On Indexing Mobile Objects

George Kollios, Polytechnic University
Dimitrios Gunopulos, UC Riverside
Vassilis J. Tsotras, UC Riverside
gkollios@cs.ucr.edu, dg@cs.ucr.edu, tsotras@cs.ucr.edu

Technical Report UCR-CS-98-06
University of California, Riverside
Department of Computer Science

Abstract

We show how to index mobile objects in one and two dimensions using efficient dynamic external memory data structures. The problem is motivated by real life applications in traffic monitoring, intelligent navigation and mobile communications domains. For the 1-dimensional case, we give a dynamic, external memory algorithm with guaranteed worst case performance and linear space and a practical approximation algorithm also in the dynamic, external memory setting, which has linear space and expected logarithmic query time. We also give an algorithm with guaranteed logarithmic query time for a restricted version of the problem. We present extensions of our techniques to two dimensions. In addition we give a lower bound on the number of I/O's needed to answer the d-dimensional problem. Initial experimental results and comparisons to traditional indexing approaches are also included.

1 Introduction

Traditional database management systems assume that data stored in the database remain constant until explicitly changed through an update. While this model serves well many applications where data changes in discrete steps, it is not appropriate for applications with continuously changing data. One such application is a "motion" database that stores the location of mobile objects (e.g. cars). Since objects change location continuously, one would have to update the database at every unit of time. This is clearly an inefficient and infeasible solution considering the prohibitively large update overhead.

A better approach is to abstract each object's location as a function of time $f(t)$, and update the database only when the parameters of f change (for example when the velocity or the direction of a car changes). Using $f(t)$ the "motion" database can compute the location of the mobile object at any time in the future. While this approach minimizes the update overhead, it introduces a variety of novel problems (such as the need for appropriate data models, query languages and query processing and optimization techniques) since the database is not directly storing data values but functions to compute these values. Motion database problems have recently attracted the interest of the research community: ([37, 40, 41]) present the Moving Objects Spatio-Temporal (MOST) model and a language (FTL) for querying the current and future locations of mobile objects; ([15]) proposes a model that tracks and queries the history (past routes) of mobile objects, based on new spatio-temporal data types.

In this paper we focus on the problem of indexing mobile objects. In particular we examine how to efficiently address range queries over the object locations into the future. An example of such a spatiotemporal query is: "Report all the objects that will be inside a query region P after 10 minutes from now". Note that the answer to this query is tentative in the sense that it is computed based on the current knowledge stored in the database about the mobile objects' location functions. In the near future this knowledge may change, which implies that the same query could have a different answer. Spatiotemporal queries about mobile objects have important applications in traffic monitoring, intelligent navigation and mobile communications domains (for example we can allocate more bandwidth for areas where high concentration of mobile phones is approaching). As the number of mobile objects in these applications can be rather large we are interested in external memory solutions.

While in general an object could move anywhere in 3-dimensional space using some rather complex motion, we limit our treatment to objects moving in 1- and 2-dimensional spaces and whose location is described by a linear function of time. There is a strong motivation for such an approach based on the real-world applications we have in mind: straight lines are usually the faster way to get from one point to another; cars move in networks of highways which can be approximated by connected straight line segments on a plane; this is also true for routes taken by airplanes or ships. In addition, solving these simpler 1- and 2-dimensional problems may provide intuition for addressing the more difficult problem of indexing general multidimensional functions.

2 Problem Description

We consider a database that keeps track of mobile objects moving in one and two dimensions. We model the objects as points that move with a constant velocity starting from a specific location at a specific time instant. Using this information we can compute the location of the object at any time in the future for as long as its movement characteristics remain the same. In one dimension, an object started from location y_0 at time t_0 with a velocity v (v can be positive or negative) will be in location $y_0 + v(t - t_0)$ at time $t > t_0$. Similarly for objects moving in two dimensions. We assume that the objects can move inside a finite terrain (a line segment in one dimension or a rectangle in two). Thus when an object has reached the limits, it has to change its velocity. Objects are responsible to update their motion information, every time when the velocity or direction changes. Also we allow to insert a new object or to delete an old one, eg. the system is dynamic.

We would like to answer efficiently proximity queries among the mobile objects. In particular, we are interesting to answering queries of the form: "Report the objects that reside inside the interval $[y_{1q}, y_{2q}]$ (or the rectangle $[x_{1q}, x_{2q}] \times [y_{1q}, y_{2q}]$ in two dimensions) at the time instants between time t_{1q} and t_{2q}, (where $t_{now} \leq t_{1q} \leq t_{2q}$), given the current motion information of all objects". We call this type of query *mobile objects range (MOR) query.*

We consider the problem in the standard external memory model of computation[3]. In this model each disk access (an I/O) transmits in a single operation B units of data. We call B the page capacity. We measure the efficiency of an algorithm in terms of the number of I/O's to perform an operation. If

N is the number of the mobile objects and K is the number of objects reported by the MOR query, then the minimum number of pages to store the database is $n = \lceil \frac{N}{B} \rceil$ and the minimum number of I/O's to report the answer is $k = \lceil \frac{K}{B} \rceil$. We say that an algorithm uses linear space, if it uses $O(n)$ disk pages, and that it uses logarithmic time to answer a query if it needs to execute $O(\log_B(n) + k)$ I/O's.

3 Summary of results

In the next section we consider the problem of indexing mobile points on a line. First we present two geometric representations of the problem (sections 4.1, 4.2) and discuss why storing the trajectories of mobile objects as lines in traditional indexing methods (like R-trees) does not work. We then give a lower bound on the static version of the problem by reducing it to simplex range searching in two and four dimensions respectively (section 4.3). We also give external memory, dynamic algorithms with matching upper bounds that are based on partition trees [29] (section 4.4). Unfortunately these algorithms are not practical because they have large hidden constant factors. In addition, the lower bound for simplex range searching in two or more dimensions is $O(\sqrt{n})$ if we allow only linear storage. Accordingly we turn our attention into algorithms designed to use linear space and to work well in the average case. We present two approaches, a simple one based on *kd*-trees and a more sophisticated one based on $B+$-trees (section 4.5). Our experimental results (section 6) show that the $B+$-trees based approach outperforms the *kd*-tree and the simple R^*-tree approach. Only if we restrict the problem there is hope that we can achieve logarithmic query time and linear space. In section 4.6, we impose the restriction that we can only answer queries within a fixed time window in the future. For this setting we present a data structure with logarithmic query time. The space requirement of our method varies between linear and quadratic, depending on the size of the time window and the distribution of the velocities of the points.

In section 5 we extend the previous results for two versions of the 2-dimensional case. In the first version we make the assumption, motivated from practice, that the objects move on a network of 1-dimensional routes (we call this version, the 1.5-dimensional problem). In the second version the objects are allowed to move arbitrarily on a plane (the general 2-dimensional problem).

4 Indexing in one dimension

We begin with the simpler problem of objects moving on a 1-dimensional line. Firstly, we partition the mobile objects into two categories, the objects with low speed $v \approx 0$ and the objects with speed between a minimum v_{min} and maximum speed v_{max}. We consider here the "moving" objects, eg. the objects with speed greater than v_{min}. We discuss the case of slowly moving objects in section 4.6.

The problem is to index the mobile objects in order to efficiently answer range queries over their locations into the future. The location of each object is described as a linear function of time, namely the location $y_i(t)$ of the object o_i at time t is equal to $v_i(t - t_{i_0}) + y_{i_0}$, where v_i is the velocity of the object and y_{i_0} is its location at t_{i_0}. We assume that objects move on the $y-$axis between 0 and Y_{max} and that an object can update its motion information whenever it changes. We treat an update as a deletion of the old information and an insertion of the new one. Next we give different geometric representations of the problem and for each one we discuss access structures to efficiently address MOR queries.

