In Advances in Conceptual Modeling:
ER 2014 Workshops,:
ENMO, MoBiD, MReBA, QMMQ, SeCoGIS, WISM, and ER Demos, Atlanta, GA, USA, October 27-29, 2014.

A Model of Aggregate Operations for Data Analytics over Spatiotemporal Objects

Logan Maughan, Mark McKenney & Zachary Benchley

Department of Computer Science, Southern Illinois University Edwardsville
{lmaugha,marmcke,zbenchl}@siue.edu

Abstract. In this paper, we identify a conceptual framework to explore notions of spatiotemporal aggregate operations over moving objects, and use this framework to discover novel aggregate operators. Specifically, we provide constructs to discover temporal and spatial coverage of a query window that may itself be moving, and identify quantitative properties of entropy relating to the motion of objects.

1 Introduction

Spatiotemporal datasets represent a valuable resource for business, industry, government, and scientific entities. Specifically, spatiotemporal data in the form of moving objects encode useful information in the form of spatial location, extent, and temporal variance that is independent of traditional thematic data that may be associated with the objects. For example, a moving region (i.e., a region that changes shape and/or position over time) might represent the aggregate position of a cluster of vehicles traveling along a highway. Despite the sensor data providing information, the spatial and temporal characteristics of the moving region provide information as well; information such as the relative speed of the region, the ability of the group to maintain proximity (forcing a smaller region), and the change of the shape of the group over time (presumably impacted by traffic). Furthermore, if many such moving regions exist in a database, aggregates of this information may be computed to gain large-scale data analysis. Aggregate operations could identify areas of greatest movement activity, areas forcing the most regions to become thin and stretched, or areas forcing the greatest or least amount of dispersion of vehicles in a group.

In this paper, we i) identify novel constructs to quantify spatiotemporal interactions of moving objects with query windows that provide insights into the duration and rate of qualitative change of such interactions, ii) we present a novel conceptual framework for discovering and classifying new types of queries over moving objects, and iii) we identify novel aggregate operations based on the discovered queries for use in large scale analytics of moving objects.

2 Related Work

Aggregate operations on spatiotemporal databases have been studied via two main approaches. The first approach uses spatial operations on spatial data to generate traditional numeric data, which is then aggregated [3]. In this paper, we are concerned about spatiotemporal aggregates that operate on purely spatial and temporal properties of moving objects, and not the related attribute values. The other approach has focused on discovering aggregate information of objects in query windows, such as the maximum or minimum count of spatiotemporal objects in a query window over time [9, 2, 5]. The operations presented in this paper differ from those in that we are aggregating spatial and temporal values derived from the intersection of moving objects and irregular query windows.

Much work has been done in the literature to identify and define spatiotemporal data types and operations to form spatiotemporal algebras [1, 7, 8]. The traditional spatial types consist of complex points, lines, and regions. A complex point can contain multiple individual points, a complex line can contain multiple lines that may branch, and a complex region is defined as a closed area that may contain multiple faces and holes: for example, Italy can be represented as a region with its mainland and islands each forming a face in the region, and the area where Vatican City lies forming a hole that does not belong to Italy [6]. A *moving object* is defined as a mapping from time to instances of a traditional spatial type, resulting in a type that changes position and shape over time, but that defines a valid spatial type at every time instant. Moving types form 3-dimensional structures when plotted in two spatial dimensions and a temporal dimension; for example, a moving region forms a volume in 3-dimensional space in which space forms the first two dimensions, and time the third.

3 Definitions

For the purposes of this paper, a *spatial object* is a point, line, or region and the type of all valid spatial objects is denoted as the set $[S]$ [6]; a *moving object* is a moving point, moving line, or moving region as defined in [1] and the set $[M]$ represents the type of all valid moving objects. A *query window* is an instance of a moving object such that the type of query windows is equivalent to $[M]$. The type of time is $[T] = \mathbb{R}$. We are concerned with interpreting the interaction of a query window and a set of moving objects in order to compute aggregate operations relating to the spatial and temporal properties of moving objects.

The intersection of a moving object and a query object defines a *motion window*. For example, Figure 1a depicts a moving region with its motion indicated by the arrow, and a 1-dimensional query window that the moving region crosses as it progresses through time. Figure 1b shows a 3-dimensional representation of the moving region and the query window, with space as the first two dimensions and time as the third dimension, as they exist over a time interval. The motion window, the intersection of the moving region and the query window, is shown as the vertical oval contained in the query window.

The function *extract* : $[M] \times [T] \to [S]$ takes the spatial object at a given time instant from a moving object. Let m be a moving object that interacts with a query window q for a time interval from t_0 to t_1, and let w be the motion window defined by the intersection of m and q. For $t_i, t_j | t_0 \leq t_i <$

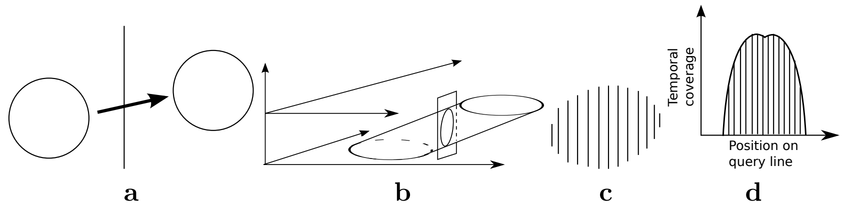


Fig. 1. Visualizations of a moving region, a query window, and their interactions.

