων γονικών βαρών:

$$\tilde{w}_i = \alpha w_i^{(1)} + (1 - \alpha) w_i^{(2)} \quad (3.24)$$

όπου $\alpha \in [0,1]$ είναι τυχαία ή προκαθορισμένη παράμετρος μίξης, και $w_i^{(1)}$, $w_i^{(2)}$ είναι τα βάρη του τίτλου ίσους δύο γονείς, με την παραδοχή ότι αν ένας τίτλος δεν υπάρχει σε έναν γονέα, τότε το αντίστοιχο βάρος λαμβάνεται ίσο με μηδέν. Η σχέση αυτή επιτρέπει στον απόγονο να κληρονομεί σταδιακά χαρακτηριστικά από αμφότερους τους γονείς, χωρίς να αντιγράφει μηχανιστικά κανέναν από τους δύο.

Τα προσωρινά αυτά βάρη δεν είναι απαραίτητα άμεσα εφικτά, καθώς μπορεί να μην αθροίζονται στη μονάδα ή να παραβιάζουν τα επιτρεπτά όρια συμμετοχής. Για τον λόγο αυτό, μετά τη διασταύρωση εφαρμόζεται στάδιο κανονικοποίησης. Πρώτα, τα βάρη επανακλιμακώνονται έτσι ώστε να ισχύει:

$$w'_i = \frac{\tilde{w}_i}{\sum_{j \in S^*} \tilde{w}_j}, \forall i \in S^* \quad (3.25)$$

ώστε να διασφαλίζεται ότι:

$$\sum_{i \in S^*} w'_i = 1 \quad (3.26)$$

Στη συνέχεια, ελέγχεται αν κάθε βάρος w'_i βρίσκεται μέσα στα επιτρεπτά όρια $[l_i, u_i]$. Αν κάποιο βάρος είναι μικρότερο από το κατώτατο επιτρεπτό όριο ή μεγαλύτερο από το ανώτατο, εφαρμόζεται διαδικασία επιδιόρθωσης, η οποία αναπροσαρμόζει τα βάρη ώστε να επανέλθουν εντός ορίων και στη συνέχεια επανακανονικοποιεί το χαρτοφυλάκιο.

Η παραπάνω προσέγγιση έχει σημαντικά πλεονεκτήματα. Πρώτον, η διασταύρωση δεν αντιμετωπίζει το χρωμόσωμα ως απλή ακολουθία τυχαίων γονιδίων, αλλά ως δομημένη αναπαράσταση χαρτοφυλακίου. Αυτό σημαίνει ότι ο συνδυασμός των γονέων πραγματοποιείται σε επίπεδο με πραγματικό επενδυτικό νόημα. Δεύτερον, η χρήση του ενωσιακού συνόλου τίτλων και της μίξης βαρών επιτρέπει τη σύνθεση νέων χαρτοφυλακίων που αξιοποιούν πληροφορία και από τους δύο γονείς, ενισχύοντας έτσι τη δυνατότητα δημιουργίας ποιοτικών απογόνων. Τρίτον, η κανονικοποίηση και η επιδιόρθωση συμβάλλουν στη διατήρηση της εφικτότητας, περιορίζοντας τον αριθμό λύσεων που θα έπρεπε αργότερα να απορριφθούν ή να επιβαρυνθούν έντονα από ποινές.

Σημαντικό ρόλο στην αποτελεσματικότητα του τελεστή παίζει και η πιθανότητα διασταύρωσης, η οποία συνήθως συμβολίζεται με p_c . Η παράμετρος αυτή εκφράζει την πιθανότητα με την οποία δύο επιλεγμένοι γονείς θα υποβληθούν πράγματι στη διαδικασία διασταύρωσης. Αν η διασταύρωση δεν πραγματοποιηθεί, οι γονείς μπορούν να μεταφερθούν αυτούσιοι ή σχεδόν αυτούσιοι στην επόμενη φάση του αλγορίθμου. Η επιλογή της τιμής του p_c είναι σημαντική, καθώς επηρεάζει τον βαθμό στον οποίο ο αλγόριθμος βασίζεται στον συνδυασμό υφιστάμενων λύσεων για τη δημιουργία νέων. Τυπικά, υψηλές τιμές του p_c ενισχύουν τον εξελικτικό χαρακτήρα του αλγορίθμου, αλλά πρέπει να συνδυάζονται με κατάλληλους μηχανισμούς ελέγχου ώστε να μην οδηγούν σε υπερβολικά ασταθείς απογόνους.

Πρέπει επίσης να επισημανθεί ότι ο τελεστής διασταύρωσης δεν λειτουργεί απομονωμένα, αλλά σε άμεση συνεργασία με τον τελεστή μετάλλαξης και με τους μηχανισμούς επιδιόρθωσης. Η διασταύρωση είναι ο κύριος μηχανισμός ανασυνδυασμού υπάρχουσας πληροφορίας, όμως δεν είναι από μόνη της επαρκής για να εξασφαλίσει υψηλή ποικιλία στον πληθυσμό. Η μετάλλαξη είναι εκείνη που θα επιτρέψει την είσοδο νέων χαρακτηριστικών και μικρών αλλά κρίσιμων αποκλίσεων από τα γονικά πρότυπα. Για τον λόγο αυτό, η ποιότητα του συνολικού αλγορίθμου εξαρτάται από την ισορροπία μεταξύ των δύο αυτών τελεστών.

3.13 Τελεστής μετάλλαξης

Ο τελεστής μετάλλαξης αποτελεί ένα από τα βασικά στοιχεία της εξελικτικής διαδικασίας, καθώς λειτουργεί ως μηχανισμός εισαγωγής νέας πληροφορίας στον πληθυσμό και συμβάλλει αποφασιστικά στη διατήρηση της ποικιλίας των λύσεων. Ενώ ο τελεστής διασταύρωσης συνδυάζει υπάρχοντα χαρακτηριστικά από δύο γονικά χρωμοσώματα, η

μετάλλαξη επιτρέπει την παραγωγή νέων χαρακτηριστικών που δεν προκύπτουν απαραίτητα από άμεσο ανασυνδυασμό. Στο πρόβλημα της αποτύπωσης δείκτη, η σημασία της μετάλλαξης είναι ιδιαίτερα μεγάλη, διότι ο χώρος λύσεων είναι σύνθετος, συνδυαστικός και υψηλής διάστασης, γεγονός που αυξάνει τον κίνδυνο πρόωρης σύγκλισης του αλγορίθμου σε τοπικά βέλτιστα.

Η βασική λειτουργία της μετάλλαξης είναι να πραγματοποιεί μικρές, ελεγχόμενες τροποποιήσεις σε ένα χρωμόσωμα, έτσι ώστε να δημιουργούνται νέες παραλλαγές υποψήφιων χαρτοφυλακίων. Οι τροποποιήσεις αυτές δεν πρέπει να είναι υπερβολικά μεγάλες, διότι τότε θα κατέστρεφαν τη συσσωρευμένη πληροφορία που έχει αποκτήσει ο πληθυσμός μέσω της επιλογής και της διασταύρωσης. Από την άλλη πλευρά, δεν πρέπει να είναι και τόσο περιορισμένες ώστε να μην επιφέρουν ουσιαστική διαφοροποίηση. Συνεπώς, ο τελεστής μετάλλαξης πρέπει να ισορροπεί ανάμεσα στην εξερεύνηση και στη σταδιακή βελτίωση, διατηρώντας τον πληθυσμό αρκετά ευέλικτο ώστε να ανακαλύπτει νέες περιοχές του χώρου λύσεων.

Στην παρούσα εργασία, η μετάλλαξη πρέπει να σχεδιαστεί έτσι ώστε να ανταποκρίνεται στη μικτή φύση του χρωμοσώματος. Εφόσον κάθε χρωμόσωμα περιλαμβάνει ταυτόχρονα πληροφορία για την επιλογή τίτλων και για τα βάρη τους, ο τελεστής μετάλλαξης δεν μπορεί να περιορίζεται μόνο σε μία από αυτές τις δύο διαστάσεις. Αντίθετα, πρέπει να δίνεται η δυνατότητα είτε τροποποίησης της σύνθεσης του χαρτοφυλακίου είτε ανακατανομής των βαρών μεταξύ των ήδη επιλεγμένων τίτλων είτε συνδυασμού των δύο. Με τον τρόπο αυτό, η μετάλλαξη θα μπορεί να επηρεάζει ουσιαστικά τόσο τη διακριτή όσο και τη συνεχή διάσταση του προβλήματος.

Μια πρώτη μορφή μετάλλαξης αφορά τη μεταβολή της σύνθεσης του χαρτοφυλακίου. Σε αυτή την περίπτωση, ένας ενεργός τίτλος μπορεί να αφαιρεθεί από το χαρτοφυλάκιο και να αντικατασταθεί από έναν άλλον, ο οποίος μέχρι εκείνη τη στιγμή δεν συμμετείχε. Αν δείναι το σύνολο των ενεργών τίτλων ενός χρωμοσώματος και $I \setminus S$ το σύνολο των μη ενεργών τίτλων, τότε μια τέτοια μετάλλαξη μπορεί να αποδοθεί εννοιολογικά ως εξής:

$$S' = (S \setminus \{i\}) \cup \{j\} \quad (3.27)$$

όπου $i \in S$ και $j \in I \setminus S$. Η κίνηση αυτή αλλάζει τη σύνθεση του χαρτοφυλακίου, αλλά διατηρεί σταθερό το πλήθος των επιλεγμένων τίτλων, γεγονός που είναι ιδιαίτερα χρήσιμο όταν υπάρχει ενεργός cardinality constraint.

Μια δεύτερη μορφή μετάλλαξης αφορά τη μεταβολή των βαρών των επιλεγμένων τίτλων χωρίς αλλαγή της σύνθεσης. Σε αυτή την περίπτωση, μπορεί να επιλεγεί τυχαία ένας ή περισσότεροι ενεργοί τίτλοι και να μεταβληθούν ελαφρώς τα βάρη τους. Για παράδειγμα, αν ένας τίτλος i είναι ενεργός, το βάρος του μπορεί να τροποποιηθεί ως:

$$\tilde{w}_i = w_i + \delta \quad (3.28)$$

όπου δείναι μια μικρή τυχαία διαταραχή, θετική ή αρνητική. Προκειμένου να διατηρηθεί ο περιορισμός προϋπολογισμού, η μεταβολή αυτή πρέπει να συνοδεύεται από αντίστοιχη αναπροσαρμογή στα βάρη των υπόλοιπων ενεργών τίτλων. Μετά τη διαδικασία αυτή, απαιτείται εκ νέου κανονικοποίηση ώστε να ισχύει:

$$\sum_{i=1}^n w_i = 1 \quad (3.29)$$

και ταυτόχρονα να τηρούνται τα επιτρεπτά όρια συμμετοχής.

Μια τρίτη, πιο σύνθετη μορφή μετάλλαξης, συνδυάζει τα δύο παραπάνω στοιχεία. Δηλαδή, επιτρέπει τόσο την αντικατάσταση ενός τίτλου όσο και την επακόλουθη ανακατανομή των βαρών. Η μορφή αυτή είναι ιδιαίτερα κατάλληλη για το πρόβλημα της αποτύπωσης δείκτη, διότι η αλλαγή ενός μόνο τίτλου χωρίς προσαρμογή των βαρών μπορεί να μην είναι αρκετή για να αποδώσει τη βέλτιστη νέα σύνθεση. Με τη συνδυαστική αυτή μετάλλαξη, ο απόγονος μπορεί να αποκτήσει ουσιαστικά νέα δομή χαρτοφυλακίου, χωρίς όμως να απομακρύνεται υπερβολικά από τη γονική του λύση.

Η πιθανότητα εφαρμογής του τελεστή μετάλλαξης συμβολίζεται συνήθως με p_m . Η παράμετρος αυτή εκφράζει την πιθανότητα με την οποία ένα χρωμόσωμα ή ένα επιμέρους γονίδιο θα υποστεί μετάλλαξη κατά τη διαδικασία εξέλιξης. Η επιλογή της τιμής του p_m είναι ιδιαίτερα κρίσιμη. Αν η πιθανότητα μετάλλαξης είναι πολύ μικρή, τότε ο αλγόριθμος μπορεί να χάσει τη δυνατότητα εξερεύνησης νέων περιοχών και να συγκλίνει πολύ γρήγορα γύρω από περιορισμένο σύνολο λύσεων. Αντίθετα, αν η πιθανότητα είναι υπερβολικά μεγάλη, τότε ο αλγόριθμος ενδέχεται να λειτουργεί περισσότερο σαν τυχαία αναζήτηση, καταστρέφοντας συστηματικά τις καλές δομές που έχουν ήδη σχηματιστεί μέσω επιλογής και διασταύρωσης. Συνεπώς, απαιτείται προσεκτική παραμετροποίηση ώστε η μετάλλαξη να παίζει τον σωστό ενισχυτικό και όχι αποδιοργανωτικό ρόλο.

Στην παρούσα μεθοδολογία, η μετάλλαξη δεν αντιμετωπίζεται ως απλή πηγή τυχειότητας, αλλά ως στοχευμένος μηχανισμός ελεγχόμενης διαφοροποίησης. Οι αλλαγές που προκαλεί πρέπει να είναι τέτοιες ώστε να διατηρούν το χρωμόσωμα μέσα σε εύλογη απόσταση από τη γονική λύση, αλλά και να είναι αρκετά ουσιαστικές ώστε να επιτρέπουν την ανακάλυψη νέων χαρτοφυλακίων με χαμηλότερο tracking error. Για τον λόγο αυτό, η μετάλλαξη εφαρμόζεται συνήθως σε μικρό ποσοστό των λύσεων ή των γονιδίων, και ακολουθείται από έλεγχο εφικτότητας και, όπου χρειάζεται, από διαδικασία επιδιόρθωσης.

Η ανάγκη επιδιόρθωσης μετά τη μετάλλαξη είναι ιδιαίτερα σημαντική στο πρόβλημα της αποτύπωσης δείκτη. Η αντικατάσταση τίτλων ή η μεταβολή βαρών μπορεί να οδηγήσει σε προσωρινή παραβίαση των περιορισμών του προβλήματος. Για παράδειγμα, μια μικρή διαταραχή στα βάρη μπορεί να έχει ως αποτέλεσμα κάποιο βάρος να γίνει μικρότερο από το κατώτατο επιτρεπτό όριο ή μεγαλύτερο από το ανώτατο. Αντίστοιχα, η εισαγωγή ενός νέου τίτλου μπορεί να απαιτεί ανακατανομή του συνολικού βάρους ώστε να εξακολουθεί να ισχύει ο περιορισμός προϋπολογισμού. Επομένως, μετά από κάθε πράξη μετάλλαξης είναι αναγκαίος ένας μηχανισμός κανονικοποίησης και ελέγχου, ώστε το τελικό χρωμόσωμα να παραμένει, στο μέτρο του δυνατού, εφικτό και χρηματοοικονομικά ερμηνεύσιμο.

Από αλγοριθμική άποψη, η μετάλλαξη παίζει επίσης σημαντικό ρόλο στη μακροχρόνια δυναμική του πληθυσμού. Σε πρώιμα στάδια της εξέλιξης, όταν ο πληθυσμός είναι ακόμη αρκετά διαφοροποιημένος, ο ρόλος της μπορεί να είναι σχετικά περιορισμένος, καθώς η ποικιλία εξασφαλίζεται ήδη σε σημαντικό βαθμό από την αρχικοποίηση και τη διασταύρωση. Καθώς όμως οι γενιές προχωρούν και ο πληθυσμός αρχίζει να ομογενοποιείται, η μετάλλαξη αποκτά όλο και μεγαλύτερη σημασία ως εργαλείο αποφυγής

πρόωρης σύγκλισης. Υπό αυτή την έννοια, λειτουργεί ως μηχανισμός «ανανέωσης» της εξελικτικής διαδικασίας.

Σημαντικό είναι επίσης ότι η μετάλλαξη μπορεί να ενισχύσει τη δυνατότητα εντοπισμού λεπτών βελτιώσεων σε χαρτοφυλάκια που είναι ήδη αρκετά καλά. Σε ένα στάδιο όπου οι περισσότερες λύσεις του πληθυσμού έχουν παρόμοια ποιότητα, μικρές μεταβολές στη σύνθεση ή στην κατανομή βαρών μπορεί να είναι εκείνες που θα οδηγήσουν σε ουσιαστική διαφορά στην τιμή του tracking error. Έτσι, η μετάλλαξη δεν συμβάλλει μόνο στην εξερεύνηση του χώρου λύσεων, αλλά και στη λεπτομερή τοπική βελτιστοποίηση όταν ο αλγόριθμος βρίσκεται ήδη κοντά σε καλές περιοχές του χώρου αναζήτησης.

3.14 Μηχανισμός επιδιόρθωσης και διαχείρισης περιορισμών

Η λειτουργία των τελεστών διασταύρωσης και μετάλλαξης, παρότι είναι απαραίτητη για την παραγωγή νέων και διαφοροποιημένων λύσεων, συνοδεύεται από έναν βασικό κίνδυνο: τη δημιουργία χρωμοσωμάτων που δεν ικανοποιούν πλήρως τους περιορισμούς του προβλήματος. Στο πρόβλημα της αποτύπωσης δείκτη, η δυσκολία αυτή είναι ιδιαίτερα έντονη, διότι το χαρτοφυλάκιο πρέπει ταυτόχρονα να σέβεται τον περιορισμό προϋπολογισμού, τον περιορισμό πληθικότητας, τα κατώτατα και ανώτατα όρια συμμετοχής και, ενδεχομένως, περιορισμούς που συνδέονται με το κόστος συναλλαγών. Για τον λόγο αυτό, ο προτεινόμενος Γενετικός Αλγόριθμος πρέπει να περιλαμβάνει ειδικό μηχανισμό επιδιόρθωσης και διαχείρισης περιορισμών, ώστε οι υποψήφιες λύσεις να παραμένουν όσο το δυνατόν πιο κοντά στον εφικτό χώρο του προβλήματος.

Ο βασικός σκοπός του μηχανισμού αυτού είναι τριπλός. Πρώτον, να εντοπίζει γρήγορα τις παραβιάσεις περιορισμών που εμφανίζονται μετά από πράξεις διασταύρωσης ή μετάλλαξης. Δεύτερον, να προσαρμόζει το χρωμόσωμα με τέτοιο τρόπο ώστε να αποκαθίσταται η συμβατότητά του με το μοντέλο, χωρίς όμως να καταστρέφεται πλήρως η πληροφορία που έχει προκύψει από τη γενετική διαδικασία. Τρίτον, να συντονίζει τη σχέση ανάμεσα στις ρητές διορθώσεις, στους κανόνες εφικτότητας και στις ποινές της συνάρτησης καταλληλότητας, έτσι ώστε η συνολική διαχείριση των περιορισμών να είναι συνεπής και αποδοτική. Η δεύτερη αυτή διάσταση είναι ιδιαίτερα σημαντική, διότι ένας υπερβολικά επιθετικός μηχανισμός επιδιόρθωσης θα μπορούσε να αναιρεί ουσιαστικά τη λειτουργία των γενετικών τελεστών, επαναφέροντας διαρκώς τις λύσεις σε παρόμοια μορφή.

Στην παρούσα εργασία, η διαχείριση περιορισμών βασίζεται σε τρία συμπληρωματικά επίπεδα. Το πρώτο επίπεδο είναι η προληπτική διαχείριση, δηλαδή ο σχεδιασμός τελεστών και διαδικασιών αρχικοποίησης που παράγουν κατά το δυνατόν εφικτές λύσεις. Το δεύτερο επίπεδο είναι η διορθωτική διαχείριση, μέσω της οποίας οι απόγονοι που προκύπτουν από τους γενετικούς τελεστές επιδιορθώνονται ώστε να ικανοποιούν τους βασικούς περιορισμούς. Το τρίτο επίπεδο είναι η ποινική διαχείριση, δηλαδή η επιβολή ποινών στη συνάρτηση καταλληλότητας όταν μια λύση παραμένει μερικώς μη εφικτή ή όταν κρίνεται προτιμότερο να μην απορρίπτεται άμεσα αλλά να αξιολογείται δυσμενέστερα. Ο συνδυασμός αυτών των τριών επιπέδων επιτρέπει πιο ευέλικτη και ρεαλιστική διαχείριση του προβλήματος από ό,τι μια απόλυτα άκαμπτη προσέγγιση.

Η πρώτη και πιο βασική μορφή ελέγχου αφορά τον περιορισμό προϋπολογισμού. Μετά από κάθε παραγωγή απογόνου, πρέπει να ελέγχεται αν το άθροισμα των βαρών ισούται με τη μονάδα. Αν δηλαδή για ένα χρωμόσωμα ισχύει:

$$\sum_{i=1}^n w_i \neq 1 \quad (3.30)$$

τότε απαιτείται κανονικοποίηση. Η απλούστερη μορφή επιδιόρθωσης είναι να επανακλιμακωθούν όλα τα ενεργά βάρη ως εξής:

$$w'_i = \frac{w_i}{\sum_{j=1}^n w_j}, \forall i \in I \text{ με } w_i > 0 \quad (3.31)$$

ώστε να διασφαλίζεται ότι:

$$\sum_{i=1}^n w'_i = 1 \quad (3.32)$$

Η κανονικοποίηση αυτή είναι απλή, υπολογιστικά αποδοτική και επιτρέπει τη διατήρηση της σχετικής δομής των βαρών, διορθώνοντας ταυτόχρονα τη συνολική τους κλίμακα.

Ο δεύτερος κρίσιμος έλεγχος αφορά τη συνέπεια ανάμεσα στις δυαδικές μεταβλητές επιλογής και στα αντίστοιχα βάρη. Αν ένας τίτλος δεν έχει επιλεγεί, δηλαδή αν $z_i = 0$, τότε το βάρος του πρέπει να είναι μηδενικό. Αντίστοιχα, αν $z_i = 1$, τότε ο τίτλος πρέπει να έχει βάρος συμβατό με τα επιτρεπτά όρια. Συνεπώς, μετά από κάθε γενετική πράξη πρέπει να ελέγχεται αν ισχύει η συνθήκη:

$$z_i = 0 \Rightarrow w_i = 0 \quad (3.33)$$

Αν διαπιστωθεί ότι κάποιος τίτλος έχει θετικό βάρος χωρίς να είναι ενεργός, τότε είτε το βάρος του μηδενίζεται είτε η μεταβλητή επιλογής του ενεργοποιείται, ανάλογα με τη γενικότερη λογική της επιδιόρθωσης. Στην παρούσα εργασία προτιμάται η πρώτη λύση, δηλαδή η άμεση μηδενικοποίηση του βάρους μη ενεργών τίτλων, καθώς αυτή είναι πιο συνεπής με τη λογική του περιορισμού πληθικότητας και της δομής του χρωμοσώματος.

Ο περιορισμός πληθικότητας απαιτεί επίσης ιδιαίτερη διαχείριση. Αν μετά από διασταύρωση ή μετάλλαξη προκύψει χρωμόσωμα με αριθμό ενεργών τίτλων μεγαλύτερο από το επιτρεπτό όριο K , δηλαδή αν:

$$\sum_{i=1}^n z_i > K \quad (3.34)$$

τότε πρέπει να αφαιρεθούν ορισμένοι τίτλοι από το χαρτοφυλάκιο. Η απομάκρυνση αυτή δεν πρέπει να γίνεται αυθαίρετα. Μια εύλογη στρατηγική είναι να αφαιρούνται πρώτα οι τίτλοι με τα μικρότερα βάρη, καθώς αυτοί είναι πιθανό να έχουν μικρότερη συμβολή στη συνολική συμπεριφορά του χαρτοφυλακίου. Δηλαδή, ταξινομούνται οι ενεργοί τίτλοι κατά αύξουσα σειρά βαρών και απομακρύνονται διαδοχικά μέχρι να ικανοποιηθεί ο περιορισμός:

$$\sum_{i=1}^n z_i \leq K \quad (3.35)$$

Μετά την αφαίρεση των τίτλων αυτών, τα υπόλοιπα βάρη επανακανονικοποιούνται ώστε να διατηρείται ο περιορισμός προϋπολογισμού. Αντίστροφα, σε περίπτωση που το χρωμόσωμα περιέχει λιγότερους τίτλους από το επιθυμητό πλήθος, εφόσον το μοντέλο απαιτεί ακριβώς K τίτλους, μπορεί να εισάγονται νέοι τίτλοι με πολύ μικρά αρχικά βάρη και στη συνέχεια να γίνεται επανακανονικοποίηση και έλεγχος ορίων.

Ανάλογα σημαντικός είναι και ο έλεγχος των κατώτατων και ανώτατων ορίων συμμετοχής. Για κάθε ενεργό τίτλο πρέπει να ισχύει:

$$l_i z_i \leq w_i \leq u_i z_i \quad (3.36)$$

Αν κάποιος τίτλος έχει βάρος μικρότερο από το κατώτατο επιτρεπτό όριο, τότε υπάρχουν δύο βασικές επιλογές επιδιόρθωσης. Η πρώτη είναι να αυξηθεί το βάρος του μέχρι το επίπεδο l_i , με αντίστοιχη μείωση βαρών άλλων τίτλων. Η δεύτερη είναι να αφαιρεθεί πλήρως από το χαρτοφυλάκιο, θέτοντας $z_i = 0$ και $w_i = 0$. Στην παρούσα εργασία, όταν το βάρος ενός τίτλου είναι πολύ μικρό και η παρουσία του δεν είναι ουσιαστική, θεωρείται προτιμότερο να αφαιρείται από το χαρτοφυλάκιο, καθώς αυτό είναι πιο συνεπές με τη λογική αποφυγής αμελητέων θέσεων. Αντίθετα, όταν ένας τίτλος υπερβαίνει το ανώτατο όριο, το βάρος του περιορίζεται στην τιμή u_i , και το πλεονάζον ποσό ανακατανέμεται στους υπόλοιπους ενεργούς τίτλους.

Μια γενική διαδικασία επιδιόρθωσης για τα όρια βαρών μπορεί να περιγραφεί ως εξής:

1. Εντοπισμός τίτλων με $w_i < l_i$ ή $w_i > u_i$.
2. Για τίτλους με υπέρβαση του ανώτατου ορίου, θέτουμε $w_i = u_i$.
3. Για τίτλους με πολύ μικρά βάρη, είτε θέτουμε $w_i = l_i$ είτε αφαιρούμε τον τίτλο από το χαρτοφυλάκιο.
4. Επανακανονικοποίηση των υπολοίπων βαρών.
5. Επαναληπτικός έλεγχος μέχρι να μην υπάρχουν παραβιάσεις.