4.1 Space-time representation

In this representation we plot the trajectories of the mobile objects as lines in the time-location (t, y) plane. The equation of each line is $y(t) = vt + a$ where v is the slope (the velocity in our case) and a is the intercept, that can be computed by the motion information. In fact a trajectory is not a line but a semi-line starting from the point (y_i, t_i). However since we ask queries for the present or for the future, assuming that the trajectory is a line does not affect the correctness of the answer. Figure1 shows a number of trajectories in the plane.

The query is expressed as a 2-dimensional interval $[(y_{1q}, y_{2q}), (t_{1q}, t_{2q})]$. The answer is the set of objects that correspond to lines that intersect the query rectangle.

While the space-time representation is quite intuitive, it leads to indexing long lines, a situation that causes significant shortcomings to traditional indexing techniques.

One way is to index the lines using a Spatial Access Method (SAM). For example, each line could be approximated by a minimum bounding rectangle (MBR) which is then indexed using a R-tree[20] or a R*-tree[8]. However, this approach is problematic. First, an MBR assigns to the mobile object

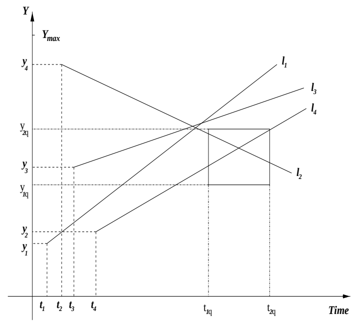

Figure 1: Trajectories and query in (t, y) plane.

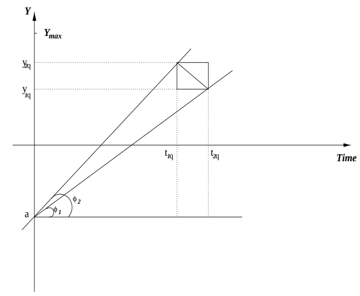

Figure 2: Query in the primal plane.

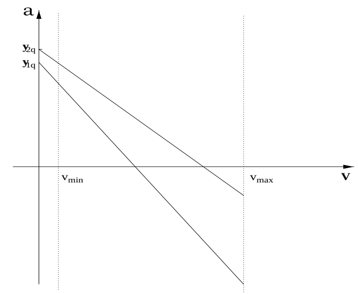

Figure 3: Query in the dual plane.

a much larger area than a line has, which implies that the index selectivity will be poor (due to the increased overlapping among MBRs). Second, the problem with SAMs is that they cannot represent infinite objects. Note that in our environment, objects retain their trajectory until being updated; this implies that all lines in Figure 1 extend to "infinity" on the time dimension. This common "ending" among lines will introduce additional overlapping among MBRs. Another traditional approach to deal with lines is to map each line segment as a point in four dimensions, by taking the coordinates of the end points. This may minimize the MBR overlapping but objects still share their common end point to "infinity". Even if we partition the time dimension in time intervals ("sessions") of length ΔT (as in [39]) and index the part of each trajectory that falls in the current session, we still have segments with a common endpoint (the end time of the current session). Also, in this approach, the SAM can only address queries until the end of the current session.

A different approach is to decompose the data space into disjoint cells and store with each cell the set of lines that intersect it (R+-tree[35], cell-tree[19], PMR-quadtree[34]). The main drawback for these methods is that every line can have many copies; this becomes worse in our environment since lines are really large. Storing many copies affects both the update performance (when an object changes its trajectory, its previous route has to be deleted from all cells it was contained), as well as space. [1]

In [25] a method is proposed to index line segments based on the dual transformation. In the next section we consider this approach in our setting, namely using the dual transformation to index mobile objects.

4.2 The dual space-time representation.

Duality is a powerful and useful transform frequently used in the computational geometry literature; in general it maps a hyper-plane h from R^d to a point in R^d and vice-versa. The duality transform is useful because it allows to formulate a problem in a more intuitive manner.

In our case we can map a line from the *primal* plane (t, y) to a point in the *dual* plane. There is no unique duality transform however, but a class of transforms with similar properties. Sometimes one transform is more convenient than another.

Consider a dual plane where one axis represents the slope of an object's trajectory and the other axis its intercept. Thus the line with equation $y(t) = vt + a$ is represented by the point (v, a) in the dual space (this is called the Hough-X transform in [25]). While the values of v are between $-v_{max}$ and v_{max}, the values of the intercept are depended on the current time. If the current time is t_{now} then the range for a is $[-v_{max} \times t_{now}, y_{max} + v_{max} \times t_{now}]$.

The query is transformed in a polygon in the dual space. We can express this polygon using a linear constraint query [18].

Proposition 1 *The MOR query is expressed in the dual plane as follows:*

[1] In [39] uses a method based on the PMR-quadtree; their experiments show that even for small number of mobile objects (50K) the number of copies can become quite large (about 250 copies/object).

- *For $v > 0$ the query is : $Q = C_1 \wedge C_2 \wedge C_3 \wedge C_4$, where: $C_1 = v \geq v_{min}$, $C_2 = v \leq v_{max}$, $C_3 = a + t_{2q}v \geq y_{1q}$ and $C_4 = a + t_{1q}v \leq y_{2q}$.*
- *For $v < 0$ the query is : $Q = D_1 \wedge D_2 \wedge D_3 \wedge D_4$, where: $D_1 = v \leq -v_{min}$, $D_2 = v \geq -v_{max}$, $D_3 = a + t_{1q}v \geq y_{1q}$ and $D_4 = a + t_{2q}v \leq y_{2q}$.*

Proof. Let $v > 0$. The C_1 and C_2 constraints are straightforward by the problem setting. We will show that the other two are also necessary. Let that a is a particular $a \leq y_{2q}$. Then a line with intercept a can intersect the query rectangle iff the slope has values, (see Figure 2):

$$\tan(\phi_1) \leq v \leq \tan(\phi_2) \Rightarrow \frac{y_{1q} - a}{t_{2q}} \leq v \leq \frac{y_{2q} - a}{t_{1q}}$$

From the above inequalities we get the two other constraints. Similarly for $v < 0$. □

Since the query is different for positive and negative slopes, we can use two structures to store the dual points. It is easy to see that the range of the a's values is now $[-v_{max} \times t_{now}, y_{max} - v_{min} \times t_{now}]$.

However since the time is monotonically increasing, the values of the intercept are not bounded. If the value of the maximum speed is significant, the values of the intercept can become very large and this potentially can be a problem (i.e., representing unbounded ranges of real numbers).

A solution is to start a new index after $T_{period} = \frac{Y_{max}}{v_{min}}$ time instants. Hence at each time we keep two indexes and at each time instant a given point can be stored in one of the two indexes. Initially, all points inserted from time t=0 until T_{period} have their intercept at $t = 0$ stored in the first index. Points inserted after time T_{period} have their intercept at t=T_{period} stored in the second index. Points of the first index that are updated after time T_{period} are deleted and transferred to the second index. The idea is that after T_{period} time instants we are sure that there is no object that has not updated its motion information. After t=T_{period} the first index is removed and the $2T_{period}$ index starts. Using this approach, the (absolute) values of intercept are always between 0 and $v_{max} \times T_{period}$. The queries are directed to both indexes.

Another way to represent a line $y = vt + a$, is to write the equation as $t = \frac{1}{v}y - \frac{a}{v}$. Then we can map this line to a point in the dual plane with coordinates $n = \frac{1}{v}$ and the $b = -\frac{a}{v}$ (Hough-Y in [25]). Note that b is the point where the given line intersects the line $y = 0$. Note also that this transform cannot represent horizontal lines (similarly, the Hough-X transform cannot represent vertical lines). However, this is not a problem since our lines have a minimum and a maximum slope. The only difference between the two transforms is the different values of the coordinates for a specific line. We will use one of the two, depending on the magnitude of the coordinates.

4.3 Lower Bounds

The dual space-time representation transforms the problem of indexing mobile objects on a line to the problem of simplex range searching in two dimensions.

