SANDIA REPORT

SAND2016-11037
Unlimited Release
Printed October 2016

Review of Tracktable for Satellite Maneuver Detection

Erin C.S. Acquesta, Christopher G. Valicka, Mark B. Hinga and Carollan B. Ehn

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online ordering:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2016-11037
Unlimited Release
Printed October 2016

Review of Tracktable for Satellite Maneuver Detection

Erin C.S. Acquesta
ISR Systems Engineering and Decision Support
eacques@sandia.gov

Christopher G. Valicka
Tasking, Planning, Mission Management
cgvalic@sandia.gov

Mark B. Hinga
Tasking, Planning, Mission Management
mbhinga@sandia.gov

Carollan B. Ehn
Analytics and Cryptography
cbehn@sandia.gov

Abstract

As a tool developed to translate geospatial data into geometrical descriptors, Tracktable offers a highly efficient means to detect anomalous flight and maritime behavior. Following the success of using geometrical descriptors for detecting anomalous trajectory behavior, the question of whether Tracktable could be used to detect satellite maneuvers arose. In answering this question, this report will introduce a brief description of how Tracktable has been used in the past, along with an introduction to the fundamental properties of astrodynamics for satellite trajectories. This will then allow us to compare the two problem spaces, addressing how easily the methods used by Tracktable will translate to orbital mechanics. Based on these results, we will then be able to outline the current limitations as well as possible path forward for using Tracktable to detect satellite maneuvers.

Acknowledgment

First we want to thank David Cox, Kristina Czuchlewski and Drew Woodbury for presenting the idea and research opportunity to explore the potential benefits of using Tracktable for satellite maneuver detection. In addition, we want to acknowledge Danny Rintoul and Andy Wilson for the support they offered as we worked toward building Tracktable onto a local system so that we could run our analysis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Tracktable for Geospatial Trajectory Analysis

In this section we will introduce a brief description for Tracktable's current functionality. For more information regarding Tracktable please reference Danny Rintoul *et al*'s Sandia report, *PANTHER: Trajectory Analysis* [4].

The idea behind converting the position data into geometric descriptors was presented as a means to handle the inefficiency for which trajectory analysis has been evaluated in the past. The earlier methods required a one-to-one comparison between each pair of trajectories. Metrics comparing the distance between all of constituent position data were used to conclude that two trajectories were similar. This can clearly become computationally expensive very quickly. Alternatively, with the introduction of Tracktable the process for considering whether trajectories are similar follows the given generalized procedure.

Tractable Trajectory Comparison Method:

1. The raw input data for aircrafts and ships, given by geospatial position data that varying over time, is ingested and processed from a database of flight or maritime tracking information.

   Geospatial Position Data:

   (a) timestamp
   (b) latitude
   (c) longitude
   (d) altitude (sometimes, but often unreliable)

2. For an aircraft or ship trajectory, with a unique identifier, a trajectory is built by organizing the corresponding position data chronologically.

3. For any given trajectory, a set of geometrical descriptors can be computed to uniquely describe its shape or other features. This allows for a much faster means to detect similar, and also anomalous, behavior.

   Examples of Geometrical Descriptors:

(a) Total distance traveled

(b) Total curvature

(c) Eccentricity of the convex hull

(d) Start/Stop point

(e) ect...

4. With each trajectory now described by a set of geometric descriptors, a spacial indexing scheme is used to store this information to allow for quick processing.

5. Clustering based on geometric descriptors results in similar trajectories getting grouped together and trajectories not sharing geometric descriptions with other trajectories being labeled as outliers. These outliers then determine the set of anomalous flight or maritime trajectories.

In addition to introducing the conversion from geospatial position data to geometrical descriptors, Rintoul et al also introduce a variation to traditional DBSCAN (Density-Based Spatial Clustering of Applications with Noise). The new method, implemented by Tracktable, is referred to as box DBSCAN where clustering is determined by feature specific parameter definitions as opposed to a single spherical neighborhood defined by traditional DBSCAN. The idea is as follows. Instead of scaling the data so that a single parameter can be used to define proximity across multiple dimensions, box DBSCAN defines a parameter for each dimension of the data. Thus, we end up with a neighborhood shaped as a hyperrectangle (or box in 3-dimensions). As we consider using Tracktable for satellite maneuver detection, we will start by implementing box DBSCAN. To do so more effectively we will need to introduce the fundamentals of astrodynamics for satellite trajectories in the following section.

Astrodynamics of Satellite Trajectories

"Astrodynamics is the study of the motion of man-made objects in space, subject to both natural and artificialy induced forces" -Vallado [5].

The man-made objects Vallado refers to are in reference to rockets, spaceships and satellites, but the focus of our research is specific to the trajectories that satellites take once they are in orbit. Our discussion regarding astrodynamics will focus on the elliptical orbits that satellites follow around an astronomical body. More specifically, we focus on low earth orbiting (LEO) satellites, primarily due to the nature of the data accessible for our analysis.

To lay a foundation for detecting satellite maneuvers, we will begin with a brief description of astrodynamics. If more detail regarding astrodynamics is needed, please reference Vallado's book *Fundamentals of Astrodynamics and Applications* [5].

To begin we will start with the six standard orbital elements (Keplerian elements), required to interpret the orbit of a satellite.

1. Semi-Major Axis, a

2. Eccentricity, e

3. Inclination, i

4. Argument of Periapsis/Perigee, ω

5. True Anomaly, ν

6. Right Ascension of the Ascending Node (RAAN), Ω

All of which are illustrated in figure 1.1. These six orbital elements are required for representing, unambiguously, the satellite in its unperturbed ideal orbit in three dimensional inertial space and time. These six elements correspond and can be converted to another expression, in six degrees of freedom, namely three spatial (position) dimensions and three displacement (velocity) dimensions. (Specifically and respectively, X position, Y position, Z position, velocity in the X, Y, and Z directions.)

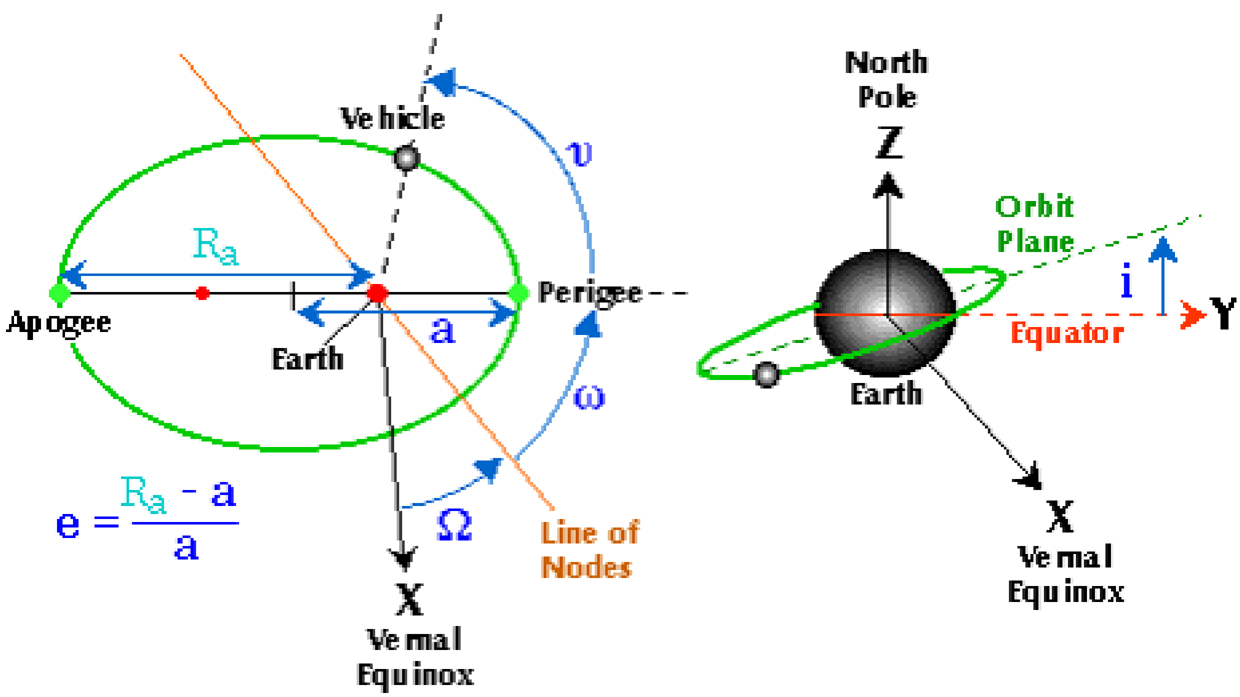

Figure 1.1. Illustrative interpretation of the orbital elements.

