11th International Conference on Intelligent Control and Information Processing
Dali, China, December 3-7, 2021

Index Tracking Based on Dynamic Time Warping and Constrained k-medoids Clustering

Ran Zhang
Department of Computer Science
City University of Hong Kong
Kowloon, Hong Kong
rzhang226-c@my.cityu.edu.hk

Hongzong Li
Department of Computer Science
City University of Hong Kong
Kowloon, Hong Kong
hongzli2-c@my.cityu.edu.hk

Jun Wang
Department of Computer Science
& School of Data Science
City University of Hong Kong
Kowloon, Hong Kong
jwang.cs@cityu.edu.hk

2021 11th International Conference on Intelligent Control and Information Processing (ICICIP) | 978-1-6654-2515-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICIP53388.2021.9642192

***Abstract*—Index tracking is a passive investment strategy by replicating a financial market index using its constituents. In this paper, index tracking is addressed based on k-medoids clustering. k-medoids clustering is formulated as a valuation-constrained k-median problem to cluster index constituents. The dissimilarity coefficients among stocks are measured by using dynamic time warping. Experimental results of index tracking on four major indices are elaborated to demonstrate that the tracking performance of the proposed method with dynamic time warping is superior to that with Pearson correlation coefficients.**

***Index Terms*—Index tracking, k-medoids clustering, dynamic time warping**

I. Introduction

As a passive investing strategy, index tracking is carried out by selecting a subset of index constituents via the minimization of tracking errors of the returns between the portfolio and a target index. Index tracking has a desirable feature of infrequent transactions resulting in relatively stable performance. Some studies further indicate that merely following broad market indices is an ideal investment strategy in stock markets [1], [2].

Index tracking with all index constituents (also known as full replication) is usually impracticable. It is infeasible in some markets like Hang Seng, where shares are traded in lots. Besides, it is not advisable because any change of index constituents over an investment horizon requires the rebalancing of the portfolio. In view of the shortcomings, index-tracking with selected index constituents (also known as partial replication or index sampling) is preferable and commonly practised [3].

Existing studies mainly concentrate on cardinality-constrained index-tracking that selects a given number of subsets of stocks bounded by a given cardinality. There are two issues in cardinality-constrained index tracking. One is the selection of constituents, and the other is their corresponding proportions. Numerous approaches for cardinality-constrained index tracking have been studied in related literature [4], [5] such as regression and clustering [6], [7]. Cluster-based cardinality-constrained index tracking has a wide application in industry practice for the advantage of reducing turnover and transaction costs and having comparatively high liquidity [8]. A stock selection approach based on clustering is developed in [7] to track the S&P 500 index. However, those methods are validated only on a specified index, not broadly applicable to other indices. More recently, an index tracking algorithm that optimizes both diversity and sparsity is proposed in [9]. A quadratic unconstrained binary optimization (QUBO) problem for k-medoids clustering [10] formulates the cluster-based index tracking with a trade-off of centrality and sparsity [8].

Dynamic time warping (DTW) is an effective algorithm for measuring the similarity between two temporal sequences with widespread applications in financial data analysis. For example, it is found that DTW captures the pattern of selecting a portfolio of huge stocks to reduce the risks by hierarchical clustering [11]. Entropic DTW kernels for stock price analysis are developed in [12]. DTW is also used to extract the representative price fluctuation patterns for cluster-based stock price prediction [13].

In this paper, the cardinality-constrained index tracking is addressed based on k-medoids clustering. Constrained k-median clustering formulations are proposed. DTW is used to measure the similarities of stocks over periods. To capture the momentum of stock movements, a moving window with a forgetting factor is used. Index constituents are clustered into k clusters via k-medoids clustering and the medoids are used as the exemplars for index tracking.

The remainder of this paper is organized as follows. Section II provides preliminary information. Section III states the problem formulations. Section IV reports the experimental results of cluster-based index tracking on four stock market datasets. Section VI concludes the paper.

II. Preliminaries

A. k-medoids Clustering

k-medoids clustering represents a class of clustering algorithms that group data by minimizing the within-cluster dissimilarities with k data as the cluster medoids. Among

This work was supported in part by the Research Grants Council of the Hong Kong Special Administrative Region of China under General Research Fund Grant 11202019; and in part by the Laboratory for AI-Powered Financial Technologies.

several existing formulations, k-medoids clustering can be formulated as a k-median problem [14], [15] as follows:

$$\min \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}x_{ij}, \tag{1a}$$

$$\text{s.t.} \sum_{j=1}^{n} x_{ij} = 1,\ i = 1,2,...,n, \tag{1b}$$

$$\sum_{j=1}^{n} x_{jj} = k, \tag{1c}$$

$$x_{ij} < x_{jj},\ i,j = 1,2,...,n, \tag{1d}$$

$$x_{ij} \in \{0,1\},\ i,j = 1,2,...,n, \tag{1e}$$

where d_{ij} denotes the dissimilarity coefficient between datum i and j, x_{ij} is the binary decision variable denoted as $x_{ij} = 1$ if datum i is assigned to the cluster in which datum j is the cluster medoid or $x_{ij} = 0$ otherwise, and n denotes the amount of data. Objective function (1a) is the sum of intra-cluster dissimilarities. Its value depends on which data are selected as cluster medians, and their corresponding dissimilarity coefficients [15]. Constraints (1b) ensure that each datum belongs to exactly one cluster. Constraint (1c) specifies the required number of clusters. Constraints (1d) ensure that datum i belongs to cluster j only if datum j is the cluster medoid. Constraints in (1e) enforce the solution to be binary.