In simplex range searching we are given a set S of 2-dimensional points, and we want to answer efficiently queries of the following form: given a set of linear constraints $a\dot{x} \leq b$, find all the points in S that satisfy all the constraints. Geometrically, the constraints form a polygon on the plane, and we want to find the points in the interior of the polygon. This problem has been extensively studied before in the static, main-memory setting (see for example the surveys of Agarwal and Erickson [2], or Matusek [28] and the related work section).

The only known lower bound for simplex range searching, if we want to report all the points that fall in the query region rather than their number, is due to Chazelle and Rosenberg ([11]). They show that simplex reporting in d-dimensions with a query time of $O(N^{\delta} + K)$, where N is the number of points, K is the number of the reported points and $0 < \delta \leq 1$, requires space $\Omega(N^{d(1-\delta)-\epsilon})$, for any fixed ϵ. This result is shown for the pointer machine model of computation. The bound holds for the static case, even if the query region is the intersection of just two hyper-planes. Since ϵ can be arbitrary small, any algorithm that uses linear space for d-dimensional simplex range searching has worst case query time of $O(N^{(d-1)/d} + k)$.

Here we show that a similar bound holds for the input-output complexity of simplex searching. Following the approach in [38] we use the external memory pointer machine as our model of computation. This is a generalization of the pointer machine suitable for analyzing external memory algorithms. In this model, a data structure is modeled as a directed graph $G = (V, E)$, with a source w. Each node of the graph represents a disk block and is therefore allowed to have B data and pointer fields. The points are stored in the nodes of G. Given a query, the algorithm traverses G starting from w, examining the points at the nodes it visits. The algorithm can only visit nodes that are neighbors of already visited nodes (with the exception of the root) and, when it terminates the answer to the query must be contained in the set of visited nodes. The running time of the algorithm is the number of nodes it visits.

Theorem 1 *Simplex reporting in d-dimensions with a query time of $O(n^{\delta} + k)$ I/O's, where N is the number of points, $n = N/B$, K is the number of the reported points, $k = K/B$, and $0 < \delta \leq 1$, requires $\Omega(n^{d(1-\delta)-\epsilon})$ disc blocks, for any fixed ϵ.*

Proof. See Appendix. □

A corollary of the theorem is that in the worst case a linear size 2-dimensional data-structure requires $O(\sqrt{n}+t)$ I/O's to answer simplex range queries. In the next section we give a dynamic, external-memory algorithm that achieves almost optimal query time with linear space. As we shall see however this algorithm is not practical so we also consider faster algorithms to approximate the queries. Finally, we give a worst case logarithmic query time algorithm for a restricted but practical version of our problem.

4.4 An (Almost) Optimal Solution

Matousek ([29]) gave an almost optimal algorithm for simplex range searching, given a static set of points. This main memory algorithm is based on the idea of simplicial partitions.

We briefly describe this approach here. For a set S of N points, a simplicial partition of S is a set $\{(S_1, \Delta_1), \ldots (S_r, \Delta_r)\}$ where $\{S_1, \ldots, S_r\}$ is a partitioning of S, and Δ_i is a triangle that contains all the points in S_i. If $\max_i |S_i| < 2 \min_i |S_i|$, we say that the partition is balanced. Matousek ([29]) shows that, given a set S of N points, and a parameter s (where $0 < s < N/2$), we can construct in linear time, a balanced simplicial partition for S of size $O(s)$ such that any line crosses at most $O(\sqrt{s})$ triangles in the partition.

This construction can be used recursively to construct a partition tree for S. The root of the tree contains the whole set S, and a triangle that contains all the points. We find a balanced simplicial partition of S of size $\sqrt{|S|}$. Each of the children of the root are associated with a set S_i from the simplicial partition, and the triangle Δ_i that contains the points in S_i. For each of the S_i's we find simplicial partitions of size $\sqrt{|S_i|}$, and continue until each leaf contains a constant number of points. The construction time is $O(N \log_2 N)$.

To answer a simplex range query, we start at the root. We take each of the triangles in the simplicial partition at the root and check if it is inside the query region, outside the query region, or intersects one of the lines that define the query. In the first case all points inside the triangle are reported, in the second case the triangle is discarded, and in the third case we recurse on the triangle. The number of triangles that the query can cross is bounded however, since each line crosses at most $O(|S|^{1/4})$ triangles at the root. The query time is $O(N^{1/2+\epsilon} + T)$, with the constant factor depending on the choice of ϵ.

Agarwal et. al. [1] give an external memory version of static partition trees that answers queries in $O(n^{1/2+\epsilon} + k)$ I/Os. To adapt this stucture to our environment, we have to make it dynamic. We show the following lemma:

Lemma 1 *We can insert or delete points in a partition tree in $O(\sqrt(N)\log(N))$ I/Os, and answer simplex queries in $O(n^{1/2+\epsilon} + k)$ I/Os.*

Proof. See Appendix. □

Alternatively we can make the partition tree structure dynamic using a standard technique by Overmars ([30]) for decomposable problems. It is easy to see that partition trees can be modified to allow additions or deletions in $O(\log^2 n)$ update time and $O(n^{1/2+\epsilon} \log n)$ query time.

4.5 Improving the average query time.

Partition trees are not very useful in practice because the query time is $O(n^{1/2+\epsilon} + k)$ for every query, and the hidden constant factor becomes large if we chose a small ϵ. In this section we present two different approaches that are designed to improve the average query time.

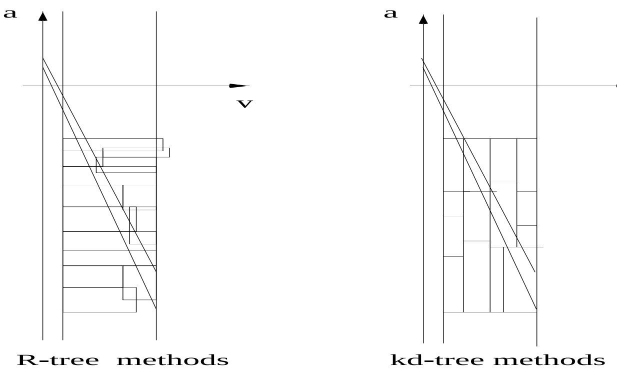

Figure 4: Data regions for R-tree like and *kd*-tree like methods .

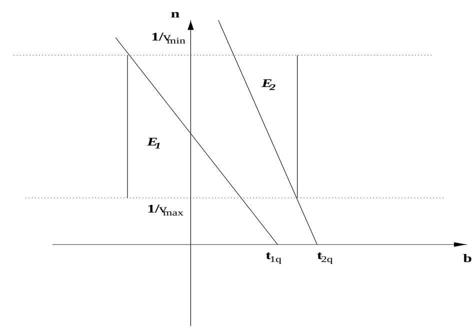

Figure 5: Query in the dual plane.

4.5.1 Using Point Access Methods

There is a large number of access methods that have been proposed to index point data[17]. All these structures were designed to address *orthogonal* range queries, eg. a query expressed as a multidimensional hyper-rectangle. However, most of them can be easily modified to address non-orthogonal queries like simplex queries.

Recently, Goldstein at al. [18] presented an algorithm to answer simplex range queries using R-trees. The idea is to change the search procedure of the tree. In particular they gave simple methods to test whether a linear constraint query region and a hyper-rectangle overlap. As mentioned in [18] this method is not only applicable to the R-tree family, but to other access methods as well.

We use this approach to answer the MOR query in the dual space. However it is not clear what structure would be more suitable here, given that the distribution of points in the dual space is highly skewed. We argue that an index structure based on *kd*-trees (like the LSD-tree [22] and the hB^{Π}-tree [16]) is more suitable than a method based on R-trees. The reason is that since R-trees try to cluster data points into squarish regions [26, 31], they will split using only one dimension (the intercept). On the other hand a *kd*-tree based method will use both dimensions to split (see Figure 4). Thus it is expected to have better performance for the MOR query.

