DEVELOPMENT OF AN INITIAL TAXONOMY AND CLASSIFICATION SCHEME FOR ARTIFICIAL SPACE OBJECTS

C. Früh

Mechanical Engineering Department, University of New Mexico / Air Force Research Laboratory, Space Vehicles Directorate, NM 87117, USA

ABSTRACT

As space gets more and more populated, a classification scheme based upon scientific taxonomy is needed to properly identify, group, and discriminate space objects. Using artificial space object taxonomy also allows for scientific understanding of the nature of the space object population and the processes, natural or not, that drive changes of an artificial space object class from one to another.

In a first step, an ancestral-dynamic hierarchical tree based on a priori knowledge is established, motivated by taxonomy schemes used in biology. In a second step, available orbital element data has been clustered. Therefore, a normalization of a reduced orbital element space has been established to provide a weighting of the input values. The clustering in the five dimensional normalized parameter space is divided in two sub-steps. In a first sub-step, a pre-clustering in a modified cluster-feature tree has been applied, to initially group the objects and reduce the sheer number of single entities, which need to be clustered. In a second sub-step, a Euclidean minimal tree algorithm has been applied, to determine arbitrarily shaped clusters. The clusters also allow determination of a passive hazard value for the single clusters, making use of their closest neighbors in the minimal tree and the radar cross section of the cluster in question.

Key words: taxonomy; classification; space debris.

1. INTRODUCTION

A study of a set of different objects leads to a specific set of parameters describing the characteristics of those objects. With the increasing number of objects and accuracy to capture all possible characteristics, a large parameter space is utilized. In order to make the data set accessible and manageable, it is desired to reduce the parameter space to significant quantities which allow determination of differences and similarities between different objects and to group and classify them accordingly. However, the aim is not to introduce a random grouping, but to find a taxonomy of significant parameters corresponding to an actual physical and behavioral (e.g. dynamic) attributes.

Currently, about 20,000 objects are cataloged in the publically available USSTRATCOM catalog, whereas in situ measurements suggest around 300,000 objects to be in orbit around the Earth. So far, only a very broad taxonomy has been applied in different orbital regions, such as the orbital classifications of geostationary, geostationary transfer orbits, Molniya, low Earth orbits, which have been applied ad hoc. Another classification scheme is based subsets of objects, such as the ESA Classification of Geosynchronous objects, sorting objects by their orbital evolution, such as objects in drift orbits, around libration points, or controlled orbits. Another classification that has been readily adopted is the discrimination in classified and unclassified objects. If in the following the term classification is used, it prescribes the scientific terminology and shall not be confused with a security relevant grouping of objects. But discussions are ongoing about that a refinement of this structure is needed. In the following, only unclassified objects are taken into account. In this paper the focus is on orbital element classification. The important topic of further means of characterization and classification based on the inclusion of spectral and light curve measurements is not discussed here.

The oldest taxonomic systems are rooted in biology primarily established by Aristotle. Biological taxonomy orders plants and animals into an organized system that includes species, genera, families and higher forms of taxonomy. The system as applied to biology also shows that taxonomies are not static, but subject to change over time as new knowledge arises. Mayr defines the crucial steps [7] in building a taxonomy of any kind: The first step is (1) the *collection* of possible data, as a second step he defines (2) the *identification*. At the identification step the individual objects are sorted in groups. The challenge is to select the relevant groups, which are as broad as possible, while not overlooking distinguishing features; the identification is, in general, the analytical taxonomy step. The identification also includes the process of naming the groups that have

Proc. '6th European Conference on Space Debris'
Darmstadt, Germany, 22–25 April 2013 (ESA SP-723, August 2013)

been identified with a useful term that is precise enough to represent the group, but also short enough to be useful. As a third step, (3) the *classification* follows as a synthetic taxonomy categorization. In this step the different identified genera and species are ordered. Available a priori knowledge can be fed in, in addition to assessing the physical reality of the defined classes. The aim is to find an ancestral descent of the different genera and species, and their interrelations. Using traditional morphological taxonomy, convergence to a habitual state is sought, e.g.. As it is easily conceived, the three steps are highly interdependent. The data at hand determines and limits the identification that is actually possible, and identification is reiterated depending on the classification step. As new knowledge independent of the initial data set is added, classification can change, which traces back to the identification step. This complete interdependent system of data, identification and classification is named the taxonomy.

