

WAZI: A Learned and Workload-aware Z-Index

Sachith Pai
University of Helsinki
Helsinki, Finland
sachith.pai@helsinki.fi

Michael Mathioudakis
University of Helsinki
Helsinki, Finland
michael.mathioudakis@helsinki.fi

Yanhao Wang
East China Normal University
Shanghai, China
yhwang@dase.ecnu.edu.cn

ABSTRACT

Learned indexes fit machine learning (ML) models to the data and use them to make query operations more time and space-efficient. Recent works propose using learned spatial indexes to improve spatial query performance by optimizing the storage layout or internal search structures according to the data distribution. However, only a few learned indexes exploit the query workload distribution to enhance their performance. In addition, building and updating learned spatial indexes are often costly on large datasets due to the inefficiency of (re)training ML models.

In this paper, we present WAZI, a learned and workload-aware variant of the Z-index, which jointly optimizes the storage layout and search structures, as a viable solution for the above challenges of spatial indexing. Specifically, we first formulate a cost function to measure the performance of a Z-index on a dataset for a range-query workload. Then, we optimize the Z-index structure by minimizing the cost function through adaptive partitioning and ordering for index construction. Moreover, we design a novel page-skipping mechanism to improve the query performance of WAZI by reducing access to irrelevant data pages. Our extensive experiments show that the WAZI index improves range query time by 40% on average over the baselines while always performing better or comparably to state-of-the-art spatial indexes. Additionally, it also maintains good point query performance. Generally, WAZI provides favorable tradeoffs among query latency, construction time, and index size.

1 INTRODUCTION

Spatial query processing has attracted significant interest in the database community over the last three decades with the proliferation of location-based services (LBSs). Although numerous index structures such as R-trees [17] and k-d trees [6] have been deployed in database systems to improve spatial query performance, their efficiency still cannot fully satisfy the requirements of real-world applications due to the rapid growth in the volume of spatial data.

The seminal work of Kraska et al. [22] inspired the introduction of machine learning (ML) based indexes (i.e., *learned indexes*) [7, 10, 18, 22, 27, 36, 43, 46] to replace their traditional counterparts as a way of improving query performance by exploiting data or query patterns or both while reducing space usage. Most of the learned indexes developed are specific for one-dimensional data [7, 10, 13, 18, 22], which follows the abstraction that an index is a model that predicts the position of an element in a sorted array. Such an abstraction is made possible only by the prerequisite that the data resides in a sorted array. Therefore, any model that accurately and efficiently learns the data’s cumulative density function (CDF) can serve as an index.

However, the above abstraction is not squarely applicable to spatial indexes since there is no predefined ordering between data points. Nonetheless, a few learned spatial indexes have attempted to overcome this challenge by using space-filling curves (SFCs) to linearize the spatial indexing task [36, 38, 43]. The construction procedure for such indexes proceeds in two phases. The first phase uses space-filling curves for data linearization. The second phase performs the transformed one-dimensional indexing task. These indexes usually optimize the second phase of the construction while using standard methods for the first phase (linearization), missing out on the opportunity to tailor the linearization to fit the data and workload better.

To remedy the problem, in this work, we present a generalized variant of the Z-index that is both data- and workload-aware. A Z-index is an intuitively simple spatial index structure with a long history in database management systems [39]. It uses a *Z-curve* (a.k.a. *Z-order curve*) to compute a sort order, denoted by ORD, for multi-dimensional data points. Figure 1a shows an example of the Z-curve and the corresponding Z-index. The Z-curve visits the data points according to a hierarchical partitioning of the data space into cells and a specific ordering of these cells. For instance, at the top level, the space is partitioned into four cells, namely A, B, C, and D; at the second level, each of these cells is partitioned into four sub-cells, and so on, with the partitioning happening at the coordinates corresponding to the median of the data along each axis. Within each cell, the ordering of its sub-cells consistently follows the same ‘Z’-like pattern: cells are ordered as “ABCD” for the first level, with higher-level cells having higher priority in the point ordering. Each cell is further partitioned and ordered until the cells reach a predetermined size L .

The Z-index has the desirable property of *monotonicity*; that is, any point “dominated” (i.e., having smaller values in both dimensions) by point a in the data space will always have a smaller sort order ORD than a . This monotonicity property facilitates the processing of range queries, represented as rectangles in the data space. A standard range-query algorithm is to obtain the locations in ORD of the bottom-left and top-right points of the query rectangle, scan the data entries between the two, and filter the data points that satisfy the query range. However, the efficiency of such a range-query algorithm could still be vastly improved. For example, let us consider the range query with rectangle R in Figure 1b: the algorithm processes many data entries outside the result set (e.g., all points in cell B), which could lead to high query latencies.

Our Contributions. We propose a generalization of the Z-index that adapts gracefully to both the distribution of spatial data and the workload of range queries, mitigating the retrieval of redundant data entries, and thus leading to improved query latencies at little additional cost to index construction. Specifically, we present a method to construct a learned and workload-aware space-filling curve to improve the range-query performance of a generalized Z-index variant. The Workload-aware Z-Index (WAZI) variant we propose is flexible in two ways compared to the base Z-index: the partition location and the ordering of

© 2024 Copyright held by the owner/author(s). Published in Proceedings of the 27th International Conference on Extending Database Technology (EDBT), 25th March-28th March, 2024, ISBN 978-3-89318-095-0 on OpenProceedings.org. Distribution of this paper is permitted under the terms of the Creative Commons license CC-by-nc-nd 4.0.

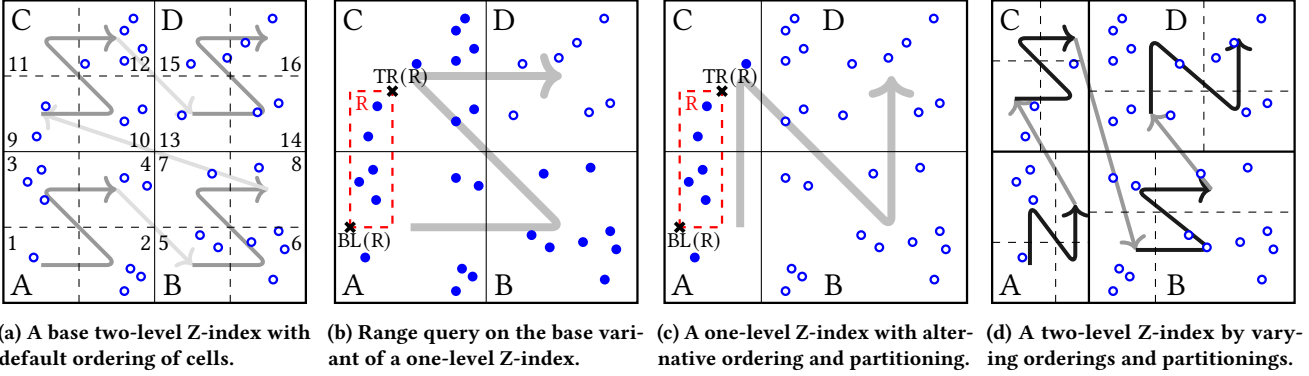


Figure 1: Illustrations of the base Z-curve and Z-index and their variants proposed in this work.

child cells. Intuitively, it pays off to partition the data space so that cells correspond to regions that are fetched by several similar range queries. Such a partitioning allows the index to avoid processing points in other cells during the execution of these queries. For example, the Z-index with an alternative ordering and partitioning shown in Figure 1c processes fewer redundant data entries than the base Z-index in Figure 1b. Therefore, if similar queries dominate the anticipated query workload, the alternative ordering and partitioning of cells can reduce query latency. The Z-index variant we propose is general enough to allow different ordering and partitioning for each cell across the index hierarchy, as shown in Figure 1d. Additionally, by harnessing the structural properties of the index, we develop and incorporate an efficient mechanism to reduce the computation required by skipping non-overlapping pages during range query processing.

In the experiments, we compare WAZI with four state-of-the-art spatial indexes including FLOOD [29], STR [25], a query-aware incremental index QUASII [35], and a cost-optimized variant of R-tree CUR [40]. The baselines considered cover the scope of indexes regarding query-awareness and the use of learned components. The results show that WAZI performs better than these baselines in terms of range and point query latency while exhibiting favorable tradeoffs regarding construction time and index size. In the critical metric of range query latency, WAZI significantly outperforms all the baselines. The main contributions of this paper include:

- 1) formalizing the problem of optimizing the partitioning and Z-ordering for the given data and anticipated query workload;
- 2) providing an index construction algorithm that minimizes the retrieval cost;
- 3) presenting a mechanism to efficiently skip over large regions of irrelevant pages to avoid the computational cost associated with skipped pages;
- 4) providing an experimental evaluation of the proposed index against existing approaches, addressing the cost redemption behavior and effects of workload changes, as well as an ablation study to demonstrate the benefit from different components of our approach.

An extended abstract associated with this work was presented at a non-archival workshop [33].

2 RELATED WORK

In this section, we review the existing literature on *learned indexes* and *spatial indexes*, which are closely relevant to the problem we study in this paper.

Traditional Spatial Indexes. Spatial indexes [12] were well studied across several decades in the database community. Existing (non-learned) spatial indexes are generally categorized into three classes. The first category is *space partitioning-based indexes*, e.g., k-d trees [6], Quad-trees [11], and Grid Files [30], which recursively split the data space into sub-regions and then index each sub-region hierarchically. The second category is *data partitioning-based indexes*, including R-tree [17] and its variants [2, 4, 5, 20, 41], which recursively divide the dataset into subsets and then index each subset hierarchically. The differences between different R-tree variants lie in how they evaluate the goodness of data partitioning and the algorithms for index construction based on the partitioning scheme. The third category is *data transformation-based indexes*, which transform multi-dimensional data into one dimension and then utilize a one-dimensional index, e.g., B-trees. *Space-filling curves* (SFCs) are the typical transformation method. Such data transformations act as a hybrid between data and space partitioning methods where the space partitioning nature of SFC is utilized along with data partitioning methods of the one-dimension index. The Z-index [3, 39] is a typical index of this kind, based on the Z-curve for data transformation. The index construction for Z-index most commonly proceeds by sorting all the points using the SFC sort-order, packing them into leaves, and then building the index bottom-up level-by-level. Some mechanisms that utilize properties of the SFCs have been used to improve range query performance in such indexes: like the *BIGMIN* algorithm associated with Z-order curves [42] or the *calculate_next_match* function for Hilbert curves [23]. However, none of the above spatial indexes can be adapted to data or query patterns, and thus they often suffer from inferior query performance compared to their learned counterparts.