Η επαναληπτική φύση αυτής της διαδικασίας είναι αναγκαία, διότι η διόρθωση ενός περιορισμού μπορεί προσωρινά να οδηγήσει σε νέα παραβίαση κάποιου άλλου. Για παράδειγμα, η αύξηση ενός βάρους ώστε να φτάσει το κατώτατο όριο μπορεί να επηρεάσει το συνολικό άθροισμα των βαρών ή να προκαλέσει υπέρβαση σε άλλο τίτλο μετά την ανακατανομή. Συνεπώς, ο μηχανισμός επιδιόρθωσης πρέπει να λειτουργεί ως ακολουθία ελέγχων και αναπροσαρμογών και όχι ως μία μόνο μεμονωμένη πράξη.

Αν το μοντέλο περιλαμβάνει περιορισμό κόστους συναλλαγών, τότε απαιτείται και αντίστοιχη διαχείριση του συνολικού κόστους της μετάβασης από το αρχικό στο νέο χαρτοφυλάκιο. Αν ισχύει:

$$\sum_{i=1}^n c_i |w_i - w_i^0| > C_{\max} \quad (3.37)$$

τότε το χρωμόσωμα παραβιάζει το αποδεκτό επίπεδο κόστους. Σε αυτή την περίπτωση, η επιδιόρθωση μπορεί να βασιστεί στη σταδιακή μετακίνηση του νέου χαρτοφυλακίου προς την αρχική δομή w_i^0 , μειώνοντας το μέγεθος των μεταβολών μέχρι να ικανοποιείται το όριο. Μια απλή μορφή τέτοιας προσαρμογής είναι:

$$w_i' = w_i^0 + \beta(w_i - w_i^0), 0 < \beta < 1 \quad (3.38)$$

όπου ο συντελεστής β επιλέγεται έτσι ώστε να περιορίζεται η συνολική συναλλακτική επιβάρυνση. Η προσέγγιση αυτή μειώνει την απόσταση μεταξύ νέας και αρχικής λύσης χωρίς να ακυρώνει εντελώς τη βελτιστοποιητική παρέμβαση.

Πέρα από τον μηχανισμό επιδιόρθωσης, η διαχείριση περιορισμών περιλαμβάνει και ρητό έλεγχο εφικτότητας μετά από κάθε γενετική πράξη. Δηλαδή, κάθε απόγονος που παράγεται από διασταύρωση ή μετάλλαξη πρέπει να ελέγχεται με βάση το πλήρες σύνολο περιορισμών πριν εισέλθει στον νέο πληθυσμό. Μπορεί να οριστεί μια λογική μεταβλητή εφικτότητας $\phi(X)$, τέτοια ώστε:

$$\phi(X) = \begin{cases} 1, & \text{αν το χρωμόσωμα } X \text{ είναι εφικτό} \\ 0, & \text{διαφορετικά} \end{cases} \quad (3.39)$$

Ο έλεγχος αυτός επιτρέπει στον αλγόριθμο να γνωρίζει αν μια λύση μπορεί να γίνει αποδεκτή αυτούσια, αν πρέπει να επιδιορθωθεί ή αν πρέπει να παραμείνει στον πληθυσμό μόνο με επιβάρυνση ποινής στη συνάρτηση καταλληλότητας.

Στο σημείο αυτό φαίνεται καθαρά η σημασία της συνολικής διαχείρισης περιορισμών και όχι μόνο της επιδιόρθωσης. Δεν είναι πάντα επιθυμητό να διορθώνονται όλες οι παραβιάσεις με απόλυτο τρόπο. Σε ορισμένες περιπτώσεις, μια λύση μπορεί να απέχει ελάχιστα από την εφικτότητα αλλά να είναι ιδιαίτερα υποσχόμενη ως προς το tracking error. Αν ο μηχανισμός επιδιόρθωσης είναι υπερβολικά αυστηρός, μπορεί να αλλοιώσει σημαντικά αυτή τη λύση. Για τον λόγο αυτό, η παρούσα εργασία υιοθετεί συνδυαστική στρατηγική: οι βασικές και εύκολα διορθώσιμες παραβιάσεις, όπως το άθροισμα βαρών ή η ασυνέπεια ανάμεσα σε z_i και w_i , διορθώνονται άμεσα, ενώ οι πιο σύνθετες ή ήπιες παραβιάσεις μπορούν να λαμβάνονται υπόψη και μέσω ποινικών όρων στη συνάρτηση καταλληλότητας.

Η χρήση μηχανισμού επιδιόρθωσης σε συνδυασμό με ποινές στη fitness function θεωρείται ιδιαίτερα αποτελεσματική. Αν βασιζόταν κανείς αποκλειστικά σε ποινές, ο αλγόριθμος θα μπορούσε να αναλώνεται συχνά σε λύσεις με χαμηλή πρακτική αξία. Αν, αντίθετα, επιχειρούνταν απόλυτη επιδιόρθωση όλων των χρωμοσωμάτων χωρίς καμία ανοχή, η διαδικασία θα μπορούσε να γίνει υπερβολικά άκαμπτη και να καταστρέφει χρήσιμη γενετική πληροφορία. Για τον λόγο αυτό, στην παρούσα εργασία επιλέγεται μια ισορροπημένη προσέγγιση: πρώτα επιχειρείται επιδιόρθωση των βασικών παραβιάσεων και στη συνέχεια, όπου απαιτείται, οι υπολειπόμενες ατέλειες λαμβάνονται υπόψη στη συνάρτηση καταλληλότητας μέσω ποινικών όρων.

3.15 Κριτήριο τερματισμού

Το κριτήριο τερματισμού αποτελεί βασικό στοιχείο στη σχεδίαση του προτεινόμενου Γενετικού Αλγορίθμου, καθώς καθορίζει το σημείο στο οποίο η εξελικτική διαδικασία

θεωρείται ότι έχει ολοκληρώσει ικανοποιητικά την πορεία της. Η επιλογή του κατάλληλου κριτηρίου είναι ιδιαίτερα σημαντική, διότι επηρεάζει άμεσα τόσο την ποιότητα της τελικής λύσης όσο και τον απαιτούμενο υπολογιστικό χρόνο. Αν ο αλγόριθμος τερματιστεί πολύ νωρίς, υπάρχει κίνδυνος να μην έχει προλάβει να εξερευνήσει επαρκώς τον χώρο λύσεων και να επιστρέψει χαρτοφυλάκιο κατώτερης ποιότητας. Αντίθετα, αν συνεχιστεί υπερβολικά για μεγάλο χρονικό διάστημα, τότε μπορεί να αυξηθεί σημαντικά το υπολογιστικό κόστος χωρίς ουσιαστική βελτίωση του αποτελέσματος.

Στο πρόβλημα της αποτύπωσης δείκτη, το ζήτημα αυτό είναι ιδιαίτερα σημαντικό, επειδή η ποιότητα μιας λύσης εξαρτάται από σύνθετη αλληλεπίδραση παραμέτρων, περιορισμών και γενετικών τελεστών. Δεν είναι πάντοτε δυνατό να γνωρίζει κανείς εκ των προτέρων πόσες γενιές απαιτούνται ώστε ο πληθυσμός να φτάσει σε ικανοποιητικό επίπεδο σύγκλισης. Συνεπώς, το κριτήριο τερματισμού πρέπει να είναι σχεδιασμένο έτσι ώστε να ανταποκρίνεται τόσο στη δυναμική της εξελικτικής διαδικασίας όσο και στις πρακτικές απαιτήσεις της υπολογιστικής εφαρμογής.

Το απλούστερο και συνηθέστερο κριτήριο τερματισμού είναι ο προκαθορισμένος μέγιστος αριθμός γενεών. Αν $G_{\max}$ είναι ο μέγιστος επιτρεπτός αριθμός γενεών, τότε ο αλγόριθμος τερματίζεται όταν συμπληρωθεί η γενιά $g = G_{\max}$. Δηλαδή:

$$g \geq G_{\max} \quad (3.40)$$

Η επιλογή αυτού του κριτηρίου έχει το πλεονέκτημα της απλότητας και του πλήρους ελέγχου του υπολογιστικού κόστους. Ο ερευνητής γνωρίζει εκ των προτέρων το ανώτατο πλήθος επαναλήψεων και μπορεί να οργανώσει αναλόγως το πειραματικό πλαίσιο. Ωστόσο, το μειονέκτημά του είναι ότι δεν λαμβάνει υπόψη τη δυναμική της σύγκλισης. Ο αλγόριθμος μπορεί να έχει ουσιαστικά σταθεροποιηθεί πολύ πριν φτάσει στη γενιά $G_{\max}$, ή αντίθετα να χρειάζεται περισσότερες γενιές για να συνεχίσει να βελτιώνεται.

Για τον λόγο αυτό, ένα πιο λειτουργικό κριτήριο βασίζεται στην απουσία ουσιαστικής βελτίωσης της καλύτερης λύσης για ορισμένο αριθμό διαδοχικών γενεών. Αν $F_{\text{best}}^{(g)}$ είναι η καλύτερη τιμή καταλληλότητας στη γενιά g , τότε μπορεί να οριστεί ότι ο αλγόριθμος τερματίζεται όταν η μεταβολή της βέλτιστης τιμής παραμένει μικρότερη από ένα πολύ μικρό κατώφλι ε για H συνεχόμενες γενιές. Η συνθήκη αυτή μπορεί να γραφτεί εννοιολογικά ως:

$$|F_{\text{best}}^{(g)} - F_{\text{best}}^{(g-h)}| < \varepsilon, \forall h = 1, 2, \dots, H \quad (3.41)$$

Η λογική αυτού του κριτηρίου είναι ότι, όταν η καλύτερη λύση δεν βελτιώνεται ουσιαστικά για αρκετό διάστημα, τότε ο πληθυσμός έχει κατά πάσα πιθανότητα συγκλίνει σε σταθερή περιοχή του χώρου λύσεων και η περαιτέρω εκτέλεση του αλγορίθμου είναι απίθανο να αποφέρει σημαντικό όφελος.

Η χρήση κατωφλίου βελτίωσης είναι ιδιαίτερα χρήσιμη στο πρόβλημα της αποτύπωσης δείκτη, όπου μικρές αριθμητικές μεταβολές στη συνάρτηση καταλληλότητας μπορεί να μην έχουν ουσιαστική σημασία σε οικονομικούς όρους. Αν, για παράδειγμα, το tracking error βελτιώνεται μόνο οριακά μετά από μεγάλο αριθμό γενεών, τότε η συνέχιση της εξελικτικής διαδικασίας μπορεί να μην δικαιολογείται από άποψη υπολογιστικού κόστους. Συνεπώς, η

έννοια της “ουσιαστικής βελτίωσης” είναι πιο κατάλληλη από την απλή τυπική διαφοροποίηση τιμών.

Ένα τρίτο πιθανό κριτήριο αφορά τη σύγκλιση του ίδιου του πληθυσμού. Στην περίπτωση αυτή, ο αλγόριθμος τερματίζεται όταν τα άτομα του πληθυσμού γίνουν υπερβολικά όμοια μεταξύ τους, είτε ως προς τη δομή των χρωμοσωμάτων είτε ως προς τις τιμές καταλληλότητας. Η λογική αυτού του κριτηρίου είναι ότι ένας πληθυσμός με πολύ μικρή ποικιλία έχει περιορισμένη εξελικτική ικανότητα και, κατά συνέπεια, η περαιτέρω συνέχιση του αλγορίθμου είναι λιγότερο πιθανό να οδηγήσει σε νέες και καλύτερες λύσεις. Παρότι η προσέγγιση αυτή είναι θεωρητικά ενδιαφέρουσα, στην πράξη είναι συχνά πιο δύσκολο να οριστεί με σαφή και σταθερό τρόπο, ιδίως σε προβλήματα όπως το index tracking όπου δύο χαρτοφυλάκια μπορεί να είναι παρόμοια ως προς την ποιότητα αλλά διαφορετικά ως προς τη σύνθεσή τους.

Στην παρούσα εργασία, προτείνεται η χρήση συνδυαστικού κριτηρίου τερματισμού, το οποίο ενσωματώνει τόσο τον μέγιστο αριθμό γενεών όσο και την απουσία ουσιαστικής βελτίωσης. Πιο συγκεκριμένα, ο αλγόριθμος τερματίζεται όταν ισχύσει μία από τις δύο παρακάτω συνθήκες:

1. έχει συμπληρωθεί ο μέγιστος επιτρεπτός αριθμός γενεών $G_{\max}$,
2. ή η καλύτερη τιμή καταλληλότητας δεν έχει βελτιωθεί περισσότερο από εγία H διαδοχικές γενιές.

Με συμβολικό τρόπο, ο τερματισμός επέρχεται όταν:

$$g \geq G_{\max} \quad (3.42)$$

ή όταν:

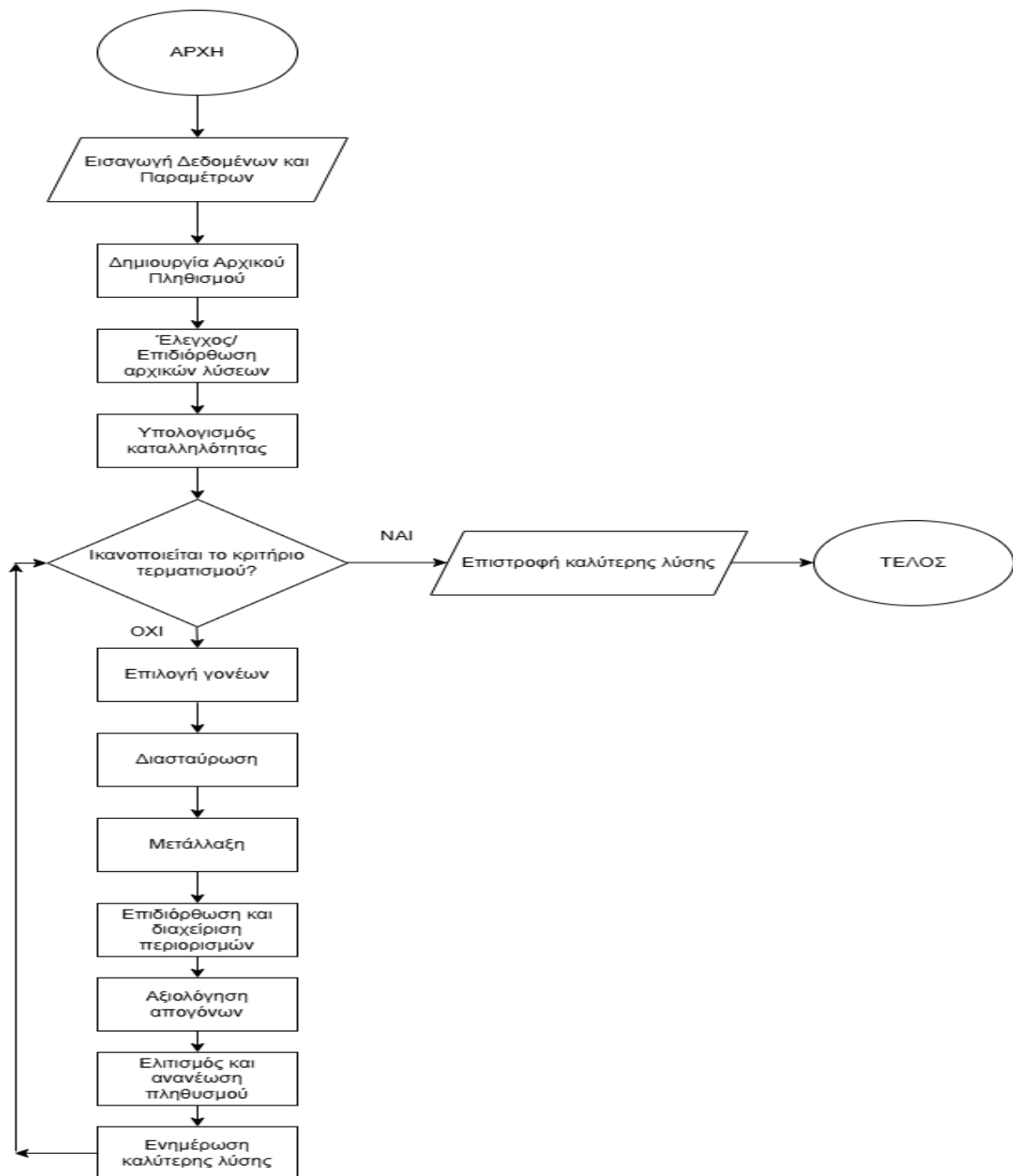
$$|F_{\text{best}}^{(g)} - F_{\text{best}}^{(g-h)}| < \varepsilon, \forall h = 1, 2, \dots, H \quad (3.43)$$

Η συνδυαστική αυτή λογική είναι προτιμότερη, επειδή εξισορροπεί δύο διαφορετικές ανάγκες. Από τη μία πλευρά, ο μέγιστος αριθμός γενεών παρέχει σαφές άνω όριο στον υπολογιστικό χρόνο. Από την άλλη, το κριτήριο στασιμότητας επιτρέπει την πρόωρη παύση της διαδικασίας όταν έχει ήδη επιτευχθεί ικανοποιητική σύγκλιση. Έτσι, αποφεύγεται τόσο ο πρόωρος τερματισμός όσο και η άσκοπη παράταση της αναζήτησης.

Η εφαρμογή του κριτηρίου στασιμότητας απαιτεί επίσης τον καθορισμό δύο κρίσιμων παραμέτρων: του κατωφλίου ε και του μήκους παρακολούθησης H . Η παράμετρος ε καθορίζει ποια βελτίωση θεωρείται ουσιαστική, ενώ η παράμετρος H καθορίζει πόσες συνεχόμενες γενιές χωρίς ουσιαστική πρόοδο απαιτούνται για να θεωρηθεί ότι ο αλγόριθμος έχει σταθεροποιηθεί. Αν το ε είναι πολύ μικρό, ο αλγόριθμος ενδέχεται να συνεχίζει να εκτελείται για πολύ μικρές βελτιώσεις χωρίς πρακτική σημασία. Αν είναι πολύ μεγάλο, υπάρχει κίνδυνος πρόωρου τερματισμού. Αντίστοιχα, πολύ μικρό H καθιστά το κριτήριο υπερβολικά ευαίσθητο, ενώ πολύ μεγάλο H μπορεί να καθυστερεί αδικαιολόγητα τη λήξη της διαδικασίας.

3.16 Διάγραμμα ροής της προτεινόμενης μεθόδου

Η ροή του προτεινόμενου αλγορίθμου ξεκινά από την εισαγωγή των δεδομένων του προβλήματος, δηλαδή των αποδόσεων των τίτλων, των αποδόσεων του δείκτη και των βασικών παραμέτρων του μοντέλου και του αλγορίθμου. Ακολουθεί η δημιουργία του αρχικού πληθυσμού, στο οποίο κάθε χρωμόσωμα αντιστοιχεί σε υποψήφιο χαρτοφυλάκιο. Στη συνέχεια, κάθε άτομο του πληθυσμού αξιολογείται μέσω της συνάρτησης καταλληλότητας, η οποία βασίζεται στο tracking error και στη διαχείριση των περιορισμών. Μετά την αρχική αυτή αξιολόγηση, η διαδικασία εισέρχεται στον κύριο εξελικτικό βρόχο, όπου επαναλαμβάνονται η επιλογή γονέων, η διασταύρωση, η μετάλλαξη, η επιδιόρθωση, η αξιολόγηση των απογόνων και η ανανέωση του πληθυσμού. Ο κύκλος αυτός συνεχίζεται μέχρι να ικανοποιηθεί το κριτήριο τερματισμού, οπότε και επιστρέφεται η καλύτερη λύση που βρέθηκε.



3.17 Υπολογιστική πολυπλοκότητα και σχολιασμός

Η ανάλυση της υπολογιστικής πολυπλοκότητας του προτεινόμενου Γενετικού Αλγορίθμου είναι ιδιαίτερα σημαντική, διότι επιτρέπει την κατανόηση του υπολογιστικού κόστους της μεθόδου και βοηθά στην ερμηνεία της πρακτικής της αποδοτικότητας όταν εφαρμόζεται σε πραγματικά χρηματιστηριακά δεδομένα. Στο πρόβλημα της αποτύπωσης δείκτη, η ανάγκη αξιολόγησης πολλών υποψήφιων χαρτοφυλακίων σε διαδοχικές γενιές, σε συνδυασμό με την ύπαρξη περιορισμών και μηχανισμών επιδιόρθωσης, καθιστά απαραίτητη μια σαφή εκτίμηση του τρόπου με τον οποίο αυξάνεται ο απαιτούμενος υπολογιστικός φόρτος καθώς αυξάνεται το μέγεθος του προβλήματος.

Για την ανάλυση αυτή, θεωρούμε ότι:

- n είναι το πλήθος των διαθέσιμων τίτλων,
- m είναι το πλήθος των χρονικών περιόδων των δεδομένων,
- N είναι το μέγεθος του πληθυσμού,
- G είναι ο αριθμός γενεών που τελικά εκτελούνται,
- και K είναι το μέγιστο πλήθος τίτλων που επιτρέπεται να συμμετέχουν σε κάθε χαρτοφυλάκιο.

Η βασική και πιο απαιτητική υπολογιστική διαδικασία σε κάθε γενιά είναι η αξιολόγηση της συνάρτησης καταλληλότητας. Για κάθε χρωμόσωμα πρέπει να υπολογιστεί η απόδοση του χαρτοφυλακίου για κάθε χρονική περίοδο και στη συνέχεια το αντίστοιχο tracking error. Αν η αναπαράσταση αξιολογείται με βάση όλα τα διαθέσιμα βάρη, τότε για κάθε χρωμόσωμα απαιτείται υπολογισμός της μορφής:

$$P_t = \sum_{i=1}^n w_i r_{it}, \forall t \in T \quad (3.44)$$

Αυτό σημαίνει ότι ο υπολογισμός της απόδοσης του χαρτοφυλακίου σε όλες τις χρονικές περιόδους απαιτεί τάξη μεγέθους $O(nm)$ ανά χρωμόσωμα. Εφόσον η ίδια διαδικασία επαναλαμβάνεται για όλα τα N άτομα του πληθυσμού, η συνολική πολυπλοκότητα της αξιολόγησης σε μία γενιά είναι:

$$O(Nnm)$$

Αν ληφθεί υπόψη ότι ο αλγόριθμος εκτελείται για G γενιές, τότε η συνολική πολυπλοκότητα της διαδικασίας αξιολόγησης σε ολόκληρη την εκτέλεση είναι:

$$O(GNnm)$$

Η παραπάνω σχέση δείχνει ότι το κόστος της μεθόδου αυξάνεται γραμμικά ως προς το πλήθος των γενεών, το μέγεθος του πληθυσμού, τον αριθμό των διαθέσιμων τίτλων και το μήκος της χρονικής σειράς. Από τις τέσσερις αυτές παραμέτρους, ιδιαίτερα σημαντική επίδραση έχουν το n και το m , καθώς αυτά εξαρτώνται άμεσα από το μέγεθος του πραγματικού χρηματοοικονομικού προβλήματος.

Στην πράξη, όμως, λόγω του περιορισμού πληθικότητας, κάθε χαρτοφυλάκιο περιλαμβάνει το πολύ K ενεργούς τίτλους, όπου συνήθως $K \ll n$. Σε αυτή την περίπτωση, αν η υλοποίηση αξιοποιεί την αραιή δομή της λύσης και υπολογίζει την απόδοση του χαρτοφυλακίου μόνο με βάση τους ενεργούς τίτλους, τότε ο υπολογισμός της απόδοσης ανά χρωμόσωμα μπορεί να μειωθεί από $O(nm)$ σε:

$$O(Km)$$

και επομένως η συνολική αξιολόγηση ανά γενιά μπορεί να γίνει:

$$O(NKm)$$

ενώ για ολόκληρη την εκτέλεση:

$$O(GNKm)$$

Η παρατήρηση αυτή είναι ιδιαίτερα σημαντική, διότι δείχνει ότι ο ίδιος ο περιορισμός πληθικότητας, πέρα από το επενδυτικό του νόημα, μπορεί να συμβάλει και στη μείωση του υπολογιστικού κόστους όταν η υλοποίηση είναι σωστά σχεδιασμένη.

Ο τελεστής επιλογής μέσω τουρνουά έχει σχετικά μικρό υπολογιστικό κόστος σε σύγκριση με την αξιολόγηση. Αν το μέγεθος του τουρνουά είναι σταθερό και μικρό, τότε η επιλογή ενός γονέα έχει κόστος $O(1)$, και η συνολική επιλογή των γονέων για μία γενιά έχει γραμμική πολυπλοκότητα ως προς το πλήθος των απογόνων, δηλαδή περίπου:

$$O(N)$$

Αντίστοιχα, οι τελεστές διασταύρωσης και μετάλλαξης έχουν κόστος που εξαρτάται από το μήκος του χρωμοσώματος ή από το πλήθος των ενεργών στοιχείων που επηρεάζονται. Σε μια γενική διατύπωση, το κόστος τους μπορεί να θεωρηθεί της τάξης $O(n)$ ανά απόγονο, ή πιο ρεαλιστικά $O(K)$ όταν η παρέμβαση γίνεται κυρίως στους ενεργούς τίτλους. Επομένως, για ολόκληρο τον πληθυσμό μιας γενιάς, το συνολικό τους κόστος παραμένει περίπου:

$$O(Nn) \text{ ή } O(NK)$$

ανάλογα με την υλοποίηση.

Ο μηχανισμός επιδιόρθωσης και διαχείρισης περιορισμών έχει επίσης υπολογιστικό κόστος που εξαρτάται από το πλήθος των τίτλων και από το πόσοι περιορισμοί παραβιάζονται. Η κανονικοποίηση των βαρών είναι γραμμική διαδικασία, ενώ ο έλεγχος cardinality και ο έλεγχος ορίων συμμετοχής απαιτούν επίσης κατά κανόνα γραμμικό χρόνο ως προς το μήκος του χρωμοσώματος ή το πλήθος των ενεργών τίτλων. Αν απαιτείται ταξινόμηση των ενεργών τίτλων, όπως στην περίπτωση απομάκρυνσης τίτλων με πολύ μικρά βάρη, τότε το αντίστοιχο κόστος μπορεί να φτάσει:

$$O(K \log K)$$

ανά χρωμόσωμα. Ωστόσο, ακόμη και σε αυτή την περίπτωση, η κυρίαρχη συνιστώσα του συνολικού κόστους παραμένει συνήθως η αξιολόγηση της συνάρτησης καταλληλότητας, ιδιαίτερα όταν το *μείναι* μεγάλο.

