Article

Strategy of Multi-Beam Spot Allocation for GEO Data Relay Satellite Based on Modified K-Means Algorithm

Huiliang Liu 1,2, Yao Chu 3,* 4, Yulong Zhang 4, Weiguo Hou 1,2, Yinqiao Li 1,2, Yuan Yao 1,2 and Yaxing Cai 1,2

1 Institute of Telecommunication and Navigation Satellites, China Academy of Space Technology, Beijing 100094, China; liuhuiliang88@126.com (H.L.); houweiguo88@163.com (W.H.); lyq_bit@126.com (Y.L.); yaoqianli19900721@163.com (Y.Y.); 13810078786@163.com (Y.C.)
2 Innovation Center of Satellite Communication System, China National Space Administration, Beijing 100094, China
3 Department of Precision Instrument, Tsinghua University, Beijing 100084, China
4 School of Integrated Circuits, Tsinghua University, Beijing 100084, China; zhangy2020@mail.tsinghua.edu.cn
* Correspondence: chuyao@mail.tsinghua.edu.cn

Abstract: With the booming development of satellite applications, the giant constellations of low Earth orbit (LEO) satellites have introduced challenges for the data relay service. The multi-beam satellite not only offers concurrent access to a large number of objects, but can also meet the high data requirements toward specific coverage of the LEO constellation. However, the multi-beam satellite often faces the mismatch problem of spot allocation and data requirements, which can cause an overload traffic jam or a waste of resources. An optimization algorithm on spot beam allocation is necessary to automatically place the spot centers with appropriate beam widths in line with the density of the traffic demands and to realize the uniformity of the beam occupation. Compared with the conventional K-means algorithm, two adjustable parameters $\alpha$ and $\beta$ are introduced: one for tuning the ratio of two components making up the distance matrix, and the other for setting the obligatory minimum number of objects per beam. In this paper, the whole process of the proposed method is demonstrated, including the establishment of the low-orbit satellite constellation model, the extraction of the distribution features, and the implementation and evaluation of the modified K-means algorithm. The results prove the validity of the proposed algorithm. A larger value of $\beta$ with a relative smaller value of $\alpha$ tends to obtain the uniformity of beam occupation; the minimum standard deviation of objects per beam is achieved when $\alpha$ is 0.2 and $\beta$ is 0.8. This demonstrates that the uniformity of objects per beam can be realized by adjusting the parameters of the distance determination matrix and the obligatory minimal number of objects in each beam. The impact of parameter range on the results is also analyzed.

Keywords: data relay service; multi-beam satellite; beam spot allocation; modified K-means

1. Introduction

Satellite applications play a fundamental role in modern society, offering practical services, including television broadcasting, Earth observation, navigation and communication [1]. The last decade has witnessed the booming revolution of satellite communication from fixed communication and mobile communication to high-throughput communication [2]. Several large constellations of low Earth orbit (LEO) satellites with altitudes of 1500 km or less were proposed to provide global wideband with cutting-edge technologies, shorter propagation delay and rapid global deployment [3,4]. Meanwhile, the control and operation of LEO constellations offers challenges to the ground stations. Due to their orbital periods of approximately 100 min, they can only contact a single ground station for a few minutes per pass [5]. Since the capability to transmit information with the ground station in time is crucial for stable in-orbit operations, a network of ground stations spread over the world would be considered to reduce the gaps between the contact times, which
results in additional costs and complexity of the system [6]. To overcome the problem, space-based data relay satellite was invented to offer data relay service at geostationary Earth orbit (GEO). With an altitude of 35,786 km, a GEO satellite can cover up to 69% of LEO satellites. Uninterrupted in-time connection toward LEO satellites can be realized with two GEO data relay satellites to provide the data path to the ground station [7].

The existing data relay satellites were initially designed and optimized with target objects of the International Space Station, Earth observation satellites and other space shuttles or launch vehicles with relatively lower volume [8]. These kinds of classical data relay satellites include the Tracking Data Relay Satellite (TDRS) of America, the Luch satellite of Russia, and the European Data Relay Satellite (EDRS) [9,10]. The principal communication link can be established based on the steerable antenna tracing the user spacecraft. The single data relay satellite can only provide concurrent access to less than 10 LEO objects in general [11]. So, the use of the conventional data relay service is mainly the high-value LEO spacecraft.