Learned Indexes. Kraska et al. [22] first proposed the *learned indexes*, which utilize machine learning models to enhance or replace traditional indexes for data access in databases. The abstraction used in [22] to motivate such methods was that an index is essentially a structure that predicts the location of an item in a sorted array, or a structure that predicts the cumulative density function of the underlying data. They propose the recursive model index (RMI), which consists of a hierarchy of regression models for capturing the relationship between sorted keys and their ranks in the dataset. According to [22], the benefits of learned indexes lie in smaller index sizes and lower query latency. After this seminal work, many learned indexes were proposed in the last four years, such as PGM-Index [10], ALEX [7],

RadixSpline [21], and Shift-Table [18]. They outperformed RMI by using simpler linear spline models and supporting efficient index updates. Nevertheless, all the learned indexes mentioned above are specific to one-dimensional data and cannot be directly used for spatial query processing.

Learned Spatial Indexes. The most relevant studies to our problem are the ones on learned spatial indexes [1, 8, 9, 16, 27–29, 36, 43, 45, 46]. The learned spatial indexes can be categorized into two categories. The first category is *learned search structure-based* indexes, where the methods use learned models to improve search performance on default data layouts. Wang et al. [43] proposed the ZM-index, an extension of RMI [22] using a linear layout based on Z-order values for spatial data points. Qi et al. [36] proposed a recursive spatial model index (RSMI) to build an index utilizing the Z-order values and using neural networks to infer data partitioning. The intuition behind RSMI is that the application of neural networks to learn fine-grained Z-order mappings would result in partitions with better locality. While ZM and RSMI seem conceptually similar to WAZI, they have three main differences from WAZI. First, ZM and RSMI utilize Z-ordering in the rank space of the data, whereas WAZI avoids the rank space projection by operating in the original data space. Second, ZM and RSMI use default Z-ordering for the data layout and a hierarchy of learned models for query processing. However, WAZI optimizes the data layout using learned models only in the training phase but does not use them to process queries. Third, both ZM and RSMI (as with other learned search structure-based indexes) are workload-agnostic, while WAZI is workload-aware. Li et al. [27] proposed LISA, a spatial index that divides the data space into a grid-like layout numbered by a partially monotonic function and learns functions to map data points to a grid cell and subsequently its respective page. Zhang et al. [46] proposed SPRIG, a learned spatial index that uses the spatial interpolation function as a learned model to efficiently filter the cells in the grid file. Al-Mamun et al. [1] proposed a method to improve the query performance of R-trees by using multi-label classification models to identify the required leaf nodes for a query incurring high tree-traversal cost.

The second category is *layout optimization-based* indexes, where data layout is optimized using ML models to improve query performance. Nathan et al. [29] proposed a learned multi-dimensional index called FLOOD based on the Grid File index [30] using machine-aware optimizations for grid layouts. FLOOD is constructed by training a performance prediction model based on randomly sampled layouts, and then using the prediction model to select the best layout from a set of candidate grid layouts. The main limitation is that FLOOD is optimized for an average query and performs poorly on skewed query workloads. This limitation was addressed in Tsunami [8], a learned multi-dimensional index that partitions the data space based on statistical tests of query distribution and constructs FLOOD indexes over each partition. The statistical tests for query distribution changes are based on the shapes and sizes of queries and hence do not apply to our setting in this paper. Yang et al. [45] proposed an index-like structure called Qd-tree to optimize multi-dimensional data layouts using deep reinforcement learning (RL). Gu et al. [16] proposed a technique to construct R-trees by training RL-methods to perform ChooseSubtree and Split operations in required for sequential index construction. Dong et al. [9] proposed an R-tree variant that utilizes ML models to partition the dataset into homogeneous zones.

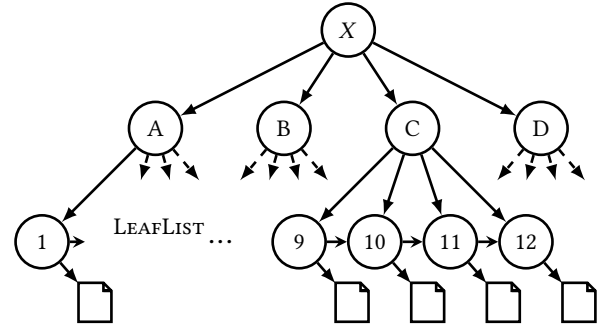


Figure 2: Illustration of the quaternary tree structure of a Z-index. Each leaf node holds a pointer to its subsequent leaf node as well as a pointer to the page containing its corresponding data points.

3 THE BASE Z-INDEX

In this section, we provide the basic background on Z-index. A Z-index is a space-partitioning index that divides the data space hierarchically. At each level, the corresponding data space is partitioned into four cells along the median of data points, and the four child cells are ordered in a ‘Z’-pattern. The complete Z-index thus enforces an ordering on cells of the leaf nodes, hence imposing a partial-sort order among the data points.

Two elements define a Z-index. The first element is a *hierarchical partitioning* of the data space into cells, with each cell partitioned into four child cells down to a predetermined granularity of leaf-cell size L . Corresponding split points \mathbf{h} determines the hierarchical partitioning: for the base Z-index variant, these occur at the median of the x - and y -coordinates of all data points. The second element is the *ordering* \mathbf{o} of the cells at each level. The four child cells of each parent cell of the base Z-index variant are ordered according to the ‘‘ABCD’’ pattern, whereby higher-level cells have higher ordering importance, as shown in Figure 1a. The curve order follows the higher-level ordering to place all leaf cells within cell B after all the leaf cells within cell A. The relative order of the leaf cells in A (or any other higher-level cell) is then determined by the second level of partitioning and ordering. Therefore, a Z-index instance \mathcal{Z} is defined as a set of partition points $\mathbf{h} = (x, y)$ and orderings \mathbf{o} associated with each of the internal nodes of a quaternary tree, as shown in Figure 2.

Each *internal node* stores the coordinates of the split point, according to which the corresponding cell is partitioned into its child cells and the ordering of the child cells. During tree traversal, say for a point query \mathbf{p} , we traverse down the tree by identifying the relevant child node at each internal node. The relevant child node is computed by first comparing the query point \mathbf{p} with the partition point \mathbf{h} of the current node to identify whether the \mathbf{p} resides in quadrants A, B, C, or D (Lines 4–5 in Algorithm 1) and picking the appropriate pointer based on the ordering \mathbf{o} used (Lines 6–9).

The *leaf nodes* of a Z-index contain a bounding rectangle (*bbs*) for the area spanned by the leaf and a pointer to a page with at most L elements. Each leaf node also contains a pointer to the next leaf node defined by the sort order, creating a linked list structure at the leaf layer of the tree (LEAFLIST). Note that Z-index is assumed to be *clustered*, with data points corresponding to consecutive leaf nodes stored in consecutive pages. We consider the data points within a page to be stored in random order.

Algorithm 1: TREE TRAVERSAL

Input : Z-index \mathcal{Z} , node X in \mathcal{Z} , point query p
Output: Boolean B whether p exists

```
1 if  $X$  is a leaf node then
2   return  $X$ .page
3 else
4    $bit_x = p.x > X.x$ 
5    $bit_y = p.y > X.y$ 
6   if  $X.o == \text{"ABCD"}$  then
7      $c_{id} = 2bit_y + bit_x$ 
8   else
9      $c_{id} = 2bit_x + bit_y$ 
10  return TREE TRAVERSAL( $X.child[c_{id}]$ )
```

Algorithm 2: RANGE-QUERY

Input : Z-index \mathcal{Z} , range query R
Output: Set of points $result$ in range R

```
1 Page LOW = TREE TRAVERSAL (BL (R))
2 Page HIGH = TREE TRAVERSAL (TR (R))
3 Initialize  $result = \emptyset$  and  $P = LOW$ 
4 while  $P \leq HIGH$  do
5    $overlap = PAGEQUERYOVERLAP (P, R)$ 
6   if  $overlap == true$  then
7      $result \leftarrow result \cup FILTER (P)$ 
8    $P = NEXTPAGE (P, overlap)$ 
9 return  $result$ 
```

Crucially, note that the ABCD ordering guarantees *monotonicity* for points that fall within different leaf cells: if point a in page X is dominated by point b in page $Y \neq X$, then leaf X will appear earlier in the LEAFLIST than Y . Here, point a is dominated by point b if $a.x \leq b.x$ and $a.y \leq b.y$, and at least one coordinate of a is strictly smaller than that of b . This monotonicity property is utilized in a Z-index for a specific range query processing mechanism. A range query R , defined by two points BL(R) and TR(R) (see Figure 1b), is processed by Z-index in two phases. First, the leaf nodes enclosing two query extremes, BL(R) and TR(R), within the Z-index are identified (Lines 1–2 of Algorithm 2) in the *index lookup phase*. Let us refer to these leaf nodes as LOW and HIGH. These represent the first and last possible leaf nodes in the LEAFLIST that may overlap with R . Second, in the *scanning phase*, we check the leaf nodes within the range [LOW : HIGH] for overlap with R based on their *bbs*, and pages of overlapping nodes are scanned to filter the query results (Lines 3–8 of Algorithm 2).

4 THE WAZI INDEX

This section presents our method to build a workload-aware variant of the Z-index, i.e., WAZI. Within a Z-index, the scanning phase completely dominates the query latency. We aim to minimize the number of points accessed during the scanning phase. Towards this end, we present a cost formulation for the number of points accessed during range query processing (Section 4.1) and an approach to building WAZI based on minimizing the cost function (Section 4.3).

