Probabilistic Kinetic Data Structures
[Abstract]

F. Betul Atalay
Dept. of Computer Science
Saint Joseph's University
5600 City Avenue
Philadelphia, PA 19131
fatalay@sju.edu

Sorelle A. Friedler
Dept. of Computer Science
Haverford College
370 W. Lancaster Avenue
Haverford, PA 19041
sorelle@cs.haverford.edu

Dianna Xu
Dept. of Computer Science
Bryn Mawr College
101 N. Merion Avenue
Bryn Mawr, PA 19010
dxu@cs.brynmawr.edu

Abstract

We introduce a probabilistic kinetic data structure that uses a motion model to maintain a specific geometric structure to within a user-specified confidence value. We also introduce a certificate-based error model and accompanying problem-specific proofs that the robust error model is implied by this new, more geometrically generalizable, error model. We show that the probabilistic kinetic data structure translation from a classic kinetic data structure (KDS) is efficient and approximately correct under the certificate-based error model based on the user-given confidence value. Significantly, this means that any existing KDS is efficient under this probabilistic framework.

Introduction

Given objects in motion, such as people on cell phones, cars, etc., we consider the question of how to calculate and update information about the resulting geometric structures, such as the convex hull, as they change with the object motion over time. We focus on a theoretical framework for this question that incorporates real-world elements of related topics from outside of the computational geometry community by allowing the point motion to be predicted probabilistically. We build on a large body of work on kinetic data structures in such a way that existing solutions can be immediately adapted for use in this probabilistic setting.

The motion framework that has gained the most traction within the computational geometry community is the kinetic data structures framework (KDS), first introduced by Basch, Guibas, and Hershberger in 1997 [1]. KDS assumes that points follow flight plans given as polynomial curves (over space and time) of bounded degree. These polynomials are known in advance. *Certificates*, representing fragments of knowledge about the geometric structure under calculation, guarantee that points participate in a certain geometric relationship with each other. *Events* are scheduled at times when certificates are known to fail based on the given individual point trajectories. The geometric maintenance is then accomplished by moving time forward and handling events via correcting the failing certificates so that the geometric structure is correctly maintained at all times. The maintenance of many geometric structures under motion has been considered within this framework, including the convex hull, minimum spanning tree, kd-trees and other hierarchical data structures, clustering problems, medians, spanners, and the Delaunay triangulation [3].

While the KDS framework has been extremely successful, it still presents many issues in modeling real-world motion. Foremost among these are the issues that points are assumed to follow flight plans that are exact and known in advance, both of which obviously limit the use of this framework in real-world experimental settings. Additionally, one issue with requiring the point motion to be known in advance is that this prevents the KDS framework from being used in an online manner, despite its inherent relationship to online algorithms seen in its incremental structural updates over time. Previous effort has attempted to address these issues by considering work analyzed per time step under assumptions such as the maximum displacement of a point over a single time step [2]. Here, we are interested in creating a framework that assumes limited knowledge of point motion in advance and importantly allows the geometric structure to be updated only when we predict that an update is required. The tradeoff is that the structure will be maintained approximately instead of exactly.

Framework and Results

We introduce *probabilistic kinetic data structures* (PKDS), a modification of *kinetic data structures* (KDS) that operates under probabilistic knowledge of point motion. We broaden the assumptions about the trajectories that points follow to include motion as understood by a probabilistic model of the type widely used within the machine learning community to predict object locations over time [4]. Specifically,

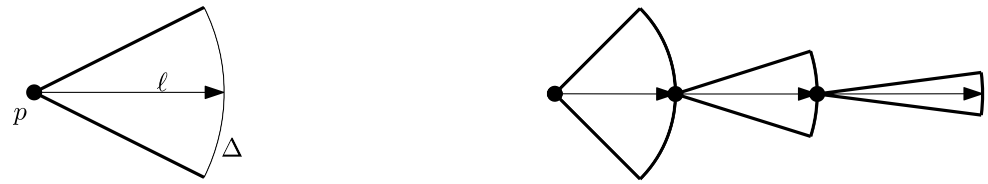

Figure 1: Left: A single pie slice showing the predicted direction (implying a linear trajectory) with highest confidence ℓ beginning from the current position p and continuing for as long as the model is confident until time Δ. Right: The improving predicted pie slices from a model as a point with a linear trajectory moves from left to right.

we assume that when given situation-dependent observations about the surroundings, such a *model* can predict future object motion according to a Gaussian distribution over the direction of the object and that this predicted direction decays according to another Gaussian distribution over the time from the last query to the model. When considering only predictions that are confident up to some probability, we can visualize this output from the model as a *pie slice* where the direction of highest confidence is the lengthwise center line of the pie slice ℓ, the current predicted position is the point p, and the curved far edge Δ represents the limits of our confidence in the model as it decays with time (see Figure 1). We assume that queries to this model take logarithmic time, i.e., they are fast enough that they can be made frequently, but not so fast that it would make sense to continuously query for their predicted position. We also assume that the model takes history into account, so that it becomes more confident as the motion is more predictable over time.

In conjunction with our point prediction model, we allow the user of the system to input a desired *confidence value*, ϕ, indicating the probability that the resulting geometric structure will be correct at any time. The correctness of the structure is determined according to a *robust error model*, more commonly known as a robust statistical estimator, which allows a structure to be robust to outliers up to some breakdown point [5]. Under a *certificate-based error model* that we introduce here, a structure is considered some percentage, ϕ, correct if ϕ percent of the total certificates used to calculate the structure are correct. We show that a solution that is ϕ-correct within the certificate-based error model is also ϕ-correct within the robust error model for the problems of maintaining the 1D maximum as well as the convex hull. Importantly, while the robust error model must be defined based on the characteristics of each problem's geometric structure, the certificate-based error model is directly understood from the certificates.

Within the PKDS model, we modify the KDS idea of events and certificates to include both structural events and time events. *Structural events* are the same as the original KDS events in the sense that they represent changes in the geometric structure under maintenance and imply that a repair of a certificate may be required. Structural events are scheduled by considering the pie slices as if all contained point locations are possible. The certificate failure that is at the minimum time over all possible certificate failures when all such points are considered is the structural event that is scheduled. We show that scheduling structural events in this way ensures that the geometric structure is updated enough so that it is expected to be correct up to the given confidence value under a certificate-based error model.

In addition, our framework includes *time events*, which are indications that our certainty about the point's location has become too low and another query to the model is required. We demonstrate that the cost incurred by adding these time events and by making the structural events dependent on the outline of the pie slice and not a single trajectory is not too much in the sense that solutions to the original KDS framework can be shown to work under the PKDS framework with only small modifications and efficiency cost.

References

[1] J. Basch, L. J. Guibas, and J. Hershberger. Data structures for mobile data. In *Proc. of the Symposium on Discrete Algorithms*, pages 747–756, 1997.

[2] M. de Berg, M. Roeloffzen, and B. Speckmann. Kinetic convex hulls, delaunay triangulations and connectivity structures in the black-box model. *Journal of Computational Geometry*, 3(1):222–249, 2012.

[3] L. Guibas. Kinetic data structures. In D. Mehta and S. Sahni, editors, *Handbook of Data Structures and Applications*, pages 23–1–23–18. Chapman and Hall/CRC, 2004.

[4] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, 2001.

[5] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection.* Wiley & Sons, New York, 2005.