The taxonomic classification used in astronomy has the most overlap with the problem of artificial space objects are perhaps asteroids. Taxonomy systems of asteroids are traditionally based on color measurements (filter UVB and spectroscopic measurements) and albedo (including polarimetry). Earliest classifications of asteroids [12] were based on the filter similarities of the asteroid colors to K0 to K2V stars. The first more complete asteroid taxonomy was based on a synthesis of polarimetry, radiometry, and spectrophotometry, using a survey of 110 asteroids [1]. The defined a class C for dark carbonaceous objects, a class with the label S for silicaceous objects, and U for objects that did not fit either class. This system has the disadvantage that it was not detailed enough and is based on the exclusion principle. The groundwork for the most complete taxonomy, that significantly expands the previous taxonomy has been proposed by Tholen [11]. This taxonomy is, with small modifications, still in use today. Tholen established several asteroid classes, which are based on the albedo as well as eight channel color indices. The overall albedo, as well as the spread and inclination of the color values, distinguishes the classes. Tholen based his taxonomy on the cluster analysis of a normalized set of the color and albedo indices. The aim of asteroid taxonomy is to link those classes to heliocentric distance, diameter, and rotation rate, but also to the evolution, creation and dynamic (orbit-attitude) long term behavior of asteroids. Relatively independently a second asteroid classification has been induced recently. These second classification is a risk assessment of asteroids, with respect of their miss-distance to the Earth.

A second increasingly relevant task of an asteroid taxonomy is the quantification of the impact risk. The impact risk according to the so-called Palermo scale [2] is based on the expected energy of in impact, normalized with the background risk and the time frame of the collision. The energy of the impact is correlated to the mass and the impact velocity. The impact mass is traced back to the albedo and density (material properties) of the object and hence correlates to the Tholen albedo-color taxonomy. This is also a hint that the spectral classes correspond to a physical reality.

In the case of artificial objects, a historically new situation arises. For the first time the ancestral state is theoretically and a priori fully known as the objects originate from known man-made objects and materials. This is a state that is normally determined at the end of the development of a taxonomy. However, it does not eliminate the task of identification as derived from the data that is collected with the means at hand. Hence, in this case the taxonomy is done in reverse as compared to other applications in science or other disciplines. The question is, how does the physical reality influence the observed dynamical state the objects are in, nominal orbital state, as well as quantities that are of interest of us, such as the potential hazard those objects bear for the preservation of space environment. Furthermore, how the known ancestral state is inferred from survey data. Those are general questions and much broader than the scope of this paper. An initial task can determine the grouping of available a priori data. A second step is to establish groups of objects according to predefined similarity criteria; the resulting taxonomy is the overall system of these different object groups. The difficulty is therefore to find useful, physically relevant criteria. This requires a normalized parameter space, which provides a weighting of the different input parameters that allows the grouping of the classes and to sort them in a useful taxonomy. Thirdly, the system is used to establish a measure for the passive hazard a number for the groups of space objects according the risk they pose for crossing space objects.

2. A PRIORI ANCESTRAL-MATERIAL-DYNAMIC TAXONOMIC CLASSIFICATION

Objects in near Earth space are either natural objects or artificial. Making use of the fact that the creation and origin of all artificial space objects is in principle known, a new taxonomy has been developed. The term *in principle known*, shall not obstruct the fact that knowledge is missing for single objects, for which the path from their creation to their current state cannot be traced back. It is rather understood in a general sense. Fig.2 shows the outline of the taxonomy in graphical image. For the natural objects the reflection classes with their name and components are given; for the M-, E-, S- and C-class together the Palermo estimation values on albedo and diameter are given [2].

For artificial objects, the list of possible materials is limited. Material degeneration may occur, but no unknown materials per se are possible. The primary branches of the tree constitute two main kingdoms, which consist of (1) objects created in a controlled launch, and (2) objects created in an uncontrolled manner. The group of space debris objects is present in both kingdoms, the group of operational space assets only in the former. Objects that are created by some kind of controlled launch process typically have a predictable shape in the sense that it has been manufactured based on common design practices and so is a priori known. The objects resulting from a controlled launch can be subdivided in two different classes: those that are operational and those that are non-operational, where the non-operational objects are one part of the space debris kingdom. The non-operational objects either never made it into an operational orbit, have been decommissioned and are now in a graveyard orbit, or the life ended and they remained where they were operationally. In this latter category, in the absence of station keeping maneuvers, the orbit evolves based on natural perturbations.