4.1 Adaptive Partitioning and Ordering

The two defining elements of a Z-index, partitioning and ordering, are computed using fixed heuristics for the base variant. By contrast, we propose a generalized variant WAZI of the Z-index,

for which the partitioning and ordering can vary for each node, as illustrated in Figure 1d. First, whereas for the base variant, the split points are predetermined to be placed at the median of the data along the x and y axes, for WAZI, we propose that the split points can be placed anywhere within the data range, allowing for more flexible data partitioning. Second, whereas for the base variant, the ordering of the child cells is predetermined to follow the “ABCD” pattern for every parent cell, for WAZI, we propose the order is allowed to be either “ABCD” or “ACBD”, as both orderings preserve monotonicity.

Our aim in making these changes in WAZI is for the index to be adaptive to the given data and anticipated range queries. For example, a Z-index with the alternative partitioning and ordering shown in Figure 1c would be better adapted to the range query R shown therein than the one shown in Figure 1b, as it would retrieve fewer data points in query processing.

Specifically, for a given dataset \mathcal{D} and a set of range queries Q , we are interested in building an instance of the generalized Z-index to minimize a corresponding *retrieval cost*. The retrieval cost for a range query is measured by the number of data points compared against a query box R in the filtering phase. In practice, Q can be obtained from historical logs of range queries, as “representatives” for the application at hand, or in general, as anticipated range queries for which the index should be optimized.

4.2 Retrieval Cost

Here we present the objective function we aim to optimize when building a Z-index, as described above. For a given Z-index and query R , let $\delta_{R \in XY}$ be the function to indicate whether R has its bottom-left vertex in X and its top-right vertex in Y . As a simple illustration case, let us consider a single-level Z-index with cells ordered according to the “ABCD” order, and the splits occur as in Figure 1b. Following the range query processing discussed earlier, the retrieval cost for the chosen split and ordering equals

$$\begin{aligned} cost_X(R | x, y; ABCD) = & \delta_{R \in AD}(n_A + n_B + n_C + n_D) + \\ & \delta_{R \in AC}(n_A + an_B + n_C) + \delta_{R \in BD}(n_B + an_C + n_D) + \\ & \delta_{R \in AB}(n_A + n_B) + \delta_{R \in CD}(n_C + n_D) + \\ & \delta_{R \in AA}n_A + \delta_{R \in BB}n_B + \delta_{R \in CC}n_C + \delta_{R \in DD}n_D, \end{aligned} \quad (1)$$

where n_X denotes the number of data points in each cell. Notice that n_X and $\delta_{R \in XY}$ depend on the split location (x, y) ; however, we omit the dependency from our notation for simplicity. To see why the formula holds, note that when $R \in AB$, the Z-index retrieves all points only from cells A and B, as no other cells fall between A and B in the “ABCD” ordering of cells. However, when $R \in AC$ (as in Figure 1b), in addition to filtering points from A and C, the Z-index also compares *bbs* and skips over the non-overlapping leaf node B as it falls between A and C in the “ABCD” ordering. The impact of skipping over leaf cells is represented in our cost as a fraction $\alpha < 1$ of the number of points.

The rest of the cases follow similarly. Notice that if the ordering of the cells is “ACBD” instead, the cost formula will differ from Eq. 1:

$$\begin{aligned} cost_X(R | x, y; ACBD) = & \delta_{R \in AD}(n_A + n_B + n_C + n_D) + \\ & \delta_{R \in AB}(n_A + an_B + n_C) + \delta_{R \in CD}(n_B + an_C + n_D) + \\ & \delta_{R \in AC}(n_A + n_C) + \delta_{R \in BD}(n_B + n_D) + \\ & \delta_{R \in AA}n_A + \delta_{R \in BB}n_B + \delta_{R \in CC}n_C + \delta_{R \in DD}n_D \end{aligned} \quad (2)$$

More generally, when the Z-index consists of more than one level of partitions, the retrieval cost is defined recursively, as the

Z-index structure of second-level partitions (for each of A, B, C, and D) affects the total cost of retrieval. Defining the retrieval cost of a Z-index with two levels would involve recursively substituting Eq. 1 or 2 for each term of the form $\delta_{R \in XX} n_X$ based on the ordering o_X under consideration:

$$\begin{aligned} cost_X(R | x, y; ACBD) &= \delta_{R \in AD}(n_A + n_B + n_C + n_D) + \\ &\delta_{R \in AB}(n_A + \alpha n_B + n_C) + \delta_{R \in CD}(n_B + \alpha n_C + n_D) + \\ &\delta_{R \in AC}(n_A + n_C) + \delta_{R \in BD}(n_B + n_D) + \\ &cost_A(R | x, y; o_A) + cost_B(R | x, y; o_B) + \\ &cost_C(R | x, y; o_C) + cost_D(R | x, y; o_D) \end{aligned} \quad (3)$$

For a set of queries Q and a Z-index Z with cell X , the total cost of all queries is aggregated to form the full cost C . Specifically, if the ordering of X is “ABCD”, the cost function is

$$\begin{aligned} C_X(Q | x, y; ABCD) &= \sum_{R \in Q} cost_X(R | x, y; ABCD) \\ &= q_{AD}(n_A + n_B + n_C + n_D) + \\ &q_{AC}(n_A + \alpha n_B + n_C) + q_{BD}(n_B + \alpha n_C + n_D) + \\ &q_{AB}(n_A + n_B) + q_{CD}(n_C + n_D) + \\ &C_A(R | x, y; o_A) + C_B(R | x, y; o_B) + \\ &C_C(R | x, y; o_C) + C_D(R | x, y; o_D), \end{aligned} \quad (4)$$

with the terms $q_{XY} = \sum_{R \in Q} \delta_{R \in XY}$. A similar equation is also present for the alternative ordering of “ACBD”.

4.3 Index Construction

The formulation of retrieval cost in Eq. 4 leads to a cost function that exhibits an optimal substructure, where optimal configuration and the associated ordering for all possible child cell combinations should be known before one can compute the optimal configuration at a given node. Finding the optimal solution under this cost formulation using dynamic programming has a complexity of $O(N^4)$. The complexity follows because, for N points in two dimensions, there are N^4 rectangles enclosing unique subsets of points. Hence, the state space for dynamic programming is at most N^4 . Obviously, such an approach is infeasible even for moderately sized datasets.

Instead, we adopt a GREEDY algorithm for index construction. The GREEDY algorithm simplifies Eq. 4 by formulating the cost for lower levels as $q_{XX} n_X$. This simplification uses an upper bound on the possible retrieval cost of each sub-partition in place of the recursive cost definition of Eq. 4. Following this intuition also yields an approach that allows for optimization steps to be performed greedily for each layer hierarchically. Intuitively, the greedy algorithm proceeds at one level at a time, from top to bottom, selecting the partition point and ordering using the alternative cost function:

$$\begin{aligned} C_X(Q | x, y; ABCD) &= \sum_{R \in Q} cost_X(R | x, y; ABCD) \\ &= q_{AD}(n_A + n_B + n_C + n_D) + \\ &q_{AC}(n_A + \alpha n_B + n_C) + q_{BD}(n_B + \alpha n_C + n_D) + \\ &q_{AB}(n_A + n_B) + q_{CD}(n_C + n_D) + \\ &q_{AA} n_A + q_{BB} n_B + q_{CC} n_C + q_{DD} n_D \end{aligned} \quad (5)$$

The pseudocode for index construction is presented in Algorithm 3. The steps for index construction are described at three levels of detail. First, our algorithm proceeds greedily, determining the partitioning and ordering of the cells within the same

Algorithm 3: GREEDY

```

Input :Z-index  $Z$ , node  $X$  in  $Z$ , workload  $Q$ , data  $\mathcal{D}$ 
1 if  $n_X < L$  then return;
   /* Draw  $\kappa$  candidate split points */
2  $XY := \text{UniformSample}(X, \kappa)$ 
   /* Select split and ordering to minimize Eq. 5 */
3  $(x, y, o) = \arg \min_{(x, y) \in XY; o \in \{ABCD, ACBD\}} C(Q | x, y; o)$ 
   /* Define cells w.r.t. split point */
4 Cells  $A, B, C, D := \text{Split}(X, (x, y))$ 
5 Add cells  $A, B, C, D$  and ordering  $o$  to  $Z$ 
   /* Apply Greedy to child cells */
6 foreach Cell  $Y \in \{A, B, C, D\}$  do
7    $\lfloor$  GREEDY( $Z, Y, Q, \mathcal{D}$ ; SolveCell)

```

level, one level at a time, from top (root) to bottom (leaf). Therefore, every time a cell X needs to be split and its children ordered, we use the configuration that minimizes the objective C (Eq. 5).

Second, we choose the partitioning and ordering minimizing the objective C for a given cell X by uniformly sampling the candidate split points and selecting the one that minimizes our objective. More specifically, for each cell that it considers for splitting, GREEDY samples κ candidate split points uniformly at random from the region covered by the cell, evaluates the objective C for each candidate split point for both possible orderings, and returns the split-point and ordering with the minimum C . We choose sampling-based optimization over more complicated optimizers (like Bayesian optimization) to avoid the high computation overhead incurred by such optimizers. More importantly, we observed that each iteration of an optimizer incurs computational cost several magnitudes higher than computing the objective function. In our experiments, we observe no performance improvement over sampling-based optimization to justify the added computational cost of said optimizers.

Third, we approximate the exact data distribution \mathcal{D} and range query distribution Q with approximate distributions by ML models (i.e., “learned” approaches). These approximate distributions allow for efficient estimations of the number of data points (and queries) that fall within each of the child cells resulting from a candidate partitioning of a given cell, as needed to compute the objective C . We used Random Forest Density Estimation models (RFDE) [44] for our density estimation. Specifically, we construct a forest of k-d trees constructed with randomized split dimensions at each node. Each node stores the number of data points (*cardinality*) contained within the region. Density estimation in an RFDE model is performed as a tree traversal, collecting the cardinality information from nodes overlapping the density estimation query.