This latter form of describing the orbit is often termed the "orbital state vector". Both forms of representation require a so called "epoch". Some may call it a seventh element. This is a time stamp that declares the moment in history when these six variables were computed to describe the satellite's orbit and where in the orbit the satellite is located.

Because a satellite also experiences non-conservative forces, i.e. solar pressure, atmospheric drag, perturbations from distant celestial bodies, etc, the motion of the satellite is perturbed away from its ideal Keplerian motion. Therefore, this two-line element set (TLE) (which is an ideal orbit solution) will become “stale” after a certain time, requiring new data to be collected and a new TLE to be estimated for the satellite of interest. An example with a description for interpreting a TLE is presented in Table 1.1

0 Envisat
1 27386U 02009A 10001.24505451 -.00000032 00000-0 42006-5 0 9998
2 27386 098.5500 070.7584 0001076 091.7123 358.2472 14.32247531409841

Line 1

Line number	1
Satellite number	27386
International Designator (launch year)	02
International Designator (launch # of the year)	009
International Designator (piece of teh launch)	A
Epoch Year (last two digits)	10
Epoch (fractional part of the day)	001.24505451
First Time Derivative of Mean Motion divided by 2	-.00000032
Second Time Derivative of Mean Motion divided by 6	00000-0
BSTAR drag term (decimal assumed)	42006-5
The number 0 originally ”Ephemeris type”	0
Element set number	999
Checksum(modulo 10)	8

Line 2

Line number	2
Satellite number	27386
Inclination (degrees)	098.5500
Right ascension of the ascending node (degrees)	070.7584
Eccentricity (decimal assumed)	0001076
Argument of Perigee (degrees)	091.7123
Mean Anomaly (degrees)	358.2472
Mean Motion (revolutions per day)	14.32247531
Revolutions number at Epoch (revolutions)	40984
Checksum (modulo 10)	1

Table 1.1. Description of the components in a TLE.

From the TLE data we are able to collect all six of the orbital elements. Some directly, while true anomaly and the semi-major axis can be derived from mean motion and the mean anomaly. Although it can be argued, from a data analysis perspective, deriving true anomaly and the semi-major axis are just translations of the data and that processing mean motion and mean anomaly may be sufficient. This can be explored in more detail when we consider feature and parameter selections for a given detection algorithm.

Considering the earlier discussion regarding geometric descriptors, it does seem as though the orbital elements align nicely with the methods used for geospatial trajectory analysis. Although in this case, we will emphasize that the raw data collected for satellite tracking is already in the form of geometric descriptors. This implies that it is not necessary to process the data for determining a trajectory as well as generate geometric descriptors. Therefore, we will start the preliminary analysis by implementing box DBSCAN for detecting outliers from a collection of features from the raw TLE data in chapter 3.

This page intentionally left blank.

Chapter 2

Problem Space Comparison

Anomalous Flight Trajectories

To give a more concise description between the two problem spaces we will focus primarily on detection of anomalous flight trajectories. The consideration for maritime trajectories has its own distinctions from flight trajectories, but none that align with the problem space for satellite trajectories. Therefore, we have determined it is sufficient to focus on the comparison between the detection of anomalous flight trajectories with the detection of satellite maneuvers.

In figure 2.1 we can see the results of detecting anomalous flight patterns from a database of geospatial position data for archived flight tracking information. Each one is very unique in its pattern. These are the flights that can later be addressed in more detail allowing an analyst to determine exactly why these flights do not match more common patterns found within other flight trajectories. This gives us the first point we want to emphasize.

> The classification for an anomalous flight trajectory is defined by the shape of the trajectory, as a whole.

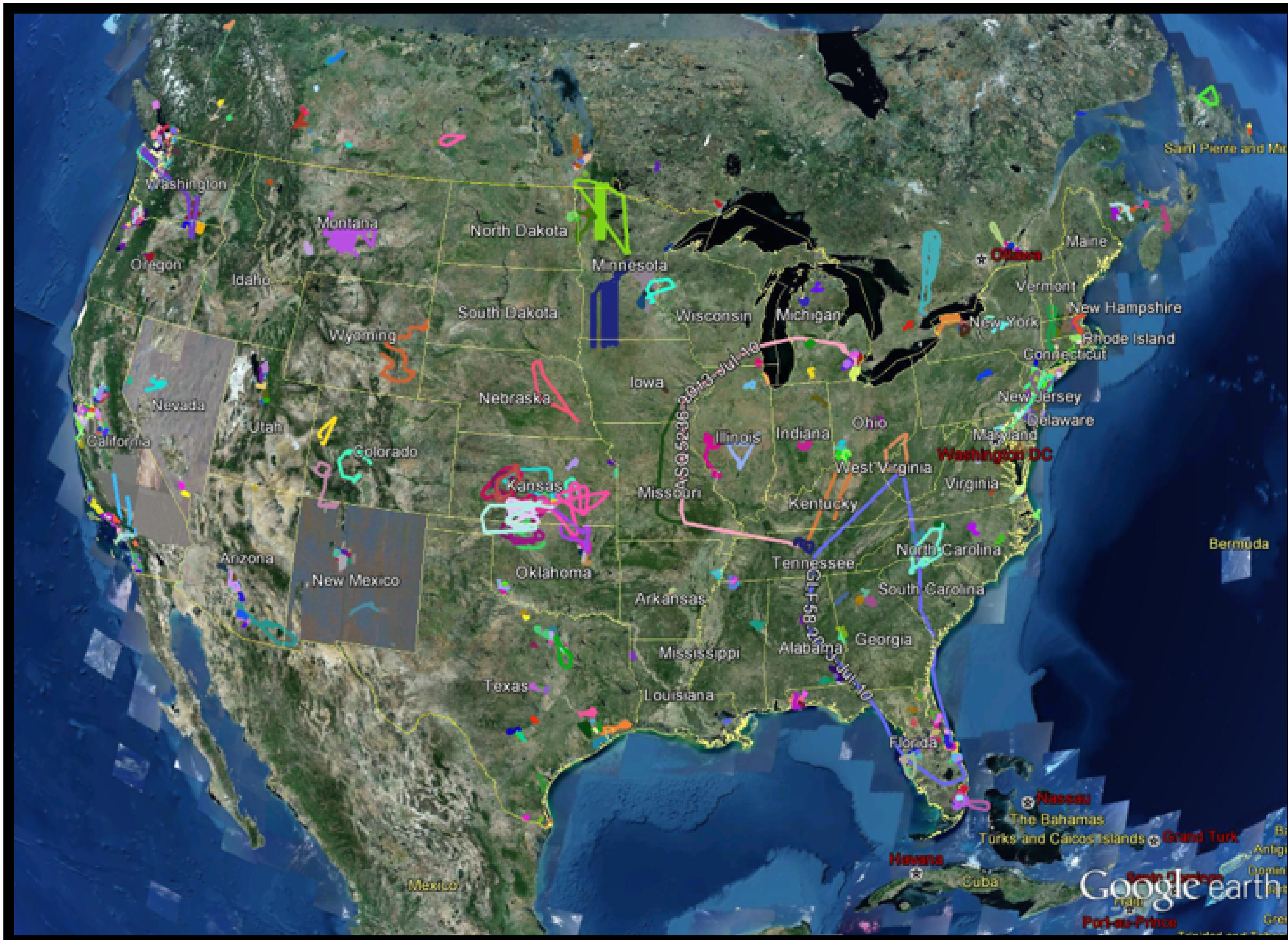

Figure 2.1. Each of the flight trajectories plotted on the graph were detected and labeled anomalous flights using Tracktable.

Satellite Maneuver Detection

With regards to detecting satellite maneuvers, we are asking the following question.

> Considering one particular satellite's orbit, where does the natural trend in the TLE data break?