In the last decade, several giant LEO satellite constellations were proposed, which spawned extensive demand of data relay between space and ground segments [12]. Novel types of data relay satellites should be considered to meet the pent-up demand for continuous, multi-user and cost-effective access to the growing volume of data from space [13,14]. Several commercial companies raised feasible proposals to build a data relay system through multi-beam satellites. The Inter-satellite Data Relay System (IDRS™) directly brings the capabilities of the Inmarsat Broadband Global Area Network (BGAN) system to space by integrating the IDRS™ terminal into a LEO spacecraft. While in orbit, the communication link to the Inmarsat BGAN Ground Network can be switched and maintained between beam spots of the Inmarsat-4 GEO satellites, which are multi-beam global mobile satellites [15]. On the one hand, leveraging upon the existing BGAN service portfolio, this proposal provides a range of always-on, on-demand, real-time connectivity, thus resolving the need to wait for the LEO satellite to pass above a ground station. On the other hand, the BGAN system was initially designed for ground and flight applications. There is no further optimization of the beam coverage and spot allocation focused on the LEO spacecraft application of the data relay service [16]. For another example, the data relay system, named SpaceLink, is customized toward spaceborne objects. The satellites also implement a fixed multi-beam antenna to realize the coverages of the LEO constellation [17]. The only difference is that the SpaceLink system selects medium Earth orbit (MEO), which helps reduce latency by about half, compared to a similar service of GEO.

In the case of the multi-beam satellite, the system should aim at designs to maximize the efficiency of resources in terms of frequency reuse, power allocation and especially the quality of the beam layout [18]. Great attention is drawn to the antenna design and optimization to further increase the total throughput of the satellite by using a limited number of beams [19]. Conventionally, the allocation of beam spots was carried out based on the supposed service requirements of ground and flight applications [20]. It is common for design tools to provide regular coverage based on the assumption of traffic demand partition [21]. Therefore, the solution is often suboptimal with inappropriate beam widths, overprovisioned or unprofitable beams. Several optimization algorithms were taken into consideration to build a non-uniform layout and further improve the systematic performance by the local search method or simple enumeration [22,23]. Previously, in the study of the communication system, there were studies focusing on dealing with the problem of uniformity, which usually takes several factors, such as energy availability, interest, physical ties and so on, into consideration to carry out the coalition and selection process. To implement the dynamic practice, the individual device is associated with a holistic utility function, which appropriately represents its degree of satisfaction with respect to the Quality of Service (QoS) prerequisites fulfillment [24]. This kind of process is also applied among a variety of wireless access technologies supporting IP connectivity as the handover function, which is a key enabler for seamless mobility and service continuity [25]. However, these frameworks are all used in the scenario with fixed hubs, and the aim is to lay out the
behavior of mobile terminals and chase the optimization of utility function. In this work, the task is to allocate the beam spot centers with a predictable distribution of dynamic terminals without the prerequisite of the centroid positions.

In this work, the service objects are LEO satellite constellations, which have predictable configuration and specific movement characteristics. Different from the ground and flight applications, the data requirement assumption of LEO objects is relatively realistic [26]. It offers the opportunity to globally optimize the beam layout, relying on truthful data. Figure 1 illustrates the schematic of a GEO data relay satellite covering the target LEO satellite constellation with spot beams. The LEO satellites with altitude lower than 1500 km are located within a field of view (FOV) of approximately $\pm 11^\circ$. The number in the center of each circle indicates the objective LEO satellites in the coverage of the certain beam. For any algorithm, the challenge is to appropriately place the spot centers and choose the beam widths in line with the density of the traffic demands. This paper aims at proposing a mathematical model to automatically realize the global optimization of beam layout based on a modified K-means approach.

![Figure 1. Schematic of a GEO data relay satellite covering the target LEO satellite constellation with spot beams. The LEO satellites are located within a field of view (FOV) of approximately $\pm 11^\circ$. The number in the center of each circle indicates the objective LEO satellites in the coverage of the certain beam (Rearth is the radius of the Earth).](image)

**2. Methods**

Standard K-means clustering is a method of vector quantization that aims to partition $n$ observations into $k$ clusters in which each observation belongs to the cluster with the nearest distance (cluster centers or cluster centroid) [27]. In the design of the satellite beam layout, the task is also to partition $n$ objects (LEO satellites) into $k$ clusters (spot beams). However, the evaluation criterion is not only to minimize the within-cluster variances of squared Euclidean distance, but also balance the number of objects in each beam [28,29]. It is crucial to mitigate the traffic jam induced by overprovisioned beams with overload objects, or the waste of resources caused by unprofitable beams with few requirements.

It is worth noting that K-means with a fixed number of cluster members was studied previously [30,31]. The data samples were clustered by standard K-means first, and the number of group elements was adjusted locally by employing a greedy strategy. In that method, the predefined fixed number of cluster members is necessary to guide the following recruitment or discard process on each cluster. The situation is partially different from the beam spot allocation process since the multi-beam satellite system has a certain amount of tolerance on the variance of object number per beam. The optimization process
should also take the size of the cluster into account, which determines the uniformity of the physical beam width. In other words, we should not only chase the uniformity of beam occupation, but also balance the uniformity of the beam width. That is why we cannot offer a certain predefined fixed number of cluster members before the clustering process. In this work, we propose a modified K-means algorithm to solve the aforementioned problem with introduced parameters to guarantee the obligatory minimum number per cluster and globally optimize the number of objects in each beam without locally implementing recruitment or the discard process. The tradeoff between the uniformity of the beam occupation and the uniformity of the beam width is also considered and discussed in the following sections.