5 SKIPPING MECHANISM

Processing range queries involves scanning and filtering points from an interval [LOW : HIGH] of leaf nodes within the LEAF LIST ordered by a Z-index. Figure 3a illustrates such an ordering (solid grey arrows) of leaf nodes. The leaf nodes have been named a through q following the ordering shown in grey. Given a range query R (red box), the query processing algorithm will compare all leaf nodes within the interval $[a : m]$ (shaded in red) and return points from relevant leaf nodes (i.e., those that overlap with the range query). For the example in Figure 3a, the relevant leaf nodes are a, j, k, l , and m , whereas the pages b through i are irrelevant. We perform a bounding box comparison of the pages

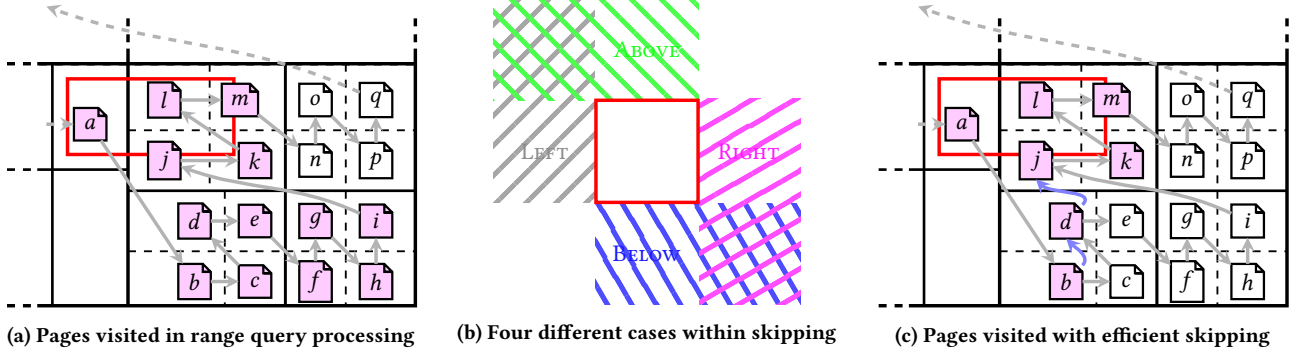


Figure 3: Illustration of skipping during range query processing; (a) Standard range query processing of range query R (red) processes pages in range $[a : m]$; (b) The four different irrelevancy criteria explained. (c) Motivating example for efficient skipping. As we process page b , we know that it does not overlap the query because of **BELOW**. We also know that the next page in the sort-order that may satisfy the criterion is d . Similarly, at page d we can skip ahead to page j , saving the computation required to process pages e through i ;

Algorithm 4: Look-ahead Pointer Construction

```

Input: Z-index  $\mathcal{Z}$ 
/* Initialize all lookahead pointer */
1 InitializeLookahead(LEAFLIST)
/* Iterate backward through leaf nodes */
2 foreach node  $P$  in Reverse(LEAFLIST) do
3   foreach CASE in [BELOW, ABOVE, LEFT, RIGHT] do
4     /* Init with the next ptr of LeafList */
5      $P.CASE = next(P)$ 
6     while CASE not improved do
7        $P.CASE = (P.CASE).CASE$ 

```

before scanning points stored within them and hence data on pages b through i are not scanned. But, the number of points r returned by the algorithm for a range query is often comparable (or significantly smaller in case of low selectivity queries) than the total number of leaf nodes s for which we compare overlap, that is, $s \gg r$. In such cases, the redundant computation of checking overlap between bounding boxes of irrelevant leaf nodes and the range query can become a bottleneck. Our solution to address this issue is skipping over leaf nodes irrelevant to the query. For the example above, Figure 3c depicts two practical instances of skipping from b to d and d to j (blue arrows), reducing the number of leaf nodes visited (shaded red) from 13 to 7.

We propose our novel mechanism for efficient skipping in two parts. First, we design a skipping mechanism that operates on any range query using *look-ahead pointers*. Second, we present the algorithm to precompute the look-ahead pointers during index construction. Finally, we modify the retrieval cost from Eq. 5 to accurately incorporate the impact of skipping.

5.1 Look-ahead Pointers

In standard processing (cf. Section 3) for a range query R , we check whether the bounding box of any leaf node overlaps with R . If so, we read and filter points stored in the page associated with that leaf node. Otherwise, we follow the pointer to the next leaf node in the LEAFLIST, iteratively proceeding until a relevant page is found or we reach the end of the interval $[LOW : HIGH]$. As discussed earlier, this may be inefficient. Our solution introduces four additional look-ahead pointers for each leaf node, allowing

us to skip irrelevant pages accessed with the standard query processing algorithm.

The four look-ahead pointers map to the four possible criteria under which a leaf node may be irrelevant; we name them **BELOW**, **ABOVE**, **LEFT**, and **RIGHT**. For instance, **BELOW** indicates that the y -coordinate of the top-right of the leaf node P , represented by $TR(P).y$, is lower than the y -coordinate of bottom-left of the query R , $TR(P).y < BL(R).y$. Or put simply, the area covered by P lies entirely below the range query R . The other criteria follow similarly. Figure 3b shows the four criteria for a candidate R . For example, in Figure 3a, the leaf node represented by d is irrelevant to the range query shown in red as it would satisfy **BELOW** criteria, where f satisfies both **RIGHT** and **BELOW** criteria. Note that pages within the interval $[LOW : HIGH]$ of a query R cannot lie in the bottom-left and top-right sections due to monotonicity constraint.

A look-ahead pointer associated with each of the four criteria points to the next possible leaf node that could be relevant, skipping over leaf nodes that are guaranteed to be deemed irrelevant due to the same criteria. Specifically, consider a leaf node P_1 whose look-ahead pointer associated with **BELOW** points to P_2 . P_2 is the earliest leaf node in LEAFLIST that satisfies the conditions $TR(P_1).y < TR(P_2).y$ and $ORD(P_1) < ORD(P_2)$. For example, leaf node d in Figure 3a is irrelevant due to the **BELOW** criterion. Consequently, nodes $[e : i]$ are guaranteed to be irrelevant to R due to **BELOW** as $\forall x \in [e : i], TR(x).y \leq TR(d).y$. Therefore, we know that for any query R that disqualifies d due to **BELOW**, nodes $[e : i]$ will also be deemed irrelevant.

We utilize look-ahead pointers to modify the range-query processing algorithm. If we identify that a leaf node P does not overlap with R , we now follow a look-ahead pointer instead of following the next pointer. The choice of look-ahead pointer is made by discerning the criteria under which leaf node P was deemed irrelevant. If an irrelevant leaf node satisfies multiple criteria, we pick the look-ahead pointer that skips over the greatest number of nodes.

5.2 Building Look-ahead Pointers

The algorithm for computing the look-ahead pointers is presented in Algorithm 4. The look-ahead pointers are constructed in the final phase of index construction for a Z-index, where the hierarchical structure already imposes an ordering ORD on the

leaf nodes. The construction of look-ahead pointers considers leaf nodes in the reverse order of the LEAFLIST. The look-ahead pointers for the last leaf node point to a dummy page signifying the end of LEAFLIST. For each subsequent (i.e., earlier in LEAFLIST) leaf node P , we utilize the constructed look-ahead pointers within the suffix of LEAFLIST to compute the look-ahead pointer. To construct a look-ahead pointer associated with a given criterion (say, BELOW), we temporarily assign the corresponding look-ahead pointer $P.BELOW$ to the next pointer of LEAFLIST. We then recursively check whether the node pointed to by $P.BELOW$ improves the criterion. Improving the criterion in this case refers to having an improved value for the corresponding disqualifying comparison. In the case of BELOW, the improved value would be if the pointer $P.BELOW$ points to a page with a greater upper bound value along the y -coordinate; put simply, it is higher than leaf node P . If the check for improving criterion fails, we follow the corresponding nodes’s pointer for the criterion, thus finding later nodes recursively (Line 6 of Algorithm 4). For each leaf node, the look-ahead construction algorithm performs at most $\sqrt{|\text{LEAFLIST}|}$ recursion steps. Therefore, the complexity of Algorithm 4 is $O(|\text{LEAFLIST}|^{3/2})$, where $|\text{LEAFLIST}| \approx N/L$.

The retrieval cost formulation presented in Eq. 5 accounts for the skipping cost of irrelevant leaf nodes using α . We can now update the cost formulation to accurately reflect the retrieval cost of a given query in light of the skipping mechanism mentioned above and account for the redundant quadrant fetched for processing queries by setting the α value to a small constant. In our experiments, we set α to 10^{-5} . The small value of α will more accurately reflect the cost of skipping over irrelevant pages within redundant quadrants using the look-ahead pointers. The index construction (Algorithm 3) remains unchanged except for the fact that one would utilize a smaller α value for Eq. 5 in Line 3 of Algorithm 3 when used along with look-ahead pointers.

6 EXPERIMENTS

In this section, we first describe the setup of our experiments. Then, we present the experimental results with discussions.

6.1 Baselines

We compare the learned workload-aware Z-Index WAZI with the following competitors:

- 1) Sort Tile Recursive R-tree (STR) [25]: a simple R-tree packing method based on tiling the data space into vertical or horizontal slices recursively to construct an R-tree.
- 2) Cost-based unbalanced R-trees (CUR) [40]: a query-aware unbalanced R-tree construction algorithm which places nodes higher in the tree based on the expected number of access under a given workload.
- 3) Flood (FLOOD) [29]: a grid-based index built to efficiently process range queries by optimizing the grid structure based on the estimated query processing cost.
- 4) QUery-Aware Spatial Incremental Index (QUASII) [35]: a query-aware index which applies cracking [19] during query processing to adapt the index to the query workload.
- 5) Base Z-Index (BASE): the base Z-index built by partitioning the data points using the median along each axis and “ABCD” ordering of children at each level as presented in Section 3.

We also implemented the greedy variant of Qd-tree (QD-GR) [45] as the RL-variant requires a large action space which is infeasible in our setting. The cost formulation of QD-GR creates unbalanced trees tailored for disk-based indexes. Additionally,

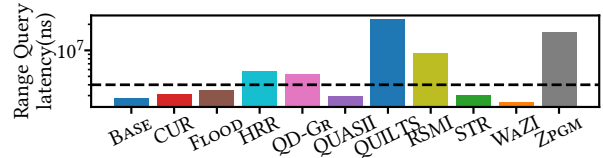


Figure 4: Average range query performance of all indexes considered in the experiments.