The other group of non-operational objects are mission related objects, which never have been operational or in a controlled orbit. Mission related objects are likely to consist of very few different exterior materials, whereas decommissioned satellites are likely to consist of a complex variety of different materials. Those objects are having low area-to-mass ratios (LAMR), which are defined with an area-to-mass value (AMR) of below 0.1 square meter per kilogram, might either have been in a spin-stabilized attitude state (a priori spun) or initially despun, in the case where it was not spin-stabilized. Formerly operational objects, which were not in a controlled orbital state are in general smaller satellites meaning that they fall in the class of medium sized objects, consisting of an edge length of smaller or equal than 1.5 meter. Mission related objects, are either also LAMR objects, which can be originally spun, such as PAM upper stages, or originally despun, such as other rocket body types. Other mission related objects, fall in the class of medium area-to-mass ratio objects (MAMR), such as covers, which have AMR values between one and above 0.1 square meter per kilogram. Those can also be either originally spun, depending on how those covers were disconnected from the parent object, or despun. Mission related objects can either be larger in diameter than 1.5 meter or medium, below 1.5 meter, or small, which are below ten centimeter. That means, cubesats consisting of a single cube are just large enough to count to the smallest medium size objects.

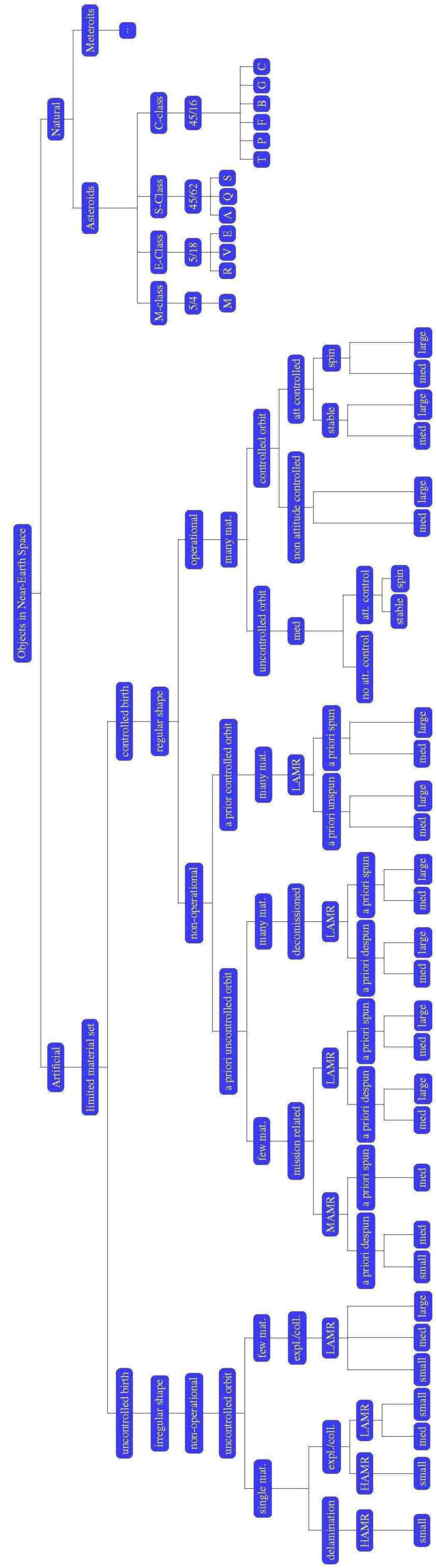
The second category of the space debris kingdom are those objects, created in uncontrolled birth, that are non-operational and have irregular shapes. They can either be created by delamination, whereas they consist of a single material, have high area-to-mass ratios and are small in dimension. Another creation process is by explosion or collision, leading to either small or medium sized objects. They typically consist of single materials which can have high or low area-to-mass ratios. These objects are not in any specific attitude state. This implies that the kingdom of space debris objects consist of three main groups, two of which result from controlled birth: those that were formerly operational, mission related objects, which never have been operational and were used in the creation of the operational objects (e.g. upper stages and boosters). The third group, the fragment group, is created in an uncontrolled birth process. The birth process has significant effects on the possible dynamical state the objects can be in, the size as well as the number of different materials a member object can be expected to consist of.

The kingdom of the operational classes of space objects can only stem from the controlled birth. These objects consist of many different materials, and are in either an active or passive attitude control state in a controlled or uncontrolled orbit. If there is attitude control, it can either be actively spin stabilized, hence, spun up, or in a stable attitude, achieved by three axis or other passive attitude stabilization mode (e.g. gravity gradient). The objects that are in controlled orbits are typically large or medium in size. Objects in uncontrolled orbits are medium in size, and can either be attitude controlled (spin or three-axis stable) or without attitude control. The ancestral-dynamic tree leads to the assignment of different size classes, which are linked to the different ways of their creation.