This is very different from determining what trajectories have dissimilar shapes. Instead of looking at the whole picture, where the individual data points collectively describe a trajectory, we are now considering the individual data points and the patterns that are traversed as a single trajectory evolves. Figure 2.2 shows this concept as a graph comparing detected maneuvers to known maneuvers as plot of mean motion against epoch time.

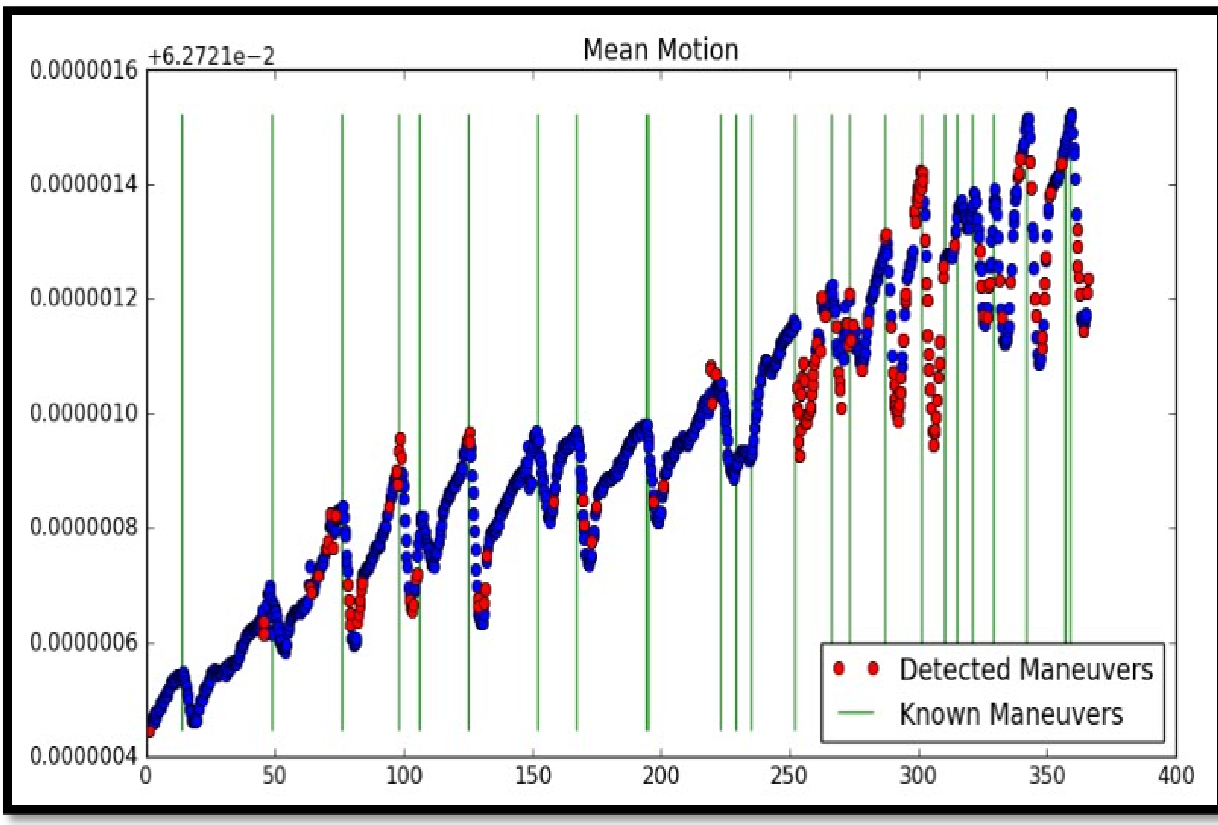

Figure 2.2. The outliers (labeled red) from the sequence of TLE data, plotted in terms of mean motion (radians per minute) against epoch day.

Can Tracktable be used to Detect Satellite Maneuvers?

Tracktable does have the potential to be a viable tool for detecting satellite maneuvers, especially since there are already discussions within the Tracktable development community to consider the sub-trajectory problem. The sub-trajectory problem considers how to break-up a trajectory to detect particular patterns that occur within the trajectories course. This type of analysis will align much better with the satellite maneuver detection problem, but it is not a current feature of the tool. Therefore, as a trajectory analysis tool today, Tracktable does not have the required functionality to detect satellite maneuvers. Although, there does seem to be a common path forward for which future functionality of Tracktable aligns well with the satellite maneuver detection problem.

We have already mentioned some concerns in the introduction and the comparison of the two problem domains. As a comprehensive list of the concerns we are currently aware of we have outlined the following

1. The problem space.

   Current Functionality: When detecting anomalous flight trajectories, an outlier is a flight trajectory that does not meet a particular similarity index when compared to thousands of other flight trajectories.

   Required Functionality: For satellite maneuver detection, an outlier is a single point within the tracking data for a particular satellite that indicates a sufficiently large[1] change in the orbital elements (or state vector).

   The Distinction: The classification of an outlier as an anomalous flight trajectory translates well since the conversion made to the geometric descriptors space implies that an outlier corresponds to the flight trajectory as a whole. The fact that the pattern for an anomalous flight does not agree with any other flight pattern in the database allows us to make this consideration that an outlier is an anomalous behavior.

   This concept does not translate well into the TLE descriptor space where the consideration for clustering is used to detect satellite maneuvers. As the raw data being ingested for analysis, the TLE data can unfortunately be contaminated with noise during the process of its calculations (by the NORAD ground segment). Therefore, assuming that an outlier translates directly to a maneuver, without pre-processing the data (or post-processing the results), will result with legitimately 'noisy' data points classifying maneuvers.

2. Data processing.

   Current Functionality:

   $$\boxed{\text{Position Data}} \Rightarrow \text{Trajectory Generated} \Rightarrow \text{Geometric Descriptors}$$

[1] What it means to be sufficiently large has yet to be defined.

Required Functionality:

$$\cancel{\text{Position Data}} \not\Rightarrow \cancel{\text{Trajectory Generated}} \not\Rightarrow \boxed{\text{Geometric Descriptors}}$$

If change in velocity is the primary focus for the detection algorithm, then the required functionality will extend to,

$$\boxed{\text{Geometric Descriptors}} \Rightarrow \text{SGP4} \Rightarrow \text{Position Data.}$$

The Distinction: Since the TLE data is already storing geometric descriptors, the satellite maneuver detection problem does not require the processing steps to generate a trajectory and derive the geometric descriptors. It can be argued that the 'noise' in the raw data for flight trajectories is managed in the process of converting geospatial position data into a trajectory. Alternatively, receiving the geometric descriptors in raw data form and jumping into clustering algorithms, the 'noisy' data will not have been filtered. Therefore, additional consideration of this issue has to be made when we use clustering techniques to detect satellite maneuvers.

With these caveats in mind we will consider a preliminary analysis of the box DBSCAN clustering approach to get a better understanding of this geometric descriptor type data, the pitfalls of the clustering method and develop an insight into possible paths forward.

This page intentionally left blank.

Chapter 3

Satellite Maneuver Detection

The initial research for detecting satellite maneuvers began with reviewing Kelecy *et al*'s paper *Satellite Maneuver Detection Using Two-line Element (TLE) Data* [2]. Instead of using a clustering approach, the authors consider a sliding window where a determination of possible maneuvers is defined by large changes in smoothed data (approximated by polynomial fits) between adjacent segments. These possible maneuvers are then run through a second round of analysis, where the differences derived in the first run of the data are then plotted against time. The peaks of this plot then determine the set of detections for a satellite's maneuvers.

Although we are considering a different approach, Kelecy *et al*'s paper gives a solid foundation for understanding the problem and outlining a methodology for analysis.

Box DBSCAN Approach to Satellite Maneuver Detection

Realizing that a clustering approach was successful in detecting anomalous flight behavior, we began our analysis of satellite maneuver detection by implementing a similar approach to analyzing TLE data. Recall that raw TLE data stores geometric descriptors regarding the orbits of satellites. Therefore we started with a direct clustering evaluation for satellite maneuver detection by determining large changes in a subselection of orbital elements.

Description of Box DBSCAN

The box DBSCAN algorithm is very similar to traditional DBSCAN with the one exception that a neighborhood for the box DBSCAN method is defined by a hyperrectangle, instead of a hypersphere. To implement box DBSCAN requires defining the ε-hyperrectangle neighborhood as well as $MinPts$.