Από τα παραπάνω προκύπτει ότι η συνολική πολυπλοκότητα του προτεινόμενου αλγορίθμου καθορίζεται κυρίως από τη διαδικασία αξιολόγησης του πληθυσμού. Σε πρακτικούς όρους, αυτό σημαίνει ότι η απόδοση του αλγορίθμου εξαρτάται σε μεγάλο βαθμό από:

- το μέγεθος του πληθυσμού N ,
- το πλήθος των γενεών G ,
- το μέγεθος του ιστορικού παραθύρου m ,
- και το αν η υλοποίηση αξιοποιεί ή όχι την αραιότητα του χαρτοφυλακίου μέσω του K .

Η παρατήρηση αυτή είναι σημαντική και για τον πειραματικό σχεδιασμό της εργασίας, διότι δείχνει ότι η επιλογή υπερβολικά μεγάλου πληθυσμού ή πολύ μεγάλου αριθμού γενεών μπορεί να αυξήσει σημαντικά τον χρόνο εκτέλεσης χωρίς απαραίτητα ανάλογη βελτίωση των αποτελεσμάτων. Συνεπώς, η ρύθμιση των παραμέτρων του αλγορίθμου δεν είναι μόνο ζήτημα ποιότητας λύσεων αλλά και ζήτημα υπολογιστικής οικονομίας.

Από θεωρητική άποψη, η χρήση Γενετικού Αλγορίθμου δεν εξαλείφει τη δυσκολία του προβλήματος. Το πρόβλημα της αποτύπωσης δείκτη με περιορισμούς πληθικότητας, ορίων βαρών και κόστους συναλλαγών παραμένει σύνθετο συνδυαστικό πρόβλημα. Η συμβολή της προτεινόμενης μεθόδου δεν είναι ότι μετατρέπει το πρόβλημα σε απλό, αλλά ότι προσφέρει μια πρακτικά διαχειρίσιμη διαδικασία αναζήτησης υψηλής ποιότητας λύσεων χωρίς να απαιτεί εξαντλητική διερεύνηση όλων των δυνατών συνδυασμών. Με αυτή την έννοια, ο Γενετικός Αλγόριθμος λειτουργεί ως υπολογιστικά εφικτή προσέγγιση για ένα πρόβλημα του οποίου η ακριβής επίλυση θα ήταν ιδιαίτερα δύσκολη σε ρεαλιστική κλίμακα.

Ιδιαίτερο ενδιαφέρον παρουσιάζει και η σχέση ανάμεσα στην υπολογιστική πολυπλοκότητα και στην ποιότητα της λύσης. Η αύξηση του πληθυσμού ή του αριθμού γενεών τείνει γενικά να ενισχύει τη διερευνητική ικανότητα του αλγορίθμου και να αυξάνει την πιθανότητα εντοπισμού καλύτερων λύσεων. Ωστόσο, η βελτίωση αυτή δεν είναι απεριόριστη. Μετά από ένα σημείο, η προσθήκη επιπλέον υπολογιστικού χρόνου οδηγεί σε όλο και μικρότερα κέρδη ως προς το tracking error. Αυτό σημαίνει ότι στην πράξη απαιτείται μια ισορροπία ανάμεσα στην ακρίβεια της αναζήτησης και στην υπολογιστική επιβάρυνση, κάτι που θα διερευνηθεί και πειραματικά στα επόμενα κεφάλαια.

Ένας ακόμη σημαντικός σχολιασμός αφορά την κλιμάκωση της μεθόδου σε πραγματικά δεδομένα. Ο αλγόριθμος είναι ιδιαίτερα κατάλληλος για εφαρμογές όπου το πρόβλημα περιλαμβάνει αρκετούς τίτλους, αλλά όχι τόσο μεγάλους ώστε να καθίσταται απαγορευτικός ακόμη και ο επαναλαμβανόμενος υπολογισμός της fitness function. Η χρήση πραγματικών δεικτών όπως ο DAX, ο FTSE ή ο S&P, με κατάλληλο cardinality constraint, θεωρείται συμβατή με την προτεινόμενη μεθοδολογία. Ωστόσο, όσο αυξάνεται ο αριθμός των τίτλων και το μήκος των ιστορικών σειρών, γίνεται όλο και πιο σημαντική η ανάγκη αποδοτικής υλοποίησης, διανυσματοποίησης υπολογισμών και ενδεχομένως παραλληλοποίησης των αξιολογήσεων.

ΚΕΦΑΛΑΙΟ 4: ΥΛΟΠΟΙΗΣΗ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΜΕΘΟΔΟΥ

4.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο παρουσιάστηκαν αναλυτικά οι θεωρητικές αρχές των μεθόδων που εξετάζονται για το πρόβλημα του index tracking. Στο παρόν κεφάλαιο το ενδιαφέρον μετατοπίζεται από τη θεωρητική διατύπωση στην υπολογιστική υλοποίηση. Στόχος είναι να αποσαφηνιστεί πώς οι βασικές έννοιες του προβλήματος μεταφράζονται σε συναρτήσεις, κλάσεις και διαδοχικά βήματα εκτέλεσης σε Python.

Η υλοποίηση είναι οργανωμένη γύρω από τέσσερις μεθόδους: Genetic Algorithm (GA), Random Search, Corr-EW και Corr-Opt. Κοινό σημείο των μεθόδων αποτελεί η αναζήτηση χαρτοφυλακίου με ακριβώς K ενεργά περιουσιακά στοιχεία, ώστε το χαρτοφυλάκιο να προσεγγίζει τις αποδόσεις του δείκτη αναφοράς. Η ίδια υποδομή χρησιμοποιείται για την επεξεργασία των δεδομένων, την αξιολόγηση των χαρτοφυλακίων, την εφαρμογή των περιορισμών και την αποτίμηση της out-of-sample επίδοσης.

Η παρουσίαση που ακολουθεί αποφεύγει την επανάληψη της θεωρίας και επικεντρώνεται στη λειτουργική αντιστοίχιση «μέθοδος $\rightarrow$ συνάρτηση/κλάση Python $\rightarrow$ υπολογιστικό βήμα $\rightarrow$ χρήση στο πείραμα». Τα αποσπάσματα κώδικα που παρατίθενται προέρχονται από την τελική υλοποίηση του notebook.

4.2 Αρχιτεκτονική της υπολογιστικής υλοποίησης

Η εφαρμογή ακολουθεί μια αρθρωτή αρχιτεκτονική. Στο επίπεδο των δεδομένων πραγματοποιείται φόρτωση και μετατροπή των τιμών σε αποδόσεις. Στη συνέχεια οι αποδόσεις των επιμέρους τίτλων και του δείκτη ευθυγραμμίζονται χρονικά. Πάνω σε αυτή τη δομή εφαρμόζονται οι περιορισμοί και η συνάρτηση αξιολόγησης. Τέλος, κάθε μέθοδος παράγει ένα διάνυσμα βαρών, το οποίο αξιολογείται τόσο στο δείγμα εκπαίδευσης όσο και στο δείγμα ελέγχου.

Επίπεδο	Βασική υλοποίηση	Ρόλος
Δεδομένα	<code>`load_returns_csv()`</code> , <code>`load_index_returns_csv()`</code>	Φόρτωση και προετοιμασία χρονοσειρών
Αποδόσεις	<code>`compute_simple_returns()`</code> , <code>`compute_simple_returns_series()`</code>	Μετατροπή τιμών σε απλές αποδόσεις
Ευθυγράμμιση	<code>`align_asset_and_index_returns()`</code>	Κοινό χρονικό δείγμα
Περιορισμοί	<code>`ConstraintConfig`</code> , <code>`ConstraintHandler`</code>	Έλεγχος και επιδιόρθωση εφικτότητας
Αξιολόγηση	<code>`FitnessEvaluator`</code> , <code>`tracking_mse()`</code>	Υπολογισμός tracking MSE και fitness
Μέθοδοι	<code>`corr_ew_method()`</code> , <code>`corr_opt_method()`</code> , <code>`random_search_method()`</code> , <code>`ga_method()`</code>	Παραγωγή βαρών χαρτοφυλακίου

Πείραμα	<code>`run_gui_all_methods_benchmark()`</code>	Εκτέλεση όλων των σεναρίων και αποθήκευση αποτελεσμάτων
---------	--	---

4.3 Εισαγωγή και προεπεξεργασία δεδομένων

- Φόρτωση και ενοποίηση των δεδομένων

Η υλοποίηση υποστηρίζει τόσο αρχεία CSV όσο και τα ενσωματωμένα δεδομένα των benchmarks του OR-Library. Η επιλογή αυτή επιτρέπει την αναπαραγωγή των πειραμάτων χωρίς να απαιτείται εξωτερική λήψη των δεδομένων. Τα τρία σύνολα δεδομένων που χρησιμοποιούνται στην εφαρμογή αντιστοιχούν στους δείκτες Hang Seng, FTSE 100 και Nikkei 225.

Μετά τη φόρτωση, η συνάρτηση ``align_asset_and_index_returns()`` ελέγχει ότι υπάρχουν μη κενές χρονοσειρές, ότι τα ονόματα των στηλών είναι μοναδικά και ότι υπάρχει κοινό σύνολο χρονικών παρατηρήσεων μεταξύ των assets και του benchmark. Έτσι, οι δύο πλευρές της συνάρτησης `tracking error` χρησιμοποιούν ακριβώς τις ίδιες χρονικές παρατηρήσεις.

- Υπολογισμός απλών αποδόσεων

Όταν τα δεδομένα εισάγονται ως τιμές, η μετατροπή σε απλές αποδόσεις πραγματοποιείται με τη ``compute_simple_returns()``. Η συνάρτηση χρησιμοποιεί την ποσοστιαία μεταβολή μεταξύ διαδοχικών παρατηρήσεων και απορρίπτει τις μη διαθέσιμες τιμές που προκύπτουν από την πρώτη παρατήρηση.

Κώδικας 4.3.1: Υπολογισμός απλών αποδόσεων στην υλοποίηση Python

```
def compute_simple_returns(price_df: pd.DataFrame) -> pd.DataFrame:
    if price_df.empty:
        raise ValueError("Empty price data.")
    return price_df.pct_change(fill_method=None).dropna(how="all")
```

Η αντίστοιχη λογική εφαρμόζεται για τον δείκτη μέσω της ``compute_simple_returns_series()``. Η διάκριση μεταξύ DataFrame και Series επιτρέπει την ίδια υποδομή να χρησιμοποιείται τόσο για πολλαπλές μετοχές όσο και για τη μοναδική χρονοσειρά του benchmark.

- Διαχωρισμός δείγματος εκπαίδευσης και ελέγχου

Η αξιολόγηση των μεθόδων πραγματοποιείται με διαχωρισμό της χρονοσειράς σε δύο ίσα τμήματα. Η ``split_train_test()`` υπολογίζει το σημείο διαχωρισμού ως ``n // 2``, χρησιμοποιώντας το πρώτο μισό για εκπαίδευση/βελτιστοποίηση και το δεύτερο μισό για out-of-sample αξιολόγηση.

Κώδικας 4.3.2: Διαχωρισμός των δεδομένων σε train και test δείγμα.

```
def split_train_test(asset_returns: pd.DataFrame, index_returns: pd.Series):
    n = len(index_returns)
    if n < 4:
        raise ValueError("Need at least four observations for a train/test
split.")
    split = n // 2
    return (
        asset_returns.iloc[:split].copy(),
        index_returns.iloc[:split].copy(),
        asset_returns.iloc[split:].copy(),
        index_returns.iloc[split:].copy(),
    )
```

4.4 Διαχείριση περιορισμών

Η τήρηση των περιορισμών αποτελεί κοινό υπολογιστικό στοιχείο όλων των μεθόδων. Η κλάση `ConstraintConfig` συγκεντρώνει τις βασικές παραμέτρους, όπως το μέγιστο πλήθος ενεργών assets, το ελάχιστο και μέγιστο βάρος και, στην παρούσα εφαρμογή, τον περιορισμό ακριβώς K ενεργών assets.

Κώδικας 4.4.1: Ορισμός της παραμέτρησης των περιορισμών.

```
@dataclass
class ConstraintConfig:
    max_assets: int
    min_weight: float
    max_weight: float
    exact_assets: Optional[int] = None
    use_transaction_cost: bool = False
```

Η `ConstraintHandler` αναλαμβάνει τον πρακτικό έλεγχο της εφικτότητας. Περιλαμβάνει λειτουργίες για την καταμέτρηση των ενεργών θέσεων (`count\_active()`), τον έλεγχο του budget, της cardinality και των ορίων, καθώς και για την παραγωγή ή επιδιόρθωση εφικτών διανυσμάτων βαρών. Με τον τρόπο αυτό οι περιορισμοί δεν επαναλαμβάνονται ξεχωριστά σε κάθε αλγόριθμο.

- Προβολή στο εφικτό σύνολο

Για την εφαρμογή των ορίων και του περιορισμού αθροίσματος των βαρών χρησιμοποιείται η `project\_bounded\_simplex()`. Η συνάρτηση δέχεται ένα αρχικό διάνυσμα, κάτω και άνω όριο και επιθυμητό άθροισμα, και επιστρέφει ένα διάνυσμα που βρίσκεται στο αντίστοιχο περιορισμένο simplex. Η λειτουργία αυτή είναι ιδιαίτερα σημαντική στη Corr-Opt, αλλά χρησιμοποιείται και στο γενικό μηχανισμό επιδιόρθωσης.

Κώδικας 4.4.2: Βασική συνάρτηση προβολής στο bounded simplex.

```
def project_bounded_simplex(
    values: np.ndarray,
    lower: float | np.ndarray,
    upper: float | np.ndarray,
    target_sum: float = 1.0,
    max_iter: int = 120,
) -> np.ndarray:
    """Euclidean projection onto lower <= x <= upper and sum(x)=target_sum."""
    v = np.asarray(values, dtype=float)
```

```

lo = np.broadcast_to(np.asarray(lower, dtype=float), v.shape).copy()
hi = np.broadcast_to(np.asarray(upper, dtype=float), v.shape).copy()
if np.any(lo > hi):
    raise ValueError("Projection lower bounds exceed upper bounds.")
if lo.sum() > target_sum + 1e-12 or hi.sum() < target_sum - 1e-12:
    raise ValueError("Bounded simplex is infeasible for the requested target
sum.")

left = float(np.min(v - hi)) - 1.0
right = float(np.max(v - lo)) + 1.0
for _ in range(max_iter):
    mid = 0.5 * (left + right)
    x = np.clip(v - mid, lo, hi)
    if x.sum() > target_sum:
        left = mid
    else:
        right = mid
x = np.clip(v - 0.5 * (left + right), lo, hi)

# Remove tiny floating residual without violating the box.
residual = target_sum - float(x.sum())
if abs(residual) > 1e-11:
    if residual > 0:
        room = hi - x
    else:
        room = x - lo
    candidates = np.where(room > 1e-14)[0]
    for idx in candidates:
        delta = math.copysign(min(abs(residual), room[idx]), residual)
        x[idx] += delta
        residual -= delta
    if abs(residual) <= 1e-12:
        break
return x

```

4.5 Συνάρτηση αξιολόγησης και *tracking MSE*

Κεντρικό μέγεθος της υλοποίησης είναι το mean squared tracking error, το οποίο στο notebook ονομάζεται ρητά `'tracking_mse'`. Για ένα διάνυσμα βαρών w , η απόδοση του χαρτοφυλακίου υπολογίζεται ως γραμμικός συνδυασμός των αποδόσεων των assets. Στη συνέχεια αφαιρείται η απόδοση του benchmark και υπολογίζεται ο μέσος όρος του τετραγώνου της ενεργού απόδοσης.

Κώδικας 4.5.1: Υπολογισμός του tracking MSE.

```

def tracking_mse(weights: np.ndarray, asset_returns: pd.DataFrame, index_returns:
pd.Series) -> float:
    p = asset_returns.to_numpy(dtype=float) @ np.asarray(weights, dtype=float)
    b = index_returns.to_numpy(dtype=float)
    d = p - b
    return float(np.mean(d * d))

```

Σε επίπεδο αντικειμενοστρεφούς σχεδίασης, η ίδια λογική υλοποιείται στην `'FitnessEvaluator'`, όπου η `'portfolio_returns()'` υπολογίζει την απόδοση του χαρτοφυλακίου και η `'tracking_mse()'` την απόσταση από το benchmark. Η `'evaluate()'` συνδυάζει την απόδοση αυτή με τους σχετικούς όρους ποινής και παράγει το αντικείμενο αποτελέσματος που χρησιμοποιείται από τον γενετικό αλγόριθμο.

4.6 Υλοποίηση της Corr-EW

Η Corr-EW αποτελεί την απλούστερη από τις δύο correlation-based μεθόδους. Πρώτα καλείται η `select_by_correlation()`, η οποία υπολογίζει τη συσχέτιση κάθε asset με τον δείκτη, λαμβάνει την απόλυτη τιμή της συσχέτισης και επιλέγει τα K assets με τις υψηλότερες τιμές. Στη συνέχεια η `corr_ew_method()` τοποθετεί ίσο βάρος $1/K$ σε κάθε επιλεγμένο asset και μηδενικό βάρος στα υπόλοιπα.

Κώδικας 4.6.1: Υλοποίηση της Corr-EW.

```
def corr_ew_method(asset_train: pd.DataFrame, index_train: pd.Series, k: int) ->
    np.ndarray:
    selected = select_by_correlation(asset_train, index_train, k)
    w = np.zeros(asset_train.shape[1], dtype=float)
    pos = {c: i for i, c in enumerate(asset_train.columns)}
    for col in selected:
        w[pos[col]] = 1.0 / k
    return w
```

4.7 Υλοποίηση της Corr-Opt

Η Corr-Opt χρησιμοποιεί το ίδιο στάδιο επιλογής των K assets βάσει συσχέτισης, αλλά διαφοροποιείται στο στάδιο κατανομής των βαρών. Η `corr_opt_method()` δημιουργεί τον υποπίνακα αποδόσεων των επιλεγμένων assets και καλεί την `_projected_least_squares_simplex()`. Η δεύτερη υπολογίζει αρχικά μια λύση ελαχίστων τετραγώνων και στη συνέχεια εφαρμόζει επαναληπτικά projected gradient steps, με προβολή στο bounded simplex.

Κώδικας 4.7.1: Υλοποίηση της Corr-Opt και διατήρηση ακριβώς K ενεργών θέσεων.

```
def corr_opt_method(asset_train: pd.DataFrame, index_train: pd.Series, k: int) ->
    np.ndarray:
    """Correlation selection + long-only constrained LS with exactly K
    positives."""
    selected = select_by_correlation(asset_train, index_train, k)
    X = asset_train[selected].to_numpy(dtype=float)
    b = index_train.to_numpy(dtype=float)
    local = _projected_least_squares_simplex(X, b, EXACT_ACTIVE_FLOOR, 1.0)
    w = np.zeros(asset_train.shape[1], dtype=float)
    pos = {c: i for i, c in enumerate(asset_train.columns)}
    for col, weight in zip(selected, local):
        w[pos[col]] = weight
    if int(np.sum(w > ACTIVE_TOL)) != k:
        raise RuntimeError("Corr-Opt internal error: exact-K support was not
        preserved.")
    return w
```

Η τελευταία συνθήκη της συνάρτησης ελέγχει ρητά ότι το τελικό διάνυσμα διατηρεί ακριβώς K θετικά βάρη. Σε περίπτωση παραβίασης, η υλοποίηση διακόπτει την εκτέλεση με `RuntimeError`, αντί να επιστρέψει σιωπηρά μη έγκυρη λύση.

4.8 Υλοποίηση της Random Search

Η Random Search λειτουργεί ως άμεση στοχαστική μέθοδος αναζήτησης. Σε κάθε δοκιμή επιλέγονται τυχαία K assets χωρίς επανάθεση και δημιουργούνται τυχαία βάρη μέσω κατανομής Dirichlet. Για κάθε υποψήφιο χαρτοφυλάκιο υπολογίζεται το tracking MSE και διατηρείται η καλύτερη λύση που έχει παρατηρηθεί μέχρι εκείνη τη στιγμή.

Κώδικας 4.8.1: Υλοποίηση της Random Search.

```
def random_search_method(
    asset_train: pd.DataFrame,
    index_train: pd.Series,
    k: int,
    seed: int,
    n_trials: int = 2500,
) -> np.ndarray:
    if not 1 <= k <= asset_train.shape[1]:
        raise ValueError("Invalid k for random search.")
    if n_trials < 1:
        raise ValueError("n_trials must be positive.")
    rng = np.random.default_rng(seed)
    n_assets = asset_train.shape[1]
    R = asset_train.to_numpy(dtype=float)
    b = index_train.to_numpy(dtype=float)
    best_w = None
    best_mse = np.inf
    for _ in range(n_trials):
        selected = rng.choice(n_assets, size=k, replace=False)
        local = rng.dirichlet(np.ones(k))
        w = np.zeros(n_assets, dtype=float)
        w[selected] = local
        d = R @ w - b
        mse = float(np.mean(d * d))
        if mse < best_mse:
            best_mse = mse
            best_w = w
    assert best_w is not None
    return best_w
```

4.9 Υλοποίηση του Genetic Algorithm

Ο Genetic Algorithm είναι η πιο σύνθετη υπολογιστική συνιστώσα της εφαρμογής και υλοποιείται μέσω της κλάσης `GeneticIndexTracker`. Η κλάση συγκεντρώνει τα δεδομένα, τις ρυθμίσεις του GA, τον μηχανισμό περιορισμών και τον evaluator. Κάθε υποψήφια λύση αναπαρίσταται από ένα αντικείμενο `Individual`, το οποίο αποθηκεύει τα βάρη, το tracking MSE, την ποινή, το fitness και την εφικτότητα.

Κώδικας 4.9.1: Αρχικοποίηση της κλάσης `GeneticIndexTracker`.

```
class GeneticIndexTracker:
    def __init__(
        self,
        asset_returns: pd.DataFrame,
        benchmark_returns: pd.Series,
        ga_config: GAConfig,
        constraint_config: ConstraintConfig,
        fitness_config: FitnessConfig,
    ):
        asset_returns, benchmark_returns =
align_asset_and_index_returns(asset_returns, benchmark_returns)
        validate_configs(ga_config, constraint_config, fitness_config,
```

```

asset_returns.shape[1])
    self.asset_returns = asset_returns
    self.benchmark_returns = benchmark_returns
    self.asset_names = list(asset_returns.columns)
    self.n_assets = asset_returns.shape[1]
    self.ga_config = ga_config
    self.rng = np.random.default_rng(ga_config.random_seed)
    self.constraint_handler = ConstraintHandler(constraint_config)
    self.fitness_evaluator = FitnessEvaluator(
        asset_returns, benchmark_returns, self.constraint_handler,
        fitness_config
    )
    self.history: Dict[str, List[float]] = {
        "best_fitness": [],
        "best_tracking_mse": [],
        "avg_fitness": [],
        "avg_tracking_mse": [],
    }
    self.best_individual: Optional[Individual] = None

```

Η αρχικοποίηση του πληθυσμού πραγματοποιείται με `initialize_population()`. Κάθε άτομο παράγεται από την `random_feasible_weights()` του `ConstraintHandler` και αξιολογείται αμέσως από την `FitnessEvaluator`.

Κώδικας 4.9.2: Δημιουργία του αρχικού πληθυσμού.

```

def initialize_population(self) -> List[Individual]:
    return [

        self.evaluate_individual(self.constraint_handler.random_feasible_weights(self.n_a
            ssets, self.rng))
        for _ in range(self.ga_config.population_size)
    ]

```

Η επιλογή γονέων υλοποιείται με tournament selection. Επιλέγεται τυχαίο υποσύνολο ατόμων μεγέθους ίσου με το `tournament_size` και από αυτά διατηρείται εκείνο με το υψηλότερο fitness.

Κώδικας 4.9.3: Tournament selection.

```

def tournament_selection(self, population: List[Individual]) -> Individual:
    idx = self.rng.choice(len(population),
        size=self.ga_config.tournament_size, replace=False)
    return max((population[i] for i in idx), key=lambda ind: ind.fitness)

```

Η διασταύρωση υλοποιείται με αριθμητικό convex combination των δύο διανυσμάτων βαρών. Αν δεν ενεργοποιηθεί το crossover, οι γονείς αντιγράφονται ως έχουν.

Κώδικας 4.9.4: Crossover.

```

def crossover(self, p1: Individual, p2: Individual) -> Tuple[np.ndarray,
    np.ndarray]:
    if self.rng.random() >= self.ga_config.crossover_rate:
        return p1.weights.copy(), p2.weights.copy()
    alpha = float(self.rng.random())
    return (
        alpha * p1.weights + (1.0 - alpha) * p2.weights,
        alpha * p2.weights + (1.0 - alpha) * p1.weights,
    )

```

Η μετάλλαξη μπορεί είτε να αντικαταστήσει ένα ενεργό asset με ένα ανενεργό είτε να μεταβάλει τα βάρη ενεργών assets. Μετά τη μετάλλαξη, το διάνυσμα οδηγείται στον μηχανισμό `repair\_weights()`, ώστε να επανέλθει στο εφικτό σύνολο.

Κώδικας 4.9.5: Mutation και διαχείριση ενεργών/ανενεργών θέσεων.

```
def mutate(self, weights: np.ndarray) -> np.ndarray:
    w = np.asarray(weights, dtype=float).copy()
    if self.rng.random() >= self.ga_config.mutation_rate:
        return w
    c = self.constraint_handler.config
    active = np.where(w > c.active_tolerance)[0]
    inactive = np.where(w <= c.active_tolerance)[0]

    if len(active) and len(inactive) and self.rng.random() < 0.5:
        remove_idx = int(self.rng.choice(active))
        add_idx = int(self.rng.choice(inactive))
        w[add_idx] = max(w[remove_idx], _effective_lower(c, c.exact_assets is
not None))
        w[remove_idx] = 0.0
    elif len(active):
        idx = int(self.rng.choice(active))
        w[idx] = max(0.0, w[idx] + float(self.rng.normal(0.0, 0.05)))
        if len(active) > 1 and self.rng.random() < 0.5:
            others = active[active != idx]
            idx2 = int(self.rng.choice(others))
            w[idx2] = max(0.0, w[idx2] + float(self.rng.normal(0.0, 0.03)))
    return w
```

Η πλήρης εξελικτική διαδικασία βρίσκεται στην `evolve()`. Σε κάθε γενιά αποθηκεύονται το καλύτερο και το μέσο fitness καθώς και το αντίστοιχο tracking MSE. Στη συνέχεια εφαρμόζονται elitism, tournament selection, crossover και mutation. Η διαδικασία τερματίζεται είτε όταν ολοκληρωθούν οι γενιές είτε όταν επιτευχθεί το καθορισμένο όριο stagnation.