4.5.2 A Query Approximation Algorithm.

A different solution is based on query approximation using the Hough-Y dual plane (Figure 5). In general, the b coordinate can be computed at different horizontal lines (y=y_r). The query region is described by the intersection of two half-space queries. The one line intersects the line $n = \frac{1}{v_{max}}$ at the point $(t_{1q} - \frac{y_{2q}-y_r}{v_{max}}, \frac{1}{v_{max}})$ and the line $n = \frac{1}{v_{min}}$ at point $(t_{1q} - \frac{y_{2q}-y_r}{v_{min}}, \frac{1}{v_{min}})$. Similarly the other line that defines the query intersects the horizontal lines at $(t_{2q} - \frac{y_{1q}-y_r}{v_{max}}, \frac{1}{v_{max}})$ and $(t_{2q} - \frac{y_{1q}-y_r}{v_{min}}, \frac{1}{v_{min}})$.

Suppose that we approximate the simplex query with a rectangular query. For example in Figure 5 the query rectangle will be $[(t_{1q} - \frac{y_{2q}-y_r}{v_{min}}, t_{2q} - \frac{y_{1q}-y_r}{v_{max}}), (\frac{1}{v_{max}}, \frac{1}{v_{min}})]$. Note that the query area is enlarged by the area $E = E_1 + E_2$. Computing this area we have:

$$E = \frac{1}{2}(\frac{v_{max} - v_{min}}{v_{min} \times v_{max}})^2(|y_{2q} - y_r| + |y_{1q} - y_r|) \tag{1}$$

For simplicity consider only the points with positive speed and assume that they are stored in c data structures (where c is a constant). Each data structure stores exactly the same points but with their b-coordinate computed at different y=y_r lines; that is, the $i-th$ structure stores the b coordinates using the line $y = \frac{Y_{max}}{c} \times i$ $(i = 0, .., c-1)$. Conceptually, all such structures contain the same information; however, to answer a query we pick the structure that minimizes the area E. We consider only the points with positive speed.

Since the query is a rectangle with one side $(\frac{1}{v_{max}}, \frac{1}{v_{min}})$ we can use a B+-tree [13] and store only the b coordinate of the points. Then the original (enlarged) query is transformed to a simple range query on a B+-tree structure.

Now we consider two cases for the query interval $[y_{1q}, y_{2q}]$:

$y_{2q} - y_{1q} \leq \frac{Y_{max}}{c}$: It is easy to see that using one of the c structures (the one that minimizes the $|y_{2q} - y_r| + |y_{1q} - y_r|$) the following holds:

$$E \leq \frac{1}{2}\left(\frac{v_{max} - v_{min}}{v_{min} \times v_{max}}\right)^2 \frac{Y_{max}}{c} \tag{2}$$

Thus the enlarged query area is bounded and by using a larger c (i.e., more B+-trees) we can decrease this area.

$y_{2q} - y_{1q} \geq \frac{Y_{max}}{c}$: First we partition the terrain $[0, Y_{max}]$ into c sub-terrains P_i for $0 \leq i \leq c-1$, where the sub-terrain P_i is the interval $[\frac{Y_{max}}{c} \times i, \frac{Y_{max}}{c} \times i+1]$. Then we use c external memory Interval trees[4], one for each P_i and we store for each object the time interval for which this object was inside the sub-terrain P_i. Now, we decompose the original query to a number of queries of the form $[P_j, (t_{1q}, t_{2q})]$ and two more queries when the y_{1q} and y_{2q} are not end points of a sub-terrain. To answer the query $[P_j, (t_{1q}, t_{2q})]$, we need to perform a range query $[(t_{1q}, t_{2q})]$ to IT_j Interval tree. The answer is obtained in optimal time $O(\log_B(n) + K_i/B)$, where K_i is the size of the answer set (part of the answer of the original query). An object can be in many answers but we can report it only once, since we have all the information we need, namely speed and direction. Now to answer the two queries corresponding to the end points, we can choose 2 of the c B+-trees such that the inequality 2 holds. The same solution can be used for the objects with negative speed. Thus we have the following lemma:

Lemma 2 *The MOR query can be answered in time $O(\log_B(n) + (K + K')/B)$, where K' is the approximation error. The space used is $O(cn)$ where c is a parameter, and the update is $O(c \log_B(n))$.*

Note that assuming that the points are distributed uniformly over the b-axis, then the approximation error is bounded by $1/c$, eg. $K' = O(1/c)$.

4.6 Achieving Logarithmic Query Time

In many practical cases, the data structures we discussed earlier are not optimal. Consider for example the case where points are moving very slowly, or with approximately the same velocity. In this case the lines in the time-space plane do not cross until well forward in the future. As a result, the relative positions of the points does not change for a long time. If we restrict our queries to occur before the first time that a point overtakes (passes) another, the original problem is equivalent to 1-dimensional range searching.

This is one of our motivations to consider a restricted version of the original problem, namely, to index mobile points in a bounded time interval T in the future. As we have seen, there exist lower bounds for the original problem that show that we cannot achieve query times better than $\Omega(\sqrt{n})$ given linear space. Using our restriction, we achieve a logarithmic query time, with space that can be quadratic in the worst case but can be expected to be linear in practice.

Formally, the problem we are considering in this section is the following: given a set of points that are moving on a line, and a time limit T, find all the points that lie in the segment $[x_{l_q}, x_{r_q}]$ at time t_q (where $t_0 \leq t_q \leq t_0 + T$). Equivalently, this is a standard MOR query where $t_{1_q} = t_{2_q}$. We will call it a MOR1 query.

Our method is to find all the times when a point overtakes another. These events correspond to line segment crosses in the time-space plane. Note that between two consecutive crossing events the relative ordering of the points on the plane remains the same.

First we show the following lemma:

Lemma 3 *If we have the relative ordering of all the points at time t_q, the position of the points at time t_c that corresponds to the closest crossing event before t_q, and the speed of the points, we can find the points that are in $[x_l, x_r]$ in $O(\log_2 N + K)$ time.*

Proof. See Appendix. □

The following lemma finds all the crossings of points efficiently.

Lemma 4 *We can find all the crossings of points in time $O(N \log_2 N + M \log_2 M)$, where M is the number of crosses in the time period $[0, T]$.*

Proof. See Appendix. □

These M crosses define M ordered lists of the N points. Each two consecutive lists differ in exactly two positions, the positions that correspond to the points that cross. The total sum of the differences between consecutive lists is therefore $O(M)$. In the appendix we show how to store these lists in space $O(n+m)$ in external memory, such that we can perform binary search in the i-th list in time $O(\log_B(n+m))$.

The following theorem follows from the previous lemmas and the discussion in the appendix:

Theorem 2 *Given N points and a time limit T, a MOR1 query can be answered in time $\log_B(n+m)$ using space $O(n+m)$, where $m = \frac{M}{B}$ and M is the number of crosses of points in the time limit T.*

To solve the problem of answering queries within a time interval T into the future, we stagger the construction of our data structure. Thus, at time t_0 we construct a data structure that will answer queries in the time interval $[t_0, t_0 + 2T]$, and at time $t_0 + iT$ we construct a data structure that will answer queries in the time interval $[t_0 + (i+1)T, t_0 + (i+2)T]$

Our approach works for any value of T. If the time limit is set too large however, all pairs of points may cross, in which case the size of the data structure will be quadratic. It is therefore important to set the time limit appropriately so that only approximately a linear number of crossings occur. Fortunately, in practice it is often true that many objects move with approximately equal speeds (one example is cars on a highway) and therefore do not cross very often.