k-medoids clustering can be also formulated as a QUBO problem [10] as follows:

$$\min_{z} \left\{\beta z^T De - \alpha\frac{1}{2}z^T Dz + \theta(z^T e - k)^2\right\},$$
$$\text{s.t. } z = [z_1, z_2, z_3, \dots, z_n],$$
$$z_j \in \{0,1\},\ \forall j \in V,$$

where D denotes dissimilarity coefficient matrix, k is the number of exemplars to seek, z_j is the binary decision variable denoted as $z_j = 1$ if datum j is an exemplar or $z_j = 0$ otherwise, the trade-off parameter α weigh the k nodes that are the most dispersed from other clusters, parameter β weigh the k nodes that are the most central intra-cluster, θ is the penalty coefficient to enforce feasibility, and $e = [1,1,1,\dots,1] \in \mathbb{R}^n$. The last term is the penalty term that pushes the solution z^* to k non-zero entries. The QUBO problem is to find an optimal combination of dispersed and central data using trade-off parameters.

B. Dissimilarity coefficients

There are numerous measures of similarity or dissimilarity coefficients for financial data processing; e.g., Pearson correlation coefficients (PCC) and DTW-based dissimilarity coefficients.

1) Pearson Correlation Coefficients: Given a pair of random variables (X, Y), the Pearson correlation coefficient ρ is defined as the covariance of (X, Y) divided by the product of their standard deviations [16] as follows:

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y},$$

where σ_X and σ_Y are the standard deviations of X and Y, respectively; cov is the covariance defined as

$$cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)],$$

E is the expectation operator; μ_X and μ_Y are the means of X and Y, respectively. Pearson correlation coefficient $\rho_{XY} \in [-1,1]$, where the zero value of ρ_{XY} means that there is no linear correlation between X and Y. A dissimilarity coefficient d_{XY} is defined as $d_{XY} = 1 - \rho_{XY}$.

2) Dynamic Time Warping: DTW is a well-known method for measuring similarities of two sequences time with optimal alignment based on the Levenshtein distance (also called the edit distance) [17], [18]. The distance of the optimal alignment is recursively defined in a dynamic programming framework by:

$$d(A_i, B_j) = \delta(a_i, b_j) + \min\left\{\begin{array}{l} d(A_{i-1}, B_{j-1}), \\ d(A_i, B_{j-1}), \\ d(A_{i-1}, B_j) \end{array}\right\},$$

where A_i is the subsequence $(a_1, \dots,\ a_T)$, B_j is the subsequence $(b_1, \dots,\ b_T)$, and δ is a distance between coordinates of sequences. The DTW dissimilarity coefficient of two time series is given by $d(A_{|A|},\ B_{|B|}) = d(A_T,\ B_T)$.

III. Problem Formulations

This section describes the moving window settings and the constrained k-median problem formulation.

A. Moving Window Settings

To make the most use of available historical data, a moving window is used instead of simply halving the dataset. A forgetting factor is also used to filter the data with the moving window, since historical stock data are time-sensitive, more recent data should be given greater weight. Therefore, an attenuating recursive window is used for setting training datasets and test datasets, as illustrated in Fig. 1.

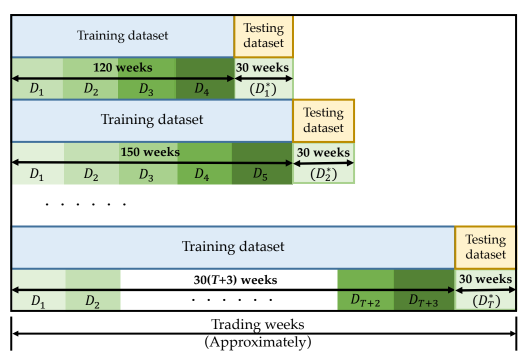

Fig. 1. Partition of training dataset and test dataset.

For period τ, the dissimilarity coefficient of each stock pair is calculated by DTW and forms the DTW dissimilarity matrix D_τ. Each D_τ is endowed with a forgetting factor. The total dissimilarity matrix D_τ^* of the training sets at period τ is defined as:

$$D_\tau^* = D_1\gamma^{\tau+2} + D_2\gamma^{\tau+1} + \ldots + D_{\tau+3}\gamma^0, \quad \tau = 1, 2, \ldots, T$$

where D_τ denotes the dissimilarity matrix of period τ, T denotes the total period amount, and the forgetting factor $\gamma \in [0, 1]$ is a control parameter in the attenuating recursive window. If $\gamma = 0$, it is a rolling window that only uses the last 30-weeks historical data for exemplar selection. If $\gamma = 1$, it is a recursive window that uses all historical data for exemplar selection.