Κώδικας 4.9.6: Κύριος βρόχος εξέλιξης του Genetic Algorithm.

```
def evolve(
    self,
    verbose: bool = False,
    progress_callback: Optional[Callable[[int, int, float], None]] = None,
) -> Dict[str, object]:
    population = self.initialize_population()
    population.sort(key=lambda x: x.fitness, reverse=True)
    self.best_individual = self._copy_individual(population[0])
    stagnation = 0

    for generation in range(self.ga_config.max_generations):
        best = max(population, key=lambda x: x.fitness)
        self.history["best_fitness"].append(best.fitness)
        self.history["best_tracking_mse"].append(best.tracking_mse)
        self.history["avg_fitness"].append(float(np.mean([x.fitness for x in
population])))
        self.history["avg_tracking_mse"].append(float(np.mean([x.tracking_mse
for x in population])))

        if best.fitness > self.best_individual.fitness +
self.ga_config.improvement_tolerance:
            self.best_individual = self._copy_individual(best)
            stagnation = 0
        else:
            stagnation += 1

        if verbose and (generation == 0 or (generation + 1) % 25 == 0):
            print(f" generation={generation + 1:4d}
best_tracking_mse={best.tracking_mse:.12g}")
```

```

        if progress_callback is not None:
            progress_callback(
                generation + 1,
                self.ga_config.max_generations,
                float(best.tracking_mse),
            )
        if stagnation >= self.ga_config.stagnation_generations:
            break

        population.sort(key=lambda x: x.fitness, reverse=True)
        new_population = [self._copy_individual(x) for x in population[:
self.ga_config.elitism_count]]
        while len(new_population) < self.ga_config.population_size:
            p1 = self.tournament_selection(population)
            p2 = self.tournament_selection(population)
            c1, c2 = self.crossover(p1, p2)
            c1 = self.constraint_handler.repair_weights(self.mutate(c1),
self.rng)
            c2 = self.constraint_handler.repair_weights(self.mutate(c2),
self.rng)
            new_population.append(self.evaluate_individual(c1))
            if len(new_population) < self.ga_config.population_size:
                new_population.append(self.evaluate_individual(c2))
        population = new_population

```

4.10 Παραμετροποίηση του Genetic Algorithm

Η τελική πειραματική συνάρτηση `ga\_method()` δημιουργεί τις παραμέτρους του GA και ταυτόχρονα ενεργοποιεί τον περιορισμό ακριβώς K ενεργών assets. Για την πλήρη εκτέλεση χρησιμοποιούνται πληθυσμός 80 ατόμων, έως 150 γενιές, crossover rate 0.9, mutation rate 0.25, tournament size 3 και elitism count 4. Το όριο stagnation είναι 30 γενιές, ενώ η Random Search εκτελεί 2.500 δοκιμές ανά σενάριο.

Κώδικας 4.10.1: Πειραματική παραμετροποίηση του GA και κοινή επιβολή του ακριβούς K.

```

def ga_method(
    asset_train: pd.DataFrame,
    index_train: pd.Series,
    k: int,
    seed: int,
    quick: bool = False,
) -> Tuple[np.ndarray, pd.DataFrame]:
    ga = GAConfig(
        population_size=36 if quick else 80,
        max_generations=45 if quick else 150,
        crossover_rate=0.9,
        mutation_rate=0.25,
        tournament_size=3,
        elitism_count=4,
        stagnation_generations=15 if quick else 30,
        improvement_tolerance=1e-10,
        random_seed=seed,
    )
    constraints = ConstraintConfig(
        max_assets=k,
        exact_assets=k, # FIX: exactly K, not merely <= K
        min_weight=0.0,
        max_weight=1.0,
        use_transaction_cost=False,
    )
    fitness = FitnessConfig(
        lambda_budget=1e4,
        lambda_cardinality=1e4,
        lambda_bounds=1e4,
        lambda_transaction=0.0,
    )

```

```
result = GeneticIndexTracker(asset_train, index_train, ga, constraints,
fitness).evolve()
return result["best_weights"].to_numpy(dtype=float), result["history"]
```

4.11 Ενιαία διαδικασία εκτέλεσης των πειραμάτων

Η σύγκριση των τεσσάρων μεθόδων δεν πραγματοποιείται με ανεξάρτητες διαδικασίες, αλλά μέσω της `run_gui_all_methods_benchmark()`. Η συνάρτηση δέχεται το σύνολο των τιμών K και τις παραμέτρους των αλγορίθμων, εκτελεί κάθε μέθοδο σε κάθε dataset και αποθηκεύει τα αποτελέσματα σε δομημένα αρχεία. Με αυτό τον τρόπο διασφαλίζεται ότι τα διαφορετικά μοντέλα αξιολογούνται στο ίδιο πλαίσιο.

Κώδικας 4.11.1: Κεντρική συνάρτηση benchmark.

```
def run_gui_all_methods_benchmark(
    k_values: Sequence[int],
    population_size: int,
    max_generations: int,
    crossover_rate: float,
    mutation_rate: float,
    tournament_size: int,
    elitism_count: int,
    stagnation_generations: int,
    random_trials: int,
    seed: int,
    status_callback: Optional[Callable[[str], None]] = None,
    ga_progress_callback: Optional[Callable[[str, int, int, float], None]] =
None,
    scenario_done_callback: Optional[Callable[[int, int], None]] = None,
    output_dir: str | Path = "gui_comparison_results",
) -> Tuple[pd.DataFrame, pd.DataFrame, pd.DataFrame, str, List[Tuple[str,
plt.Figure]]]:
```

4.12 Αξιολόγηση in-sample και out-of-sample

Για κάθε συνδυασμό dataset και K παράγονται βάρη από το train δείγμα. Τα βάρη αυτά αξιολογούνται αρχικά στο ίδιο δείγμα και στη συνέχεια εφαρμόζονται στο test δείγμα. Η δεύτερη μέτρηση αποτελεί το out-of-sample tracking MSE και χρησιμοποιείται για την τελική σύγκριση των μεθόδων. Η επιλογή αυτή είναι ιδιαίτερα σημαντική, επειδή αποφεύγει την ταύτιση της ποιότητας της λύσης με την επίδοση μόνο στο δείγμα πάνω στο οποίο έγινε η βελτιστοποίηση.

4.13 Αποθήκευση και οργάνωση αποτελεσμάτων

Η εφαρμογή αποθηκεύει τα αποτελέσματα σε ξεχωριστά αρχεία CSV, ώστε να είναι δυνατή η μεταγενέστερη στατιστική επεξεργασία και η δημιουργία πινάκων και γραφημάτων. Μεταξύ των βασικών εξαγόμενων αρχείων περιλαμβάνονται τα `'all_method_results.csv'`, `'overall_method_summary.csv'`, `'scenario_winners.csv'` και `'oos_tracking_mse_comparison.csv'`. Παράλληλα αποθηκεύονται οι πίνακες βαρών των χαρτοφυλακίων και τα ιστορικά εξέλιξης του GA.

Αρχείο	Περιεχόμενο
all_method_results.csv	Πλήρης σύγκριση ανά dataset, K και μέθοδο

overall_method_summary.csv	Συνοπτικά μέτρα σύγκρισης των μεθόδων
scenario_winners.csv	Μέθοδος με το μικρότερο OOS tracking MSE ανά σενάριο
oos_tracking_mse_comparison.csv	Πίνακας OOS MSE για εύκολη σύγκριση

4.14 Γραφική παρακολούθηση της εξέλιξης

Η υλοποίηση παράγει επίσης γραφικές απεικονίσεις της εξέλιξης του GA και της σύγκρισης των μεθόδων. Για τον GA καταγράφονται σε κάθε γενιά το best tracking MSE και το average tracking MSE, επιτρέποντας την παρακολούθηση της σύγκλισης. Για τη δια-μεθοδολογική σύγκριση χρησιμοποιούνται γραφήματα του OOS tracking MSE και κανονικοποιημένης επίδοσης.

Η γραφική απεικόνιση λειτουργεί συμπληρωματικά προς τους αριθμητικούς πίνακες: δεν αποτελεί ξεχωριστό στάδιο βελτιστοποίησης, αλλά εργαλείο ελέγχου της συμπεριφοράς και της σύγκλισης των αλγορίθμων.

4.15 Αντιστοίχιση θεωρίας και υλοποίησης

Θεωρητικό στοιχείο	Υλοποίηση Python	Υπολογιστικός ρόλος
Απόδοση χαρτοφυλακίου	`FitnessEvaluator.portfolio_returns()`	Γραμμικός συνδυασμός asset returns με τα βάρη
Tracking MSE	`FitnessEvaluator.tracking_mse()` / `tracking_mse()`	Κύριο μέτρο απόστασης από τον δείκτη
Περιορισμοί	`ConstraintConfig`, `ConstraintHandler`	Έλεγχος εφικτότητας και επιδιόρθωση
Επιλογή με συσχέτιση	`select_by_correlation()`	Επιλογή των K assets
Corr-EW	`corr_ew_method()`	Ίσα βάρη στα επιλεγμένα assets
Corr-Opt	`corr_opt_method()`	Βελτιστοποίηση βαρών στο επιλεγμένο support
Random Search	`random_search_method()`	Τυχαία εξερεύνηση υποψηφίων χαρτοφυλακίων
GA πληθυσμός	`initialize_population()`	Αρχικοποίηση εφικτών ατόμων
Selection	`tournament_selection()`	Επιλογή γονέων
Crossover	`crossover()`	Συνδυασμός γονικών βαρών
Mutation	`mutate()`	Εξερεύνηση νέων λύσεων
Εξέλιξη	`evolve()`	Επαναληπτική βελτιστοποίηση
Benchmark	`run_gui_all_methods_benchmark()`	Ενιαία εκτέλεση όλων των σεναρίων

ΚΕΦΑΛΑΙΟ 5: ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται η πειραματική αξιολόγηση της εφαρμογής παρακολούθησης χρηματιστηριακών δεικτών με περιορισμό ακριβούς πληθικότητας χαρτοφυλακίου. Η αξιολόγηση πραγματοποιείται συγκριτικά για τέσσερις μεθόδους επιλογής και στάθμισης τίτλων: τον Γενετικό Αλγόριθμο, την Τυχαία Αναζήτηση (Random Search), τη μέθοδο Corr-EW και τη μέθοδο Corr-Opt. Η σύγκριση εκτελείται αυτόματα και στους τρεις διαθέσιμους δείκτες, Hang Seng, FTSE 100 και Nikkei 225, με κοινές συνθήκες εκπαίδευσης και ελέγχου, ώστε η διαφορά στην επίδοση να αποδίδεται κατά το δυνατόν στη μέθοδο βελτιστοποίησης και όχι σε διαφορετική πειραματική μεταχείριση.

Για την εμπειρική αξιολόγηση των μεθόδων χρησιμοποιήθηκαν τρία από τα κλασικά διεθνή datasets του προβλήματος του index tracking, τα οποία διατίθενται στο πλαίσιο της OR-Library του Beasley. Συγκεκριμένα, επιλέχθηκαν τα Hang Seng (indtrack1), FTSE 100 (indtrack3) και Nikkei 225 (indtrack5). Τα datasets περιλαμβάνουν ιστορικές εβδομαδιαίες τιμές του αντίστοιχου χρηματιστηριακού δείκτη και των επιμέρους μετοχών που χρησιμοποιούνται ως επενδυτικό σύνολο. Για τα συγκεκριμένα benchmark instances, το indtrack1 περιλαμβάνει 31 μετοχές, το indtrack3 89 μετοχές και το indtrack5 225 μετοχές.

- Hang Seng Index (HSI)

Ο Hang Seng Index (HSI) αποτελεί τον βασικό χρηματιστηριακό δείκτη της αγοράς του Χονγκ Κονγκ και χρησιμοποιείται για την αποτύπωση της πορείας σημαντικών εισηγμένων εταιρειών του Χρηματιστηρίου του Χονγκ Κονγκ. Η σύνθεσή του περιλαμβάνει επιχειρήσεις από διαφορετικούς κλάδους, όπως ο χρηματοπιστωτικός τομέας, το εμπόριο και οι καταναλωτικές υπηρεσίες, η τεχνολογία και οι τηλεπικοινωνίες. Ενδεικτικά παραδείγματα εταιρειών που έχουν συμμετάσχει στη σύνθεση του δείκτη είναι η HSBC Holdings, η AIA Group, η Tencent Holdings και η China Mobile, οι οποίες καταδεικνύουν τη σημαντική παρουσία χρηματοπιστωτικών, τεχνολογικών και τηλεπικοινωνιακών επιχειρήσεων στην αγορά του Χονγκ Κονγκ.

Στο benchmark dataset indtrack1 χρησιμοποιούνται 31 επιμέρους μετοχές μαζί με τη χρονοσειρά του Hang Seng Index. Συνεπώς, το συγκεκριμένο dataset αντιπροσωπεύει ένα σχετικά περιορισμένο επενδυτικό σύνολο και επιτρέπει την εξέταση του προβλήματος index tracking σε μία σημαντική ασιατική χρηματιστηριακή αγορά. Στο πλαίσιο της παρούσας εργασίας, στόχος δεν είναι η αγορά του συνόλου των τίτλων του δείκτη, αλλά η επιλογή ενός μικρότερου χαρτοφυλακίου από το διαθέσιμο σύνολο, το οποίο θα αναπαράγει όσο το δυνατόν καλύτερα τη συμπεριφορά του δείκτη.

- FTSE 100

Ο FTSE 100 αποτελεί έναν από τους σημαντικότερους χρηματιστηριακούς δείκτες του Ηνωμένου Βασιλείου και περιλαμβάνει 100 από τις μεγαλύτερες επιλέξιμες εταιρείες που είναι εισηγμένες στο London Stock Exchange, με βάση τη χρηματιστηριακή τους κεφαλαιοποίηση. Η σύνθεση του δείκτη καλύπτει ένα ευρύ φάσμα οικονομικών δραστηριοτήτων. Ενδεικτικά, μεταξύ των γνωστών εταιρειών που συνδέονται με τον FTSE

100 περιλαμβάνονται η HSBC Holdings στον χρηματοπιστωτικό κλάδο, η Shell στον ενεργειακό τομέα, η AstraZeneca στον φαρμακευτικό κλάδο και η Unilever στον κλάδο καταναλωτικών αγαθών.

Η παρουσία εταιρειών από διαφορετικούς κλάδους έχει ιδιαίτερη σημασία για το πρόβλημα του index tracking, καθώς το χαρτοφυλάκιο που επιδιώκει να αναπαράγει τον δείκτη πρέπει να αποτυπώνει τη συμπεριφορά μιας αγοράς με σημαντική κλαδική διαφοροποίηση. Επιπλέον, πολλές εταιρείες του FTSE 100 έχουν έντονη διεθνή δραστηριότητα, επομένως η συμπεριφορά του δείκτη δεν αντανακλά αποκλειστικά την εγχώρια οικονομική δραστηριότητα του Ηνωμένου Βασιλείου.

Στο benchmark dataset indtrack3 περιλαμβάνονται 89 μετοχές και η αντίστοιχη χρονοσειρά του FTSE 100. Η μεγαλύτερη διάσταση του επενδυτικού συνόλου, σε σύγκριση με το indtrack1, επιτρέπει την εξέταση του κατά πόσο οι χρησιμοποιούμενες μέθοδοι μπορούν να εντοπίσουν ένα μικρό υποσύνολο τίτλων που είναι επαρκές για την αναπαραγωγή της συμπεριφοράς ενός ευρύτερου χρηματιστηριακού δείκτη.

- Nikkei 225

Ο Nikkei 225, γνωστός και ως Nikkei Stock Average, αποτελεί έναν από τους σημαντικότερους χρηματιστηριακούς δείκτες της Ιαπωνίας και περιλαμβάνει 225 μετοχές εισηγμένες στην ιαπωνική χρηματιστηριακή αγορά. Ο δείκτης παρουσιάζει σημαντική κλαδική διαφοροποίηση και περιλαμβάνει επιχειρήσεις από κλάδους όπως η αυτοκινητοβιομηχανία, η ηλεκτρονική και τεχνολογία, οι χρηματοπιστωτικές υπηρεσίες, το λιανικό εμπόριο και η βιομηχανία.

Ενδεικτικά παραδείγματα εταιρειών που περιλαμβάνονται στον Nikkei 225 είναι η Toyota Motor, η Honda Motor και η Nissan Motor από τον κλάδο της αυτοκινητοβιομηχανίας, η Sony Group, η Panasonic Holdings, η Hitachi και η Tokyo Electron από τους κλάδους ηλεκτρονικής και τεχνολογίας, καθώς και η Mitsubishi UFJ Financial Group από τον χρηματοπιστωτικό τομέα. Η επίσημη σύνθεση του δείκτη καταγράφει επίσης εταιρείες όπως η Fast Retailing, η Canon, η Denso και η Mitsubishi Corporation, γεγονός που αναδεικνύει την ευρεία κλαδική κάλυψή του.

Στο benchmark dataset indtrack5 περιλαμβάνονται 225 επιμέρους μετοχές και η αντίστοιχη χρονοσειρά του Nikkei 225. Επομένως, το συγκεκριμένο dataset αποτελεί το μεγαλύτερο από τα τρία επενδυτικά σύνολα που χρησιμοποιούνται στην παρούσα εργασία. Η διάστασή του είναι ιδιαίτερα χρήσιμη για την αξιολόγηση των αλγοριθμικών μεθόδων, καθώς η αναζήτηση ενός μικρού χαρτοφυλακίου που αναπαράγει έναν δείκτη από ένα σύνολο 225 διαθέσιμων τίτλων αποτελεί σαφώς πιο απαιτητικό υπολογιστικό πρόβλημα.

Συνοπτικά, χρησιμοποιήθηκαν τρία από τα κλασικά διεθνή benchmark datasets της OR-Library (J.E.Beasley, 1990, link: <https://people.brunel.ac.uk/~mastjjb/jeb/info.html>): το Hang Seng (indtrack1), το FTSE 100 (indtrack3) και το Nikkei 225 (indtrack5). Κάθε dataset περιλαμβάνει 291 εβδομαδιαίες τιμές δείκτη και constituent stocks, από τις οποίες προκύπτουν 290 εβδομαδιαίες αποδόσεις. Ακολουθήθηκε συμμετρικός διαχωρισμός 145 περιόδων για in-sample εκπαίδευση και 145 περιόδων για out-of-sample αξιολόγηση.

Ο όρος in-sample αναφέρεται στο τμήμα των δεδομένων που χρησιμοποιείται κατά τη διαδικασία βελτιστοποίησης μιας μεθόδου. Στην παρούσα εργασία, τα in-sample δεδομένα χρησιμοποιούνται από τις εξεταζόμενες μεθόδους για τον προσδιορισμό του χαρτοφυλακίου που επιτυγχάνει την καλύτερη δυνατή παρακολούθηση του αντίστοιχου χρηματιστηριακού δείκτη.

Αντίθετα, τα out-of-sample δεδομένα είναι το τμήμα των δεδομένων που δεν χρησιμοποιείται κατά τη διαδικασία βελτιστοποίησης. Το χαρτοφυλάκιο που προκύπτει από τα in-sample δεδομένα εφαρμόζεται στη συνέχεια στα out-of-sample δεδομένα, προκειμένου να αξιολογηθεί η απόδοσή του σε δεδομένα που δεν είχε προηγουμένως χρησιμοποιήσει.

Η διάκριση αυτή επιτρέπει να εξεταστεί κατά πόσο μια μέθοδος μπορεί να διατηρήσει την αποτελεσματικότητά της πέρα από τα δεδομένα στα οποία πραγματοποιήθηκε η βελτιστοποίηση και να περιοριστεί ο κίνδυνος υπερπροσαρμογής (overfitting).

Η κύρια έμφαση δίνεται στην ικανότητα των μεθόδων να γενικεύουν σε δεδομένα που δεν χρησιμοποιήθηκαν κατά την κατασκευή του χαρτοφυλακίου. Για τον λόγο αυτό, ως βασικό κριτήριο χρησιμοποιείται το εκτός δείγματος Μέσο Τετραγωνικό Σφάλμα παρακολούθησης (out-of-sample tracking MSE). Επιπλέον, εξετάζονται ο αριθμός νικών ανά σενάριο, η μέση κατάταξη, η σχετική επίδοση ως προς τον νικητή κάθε σεναρίου και η συμπεριφορά σύγκλισης του Γενετικού Αλγορίθμου.

5.2 Πειραματικός Σχεδιασμός

Τα δεδομένα κάθε δείκτη χωρίζονται χρονολογικά σε δύο ίσα τμήματα. Το πρώτο 50% των παρατηρήσεων χρησιμοποιείται ως σύνολο εκπαίδευσης, πάνω στο οποίο γίνεται η επιλογή τίτλων και ο υπολογισμός των βαρών, ενώ το δεύτερο 50% χρησιμοποιείται αποκλειστικά για εκτός δείγματος αξιολόγηση. Η επιλογή αυτή αποφεύγει την τυχαία ανάμειξη παλαιότερων και νεότερων παρατηρήσεων και προσομοιώνει καλύτερα τη λογική μιας επενδυτικής εφαρμογής, στην οποία ένα χαρτοφυλάκιο που κατασκευάζεται από ιστορικά δεδομένα αξιολογείται σε μεταγενέστερη περίοδο.

Σε όλα τα πειράματα επιβάλλεται ακριβής πληθικότητα K , δηλαδή κάθε μέθοδος οφείλει να επιστρέφει ακριβώς K ενεργούς τίτλους. Εξετάζονται οι τιμές $K=5$, $K=10$ και $K=15$. Με αυτόν τον τρόπο διασφαλίζεται ότι οι μέθοδοι συγκρίνονται σε χαρτοφυλάκια ίδιας πολυπλοκότητας και ότι μια πιθανή υπεροχή δεν οφείλεται απλώς στη χρήση μεγαλύτερου αριθμού τίτλων.

Πίνακας 5.2.1 Βασικές παράμετροι της τελικής πειραματικής εκτέλεσης.

Παράμετρος	Τιμή
Δείκτες	Hang Seng, FTSE 100, Nikkei 225
Τιμές K	5, 10, 15
Διαχωρισμός δεδομένων	50% εκπαίδευση / 50% έλεγχος
Μέγεθος πληθυσμού ΓΑ	80
Μέγιστες γενιές ΓΑ	150

Παράμετρος	Τιμή
Πιθανότητα διασταύρωσης	0,90
Πιθανότητα μετάλλαξης	0,25
Tournament size	3
Elitism	4 άτομα
Early stopping	30 γενιές χωρίς ουσιαστική βελτίωση
Random Search	2.500 δοκιμές ανά σενάριο
Βασικός seed	42
Περιορισμός χαρτοφυλακίου	Long-only, ακριβώς K ενεργοί τίτλοι, άθροισμα βαρών = 1
Όρια βαρών	$0 \leq w_i \leq 1$
Κόστη συναλλαγών	Απενεργοποιημένα στα πειράματα του κεφαλαίου
Tolerance βελτίωσης	1×10^{-10}
Επαναλήψεις ΓΑ	5 ανεξάρτητα runs ανά σενάριο
Base seeds επαναλήψεων	42, 43, 44, 45, 46
Runs ανάλυσης σταθερότητας ΓΑ	45 (3 δείκτες $\times$ 3 τιμές K $\times$ 5 επαναλήψεις)
Runs ελέγχου ευαισθησίας	35 (7 ρυθμίσεις $\times$ 5 επαναλήψεις, Hang Seng K=10)
Συνολικά τεκμηριωμένα runs ΓΑ	80

Η πειραματική διαδικασία οργανώθηκε έτσι ώστε να διαχωρίζεται σαφώς η φάση βελτιστοποίησης από τη φάση αξιολόγησης. Για κάθε έναν από τους τρεις δείκτες οι αποδόσεις διαχωρίστηκαν χρονολογικά σε 50% δεδομένα εκπαίδευσης και 50% δεδομένα ελέγχου. Ο Γενετικός Αλγόριθμος είχε πρόσβαση αποκλειστικά στο πρώτο τμήμα κατά την αναζήτηση των τίτλων και των βαρών. Το out-of-sample MSE υπολογίστηκε μόνο μετά την ολοκλήρωση της βελτιστοποίησης, χωρίς επαναπροσαρμογή στο σύνολο ελέγχου.

Για κάθε τιμή $K \in \{5, 10, 15\}$ ο αρχικός πληθυσμός αποτελείται από 80 εφικτά χαρτοφυλάκια. Κάθε άτομο του πληθυσμού είναι ένα διάνυσμα βαρών μήκους ίσου με τον αριθμό των τίτλων του δείκτη, αλλά ακριβώς K βάρη είναι ενεργά, όλα τα βάρη είναι μη αρνητικά και το άθροισμά τους είναι ίσο με 1. Η συνάρτηση καταλληλότητας προκύπτει από το tracking MSE και από ποινές για τυχόν παραβιάσεις περιορισμών. Επειδή μετά από κάθε γενετικό τελεστή εφαρμόζεται διαδικασία repair, τα άτομα που εισέρχονται στον πληθυσμό είναι εφικτά και οι ποινές είναι πρακτικά μηδενικές.

Σε κάθε γενιά εφαρμόζεται tournament selection με μέγεθος 3. Με πιθανότητα 0,90 δύο γονείς συνδυάζονται μέσω αριθμητικής διασταύρωσης, ενώ με πιθανότητα 0,25 εφαρμόζεται μετάλλαξη. Η μετάλλαξη είτε αντικαθιστά έναν ενεργό τίτλο με έναν ανενεργό είτε διαταράσσει τα βάρη ενεργών τίτλων με μικρή κανονική μεταβολή. Μετά τη μετάλλαξη, η συνάρτηση repair επαναφέρει ακριβώς K ενεργούς τίτλους, μη αρνητικά βάρη και άθροισμα βαρών 1. Οι τέσσερις καλύτερες λύσεις περνούν αυτούσιες στην επόμενη γενιά (elitism). Η εκτέλεση σταματά στις 150 γενιές ή νωρίτερα εάν δεν εμφανιστεί βελτίωση μεγαλύτερη από 10^{-10} για 30 συνεχόμενες γενιές.