5 Indexing in two dimensions

In this section we consider the problem of mobile objects in the plane. Again we consider only "moving" objects, namely objects with a speed between v_{min} and v_{max}. We assume that objects move in the (x, y) plane inside the finite terrain $[(0, X_{max})(0, Y_{max})]$. The initial location of the object o_i is (x_{i_0}, y_{i_0}) and its velocity is a vector $\vec{v} = (v_x, v_y)$.

We distinguish two important cases. The first considers objects moving in the plane but their movement is restricted on using a given collection of routes (roads) on the finite terrain. Due to its restriction, we call this case the 1.5-dimensional problem. There is a strong motivation for such an environment; for the applications we have in mind, objects (cars, airplanes etc.) move on a network of specific routes (highways, airways).

5.1 The 1.5-dimensional problem

The 1.5-dimensional problem can be reduced to a number of 1-dimensional queries. In particular, we propose representing each predefined route as a sequence of connected (straight) line segments. The positions of these line segments on the terrain are indexed by a standard SAM. (Maintaining this SAM does not introduce a large overhead since for most practical applications: (a) the number of routes is much smaller than the number of objects moving on them, (b) each route can be approximated by a small number of straight lines, and, (c) new routes are added rather infrequently.) Indexing the points moving on a given route is a 1-dimensional model and will use techniques from the previous section.

Given a MOR query, the above SAM identifies the intersection of the routes with the query's spatial predicate (the rectangle $[x_{1q}, x_{2q}] \times [y_{1q}, y_{2q}]$). Since each route is modeled as a sequence of line segments, the intersection of the route and the query's spatial predicate is also a set of line segments, possibly disconnected. Each such intersection corresponds to the spatial predicate of a 1-dimensional query for this route. In this setting we assume that when routes intersect, objects remain in the route previously traveled (otherwise an update is issued).

An interesting open problem is to index such mobile objects when a probabilistic distribution is assigned at route intersections (i.e., an object arriving at the intersection between routes A,B and C, traveling from route A, has probability p to remain on route A, q to continue on B and $1-p-q$ to continue on C). This of course will lead to probabilistic query answers [40].

5.2 The 2-dimensional problem

The full 2-dimensional problem (i.e., allowing objects to move anywhere on the finite terrain) is more difficult. As with the 1-dimensional case, we discuss different representations of the problem and we propose methods to address the 2-dimensional MOR query.

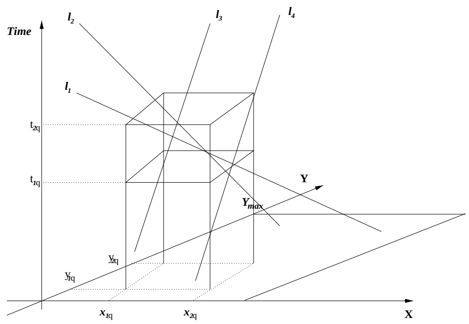

Figure 6: Trajectories and query in (x, y, t) plane.

Parameter	Values
N	40k, 60k, 80k, 100k 120k
Y_{max}	3000
v_{min}	1
v_{max}	30
$VarSpeed$	5
L	30
$MAXT$	5000
NPQ	10
$YQMAX$	300
TW	40

Figure 7: Parameter values.

In the space-time representation the trajectories of the mobile objects are lines in the space. The lines can be computed by the motion informations of each object. In this case the MOR query is expressed as a cube in the 3-dimensional (x, y, t) space and the answer is the set of objects with lines that cross the query cube.

Algorithms that are applied directly to the time-space representation do not work well in one-dimension, so the performance is likely to be even worse in two dimensions. Unfortunately we cannot use directly the dual transformations of the previous section, since these transforms map a hyper-plane in the space into a point and vice-versa, where here we have lines. We point out that the problems with lines in the space are much harder than lines in the plane [10]. The reason is that a line in space has 4 degrees of freedom and therefore taking the dual we jump to a 4-dimensional space, that is, the problem is inherently 4-dimensional. To get the dual we project the lines on the (x, t) and (y, t) planes and then take the duals for the two lines on these planes. Thus now a line can be represented by the 4-dimensional point (v_x, a_x, v_y, a_y), where the v_x and v_y are the slopes of the lines on the (x, t) and (y, t) planes and the a_x and a_y are the intercepts respectively.

Now the MOR query is mapped to a simplex query to the dual space. This query is the intersection of four 3-d hyper-planes and the projection of the query to (t, x) and to (t, y) plane is a wedge, as in the 1-dimensional case. Thus we can use a 4-dimensional partition tree (section 4.4) and answer the MOR query in $O(n^{0.75+\epsilon} + t)$ I/O's that almost matches the lower bound for four dimensions.

A simple approach to solve the 4-dimensional problem is to use an index based on the $kd-tree$. An alternative approach is to decompose the motion of the object into two independent motions, one in the x-axis and the other in the y-axis. For each axis we can use the methods for the 1-dimensional case and answer two 1-dimensional MOR queries. We must then take the intersection of the two answers to find the answer to the initial query. This method allows us to use the algorithms for the 1-dimensional case.

6 A Performance Study

We present initial results comparing our 1-dimensional structures and a traditional R-tree based approach. First we describe the way experimental data is generated. At time $t = 0$ we generated the initial locations of N mobile points uniformly distributed on the interval $[0, Y_{max}]$. The initial speeds were generated uniformly from v_{min} to v_{max} and the direction randomly positive or negative. Thus almost half of the points are moving in one direction. Then points start moving. When a point reaches a border simply it changes its direction. At each time instant we choose L objects randomly and we change their speed and/or direction. The change of the speed is chosen uniformly between 0 and $VarSpeed$ and the direction again randomly between $+1$ and -1. We generate NPQ different time instants that represent the times when queries are executed. At each such time instant we actually execute 200 queries, where the length of the y-range is chosen uniformly between 0 and $YQMAX$ and the length of the time range between 0 and TW. The beginning of the y-range is uniformly distributed in the interval $[0, Y_{max}]$, while the start time of the query time range (t_{1q}) is chosen randomly between t_{now} and $t_{now} + 30$. We run this scenario using a particular access method until the $MAXT$. The values of the parameters in the experiments are shown on Figure 7.

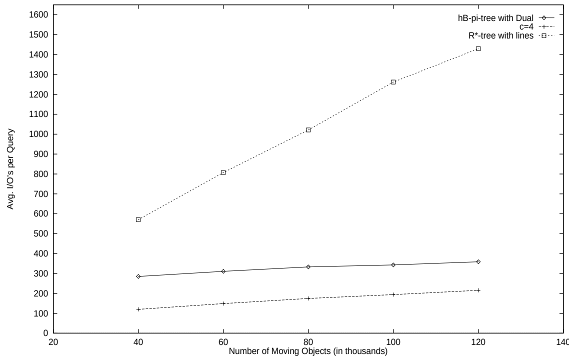

Figure 8: Query Performance.

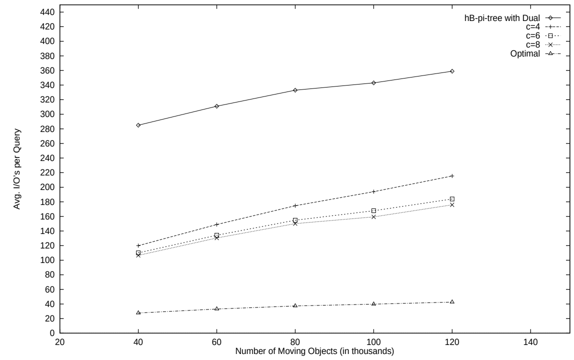

Figure 9: Query Performance.