B. Constrained k-median problem formulation

As mentioned in Subsection II, cluster medoids can be determined by solving the k-median or QUBO problems. In the context of index tracking, these medoids can be used as the stock exemplars to represent the clustered index constituents for tracking the index. The weight of exemplar j in a tracking portfolio is defined as [1]:

$$w_j = \frac{\sum_{i=1}^{n} P_i x_{ij}}{\sum_{j=1}^{n}\sum_{i=1}^{n} P_i x_{ij}}, \; i, j = 1, 2, ..., n,$$

where P_i denotes the market value of stock i, $\sum_{i=1}^{n} P_i x_{ij}$ is the total market value of the clustered stocks represented by exemplar j on the the last day in the training set, and $\sum_{j=1}^{n}\sum_{i=1}^{n} P_i x_{ij}$ is the total value of the whole market on the the last day in the training set.

A limitation of index tracking based on the k-median problem is that there is no control of the unsystematic risk of the portfolio. The unsystematic risk is the difference between total portfolio variation and systematic variation. In other words, it is the variation due to attributes of individual stocks such as financial fraud or crucial executive turnover. The k-median problem does not consider a potential imbalance of market valuation among clusters. If the market value falls heavily in some clusters, the stock exemplars within this cluster are endowed with high weights and the unsystematic risk of the whole portfolio is greatly increased.

To limit the market value of each cluster, a valuation constraint is defined for each cluster:

$$\sum_{i=1}^{n} w_i x_{ij} \leq \frac{1}{k - k'} x_{jj}, j = 1, 2, \ldots, k, \tag{2}$$

where w_i denotes the weight of exemplar stock i, $\frac{1}{k-k'}$ denotes the upper bound weight for cluster containing stock j. This constraint is effective if and only if x_{jj} is a cluster median. Parameter $k' \in (0, k)$, and k' can be adjusted flexibly according to the datasets. If k' is too small, the problem would be infeasible. If k' is too large, constraint (2) will lose its effect. The dissimilarity d_{ij} in the objective function (1a) is measured by two measures: DTW, the proposed measure in this paper, and PCC, the ordinary measure of dissimilarity. So far, the k exemplar stocks and their corresponding cluster members have been identified. With constraints (2), the problem is named constrained k-median problem. Above all, the exemplars and their corresponding proportions are determined.

IV. Experimental Results

A. Benchmark Datasets

The experiments are based on the public datasets in Beasley's OR-Library [6]. These curated datasets contain 290 weekly price observations of constituent stocks. The stocks that are not index constituents throughout the investment horizon are removed. The market indices and their respective number of constituents after data cleaning are shown in Table I.

TABLE I
Number of constituents in benchmark datasets

Index	Number of stocks
Hang Seng	31
S&P 100	98
FTSE 100	89
DAX 100	85

B. Setups

To set up multiple experiments for verification, a dataset is divided into multiple subsets based on two considerations. First, the length of the input data should be fixed to ensure the fairness of the data. Second, more datasets should be divided to verify the applicability of the proposed method. If a dataset is split into one training set and one test set, the experiment can only be conduct once which increases the randomness of the results. Based on these reasons, the weekly (daily) price data of a constituent stock is divided into 30-week sequences as periods. Then, to cluster the index constituents with similar temporal behaviors, the 30-week price sequences are transformed to cumulative returns for measuring dissimilarities. The portfolio is supposed to be bought in the first week of the test period and held until the end.

The cardinality in the experiments is ten; i.e., $p = 10$, the same as those of the data providers [6]. For experiments of the constrained k-median problem, the parameter p' is set based on previous experience which is initially set to be three to reach a feasible solution and constrain the valuation on all datasets. An exception happens for Hang Seng Index which has 31 constituents only and it is infeasible if $p' = 3$, so p' is chosen to be five to obtain feasible solutions.

The experiments are based on two tools. One is tslearn [19], a python package used for DTW. The other is Gurobi, a commercial mathematical optimization solver.