Figure 2 illustrates the step-by-step process proposed in this paper. The whole process is introduced in detail in the following subsections.

Figure 2. Flow chart of proposed methods.

2.1. Low-Orbit Satellite Constellation Model

The first step is to establish the low-orbit satellite constellation model. In this work, a typical LEO satellite constellation is considered, including 324 satellites in 18 orbit planes at 1200 km with inclination of 88° and additional 576 satellites in 24 orbit planes at 960 km with inclination of 55°. The satellites in the same orbital plane are uniformly distributed. As shown in Figure 1, the satellite constellation is modeled in a coordinate system with relatively stationary Earth and the GEO data relay satellite. Then, the multi-beam coverage is modeled in the perspective of the GEO satellite. The coverage of a certain beam is modeled with the half-cone angle and the beam center in terms of the azimuth and elevation degrees. Finally, the coordinates of a certain satellite are projected into the multi-beam coverage area to check the service access toward a certain beam. It is worth noting that the Earth as an obstruction should be taken into account to check the line-of-sight propagation between the object and the GEO data relay satellite.

2.2. Feature of Density Distribution

The second step is to extract the feature of density distribution since the design of the multi-beam GEO data relay satellite requires the knowledge of the LEO statistics and density distribution. As shown in Figure 3, the characteristics of the object distribution in the FOV of the GEO data relay satellite are extracted with 2° uniform spot beams. Figure 3a illustrates LEO satellite occupation for a ±11° FOV subdivided into hexagonal cells of 2° diameter. The number in the center of each cell indicates the transient LEO satellites in the coverage of the certain beam. The blue dash line sketches the contour of Earth’s surface. Figure 3b shows the statistics of the cell occupation. It is obvious that the number of satellites per cell is larger in the peripheral FOV compared with the central ±8.5° FOV. The area surpassing the contour of earth surface offers more space to satellites in adjacent orbits with line-of-sight propagation toward the GEO data relay satellite. There are eight cells in which the number of satellites exceeds 16. Four of them lie around the intersection of the orbit panels with an inclination of 88°, and the other four cells are around the corners, where the orbit panels with an inclination of 55° overlap. For most of the area within the ±8.5° FOV, the number of satellites per cell is less than 5.
To further feature the details, the diameter of the cell decreases from 2° to 1°. Figure 4a illustrates the LEO satellite occupation for a ±11° FOV subdivided into hexagonal cells of 1° diameter. Figure 4b shows the statistics of the cell occupation. There exist hollow cells within the ±8.5° FOV. The satellites also tend to concentrate around the intersection of orbit panels with inclinations of 88° and 55°.

Figure 3. Characteristics of the object distribution in FOV of GEO data relay satellite with 2° spot beams: (a) LEO satellite occupation of spot beams. The number in the center of each cell indicates the transient LEO satellites in the coverage of a certain beam. The blue dash line sketches the contour of Earth’s surface; (b) statistics of the cell occupation.

Figure 4. Characteristics of the object distribution in FOV of GEO data relay satellite with 1° spot beams: (a) LEO satellite occupation of spot beams. The number in the center of each cell indicates the transient LEO satellites in the coverage of a certain beam. The blue dash line sketches the contour of Earth’s surface; (b) statistics of the cell occupation.

To build up the data set for analysis, a finer cell is necessary to depict the distribution more accurately and help advance the precision of the output beam width and beam centers. As shown in Figure 5, the characteristics of the object distribution in the FOV of the GEO data relay satellite are extracted with 0.5° uniform hexagonal cells along continuous three moments with a one-second interval to mitigate the impact of randomness.
Figure 5. Characteristics of the object distribution in FOV of GEO data relay satellite with 0.5° spot beams: (a) LEO satellite occupation of spot beams. The number in the center of each cell indicates the transient LEO satellites in the coverage of a certain beam. The blue dash line sketches the contour of Earth’s surface; (b) statistics of the cell occupation.

2.3. A modified K-Means Clustering Algorithm

As shown in Figure 2, the third step is to allocate the beam spots by the proposed clustering algorithm.

The following are the inputs of the clustering algorithm:

1. Number of beams \( k \).
2. Number of objects \( n \), which have line-of-sight propagation with the GEO data relay satellite as shown in Figure 2.

Tolerable variance range of object number in each beam. In an ideal case, the objects in each beam should be equal. In reality, a variance in the number of objects per beam is tolerable. In this study, minimum number \( N_{BL} \) of objects in a single beam should be guaranteed.

The process of the proposed clustering algorithm is as follows.

Step 1. Initializing the beam centers.