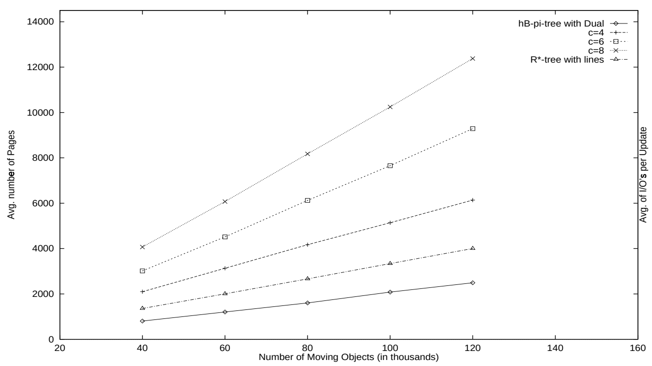

Figure 10: Space Consumption.

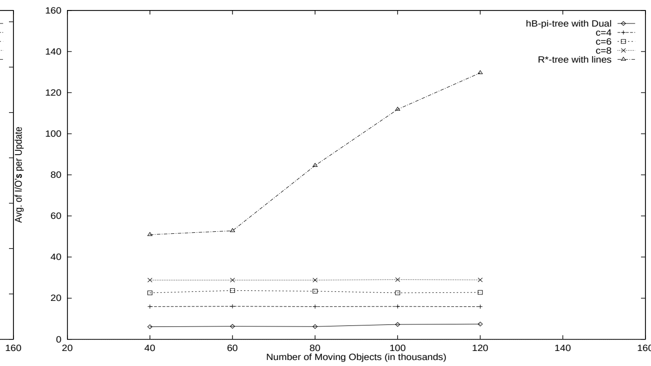

Figure 11: Update Performance.

To verify that indexing mobile objects as line segments is not efficient, we stored the trajectories in a R*-tree. We fixed the page size to 1024 bytes. To represent a line segment in an R*-tree we used four 4-byte numbers (the two end points) and one more number as a pointer to the real object, resulting in a page capacity of $B = 51$. For the B+-tree we used one 4-byte number to represent the b-coordinate and another for the pointer, so the page capacity was $B = 120$. In the hB-tree an object is stored as two numbers (slope and intercept) and one pointer; this results in page capacity $B = 85$. However, each h-B-tree page reserves some space for internal structural data. We consider a simple buffering scheme for the results we present here. For each tree we buffer the path from the root to a leaf node, thus the buffer size is only 3 or 4 pages. For the queries we always clear the buffer pool before we run a query. An update is performed when the motion information of an objects changes.

In Figure 8 we present the results for the average number of I/O's per query for the various methods. The approximation method used $c = 4$ B+-trees. As anticipated, the line segments method with R*-trees has the worse performance. Also, the approximation method is better than the $(hB^{\Pi} - tree$. To examine the effect of using more B+-trees in the approximate method, Figure 9 shows the query performance for c= 4,6 and 8 (the R*-tree approach is not shown). We also plot the minimal number of pages needed to answer the query by an "optimal" method that only fetches the answer in full pages.

In Figures 10,11 we plot the space consumption and the average number of I/O's per update respectively. The update and space performance of the $hB^{\Pi} - tree$ is better than the other methods since its objects are stored only once and better clustered than the R*-tree. As we can observe however the space overhead of the approximate methods is not high and also the update overhead is not changing for different number of mobile objects. Note also the increase in the update performance of the R*-tree as the number of objects increases. This is due to the reinsertions that this method uses in order to achieve better clustering.

7 Related Work

The problem of indexing mobile points is novel; we are not aware of any other related work except [39] where a method to index mobile points based on the PMR-quadtree is presented. However as we mentioned earlier, this approach has large space and update overhead.

Mobility in a geographic system is addressed in [36] where the aim is to map close points in space to adjacent disks so as collision detection queries are optimized.

The queries we examine have also a temporal component. There has been a lot of research in temporal indexing [33], however it has focused on queries about the past and not the future as in our case.

Representing the trajectories as line segments in two and three dimensions, also relates to spatial indexing [17]. [24] presents a qualitative comparison of three spatial access methods for a line segments database is presented. In particular they consider the R*-tree, the R+-tree and the PMR-quadtree. The result is that the methods are comparable and no one seems to be superior than the other.

A method to index line segments on the plane is presented by Jagadish in [25]. A line segment is represented by the slope and the intercept of the line obtained by extending the line segment and by the range of the projection of the segment to one axis. Using this mapping, a line segment on the plane is mapped to a vertical line segment in three dimensions. Then a standard spatial access method can be used to index the new segments. It is shown analytically and experimentally that the queries in the transformation space have better selectivity that in the original space.

An interesting approach to index constraint databases is presented by Bertino et al. in [9]. In particular they address the problem of indexing conjunction of linear constraints with two variables, in order to answer ALL and EXIST queries (variations of the half-plane query). They use the dual transformation and they reduce the problem to a point location problem. Then, if the line that defines the query has slope from a predefined set of slopes, an optimal solution can be derived using the external memory Interval tree[4]. Also other works on indexing constraint databases include [4, 23, 32]. All these approaches reduce the problem of indexing constraint to a dynamic interval management problem or to a special case of two-dimensional range searching, and therefore are not applicable to our problem.

The issue of mobility and maintenance of a number of configuration functions among continuously moving objects has been addressed by Basch et al. in [5]. Such functions are the convex hull, the closest pair and the minimum spanning tree. They propose a framework to transform a static data structure into a *kinetic* data structure (KDS) that maintains an attribute of interest for a set of mobile objects and they give a number of criteria for the quality of such structures.The key structure is an event queue that contains events corresponding to times where the value of the configuration function (may) change. This event queue is the interface between the data structure and the mobile objects. All these structures are main memory data structures. It will be an interesting problem to investigate how these structures can be implemented efficiently in external memory.

8 Conclusions and Future Work

Indexing mobile objects is a novel problem motivated by real life applications. We study the one and 2-dimensional versions of the problem. For the 1-dimensional case, we give an dynamic, external memory algorithm with guaranteed worst case performance and linear space. We also give a practical approximation algorithm also in the dynamic, external memory setting, which has linear space and expected logarithmic query time. Finally we give an algorithm with guaranteed logarithmic query time for a restricted version of the problem. We also extend some of our results into two dimensions. First we consider the case where objects move in 2-dimensional networks of 1-dimensional routes. In this case we can effectively apply our 1-dimensional algorithms. We also consider points that move on a plane, and we discuss extensions of our techniques to two dimensions.

Future work includes a variety of interesting problems. In addition to performing a more complete performance study (using various data distributions) we plan to address restricted versions of the 2-dimensional problem using realistic assumptions. One idea is to cluster similarly moving points into representative clusters. If query response is time critical, main-memory database techniques need to be involved. We are currently studying the problem of indexing mobile points with probabilistic route choices. A generalization of the 1.5-dimensional problem is when the terrain is subdivided into areas with various speed limits (or terrain abnormalities that limit movement according to direction). Other interesting queries are near-neighbor queries and joins among relations of mobile objects. Some applications may require keeping the history of mobile objects (for traffic analysis etc.); then the indexes presented need to support historical queries. This probably requires making the presented structures partially persistent [27, 7]. While in this paper we restricted the object movement to simple (linear) functions, it is a first step at examining ways to index more complex functions.

References

[1] P.K. Agarwal, L. Arge, J. Erickson, P. Franciosa and J.S. Vitter. Efficient Searching with Linear Constraints In *Proc. 17th ACM PODS Symposium on Principles of Database Systems*, Seattle, 1998.

[2] P.K. Agarwal, and J. Erickson. Geometric range searching and its relatives. In *Discrete and Computational Geometry: Ten Years Latter.*, (B. Chazelle, E. Goodman, and R. Pollack eds.), American Mathematical Society, Providence, 1998.

[3] A. Aggarwal and J. S. Vitter. The Input/Output complexity of sorting and related problems. In *Communications of the ACM*, 31(9):1116-1127, 1988.

[4] L. Arge and J. S. Vitter. Optimal Dynamic Interval Management in External Memory. In *Proc. 37th Annual Symposium on Foundations of Computer Science*, 1996.