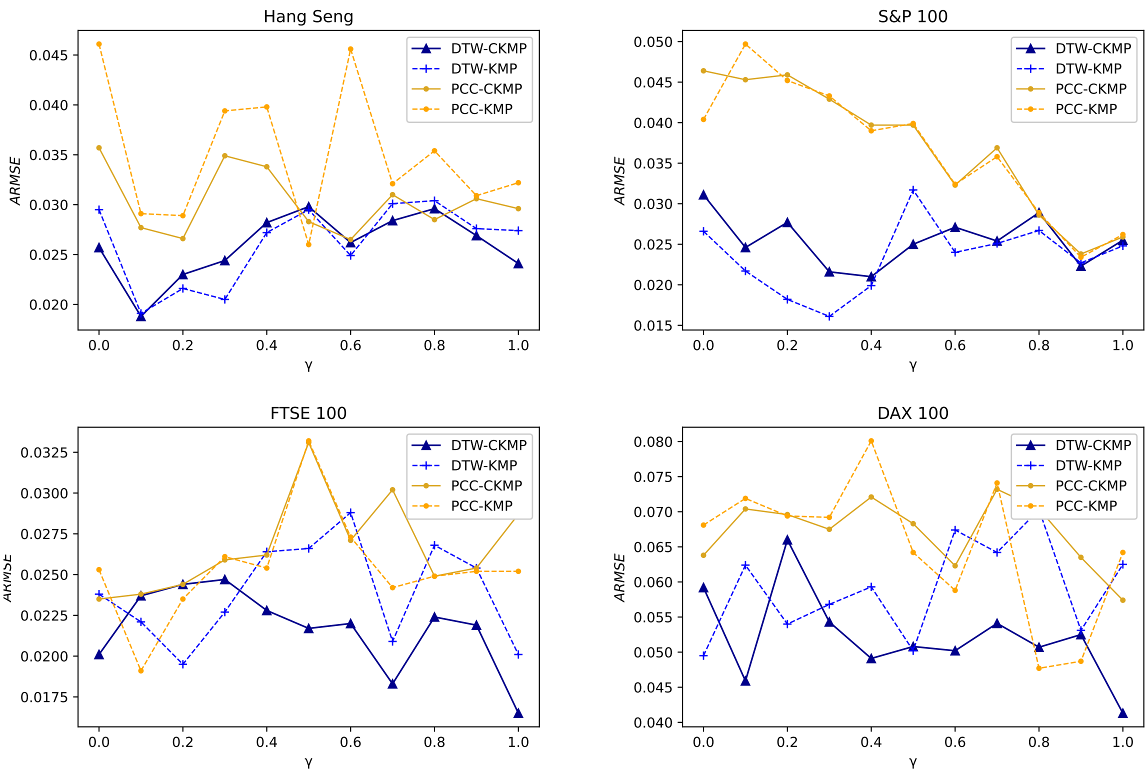

Fig. 2. Performance in terms of $ARMSE$ with different values of forgetting factor γ on four benchmark datasets.

TABLE II
TRACKING ERRORS ON HANG SENG

γ	KMP			CKMP		
	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ
0.0	0.0295	0.0461	0.0166	0.0257	0.0357	0.0100
0.1	0.0191	0.0291	0.0100	0.0188	0.0277	0.0088
0.2	0.0216	0.0289	0.0074	0.0230	0.0266	0.0036
0.3	0.0205	0.0394	0.0189	0.0244	0.0349	0.0106
0.4	0.0272	0.0398	0.0126	0.0282	0.0338	0.0056
0.5	0.0295	0.0260	-0.0035	0.0298	0.0283	-0.0015
0.6	0.0249	0.0456	0.0207	0.0262	0.0265	0.0003
0.7	0.0301	0.0321	0.0020	0.0284	0.0310	0.0026
0.8	0.0304	0.0354	0.0050	0.0296	0.0285	-0.0010
0.9	0.0276	0.0309	0.0033	0.0269	0.0306	0.0037
1.0	0.0274	0.0322	0.0048	0.0241	0.0296	0.0054
average	0.0262	0.0350	0.0089	0.0259	0.0303	0.0044

TABLE III
TRACKING ERRORS ON S&P 100

γ	KMP			CKMP		
	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ
0.0	0.0266	0.0404	0.0137	0.0311	0.0464	0.0152
0.1	0.0217	0.0497	0.0280	0.0246	0.0453	0.0207
0.2	0.0182	0.0452	0.0271	0.0277	0.0459	0.0182
0.3	0.0161	0.0433	0.0271	0.0216	0.0429	0.0213
0.4	0.0199	0.0390	0.0191	0.0210	0.0397	0.0187
0.5	0.0317	0.0399	0.0082	0.0250	0.0397	0.0146
0.6	0.0240	0.0324	0.0084	0.0271	0.0323	0.0052
0.7	0.0251	0.0358	0.0107	0.0254	0.0369	0.0114
0.8	0.0267	0.0289	0.0022	0.0289	0.0286	-0.0003
0.9	0.0227	0.0234	0.0008	0.0223	0.0238	0.0015
1.0	0.0248	0.0262	0.0014	0.0255	0.0259	0.0004
average	0.0234	0.0367	0.0133	0.0255	0.0370	0.0115

TABLE IV
TRACKING ERRORS ON FTSE 100

γ	KMP			CKMP		
	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ
0.0	0.0238	0.0253	0.0016	0.0201	0.0235	0.0035
0.1	0.0221	0.0191	-0.0030	0.0237	0.0238	0.0001
0.2	0.0195	0.0235	0.0039	0.0244	0.0244	0.0000
0.3	0.0227	0.0261	0.0034	0.0247	0.0259	0.0012
0.4	0.0264	0.0254	-0.0010	0.0228	0.0262	0.0034
0.5	0.0266	0.0332	0.0067	0.0217	0.0331	0.0114
0.6	0.0288	0.0273	-0.0015	0.0220	0.0271	0.0051
0.7	0.0209	0.0242	0.0033	0.0183	0.0302	0.0119
0.8	0.0268	0.0249	-0.0019	0.0224	0.0249	0.0025
0.9	0.0254	0.0252	-0.0003	0.0219	0.0254	0.0035
1.0	0.0201	0.0252	0.0051	0.0165	0.0287	0.0122
average	0.0239	0.0254	0.0015	0.0217	0.0267	0.0050