1. ε-hyperrectangle, defines the neighborhood surrounding a point used to collect other 'nearby' points.

2. $MinPts$, is the minimum number of points within the ε-hyperrectangular neighborhood (including the point itself) needed to initiate a cluster.

As the data is processed and clusters are evaluated, each point is classified as one of the following

1. *Core point:* a point is evaluated to be a core point if it has at least $MinPts$ in its ε-hyperrectangle neighborhood.
2. *Reachable point:* is a non-core point for which a path of core points connected by ε-hyperrectangle neighborhoods can be used to reach it.
3. *Outlier:* Any point that does not have at least $MinPts$ in its ε-hyperrectangle neighborhood and can not be reached by any sequence of core points.

Clusters are defined by the core points that are a collection of overlapping neighborhoods, along with any reachable points. Then the outliers, or the points that are not reachable from any cluster, are determined to be the detected maneuvers.

The Satellite Data

To begin our analysis required both TLE data for an unclassified satellite and the known maneuvers for the same period of time. The collection of TLE data for the satellites is currently collected from TARDIS (TLE Analysis and Research Database Information Service) at `tardis.sandia.gov`[1]. The known maneuvers, for a handful of satellites, can be found on 'http://ilrs.gsfc.nasa.gov/data_and_products /predictions/maneuver.html'. Using these two sites to collect our data we are restricted to running our analysis for the Envisat LEO satellite (27386) from 2010 to 2012. Although we evaluated the Envisat TLE data from 2010 to 2012 and compared various orbital parameters to known maneuvers for all three years, the bulk of our research focuses on the year 2011.

Box DBSCAN Results

Features: Perigee, Inclination, Eccentricity and Mean Motion

For the feature space [Eccentricity, Inclination, Perigee, Mean Motion], we set the following parameters to detect satellite maneuvers using box DBSCAN:

$$\varepsilon_{[E,I,P,MM]} = [1e-05, 5e-05, 1e-01, 1e-08] \tag{3.1a}$$
$$MinPts = 5 \tag{3.1b}$$

[1]This is an internal site managed by Sandia Natinoal Laboratories and requires permission for access.

Confusion Matrix

	Positive	Negative	
True	18	1005	1023
False	175	7	185
	193	1012	1205

Table 3.1. The above confusion matrix shows the results of the detection algorithm when we use box DBSCAN with the feature space [Eccentricity, Inclination, Perigee, Mean Motion] with the epsilon hyperrectangle = [1e-05,5e-05,1e-01,1e-08]

Sensitivity (Recall)	72%
Precision	9%
Specificity	85%
Accuracy	85%
F_1 score	16%

Table 3.2. Evaluations of the detection algorithm using box DBSCAN with the feature space [Eccentricity, Inclination, Perigee, Mean Motion] with the epsilon hyperrectangle = [1e-05,5e-05,1e-01,1e-08]

From the results laid out in table 3.1, a determination for how well the detection algorithm did can be evaluated using a number of various metrics derived from the confusion matrix. For our analysis, we evaluated sensitivity/recall (proportion of true positives relative to all known maneuvers), precision (proportion of true positives relative to all detections), specificity (true negatives relative to the total non-detections expected), accuracy (total true positives and true negatives relative to the total data processed) and the F_1 score (harmonic mean of precision and recall). Typically in evaluating detection algorithms, sensitivity and specificity will usually be associated. These two metrics together give us an understanding for how well the algorithm classified the data. Alternatively, as another common means to evaluate detection algorithms, precision and recall can be considered, which gives an indication for how reliable the results are. We present these five metrics, sensitivity (recall), precision, specificity, accuracy and the F_1 score to give us a full understanding of the results.

From the results in table 3.2, we can see that the detection method using box DBSCAN with the parameters defined by 3.1a and 3.1b generated reasonable results for sensitivity (recall), specificity and accuracy. There were, however, highly inaccurate results for precision and consequently the F_1 score. As an interesting alternative to precision, if we instead consider the ratio of true positive results to false positive we get approximately 1:10. This implies that for every true positive, we get an average of 10 false positive detections. For this reason we will discuss the consideration of running a layered analysis approach in chapter 4.

Data and Cluster Analysis

In an attempt to better understand the results we received, we further analyzed the data as it relates to the known maneuvers. Figure 3.1 plots each of the analyzed TLE features against epoch time, where the vertical lines indicate the known maneuvers for the Envisat satellite in 2011. The most prevalent pattern that we have noticed from these results exists within the variation in mean motion, presented in figure 3.2. From the TLE data for the Envisat satellite in 2011, the pattern associated with a known maneuver consistently shows a decrease in mean motion for an average of 4-5 days immediately after the known maneuver occurs.

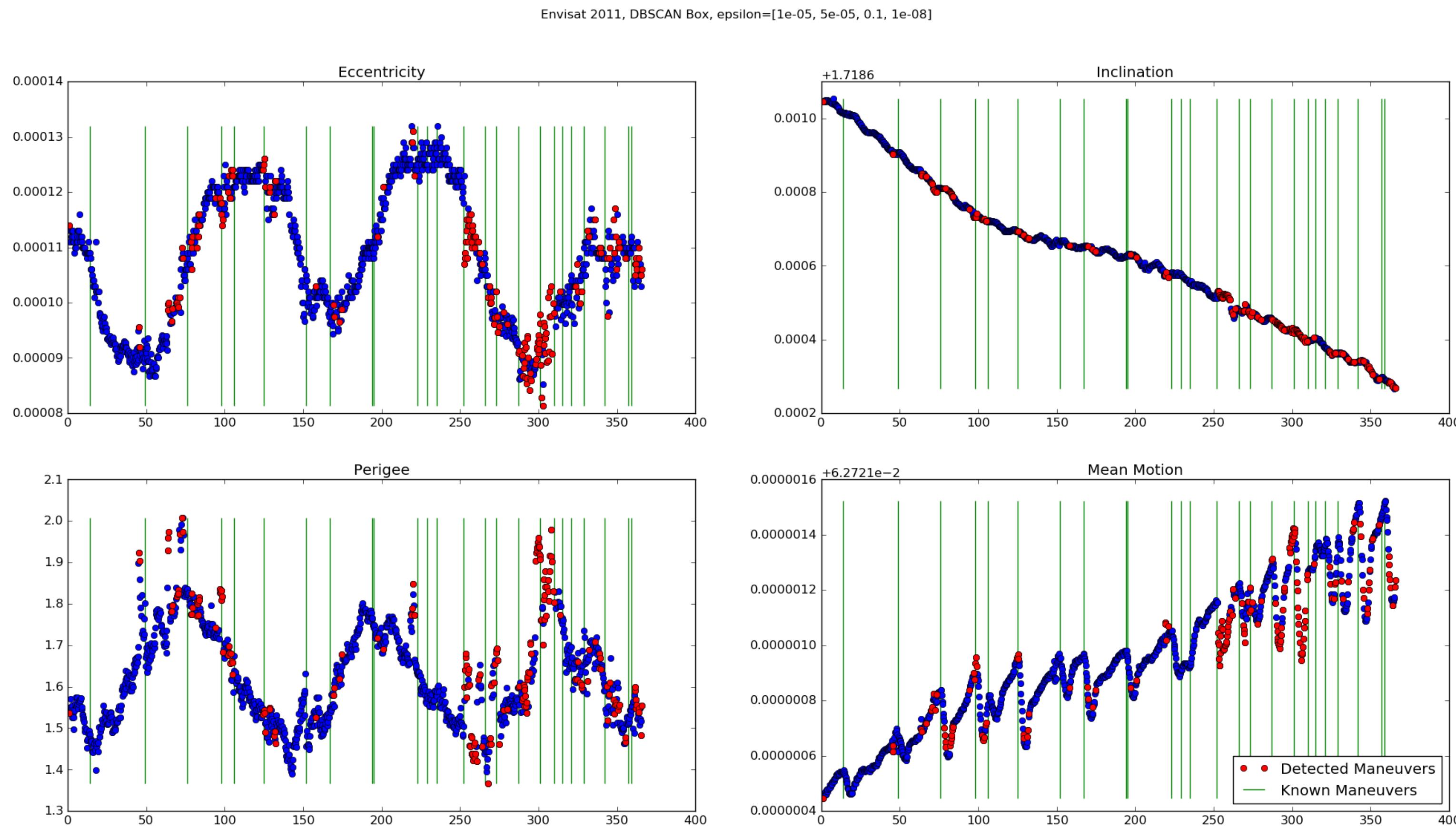

Figure 3.1. The graphs above show the relationship between the results of the satellite detection algorithm using box DBSCAN. Each red point shows a maneuver detection. The green bars indicate the days on which a known maneuver occurred. The blue points represent the rest of the TLE data that were not classified as outliers.(Note: Mean Motion is measured in radians per minute.)