Table 1: Key properties of indexes in the experiments.

Index	SFC-based	Query-Aware	Learned
STR	✗	✗	✗
CUR	✗	✓	✗
FLOOD	✗	✓	✓
QUASII	✗	✓	✗
BASE	✓	✗	✗
WAZI	✓	✓	✓

we experimented with ZPGM [10, 42], HRR [37, 38], QUILTS [31] and RSMT [36]. However, these five indexes performed significantly worse than the other baselines considered (as shown in Figure 4). Therefore, we chose to not include them in more detailed experimentation later. Interestingly, four out of the five indexes we discarded (ZPGM, HRR, QUILTS, and RSMT) are based on Z-order curves within the rank space. Additionally, we also attempted to use BMTREE [26] and LMSFC [15] as baselines in our experiments. However, BMTREE construction failed to finish in a reasonable duration and LMSFC does not have publicly available implementation.

The indexes considered show a balanced representation along three different categorizations, as presented in Table 1. The ‘SFC-based’ column indicates if the construction algorithm of the indexes utilized space filling curves. The ‘Query-Aware’ column indicates if the index construction tailors for a specific query workload. Finally, the ‘Learned’ column refers to indexes that utilize any form of learning algorithms (to model cost function, density estimates, etc.) within the index or during construction.

Implementation. BASE and WAZI are constructed with specific algorithms for finding their respective partition point and ordering. The BASE version utilizes the naive method of comparing bounding boxes to decide if a given page overlaps with the range query before filtering points within the page. Whereas WAZI utilizes the efficient skipping mechanism mentioned in Section 5. We implement the two variants of Z-index in C++.¹

Similarly, we also implement STR [25] and QUASII [35] as described by the original authors. We utilize a converged QUASII index for our evaluations, i.e., an index that has completely adapted to the query distribution without the need for further incremental updates. A simplified FLOOD-Index for 2D indexing (FLOOD) is implemented in C++, where we identify the optimal grid structure by evaluating the performance of candidate grid partitions on a sub-sample of range queries. We adapted the CUR tree index to point data by using a ‘weighted’ RFDE estimator to hold a weighted sum of points covering each node. Each point is weighted by the number of distinct queries fetching the point. Finally, we use weighted density estimates to select partitions following the Sort Tile Recursive algorithm to construct CUR.

¹Our implementation is published at <https://version.helsinki.fi/ads/learned-zindex>.

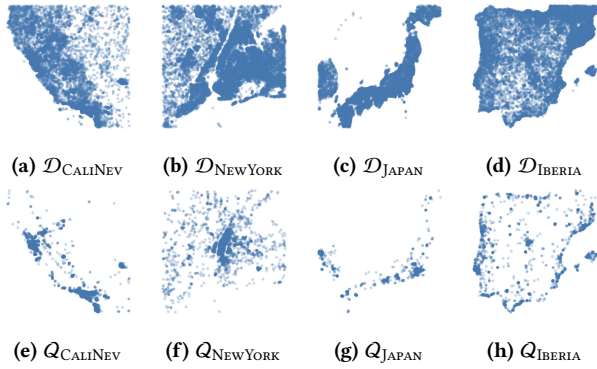


Figure 5: Datasets and query workloads in the experiments.

We conducted all the experiments on a server with an Intel[®] Xeon[®] E5-2680v4 CPU @ 2.40GHz and 16 GB memory. The experiments were performed on a single thread without multithreading or GPUs. All methods were compiled with the `-O3` optimizer flag. Finally, we set the leaf node capacity L to 256 in our experiments.

6.2 Datasets and Query Workloads

We utilize real-world datasets along with skewed semi-synthetic query workloads. The real-world datasets consist of points of interest (POI) from OpenStreetMap [32] for selected regions. The four regions we chose are: California Coast (CALINEV), New York City (NEWYORK), Japan (JAPAN), and Iberian Peninsula (IBERIA). We extract all POIs from the regions and sample these to produce datasets of appropriate size when required. The distribution of data points \mathcal{D} are presented in Figures 5a–5d.

We focus on skewed query workloads varying from the underlying data distribution for insightful analysis. This setup differs from past works showcasing experiments that use query distributions that follow the data distribution [36]. Alternatively, we present results on query workloads with real-world inspired skewness on the underlying data distribution. We generate semi-synthetic query workloads for the real-world datasets by utilizing check-in information in the Gowalla dataset [24]. Specifically, we extract check-ins that lie within the corresponding region and utilize these locations to generate a query workload resembling the distribution of check-in locations. The process for generating range queries proceeds by sampling the centers of query rectangles from the set of check-in locations and growing along the four directions such that the query covers a portion of data space equivalent to the required selectivity. We represent selectivity as a percentage of data space. Figure 5 presents the query distribution Q (i.e., the check-in locations) for each of the real-world datasets. The distributions of check-in locations are skewed towards popular locations compared to the underlying data distribution \mathcal{D} . We experiment with datasets ranging from 4 to 64 million data points and range query workloads of varying selectivity (ranging from 0.0004% to 0.1024% of data space). The range of parameters used for data generation are presented in Table 2.

6.3 Range Query Performance

Figure 6 presents the average range query latencies of different indexes at four levels of query selectivities (low to high). These index latencies are evaluated over range-query workloads of 20,000 queries. We observe that WAZI clearly outperforms all other indexes in most cases. The closest competition for WAZI

Table 2: Parameter setting.

Parameter	Values (default in bold)
Dataset size ($\times 10^6$)	[4, 8, 16, 32 , 64]
Query selectivity (%)	[0.0016, 0.0064, 0.0256 , 0.1024]
Leaf-node size	256
Range-query workload size	20,000

is from QUASII on the JAPAN dataset for the selectivity values 0.0064% (top-right in Figure 6) and 0.1024% (bottom-right). Nevertheless, WAZI always performs similarly or (often substantially) better than QUASII. The improvements are most noticeable for the lower selectivity queries where WAZI is better by about 2.3 \times to 8.1 \times over QUASII. We also note that the indexes showcase different scales of range query latencies over different datasets.

To further analyze the advantages of WAZI over BASE, we present a percentage improvement plot in Figure 7. The percentages are calculated in comparison with the performance of BASE for each dataset-workload pair. We present the average improvements of indexes over different datasets and selectivities. Across different datasets, WAZI showcases improvements between 12.8% for the JAPAN dataset and 47.6% for the NEWYORK dataset (top in Figure 7). CUR and STR showcase improvements over BASE for two out of the four datasets. We also see that the percentage differences between indexes decrease as the selectivity increases. The improvements by WAZI linearly decrease from 31% for 0.0016% selectivity queries to 14% for 0.1024% selectivity queries. This indicates that, as the selectivity increases, query processing encounters fewer false positives relative to the actual query result size. Therefore, a learned and workload-aware index is more suited for settings with low-selectivity queries. Finally, WAZI is the only index that shows a consistent improvement over BASE.

Range Query over Different Dataset Sizes. Additionally, we plot the average range query latencies in Figure 8 to analyze the performance of different indexes across varying dataset sizes. The query latencies are recorded for a query workload with a selectivity of 0.0256%. We observe that the performance of all indexes scales nearly linearly with the dataset size, with WAZI outperforming all competing baselines over different data sizes. Furthermore, the difference in performance between the competitors and WAZI remains somewhat constant with an increasing dataset size.

Projection vs Scanning. It is useful to analyze the query processing latency of an index in two parts. In the first part, termed Projection, the index traverses its search structure to identify all the leaf nodes (and hence the pages) overlapping a given range query. Second, we scan the data points within the projected pages to filter the result set. Such analysis separates the performance overhead of the internal search structure from the performance overheads originating from the data layout of the constructed index.

Figure 9 shows the two parts of query time for a 32 million-sized dataset under a mid-selectivity query workload. In terms of projection (top in Figure 9), CUR, QUASII, and STR perform worst due to its tree-traversal cost, followed by Z-indexes BASE and WAZI. WAZI performs 4.3 \times faster projection than BASE. This speedup could mostly be attributed to the efficient skipping mechanism. FLOOD performs the fastest projection (about 20 \times faster than BASE) of the query to pages, as it does not perform a tree

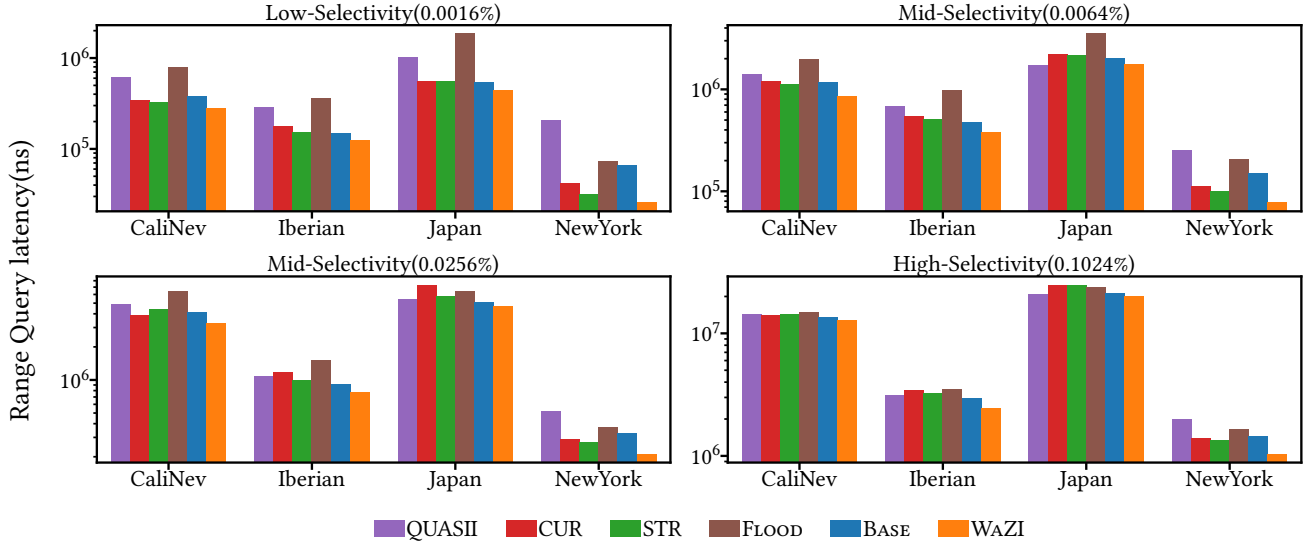


Figure 6: Average range query latency for all the indexes over four selectivity ranges. We observe that WAZI consistently outperforms all the competing baselines.