3. A CLUSTER BASED CLASSIFICATION ANALYSIS

A cluster analysis groups data by means of a similarity criterion in a selected feature space. The cluster analysis allows the determination of interrelationships within sample data. In general, one distinguishes between hierarchical clustering, heuristic segmentation methods and partitioning methods involving objective functions. Whereas the latter two methods produce clustered data on the same level, hierarchical clustering methods arrange data in nested sequences of groups, and can be displayed in a dendrogram, or tree structure.

Cluster analyses have a long tradition in taxonomy of natural objects. In biological taxonomy of bacteria, minimum distance sorting in a equal weighted feature space has been determined to lead to a classification [10]. A similar approach has been applied to a small set of asteroid spectra [3]. Tholen has used minimal spanning trees to establish an asteroid taxonomy [11]. The results of Tholen were confirmed using the same data and artificial neural network clustering [6]. Zahn compared different nearest neighbor algorithms to the minimal tree structure for clustering, proving the efficiency for cluster detection and pattern recognition, especially for irregular patterns [13]. Extensive research has been done on minimum spanning Euclidean trees, where the distance in a Euclidean parameter space is used as the measure of similarity [5, 8]. A greedy algorithm for finding a minimal tree is also being used [9], but comes with a large computational burden. In two dimensions, Delaunay triangulation provides a mean to supplement Prim's algorithm [4], though this is not feasible for higher dimensions. In this paper the focus is on hierarchical methods for clustering. Furthermore it is assumed that the number of clusters is not a priori known, and that a priori means have been defined.



The sheer amount of space resident objects, even when only cataloged objects are taken into account is very large, which makes the direct application of a minimal tree for clustering not very attractive. Hence, a two step process is proposed. In a first step, a modification of the Iterative Reducing and Clustering Using Hierarchies (BIRCH) algorithm is applied [14]. BIRCH is a very effective algorithm, leading to a hierarchical clustering in a single run. However, only the first step is used, ordering the data in a so-called CF (Cluster Feature) tree. A primary CF tree is created in a single run. The classical BIRCH algorithm then continues by slimming down the tree, with different resorting and merging techniques used to overcome some of the shortcomings of the initial CF tree. In the work presented here, a minimal tree is implemented as a second step, deviating from other CF combined approaches. Cluster features are a triple, consisting of the number of data points, which have been merged in the cluster, the linear and quadratic sum of the data points. This allows for a convenient adding of new data points into existing clusters. In the current approach the last entry is substituted by the sum of the radar cross sections (RCS). The feature consists of the three following entries:

$$CF = (N, \sum \vec{x}_i, \sum \mathrm{RCS}_i), \tag{1}$$

N is the number of objects in the cluster, $vecx_i$ is the five dimensional vector of to the object in normalized and cropped orbital element space, and RCS is the radar cross section of the single objects that have joined the cluster. The distance, which is used in determining the threshold, between two clusters or between an existing cluster and a new data point is determined as the following

$$d_{CF_1 CF_2} = \sqrt{(\frac{\sum \vec{x}_{1i}}{N_1} - \frac{\sum \vec{x}_{2i}}{N_2})^2}, \tag{2}$$

which makes direct use of the CF structure, which also makes it easy to join two clusters together, by simple vector addition. Two different algorithms have been implemented to group the elements. The first step is identical, a threshold T is determined on the leaf level for the maximum radius of a leaf-cluster.This means, these clusters have a circular structure. For both algorithms, a threshold value for the number of leafs that can be combined in one node, cannot be larger than a maximum number B , and similarly a maximum number of nodes in a root of L. This allows for deviations from the circular shape. Nodes and roots are split, when the maximum numbers are reached, sorting out the leaf or node, respectively, that is furthest from the mean, determined by the values stored in the node or root respectively. The difference between the algorithms is the following. In the classical BIRCH approach, the first node and root are filled until their threshold values are reached and then split according to the rule. This is computationally very efficient, one of the great advantages of the method. In the next step the next leaf is added to the node to which it is closest, and so on. In the real of a limited number of entries, this can lead to wrong results; the method relies on a *sufficient* amount of data points, that lead to enough splits in the nodes or roots, respectively, so no nodes and roots are kept, combining very diverse data points. *Sufficient* is determined by the amount of data in combination with the threshold values, but also depends on the diversity of the data, e.g. the number of outliers. That is the cost at which the computational savings are achieved. If the distribution would be perfectly uniform, the number of number$_{leafs}$/B and number$_{nodes}$/L are built. The number of data is nearly 20 000 objects, threshold values are chosen as low as 10 for both roots and nodes. Nevertheless, the CF tree is cross checked with a different algorithm. In that algorithm, the leafs and on the next level the nodes are combined to the absolute closest node and root, respectively. Nodes and roots are split at the threshold values at the same criterion as in the BIRCH algorithm. The modified algorithm allows a larger number of nodes and roots that are built automatically when the data is very diverse, no limits on the size of the data that is analyzed is enforced. For a uniform distribution of leafs, a number of number$_{leafs}$/ number$_{closest_neighbours}$ is built, that is in a perfectly uniform distribution number$_{closes_neighbours}$ is equal to two. If this number is larger than the threshold, this number is replaced by the threshold value and hence is equal to the BIRCH approach. To explain this further. BIRCH relies on the fact, that the nodes and roots are overpopulated in order to find meaningful nodes and roots, that represent the lower tree structures well, the alternative approach, does equal to BIRCH in the overpopulated case, but provides a more meaningful structure, in sparse regions, where overpopulation is not reached to enforce splitting in the BIRCH algorithm. It is hence more flexible to data which is not circularly shaped.