Για την κύρια σύγκριση των τεσσάρων μεθόδων χρησιμοποιείται το αναπαραγωγίμο canonical run με base seed 42. Ο πραγματικός seed κάθε σεναρίου ορίζεται ως $\text{base_seed} + 1000 \cdot d + K$, όπου $d=0$ για Hang Seng, $d=1$ για FTSE 100 και $d=2$ για Nikkei 225. Έτσι, για

παράδειγμα, το Hang Seng με $K=10$ εκτελείται με seed 52, το FTSE 100 με $K=10$ με seed 1052 και το Nikkei 225 με $K=10$ με seed 2052.

5.2 Μέθοδοι και κριτήρια αξιολόγησης

Ο Γενετικός Αλγόριθμος αναζητά ταυτόχρονα τον συνδυασμό των K ενεργών τίτλων και τα αντίστοιχα βάρη, με μηχανισμούς επιλογής, διασταύρωσης, μετάλλαξης και ελιτισμού. Η Τυχαία Αναζήτηση δημιουργεί 2.500 τυχαία ακριβώς- K χαρτοφυλάκια ανά σενάριο και διατηρεί εκείνο με το μικρότερο σφάλμα στο σύνολο εκπαίδευσης. Η Corr-EW επιλέγει τους K τίτλους με τη μεγαλύτερη απόλυτη συσχέτιση με τον δείκτη και τους αποδίδει ίσα βάρη. Τέλος, η Corr-Opt χρησιμοποιεί την ίδια επιλογή τίτλων βάσει συσχέτισης, αλλά υπολογίζει τα βάρη με περιορισμένη least-squares βελτιστοποίηση πάνω στο simplex, ώστε τα βάρη να είναι μη αρνητικά, να αθροίζουν στη μονάδα και να διατηρείται ακριβώς K ενεργή υποστήριξη.

Το tracking MSE υπολογίζεται ως ο μέσος όρος του τετραγώνου της διαφοράς μεταξύ της απόδοσης του χαρτοφυλακίου και της απόδοσης του δείκτη. Αν $r_{p,t}$ είναι η απόδοση του χαρτοφυλακίου και $r_{b,t}$ η απόδοση του δείκτη στη χρονική στιγμή t , τότε $MSE = (1/T) \cdot \sum_t (r_{p,t} - r_{b,t})^2$. Μικρότερη τιμή υποδηλώνει ακριβέστερη παρακολούθηση.

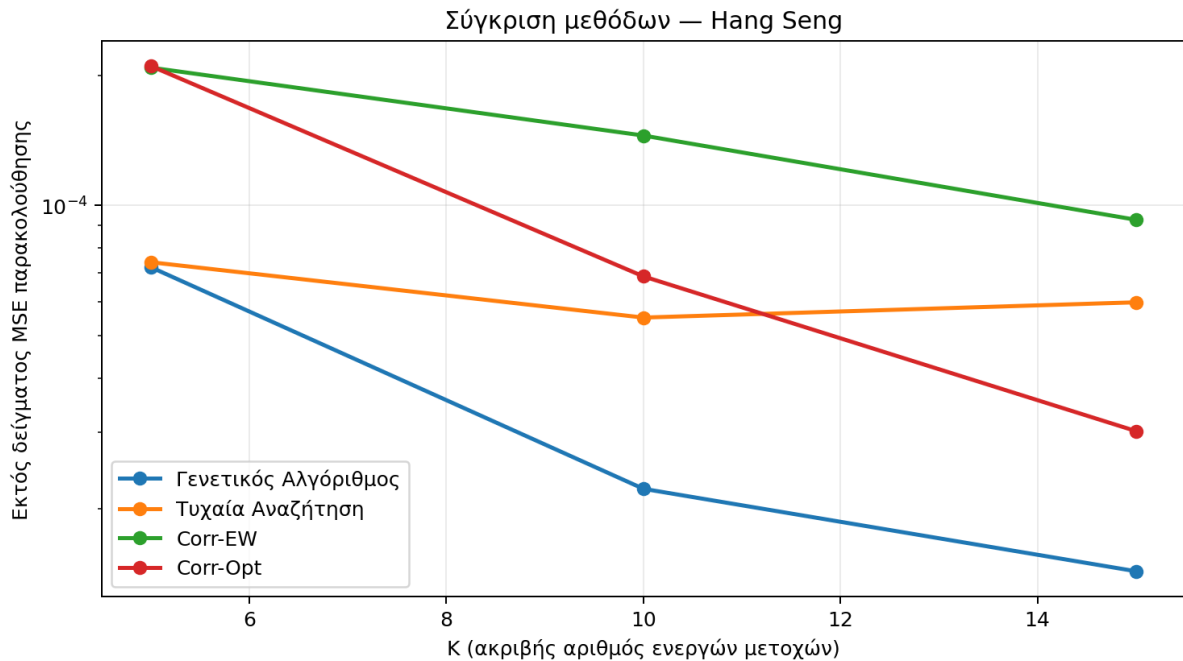
Για τη συνολική κατάταξη χρησιμοποιείται κατά κύριο λόγο η θέση κάθε μεθόδου μέσα σε κάθε ένα από τα εννέα σενάρια (3 δείκτες $\times 3$ τιμές K). Η χρήση μέσης κατάταξης είναι προτιμότερη από μια απλή σύγκριση των ακατέργαστων μέσων MSE, επειδή οι κλίμακες σφάλματος διαφέρουν μεταξύ δεικτών. Παράλληλα, υπολογίζεται το σχετικό MSE ως λόγος της επίδοσης κάθε μεθόδου προς την καλύτερη επίδοση του ίδιου σεναρίου. Τιμή 1,00 αντιστοιχεί στον νικητή, ενώ τιμή 2,00 σημαίνει διπλάσιο σφάλμα από τον νικητή.

5.3 Αποτελέσματα ανά χρηματιστηριακό δείκτη

5.3.1 Hang Seng

Πίνακας 5.3.1 Εκτός δείγματος MSE παρακολούθησης για τον δείκτη Hang Seng.

K	Γενετικός Αλγόριθμος	Τυχαία Αναζήτηση	Corr-EW	Corr-Opt	Καλύτερη μέθοδος
5	0,00007205	0,00007399	0,00020812	0,00020987	Γενετικός Αλγόριθμος
10	0,00002219	0,00005514	0,00014508	0,00006863	Γενετικός Αλγόριθμος
15	0,00001432	0,00005980	0,00009271	0,00003016	Γενετικός Αλγόριθμος



Σχήμα 5.1. Εκτός δείγματος MSE για τις τέσσερις μεθόδους στον Hang Seng.

Στον Hang Seng ο Γενετικός Αλγόριθμος καταλαμβάνει την πρώτη θέση και στις τρεις τιμές K. Για K=5 το εκτός δείγματος MSE είναι 0,00007205, οριακά χαμηλότερο από την Τυχαία Αναζήτηση (0,00007399), δηλαδή πλεονέκτημα περίπου 2,6% έναντι της δεύτερης καλύτερης μεθόδου. Η διαφορά γίνεται πολύ εντονότερη όταν αυξάνεται το K. Για K=10 το MSE του Γενετικού Αλγορίθμου μειώνεται σε 0,00002219, περίπου 59,8% χαμηλότερα από την Τυχαία Αναζήτηση, ενώ για K=15 περιορίζεται σε 0,00001432, περίπου 52,5% χαμηλότερα από την Corr-Opt.

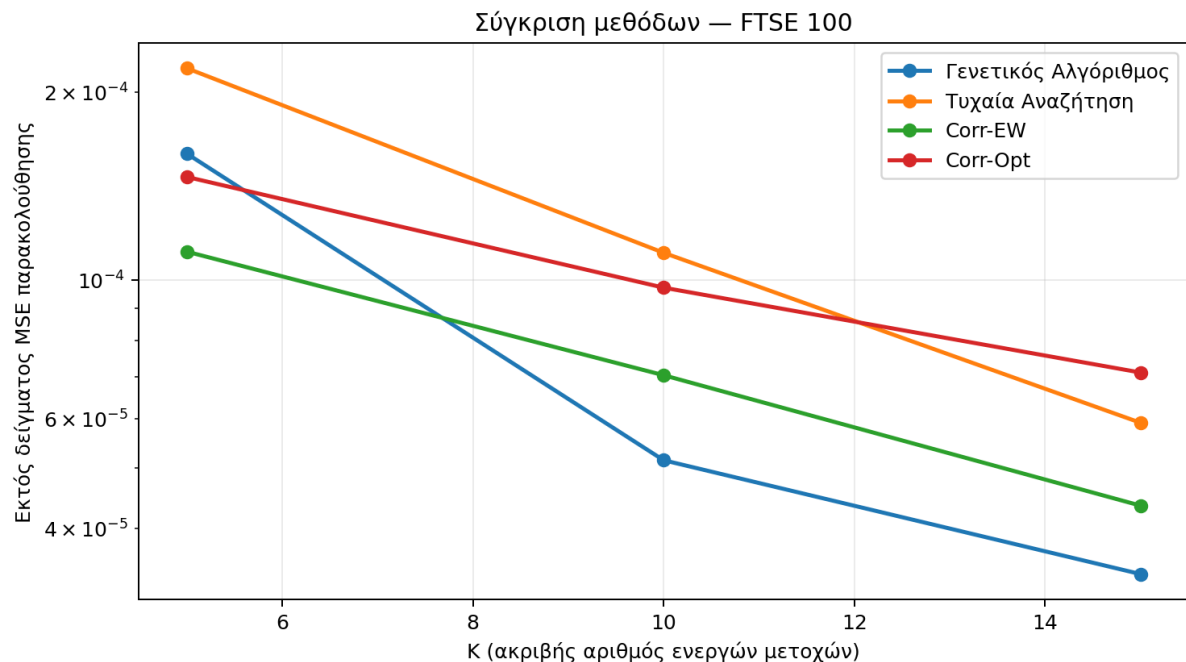
Η αύξηση του K από 5 σε 15 μειώνει το εκτός δείγματος σφάλμα του Γενετικού Αλγορίθμου κατά περίπου 80%. Η Corr-Opt επίσης επωφελείται σημαντικά από την αύξηση της πληθικότητας, αλλά παραμένει αισθητά χειρότερη από τον Γενετικό Αλγόριθμο. Αντίθετα, η Τυχαία Αναζήτηση βελτιώνεται μέχρι K=10 και στη συνέχεια παρουσιάζει μικρή επιδείνωση στο K=15, στοιχείο που δείχνει ότι η αύξηση του αριθμού των τίτλων δεν εγγυάται από μόνη της καλύτερη γενίκευση όταν η διαδικασία αναζήτησης δεν κατευθύνεται επαρκώς.

5.3.2 FTSE 100

Πίνακας 5.3.2 Εκτός δείγματος MSE παρακολούθησης για τον δείκτη FTSE 100.

K	Γενετικός Αλγόριθμος	Τυχαία Αναζήτηση	Corr-EW	Corr-Opt	Καλύτερη μέθοδος
5	0,00015930	0,00021822	0,00011096	0,00014616	Corr-EW
10	0,00005142	0,00011049	0,00007035	0,00009716	Γενετικός Αλγόριθμος

Κ	Γενετικός Αλγόριθμος	Τυχαία Αναζήτηση	Corr-EW	Corr-Opt	Καλύτερη μέθοδος
15	0,00003380	0,00005904	0,00004350	0,00007105	Γενετικός Αλγόριθμος



Σχήμα 5.2. Εκτός δείγματος MSE για τις τέσσερις μεθόδους στον FTSE 100.

Ο FTSE 100 αποτελεί τη μοναδική περίπτωση στην οποία ο Γενετικός Αλγόριθμος δεν κερδίζει όλα τα σενάρια. Για $K=5$, η Corr-EW εμφανίζει το χαμηλότερο εκτός δείγματος MSE (0,00011096), ενώ ο Γενετικός Αλγόριθμος ακολουθεί στην τρίτη θέση με 0,00015930.

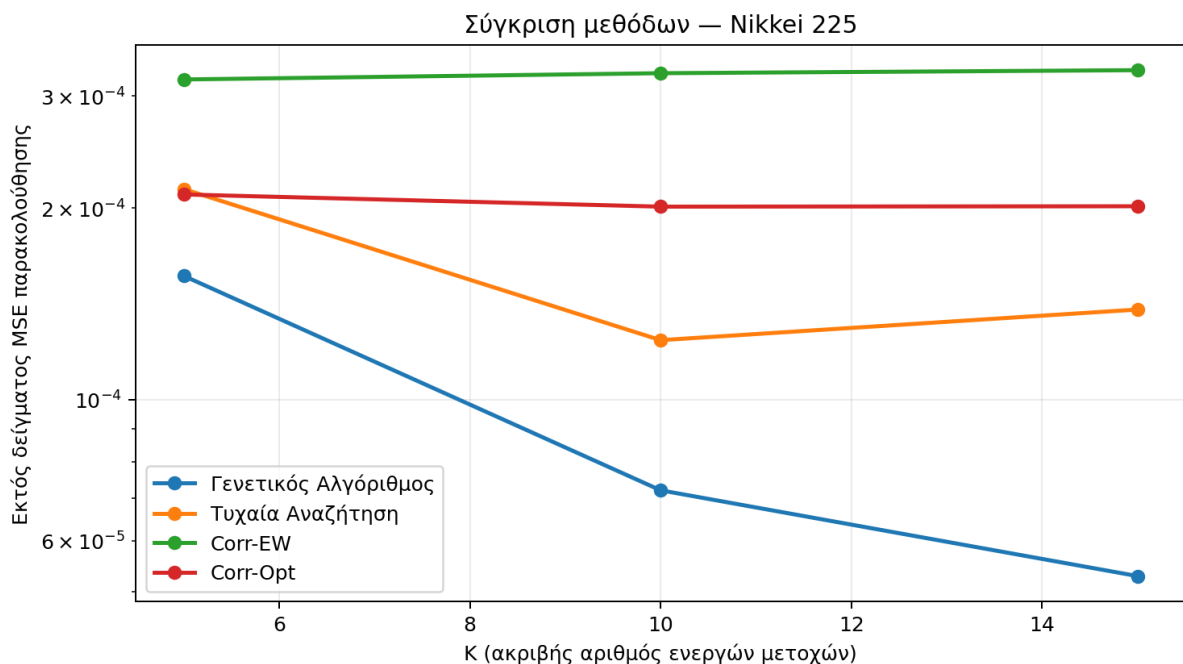
Το αποτέλεσμα αυτό είναι ιδιαίτερα ενδιαφέρον, επειδή το MSE εκπαίδευσης του Γενετικού Αλγορίθμου στο ίδιο σενάριο είναι μόλις 0,00006417, πολύ χαμηλότερο από το αντίστοιχο της Corr-EW. Συνεπώς, το $K=5$ στον FTSE 100 αποτελεί σαφή ένδειξη ότι η πολύ καλή προσαρμογή στο σύνολο εκπαίδευσης δεν μεταφράζεται πάντοτε σε καλύτερη εκτός δείγματος επίδοση.

Για $K=10$ και $K=15$ η εικόνα αντιστρέφεται. Ο Γενετικός Αλγόριθμος επιτυγχάνει MSE 0,00005142 και 0,00003380 αντίστοιχα, κερδίζοντας τη Corr-EW κατά περίπου 26,9% και 22,3%. Η σταδιακή μείωση του σφάλματος υποδεικνύει ότι στον FTSE 100 απαιτείται μεγαλύτερη πληθικότητα ώστε ο Γενετικός Αλγόριθμος να έχει αρκετή ευελιξία για να αναπαραστήσει τη συμπεριφορά του δείκτη χωρίς υπερβολική συγκέντρωση σε λίγους τίτλους.

5.3.3 Nikkei 225

Πίνακας 5.3.3 Εκτός δείγματος MSE παρακολούθησης για τον δείκτη Nikkei 225.

K	Γενετικός Αλγόριθμος	Τυχαία Αναζήτηση	Corr-EW	Corr-Opt	Καλύτερη μέθοδος
5	0,00015658	0,00021433	0,00031862	0,00021019	Γενετικός Αλγόριθμος
10	0,00007208	0,00012408	0,00032592	0,00020112	Γενετικός Αλγόριθμος
15	0,00005285	0,00013862	0,00032950	0,00020142	Γενετικός Αλγόριθμος



Σχήμα 5.3. Εκτός δείγματος MSE για τις τέσσερις μεθόδους στον Nikkei 225.

Στον Nikkei 225 ο Γενετικός Αλγόριθμος είναι ο καθαρός νικητής και για τις τρεις τιμές K. Για K=5 το MSE είναι 0,00015658, περίπου 25,5% χαμηλότερο από την Corr-Opt. Για K=10 η διαφορά έναντι της δεύτερης Τυχαίας Αναζήτησης αυξάνεται σε περίπου 41,9%, ενώ για K=15 φθάνει περίπου στο 61,9%. Η επίδοση του Γενετικού Αλγορίθμου βελτιώνεται συστηματικά όσο αυξάνεται το K, από 0,00015658 σε 0,00005285, δηλαδή μείωση περίπου 66%.

Αξιοσημείωτη είναι η συμπεριφορά της Corr-EW, της οποίας το εκτός δείγματος MSE δεν βελτιώνεται με την αύξηση του K: από 0,00031862 στο K=5 αυξάνεται σε 0,00032592 και 0,00032950. Αυτό δείχνει ότι η απλή προσθήκη περισσότερων τίτλων υψηλής συσχέτισης με ίσα βάρη μπορεί να εισάγει πλεονάζουσα ή αλληλοεπικαλυπτόμενη πληροφορία. Η Corr-Opt βελτιώνει αισθητά την Corr-EW μέσω της βελτιστοποίησης των βαρών, όμως εξακολουθεί να υστερεί από τον Γενετικό Αλγόριθμο, ο οποίος δεν περιορίζεται σε μια προεπιλεγμένη ομάδα τίτλων βάσει μεμονωμένων συσχετίσεων.

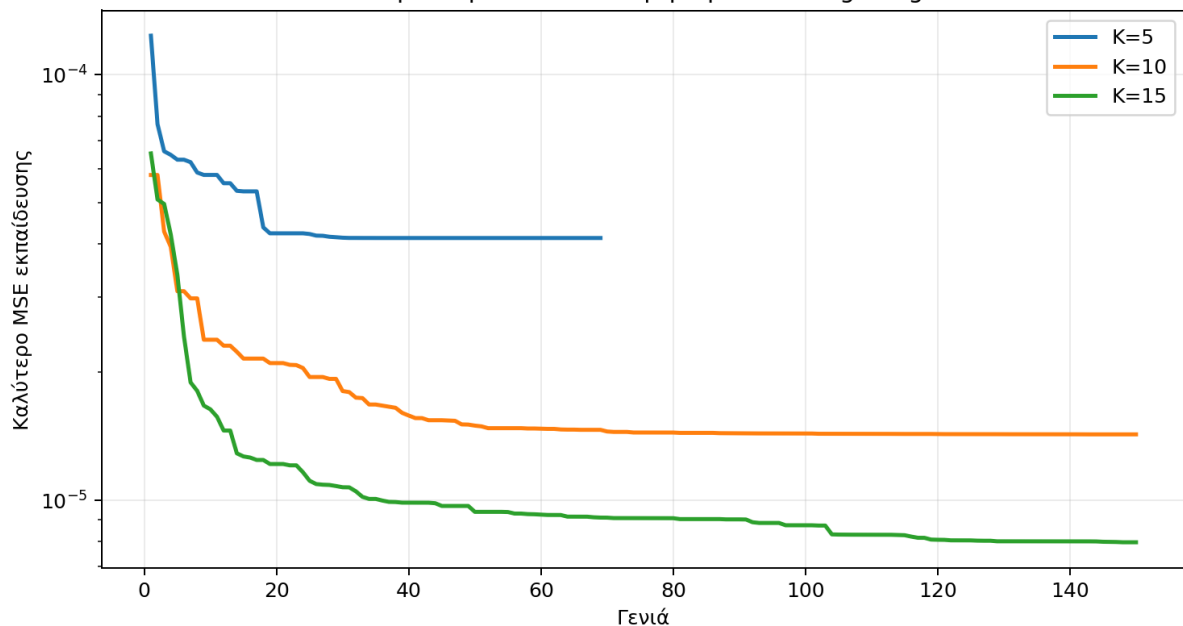
5.4 Ανάλυση σύγκλισης του Γενετικού Αλγορίθμου

Η εξέλιξη του καλύτερου MSE ανά γενιά επιτρέπει την άμεση παρατήρηση της διαδικασίας βελτιστοποίησης και αποτελεί βασικό στοιχείο της γραφικής διεπαφής. Στα περισσότερα σενάρια το σφάλμα μειώνεται απότομα στις πρώτες γενιές και στη συνέχεια η βελτίωση γίνεται σταδιακότερη, καθώς ο πληθυσμός συγκεντρώνεται γύρω από καλύτερες περιοχές του χώρου λύσεων. Αυτό είναι η αναμενόμενη συμπεριφορά ενός εξελικτικού αλγορίθμου που συνδυάζει εξερεύνηση μέσω μετάλλαξης και διασταύρωσης με εκμετάλλευση των καλύτερων λύσεων μέσω ελιτισμού.

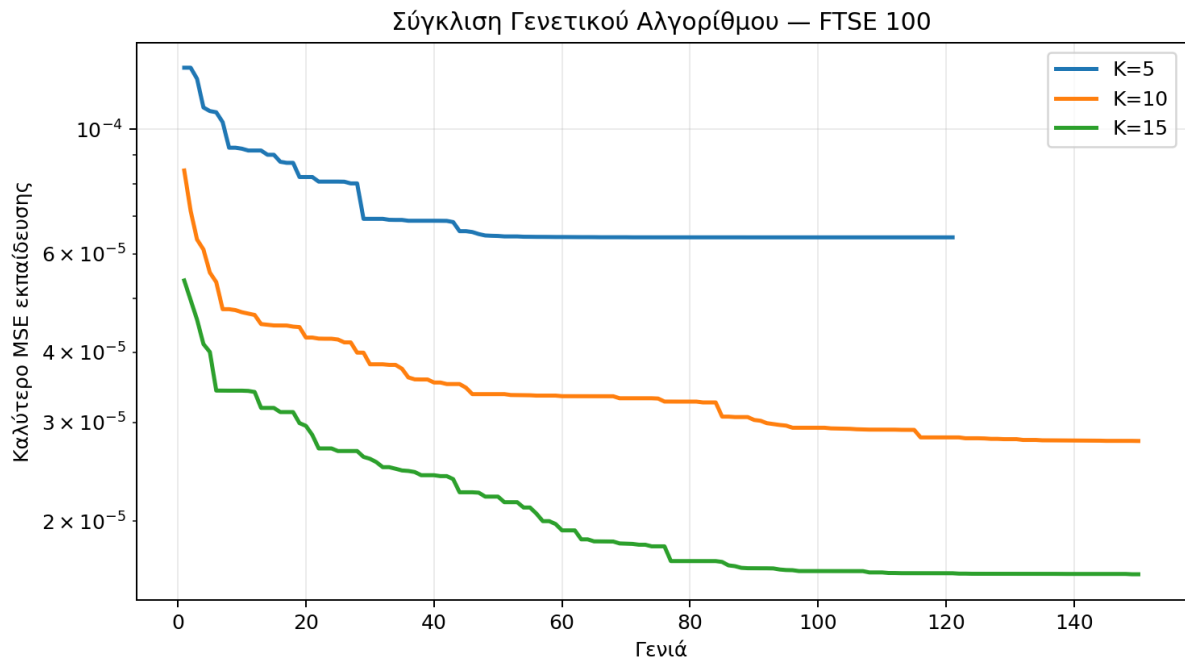
Πίνακας 5.4.1 Μεταβολή του καλύτερου MSE εκπαίδευσης κατά την εκτέλεση του Γενετικού Αλγορίθμου.

Δείκτης	K	Γενιές	Αρχικό καλύτερο MSE	Τελικό καλύτερο MSE	Μείωση
Hang Seng	5	69	0,00012356	0,00004135	66,5%
Hang Seng	10	150	0,00005815	0,00001429	75,4%
Hang Seng	15	150	0,00006534	0,00000797	87,8%
FTSE 100	5	121	0,00012883	0,00006417	50,2%
FTSE 100	10	150	0,00008445	0,00002780	67,1%
FTSE 100	15	150	0,00005375	0,00001608	70,1%
Nikkei 225	5	150	0,00014695	0,00007334	50,1%
Nikkei 225	10	150	0,00006967	0,00002593	62,8%
Nikkei 225	15	150	0,00005712	0,00001899	66,8%

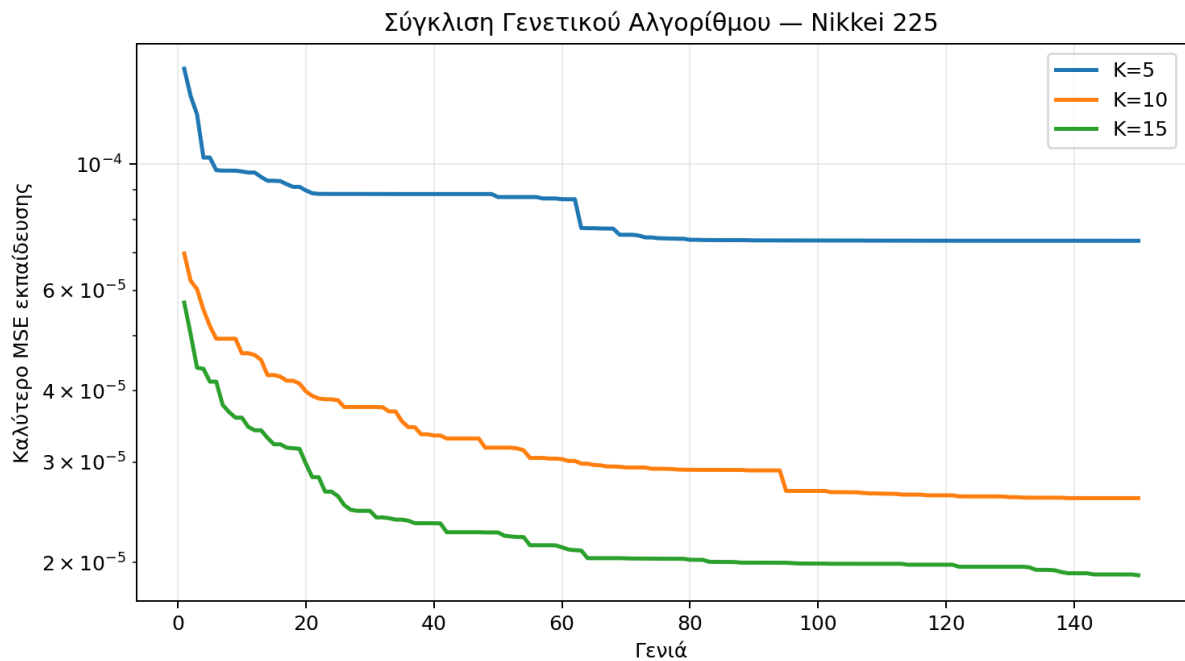
Σύγκλιση Γενετικού Αλγορίθμου — Hang Seng



Σχήμα 5.4. Σύγκλιση του Γενετικού Αλγορίθμου στον Hang Seng για $K=5$, 10 και 15.