TABLE V
TRACKING ERRORS ON DAX 100

γ	KMP			CKMP		
	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ	$ARMSE_{DTW}$	$ARMSE_{PCC}$	Δ
0.0	0.0495	0.0681	0.0186	0.0592	0.0638	0.0045
0.1	0.0624	0.0719	0.0096	0.0459	0.0704	0.0245
0.2	0.0540	0.0694	0.0154	0.0660	0.0696	0.0036
0.3	0.0568	0.0692	0.0124	0.0543	0.0675	0.0133
0.4	0.0593	0.0801	0.0208	0.0491	0.0721	0.0230
0.5	0.0502	0.0642	0.0140	0.0508	0.0683	0.0175
0.6	0.0674	0.0588	-0.0085	0.0502	0.0623	0.0121
0.7	0.0642	0.0741	0.0099	0.0541	0.0732	0.0191
0.8	0.0703	0.0477	-0.0225	0.0507	0.0704	0.0198
0.9	0.0531	0.0487	-0.0044	0.0525	0.0635	0.0110
1.0	0.0625	0.0642	0.0017	0.0413	0.0574	0.0161
average	0.0591	0.0651	0.0061	0.0522	0.0671	0.0150

C. Evaluation Criterion

Root-mean-square error (RMSE) as defined below is used to evaluate the tracking performance for each test period:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_{\text{index}} - r_{\text{portf}})^2},$$

where the cumulative returns r of the tracking-portfolio and the target index are denoted as r_{portf} and r_{index}, respectively. n is the number of observations (weeks). The cumulative returns are used for testing instead of log returns. A cumulative return r is a percentage price change of a stock and is defined as:

$$r = \frac{P_{cur} - P_{ori}}{P_{ori}},$$

where P_{cur} denotes the stock's current price, P_{ori} denotes the stock's original price.

The Average root-mean-square error (ARMSE) is defined as the average $RMSE_\tau$ over T test periods:

$$ARMSE = \frac{1}{T}\sum_{\tau=1}^{T} RMSE_\tau.$$

The $ARMSE$ is a metric to evaluate the overall tracking performance of index tracking methods.

D. Experimental Results

In the results, DTW-KMP and DTW-CKMP denote, respectively, the k-median problem and the constrained k-median problem based on DTW. PCC-KMP and PCC-CKMP denote, respectively, the k-median problem and the constrained k-median problem based on PCC.

1) DTW vs. PCC: Fig. 2 depicts the $ARMSE$ with various values of γ, where the blue lines are results using DTW, and the yellow lines are results with PCC, the dash lines are results based on KMP, and the solid lines are results based on CKMP. The four subplots show the same pattern that the blue lines are lower than the yellow ones, which means that DTW outperforms PCC. These results verify that DTW captures the features of stock series' dissimilarity better and is a more suitable measure for clustering stock price variations than PCC. The blue lines in the upper two subplots illustrate that, by using DTW, there is no significant difference between KMP and CKMP in tracking Hang Seng or S&P 100. Nevertheless, when tracking FTSE 100 and DAX 100 (results in the lower two subplots), by using DTW, the results based on CKMP (the blue dash lines) has smaller $ARMSE$ than the results based on KMP. This situation indicates that the imbalance of clusters happens in tracking these two indices, and the valuation constraint effectively reduces testing errors. Generally, DTW-CKMP (the solid blue line) has the lowest ARMSE tracking

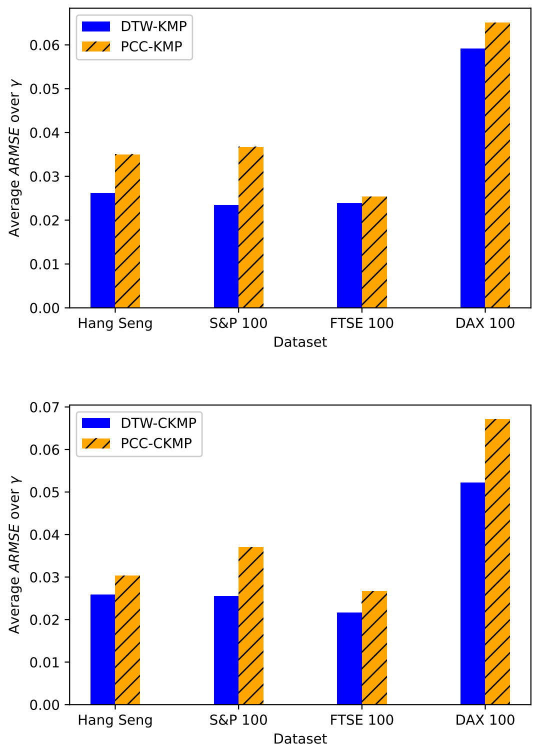

Fig. 3. Average $ARMSE$ on four benchmark datasets based on KMP (upper subplot) and CKMP (lower subplot).