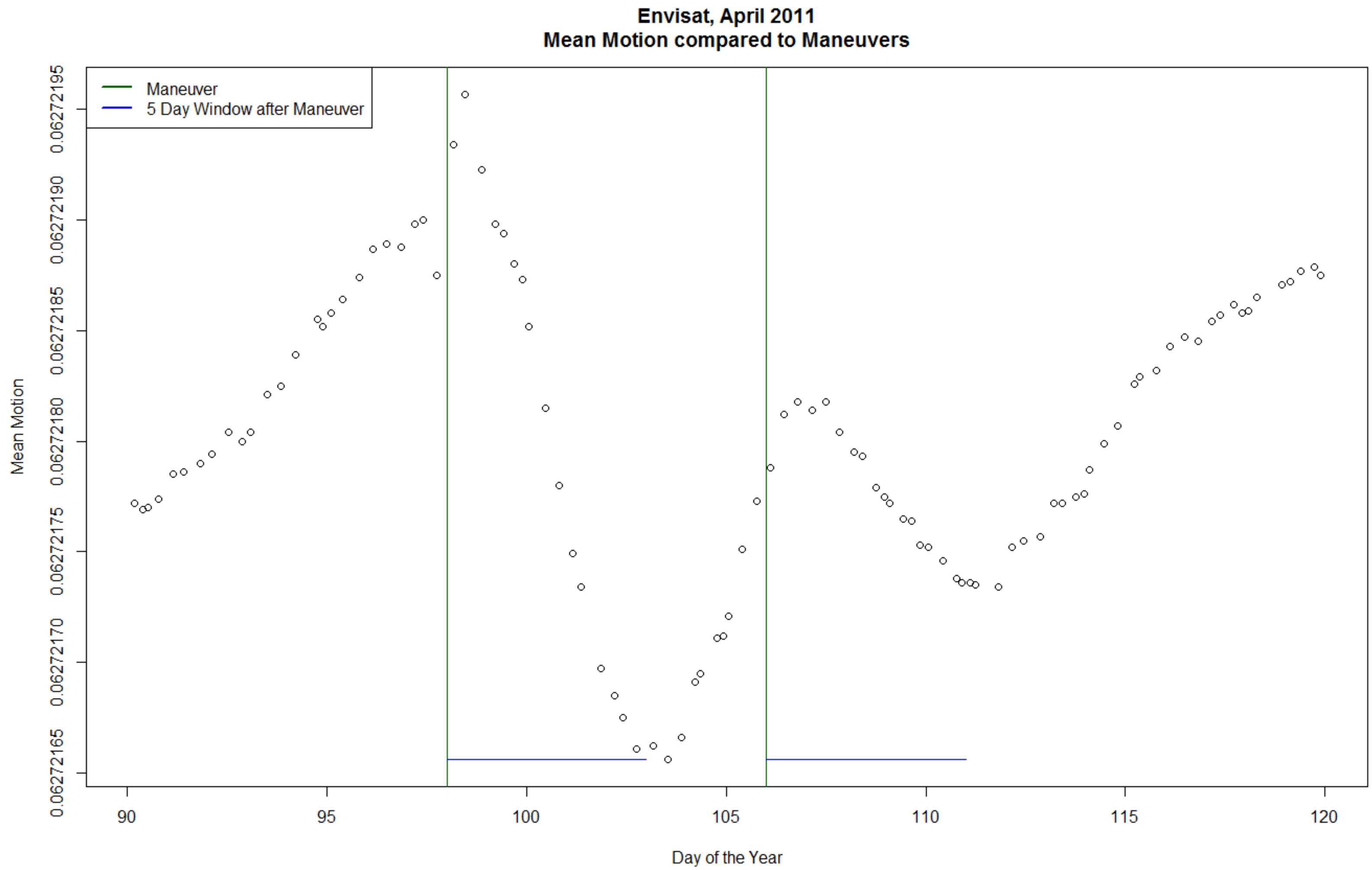

Figure 3.2. Mean motion pattern associated with known maneuvers for the Envisat satellite in 2011.(Note: Mean Motion is measured in radians per minute.)

There were 25 known maneuvers for the Envisat satellite during the year 2011, 16 of which show consistency with the pattern illustrated in figure 3.2. We chose to emphasize this pattern relationship, not because we believe it will hold true for a general detection scheme, but instead to highlight that a very clear pattern association exists between one of the clustering features and our known maneuvers. With this in mind, and the consideration that the detection algorithm results showed low precision, we considered a plot of the cluster labels against epoch time to determine if the data was clustering the way we were expecting it to. In figure 3.3, we can see that the fifth cluster includes TLE data spanning approximately 150 days. After noting this unexpected clustering, we chose to investigate the impact of adding epoch day to the feature set.

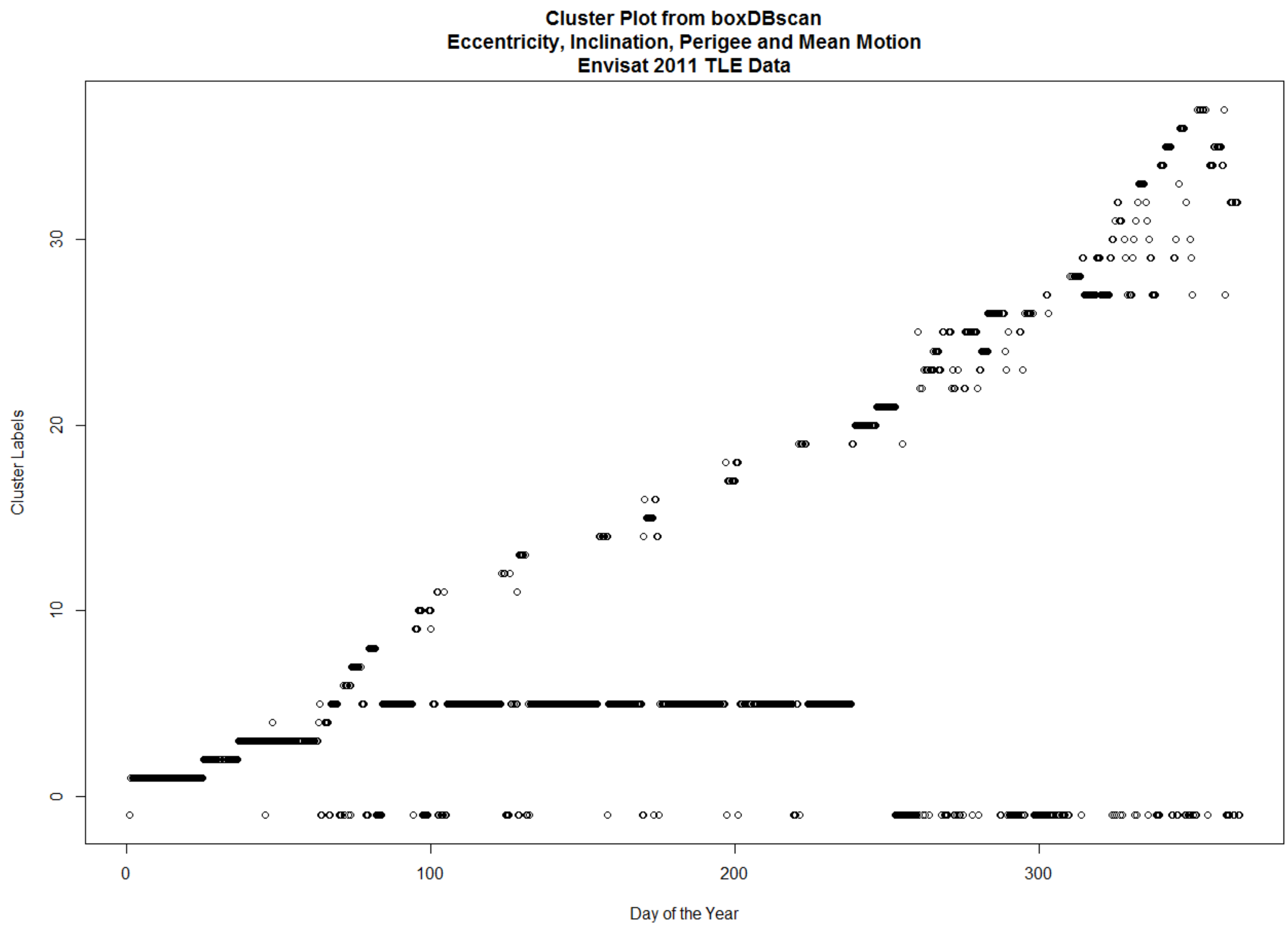

Figure 3.3. For the feature space [Eccentricity, Inclination, Perigee, Mean Motion] with the parameters defined by 3.1a and 3.1b, results from box DBSCAN labeled 37 clusters and 193 outliers. This graph shows the relationship between the cluster label for each TLE data point plotted against epoch day.