In a second step all leaf-nodes are taken as initial input data points in the minimal tree, with their centroid position as their new nominal position of a mean object. For those nodes, a minimal tree is fitted [9], which leads to the final clustering in cutting the longest links. Cutting the longest links sacrifices some of the hierarchical structure and leads to a flat clustering into same-level clouds.

For the analysis, two line elements (TLE) have been used as the source of orbital information. Radar cross sections from the satcat catalog have been used to supplement the data. In the case where no radar cross sections were available, a default value of 0.3 m^2 has been used. It is the very aim of the work to expand the classification to include light curve and spectral measurements, to supplement the orbital element classification. However, the present lack of light curves for the majority of objects made this intractable. A crucial point in even starting the clustering analysis is to express the quantities in a comparable manner, which expresses the values common units and hence makes them comparable in the first place. The second step is to weight the different orbital elements in a physical meaningful way. To find a common scale, the semi-major axis is used as a scaling factor and eccentricity, inclination, right ascension of the ascending node (RAAN), and argument of perigee are expressed in units of lengths as well, for the specific orbit. Orbital anomalies are not included for the clustering analysis. For the scaling of RAAN and argument of perigee using the approximated value from the circumference of the ellipse u, independently of the specific anomaly, the system is assumed to be locally flat. Because the accepted distances are limited by the threshold T the locally flat approach is justified. The inclination i, RAAN Ω and argument of perigee ω are scaled as the following:

$$i', \omega', \Omega' = \frac{i, \omega, \Omega}{2\pi} \cdot u, \tag{3}$$

The circumference u is approximated by the following:

$$u = \pi(a+b)\big(1 + \frac{3\lambda^2}{10 + \sqrt{4 - 3\lambda^2}}\big) \tag{4}$$

$$\lambda = \frac{a-b}{a+b}, \qquad b = \sqrt{a^2(1-e^2)}, \tag{5}$$

where a is the semi-major axis. The eccentricity e is scaled to the linear eccentricity ϵ:

$$\epsilon = e \cdot a. \tag{6}$$

The absolute scales are necessary to determine the distance in the five dimensional space, so only the relative values enter the clustering process. One way to weight the orbital elements would be an equal weight to all elements.

The topic of normalization cannot be overemphasized, since it directly determines and shapes the clustering, as it has the effect of *defining* what is regarded as dense region and close neighbors.

The scaling and weighting is done a priori, to save computational time.
Similarly to the Palermo scale, a hazard scale can be developed for space debris objects. In general the severity of a collision is linked to the energy and impulse conservation of the participants involved in a possible collision process. That is the mass of the objects involved, material properties, and their relative velocities. In contrary to the Palermo scale, which is an active scale of the approaching object, in the current paper we propose as a passive scale which depicts the hazard for any approaching object to cross a specific region.

A passive hazard value is attached to each of the clusters. It is derived from the radar cross section of all objects in the cluster and its averaged velocity as determined from the semi-major axis. For the single objects within the same cluster, a weighting can be applied with the actual weight of the object, if known, or from the inverse of the area-to-mass ratio value multiplied by the RCS. A measured albedo times a reflection value can also be used. This gives the highest passive hazard value to the largest object in a dense region, and a scaling with the semi-major axis leads to a higher assignment of passive hazard values to objects in lower orbits where absolute velocities of the cluster are largest. The scaled values of the relative distance and radar cross section of those clusters within the same cloud are added. This takes into account that the object is not only in a dense cluster, but also how densely the space outside the cluster are populated and how close other clusters are in normalized orbital element space.

$$\mathrm{h} = r \cdot n^2 \cdot \sum_{i=1}^{2} \frac{r_i}{d_i}, \qquad (7)$$

where r is the radar cross section of the whole cluster in square meters, n is the mean motion, r_i are the radar cross sections of the neighboring clusters, and d_i the distances to the center of the current cluster, d is the diameter of the original cluster. In order to evaluate the passive hazard scale, the singe leafs are used, in combination with two closest objects, that belong to the same cluster in the minimal tree.