Σχήμα 5.5. Σύγκλιση του Γενετικού Αλγορίθμου στον FTSE 100 για $K=5$, 10 και 15.



Σχήμα 5.6. Σύγκλιση του Γενετικού Αλγορίθμου στον Nikkei 225 για $K=5$, 10 και 15.

Η ποσοστιαία μείωση του καλύτερου MSE εκπαίδευσης κυμαίνεται από περίπου 50% έως 88%. Η μεγαλύτερη σχετική μείωση παρατηρείται στον Hang Seng με $K=15$ (87,8%), ενώ ακόμη και στα δυσκολότερα σενάρια η βελτίωση υπερβαίνει το 50%. Στον Hang Seng με $K=5$ ενεργοποιείται το κριτήριο πρόωρου τερματισμού μετά από 69 γενιές και στον FTSE

100 με $K=5$ μετά από 121 γενιές, ενώ τα υπόλοιπα σενάρια χρησιμοποιούν το μέγιστο των 150 γενεών. Το γεγονός αυτό δείχνει ότι ο μηχανισμός early stopping λειτουργεί ως πρακτικός περιορισμός υπολογιστικού κόστους όταν η βελτίωση έχει κορεστεί.

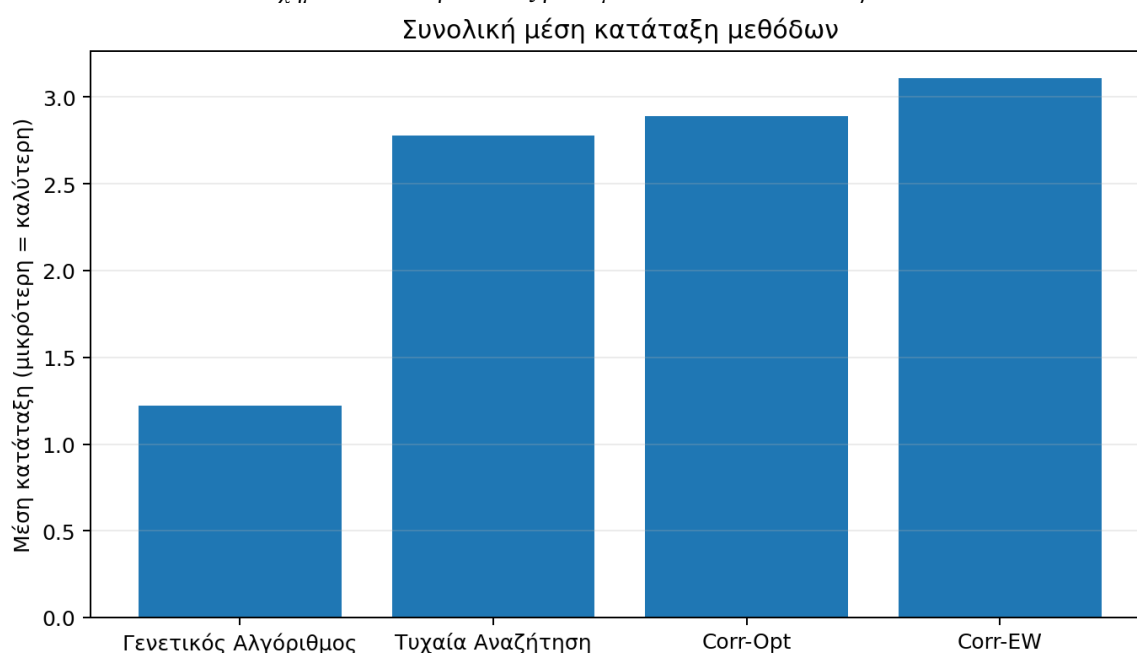
5.5 Συνολική σύγκριση των μεθόδων

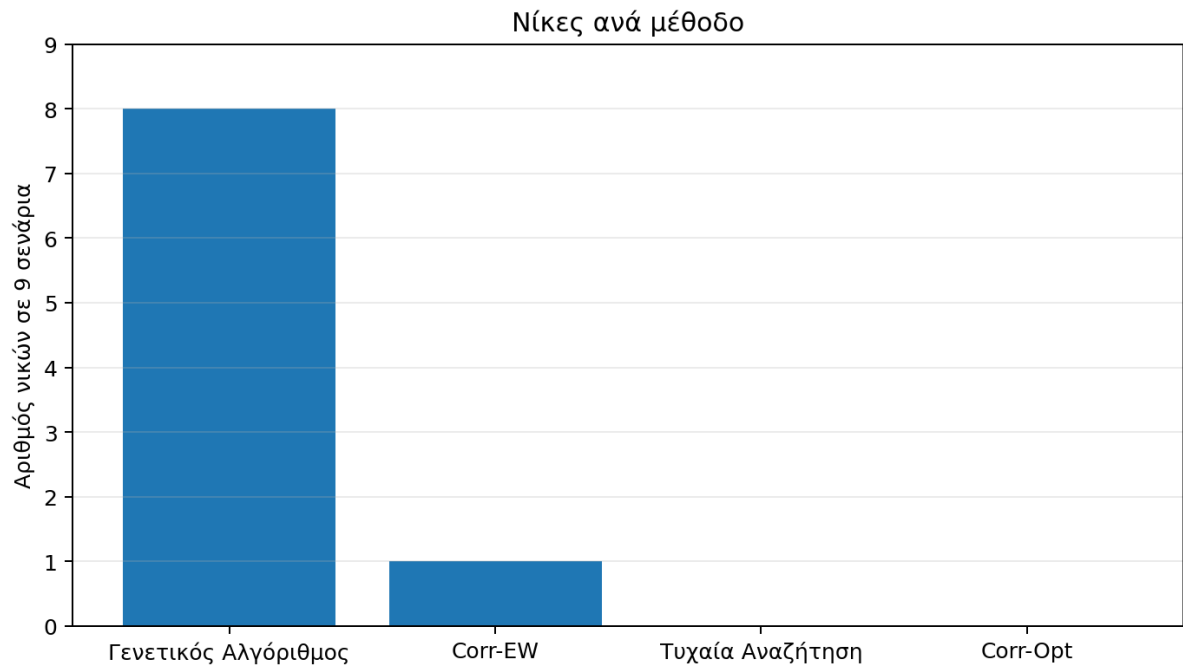
Η συνολική αξιολόγηση βασίζεται και στα εννέα σενάρια. Ο Γενετικός Αλγόριθμος κερδίζει 8 από τα 9, με μέση κατάταξη 1,22 και διάμεση κατάταξη 1. Η Τυχαία Αναζήτηση βρίσκεται συνολικά στη δεύτερη θέση με μέση κατάταξη 2,78, ακολουθεί η Corr-Opt με 2,89, ενώ η Corr-EW έχει μέση κατάταξη 3,11. Παρότι η Corr-EW κερδίζει το FTSE 100 στο $K=5$, η επίδοσή της είναι λιγότερο σταθερή στα υπόλοιπα σενάρια.

Πίνακας 5.5.1 Συνολική κατάταξη των τεσσάρων μεθόδων στα εννέα σενάρια.

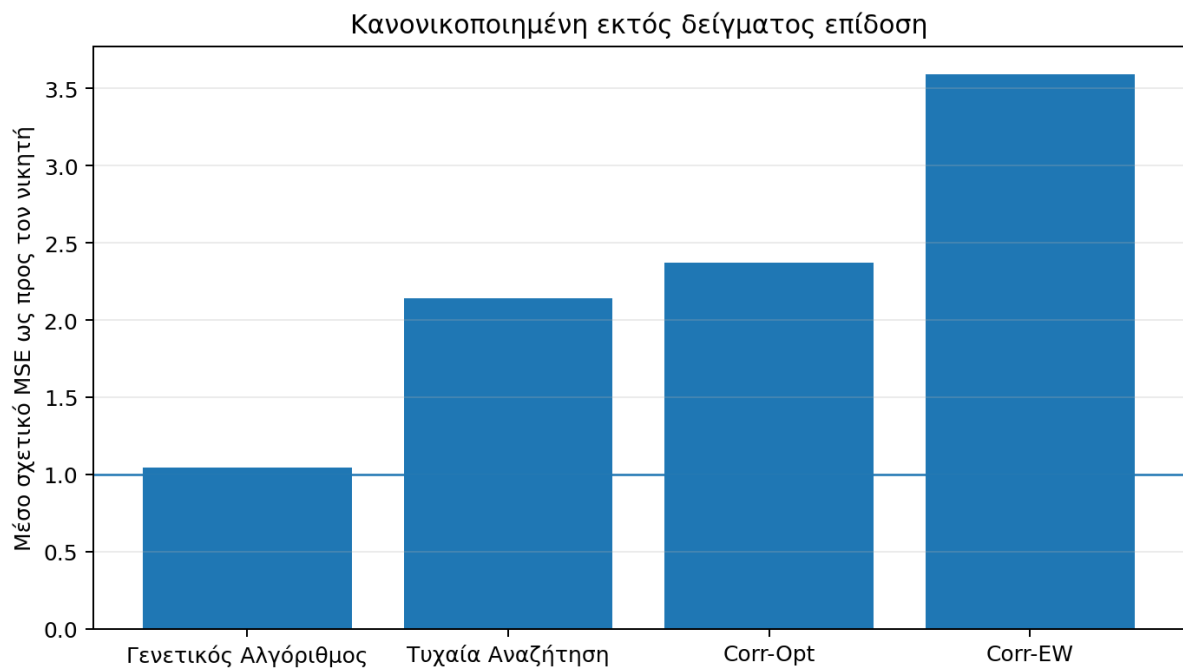
Θέση	Μέθοδος	Νίκες	Μέση θέση	Διάμεση θέση	Μέσο σχετικό MSE	Μέσο OOS MSE
1	Γενετικός Αλγόριθμος	8	1,222	1,000	1,048	0,00007051
2	Τυχαία Αναζήτηση	0	2,778	3,000	2,140	0,00011708
3	Corr-Opt	0	2,889	3,000	2,374	0,00013731
4	Corr-EW	1	3,111	4,000	3,594	0,00018275

Σχήμα 5.7. Μέση κατάταξη των μεθόδων στα εννέα σενάρια





Σχήμα 5.8. Αριθμός νικών ανά μέθοδο στα εννέα σενάρια.



Σχήμα 5.9. Μέσο σχετικό εκτός δείγματος MSE ως προς τον νικητή κάθε σεναρίου.

Το μέσο σχετικό MSE του Γενετικού Αλγορίθμου είναι 1,048. Επομένως, κατά μέσο όρο το σφάλμα του απέχει μόλις 4,8% από την καλύτερη τιμή κάθε σεναρίου — ουσιαστικά επειδή είναι ο ίδιος ο νικητής σε οκτώ περιπτώσεις και υστερεί μόνο στο FTSE 100 με $K=5$. Αντίθετα, η Τυχαία Αναζήτηση έχει μέσο σχετικό MSE 2,140, η Corr-Opt 2,374 και η Corr-EW 3,594. Με άλλα λόγια, οι τρεις εναλλακτικές μέθοδοι παρουσιάζουν κατά μέσο όρο υπερδιπλάσιο ή και υπερτριπλάσιο σφάλμα σε σχέση με τον νικητή του αντίστοιχου σεναρίου.

5.6 Έλεγχος ευαισθησίας βασικών παραμέτρων

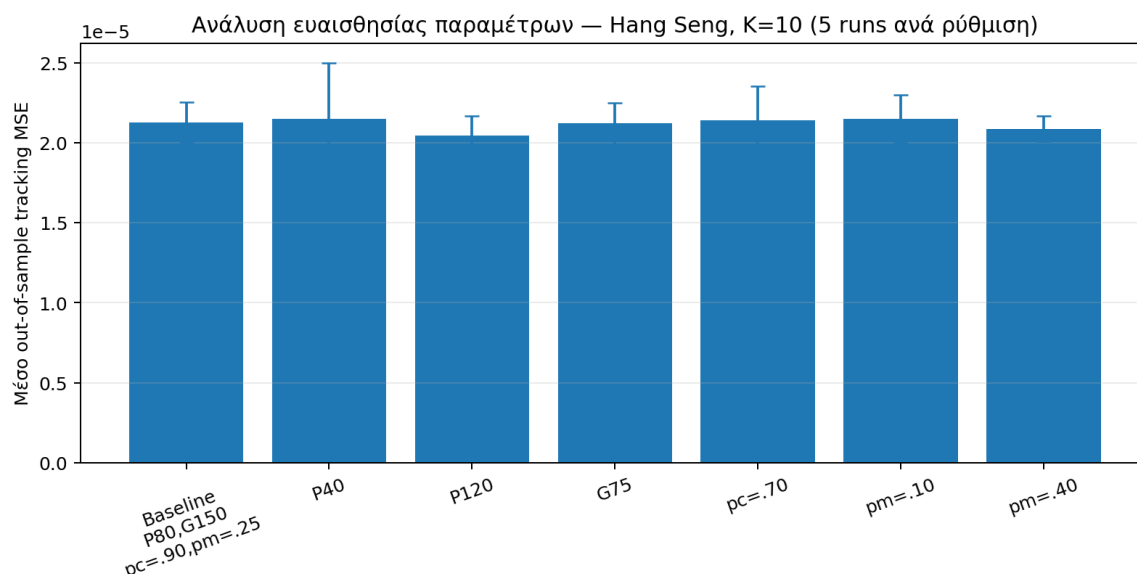
Πέρα από την τελική ρύθμιση, εξετάστηκε η ευαισθησία του ΓΑ σε βασικές υπερπαραμέτρους σε ένα αντιπροσωπευτικό σενάριο, Hang Seng με $K=10$. Η ανάλυση αυτή έχει διαγνωστικό χαρακτήρα και δεν χρησιμοποιείται για επιλογή παραμέτρων βάσει του test set. Για κάθε ρύθμιση πραγματοποιήθηκαν πέντε ανεξάρτητα runs με τα ίδια base seeds 42–46. Μεταβαλλόταν μία παράμετρος κάθε φορά, ενώ οι υπόλοιπες διατηρούνταν στις τιμές της baseline ρύθμισης ($P=80$, $G=150$, $p_c=0,90$, $p_m=0,25$). Συνολικά πραγματοποιήθηκαν 35 επιπλέον runs.

Πίνακας 5.6.1. Έλεγχος ευαισθησίας βασικών παραμέτρων στο Hang Seng, $K=10$ (5 runs ανά ρύθμιση).

Ρύθμιση	P	G	p_c	p_m	Μέσο OOS MSE	Τυπ. απόκλιση	Μέσος χρόνος (s)
Baseline	80	150	0,90	0,25	0,00002130	0,00000125	1,53
Population 40	40	150	0,90	0,25	0,00002150	0,00000349	0,68
Population 120	120	150	0,90	0,25	0,00002045	0,00000126	2,22
Generations 75	80	75	0,90	0,25	0,00002124	0,00000129	0,77
Crossover 0.70	80	150	0,70	0,25	0,00002144	0,00000211	1,43
Mutation 0.10	80	150	0,90	0,10	0,00002151	0,00000151	1,31
Mutation 0.40	80	150	0,90	0,40	0,00002089	0,00000079	1,48

Οι διαφορές είναι σχετικά μικρές σε σχέση με τη στοχαστική διακύμανση. Η αύξηση του πληθυσμού από 80 σε 120 μειώνει το μέσο OOS MSE από 0,00002130 σε 0,00002045, αλλά αυξάνει τον μέσο χρόνο εκτέλεσης από περίπου 1,53 s σε 2,22 s για το συγκεκριμένο σενάριο. Ο πληθυσμός 40 είναι ταχύτερος, όμως εμφανίζει μεγαλύτερη μεταβλητότητα.

Η μείωση του μέγιστου αριθμού γενεών σε 75 δίνει παρόμοιο μέσο αποτέλεσμα στο Hang Seng $K=10$, αλλά η τελική τιμή 150 διατηρήθηκε ώστε να παρέχεται επαρκές περιθώριο σύγκλισης και στα μεγαλύτερα προβλήματα. Αντίστοιχα, crossover 0,70 και mutation 0,10 δεν βελτιώνουν τη baseline κατά μέσο όρο, ενώ mutation 0,40 είναι ελαφρώς καλύτερο στο συγκεκριμένο σενάριο. Επειδή όμως η διερεύνηση αφορά μόνο ένα dataset/ K και δεν πρέπει να γίνεται επιλογή από το test set, η baseline ρύθμιση 0,90/0,25 διατηρείται ως σταθερή, μέτρια και συγκρίσιμη ρύθμιση για όλα τα κύρια πειράματα.



Σχήμα 5.10. Μέσο out-of-sample MSE και τυπική απόκλιση για τις επτά ρυθμίσεις ευαισθησίας (Hang Seng, K=10).

5.7 Ανεξάρτητες επαναλήψεις και έλεγχος στοχαστικής σταθερότητας

Επειδή ο Γενετικός Αλγόριθμος είναι στοχαστική μέθοδος, μία μόνο εκτέλεση δεν επαρκεί για να χαρακτηριστεί η σταθερότητά του. Για τον λόγο αυτό πραγματοποιήθηκαν πέντε ανεξάρτητες επαναλήψεις για κάθε ένα από τα εννέα σενάρια, χρησιμοποιώντας base seeds 42, 43, 44, 45 και 46. Συνολικά εκτελέστηκαν 45 runs του ΓΑ. Η πρώτη επανάληψη (base seed 42) είναι η εκτέλεση που χρησιμοποιείται στους κύριους συγκριτικούς πίνακες με τις άλλες μεθόδους, ενώ οι πέντε επαναλήψεις χρησιμοποιούνται για την εκτίμηση της διακύμανσης του αποτελέσματος.

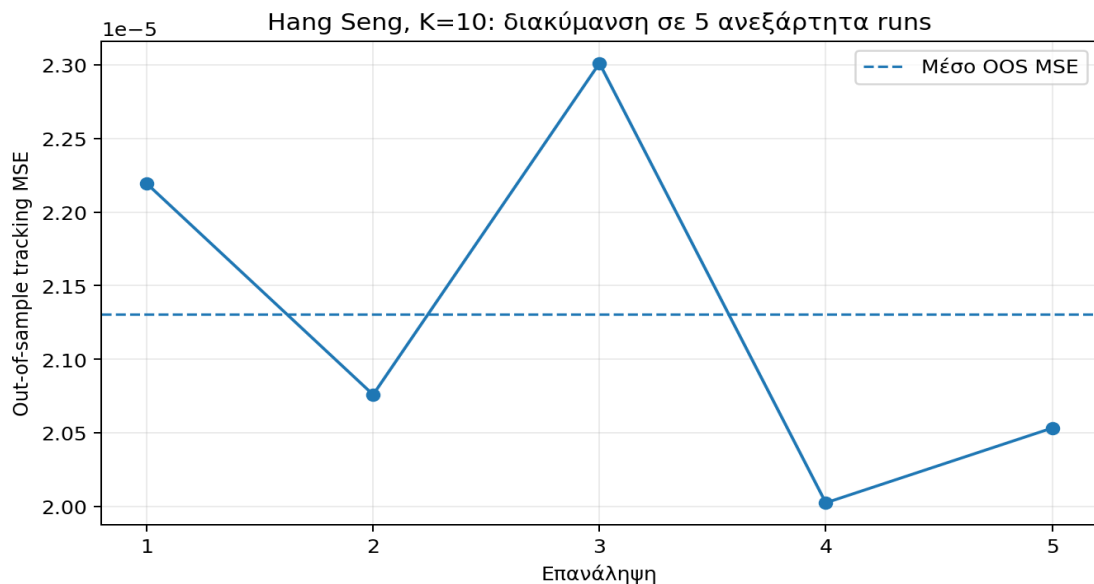
Πίνακας 5.7.1 Σταθερότητα του Γενετικού Αλγορίθμου σε 5 ανεξάρτητες επαναλήψεις ανά σενάριο.

Δείκτης	K	Runs	Μέσο OOS MSE	Τυπ. απόκλιση	Ελάχιστο	Μέγιστο	CV	Μέσες γενιές
Hang Seng	5	5	0,00007215	0,00000014	0,00007196	0,00007229	0,19%	126,40
Hang Seng	10	5	0,00002130	0,00000125	0,00002003	0,00002301	5,85%	150,00
Hang Seng	15	5	0,00001281	0,00000149	0,00001109	0,00001432	11,61%	150,00
FTSE 100	5	5	0,00013719	0,00001590	0,00011830	0,00015930	11,59%	132,20
FTSE 100	10	5	0,00005390	0,00000435	0,00004846	0,00006006	8,06%	150,00
FTSE 100	15	5	0,00004249	0,00000638	0,00003380	0,00005180	15,02%	150,00
Nikkei 225	5	5	0,00017101	0,00002917	0,00012614	0,00019283	17,06%	138,00
Nikkei 225	10	5	0,00009636	0,00002974	0,00007208	0,00014059	30,86%	150,00
Nikkei 225	15	5	0,00006935	0,00001533	0,00005285	0,00008597	22,11%	150,00

Τα αποτελέσματα δείχνουν ότι η επίδραση του seed δεν είναι ίδια σε όλα τα προβλήματα. Στον Hang Seng με $K=5$ η μεταβλητότητα είναι εξαιρετικά μικρή (CV 0,19%), γεγονός που δείχνει ότι διαφορετικές αρχικοποιήσεις οδηγούν ουσιαστικά στο ίδιο επίπεδο out-of-sample σφάλματος. Για $K=10$ και $K=15$ η μεταβλητότητα αυξάνεται σε 5,85% και 11,61% αντίστοιχα, αλλά παραμένει ελεγχόμενη. Στον FTSE 100 το CV κυμαίνεται περίπου από 8,06% έως 15,02%. Η μεγαλύτερη εξάρτηση από την αρχικοποίηση εμφανίζεται στον Nikkei 225, κυρίως για $K=10$ (CV 30,86%) και $K=15$ (22,11%). Αυτό είναι συμβατό με έναν μεγαλύτερο και πιο σύνθετο χώρο αναζήτησης, στον οποίο διαφορετικές αρχικές υποστηρίξεις μπορούν να οδηγήσουν σε διαφορετικά τοπικά άριστα.



Σχήμα 5.11 Συντελεστής μεταβλητότητας του out-of-sample MSE του ΓΑ στις πέντε επαναλήψεις.



Σχήμα 5.12. Out-of-sample MSE στις πέντε ανεξάρτητες εκτελέσεις για Hang Seng και $K=10$.

5.8 Παράδειγμα εξέλιξης του πληθυσμού του ΓΑ

Για να αποτυπωθεί σε επίπεδο χαρτοφυλακίων η εσωτερική λειτουργία του αλγορίθμου, εξετάζεται αναλυτικά το canonical run του Hang Seng με $K=10$ και seed 52. Ο πληθυσμός περιλαμβάνει 80 χαρτοφυλάκια. Στους επόμενους πίνακες παρουσιάζονται τα τρία καλύτερα άτομα του αρχικού πληθυσμού (γενιά 1), μιας ενδιάμεσης κατάστασης (γενιά 75) και της τελικής γενιάς 150. Τα βάρη εμφανίζονται μόνο για τους δέκα ενεργούς τίτλους.

Πίνακας 5.8.1 Αρχικός πληθυσμός – γενιά 1: τρία καλύτερα χαρτοφυλάκια

Θέση	Training MSE	Ενεργοί τίτλοι και βάρη
1	0,00005815	Stock_1=0.0747 Stock_4=0.0511 Stock_11=0.0213 Stock_15=0.1768 Stock_17=0.0970 Stock_18=0.1245 Stock_20=0.1697 Stock_26=0.1008 Stock_30=0.0985 Stock_31=0.0857
2	0,00007643	Stock_7=0.0259 Stock_10=0.0417 Stock_11=0.0299 Stock_12=0.1333 Stock_14=0.0644 Stock_15=0.2146 Stock_16=0.0634 Stock_20=0.2778 Stock_22=0.0127 Stock_24=0.1363
3	0,00008091	Stock_2=0.0701 Stock_3=0.1138 Stock_6=0.0978 Stock_10=0.0271 Stock_11=0.1468 Stock_13=0.1387 Stock_21=0.0248 Stock_22=0.1463 Stock_24=0.1255 Stock_26=0.1091

Πίνακας 5.8.2 Ενδιάμεσος πληθυσμός – γενιά 75: τρία καλύτερα χαρτοφυλάκια.

Θέση	Training MSE	Ενεργοί τίτλοι και βάρη
1	0,00001443	Stock_4=0.0887 Stock_11=0.1271 Stock_12=0.0721 Stock_13=0.0884 Stock_15=0.1890 Stock_20=0.0724 Stock_22=0.0733 Stock_26=0.0741 Stock_27=0.1039 Stock_28=0.1111
2	0,00001443	Stock_4=0.0887 Stock_11=0.1271 Stock_12=0.0721 Stock_13=0.0884 Stock_15=0.1890 Stock_20=0.0724 Stock_22=0.0733 Stock_26=0.0741 Stock_27=0.1038 Stock_28=0.1111

Θέση	Training MSE	Ενεργοί τίτλοι και βάρη
3	0,00001445	Stock_4=0.0878 Stock_11=0.1255 Stock_12=0.0714 Stock_13=0.0880 Stock_15=0.1872 Stock_20=0.0748 Stock_22=0.0735 Stock_26=0.0734 Stock_27=0.1084 Stock_28=0.1100

Πίνακας 5.8.3 Τελικός πληθυσμός – γενιά 150: τρία καλύτερα χαρτοφυλάκια.