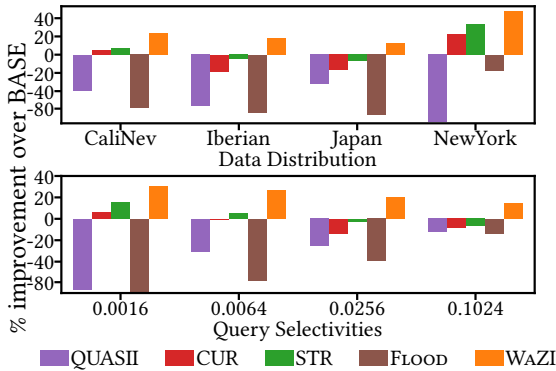


Figure 7: Percentage improvements over BASE across the different data distributions and selectivities (plotted on the symlog scale with a threshold of 50.)

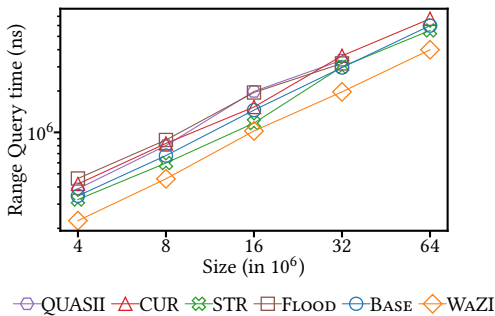


Figure 8: Range query time by varying the dataset size from 4 to 64 million for a mid-selectivity (0.0256%) query workload.

traversal for projection like the other indexes. The scan phase of the indexing dominates the performance of an index. WAZI outperforms other indexes in terms of scan time due to its optimized data layout.

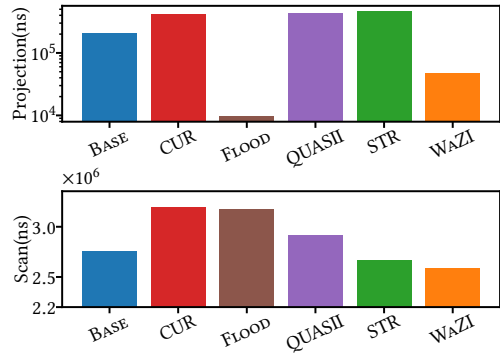


Figure 9: Range query latencies of different indexes split into the Projection and Scan phases.

Remark on kNN and Spatial-Join Queries. For spatial indexes that are not specialized to kNN or spatial-join queries (as is the case for all indexes we evaluate), kNN and spatial-join queries are typically decomposed to and processed as sets of range queries [34]. As a result, query performance exhibits the same behavior as their range query performance presented above, as we have confirmed experimentally. For this reason, we do not include additional results for kNN and spatial-join queries.

6.4 Point Query Performance

Although we focus on optimizing indexes for range-query workloads, we also show the performance of proposed indexes for point queries for completeness. We sample 50,000 point queries from the data distribution \mathcal{D} of the respective datasets. Figure 10 presents the average query time (over different datasets) for various indexes against increasing data size. Comparing indexes from best to worst, we see that WAZI and BASE outperform all other baselines. This performance is due to simpler computations required at the internal nodes of a Z-index (see Algorithm 1) to identify the relevant child nodes. FLOOD index performs similarly but slower than WAZI and BASE. STR and CUR have non-overlapping

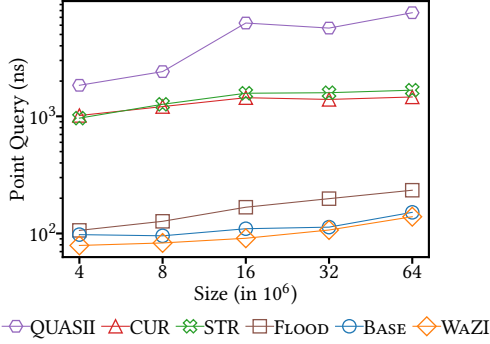


Figure 10: Point query time by varying the dataset size from 4 to 64 million. We observe that WAZI performs the best across all dataset sizes.

Table 3: Build time (in seconds) of all indexes compared.

Size ($\times 10^6$)	BASE	CUR	FLOOD	QUASII	STR	WAZI
4	7.4	23.8	4.6	123.6	1.8	25.1
8	16.2	48.5	10.1	418.1	4.0	51.8
16	35.0	97.0	25.8	4400.7	8.4	105.3
32	71.6	188.5	42.1	15781.5	19.1	203.2
64	154.6	366.9	97.3	57943.8	40.5	414.3

child nodes and perform marginally faster than CUR which allows for overlapping. Finally, QUASII suffers computationally due to the heavily fractured data layout caused by database cracking-based index construction. Generally, the performance of each index over point queries is highly correlated with its respective cost of tree traversal and the underlying data layout. We see that WAZI performs better than the baselines by a factor of 1.5–15 \times due to the efficiency of node-level computations.

6.5 Build Time

We record the build time for all the indexes we compare in Table 3. For all indexes, the build time grows linearly as the dataset size increases. STR takes the least time for index construction as the Sort-tile recursive algorithm has a log-linear complexity. Among the query-aware indexes (see Table 1), WAZI performs third best behind FLOOD and CUR. On comparison against learned indexes, WAZI performs second best behind FLOOD. WAZI requires 5.4 \times more build time than FLOOD. We mainly attribute this to the need for constructing and evaluating data density estimators required for cost computations. The build-time of WAZI is similar to the query-aware CUR index which utilizes query selectivity estimation of data points to compute an approximate cost. Therefore, compared to the baselines considered, WAZI has a reasonable build time.

Cost-redemption. Typically, learned indexes achieve improved query execution time, at the cost of computationally expensive index building. To quantify this trade-off for a given dataset and query workload, we define *cost redemption* as the number of query executions for which the cumulative (building+querying) time of an index is equal to the respective cumulative time of BASE. Specifically, the cost-redemption value of index X is defined as $red_X = \frac{X.Build - BASE.Build}{BASE.Query - X.Query}$.

Table 4: Cost-redemption value of indexes against BASE. Lower is better. Missing values indicate that the index does not redeem its high-construction time with efficient query performance compared to BASE.

Data Dist.	CUR	FLOOD	QUASII	STR	WAZI
CALINEV	(+) 550k	(-) 13k	(-)	(-) 320k	(+) 192k
IBERIA	(-)	(-) 44k	(-)	(-) 450k	(+) 746k
JAPAN	(-)	(-) 6.5k	(-)	(-) 69k	(+) 269k
NEWYORK	(+) 1300k	(-) 830k	(-)	(+)	(+) 743k

Table 5: Sizes (in MBs) of all indexes compared.

Size ($\times 10^6$)	BASE	CUR	FLOOD	QUASII	STR	WAZI
4	8.27	10.33	5.39	3.02	7.09	8.99
8	16.83	19.05	6.14	5.31	15.41	17.80
16	34.94	36.91	11.00	13.79	23.48	35.15
32	67.60	74.73	37.18	26.83	66.02	69.84
64	135.86	149.01	53.13	61.48	112.00	138.83

We showcase the cost-redemption of the considered indexes in Table 4. Cells marked with (+) correspond to cases where the index eventually redeems its higher building time after the reported number of queries. By contrast, cells marked with (-) correspond to the opposite behavior: they have faster building time than BASE, but larger query execution time, leading to larger cumulative time after the reported number of queries. Finally, a missing value indicates that the index has better (+) or worse (-) cumulative time regardless of the number of queries executed. From Table 4 we can see that FLOOD and STR are better than BASE in the short run due to their smaller build times. CUR is better than BASE in the longer run for CALINEV and NEWYORK datasets while being clearly worse for IBERIA and JAPAN. The massive build time of QUASII and the slower query performance indicate that it is the worst-performing index. Finally, WAZI is a better choice than BASE if the expected query workload has at least 10^6 queries. Therefore, this is a limitation of WAZI as it is not suited for workflows that require quick index construction to execute a few queries as found in data analysis workflows. Instead, it is suited for workflows where index construction can be performed offline using more computational resources and deployed for an extended amount of time.

6.6 Index Size

We present the size of each index built for the datasets with 32 million points in Table 5. QUILTS, FLOOD and QUASII exhibit 0.31 – 0.59 \times smaller index sizes than WAZI. The index size of WAZI is similar to the R-tree-based indexes, CUR and STR. Crucially note that the size of WAZI is nearly identical to BASE. This indicates that the methods presented in this paper construct a workload-aware version of Z-index at no additional space cost. Furthermore, we see that the size of all the indexes we utilized in our experiment grows linearly with the dataset size.

6.7 Index Update

WAZI supports index updates like other tree-based indexes. An index update operation in WAZI is performed similarly to point query processing. Inserting/deleting a point p proceeds by finding the leaf node enclosing p (using Algorithm 1) and updating the

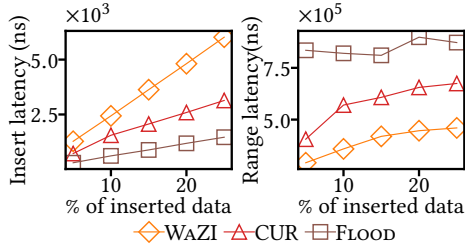


Figure 11: Insert and range query latencies over inserts.