Features: Perigee, Inclination, Eccentricity, Mean Motion and Epoch Day

Building upon our previous approach to detecting satellite maneuvers, we added epoch day to the clustering feature vector and evaluated the impact on the resultant clustering. For the feature space [Eccentricity, Inclination, Perigee, Mean Motion and Epoch Day], we set the following parameters to detect satellite maneuvers using box DBSCAN:

$$\varepsilon_{[E,I,P,MM,ED]} = [1e-05, 5e-05, 1e-01, 1e-08, 4.5] \quad labeleqn: param2a \quad (3.2a)$$

$$MinPts = 5 \quad labeleqn: param2b \quad (3.2b)$$

Confusion Matrix

	Positive	Negative	
True	25	808	833
False	372	0	372
	397	808	1205

Table 3.3. The above confusion matrix shows the results of the detection algorithm when we use box DBSCAN with the feature space [Eccentricity, Inclination, Perigee, Mean Motion and Epoch Day] with the epsilon hyperrectangle = [1e-05, 5e-05, 1e-01, 1e-08, 4.5]

From the results of adding Epoch Day, in table 3.4 we can see that sensitivity (recall) jumped to 100%, but that every other metric decreased by a reasonable amount. It is very common that feature and parameter changes for detection algorithms will have a trade-off between sensitivity (recall) with both precision and specificity. To better understand the nature of this trade-off, we can think of the following scenarios. Achieving the result where 100% of the maneuvers are detected can be easily obtained if we define an algorithm that classifies every data point as a maneuver. In this case, with 1205 data points and 25 known maneuvers, sensitivity (recall) will be 100% but precision is only 2% and specificity will be 0%. If we allow this to be the base from which we can build from, we can tailor the algorithm to focus on increasing precision or specificity. Once we start doing such, we introduce rules and/or parameters that aid in a classification process. If precision is the priority, we require high confidence that any detection received is a true maneuver. If specificity is the priority, then the alternative objective is to know when not to detect. Classification for either objective will inherently result in detection error.

Sensitivity (Recall)	100%
Precision	6%
Specificity	68%
Accuracy	69%
F_1 score	12%

Table 3.4. Evaluations of the detection algorithm using box DB-SCAN with the feature space [Eccentricity, Inclination, Perigee, Mean Motion and Epoch Day] with the epsilon hyperrectangle = [1e-05, 5e-05, 1e-01, 1e-08, 4.5]

Therefore, we want to emphasize that in adding epoch day as a feature, the algorithm detected all of the maneuvers, but we didn't get those additional detections for free. The additional correct maneuver detections came with additional incorrect detections as well. The ratio of true positive results to false positive results is now approximately 1:15. An increase of approximately 112% to the false positive detections, while true positive detections only increased by approximately 39%.

Data and Cluster Analysis

The goal in adding epoch day was to generate clustering results that seemed to align more closely with our intuition for detecting satellite maneuvers. The cluster results of doing such with an epoch day epsilon bound set at 4.5 days is shown in figure 3.4. The impact of adding epoch day as a feature generated 56 cluster labels and 397 outliers, as well as a graph of clusters that seems to make more intuitive sense.

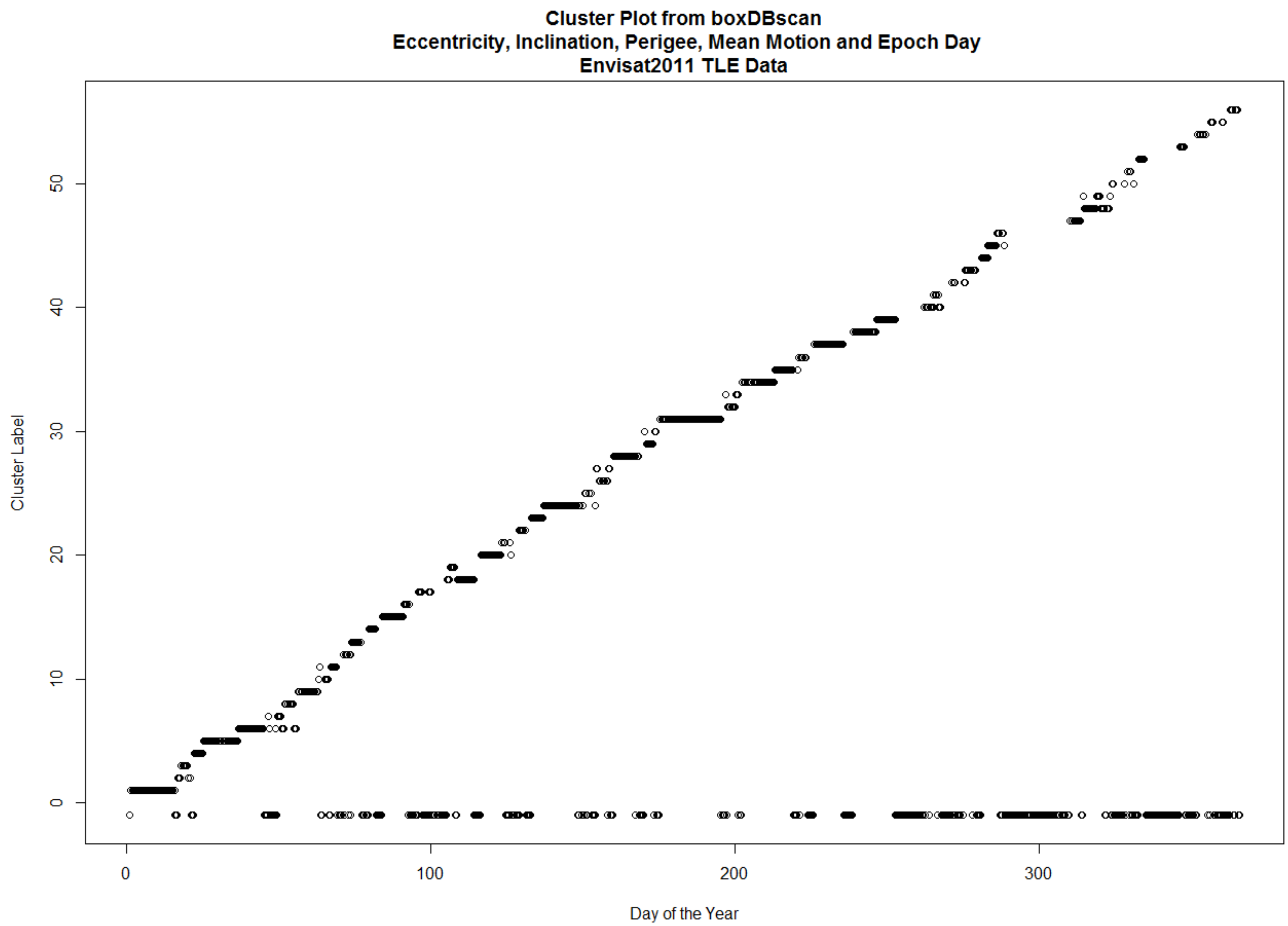

Figure 3.4. For the feature space [Eccentricity, Inclination, Perigee, Mean Motion, Epoch Day] with the parameters defined by **??** and **??** the result was 56 clusters and 397 outliers. This graph shows the relationship between the cluster label for each TLE data point plotted against epoch day.