The center coordinate of each beam \( \mu_j \) can be distributed randomly within the FOV or initialized according to specific requirements. In this study, we initialized the beam centers with a random sequence.

Step 2. Calculating the Euclidean distance.

The Euclidean distance between the objects and the center of each beam are calculated as follows.

\[
M_{ij} = \|x_i - \mu_j\| \quad (i = 1, 2, \ldots n; \quad j = 1, 2, \ldots k) \tag{1}
\]

Step 3. Sorting of distances.

The distance matrix \( M \) is sorted twice by row and column, respectively, to obtain the ranked matrices \( A \) and \( B \) as follows.

\[
A_{ij} = \arg(j = \text{arg sort}(M_{i.})) \quad B_{ij} = \arg(i = \text{arg sort}(M_{.j})) \tag{2}
\]

Step 4. Normalizing and weighting.

Matrices \( A \) and \( B \) are normalized and weighted to obtain the determination matrix \( D \) as follows. When \( \alpha \) is equal to 1, the determination distance matrix \( D \) degenerates to the standard K-means algorithm determination matrix.

\[
D_{ij} = \alpha \cdot \frac{A_{ij}}{k} + (1 - \alpha) \cdot \frac{B_{ij}}{n} \tag{3}
\]

Step 5. Setting boundary of parameter.
A specific value of $\beta$ is chosen to determine the minimum number $N_{BL}$ of objects in a single beam as follows.

$$N_{BL} = \left\lfloor \beta \cdot \left\lfloor \frac{n}{k} \right\rfloor \right\rfloor \quad (0 \leq \beta \leq 1)$$ (4)

Step 6. Clustering of objects.

A set $E = \{1, 2, \ldots, n\}$ is initialized to keep a record of the objects to be clustered. For each beam, we choose the $N_{BL}$ smallest objects in $D_j$ that is also available in set $E$. We update its cluster index $c_i$ to be $j$, and subtract the allocated objects from set $E$.

The remaining unallocated objects are allocated according to the determination distance matrix as follows.

$$c_i = \arg \min_j \{D_{ij}\}$$ (5)

Step 7. Updating the beam center.

The coordinates of the beam centers are updated as follows.

$$\mu_j = \frac{\sum_{i=1}^{n} \{c_i = j\} x_i}{\sum_{i=1}^{n} \{c_i = j\}}$$ (6)

Step 8. Repeating Step 2 to Step 7.

The process from Step 2 to Step 7 is repeated until the allocation of objects to the beams is no longer changing, or it reaches the preset maximum number of iterations.

Compared with the conventional K-means, the main difference of the proposed algorithm is the modification of the determination matrix. The proposed determination matrix is based on the sorted matrix of the Euclidean distance by row and column, respectively. To indicate clearly the distinctive ranking process of Equation (2), we raise an example with 6 objects and 3 cluster centroids and assume the initial distance as the following.

$$M = \begin{bmatrix}
0.5 & 0.8 & 0.6 \\
1.1 & 0.5 & 1.6 \\
0.7 & 0.3 & 1.0 \\
0.6 & 0.7 & 0.8 \\
0.8 & 0.6 & 0.4 \\
1.2 & 0.4 & 1.5 
\end{bmatrix}$$ (7)

The matrix $M$ is sorted by row to obtain matrix $A$ as follows.

$$A = \begin{bmatrix}
1 & 3 & 2 \\
2 & 1 & 3 \\
2 & 1 & 3 \\
3 & 2 & 1 \\
2 & 1 & 3 
\end{bmatrix}$$ (8)

It is also sorted by column to obtain matrix $B$ as follows.

$$B = \begin{bmatrix}
1 & 6 & 2 \\
5 & 3 & 6 \\
3 & 1 & 4 \\
2 & 5 & 3 \\
4 & 4 & 1 \\
6 & 2 & 5 
\end{bmatrix}$$ (9)

If we directly use matrix $A$ as the determination distance matrix to conduct the clustering process following Equation (5), the result will be the same with the conventional K-means, which guarantees the minimum sum of the distances in a cluster by allocating each object into the nearest cluster. This means that when $\alpha = 1$ and $\beta = 0$, the algorithm
degrades to the conventional K-means. In the modified K-means in this work, we introduce another matrix $B$, which is generated by sorting the distance matrix by column. If matrix $B$ dominates the clustering process, a certain object can achieve a relatively high ranking in the sparse area with fewer objects, but a relatively poor ranking in the dense area with lots of objects, even with the similar Euclidean distance toward the certain centroid. This proposed algorithm can help balance the cluster occupation and the uniformity of the cluster size.