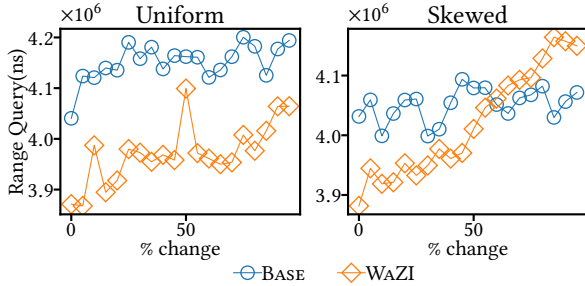


Figure 12: Range query time of WAZI and BASE over the changes in workloads. On the left, higher % change corresponds to a more uniform workload; on the right, to a workload more similar to a different skewed workload.

corresponding page. When any page overflows/underflows, we perform a node splitting/merging operation as for any tree index.

Figure 11 examines the impact of insert queries on a subset of indexes each containing 32 million data points. We sample 8 million insert query points uniformly from the data space and incrementally insert these points in five equal-sized iterations. For each iteration, we recorded the latencies of insert queries and range queries after insertions. We split any overflowing pages of WAZI along the data medians. We see that WAZI performs slowest in terms of insert latencies in comparison with CUR and FLOOD (left in Figure 11). This can be attributed to the costly update to PAGELIST during node splits and the need to recompute the look-ahead pointers. We also note that the inserts have a minor effect on range query performance, with latencies degrading logarithmically with inserts.

6.8 Effect of Workload Changes

Workload-aware indexes are susceptible to degrading query performance due to the changes in workloads. In this subsection, we compare the effects of query workload changes on BASE and WAZI. Specifically, we build BASE and WAZI using their original query workload and evaluate them over iteratively changing workloads. We experiment with both uniform and skewed workload changes. In the former case, we replace the dataset’s original workload with uniformly sampled queries over the data space. In the latter case, we replace the dataset’s original workload with the (skewed) workloads of other datasets.

The results of the experiment are presented in Figure 12. We see that the performance of BASE remains relatively stable across the range of workload changes, while WAZI degrades in performance. As the workload changes to a uniform workload (left in Figure 12), WAZI degrades gracefully and remains better than BASE. This shows that WAZI’s skipping mechanism and the layout

which is tailored to both the data and the original query workload (see Eq. 5) retains some advantage over the BASE index even over uniform queries. The weakness of WAZI is highlighted in the case where the workload changes to a differently skewed workload (right in Figure 12). The performance of WAZI degrades rapidly and becomes slower than BASE after the workload changes more than 60%. This indicates that WAZI is most suited for the settings with known query workloads which do not change drastically and it requires to be rebuilt if the workload changes sufficiently.

To our knowledge, automatically deciding when to rebuild a spatial index in case of workload drift is an open problem. However, there have been extensive studies on detecting concept drift for ML models [14]. We suggest that similar methods could be adapted for rebuilding spatial indexes. However, the exploration of such methods is beyond the scope of the paper.

6.9 Ablation Study

In this paper, we presented methods to perform adaptive partitioning and ordering (Section 4.1) and a novel skipping algorithm (Section 5) to create the WAZI index. Here we perform an ablation study for the two methods we presented. To facilitate comparison of each individual method mentioned above, we compare the BASE and WAZI index against two additional methods:

- BASE+SK: A BASE index variant with default partitioning upon which we construct and utilize the look-ahead pointer logic.
- WAZI-SK: A WAZI variant with adaptive partitioning, but no look-ahead pointers.

We perform experiments on a dataset of size 8 million and log metrics of interest. The results of the experiment are presented in Figure 13. First, we show the incremental improvements made against BASE by the additional methods in Figure 13 (top-left). We see that both WAZI-SK and BASE+SK perform similarly for low selectivity indicating both the methods contribute equally towards the improved performance. However, as selectivity increases, WAZI-SK and BASE+SK diverge. At higher selectivity workloads, BASE+SK tends towards the performance of BASE index while WAZI-SK tends towards WAZI. This trend shows that adaptive partitioning improves performance by a larger factor than the use of look-ahead pointers at high selectivity ranges.

Second, the bottom-left plot of Figure 13 shows the number of bounding boxes (*bbs*) checked for overlap with the range query while processing leaf nodes in the [LOW : HIGH] interval. We clearly see that the use of look-ahead pointers drastically reduces the number of bounding boxes compared. Indexes with look-ahead pointers perform 50–100× fewer bounding box comparisons.

Third, we show the number of excess points compared (top-right) and the number of pages scanned (bottom-right). The underlying data-layout of an index greatly impacts the performance on these two metrics. We see that BASE and BASE+SK perform worse than WAZI and WAZI-SK indicating that the adaptive partitioning is the key differentiating factor. Therefore, we infer that adaptive partitioning creates data layouts that are more efficient for range query processing in a Z-index.

From these results, we can conclude that the two methods we present in this paper address two key aspects of range query performance in a Z-index. The adaptive partitioning generates better data layouts and the look-ahead pointers facilitate quicker traversal of the layout.

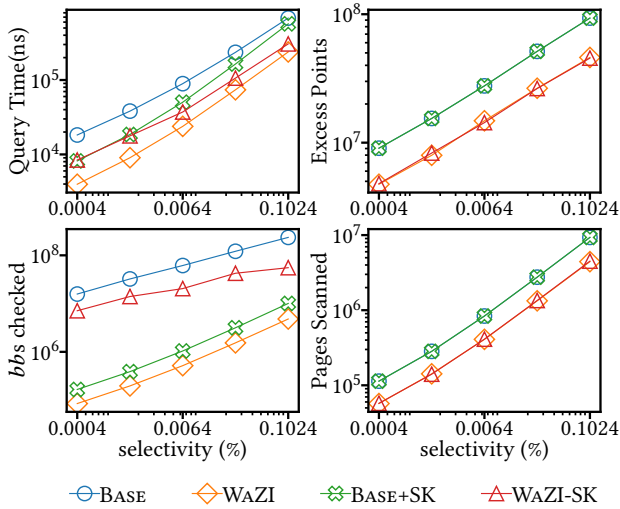


Figure 13: Results for ablation study regarding query time (top-left), excess retrievals (top-right), bounding boxes checked (bottom-left), and pages scanned (bottom-right).

7 CONCLUSION

In this paper, we presented WAZI, a learned and workload-aware variant of a Z-index that jointly optimizes the storage layout and search structures for spatial indexing. Our method adapts to data distribution and query workload to build indexes to provide better range query performance. We began by devising a metric for optimization based on the number of data points fetched during query processing, called retrieval cost. Then we proposed an algorithm for building a Z-index variant, WAZI, which minimizes the retrieval cost. Additionally, we designed efficient mechanisms for WAZI to skip irrelevant pages accessed during range query processing. Through extensive experiments, we demonstrated that WAZI improved upon existing spatial indexes, such as FLOOD and CUR, in terms of range query performance for real-world datasets. We presented a comprehensive ablation study, where we showed the relative contribution of each component towards improvements shown by WAZI. Additionally, we see that WAZI performs well on point query performance while providing favorable construction time and index size tradeoffs.

For future work, we intend to explore three aspects that are not yet covered by this paper. First, we intend to expand the cost formulation to account for other factors of index performance beyond retrieval cost so as to fit mixed types of query workloads. Second, we intend to explore ways to solve for the recursive cost formulation briefly mentioned in Section 4.3. Specifically, we aim to use machine learning models to efficiently learn the retrieval cost, which currently requires a dynamic programming algorithm with a state space of $O(N^4)$. Third, we intend to develop mechanisms to decide when to retrain an index when the data distribution and query workload change.

ACKNOWLEDGMENTS

We would like to thank Arpit Merchant and Ananth Mahadevan for their useful suggestions regarding the presentation of experimental results. Sachith Pai and Michael Mathioudakis are supported by University of Helsinki and Academy of Finland Projects MLDB (322046) and HPC-HD (347747). Yanhao Wang

was supported by the National Natural Science Foundation of China under grant no. 62202169.