[5] J. Basch, L. Guibas and J. Hershberger. Data Structures for Mobile Data. In Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, Louisiana, 1997.

[6] R. Bayer and C. McCreight. Organization and Maintenance of Large Ordered Indexes. *Acta Informatica*, 1(3):173-189, 1972.

[7] B. Becker, S. Gschwind, T. Ohler, B. Seeger and P. Widmayer. An Asymptotically Optimal Multiversion B-Tree. *VLDB Journal*, 5(4):264-275,1996.

[8] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger. The R*-tree: An Efficient and Robust Access Method For Points and Rectangles. In *Proc. ACM-SIGMOD International Conference on Management of Data*, pages 322–331, Atlantic City, May 1990.

[9] E. Bertino, B. Catania and B. Shidlovsky. Towards optimal two-dimensional indexing for constraint databases. *Information Processing Letters*, 64(1997):1-8.

[10] B. Chazelle, H. Edelsbrunner, L. J. Guibas, M. Sharir and J. Stolfi. Lines in Space: Combinatorics and Algorithms. In *Algorithmica*, 15(5): 428-447, 1996.

[11] B. Chazelle and B. Rosenberg. Lower Bounds on the Complexity of Simplex Range Reporting on a Pointer Machine. *Proc. 19th Intern. Colloquium on Automata, Languages and Programming. LNCS,* Vol. 693, Springer-Verlang, Berlin, 1992.

[12] R. Cole. Searching and Storing Similar Lists. *Journal of Algorithms*, 7(2):202-220, 1986.

[13] D. Comer. The Ubiquitous B-Tree. *Computing Surveys*, 11(2):121–137, June 1979.

[14] J. Driscoll, N. Sarnak, D. Sleator and R.E. Tarjan. Making Data Structures Persistent. In *Proc. of the Eighteenth Annual ACM Symposium on Theory of Computing*, Berkeley, California, 1986.

[15] M. Erwig, R.H. Guting, M. Schneider and M. Vazirgianis. Spatio-temporal Data Types: An Approach to Modeling and Querying Moving Objects in Databases. to appear in the proceedings of ACM GIS Symposium '98.

[16] G. Evangelidis, D. Lomet, and B. Salzberg. The hBΠ-tree: A Modified hB-tree Supporting Concurrency, Recovery and Node Consolidation. In *Proc. 21st International Conference on Very Large Data Bases*, Zurich, September 1995, pages 551–561.

[17] V. Gaede and O. Gunther. Multidimensional Access Methods. *ACM Computing Surveys*, 30(2):170-231, 1998.

[18] J. Goldstein, R. Ramakrishnan, U. Shaft and J.B. Yu. Processing Queries By Linear Constraints. In *Proc. 16th ACM PODS Symposium on Principles of Database Systems*, Tuscon, Arizona, 1997.

[19] O. Gunther. The Design of the Cell Tree: An Object-Oriented Index Structure for Geometric Databases. In *Proc. Fifth IEEE International Conference on Data Engineering*, Los Angeles, California, USA, February 1989.

[20] A. Guttman. R-Trees: A Dynamic Index Structure For Spatial Searching. In *Proc. ACM-SIGMOD International Conference on Management of Data*, pages 47–57, Boston, June 1984.

[21] J.M. Hellerstein, J. Naughton and A. Pfeffer. Generalized Search Trees for Database Systems. In *Proc. 21st International Conference on Very Large Data Bases*, pages 562–573, Zurich, Switzerland, September 1995.

[22] A. Henrich, H.-W. Six, P. Widmayer. The LSD tree: Spatial Access to Multidimensional Point and Nonpoint Objects. In *Proc. 15th International Conference on Very Large Data Bases*, pages 45–53, Amsterdam, The Netherlands, August 1989.

[23] P. Kanellakis, S. Ramaswamy, D. Vengroff and J. Vitter. Indexing for Data Models with Constraint and Classes. In *Proc. 12th ACM SIFACT-SIGMOD-SIGART Symposium on Principles of Database Systems*,pages 233-243 , Washington, D.C, 1993.

[24] E.G. Hoel and H. Samet. A Qualitative Comparison Study of Data Structures for Large Linear Segment Databases. In *Proc. ACM-SIGMOD International Conference on Management of Data*, pages 205-214, San Diego, June 1992.

[25] H. V. Jagadish. On Indexing Line Segments. In *Proc. 16th International Conference on Very Large Data Bases*, pages 614–625, Brisbane, Queensland, Australia, August 1990.

[26] I. Kamel and C. Faloutsos. On Packing R-trees. In *Proc. Second Int. Conference on Information and Knowledge Management (CIKM)*, Washington, DC, Nov. 1-5, 1993.

[27] A. Kumar, V.J. Tsotras and C. Faloutsos. Designing Access Methods for Bitemporal Databases. *IEEE Transactions on Knowledge and Data Engineering*, 10(1):1-20, 1998.

[28] J. Matousek. Geometric range searching. *ACM Computing Surveys*, 26(4):421-461, 1994.

[29] J. Matousek. Efficient Partition Trees. *Disc. Comp. Geom.* , 8 (1992), 432-448.

[30] M.H. Overmars. The Design of Dynamic Data Structures. LNCS vol. 156, Springer-Verlag, Heidelberg, West Germany, 1983.

[31] B.-U. Pagel, H.-W. Six, H. Toben, and P. Widmayer. Towards an Analysis of Range Query Performance in Spatial Data Structures. In *Proc. 12th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 214–221, Washington, D. C., May 1993.

[32] S. Ramaswamy. Efficient Indexing for Constraint and Temporal Databases. In *Proc. 6th International Conference on Database Theory*, pages 419-431, Delphi, Greece, January 8-10, 1997.

[33] B. Salzberg and V.J. Tsotras. A Comparison of Access Methods for Time-Evolving Data. *ACM Computing Surveys*, to appear.

[34] H. Samet. *The Design and Analysis of Spatial Data Structures.*, Addison-Wesley, Reading, MA, 1990.

[35] T. Sellis, N. Roussopoulos and C. Faloutsos. The R+-Tree: A Dynamic Index for Multi-Dimensional Objects. In *Proc. 13rd International Conference on Very Large Data Bases*, pages 507-518, Brighton, England, September 1987.

[36] S. Shekhar and T.A. Yang. Motion in a Geographical Database System In *2nd International Symposium on Advances in Spatial Databases*, pages 339–357, Zürich, Switzerland, August 1991.

[37] A. P. Sistla, O. Wolfson, S. Chamberlain, S. Dao. Modeling and Querying Moving Objects. In *Proc. 13th IEEE International Conference on Data Engineering*, pages 422-432, Birmingham, U.K, April 1997.

[38] S. Subramanian and S. Ramaswamy. The P-range Tree: A New Data Structure for Range Searching in Secondary Memory. In *Proceedings of the 6th Annual Symposium on Discrete Algorithms*, New York, NY, USA, 1995.

[39] J. Tayeb, O. Ulusoy, O. Wolfson. A Quadtree-Based Dynamic Attribute Indexing Method. *The Computer Journal*, 41(3):185-200, 1998.

[40] O. Wolfson, S. Chamberlain, S. Dao, L. Jiang, G. Mendez. Cost and Imprecision in Modeling the Position of Moving Objects. In *Proc. 14th IEEE International Conference on Data Engineering*, pages 588-596, Orlando, Florida, February 1998.

[41] O. Wolfson, B. Xu, S. Chamberlain, L. Jiang. Moving Objects Databases: Issues and Solutions In *Proceedings of the 10th International Conference on Scientific and Statistical Database Management*. Capri, Italy, July 1998.