Figure 6 demonstrates the comparison of the conventional K-means and proposed K-means with the implementation of the two clustering methods on the objects in Figure 5. The blue dash circles in Figure 6a indicate the clusters of conventional K-means, and the red circles are the results of the proposed K-means. It is clear that the sizes of the blue dash clusters are relatively uniform on the whole. The red circles are larger within the $\pm 8.5^\circ$ FOV, where there are fewer objects as shown in Figure 5. Figure 6b shows the statistics of the beam occupation based on the conventional K-means and the proposed K-means. It is obvious that the proposed K-means helps to improve the uniformity of the beam occupation and decreases the deviation of the number of objects per beam. Figure 6c shows the statistics of the beam width, which numerically demonstrate the underlying tradeoff between the uniformity of the beam occupation and the uniformity of the beam width.

**Figure 6.** The primary results of the beam allocation of objects in Figure 5 following the conventional K-means and proposed K-means, respectively: (a) Plot of the allocated 40 beams. The crossing indicates the geometrical center of each beam. The blue dash circles indicate the clusters of conventional K-means; the red circles are the results of the proposed K-means; (b) Statistics of the beam occupation; (c) Statistics of the beam width.

As shown in Figure 2, the final step is to verify and evaluate the results, which are introduced in the next section.

### 3. Results

In this work, a typical LEO satellite constellation is considered to generate the analysis data set, as shown in Figure 5, along continuous three moments with one-second intervals to mitigate the impact of randomness. There are 1638 objects in the FOV, which have the line-of-sight communication capability with the GEO data relay satellite. A medium-scale GEO communication satellite with 40 user beams is taken into account in the research. So, the optimization problem can be summarized as a total of 1638 objects being divided into 40 clusters by using the modified K-means clustering with adjustable parameters. Compared with the conventional K-means clustering, two parameters $\alpha$ and $\beta$ are introduced to the modified algorithm. The impact of the parameters on the clustering results is demonstrated in the following subsections.

#### 3.1. Impact of Parameter $\alpha$

In this work, the Euclidean distance $M_{ij}$ in Equation (1) is a 1638-by-40 matrix. It is sorted by row and column, respectively, to obtain $A_{ij}$, which sorts the distances to the object
from each centroid, and \( B_{ij} \), which sorts the distances to the centroid from each object. The parameter \( \alpha \) in Equation (3) is the weighting scale to mix normalized \( A_{ij} \) and \( B_{ij} \) and obtain determination matrix \( D_{ij} \).

In this subsection, the value of \( \beta \) is kept at 0, which means that there is no obligatory minimal number of objects in each beam. The allocation of objects is conducted directly by Equation (5).

Figure 7 illustrates the results of the beam allocation based on the modified K-means clustering when parameter \( \alpha \) equals 1 and \( \beta \) equals 0. Figure 7a is the plot of the allocated 40 beams. The crossing indicates the centroid of each beam. The neighboring hexagonal cells in the same beam are marked with the same color. Figure 7b is the statistics of the beam occupation. The number of objects in each beam ranges from 15 to 95 with the average number of 40.95 and standard deviation of 16.55. Figure 7c is the statistics of the beam width. The half-cone angle of beam ranges from 1.2° to 3° with the average number of 1.96° and standard deviation of 0.45°.

In this case, the determination distance matrix \( D \) degenerates to the classical K-means algorithm determination matrix since \( \alpha \) equals 1. The allocation process is entirely based on the sorting of distances to the object from each centroid, which results in the sum of distance at a minimum.

Figure 7. Optimization results of the beam allocation based on the modified K-means clustering when parameter \( \alpha \) equals 1 and \( \beta \) equals 0: (a) Plot of the allocated 40 beams. The crossing indicates the geometrical center of each beam. The neighboring hexagonal cells in same beam are marked with same color; (b) Statistics of the beam occupation; (c) Statistics of the beam width.

Figure 8 illustrates the results of the beam allocation when parameter \( \alpha \) equals 0 and \( \beta \) equals 0. Figure 8a shows the plot of the allocated 40 beams. Figure 8b shows the statistics of the beam occupation. The number of objects in each beam ranges from 20 to 85 with the average number of 40.95 and standard deviation of 12.92. Figure 8c shows the statistics of the beam width. The half-cone angle of beam ranges from 0.6° to 4.6° with the average number of 1.77° and standard deviation of 0.95°.

This is the case at the opposite extreme when we change the value of \( \alpha \) from 1 to 0. The allocation process is entirely based on the sorting of distances to the centroid from each object. This process generates relatively larger beams in the area with sparse satellites, which mitigates the standard deviation of the object number in each beam from 16.55 to 12.92.