4. RESULTS: DATA ANALYSIS

For the data analysis, a TLE data set has been selected and supplemented by the satcat catalog for the radar cross sections. In the case where no radar cross section was listed in the satcat catalogue, a default value of 0.3 m has been assumed. The CF tree has been built with a threshold value T of 1500 km, in the normalized cropped orbital element space. This distance should not be interpreted as a threshold for the state. For the branch factor and leaf factor B and L the value ten has been used.

In the initial run for building the CF tree in the classical approach, a total number of 575 roots were found with a total number of 1206 nodes holding 5229 leafs, 1964 leafs hold more than one element. With the modified algorithm 599 roots were found, with a number of 1782 nodes. On the root level, the algorithms produce nearly the same roots, on the node level, a significantly larger amount of nodes are created with the modified approach. The orbital elements of the six cluster leafs with the largest amount of data points are listed in Tab.1. Those are all high inclination low Earth orbits with small eccentricities. The RCS sum of the clusters is of the order of eight to six meters. Similar orbital regions are covered by the roots containing the non-leaf nodes and leaf nodes. The orbital elements and radar cross section of the roots with the highest number of member objects created in the modified approach is displayed in Tab.2.

Fig.1 shows the inclination as a function of the semi-major axis and the eccentricity as a function of the right ascension of the ascending node for the leaf nodes determined in the pre-clustering step. In Fig.1(a) the densely populated area is clearly visible is between 7000 and 9000 km, which contains the leafs with the largest number of objects. The geosynchronous region around 42000 km is also clearly discernible. Fig.1(b) shows the accumulation of the objects at low eccentricities and around 0.7 for all right ascension values. Fig.2 and 3 show the comparison between the classical birch clstering (noted by the subscripts *B*), and the modified algorithm. The modified algorithm tends to provides a stronger focus on the densely populated areas and seems to captures better the overall structure, which is discernible in the data. The minimal tree algorithm was applied to connect the clusters at the leaf level. Only the leafs containing more than one object are taken into account. The fifty longest connections in the tree have been cut leading to a total number of four main remaining clusters. A main cluster is defined by consisting of more than one leaf, and of more than 10 objects in total. This method works well in the case when only the leafs that include more than one object. In this case cutting the connections leads to the distinct clustering of four different regions: The sun-synchronous region in low Earth Orbit, the geosynchronous region, as well as two clusters in the medium Earth orbit, one with a low inclination and one with a higher inclination of around 60 degrees, the color coded clusters can be found in Fig.4. Depicted in black are the leafs which were cut loose, but are then neglected. They do not match the criteria for a cluster mentioned above. The method did not prove successful when including all leafs, also in particular, those which only hold one object. The selection of the

Table 1. Six most populated leafs in the circular cluster feature tree and their mid points: number of objects, semi-major axis (km), eccentricity, inclination (deg), argument of perigee (deg), RAAN (deg), radar cross section of the whole cluster (m)

# obj	**a**	**e**	**i**	ω	Ω	**RCS**
30	7177.6	0.009	98.72	85.73	40.57	9.00
29	7196.4	0.008	73.25	158.03	279.60	8.85
29	7152.8	0.004	74.04	145.01	286.79	8.70
27	7201.4	0.008	72.15	141.45	271.10	7.22
27	7144.8	0.004	74.04	143.10	299.69	8.10
26	7203.6	0.008	98.82	29.18	84.507	6.87

Table 2. Six most populated root nodes in the circular cluster feature tree and their mid points: number of objects, semi-major axis (km), eccentricity, inclination (deg), argument of perigee (deg), RAAN (deg), radar cross section of the whole cluster (m)

# obj	**a**	**e**	**i**	ω	Ω	**RCS**
197	7276.5	0.012	98.55	127.75	90.21	57.38
133	7546.6	0.056	77.07	81.70	97.40	83.75
131	7256.7	0.011	98.73	94.07	290.27	38.15
130	7361.2	0.010	97.16	73.26	341.12	41.81
129	7172.1	0.008	73.03	97.40	298.99	44.62