Summary of Preliminary Results

As a summary to the preliminary results we have indicated that epoch day, although not an orbital element, may play a critical role in generating more effective clustering when used for detecting satellite maneuvers. Although the results generated a 100% recall for all 25 maneuvers, the impact to precision and specificity were less than desirable. Therefore, more considerations regarding satellite maneuver detection have to be made. A number of possible areas for development are outlined in chapter 4.

Chapter 4

Possible Paths Forward

Feature Selection

The first path forward that we will document and address is the consideration of variations in feature and parameter selections. In chapter 1, when we introduced astrodynamics for satellite trajectories, we briefly discussed the importance of the six orbital elements.

1. Semi-Major Axis, a
2. Eccentricity, e
3. Inclination, i
4. Argument of Periapsis/Perigee, ω
5. True Anomaly, ν
6. Right Ascension of the Ascending Node (RAAN), Ω

These six orbital elements are essential to uniquely describe the trajectory of an orbiting satellite. Using these six orbital elements, we are capable of deriving additional orbital descriptors. To begin with, clustering using all six orbital elements as the feature space, along with Epoch time, should be investigated. It should be noted that these six orbital elements can also be used to derive a state vector which defines the position and velocity of the satellite at each Epoch time. This alternative interpretation of the satellite trajectory is an especially important consideration since the detection of maneuvers should ideally focus on changes in velocity, which aligns with the considerations made by Kelcey *et al* in their paper [2].

In addition to the translation from orbital elements to the position and velocity descriptors, an alternative basis of elements used to describe an orbit is the Equinoctial elements. The Equinoctial elements are useful for handling special geometric difficulties that can be introduced by the classical orbital elements [5]. Therefore, we recommend the consideration for this change of basis be explored in more detail as research continues.

Determining the ideal feature and parameter selection needed to detect satellite maneuvers is not only the first consideration we will make, but also one that is likely to continue throughout future research. As we develop a better understanding for handling the additional concerns we laid out in chapter 2, we will continually need to return to and address fine tuning of the parameters as well as feature space selection.

Alternative Clustering Methods

The consideration for introducing alternative clustering methods for processing the TLE data to detect satellite maneuvers can be best explained using the illustration in figure 4.1. We first introduced this graph in chapter 2 as a means to illustrate the difference between the current use of Tracktable as a detection tool to the problem space of satellite maneuver detection. In this chapter we are reintroducing the graph to highlight the variance in patterns associated with mean motion. In the first 200 days there is a lot of consistency to the pattern associated with the rise and fall of mean motion. After day 200, we can see a much larger variance in the mean motion values

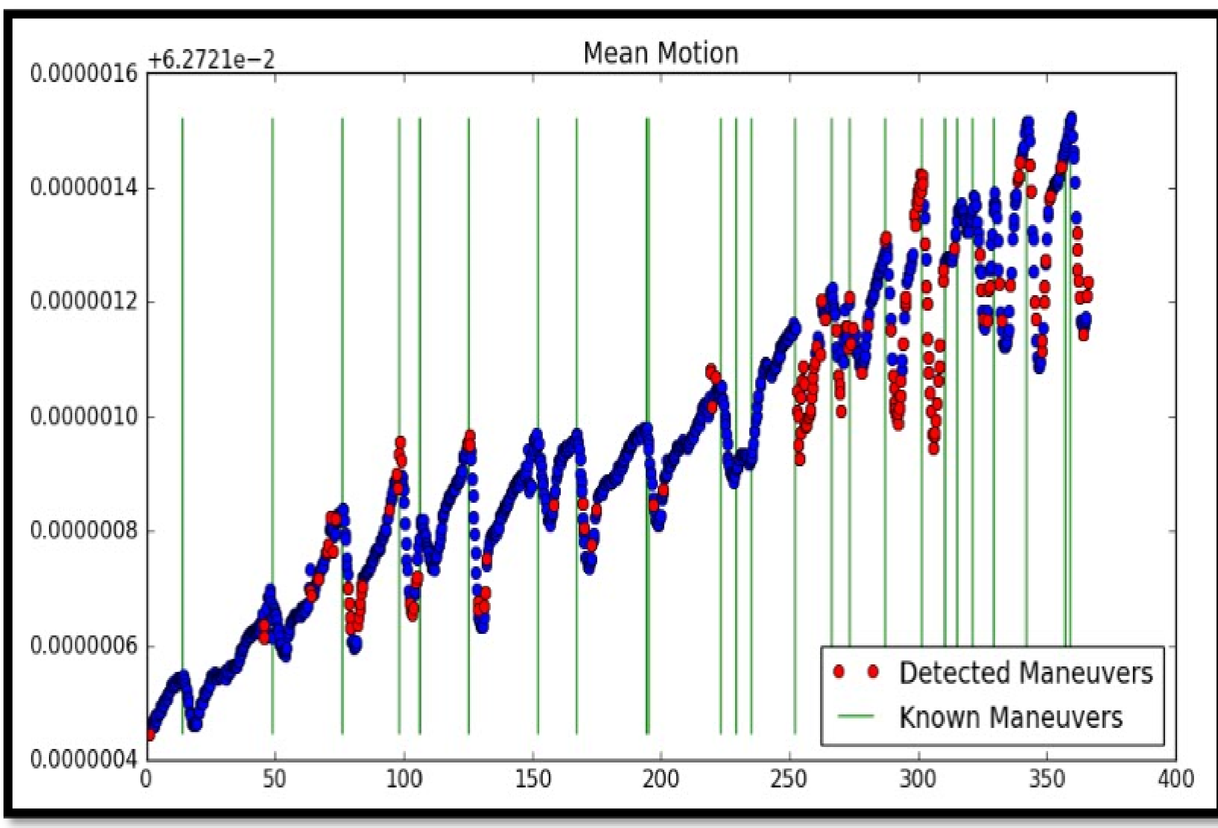

Figure 4.1. The outliers (labeled red) from the sequence of TLE data, plotted in terms of mean motion (radians per minute) against Epoch day.

although the TLE date has a fairly consistent sampling over the entire year. This generates the question of whether a fixed set of parameters is sufficient to accurately detect satellite maneuvers. The suggestions for future research to consider alternative clustering methods, such as hierarchical clustering or adaptive density based clustering, is being made to determine whether these methods would be more effective in detecting abrupt changes in the orbital elements with varying densities.

Hierarchical Clustering

The first alternative clustering approach that we want to consider is one that aligns with the considerations being made for Tracktable development. Hierarchical clustering methods are an approach that can be implemented when clusters of varying densities need to found in a data set. There are two extreme cluster levels in hierarchical clustering. The first assumes there is only one cluster and all the data belongs to it. The second extreme consideration is that every data point is its own cluster. Then there exists a breakdown of clusters in between these two extremes that have similar densities that can be illustrated using a dendrogram.

OPTICS (Ordering Points To Identify the Clustering Structure) Algorithm

OPTICS is an unsupervised clustering algorithm that is classifed as a hierarchical clustering algorithm method [1]. The development of the OPTICS algorithm was introduced shortly after DBSCAN with a similar concept in mind. Just as DBSCAN requires the notion of a neighborhood and a minimum number of points to justify the forming of a cluster OPTICS will also require these parameters to detect *core points*. In contrast to DBSCAN, OPTICS considers whether there exists a more densely connected cluster within the ε-neighborhood. Instead of classifying all of the data points as either core points, reachable points, or outliers the objective is to evaluate *core distance* and *reachability distance*, defined for each core point as the following

1. (Recall) *Core point:* a point is evaluated to be a core point if it has at least $MinPts$ in its ε-neighborhood.
2. *Core distance:* is defined for each core point as the distance to the $MinPts^{th}$ nearest neighbor.
3. *Reachablility distance:* is defined for each point within the ε-neighborhood of a core point as either the distance between the two points or the core distance, whichever is greater.

This defines a prioritized structure based on proximity between points to allow for the hierarchical clustering method to detect varying density within a data set. One of the restrictions, we noted with DBSCAN, that we will persist with the hierarchical clustering approach is the restriction to the shapes of the clusters. This leads us to the consideration of the following clustering approach as another alternative.