Θέση	Training MSE	Ενεργοί τίτλοι και βάρη
1	0,00001429	Stock_4=0.0894 Stock_11=0.1300 Stock_12=0.0783 Stock_13=0.0876 Stock_15=0.1903 Stock_20=0.0690 Stock_22=0.0752 Stock_26=0.0663 Stock_27=0.1063 Stock_28=0.1076
2	0,00001429	Stock_4=0.0894 Stock_11=0.1300 Stock_12=0.0783 Stock_13=0.0876 Stock_15=0.1903 Stock_20=0.0690 Stock_22=0.0752 Stock_26=0.0663 Stock_27=0.1063 Stock_28=0.1077
3	0,00001429	Stock_4=0.0894 Stock_11=0.1300 Stock_12=0.0783 Stock_13=0.0876 Stock_15=0.1904 Stock_20=0.0690 Stock_22=0.0752 Stock_26=0.0664 Stock_27=0.1063 Stock_28=0.1073

Στην πρώτη γενιά τα καλύτερα χαρτοφυλάκια έχουν διαφορετικές υποστηρίξεις και το καλύτερο training MSE είναι 0,00005815, ενώ ο μέσος όρος του πληθυσμού είναι 0,00018471. Μέχρι τη γενιά 75 το καλύτερο MSE έχει μειωθεί σε 0,00001443 και τα καλύτερα άτομα έχουν πλέον συγκλίνει στην ίδια περίπου δεκάδα τίτλων και σε πολύ παρόμοια βάρη. Στη γενιά 150 το καλύτερο MSE είναι 0,00001429, δηλαδή περίπου 75,4% χαμηλότερο από το καλύτερο άτομο της αρχικοποίησης. Η πολύ μικρή διαφοροποίηση μεταξύ των κορυφαίων ατόμων στο τέλος αποτελεί ένδειξη σύγκλισης του πληθυσμού. Το αντίστοιχο out-of-sample MSE του τελικού χαρτοφυλακίου είναι 0,00002219.

5.9 Παραδείγματα τελικών χαρτοφυλακίων

Ως επιπλέον παράδειγμα, ο Πίνακας 5.9.1 παρουσιάζει τα τελικά χαρτοφυλάκια του ΓΑ για $K=10$ και για τους τρεις δείκτες στο canonical run. Οι ονομασίες Stock_i είναι οι εσωτερικές ονομασίες των σειρών τίτλων στα benchmark αρχεία OR-Library. Σε όλες τις περιπτώσεις υπάρχουν ακριβώς δέκα ενεργοί τίτλοι και τα βάρη αθροίζονται σε 1.

Πίνακας 5.9.1 Παραδείγματα τελικών χαρτοφυλακίων του ΓΑ για $K=10$.

Δείκτης	Seed	IS MSE	OOS MSE	Τελικό χαρτοφυλάκιο (τίτλος: βάρος)
Hang Seng	52	0,00001429	0,00002219	Stock_15: 0.1903; Stock_11: 0.1300; Stock_28: 0.1076; Stock_27: 0.1063; Stock_4: 0.0894; Stock_13: 0.0876; Stock_12: 0.0783; Stock_22: 0.0752; Stock_20: 0.0690; Stock_26: 0.0663
FTSE 100	1052	0,00002780	0,00005142	Stock_67: 0.1601; Stock_7: 0.1104; Stock_79: 0.1051; Stock_30: 0.1009; Stock_52: 0.0978; Stock_23: 0.0958; Stock_85: 0.0914; Stock_77: 0.0838; Stock_62: 0.0806; Stock_74: 0.0741
Nikkei 225	2052	0,00002593	0,00007208	Stock_104: 0.1553; Stock_212: 0.1257; Stock_160: 0.1086; Stock_64: 0.1064; Stock_53: 0.1062; Stock_8: 0.0902; Stock_187: 0.0887; Stock_113: 0.0861; Stock_214: 0.0724; Stock_28: 0.0603

5.10 Ψευδοκώδικας της υλοποίησης του Γενετικού Αλγορίθμου

Ο ακόλουθος ψευδοκώδικας συνοψίζει την πραγματική λογική της υλοποίησης. Περιλαμβάνει την αρχικοποίηση εφικτών χαρτοφυλακίων, την αξιολόγηση, την επιλογή με tournament, τη διασταύρωση, τη μετάλλαξη, το repair των περιορισμών, τον ελιτισμό και το κριτήριο πρόωρου τερματισμού.

Αλγόριθμος Γενετικός Αλγόριθμος για index tracking με ακριβώς K τίτλους

Είσοδος: R_{train} , r_{index} , K , $P=80$, $G=150$, $pc=0.90$, $pm=0.25$, $T=3$, $E=4$, $S=30$, seed

Έξοδος: καλύτερα βάρη w^* , ιστορικό MSE ανά γενιά

1. Θέσε τη γεννήτρια τυχαίων αριθμών με το seed.
2. Δημιούργησε P εφικτά χαρτοφυλάκια:
ακριβώς K ενεργοί τίτλοι, $w_i \geq 0$ και $\sum w_i = 1$.
3. Για κάθε άτομο w υπολόγισε $MSE(w) = \text{mean}[(R_{train} \cdot w - r_{index})^2]$.
4. Θέσε $global_best =$ καλύτερο άτομο και $stagnation = 0$.
5. Για $generation = 1, \dots, G$:
 - 5.1 Κατέγραψε best MSE και average MSE του πληθυσμού.
 - 5.2 Αν το best βελτιώνει το $global_best$ περισσότερο από 10^{-10} :
 $global_best = best$, $stagnation = 0$
αλλιώς $stagnation = stagnation + 1$.
 - 5.3 Αν $stagnation \geq S$, τερμάτισε.
 - 5.4 Αντέγραψε τα E καλύτερα άτομα στον νέο πληθυσμό (elitism).
 - 5.5 Όσο ο νέος πληθυσμός έχει λιγότερα από P άτομα:
 - $p1 = \text{TournamentSelection}(\text{population}, T)$
 - $p2 = \text{TournamentSelection}(\text{population}, T)$
 - με πιθανότητα pc : $(c1, c2) = \text{ArithmeticCrossover}(p1, p2)$
 - αλλιώς: $c1 = p1$ και $c2 = p2$
 - με πιθανότητα pm : $\text{Mutate}(c1)$, $\text{Mutate}(c2)$
 - $\text{Repair}(c1)$, $\text{Repair}(c2)$ ώστε να ισχύουν ακριβώς K , $w_i \geq 0$, $\sum w_i = 1$
 - $\text{Evaluate}(c1)$, $\text{Evaluate}(c2)$ και πρόσθεσέ τα στον νέο πληθυσμό.
 - 5.6 $\text{population} =$ νέος πληθυσμός.
6. Επέστρεψε $global_best$ και το ιστορικό σύγκλισης.
7. Υπολόγισε το out-of-sample MSE του $global_best$ στο σύνολο ελέγχου.

Η δομή αυτή εξηγεί και τη γραφική παρακολούθηση που εμφανίζει η εφαρμογή: σε κάθε γενιά αποθηκεύονται τόσο το καλύτερο όσο και το μέσο tracking MSE του πληθυσμού.

Επομένως, ο χρήστης μπορεί να παρακολουθεί όχι μόνο την τελική λύση αλλά και την πορεία της αναζήτησης προς τη σύγκλιση.

5.11 Ερμηνεία των αποτελεσμάτων

Το πρώτο και σημαντικότερο συμπέρασμα είναι ότι η ταυτόχρονη αναζήτηση τόσο της υποστήριξης του χαρτοφυλακίου όσο και των βαρών του προσφέρει ουσιαστικό πλεονέκτημα. Οι μέθοδοι Corr-EW και Corr-Opt ξεκινούν από μια μονοδιάστατη κατάταξη των τίτλων βάσει συσχέτισης. Η προσέγγιση αυτή αγνοεί σε πρώτο στάδιο τις αλληλεπιδράσεις μεταξύ τίτλων και την πιθανότητα ένας τίτλος με μέτρια ατομική συσχέτιση να είναι ιδιαίτερα χρήσιμος σε συνδυασμό με άλλους. Ο Γενετικός Αλγόριθμος, αντίθετα, εξετάζει συνδυασμούς και μπορεί να ανακαλύψει τέτοιες συμπληρωματικότητες. Δεύτερον, η αύξηση του K γενικά βελτιώνει την επίδοση του Γενετικού Αλγορίθμου. Η τάση είναι ιδιαίτερα καθαρή και στους τρεις δείκτες. Με περισσότερους ενεργούς τίτλους αυξάνεται η δυνατότητα αναπαράστασης της σύνθετης κίνησης του δείκτη, αλλά ταυτόχρονα μεγαλώνει ο χώρος αναζήτησης. Το γεγονός ότι ο Γενετικός Αλγόριθμος συνεχίζει να βελτιώνεται στα $K=10$ και $K=15$ δείχνει ότι οι εξελικτικοί τελεστές μπορούν να αξιοποιήσουν αποτελεσματικά αυτή την πρόσθετη ευελιξία.

Τρίτον, τα αποτελέσματα τονίζουν τη σημασία της εκτός δείγματος αξιολόγησης. Στο FTSE 100 με $K=5$ ο Γενετικός Αλγόριθμος έχει το χαμηλότερο σφάλμα εκπαίδευσης από όλες τις μεθόδους, αλλά δεν έχει το χαμηλότερο σφάλμα ελέγχου. Το συγκεκριμένο αποτέλεσμα λειτουργεί ως πρακτικό παράδειγμα υπερπροσαρμογής: μια λύση μπορεί να ταιριάζει εξαιρετικά στην περίοδο εκπαίδευσης, χωρίς να αναπαράγει το ίδιο καλά τη μελλοντική συμπεριφορά του δείκτη. Συνεπώς, η επιλογή μεθόδου αποκλειστικά βάσει in-sample MSE θα ήταν μεθοδολογικά ανεπαρκής.

Τέταρτον, η Τυχαία Αναζήτηση λειτουργεί ως χρήσιμος baseline. Σε ορισμένα σενάρια, όπως στον Hang Seng με $K=5$, πλησιάζει πολύ τον Γενετικό Αλγόριθμο. Ωστόσο, καθώς αυξάνεται η διάσταση του προβλήματος, η διαφορά μεγαλώνει. Αυτό είναι αναμενόμενο, επειδή 2.500 ανεξάρτητες τυχαίες δοκιμές καλύπτουν ολόένα μικρότερο ποσοστό του συνδυαστικού χώρου, ενώ ο Γενετικός Αλγόριθμος επαναχρησιμοποιεί πληροφορία από καλές λύσεις προηγούμενων γενεών.

5.12 Περιορισμοί της πειραματικής αξιολόγησης

Παρότι τα αποτελέσματα είναι σαφή, πρέπει να ερμηνεύονται υπό συγκεκριμένους περιορισμούς. Πρώτον, η τελική σύγκριση χρησιμοποιεί μία βασική τυχαία αρχικοποίηση ανά σενάριο, με ντετερμινιστικό seed για αναπαραγωγικότητα. Επειδή ο Γενετικός Αλγόριθμος και η Τυχαία Αναζήτηση είναι στοχαστικές μέθοδοι, μια αυστηρότερη στατιστική αξιολόγηση θα απαιτούσε πολλαπλές ανεξάρτητες επαναλήψεις ανά σενάριο και αναφορά μέσης τιμής, τυπικής απόκλισης και κατάλληλων στατιστικών ελέγχων. Δεύτερον, το κριτήριο σφάλματος είναι το MSE ενεργών αποδόσεων. Στη βιβλιογραφία ο όρος tracking error χρησιμοποιείται επίσης για την τυπική απόκλιση των active returns ή για τη ρίζα του MSE. Στην παρούσα υλοποίηση ο όρος πρέπει να ερμηνεύεται ρητά ως tracking MSE. Τρίτον, η συγκεκριμένη σύγκριση δεν περιλαμβάνει κόστη συναλλαγών, περιορισμούς turnover ή άλλες μικροδομές αγοράς. Επομένως, τα συμπεράσματα αφορούν

κυρίως την ακρίβεια παρακολούθησης και όχι το πλήρες οικονομικό κόστος μιας πραγματικής επενδυτικής στρατηγικής.

Τέλος, ο χρονολογικός διαχωρισμός 50/50 παρέχει ένα καθαρό και εύκολα αναπαραγώγιμο πρωτόκολλο, αλλά αντιστοιχεί σε μία συγκεκριμένη περίοδο εκπαίδευσης και μία περίοδο ελέγχου. Μελλοντική επέκταση θα μπορούσε να χρησιμοποιήσει rolling ή walk-forward αξιολόγηση, ώστε να εξεταστεί η σταθερότητα των συμπερασμάτων σε περισσότερα χρονικά καθεστώτα αγοράς.

5.13 Συμπεράσματα Κεφαλαίου

Η συγκριτική πειραματική αξιολόγηση δείχνει ότι ο Γενετικός Αλγόριθμος αποτελεί την αποτελεσματικότερη από τις τέσσερις εξεταζόμενες μεθόδους για το συγκεκριμένο πρόβλημα index tracking με περιορισμό ακριβούς πληθικότητας. Κερδίζει 8 από τα 9 σενάρια, επιτυγχάνει μέση κατάταξη 1,22 και μέσο σχετικό MSE μόλις 1,048 ως προς τον καλύτερο αλγόριθμο κάθε σεναρίου.

Η υπεροχή είναι ιδιαίτερα έντονη για $K=10$ και $K=15$, όπου ο χώρος επιλογής χαρτοφυλακίου γίνεται μεγαλύτερος και η κατευθυνόμενη εξελικτική αναζήτηση αξιοποιεί αποτελεσματικότερα τη διαθέσιμη ευελιξία από τις απλούστερες εναλλακτικές.

Παράλληλα, η μοναδική ήττα στο FTSE 100 με $K=5$ αναδεικνύει ότι η εκτός δείγματος αξιολόγηση παραμένει απαραίτητη και ότι κανένας αλγόριθμος δεν πρέπει να θεωρείται καθολικά ανώτερος χωρίς έλεγχο γενίκευσης.

Συνολικά, όμως, τα αποτελέσματα υποστηρίζουν ισχυρά τη χρήση του Γενετικού Αλγορίθμου ως βασικής μεθόδου της εφαρμογής, ενώ οι Random Search, Corr-EW και Corr-Opt λειτουργούν ως χρήσιμα σημεία αναφοράς για την ποσοτική τεκμηρίωση της προστιθέμενης αξίας της εξελικτικής βελτιστοποίησης.

ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ

6.1 Συνοπτική ανακεφαλαίωση της εργασίας

Η παρούσα διπλωματική εργασία είχε ως αντικείμενο τη διερεύνηση της εφαρμογής Γενετικών Αλγορίθμων στο πρόβλημα της βελτιστοποίησης χαρτοφυλακίου με χρήση της τεχνικής αποτύπωσης δείκτη. Το πρόβλημα αυτό αποτελεί μία ιδιαίτερα σημαντική περίπτωση χρηματοοικονομικής βελτιστοποίησης, καθώς συνδέεται με την κατασκευή χαρτοφυλακίων που επιδιώκουν να αναπαράγουν όσο το δυνατόν πιστότερα την πορεία ενός χρηματιστηριακού δείκτη αναφοράς, υπό συγκεκριμένους περιορισμούς.

Αρχικά παρουσιάστηκε το θεωρητικό υπόβαθρο της βελτιστοποίησης χαρτοφυλακίου και της αποτύπωσης δείκτη, καθώς και η σχετική βιβλιογραφία γύρω από τις κλασικές, ακριβείς και μεταερευτικές μεθόδους επίλυσης του προβλήματος. Ιδιαίτερη έμφαση δόθηκε στις εξελικτικές και μεταερευτικές προσεγγίσεις, και ειδικότερα στους Γενετικούς Αλγορίθμους, οι οποίοι θεωρούνται ιδιαίτερα κατάλληλοι για προβλήματα με μικτή δομή

απόφασης, περιορισμούς πληθικότητας και απαιτήσεις ρεαλιστικής διαχείρισης χαρτοφυλακίου.

Στη συνέχεια διατυπώθηκε το πρόβλημα της αποτύπωσης δείκτη ως πρόβλημα βελτιστοποίησης, με σαφή ορισμό της αντικειμενικής συνάρτησης, των μεταβλητών απόφασης και των βασικών περιορισμών. Η αντικειμενική συνάρτηση βασίστηκε στην ελαχιστοποίηση του tracking error, ενώ ενσωματώθηκαν περιορισμοί που αφορούν τον προϋπολογισμό, το μέγιστο πλήθος επιλεγμένων τίτλων, τα κατώτατα και ανώτατα όρια συμμετοχής και, σε γενικότερο επίπεδο, τη δυνατότητα ενσωμάτωσης κόστους συναλλαγών.

Με βάση αυτή τη διατύπωση σχεδιάστηκε και αναπτύχθηκε ένας προσαρμοσμένος Γενετικός Αλγόριθμος, ο οποίος περιλαμβάνει κατάλληλη αναπαράσταση χρωμοσώματος, διαδικασία δημιουργίας αρχικού πληθυσμού, συνάρτηση καταλληλότητας, τελεστές επιλογής, διασταύρωσης και μετάλλαξης, καθώς και μηχανισμούς επιδιόρθωσης και διαχείρισης περιορισμών. Η μεθοδολογία υλοποιήθηκε σε Python και οργανώθηκε σε αρθρωτή μορφή, ώστε να υποστηρίζει τόσο τη λειτουργικότητα όσο και τον πειραματικό έλεγχο.

Τέλος, η προτεινόμενη μέθοδος αξιολογήθηκε σε πραγματικά χρηματιστηριακά δεδομένα από καθιερωμένα benchmark datasets που αντιστοιχούν στους δείκτες Hang Seng, FTSE 100 και Nikkei 225. Πραγματοποιήθηκαν πειράματα για διαφορετικά επίπεδα cardinality και τα αποτελέσματα συγκρίθηκαν με άλλες μεθόδους επιλογής και κατασκευής χαρτοφυλακίου, ώστε να εξαχθούν συμπεράσματα για την αποτελεσματικότητα, τη σταθερότητα και τη συνολική χρησιμότητα της προτεινόμενης προσέγγισης.

6.2 Κύρια συμπεράσματα

Από την ανάλυση και τα πειραματικά αποτελέσματα της εργασίας προκύπτουν ορισμένα βασικά συμπεράσματα. Πρώτον, επιβεβαιώνεται ότι το πρόβλημα της αποτύπωσης δείκτη είναι ένα σύνθετο πρόβλημα βελτιστοποίησης, στο οποίο η δυσκολία δεν προκύπτει μόνο από την ανάγκη ελαχιστοποίησης του tracking error, αλλά και από την ταυτόχρονη ύπαρξη πολλών πρακτικών περιορισμών. Ο περιορισμός πληθικότητας, τα όρια συμμετοχής των τίτλων και η ανάγκη ρεαλιστικής σύνθεσης χαρτοφυλακίου καθιστούν το πρόβλημα υπολογιστικά απαιτητικό και περιορίζουν την αποτελεσματικότητα πιο απλών ή αμιγώς κλασικών προσεγγίσεων.

Δεύτερον, η εφαρμογή του Γενετικού Αλγορίθμου έδειξε ότι η εξελικτική αναζήτηση μπορεί να αντιμετωπίσει αποτελεσματικά τη μικτή φύση του προβλήματος, δηλαδή τον συνδυασμό επιλογής τίτλων και κατανομής βαρών. Ο αλγόριθμος κατάφερε να παράγει χαρτοφυλάκια με ικανοποιητικό έως πολύ καλό βαθμό αποτύπωσης, παρά το γεγονός ότι εργάστηκε υπό περιορισμένο αριθμό τίτλων. Αυτό σημαίνει ότι η προτεινόμενη μεθοδολογία μπορεί να οδηγήσει σε συμπαγή και διαχειρίσιμα χαρτοφυλάκια χωρίς να θυσιάζεται υπερβολικά η ποιότητα παρακολούθησης του δείκτη.

Τρίτον, τα πειραματικά αποτελέσματα έδειξαν ότι η προτεινόμενη προσέγγιση εμφανίζει ιδιαίτερα καλή απόδοση σε out-of-sample αξιολόγηση, στοιχείο που είναι κρίσιμο για τη χρηματοοικονομική αξία μιας μεθόδου. Ειδικά στον δείκτη Hang Seng, ο Γενετικός Αλγόριθμος ήταν η ισχυρότερη μέθοδος σε όλα τα εξεταζόμενα επίπεδα cardinality, ενώ

συνολικά παρουσίασε την καλύτερη μέση κατάταξη μεταξύ των μεθόδων που συγκρίθηκαν. Το εύρημα αυτό είναι σημαντικό, διότι δείχνει ότι η μέθοδος δεν είναι απλώς ικανή να προσαρμόζεται καλά στα in-sample δεδομένα, αλλά και να διατηρεί ικανοποιητική συμπεριφορά εκτός δείγματος.

Τέταρτον, επιβεβαιώθηκε ότι η απόδοση μιας μεθόδου αποτύπωσης δείκτη δεν εξαρτάται μόνο από τον ίδιο τον αλγόριθμο, αλλά και από τα χαρακτηριστικά του dataset και από το επίπεδο του περιορισμού πληθικότητας. Σε ορισμένα datasets και για συγκεκριμένες τιμές του K , άλλες μέθοδοι εμφάνισαν ανταγωνιστική ή ακόμη και καλύτερη απόδοση. Αυτό σημαίνει ότι το πρόβλημα δεν επιδέχεται μία απόλυτα καθολική λύση και ότι η επιλογή μεθοδολογίας πρέπει να λαμβάνει υπόψη τη φύση του δείκτη, τη δομή των αποδόσεων και το ζητούμενο επίπεδο συμπύκνωσης του χαρτοφυλακίου.

Συνολικά, το βασικό συμπέρασμα της εργασίας είναι ότι οι Γενετικοί Αλγόριθμοι αποτελούν μια ουσιαστικά αποτελεσματική και μεθοδολογικά κατάλληλη προσέγγιση για το πρόβλημα της αποτύπωσης δείκτη, ιδίως όταν αυτό διατυπώνεται με ρεαλιστικούς περιορισμούς και όταν η αναζήτηση πρέπει να κινηθεί μέσα σε πολύπλοκο και μη γραμμικό χώρο λύσεων.

6.3 Αξιολόγηση της προτεινόμενης μεθοδολογίας

Η αξιολόγηση της προτεινόμενης μεθοδολογίας μπορεί να γίνει σε τρία επίπεδα:

- Θεωρητικό
- Υπολογιστικό
- και Πειραματικό

Σε θεωρητικό επίπεδο, η μεθοδολογία κρίνεται επιτυχής, διότι είναι πλήρως ευθυγραμμισμένη με τη φύση του προβλήματος. Το index tracking, όπως διατυπώθηκε στην εργασία, δεν είναι ένα απλό συνεχές πρόβλημα βελτιστοποίησης, αλλά ένα μικτό πρόβλημα επιλογής και κατανομής, το οποίο συνοδεύεται από πλήθος περιορισμών. Οι Γενετικοί Αλγόριθμοι είναι ιδιαίτερα κατάλληλοι για τέτοιου τύπου προβλήματα, καθώς δεν απαιτούν αυστηρές παραδοχές κυρτότητας ή διαφορισιμότητας και μπορούν να διαχειριστούν αποτελεσματικά τόσο διακριτές όσο και συνεχείς μεταβλητές.

Σε υπολογιστικό επίπεδο, η προτεινόμενη υλοποίηση αποδείχθηκε λειτουργική, σταθερή και αρκετά ευέλικτη. Η αρθρωτή οργάνωση του κώδικα, η σαφής αναπαράσταση των λύσεων, η χρήση μηχανισμών επιδιόρθωσης και ο έλεγχος της εφικτότητας επέτρεψαν τη συστηματική εφαρμογή της μεθόδου σε πραγματικά datasets. Επιπλέον, η μεθοδολογία μπορεί να επεκταθεί σχετικά εύκολα ώστε να περιλάβει περισσότερους περιορισμούς ή εναλλακτικές εκδοχές της συνάρτησης καταλληλότητας.

Σε πειραματικό επίπεδο, η μεθοδολογία αξιολογήθηκε θετικά, καθώς εμφάνισε συνολικά πολύ καλή επίδοση και ήταν η ισχυρότερη μέθοδος στη μέση κατάταξη των συγκριτικών πειραμάτων. Το γεγονός ότι ήταν η καλύτερη σε οκτώ από τα εννέα εξεταζόμενα σενάρια δείχνει ότι διαθέτει ουσιαστική ικανότητα προσαρμογής σε διαφορετικές συνθήκες. Παράλληλα, το ότι δεν κυριάρχησε απολύτως σε όλα τα σενάρια αποτελεί ρεαλιστική ένδειξη ότι πρόκειται για ισχυρή αλλά όχι καθολικά ανίκητη μέθοδο, γεγονός που ενισχύει την αξιοπιστία της αξιολόγησης.

Ένα ακόμη θετικό στοιχείο της προτεινόμενης μεθοδολογίας είναι ότι παρήγαγε χαρτοφυλάκια με μικρό πλήθος τίτλων και παράλληλα ικανοποιητική ικανότητα παρακολούθησης. Αυτό είναι ιδιαίτερα σημαντικό από πρακτική άποψη, καθώς στη διαχείριση χαρτοφυλακίου δεν ενδιαφέρει μόνο η θεωρητική ελαχιστοποίηση του σφάλματος, αλλά και η δυνατότητα δημιουργίας χαρτοφυλακίων που να είναι οικονομικά και λειτουργικά διαχειρίσιμα.

Συνεπώς, η συνολική αξιολόγηση της προτεινόμενης μεθοδολογίας είναι θετική. Η μέθοδος κρίνεται επιτυχής ως προς τη συμβατότητά της με το πρόβλημα, ως προς την υλοποιησιμότητά της και ως προς τη συνολική ποιότητα των αποτελεσμάτων που παρήγαγε.