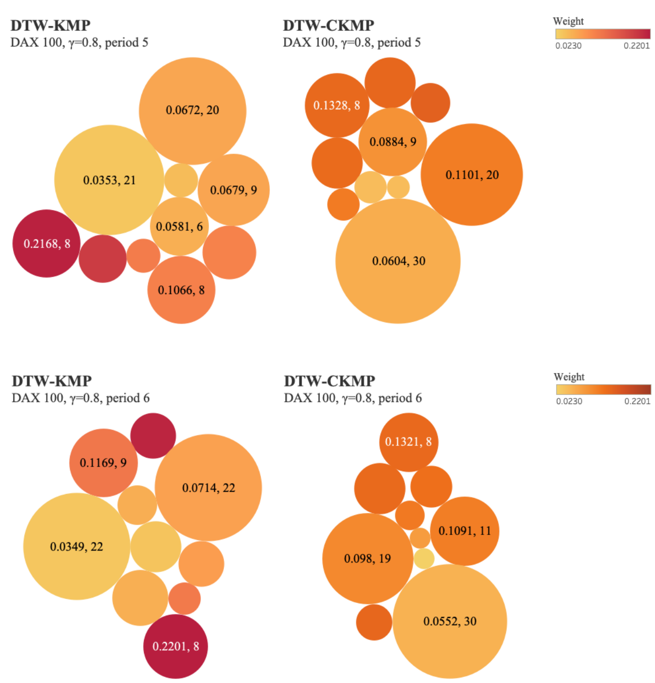

Fig. 4. Resulting clusters on tracking DAX 100 in period 5 (subplots of the upper two subplots) and period 6 (lower two subplots).

four indices. The numerical results corresponding to Fig. 2 are shown in Table II - V.

Fig. 3 depicts the averaged $ARMSE$ across all γ by using the proposed index tracking approach and its competitors on the four datasets. The upper subplot shows the results based on KMP, and the lower subplot shows the results based on CKMP. It can be observed that the measure DTW reduces the average $ARMSE$ on each dataset.

2) KMP vs. CKMP: As mentioned, CKMP is for limiting the extreme imbalance of resulting clusters. After checking each clustering result, it is found that this imbalance situation happens in tracking DAX 100 in period 5 and period 6 when $\gamma = 0.8$. Their corresponding resulting clusters are shown in Fig. 4 where the upper two subplots are for period 5, and the lower two subplots are for period 6. The color indicates exemplar's weight (the number at the front), and the circle size indicates stock amounts in one cluster (the number at the back). It can be seen that the resulting clusters based on KMP (the left two subplots) contains some large-weight clusters which are dark red. However, in CKMP (the left two subplots) results, all the clusters are orange, and no cluster is dark red. It illustrates that the weight imbalance among clusters is alleviated by adding the valuation constraint.

Fig. 5 and Fig. 6 show the cumulative returns based on KMP and CKMP with $\gamma = 0.3$, respectively, and they are for viewing the tracking performance more intuitively. Both figures illustrate that using DTW leads to lower tracking errors than using PCC.

The QUBO formulation for k-medoids clustering index-tracking [8] as a comparison is also solved by Gurobi. The tracking results of 6 test sets are shown in Fig. 7. Due to the computational resource limitation, the experiments of the QUBO formulation [8] (the PCC-QUBO in Fig. 7) is only conducted tracking Hang Seng with a forgetting factor $\gamma = 0.3$. It can be seen that the proposed DTW-KMP and DTW-CKMP are superior to PCC-QUBO [8] in terms of both solution speed and tracking accuracy.

V. Conclusion

In this paper, index tracking is carried out by means of k-medoids clustering formulated as a valuation-constrained k-median problem. The proposed valuation constraint has an effect of limiting the tracking errors to avoid imbalanced clustering results. Experimental results on four indices are elaborated to compare, analyze, and demonstrate the tracking performance of the proposed approach against baselines. Comparison studies show that the proposed approach based on constrained k-median problem with DTW-based similarity coefficients outperforms other approaches in terms of tracking accuracy. Further investigations may aim to enhanced index tracking based on neural networks [20], and multi-agent systems [21].

References

[1] G. Cornuejols and R. Tutuncu, *Optimization Methods in Finance*. Cambridge University Press, 2006.