$t_j \leq t_1$ and $a = extract(w, t_i)$ and $b = extract(w, t_j)$, the symmetric difference $a - b$ indicates the portion of the query window in the time interval (t_i, t_j) in which m was involved in a topological transition with q; in other words, the moving object made a qualitative change with respect to its topology at the query window over that time interval. For example, the boundary of the motion window in Figure 1b depicts the points in space and time when the boundary of the moving region transitioned from being disjoint with the query window to being in contact with the motion window. The area inside the motion window depicts the points in space and time when the interior of the moving region is in contact with the query window. We denote the concept of the transition of the topological relationship between a moving object and a query window *topological transition entropy (TTE).* The rate of TTE that occurs over a time interval, the *topological transition entropy rate (TTER)* is a quantifiable measure of the rate of topological transition of a moving object over a query window. A value for a TTER is implicitly represented in a motion window. Details on computing measures of TTER are discussed in the following sections.

4 Conceptual Framework

The example used in Section 3 to introduce the notion of TTER utilized a moving region and query window defined as a moving line that was static in terms of space over a time interval (i.e., the moving line did not actually move in the spatial dimensions across the interval). Because the query window is defined as a moving line, it can move through space over time; thus, we have two possible query scenarios that may affect the notion of the TTER across a query line: a query window that is spatially static or spatially dynamic. In general, we can manipulate concepts of space and time to identify four combinations of the space/time states (the states being *static* and *dynamic*) that a query window of any dimension can conform to. In Section 5, we explore the implications of each of the four space/time state combinations in terms of the semantic meaning of query windows that fit each category, the implications in terms of TTER computations, and the meaning of queries for each category.

5 Implications for Queries

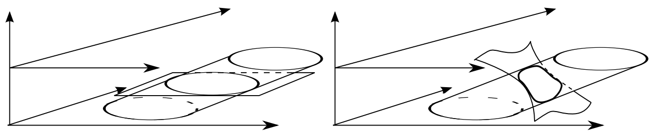
Static Space and Static Time (SS): The static space/static time combination for a query window describes a situation in which an n-dimensional query window is fixed in both space time. This corresponds to traditional window queries in which a user may ask "return the portions of all moving regions in a database that intersect with query window q at time t". Aggregate operations available in this situation involve finding the area in the query window that is covered by the highest or lowest number of objects, as defined in [4], among others.

Static Space and Dynamic Time (SD): Static space/dynamic time (SD) queries occur when a query window is held constant in space, but allowed to exist across a time interval; the example discussed in Section 3 and illustrated in Figure 1 corresponds to this type of query. This type of query is useful, for example, to a user who wishes to discover the area along a line segment which is covered for the longest time by a moving region. The query line is spatially static, but is projected through time for the duration of a user-defined interval. Figure 1 depicts an example scenario in which a region moves over time across a query line. The interaction of the query line and the object may be computed by sampling the object at the query line at discrete time intervals, or by computing the motion window induced by the interaction of the object and the query line.

The motion window induced by a query object and a query line encodes two useful pieces of information: i) the duration of coverage of the query window by the moving object, and ii) the TTE of the moving object as the query window progresses in time. The duration of coverage across the query line is computed by recording the height of the motion window in the temporal dimension. Coverage duration is most intuitively represented as a graph, such as Figure 1d for the motion geometry in Figure 1b, which we denote the *coverage duration graph*.

The TTER for the motion window in Figure 1b is most intuitively described from a discritized point of view. Consider the sampled interaction of a moving region and query window in Figure 1c: the TTER between each successive sampled interaction of the moving object and query window in Figure 1b is computed as the difference in the length of adjacent samples in Figure 1c. In essence, the amount of topological transition of the moving object across the query line is embodied in the difference of lengths, in this case, of adjacent sampled lines of the motion window. From a continuous point of view, the transitional entropy rate at any instant is computed from a motion geometry by summing the absolute value of the slope of the boundary of the motion geometry in the vertical direction wherever it intersects a specified time instant. In the general case, the motion window may be a 3D line, 3D surface, or volume depending on the dimensionality of the input moving object and query window. Thus, the general form of the computation of TTER at a particular time instant t for a motion window is the sum of the absolute values of the partial derivative in the z direction of the function describing the boundary of the motion geometry at all points at which the boundary of the motion geometry is defined at $z = t$.