6.4 Σύγκριση με άλλες μεθόδους

Η σύγκριση της προτεινόμενης μεθοδολογίας με άλλες μεθόδους δείχνει ότι ο Γενετικός Αλγόριθμος είναι σαφώς ανταγωνιστικός και σε πολλές περιπτώσεις υπερέχει. Στα πειράματα που πραγματοποιήθηκαν, συγκρίθηκε με τρεις εναλλακτικές προσεγγίσεις: μια μέθοδο ίσων βαρών βασισμένη σε συσχέτιση, μια βελτιστοποιημένη εκδοχή της μεθόδου συσχέτισης και μια διαδικασία τυχαίας αναζήτησης.

Η σύγκριση ανέδειξε ότι η Corr-EW είναι μια απλή και χρήσιμη baseline μέθοδος, αλλά γενικά εμφανίζει περιορισμένη ευελιξία. Το βασικό της πλεονέκτημα είναι η απλότητα και η χαμηλή υπολογιστική απαίτηση. Ωστόσο, η αδυναμία της να βελτιστοποιεί ενεργά τα βάρη και να εξερευνά πιο σύνθετες δομές χαρτοφυλακίου την καθιστά λιγότερο αποδοτική στις περισσότερες περιπτώσεις. Παρ' όλα αυτά, σε ένα από τα εξεταζόμενα σενάρια εμφάνισε την καλύτερη out-of-sample επίδοση, γεγονός που δείχνει ότι ακόμη και απλές μέθοδοι μπορούν υπό προϋποθέσεις να είναι ανταγωνιστικές.

Η Corr-Opt εμφανίστηκε γενικά ως βελτιωμένη εκδοχή της Corr-EW, αλλά δεν κατόρθωσε να διατηρήσει σταθερά ανώτερη απόδοση απέναντι στον Γενετικό Αλγόριθμο. Αυτό υποδηλώνει ότι η βελτιστοποίηση πάνω σε προεπιλεγμένες ή ημι-καθοδηγούμενες δομές δεν είναι πάντοτε αρκετή για να ανταγωνιστεί μια πιο ολοκληρωμένη εξελικτική διαδικασία, ιδιαίτερα όταν το πρόβλημα έχει έντονα συνδυαστικό χαρακτήρα.

Η Random Search αποδείχθηκε αρκετά ισχυρή και μάλιστα σε ορισμένα σενάρια πλησίασε πολύ τον Γενετικό Αλγόριθμο. Το εύρημα αυτό είναι ενδιαφέρον, διότι δείχνει ότι σε τέτοια προβλήματα ο μεγάλος και σύνθετος χώρος λύσεων περιέχει πολλές ανταγωνιστικές λύσεις και ότι ακόμη και μια μη κατευθυνόμενη ή ημι-κατευθυνόμενη αναζήτηση μπορεί να εντοπίσει πολύ καλές συνθέσεις. Ωστόσο, η Random Search δεν προσφέρει τον ίδιο βαθμό συστηματικής βελτίωσης και προοδευτικής μάθησης που προσφέρει ένας Γενετικός Αλγόριθμος. Έτσι, παρότι σε μεμονωμένες περιπτώσεις ήταν εξαιρετικά ανταγωνιστική, δεν είχε την ίδια συνολική σταθερότητα σε όλα τα σενάρια.

Αντίθετα, ο Γενετικός Αλγόριθμος ήταν η μέθοδος με την καλύτερη συνολική επίδοση. Η υπεροχή του δεν βασίζεται μόνο στον αριθμό των σεναρίων που κέρδισε, αλλά και στη μέση κατάταξή του, η οποία ήταν η καλύτερη μεταξύ όλων των συγκρινόμενων μεθόδων. Το στοιχείο αυτό είναι ιδιαίτερα σημαντικό, διότι υποδηλώνει ότι η μέθοδος παρουσιάζει καλύτερη συνολική ισορροπία ανάμεσα στην ποιότητα των λύσεων και στη σταθερότητα της απόδοσης.

Συνεπώς, η σύγκριση με άλλες μεθόδους ενισχύει το βασικό επιχείρημα της εργασίας: οι Γενετικοί Αλγόριθμοι συνιστούν μια ιδιαίτερα κατάλληλη προσέγγιση για το πρόβλημα της αποτύπωσης δείκτη, όχι επειδή είναι πάντοτε οι απόλυτα καλύτεροι σε κάθε επιμέρους περίπτωση, αλλά επειδή εμφανίζουν την πιο ισχυρή και σταθερή συνολική συμπεριφορά.

6.5 Περιορισμοί της εργασίας

Παρά τα θετικά ευρήματα της εργασίας, υπάρχουν ορισμένοι περιορισμοί που πρέπει να αναγνωριστούν ρητά. Ο πρώτος περιορισμός αφορά το εύρος των datasets που χρησιμοποιήθηκαν. Παρότι τα benchmarks που αξιοποιήθηκαν είναι καθιερωμένα και ιδιαίτερα σημαντικά στη σχετική βιβλιογραφία, η πειραματική αξιολόγηση δεν κάλυψε το σύνολο όλων των πιθανών αγορών, δεικτών ή χρονικών περιόδων. Επομένως, τα συμπεράσματα της εργασίας είναι ισχυρά για τα συγκεκριμένα datasets, αλλά δεν πρέπει να γενικεύονται άκριτα σε κάθε πιθανό περιβάλλον αγοράς.

Δεύτερος περιορισμός αφορά το πλήθος των συγκριτικών μεθόδων. Η εργασία περιέλαβε ουσιαστικές και χρήσιμες συγκρίσεις, αλλά δεν επεκτάθηκε σε ευρύτερο σύνολο προηγμένων μεταερευνητικών ή ακριβών μεθόδων, όπως για παράδειγμα tabu search, simulated annealing, threshold accepting, particle swarm optimization ή προηγμένα μαθηματικά μοντέλα mixed-integer optimization στην ίδια πειραματική βάση. Συνεπώς, η συγκριτική αξιολόγηση είναι επαρκής για να δείξει τη χρησιμότητα της προτεινόμενης μεθόδου, αλλά όχι εξαντλητική ως προς όλο το φάσμα της διεθνούς βιβλιογραφίας.

Τρίτος περιορισμός αφορά την παραμετροποίηση του Γενετικού Αλγορίθμου. Παρότι επιλέχθηκαν λογικές και λειτουργικές τιμές παραμέτρων, η απόδοση ενός τέτοιου αλγορίθμου επηρεάζεται σε σημαντικό βαθμό από το μέγεθος του πληθυσμού, τις πιθανότητες διασταύρωσης και μετάλλαξης, τον βαθμό ελιτισμού και το κριτήριο τερματισμού. Επομένως, τα αποτελέσματα της εργασίας αντανακλούν και τις συγκεκριμένες επιλογές παραμετροποίησης που έγιναν, και δεν αποκλείεται άλλες ρυθμίσεις να οδηγούσαν σε ακόμη καλύτερη ή διαφορετική συμπεριφορά.

Τέταρτος περιορισμός αφορά τη μορφή του κόστους συναλλαγών και γενικότερα τη ρεαλιστική πολυπλοκότητα του μοντέλου. Στην παρούσα εργασία το κόστος συναλλαγών αντιμετωπίστηκε σε επίπεδο που επιτρέπει την ενσωμάτωση στο μοντέλο, αλλά δεν αναπτύχθηκε σε πλήρως ρεαλιστική πολυπαραγοντική μορφή που να περιλαμβάνει bid-ask spread, market impact, ρευστότητα και δυναμικά στοιχεία εκτέλεσης. Κατά συνέπεια, το μοντέλο παραμένει ρεαλιστικό αλλά όχι πλήρως εξαντλητικό ως προς όλες τις διαστάσεις της πραγματικής αγοράς.

Πέμπτος περιορισμός αφορά την εστίαση σε μία κύρια μεταερευνητική προσέγγιση. Η εργασία σχεδιάστηκε γύρω από τους Γενετικούς Αλγορίθμους, κάτι που είναι πλήρως συμβατό με τον στόχο της. Ωστόσο, αυτό σημαίνει ότι δεν διερευνήθηκαν σε αντίστοιχο βάθος υβριδικά σχήματα ή πιο εξειδικευμένες εξελικτικές παραλλαγές, οι οποίες πιθανόν να προσέφεραν επιπλέον βελτιώσεις.

Συνολικά, οι περιορισμοί αυτοί δεν αναιρούν την αξία της εργασίας, αλλά οριοθετούν το πλαίσιο μέσα στο οποίο πρέπει να διαβαστούν και να ερμηνευθούν τα αποτελέσματά της.

6.6 Προτάσεις για μελλοντική έρευνα

Με βάση τα παραπάνω, προκύπτουν αρκετές ενδιαφέρουσες κατευθύνσεις για μελλοντική έρευνα. Μία πρώτη κατεύθυνση είναι η επέκταση της πειραματικής αξιολόγησης σε περισσότερα datasets, περισσότερους δείκτες και διαφορετικές αγορές. Η εφαρμογή της προτεινόμενης μεθόδου σε ευρύτερο φάσμα πραγματικών δεδομένων θα επέτρεπε την εξαγωγή πιο γενικευμένων συμπερασμάτων για τη σταθερότητα και την προσαρμοστικότητα της.

Μία δεύτερη κατεύθυνση είναι η συστηματική σύγκριση με μεγαλύτερο πλήθος ανταγωνιστικών μεθόδων. Θα ήταν ιδιαίτερα ενδιαφέρον να συγκριθεί ο προτεινόμενος Γενετικός Αλγόριθμος με πιο σύνθετες μεταερευτικές μεθόδους, όπως tabu search, simulated annealing, threshold accepting, particle swarm optimization ή υβριδικά evolutionary frameworks. Με αυτόν τον τρόπο θα μπορούσε να αποτυπωθεί ακόμη πιο καθαρά η σχετική θέση της προτεινόμενης μεθοδολογίας μέσα στο ευρύτερο πεδίο των σύγχρονων τεχνικών index tracking.

Μια τρίτη κατεύθυνση αφορά την ανάπτυξη υβριδικών προσεγγίσεων. Για παράδειγμα, θα μπορούσε να μελετηθεί ένας υβριδικός αλγόριθμος στον οποίο ο Γενετικός Αλγόριθμος αναλαμβάνει την επιλογή τίτλων και μια συμπληρωματική μέθοδος μαθηματικής βελτιστοποίησης ή τοπικής αναζήτησης αναλαμβάνει τη λεπτομερή προσαρμογή των βαρών. Τέτοιες αρχιτεκτονικές ενδέχεται να συνδυάζουν τα πλεονεκτήματα της εξελικτικής εξερεύνησης με την ακρίβεια πιο εξειδικευμένων διαδικασιών βελτιστοποίησης.

Μια τέταρτη κατεύθυνση είναι η πιο ρεαλιστική και εκτενής ενσωμάτωση του κόστους συναλλαγών και γενικότερα δυναμικών περιορισμών. Μελλοντική έρευνα θα μπορούσε να εξετάσει πιο σύνθετα μοντέλα κόστους, περιορισμούς ρευστότητας, turnover constraints και δυναμικά σχήματα αναπροσαρμογής χαρτοφυλακίου, ώστε το πρόβλημα να προσεγγίζεται ακόμη πιο κοντά στις συνθήκες της πραγματικής επενδυτικής διαχείρισης.

Μια πέμπτη ενδιαφέρουσα κατεύθυνση είναι η πολυκριτηριακή επέκταση του προβλήματος. Στην παρούσα εργασία ο βασικός στόχος ήταν η ελαχιστοποίηση του tracking error, αλλά θα μπορούσε να εξεταστεί ταυτόχρονα και η βελτιστοποίηση άλλων μεγεθών, όπως το turnover, η διαφοροποίηση, η έκθεση σε συγκεκριμένους τομείς ή ακόμα και δείκτες κινδύνου πέρα από τη μέση τετραγωνική απόκλιση. Σε αυτή την περίπτωση, πολυκριτηριακοί Γενετικοί Αλγόριθμοι θα μπορούσαν να προσφέρουν ακόμη πιο ευέλικτο πλαίσιο επίλυσης.

Τέλος, ιδιαίτερα σημαντική κατεύθυνση είναι και η βελτίωση της υπολογιστικής αποδοτικότητας. Η χρήση παραλληλοποίησης, πιο αποδοτικών δομών δεδομένων ή τεχνικών επιτάχυνσης της αξιολόγησης του πληθυσμού θα μπορούσε να επιτρέψει την εφαρμογή της μεθόδου σε ακόμη μεγαλύτερα επενδυτικά σύνολα ή σε πιο απαιτητικές εκδοχές του προβλήματος.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.
- [2] Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3), 247–257.
- [3] Perold, A. F., & Sharpe, W. F. (1988). Dynamic strategies for asset allocation. *Financial Analysts Journal*, 44(1), 16–27.
- [4] Schaerf, A. (2002). Local search techniques for constrained portfolio selection problems. *Computational Economics*, 20(3), 177–190.
- [5] 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.
- [6] Canakgoz, N. A., & Beasley, J. E. (2009). Mixed-integer programming approaches for index tracking and enhanced indexation. *European Journal of Operational Research*, 196(1), 384–399.
- [7] Steuer, R. E., Qi, Y., & Hirschberger, M. (2005). Multiple objectives in portfolio selection. *Journal of Financial Decision Making*, 1, 11–26.
- [8] Meade, N., & Salkin, G. R. (1989). Index funds—Construction and performance measurement. *Journal of the Operational Research Society*, 40(10), 871–879.
- [9] Derigs, U., & Nickel, N.-H. (2003). Meta-heuristic based decision support for portfolio optimization with a case study on tracking error minimization in passive portfolio management. *OR Spectrum*, 25(3), 345–378.
- [10] Maringer, D., & Oyewumi, O. (2007). Index tracking with constrained portfolios. *Intelligent Systems in Accounting, Finance and Management*, 15(1–2), 57–71.
- [11] Beasley, D., Bull, D. R., & Martin, R. R. (1993). An overview of genetic algorithms: Part 1, fundamentals. *University Computing*, 15(2), 56–69.
- [12] Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, 4(2), 65–85.
- [13] Chang, T.-J., Meade, N., Beasley, J. E., & Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13), 1271–1302.
- [14] 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.
- [15] Lin, C.-C., & Liu, Y.-T. (2008). Genetic algorithms for portfolio selection problems with minimum transaction lots. *European Journal of Operational Research*, 185(1), 393–404.
- [16] Chang, T.-J., Yang, S.-C., & Chang, K.-J. (2009). Portfolio optimization problems in different risk measures using genetic algorithm. *Expert Systems with Applications*, 36(7), 10529–10537.
- [17] Anagnostopoulos, K. P., & Mamanis, G. (2011). The mean–variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, 38(11), 14208–14217.
- [18] Sharpe, W. F. (1991). The arithmetic of active management. *Financial Analysts Journal*, 47(1), 7–9.
- [19] Goodwin, T. H. (1998). The information ratio. *Financial Analysts Journal*, 54(4), 34–43.
- [20] Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, 116(1), 23–45.

- [21] Arnott, R. D., Hsu, J., & Moore, P. (2005). Fundamental indexation. *Financial Analysts Journal*, 61(2), 83–99.
- [22] Perold, A. F. (2007). Fundamentally flawed indexing. *Financial Analysts Journal*, 63(6), 31–37.
- [23] Walkshäusl, C., & Lobe, S. (2010). Fundamental indexing around the world. *Review of Financial Economics*, 19(3), 117–127.
- [24] Brogaard, J., Heath, D., & Huang, D. (2026). ETF sampling and index arbitrage. *Journal of Financial and Quantitative Analysis*, 61(2), 547–579.
- [25] Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance*, 52(3), 1035–1058.
- [26] Chen, Z., & Knez, P. J. (1996). Portfolio performance measurement: Theory and applications. *The Review of Financial Studies*, 9(2), 511–555.
- [27] Jiang, P., Fu, Z., Huang, K., & Yang, H. (2021). Follow the leader: Index tracking with factor models. *European Journal of Operational Research*, 295(3), 1125–1137.
- [28] Ni, H., & Yin, H. (2013). Stock index tracking by Pareto efficient genetic algorithm. *Applied Soft Computing*, 13(12), 4519–4535.
- [29] Wu, D., Wang, Y., Oliveira, F. S., & Birge, J. R. (2019). An enhanced decision support approach for learning and portfolio optimization with transaction costs and market liquidity. *Omega*, 87, 65–81.
- [30] Coleman, T. F., Li, Y., & Henniger, J. (2006). Minimizing tracking error while restricting the number of assets. *The Journal of Risk*, 8(4), 33–55.
- [31] Sant’Anna, L. R., Filomena, T. P., Guedes, P. C., & Borenstein, D. (2020). Solving the index tracking problem based on a convex similarity measure between portfolios. *Operations Research Perspectives*, 7, 100173.
- [32] Yuen, M. C., Zhang, J., & Zhang, C. (2022). A metaheuristic-based framework for index tracking with practical constraints. *Complex & Intelligent Systems*, 8, 3379–3404.
- [33] Strub, O., & Baumann, P. (2018). Optimal construction and rebalancing of index-tracking portfolios. *European Journal of Operational Research*, 264(1), 370–387.
- [34] de Paulo, W. L., Leite, A. L. S., & Camargo, A. J. (2016). Enhanced index tracking optimal portfolio selection. *Finance Research Letters*, 16, 194–200.
- [35] Chavez-Bedoya, L. A., & Alonso-Becerra, C. A. (2014). Index tracking and enhanced indexation using a parametric approach. *Investment Management and Financial Innovations*, 11(2), 270–278.
- [36] Chen, Q., Wang, Y., & Li, X. (2022). A kind of new time-weighted nonnegative lasso index tracking model under different market constraints. *Applied Mathematics and Computation*, 431, 127330.
- [37] Roll, R. (1992). A mean/variance analysis of tracking error. *The Journal of Portfolio Management*, 18(4), 13–22.
- [38] Rudolf, M., Wolter, H.-J., & Zimmermann, H. (1999). A linear model for tracking error minimization. *Journal of Banking & Finance*, 23(1), 85–103.
- [39] Barro, D., & Canestrelli, E. (2009). Tracking error: A multistage portfolio model. *Annals of Operations Research*, 165(1), 47–66.
- [40] Huang, J., Li, Y., & Yao, H. (2018). Index tracking model, downside risk and non-parametric kernel estimation. *Journal of Economic Dynamics and Control*, 92, 103–128.
- [41] Guastaroba, G., Mansini, R., Ogryczak, W., & Speranza, M. G. (2020). Enhanced index tracking with CVaR-based ratio measures. *Annals of Operations Research*, 292(2), 883–931.

- [42] Mezali, H., & Beasley, J. E. (2014). Index tracking with fixed and variable transaction costs. *Optimization Letters*, 8(1), 61–80.
- [43] Lobo, M. S., Fazel, M., & Boyd, S. (2007). Portfolio optimization with linear and fixed transaction costs. *Annals of Operations Research*, 152(1), 341–365.
- [44] Rubio-García, Á., Ruiz-Torrubiano, R., & Suárez, A. (2024). Accurate solution of the index tracking problem with a small number of assets and transaction costs. *European Journal of Operational Research*.
- [45] García, F., Guijarro, F., & Oliver, J. (2018). Index tracking optimization with cardinality constraint: A performance comparison of genetic algorithms and tabu search heuristics. *Neural Computing and Applications*, 30(8), 2625–2642.
- [46] Ruiz-Torrubiano, R., & Suárez, A. (2009). A hybrid optimization approach to index tracking. *Annals of Operations Research*, 166(1), 57–71.
- [47] Zheng, Y., Chen, B., Hospedales, T. M., & Yang, Y. (2020). Index Tracking with Cardinality Constraints: A Stochastic Neural Networks Approach. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01), 1242–1249.
- [48] Chiam, S. C., Tan, K. C., Al Mamun, A., & Goh, C. K. (2013). Dynamic index tracking via multi-objective evolutionary algorithm. *Applied Soft Computing*, 13(7), 3392–3408.
- [49] Frino, A., Lepone, A., & Nolin, J. (2005). The index tracking strategies of passive and enhanced index equity funds. *Australian Journal of Management*, 30(1), 23–55.
- [50] Ammann, M., & Zimmermann, H. (2001). Tracking Error and Tactical Asset Allocation. *Financial Analysts Journal*, 57(2), 32–43.
- [51] Takeda, A., Niranjan, M., Gotoh, J.-y., & Kawahara, Y. (2013). Simultaneous pursuit of out-of-sample performance and sparsity in index tracking portfolios. *Computational Management Science*, 10(1), 21–49.
- [52] Bailey, D. H., Borwein, J., López de Prado, M., & Zhu, Q. J. (2014). The probability of backtest overfitting. *Journal of Computational Finance*.
- [53] Guastaroba, G., Mansini, R., & Speranza, M. G. (2012). Kernel Search: An application to the index tracking problem. *European Journal of Operational Research*, 217(1), 54–68.
- [54] Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies*, 4(2), 315–342.
- [55] Michaud, R. O. (1989). The Markowitz optimization enigma: Is ‘optimized’ optimal? *Financial Analysts Journal*, 45(1), 31–42.
- [56] Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, 37(5), 519–531.
- [57] Lai, T. L., Xing, H., & Chen, Z. (2011). Mean-variance portfolio optimization when means and covariances are unknown. *The Annals of Applied Statistics*, 5(2A), 798–823.
- [58] Bonami, P., & Lejeune, M. A. (2009). An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Operations Research*, 57(3), 650–670.
- [59] Scozzari, A., Tardella, F., & Paterlini, S. (2013). Exact and heuristic approaches for the index tracking problem with UCITS constraints. *Annals of Operations Research*, 205(1), 235–250.
- [60] De Jong, K. A. (1988). Learning with genetic algorithms: An overview. *Machine Learning*, 3(2–3), 121–138.
- [61] McCall, J. (2005). Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied Mathematics*, 184(1), 205–222.
- [62] Lin, C.-C., & Liu, Y.-T. (2008). Genetic algorithms for portfolio selection problems with minimum transaction lots. *European Journal of Operational Research*, 185(1), 393–

- [63] Beyer, H.-G., & Schwefel, H.-P. (2002). Evolution strategies – A comprehensive introduction. *Natural Computing*, 1(1), 3–52.
- [64] Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2), 159–195.
- [65] Li, Z., Deng, J., & Zhang, C. (2020). Evolution strategies for continuous optimization: A survey of the state-of-the-art. *Swarm and Evolutionary Computation*, 56, 100694.
- [66] Beyer, H.-G., Finck, S., & Breuer, T. (2014). On the design of an evolution strategy for scenario-based multi-period portfolio optimization under transaction costs. *Swarm and Evolutionary Computation*, 17, 68–78.
- [67] Kirkpatrick, S., Gelatt, C. D., Jr., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598), 671–680.
- [68] Eglese, R. W. (1990). Simulated annealing: A tool for operational research. *European Journal of Operational Research*, 46(3), 271–281.
- [69] Crama, Y., Schyns, M., & others. (2003). Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research*.
- [70] Derigs, U., & Nickel, N. H. (2004). On a local-search heuristic for a class of tracking error minimization problems in portfolio management. *Annals of Operations Research*.
- [71] Rubio-García, Á., Fernández-Lorenzo, S., García-Ripoll, J. J., & Porras, D. (2024). Accurate solution of the index tracking problem with a hybrid simulated annealing algorithm. *Physica A: Statistical Mechanics and its Applications*, 639, 129637.
- [72] Glover, F. (1989). Tabu Search—Part I. *ORSA Journal on Computing*, 1(3), 190–206.
- [73] Glover, F. (1990). Tabu Search—Part II. *ORSA Journal on Computing*, 2(1), 4–32.
- [74] Dueck, G., & Scheuer, T. (1990). Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *Journal of Computational Physics*, 90(1), 161–175.
- [75] Eglese, R. W. (1990). Simulated annealing: A tool for operational research. *European Journal of Operational Research*, 46(3), 271–281.
- [76] Gilli, M., Hysi, H., & Këllezi, E. (2010). Distributed optimisation of a portfolio's Omega. *Parallel Computing*, 36(2–3), 172–187.
- [77] Erwin, K. D., Muhammad, Y., Mayefis, D., et al. (2023). Meta-heuristics for portfolio optimization: A systematic literature review. *Soft Computing*.
- [78] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks* (Vol. 4, pp. 1942–1948). IEEE.
- [79] Cheng, S., Shi, Y., Qin, Q., Bai, R., & Liu, Y. (2018). A survey of particle swarm optimization. *Complex & Intelligent Systems*, 4, 227–239.
- [80] Cura, T. (2009). Particle swarm optimization approach to portfolio optimization. *Nonlinear Analysis: Real World Applications*, 10(4), 2396–2406.
- [81] Zhu, H., Wang, Y., Wang, K., & Chen, Y. (2011). Particle Swarm Optimization (PSO) for the constrained portfolio optimization problem. *Expert Systems with Applications*, 38(8), 10161–10169.
- [82] Corazza, M., Fasano, G., & Gusso, R. (2013). Particle Swarm Optimization with non-smooth penalty reformulation for a complex portfolio selection problem. *Applied Mathematics and Computation*, 224, 611–624.
- [83] Zhu, H., Wang, Y., Wang, K., & Chen, Y. (2010). A Particle Swarm Optimization Heuristic for the Index Tracking Problem. In *Advanced Data Mining and Applications* (Lecture Notes in Computer Science, Vol. 6440, pp. 79–91). Springer
- [84] Rudd, A. (1980). Optimal selection of passive portfolios. *Financial Management*,

9(1), 57–66.

[85] Meade, N., & Salkin, G. R. (1990). Developing and maintaining an equity index fund. *The Journal of Portfolio Management*, 16(4).

[86] 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.

[87] Canakgoz, N. A., & Beasley, J. E. (2009). Mixed-integer programming approaches for index tracking and enhanced indexation. *European Journal of Operational Research*, 196(1), 384–399.

[88] Gilli, M., & K llezi, E. (2001). The threshold accepting heuristic for index tracking. Conference/working paper version widely cited in the index tracking literature.

[89] 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.

[90] Ni, H., & Wang, Y. (2013). Stock index tracking by Pareto efficient genetic algorithm. *Applied Soft Computing*, 13(12), 4519–4535.

[91] Zheng, Y., Zhu, Y., & Li, X. (2020). Index tracking with cardinality constraints: A stochastic neural networks approach. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01), 227–234.

[92] Moeini, M. (2022). Solving the index tracking problem: A continuous optimization approach. *Central European Journal of Operations Research*, 30(2), 807–835.

[93] Yuen, M. C., Ng, A. W. H., Leung, S. C. H., & Che, Y. (2022). A metaheuristic-based framework for index tracking with practical constraints. *Complex & Intelligent Systems*.