When \( \alpha \) is equal to 0, the determination matrix \( D \) is entirely up to \( B_{ij} \) based on Equation (3). \( B_{ij} \) is the ranking results after sorting the Euclidean distance matrix \( M \) by column according to Equation (2). In other words, the sorting is executed toward the centroid from each object in the pool of samples, the final ranking, also known as the numerical order or sequence number, is normalized to generate the determination matrix \( D \) according to Equation (3). In this case, a certain object \( i \) can obtain a relatively high ranking in the sparse area with fewer objects, but a relatively poor ranking in the dense area with lots of objects, even with the similar Euclidean distance toward the certain centroid \( j \). When \( \beta \) is equal to 0, which means there is no limitation of a minimum number in each cluster, the object \( i \) with the top ranking in column \( j \) is allocated toward cluster \( j \) according to Equation
(5). So, in sparse area, the objects with a relatively larger Euclidean distance to the centroid can also obtain good ranking and be allocated to the cluster, compared with the ones in the dense area. The aforementioned process guarantees the balance of the beam occupation by the proposed ranking mechanism, which offers a greater chance of good ranking for objects with a distance disadvantage in the sparse area.

Figure 8. Optimization results of the beam allocation based on the modified K-means clustering when parameter $\alpha$ equals 0 and $\beta$ equals 0: (a) Plot of the allocated 40 beams. The crossing indicates the geometrical center of each beam. The neighboring hexagonal cells in the same beam are marked with the same color; (b) Statistics of the beam occupation; (c) Statistics of the beam width.

3.2. Impact of Parameter $\beta$

To balance the object number in each beam with a forceful method, we introduce parameter $\beta$. This helps to determine the minimum number $N_{BL}$ of objects in a single beam as shown in Equation (4). In Step 6 of the proposed algorithm, the smallest available objects $N_{BL}$ in $D_j$ are obligatorily allocated to cluster $j$.

In this subsection, the optimization results are illustrated when parameter $\beta$ equals 0.2, 0.5 and 0.8, and the other variable $\alpha$ remains at 0.5.

Figure 9 illustrates the results of the beam allocation based on the modified K-means clustering when parameter $\alpha$ equals 0.5 and $\beta$ equals 0.2. Figure 9a shows the plot of the allocated 40 beams. The crossing indicates the centroid of each beam. The neighboring hexagonal cells in the same beam are marked with the same color. Figure 9b shows the statistics of the beam occupation. The number of objects in each beam ranges from 15 to 75 with the average number of 40.95 and standard deviation of 12.55. Figure 9c shows the statistics of the beam width. The half-cone angle of beam ranges from 1.2° to 3° with the average number of 1.95° and standard deviation of 0.46°.

Figure 9. Optimization results of the beam allocation based on the modified K-means clustering when parameter $\alpha$ equals 0.5 and $\beta$ equals 0.2: (a) Plot of the allocated 40 beams. The crossing indicates the geometrical center of each beam. The neighboring hexagonal cells in the same beam are marked with the same color; (b) Statistics of the beam occupation; (c) Statistics of the beam width.
Figure 10 illustrates the results of the beam allocation when parameter $\alpha$ equals 0.5 and $\beta$ equals 0.5. Figure 10a shows the plot of the allocated 40 beams. Figure 10b shows the statistics of the beam occupation. The number of objects in each beam ranges from 20 to 75 with the average number of 40.95 and standard deviation of 12.03. Figure 10c shows the statistics of the beam width. The half-cone angle of beam ranges from $1^\circ$ to $3.2^\circ$ with the average number of $1.96^\circ$ and standard deviation of $0.49^\circ$.

Figure 11 illustrates the results of the beam allocation when parameter $\alpha$ equals 0.5 and $\beta$ equals 0.8. Figure 11a shows the plot of the allocated 40 beams. Figure 11b shows the statistics of the beam occupation. The number of objects in each beam ranges from 30 to 80 with the average number of 40.95 and standard deviation of 9.87. Figure 11c shows the statistics of the beam width. The half-cone angle of beam ranges from $1.1^\circ$ to $3.5^\circ$ with the average number of $1.98^\circ$ and standard deviation of $0.56^\circ$.

All of the aforementioned results are summarized in Table 1 and are discussed in the next section.
Table 1. Statistic of the impact of $\alpha$ and $\beta$ on the results.

| $\alpha$ | $\beta$ | Number of Objects in Each Beam | Half-Cone Angle of Beam |
|----------|---------|--------------------------------|-------------------------|
|          |         | Avg. | Std. | Avg. | Std. |
| 1        | 0       | 16.55 | 1.96° | 1.65° | 0.45° |
| 0        | 0       | 12.92 | 1.77° | 1.95° | 0.95° |
| 0.5      | 0.2     | 12.55 | 1.95° | 1.77° | 0.46° |
| 0.5      | 0.5     | 12.03 | 1.96° | 1.77° | 0.49° |
| 0.5      | 0.8     | 9.87  | 1.98° | 1.95° | 0.56° |

4. Discussion

The optimization results demonstrate the impact of parameters $\alpha$ and $\beta$ on the clustering process, respectively. The modified K-means clustering introduces a tradeoff between the uniformity of the beam occupation and the uniformity of the beam width. On the one hand, the reduction in $\alpha$ helps in tuning the determination matrix from sorting the distances to the object from each centroid to sorting the distances to the centroid from each object. On the other hand, the growth of $\beta$ keeps increasing the obligatory minimum number of objects in a single beam.