For the SD case, two types of operations are implied for a query window coupled with a single object: i) the duration of the interaction of the moving object and query window defined as the height in the z direction at all points of the resulting motion window, and ii) the TTER of the object in the motion window defined as the 1st order derivative in the z direction of all points on



a **b**

Fig. 2. Examples of a DS and DD query window.

the boundary of the motion window. Clearly, the TTER can be aggregated for a single motion window to find the average TTER, the area of max TTER, or the area of minimum TTER. In the case of applying these operations to sets of moving regions, aggregate operations are available that are currently not possible in systems: the area representing the max/min/avg amount of temporal coverage of a set of moving objects over a query window, and the area representing the max/min/avg amount of TTER for a set of moving objects over a query window. To compute the aggregate operations, the motion windows computed from each object are projected out of time into a map of regions such that each region represents a constant time coverage or TTER, depending on the query type. The procedure in [4] is then used to compute the appropriate aggregate operation.

Dynamic Space and Static Time (DS): The dynamic space/static time (DS) combination of a query window is a rather non-intuitive concept. Essentially, a query window is allowed to exist at multiple spatial locations at a single instant in time. One can conceptualize the query as a SD query in which the roles of the moving object and the query window are reversed. Figure 2a illustrates a situation in which a query window defined by a line segment that is projected through space creates a motion window with a moving region. Again, the moving object can be a moving point, line or region, and the query window can be defined as a point, line, or region projected through space. The types of queries available in this construction are similar to the queries available in the DS construction, except that the motion window in the DS construction includes a temporal component, while the query windows in the SD construction represent spatial components. Thus, instead of time coverage and TTER, the DS construction provides concepts of space coverage and spatial transition entropy rate (STER). Essentially, the query types, and associated aggregates are constructed in the same manner, but are taken in a spatial direction. Because time is one-dimensional, time coverage and TTER are computed only in the vertical direction. The corresponding concepts of space coverage and STER can be taken in any spatial direction. The concept of space coverage as defined as time coverage in a spatial dimension, is not new; however, the concept of STER is novel. The STER in a spatial dimension provides a quantitative value describing the irregularity of a moving object's boundary at a specific time instant. For example, a region whose boundary is star shaped, or that has numerous concavities will have a higher STER value than a square.

Dynamic Space and Dynamic Time (DD): The dynamic space/dynamic time (DD) combination of a query window and a moving object describes the case in which the query window is essentially allowed to be a moving object, that is, it is allowed to change over both space and time. For example, a query window might represent a cold front that moves over time, and query objects may be areas of pollution that are moving due to wind, thus the user is interested in discovering the areas of greatest movement in the pollution regions as they interact with the cold front. Because the DD construction allows a query window to move over time, the concepts of temporal coverage and TTER once again apply. Figure 2b depicts a moving region, a query window defined by a moving line, and the motion window induced by the scene. The coverage duration of the region by the query window at any point p in space is defined as the length of the intersection of a vertical line extending in the temporal dimension at point p and the motion window. Furthermore, the temporal coverage can be taken along the surface of the motion window to more closely align with the SD construct. Similarly, spatial coverage can be computed using a spatial dimension.

Similar to temporal coverage, two options exists for computing the TTER. A user can compute the TTER in the z direction for each time instant along the boundary of the motion window, resulting in a TTER representing only the changes along vertical movement. More interesting is the case when the TTER is taken along a direction parallel to the motion window; such a computation provides entropy values along the query window. Finally, because change occurs in the spatial dimension in the DD case, STER values may be computed as well.

6 Conclusion

We introduced the concepts of coverage duration, TTER and STER, and systematically explored their interpretations by manipulating the spatial and temporal components of query windows. The result are new operations and aggregates that are not possible given existing query constructs in spatiotemporal databases.

References

1. Güting, R.H., Böhlen, M.H., Erwig, M., Jensen, C.S., Lorentzos, N.A., Schneider, M., Vazirgiannis, M.: A Foundation for Representing and Querying Moving Objects. ACM Trans. Database Syst. 25(1), 1–42 (2000)
2. Lazaridis, I., Mehrotra, S.: Progressive approximate aggregate queries with a multiresolution tree structure. SIGMOD Rec. 30(2), 401–412 (May 2001)
3. Lopez, I., Snodgrass, R., Moon, B.: Spatiotemporal aggregate computation: a survey. IEEE Transactions on Knowledge and Data Engineering 17(2), 271–286 (2005)
4. McKenney, M., Olsen, B.: Algorithms for fundamental spatial aggregate operations over regions. In: Proceedings of the ACM SIGSPATIAL BigSpatial. pp. 55–64 (2013)
5. Papadias, D., Kalnis, P., Zhang, J., Tao, Y.: Efficient olap operations in spatial data warehouses. In: Proceedings of SSTD (2001)
6. Schneider, M., Behr, T.: Topological relationships between complex spatial objects. ACM Trans. Database Syst. 31(1), 39–81 (Mar 2006)
7. Wolfson, O., Sistla, P., Xu, B., Xu, J., Chamberlain, S.: Domino: Databases for moving objects tracking. SIGMOD Rec. 28(2), 547–549 (Jun 1999)
8. Worboys, M.F.: A unified model for spatial and temporal information. The Computer Journal 37(1), 26–34 (1994)
9. Zhang, D., Tsotras, V.: Improving min/max aggregation over spatial objects. In: Proceedings of ACM GIS (2001)