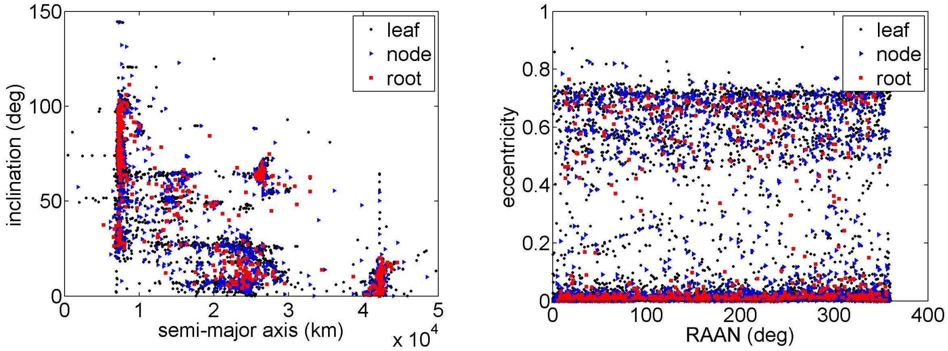


Figure 1. Cataloged space objects clustered in leafs, nodes and roots with the modified algorithm.

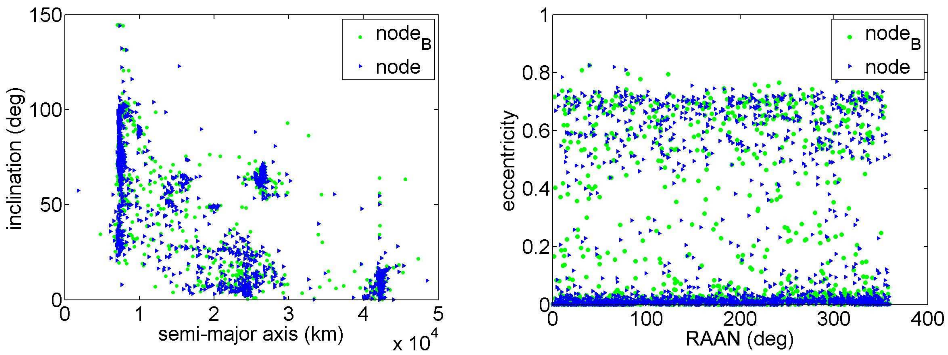


Figure 2. Leafs of clustered space objects with more than one member.

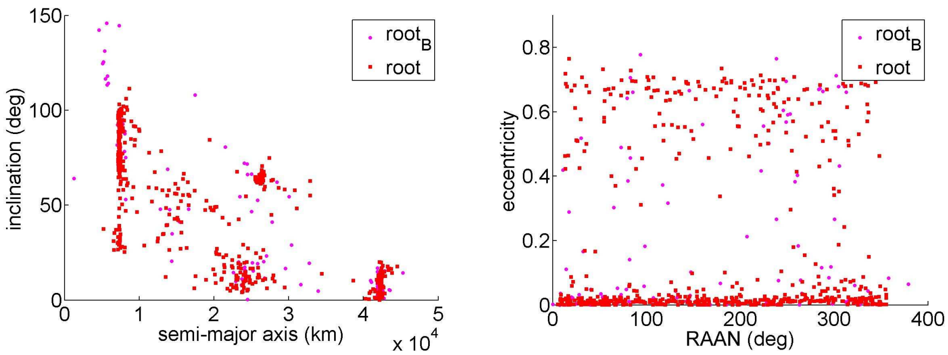


Figure 3. Leafs of clustered space objects with more than one member.

Table 3. Example cluster with large RCS and the cluster to which it is connected in the minimal tree: Number of objects in the cluster, semi-major axis (km), eccentricity, inclination (deg), argument of perigee (deg), RAAN (deg), radar cross section of the whole cluster (m), and distance in 5 parameter space between the two clusters (km)

# obj	**a**	**e**	**i**	ω	Ω	**RCS**	dist
3	7235	0.05	32.87	61.61	137.43	24.02	-
2	7444	0.06	29.48	46.86	121.47	0.40	2355.8

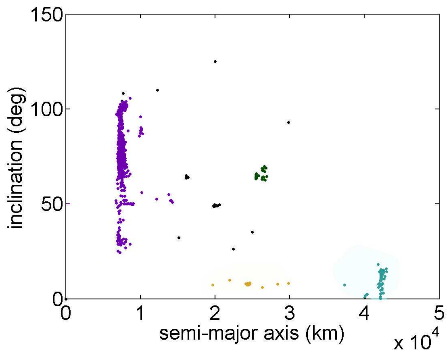


Figure 4. Color coded clusters of space objects: near sun-synchronous clusters, medium earth cluster with low inclination (iMEO), medium Earth cluster with high inclination (IMEO) and near-geosynchronous cluster.

correct link to cut is crucial and not as intuitive as in the previous case. Often single leafs are cut off rather than separating clustered regions. Further research is needed to develop an efficient mechanism for link selection.