REFERENCES

- [1] Abdullah Al-Mamun, Ch. Md. Rakin Haider, Jianguo Wang, and Walid G. Aref. 2022. The “AI + R”-tree: An Instance-optimized R-tree. In *2022 23rd IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 9–18.
- [2] Lars Arge, Mark de Berg, Herman J. Haverkort, and Ke Yi. 2004. The Priority R-Tree: A Practically Efficient and Worst-Case Optimal R-Tree. In *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data (SIGMOD '04)*. Association for Computing Machinery, New York, NY, USA, 347–358.
- [3] Rudolf Bayer. 1997. The Universal B-Tree for Multidimensional Indexing: General Concepts. In *Worldwide Computing and Its Applications – International Conference, WWCA '97, Tsukuba, Japan, March 10-11, 1997 Proceedings*. Springer, Berlin, Heidelberg, 198–209.
- [4] Norbert Beckmann, Hans-Peter Kriegel, Ralf Schneider, and Bernhard Seeger. 1990. The R*-Tree: An Efficient and Robust Access Method for Points and Rectangles. In *Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data (SIGMOD '90)*. Association for Computing Machinery, New York, NY, USA, 322–331.
- [5] Norbert Beckmann and Bernhard Seeger. 2009. A Revised R*-Tree in Comparison with Related Index Structures. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data (SIGMOD '09)*. Association for Computing Machinery, New York, NY, USA, 799–812.
- [6] Jon Louis Bentley. 1975. Multidimensional Binary Search Trees Used for Associative Searching. *Commun. ACM* 18, 9 (1975), 509–517.
- [7] Jialin Ding, Umar Farooq Minhas, Jia Yu, Chi Wang, Jaeyoung Do, Yinan Li, Hantian Zhang, Badrish Chandramouli, Johannes Gehrke, Donald Kossmann, David B. Lomet, and Tim Kraska. 2020. ALEX: An Updatable Adaptive Learned Index. In *Proceedings of the 2020 International Conference on Management of Data (SIGMOD '20)*. Association for Computing Machinery, New York, NY, USA, 969–984.
- [8] Jialin Ding, Vikram Nathan, Mohammad Alizadeh, and Tim Kraska. 2020. Tsunami: A Learned Multi-dimensional Index for Correlated Data and Skewed Workloads. *Proc. VLDB Endow.* 14, 2 (2020), 74–86.
- [9] Haowen Dong, Chengliang Chai, Yuyu Luo, Jiabin Liu, Jianhua Feng, and Chaoqun Zhan. 2022. RW-Tree: A Learned Workload-aware Framework for R-tree Construction. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*. IEEE, 2073–2085.
- [10] Paolo Ferragina and Giorgio Vinciguerra. 2020. The PGM-Index: A Fully-Dynamic Compressed Learned Index with Provable Worst-Case Bounds. *Proc. VLDB Endow.* 13, 8 (2020), 1162–1175.
- [11] Raphael A. Finkel and Jon Louis Bentley. 1974. Quad Trees: A Data Structure for Retrieval on Composite Keys. *Acta Inform.* 4 (1974), 1–9.
- [12] Volker Gaede and Oliver Günther. 1998. Multidimensional Access Methods. *ACM Comput. Surv.* 30, 2 (1998), 170–231.
- [13] Alex Galakatos, Michael Markovitch, Carsten Binnig, Rodrigo Fonseca, and Tim Kraska. 2019. FITing-Tree: A Data-aware Index Structure. In *Proceedings of the 2019 International Conference on Management of Data (SIGMOD '19)*. Association for Computing Machinery, New York, NY, USA, 1189–1206.
- [14] João Gama, André Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. 2014. A survey on concept drift adaptation. *ACM Comput. Surv.* 46, 4 (2014), 44:1–44:37.
- [15] Jian Gao, Xin Cao, Xin Yao, Gong Zhang, and Wei Wang. 2023. LMSFC: A Novel Multidimensional Index based on Learned Monotonic Space Filling Curves. *Proc. VLDB Endow.* 16, 10 (2023), 2605–2617.
- [16] Tu Gu, Kaiyu Feng, Gao Cong, Cheng Long, Zheng Wang, and Sheng Wang. 2023. The RLR-Tree: A Reinforcement Learning Based R-Tree for Spatial Data. *Proc. ACM Manag. Data* 1, 1, Article 63 (2023), 26 pages.
- [17] Antonin Guttman. 1984. R-Trees: A Dynamic Index Structure for Spatial Searching. In *Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data (SIGMOD '84)*. Association for Computing Machinery, New York, NY, USA, 47–57.
- [18] Ali Hadian and Thomas Heinis. 2021. Shift-Table: A Low-latency Learned Index for Range Queries using Model Correction. In *Proceedings of the 24th International Conference on Extending Database Technology (EDBT)*. OpenProceedings.org, 253–264.
- [19] Stratos Idreos, Martin L. Kersten, and Stefan Manegold. 2007. Database Cracking. In *3rd Biennial Conference on Innovative Data Systems Research (CIDR)*. cidrdb.org, 68–78.
- [20] Ibrahim Kamel and Christos Faloutsos. 1994. Hilbert R-tree: An Improved R-tree using Fractals. In *Proceedings of 20th International Conference on Very Large Data Bases (VLDB '94)*. Morgan Kaufmann, San Mateo, CA, USA, 500–509.
- [21] Andreas Kipf, Ryan Marcus, Alexander van Renen, Mihail Stoian, Alfons Kemper, Tim Kraska, and Thomas Neumann. 2020. RadixSpline: A Single-Pass Learned Index. In *Proceedings of the Third International Workshop on Exploiting Artificial Intelligence Techniques for Data Management (aiDM@SIGMOD '20)*. Association for Computing Machinery, New York, NY, USA, 5:1–5:5.
- [22] Tim Kraska, Alex Beutel, Ed H. Chi, Jeffrey Dean, and Neoklis Polyzotis. 2018. The Case for Learned Index Structures. In *Proceedings of the 2018 International Conference on Management of Data (SIGMOD '18)*. Association for Computing

- Machinery, New York, NY, USA, 489–504.
- [23] Jonathan K. Lawder and Peter J. H. King. 2000. Using Space-Filling Curves for Multi-dimensional Indexing. In *Advances in Databases – 17th British National Conference on Databases, BNCOD 17 Exeter, UK, July 3-5, 2000 Proceedings*. Springer, Berlin, Heidelberg, 20–35.
 - [24] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. <http://snap.stanford.edu/data>.
 - [25] Scott T. Leutenegger, J. M. Edgington, and Mario Alberto López. 1997. STR: A Simple and Efficient Algorithm for R-Tree Packing. In *Proceedings of the 13th International Conference on Data Engineering (ICDE)*. IEEE, 497–506.
 - [26] Jiangneng Li, Zheng Wang, Gao Cong, Cheng Long, Han Mao Kiah, and Bin Cui. 2023. Towards Designing and Learning Piecewise Space-Filling Curves. *Proc. VLDB Endow.* 16, 9 (2023), 2158–2171.
 - [27] Pengfei Li, Hua Lu, Qian Zheng, Long Yang, and Gang Pan. 2020. LISA: A Learned Index Structure for Spatial Data. In *Proceedings of the 2020 International Conference on Management of Data (SIGMOD '20)*. Association for Computing Machinery, New York, NY, USA, 2119–2133.
 - [28] Moin Hussain Moti, Panagiotis Simatis, and Dimitris Papadias. 2022. Waffle: A Workload-Aware and Query-Sensitive Framework for Disk-Based Spatial Indexing. *Proc. VLDB Endow.* 16, 4 (2022), 670–683.
 - [29] Vikram Nathan, Jialin Ding, Mohammad Alizadeh, and Tim Kraska. 2020. Learning Multi-Dimensional Indexes. In *Proceedings of the 2020 International Conference on Management of Data (SIGMOD '20)*. Association for Computing Machinery, New York, NY, USA, 985–1000.
 - [30] Jürg Nievergelt, Hans Hinterberger, and Kenneth C. Sevcik. 1984. The Grid File: An Adaptable, Symmetric Multikey File Structure. *ACM Trans. Database Syst.* 9, 1 (1984), 38–71.
 - [31] Shoji Nishimura and Haruo Yokota. 2017. QUILTS: Multidimensional Data Partitioning Framework Based on Query-Aware and Skew-Tolerant Space-Filling Curves. In *Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD '17)*. Association for Computing Machinery, New York, NY, USA, 1525–1537.
 - [32] OpenStreetMap 2017. OpenStreetMap. <https://www.openstreetmap.org>.
 - [33] Sachith Pai, Michael Mathioudakis, and Yanhao Wang. 2022. Towards an Instance-Optimal Z-Index. In *4th International Workshop on Applied AI for Database Systems and Applications (AIDB '22)*. 4 pages.
 - [34] Varun Pandey, Alexander van Renen, Eleni Tziritza Zacharitou, Andreas Kipf, Ibrahim Sabek, Jialin Ding, Volker Markl, and Alfons Kemper. 2023. Enhancing In-Memory Spatial Indexing with Learned Search. *CoRR abs/2309.06354* (2023). [arXiv:2309.06354](https://arxiv.org/abs/2309.06354)
 - [35] Mirjana Pavlovic, Darius Sidlauskas, Thomas Heinis, and Anastasia Ailamaki. 2018. QUASII: QUery-Aware Spatial Incremental Index. In *Proceedings of the 21st International Conference on Extending Database Technology (EDBT)*. OpenProceedings.org, 325–336.
 - [36] Jianzhong Qi, Guanli Liu, Christian S. Jensen, and Lars Kulik. 2020. Effectively Learning Spatial Indices. *Proc. VLDB Endow.* 13, 11 (2020), 2341–2354.
 - [37] Jianzhong Qi, Yufei Tao, Yanchuan Chang, and Rui Zhang. 2018. Theoretically Optimal and Empirically Efficient R-trees with Strong Parallelizability. *Proc. VLDB Endow.* 11, 5 (2018), 621–634.
 - [38] Jianzhong Qi, Yufei Tao, Yanchuan Chang, and Rui Zhang. 2020. Packing R-trees with Space-filling Curves: Theoretical Optimality, Empirical Efficiency, and Bulk-loading Parallelizability. *ACM Trans. Database Syst.* 45, 3 (2020), 14:1–14:47.
 - [39] Raghu Ramakrishnan and Johannes Gehrke. 2003. *Database Management Systems* (3rd ed.). McGraw-Hill, New York, NY, USA.
 - [40] Kenneth A. Ross, Inga Sitzmann, and Peter J. Stuckey. 2001. Cost-based Unbalanced R-Trees. In *Proceedings of the 13th International Conference on Scientific and Statistical Database Management (SSDBM)*. IEEE, 203–212.
 - [41] Timos K. Sellis, Nick Roussopoulos, and Christos Faloutsos. 1987. The R+-Tree: A Dynamic Index for Multi-Dimensional Objects. In *Proceedings of 13th International Conference on Very Large Data Bases (VLDB '87)*. Morgan Kaufmann, San Mateo, CA, USA, 507–518.
 - [42] Herbert Tropf and Helmut Herzog. 1981. Multidimensional Range Search in Dynamically Balanced Trees. *Angew. Info.* 2 (1981), 71–77.
 - [43] Haixin Wang, Xiaoyi Fu, Jianliang Xu, and Hua Lu. 2019. Learned Index for Spatial Queries. In *2019 20th IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 569–574.
 - [44] Hongwei Wen and Hanyuan Hang. 2022. Random Forest Density Estimation. In *Proceedings of the 39th International Conference on Machine Learning*. PMLR, 23701–23722.
 - [45] Zongheng Yang, Badrish Chandramouli, Chi Wang, Johannes Gehrke, Yinan Li, Umar Farooq Minhas, Per-Åke Larson, Donald Kossmann, and Rajeev Acharya. 2020. Qd-tree: Learning Data Layouts for Big Data Analytics. In *Proceedings of the 2020 International Conference on Management of Data (SIGMOD '20)*. Association for Computing Machinery, New York, NY, USA, 193–208.
 - [46] Songnian Zhang, Suprio Ray, Rongxing Lu, and Yandong Zheng. 2021. SPRIG: A Learned Spatial Index for Range and kNN Queries. In *Proceedings of the 17th International Symposium on Spatial and Temporal Databases (SSTD '21)*. Association for Computing Machinery, New York, NY, USA, 96–105.