Figure 12 illustrates the standard deviation of objects per beam in terms of $\alpha$ and $\beta$. The minimum value equals 6.95 when $\alpha$ is 0.2 and $\beta$ is 0.8. This indicates that a larger value of $\beta$ with a relative smaller value of $\alpha$ tends to obtain uniformity of the beam occupation. On the contrary, if $\beta$ is smaller than 0.5 and $\alpha$ is close to 1, the standard deviation is raised to c.a. 16, which means a greater variation in the object number per beam.

Figure 12. Standard deviation of objects per beam in terms of $\alpha$ and $\beta$.

The uniformity of the beam occupation is a key specification that determines the performance. However, to make the whole proposal feasible, other factors should also be taken into consideration, such as the key characteristics of the antenna system. Figure 13a illustrates the standard deviation of the half-cone angles in terms of $\alpha$ and $\beta$. Figure 13b illustrates the maximum of the half-cone angles in terms of $\alpha$ and $\beta$.

As mentioned above, the uniformity of the beam occupation is maximized when $\alpha$ is 0.2 and $\beta$ is 0.8. In this case, the maximum of the half-cone angles is 3.94° and the standard deviation is 0.79°. The growth of $\alpha$ and reduction of $\beta$ mitigates the large variation of the beam widths. The variance is the key characteristic that challenges the performance of the antenna and the feasibility of the proposal. The uniformity of the beam width can be obtained when $\alpha$ is 1 and $\beta$ is 0.4, which suggests that the determination matrix is degenerated to the conventional K-means with a moderate obligatory minimum number
of objects per beam. In this case, the maximum of the half-cone angles is $2.57^\circ$ and the standard deviation is $0.33^\circ$.

It is worth noting that there is a plateau when $\alpha$ is larger than 0.6 and $\beta$ is smaller than 0.4 as shown in Figures 11 and 12. The moderate obligatory minimum number of objects per beam and the modified determination matrix tune the results to be similar with conventional K-means clustering, which aims at the uniformity of the beam sizes. However, the uniformity of the beam occupation can only be obtained from the plateau area with a larger value of $\beta$ and smaller value of $\alpha$.

Finally, the computational complexity of the proposed algorithm is studied. The calculating time of Equation (1) to generate the Euclidean distance matrix is $O(nkd)$, where $n$ is the number of objects, $k$ is the number of clusters and $d$ is the dimension. This step has the same implementation cost with the conventional K-means algorithm [32]. Then, the distance matrix is sorted twice by row and column following Equation (2). The running time to generate matrix $A$ is $O(nk \log k)$ and the time to generate $B$ is $O(kn \log n)$. Finally, the allocating process costs $O(nk)$ following Equation (5) and the updating of the centroids costs $O(nd)$. To sum up, the proposed algorithm can be implemented in time $O((nkd + nk \log k + kn \log n + nk + nd)i) = O((nkd + kn \log n)i)$, where $i$ is the number of iterations needed for convergence. The time complexity can be simplified as $O(n \log n)$, if $k$, $d$, $i$ are treated as constants. This value is relatively larger than the running time $O(n)$ of the conventional K-means algorithm, as an additional ranking process is introduced to balance the uniformity of the object number in the cluster. Figure 14 demonstrates the computational time in MATLAB of the proposed algorithm in terms of the number of objects. It is worth noting that the algorithm is proposed to guide the design of the fixed beam antenna of the multi-beam satellite. It would be implemented only in the early stages of design for task analysis until the allocation of the beam spots is completed; there is no real-time requirement in this application scenario.
5. Conclusions

The multi-beam satellite often faces the mismatch problem of spot allocation and data requirements, which can cause an overload traffic jam or a waste of resources. In this work, we introduced a modified K-means method to automatically place the spot centers with appropriate beam widths in line with the density of the traffic demands and to realize the uniformity of the beam occupation. Compared with the conventional K-means algorithm, two adjustable parameters $\alpha$ and $\beta$ were introduced: one for tuning the ratio of determination matrix and the other for setting the obligatory minimum number of objects per beam. The impact of the parameter range on the results was also analyzed. A larger value of $\beta$ with a relative smaller value of $\alpha$ tends to obtain the uniformity of the beam occupation; the minimum standard deviation of objects per beam is achieved when $\alpha$ is 0.2 and $\beta$ is 0.8. When $\alpha$ is larger than 0.6 and $\beta$ is smaller than 0.4, the results tend to the uniformity of beam width, similar to conventional K-means, with a relatively larger but stable standard deviation of objects per beam. The algorithm would be theoretically degenerated to the conventional K-means when $\alpha$ equals 1. The results have proven the validity of the proposed algorithm, which helps to introduce a trade-off strategy between the uniformity of the beam occupation and the uniformity of the beam width and would be implemented in the engineering practices in the field of multi-beam satellite design. In the future, the proposed strategy to balance the uniformity of cluster occupation will be considered to transfer to other unsupervised learning algorithms, such as Fuzzy C-means, and be implemented in more comprehensive applications.