Tab.3 shows the example of the leaf with the highest radar cross section value together with its orbital elements. The leaf is connected to one other leaf only (at maximum two would be possible). The connected leaf is also shown in the table. The passive hazard value for this region can be readily determined using the Eq.7, leading to a value of h=43.8. No complete list of hazard values for all clusters is shown here because of limited space for this paper. In general, objects in the class of near sun-synchronous LEO orbits have the highest values because they have the highest radar cross section sum and the connections to the neighboring clusters as the shortest.

5. CONCLUSIONS

Initial work towards a taxonomic categorization of artificial space objects have been conducted. An ancestral-dynamic taxonomy of object creation has been outlined in this initial step, making apparent the relationship between shape, material composition and dynamical state on the creation mechanism of the different space objects. Three different space debris classes have been established, consisting of former satellites, mission related objects with a controlled birth as well as debris objects stemming from an uncontrolled birth. In a second step, the orbital element sets available from USSTRATCOM were analyzed, and cluster analysis performed. Orbital elements have been normalized, and analyzed, neglecting anomalies such as orbital maneuvers. Cluster features have been determined and circular hierarchical clustering has been performed., and 517 root nodes found that hold a total of 8090 leafs in several branches. In a follow-up step, a minimal tree has been applied and clusters have been determined. Clustering based on the minimal tree approach was established using only the leafs which consisted of more than one object. This lead to four groups of objects that can be grouped as being near sun-synchronous low Earth orbits, in highly inclined medium Earth orbits, in moderately inclined medium Earth orbits, and in near geosynchronous orbits. A mean to determine a passive hazard values has been established based on the leaf level clusters and the minimal tree, giving emphasis to the size of the cluster itself based on its radar cross section, as well as the clusters to which the region is linked the most closely in the minimal tree.

Next steps are the refinement of the clustering based on the minimal tree that includes all objects, the establishment of a better and approachable size metric for the objects, and the inclusion of the classification and taxonomy to spectral and light curve measurements.

ACKNOWLEDGMENTS

The first author would like to thank the National Science Foundation for providing the funding that supported this work. Thanks goes to Moriba Jah, Ernest Valdez, Tom Kelecy, and Paul Kervin.

REFERENCES

1. C.R. Chapman, D. Morrison, and B. Zellner. Surface properties of asteroids: A synthesis of polarimetry, radiometry, and spectrophotometry. *Icarus*, 25(1):104 – 130, 1975.
2. S.R. Chesley, P.W. Chodas, A. Milani, G.B. Valsecchi, and D.K. Yeomans. Quantifying the risk posed by potential Earth impacts. 159:423–432, 2002.
3. J.K. Davies, N. Eaton, S.F. Green, R.S. McCheyne, and A.J. Meadows. The classification of asteroids. *Vistas in Astronomy*, 26, Part 3(0):243 – 251, 1982.
4. C. Eldershaw and M. Hegland. Cluster analysis using triangulation. *Computational Techniques and Applications: CTAC97*, page 201208, 1997.
5. O. Grygorash, Yan Zhou, and Z. Jorgensen. Minimum spanning tree based clustering algorithms. In *Tools with Artificial Intelligence, 2006. ICTAI '06. 18th IEEE International Conference on*, pages 73–81, 2006.
6. E. S. Howell, E. Mernyi, and L. A. Lebofsky. Classification of asteroid spectra using a neural network. *Journal of Geophysical Research: Planets (19912012)*, 99(E5):10847–10865, 1994.
7. E. Mayr. *Systematics and the Origin of Species*. Columbia University Press, 1942. ISBN: 9780674862500.
8. F. P. Preparata and M. I. Shamos. *Computational Geometry: An Introduction*. Springer-Verlag New York, Inc., 1985.

9. R. C. Prim. Shortest connection networks and some generalizations. *Bell System Technology Journal*, 36:1389–1401, 1957.
10. P. H. A. Sneath. The application of computers to taxonomy. *Journal of General Microbiology*, 17:201–226, 1957.
11. D. J. Tholen. Asteroid taxonomy from cluster analysis of photometry. The University of Arizona, 1984. PhD Thesis.
12. I. van Houten-Groeneveld and C. J. van Houten. Photometrics Studies of Asteroids. VII. *Astrophysical Journal, vol. 127*, 127:253, March 1958.
13. C.T. Zahn. Graph-theoretical methods for detecting and describing gestalt clusters. *Computers, IEEE Transactions on*, C-20(1):68–86, 1971.
14. Tian Zhang, Raghu Ramakrishnan, and Miron Livny. Birch: an efficient data clustering method for very large databases. *SIGMOD Rec.*, 25(2):103–114, 1996.