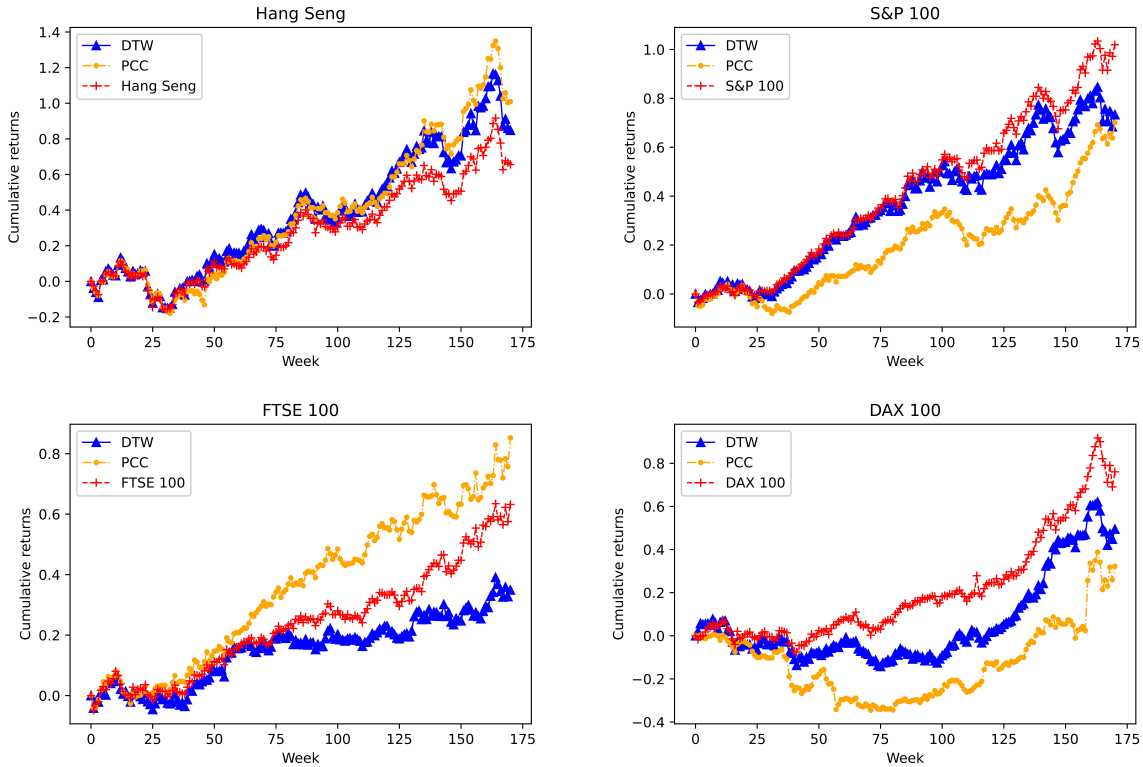

Fig. 5. Cumulative returns based on KMP for tracking four benchmark indices ($\gamma = 0.3$).

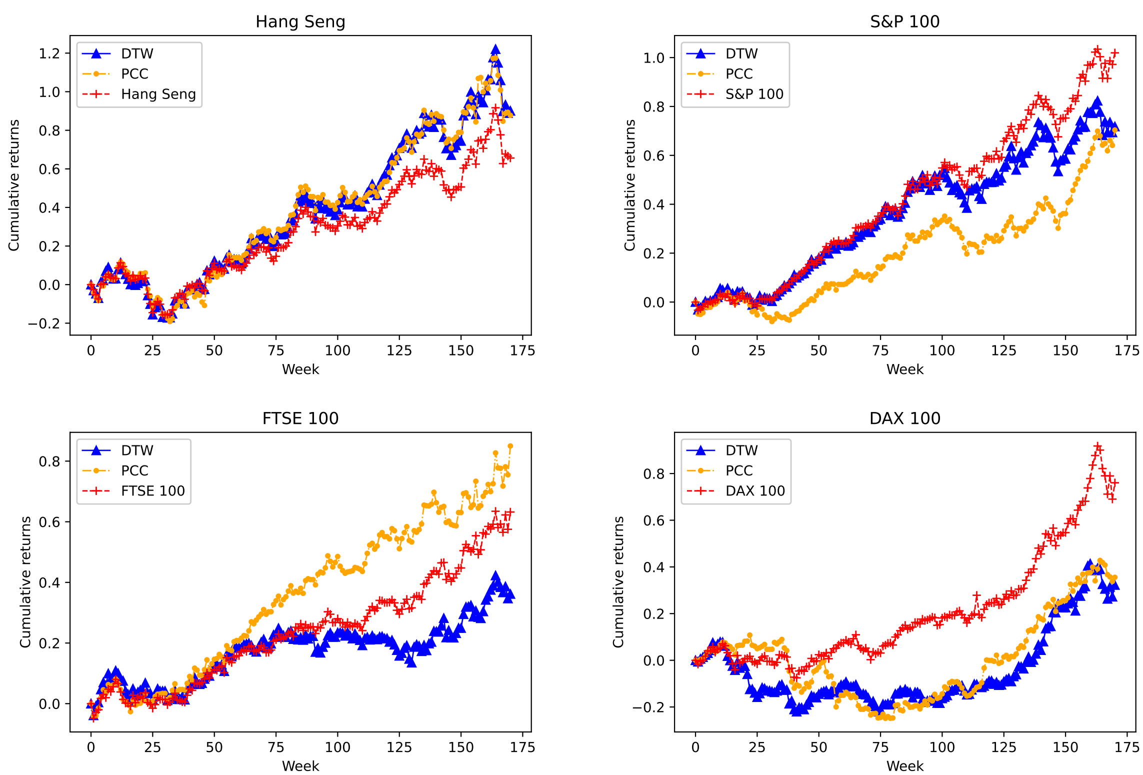

Fig. 6. Cumulative returns based on CKMP for tracking four benchmark indices ($\gamma = 0.3$).