Appendix

A Proofs

Proof of Theorem 1. (Sketch) To prove the lower bound we need to show that, given δ, there exists a set of N points, and a set of $\Omega(n^{d(1-\delta)-\delta-\epsilon})$ queries such that, each query has $\Theta(Bn^{\delta})$ points, and the intersection of any pair of query results is small. To answer a query with $\Theta(Bn^{\delta})$ points, the answering algorithm must visit $\Omega(n^{\delta})$ nodes. To answer this query in $O(n^{\delta})$ I/O's, at least a constant fraction of that many blocks have a constant fraction of their points in the answer of the query. But if the set of the queries has small intersection, it follows that to answer this set of queries in time $\Omega(n^{\delta})$ at least $\Theta(n^{\delta}) \times \Omega(n^{d(1-\delta)-\delta-\epsilon}) = \Omega(n^{d(1-\delta)-\epsilon})$ nodes have to be visited. It remains to show that such a set of queries exist. To do so we simply modify the existing construction by Chazelle and Rosenberg [11] by replacing each point in their point set by B copies. □

Proof of Lemma 1. To make the structure dynamic we note that, although a partition tree has depth $O(\log\log N)$, to insert or delete a point we have to find a triangle that contains the point at every node. The total number of triangles we have to check is $O(\sqrt{N})$. In the case of an insertion, it may be the case that the new point is contained in a node but not in any of the triangles of the simplicial partition of the node. In that case, the point is kept in a separate list of points in the node. To answer a given query, if a given node intersects the query, we check each of the points in the list separately against the query. Note that this list can contain up to $O(\sqrt{N_v})$ points without degrading the query time because there are also $O(\sqrt{N_v})$ triangles that we have to check the query against.

If after we insert or delete a point the partition remains balanced, nothing more needs to be done. If it is not, we have to rebuild the partition, including the points that are kept separately in a list. However, since a partition of size $O(N_v)$ is balanced if all the triangles contain $O(\sqrt{N_v})$, or if the number of points outside all triangles is at most $O(\sqrt{N_v})$, it will need rebuilding only after $O(\sqrt{N_v})$ insertions or deletions. Since the time to rebuild the partition is $O(N_v \log N_v)$, we can use an amortization argument and show that we can perform insertions or deletions in $O(\sqrt{N_v}\log N_v)$ time. □

Proof of Lemma 3. Assume that the points are $\{p_1, p_2, \ldots, p_N\}$, where p_i has a position x_i at time t_c and a velocity v_i. Without loss of generality, assume that, at time t_q, the relative order of the points from left to right is $p_1, p_2, \ldots, p_N$.

Store the points in a binary tree, sorted by their original positions. The root of the tree, point p_i, is going to be at position $x_i + v_i \times t_q$ at time t_q. Since the points in the binary tree are stored by order at the time t_q, if $x_i + v_i \times t_q < x_l$ then this is also true for all the points to the left child of the root, in which case we eliminate the left child and recurse on the right child. Otherwise we recurse on the left child of the tree. Thus in $O(\log N)$ time we can find the positions of x_l and x_r relative to the points at time t_q, and we report the points that lie between. □

Proof of Lemma 4. Let $\{p_1, \ldots, p_N\}$ be the ordering of the N points at time 0. At time T, the position of point i is $x_i + u_i \times T$. To find the ordering of the points at time T we have to sort their positions. Let $\{p_{t(1)}, \ldots, p_{t(N)}\}$ be the ordering of the same N points at time T. Then points i and j cross if and only if $t(j) < t(i)$.

Keep the points in a linked list, in the same order they were at time $t = 0$. Scan the sorted list of points at time T. Find point $p_{t(1)}$ in the list. This point crosses all the points ahead of it in the list. After reporting these crosses, we remove it from the list, and repeat this process with the next point. This procedure reports all the crossings in $O(N + M)$ time. After all the crossings are reported we can find when each occurs and sort them on time. □

B Storing similar lists in external memory

Let $L(t)$ be the list of points at time t. Consider $CS = t_1, \ldots, t_M$ the ordered sequence of the time instants where crossings occur during the interval $(0, T)$. The problem of storing the M ordered lists $L(t_1)$ through $L(t_M)$ can be "visualized" as storing the history of a list $L(t)$ that evolves over time, i.e., a partial persistence problem [14]. That is, list $L(t)$ starts from an initial state $L(0)$ and then evolves through consecutive states $L(t_1)$, $L(t_2)$,..., $L(t_M)$, where $L(t_{i+1})$ is produced from $L(t_i)$ by applying the crossing that occurred at t_{i+1} (i=0,...,M-1, and t_0=0). We would like to store this history in external memory using space $O(n+m)$ while being able to perform a binary search on any list $L(t)$ (t in (0,T)) in time $O(\log_B(n+m))$.

A common characteristic in the list evolution is that each $L(t)$ has exactly N positions, namely positions 1 through N, where position j stores the $j-th$ element of $L(t)$. To perform a binary search on a given $L(t)$ we could implement it using a binary tree with N nodes, where each node is numbered by a position (the root node corresponds to the middle position in the list and so on) and holds the element of $L(t)$ at that position. One obvious solution to the problem would be to store the binary tree of the original list $L(0)$ and the binary tree of each $L(t_i)$ for all t_i in CS. Then, a query about list $L(t)$ is addressed by using the binary tree of $L(t_i)$, where t_i is the largest instant in CS that is less or equal to t. While this achieves $O(\log_2(N+M))$ query time, it uses $O(MN)$ space.

To reduce the space to $O(N+M)$ we must take advantage of the fact that subsequent lists do not defer much. A main-memory solution to this problem appears in [12]. Here we present an efficient external memory solution. In particular, we first embed the binary tree structure inside a B-tree. This is easily done since the structure of the list (and its corresponding binary tree) does not change over time. Consider for example tree $B(0)$ that corresponds to the initial list $L(0)$. Tree $B(0)$ uses $O(n)$ nodes where each node can hold B entries. An entry is now a record: $(position, occupant, pointer, t)$, where *position* corresponds to a position in the list, *occupant* contains the element at that position, *pointer* points to a child node and t corresponds to the time this element was at that position, in this case t=0.

Conceptually, each B-tree node is permanently assigned B positions and is responsible for storing the occupants of these positions. Consider the evolution of such a node s through trees $B(0)$,$B(t_1)$,$B(t_2)$,...,$B(t_M)$. An obvious way to store this evolution is to store a copy of $s(0)$ and a "log" of changes that happen on the occupants of node s at later times. A change is simply another record that stores the position where a change occurred, the new occupant and the time of the change. To achieve fast access to $s(t)$ we do not allow the "log" to get too large. Every $O(B)$ changes (in practice when the log fills one or two pages) we store a new, current copy of s. If we consider the history of node s independently, we can have an auxiliary array with records (*time*, *pointer*) that point to the various copies of node s. Locating the appropriate node $s(t)$ takes $O(\log_B(m))$ time (first find the record in the auxiliary array with the largest timestamp that is less or equal to t and then we access the appropriate copy of s and probably a (constant) number of "log" pages). The space remains $O(n+m)$ since every new node copy is amortized over the $O(B)$ changes in the "log".

While this solution works nicely for the history of a given B-tree node, it would lead to $O(\log_B(n) \times \log_B(m))$ search (since finding the appropriate version of a child node, when searching the B-tree, requires $O(\log_B(m))$ search in the child node's history). Instead of using the auxiliary array to index the copies of node s we post such entries as changes in the history of the parent node p. Assume that node s is pointed by the record on position l in node p. When a new copy of node s is created, a new record is added on the "log" of p that has the same position l, but a pointer to the new copy of s and the current time. Since new node copies are added after $O(B)$ changes, the overall space remains $O(n+m)$. The query time is reduced to $O(\log_B(n+m))$ since performing a binary search on list $L(t)$ is equivalent to searching a path of $B(t)$; locating the root of $B(t)$ takes $O(\log_B(m))$ (searching the history of the B-tree root node) while all other nodes of $B(t)$ are found in $O(\log_B(n))$ using the appropriate parent to child pointers.