SAND2020-2839C

Tracktable: Trajectory Analysis & Visualization

Joe Bertino jbertin@sandia.gov, (505)844-0527

Sandia National Laboratories, Albuquerque, NM 87185-0810

ABSTRACT

Tracktable is a tool developed by Sandia National Laboratories to analyze the paths of moving objects, called "tracks" or trajectories. Tracktable transforms each trajectory into a set of numbers, called a feature vector, representing various attributes of a trajectory. Using this representation, modern machine learning techniques can be leveraged to enable efficient analysis of large sets of trajectories. A wide variety of analyses are made possible by this approach. For example, unsupervised learning can be applied via internal clustering methods to find similar types of activities with minimal user descriptions of the activities. In addition, collective behaviors can be detected, such as activities that are correlated in time and/or space (co-traveling or meetings). Another common use of Tracktable is to establish normal patterns of behavior such that outliers can be more easily detected. Finally, analyst derived patterns could be easily described and monitored for, allowing specific "trip wires" to be established as a part of a study.

Tracktable has been built and tested in a number of development environments. This includes the easy-to-use Jupyter notebook environment where widgets can be added, yielding an interface for analysts that includes easy-to-set parameters and interactive map rendering.

PROJECT DESCRIPTION

Tracktable is a set of trajectory analytics that can be used to apply advanced machine learning techniques to large trajectory data sets. It enables searches for shapes and patterns in space and/or time by providing a mathematical framework to describe such patterns, and it also enables a very fast search capability that provides a means for searching through millions of patterns quickly. Tracktable additionally provides the ability to apply traditional *unsupervised* learning of trajectory patterns that enables the computer to group similar shapes together and to find unique outlying trajectories. Additionally, its ability to treat time as a variable similar to space enables searches for collective behavior and patterns over long periods of time. Finally, fast search techniques allow observed trajectories to be searched for in historical databases in order to make predictions based on comparing the existing path. It is easily deployable through the Jupyter notebook interface as a pip-installable Python library.

Significance of Technology: Remarkable advances in artificial intelligence have enabled computers to do sophisticated analysis on text documents and images. However, ubiquitous GPS technology has recently led to a flood of data stamped with time and location. The ability to have computers search through large amounts of trajectory data quickly has the ability to significantly impact national security, along with understanding how to optimally manage the increasing motion traffic in crowded areas.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Wow! Factor: Most previous methods for analyzing trajectories were focused on complex one-to-one geometric comparisons using curve alignment types of techniques. This limited both the type and number of trajectories that could be easily analyzed. By using our patented representation techniques, we can massively increase the number of trajectories to be analyzed while at the same time enabling many new types of analysis.

Competitiveness: The notion of fundamentally representing the trajectory as a vector of "features" is the key idea that makes this approach different from previous trajectory analysis techniques. This allows us to do complex geometric comparisons by simple point operations as opposed to complex geometric operations, thus allowing us to handle up to a billion trajectories instead of hundreds or a few thousand trajectories. We have successfully patented this technique (US Patent 10,345,106) to protect this unique ability.

Use Case

The ability to represent a trajectory as a feature vector --- an n-dimensional vector of float values --- allows us to leverage a collection of geometric and temporal metrics, including distance geometry, as a succinct representation of complex paths. We can then use robust clustering techniques, such as DBSCAN, to rapidly organize and categorize millions of trajectories based on their shapes, as well as identify outliers in our dataset. Our feature set representation of trajectories also allows us to index trajectories based on their shape and perform retrieval of trajectories whose shape matches user-defined patterns.

Figure 1 (below) demonstrates Tracktable's power in finding trajectories with specific distance geometries. We did not draw these tracks---these are actual flight paths in our dataset, and we simply provided Tracktable with feature vectors of the letters ['P', 'A', 'N', etc.] and said, "can you find examples of these?"

Academic Partners

Sandia National Labs has partnered with Purdue University to provide applied trajectory problems to student residents of an on-campus data-mining program. The students gain experience solving real-world, large-scale analytic problems designed by Tracktable staff. Sandia benefits by tapping into a pool of highly motivated and technical undergraduates who have the potential to migrate to Sandia as either interns or full-time staff upon graduation.

Biography

Joe Bertino has an M.S. in Computer Science from New York University and a B.S. in Mathematics from Fordham University. He is currently developing geospatial analytics for Tracktable, focusing on (1) clustering trajectories from spatiotemporal features and (2) aiding machine learning models in classifying patterns of movement. Joe is also conducting Vulnerability Research on software written in C.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.