Adaptive Density Based Clustering

Adaptive density based clustering algorithms expand upon the concept of clustering data with varying densities to also considering varying shapes of clusters as well as non-uniform densities. From a preliminary review of the literature adaptive density based clustering is one of the more recent areas of research for unsupervised clustering methods. The benefits of considering an adaptive density based clustering approach is the freedom from user bias in the selection of parameters. The article *Automatic Clustering and Boundary Detection Algorithm based on Adaptive Influence Funtion* presented by Nosovskiy *et al* introduces a clustering approach that is parameter independent and does not assume a uniform density of the data to be clustered [3]. Alternatively, their novel approach allows for variance in the distribution of the data and can therefore be applied to both stochastic as well as deterministic data, while staying robust to noise. This opens the potential for a much wider application space for solving unsupervised clustering problems while removing human bias in the interpretation and expectations of the data. The attraction to such a algorithm is the potential it has for detecting not only the breaks in natural trends that we are expecting, but also detection for those that we did not expect.

The Nosovskiy paper is just one reference of this alternative clustering approach that has been reviewed so far. The recommendation is to further investigate additional adaptive density based clustering schemes.

Layered Analysis Approach

Another consideration for increasing precision and specificity is to perform a layered analysis. The initial analysis, much like the one presented in section 3, would ideally have a sensitivity (recall) of 100% for which we can then take the results of the detection algorithm and assume the subset of data as a set of possible maneuvers. These possible maneuvers are then run through a second layer of analysis for which a refined consideration for detecting maneuvers can be made. The idea here is to use a fast analysis on the first pass to filter the data to a smaller set of information that can be further analyzed in the second layer to achieve higher precision. Ideally, the result will be a set of detected maneuvers that we can be more confident are associated with real maneuvers.

As a quick experiment we considered exploiting the pattern associated with mean motion. We are not offering these results as a solution to the overall problem of detecting satellite maneuvers. Instead, we wanted to include these results to show the impact of running a layered analysis approach. The following script is the R code used to run a sliding window analysis to capture the rise and fall of mean motion, using a 5 day average.

R script:

```
possibleManeuvers<-data.frame(envisatData$Epoch.Day,envisatData$clusterLabel)
possibleManeuvers<-subset(possibleManeuvers,possibleManeuvers$clusterLabel==-1)

day1=0
day2=4
track<-data.frame(day = seq(1:365) , averageMeanMotion = rep(NaN,365) ,
negativeSlopeIndicator = rep(0,365) , maneuvers = rep(0,365) )

for(ii in 1:364){
windowData<-subset(possibleManeuvers,possibleManeuvers$Epoch.Day>=(day1+ii)
& possibleManeuvers$Epoch.Day<(day2+ii))
   if(nrow(windowData)>1){
   track[(ii+1),2]<-mean(diff(windowData$Mean.Motion))
     if(track[(ii+1),2]<0){
       track[(ii+1),3]<-1
     }
   }
}

for(ii in 2:365){
   if(track[(ii-1),3]==1 & track[ii,3]==1 & track[(ii+1),3]==0 ){
    track[ii,4]<-track[ii,1]
   }
}

detectedManeuvers<-subset(track,track[,4]!=0)
```

For the feature space [Eccentricity, Inclination, Perigee, Mean Motion, Epoch Day], we set the following parameters to detect the set of possible satellite maneuvers using Box DBSCAN.

$$\varepsilon_{[E,I,P,MM,ED]} = [1e-05, 5e-05, 1e-01, 1e-08, 4.5] \tag{4.1a}$$

$$MinPts = 5 \tag{4.1b}$$

The results from running Box DBSCAN for the preliminary analysis generated 397 outliers that we will now refer to as the set of possible satellite maneuvers. This set of 397 possible maneuvers is re-evaluated using the algorithm outlined above. The results of running this rudimentary approach

Confusion Matrix

	Positive	Negative	
True	19	1178	1197
False	2	6	8
	21	1184	1205

Table 4.1. The above confusion matrix shows the results of the detection algorithm when we use Box DBSCAN with the feature space [Eccentricity, Inclination, Perigee, Mean Motion, Epoch Day] with the epsilon hyperrectangle = [1e-05, 5e-05, 1e-01, 1e-08, 4.5] to determine the set of possible maneuvers. Then a sliding window used to detect the average rate of change for mean motion is used to determine where mean motion decreases consistently for 4 days.

Sensitivity (Recall)	76%
Precision	90%
Specificity	99.8%
Accuracy	99%
F_1 score	82%

Table 4.2. Evaluations of the detection algorithm using Box DBSCAN with the feature space [Eccentricity, Inclination, Perigee, Mean Motion, Epoch Day] with the epsilon hyperrectangle = [1e-05, 5e-05, 1e-01, 1e-08, 4.5] to determine the set of possible maneuvers. Then a sliding window used to detect the average rate of change for mean motion is used to determine where mean motion decreases consistently for 4 days.

to a layered analysis is presented in tables 4.1 and 4.2. From an overall comparison to the results we started with in chapter 3, the addition of epoch day as a feature and the second layer of analysis for filtering the results from box DBSCAN, we were able to increase recall by 5.5% and precision

855%. Although recall is still not very high, these are results that we could more readily except if we can devise a generalized scheme for a layered analysis approach. Therefore, as another recommendation, considerations for a layered analysis can be made as a viable path forward to developing a detection algorithm for satellite maneuvers.

This page intentionally left blank.

Chapter 5

Conclusion

Tracktable offers a novel and highly efficient means to detect anomalous flight and maritime behavior. By computing geometric descriptions of aircraft and ship trajectories, and utilizing a spatial index, patterns can be identified quickly and efficiently. We studied the related but different problem of detecting maneuvers within a satellite trajectory and investigated the suitability and reported on the performance of Tracktable towards this end. In our study, we researched the nature of satellite trajectory (Two-Line Element Set) data and astrodynamics noting that raw data represented a geometric description of a satellite orbit. This stands in contrast to raw flight and maritime data that is traditionally reported as position and sometime velocity state information. We performed several clustering analyses and through careful selection of trajectory features and use of a multi-layered analysis we improved upon previous approaches. Through a comparison across problem domains, potential limitations of box DBSCAN were identified. Paths for improvement may include utilizing hierarchical or adaptive density based clustering techniques or developing effective techniques for conducting automated sliding window, or sub-trajectory, analyses.

This page intentionally left blank.

References

[1] Mihael Ankerst, Markus M. Breunig, Hans-Peter Kriegel, and Jörg Sander. Optics: Ordering points to identify the clustering structure. *SIGMOD Rec.*, 28(2):49–60, June 1999.

[2] T Kelecy, D Hall, K Hamada, and D Stocker. Satellite maneuver detection using two-line elements data. In *Advanced Maui Optical and Space Surveillance Technologies Conference*, volume 1, page 19, 2007.

[3] Gleb V. Nosovskiy, Dongquan Liu, and Olga Sourina. Automatic clustering and boundary detection algorithm based on adaptive influence function. *Pattern Recognition*, 41(9):2757 – 2776, 2008.

[4] Mark Daniel Rintoul, Andrew T Wilson, Christopher G Valicka, W Philip Kegelmeyer, Timothy Shead, Kristina Rodriguez Czuchlewski, and Benjamin D Newton. Panther: Trajectory analysis. Technical report, Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States); Sandia National Laboratories, Livermore, CA, 2015.

[5] David A Vallado. *Fundamentals of astrodynamics and applications*, volume 12. Springer Science & Business Media, 2001.

DISTRIBUTION:

1	MS 0519	Kristina Czuchlewski, 5346
1	MS 0519	Erin Acquesta, 5346
1	MS 0980	David Cox, 5554
1	MS 0980	Drew Woodbury, 5554
1	MS 0980	Prabal Nandy, 5554
1	MS 1027	Carollan Ehn, 5635
1	MS 1243	Chris Valicka, 5523
1	MS 1243	Mark Hinga, 5523
1	MS 1326	Mark ”Danny” Rintoul, 1462
1	MS 1326	Andy Wilson, 1461
1	MS 0899	Technical Library, 9536 (electronic copy)

Sandia National Laboratories