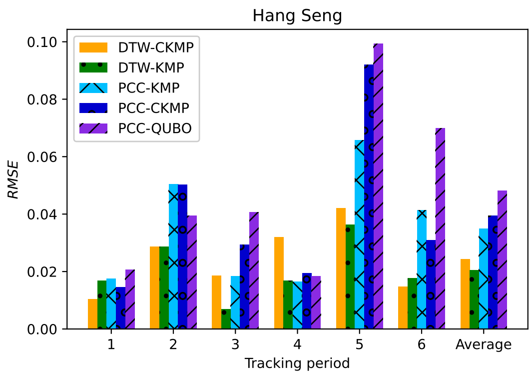

Fig. 7. RMSE in test period for tracking Hang Seng ($\gamma = 0.3$).

[2] J. L. Treynor and F. Black, "How to use security analysis to improve portfolio selection," *The Journal of Business*, vol. 46, no. 1, pp. 66–86, 1973.

[3] S. Chiam, K. Tan, and A. A. Mamun, "Dynamic index tracking via multi-objective evolutionary algorithm," *Applied Soft Computing*, vol. 13, no. 7, pp. 3392–3408, 2013.

[4] J. Beasley, N. Meade, and T.-J. Chang, "An evolutionary heuristic for the index tracking problem," *European Journal of Operational Research*, vol. 148, no. 3, pp. 621–643, 2003.

[5] B. Fastrich, S. Paterlini, and P. Winker, "Cardinality versus q-norm constraints for index tracking," *Quantitative Finance*, vol. 14, no. 11, pp. 2019–2032, 2014.

[6] N. Canakgoz and J. Beasley, "Mixed-integer programming approaches for index tracking and enhanced indexation," *European Journal of Operational Research*, vol. 196, no. 1, pp. 384–399, 2009.

[7] D. Wu, R. Kwon, and G. Costa, "A constrained cluster-based approach for tracking the S&P 500 index," *International Journal of Production Economics*, vol. 193, pp. 222–243, 2017.

[8] S. W. Hong, P. Miasnikof, R. Kwon, and Y. Lawryshyn, "Market graph clustering via QUBO and digital annealing," *Journal of Risk and Financial Management*, vol. 14, no. 1, 2021.

[9] Y. Zheng, T. M. Hospedales, and Y. Yang, "Diversity and sparsity: A new perspective on index tracking," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1768–1772.

[10] C. Bauckhage, N. Piatkowski, R. Sifa, D. Hecker, and S. Wrobel, "A QUBO formulation of the k-medoids problem," in *Proceedings of the Conference on Lernen, Wissen, Daten, Analysen, Berlin, Germany, September 30 - October 2, 2019*, ser. CEUR Workshop Proceedings, R. Jäschke and M. Weidlich, Eds., vol. 2454. CEUR-WS.org, 2019, pp. 54–63.

[11] P. E. Puspita and Zulkarnain, "A practical evaluation of dynamic time warping in financial time series clustering," in *2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2020, pp. 61–68.

[12] L. Bai, L. Cui, Z. Zhang, L. Xu, Y. Wang, and E. R. Hancock, "Entropic dynamic time warping kernels for co-evolving financial time series analysis," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2020.

[13] K. Nakagawa, M. Imamura, and K. Yoshida, "Stock price prediction using k-medoids clustering with indexing dynamic time warping," *Electronics and Communications in Japan*, vol. 102, no. 2, pp. 3–8, 2019.

[14] J. Mulvey and H. Crowder, "Cluster analysis: An application of lagrangian relaxation," *Manage. Sci*, vol. 25, pp. 329–340, 1979.

[15] J. Wang, "A linear assignment clustering algorithm based on the least similar cluster representatives," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 29, no. 1, pp. 100–104, 1999.

[16] R. E. Walpole, *Probability and Statistics for Engineers and Scientists*. Pearson, 2011.

[17] H. Sakoe and S. Chiba, "A dynamic programming approach to continuous speech recognition," *Proceedings of the Seventh International Congress on Acoustics*, vol. 3, pp. 65–69, 1971.

[18] ——, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978.

[19] R. Tavenard, J. Faouzi, G. Vandewiele, F. Divo, G. Androz, C. Holtz, M. Payne, R. Yurchak, M. Rußwurm, K. Kolar, and E. Woods, "Tslearn, a machine learning toolkit for time series data," *Journal of Machine Learning Research*, vol. 21, no. 118, pp. 1–6, 2020.

[20] M.-F. Leung and J. Wang, "Minimax and biobjective portfolio selection based on collaborative neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 2825–2836, 2021.

[21] M.-F. Leung, J. Wang, and D. Li, "Decentralized robust portfolio optimization based on cooperative-competitive multiagent systems," *IEEE Transactions on Cybernetics*, 2021, in press.