Author Contributions: Conceptualization, H.L. and Y.C. (Yaxing Cai); methodology, H.L. and Y.C. (Yao Chu); software and experiments, H.L. and Y.C. (Yao Chu); writing—original draft preparation, H.L.; writing—review and editing, Y.C. (Yao Chu) and Y.Z.; supervision, W.H., Y.L. and Y.Y.; funding acquisition, Y.C. (Yao Chu) and Y.C. (Yaxing Cai). All authors have read and agreed to the published version of the manuscript.

Funding: Y. Chu gratefully acknowledges the financial support from the China Postdoctoral Science Foundation (2020M670361, 2021T140392) and the Shimu Tsinghua Scholar Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.
References

1. Letaief, K.B.; Chen, W.; Shi, Y.; Zhang, J.; Zhang, Y.A. The Roadmap to 6G: AI Empowered Wireless Networks. *IEEE Commun. Mag.* 2019, 57, 84–90. [CrossRef]

2. Chen, S.; Sun, S.; Kang, S. System integration of terrestrial mobile communication and satellite communication—The trends, challenges and key technologies in 5G and 6G. *China Commun.* 2020, 17, 156–171. [CrossRef]

3. Leyva-Mayorga, I.; Soret, B.; Popovski, P. Inter-Plane Inter-Satellite Connectivity in Dense LEO Constellations. *IEEE Trans. Wirel. Commun.* 2021, 20, 3430–3443. [CrossRef]

4. Di, B.; Zhang, H.; Song, L.; Li, Y.; Li, G.Y. Ultra-Dense LEO: Integrating Terrestrial-Satellite Networks into 5G and beyond for Data Offloading. *IEEE Trans. Wirel. Commun.* 2019, 18, 47–62. [CrossRef]

5. Stock, G.; Fraire, J.A.; Mömke, T.; Hermanns, H.; Babayev, F.; Cruz, E. Managing Fleets of LEO Satellites: Nonlinear, Optimal, Efficient, Scalable, Usable, and Robust. *IEEE Trans. Comput. Aid. Des. Integr. Circuits Syst.* 2020, 39, 3762–3773. [CrossRef]

6. Li, J.; Ye, G.; Zhang, T.; Zhang, T. A TT&C management method for satellite constellation networking process. In Proceedings of the 2017 Intelligent Systems Conference (IntelliSys), London, UK, 7–8 September 2017; pp. 274–282.

7. Wang, L.; Jiang, C.; Kuang, L.; Wu, S.; Huang, H.; Qian, Y. High-Efficient Resource Allocation in Data Relay Satellite Systems With Users Behavior Coordination. *IEEE Trans. Veh. Technol.* 2018, 67, 12072–12085. [CrossRef]

8. Si-Wei, C.; Hui, Z.; Chao, W.; Jing, C.; Jun-Sheng, Z. Operation Planning Modeling of Tracking and Data Relay Satellite: A Sketch. In Proceedings of the 2009 Second International Conference on Intelligent Computation Technology and Automation, Changsha, China, 10–11 October 2009; pp. 144–147.

9. Brandel, D.L.; Watson, W.A.; Weinberg, A. NASA’s advanced tracking and data relay satellite system for the years 2000 and beyond. *Proc. IEEE* 1990, 78, 1141–1151. [CrossRef]

10. Berretta, G.; Agostini, A.; Dickinson, A. The European data relay system: Present concept and future evolution. *Proc. IEEE* 1990, 78, 1152–1164. [CrossRef]

11. Chen, X.; Li, X.; Wang, X.; Luo, Q.; Wu, G. Task Scheduling Method for Data Relay Satellite Network Considering Breakpoint Transmission. *IEEE Trans. Veh. Technol.* 2021, 70, 844–857. [CrossRef]

12. Elertsen, B. Ground station networks vs. GEO relay satellite systems for polar orbiting satellites. In Proceedings of the SpaceOps 2012 Conference, Stockholm, Sweden, 11–15 June 2012; American Institute of Aeronautics and Astronautics: Reston, VA, USA, 2012.

13. Katona, Z. GEO data relay for low earth orbit satellites. In Proceedings of the 2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC), Vigo, Spain, 5–7 September 2012; pp. 81–88.

14. Greder, L.A.; Knüpfel, B.; Knogl, J.S.; Heckler, M.V.T.; Bischl, H.; Dreher, A. A multibeam antenna for data relays for the German communications satellite Heinrich-Hertz. In Proceedings of the Fourth European Conference on Antennas and Propagation, Barcelona, Spain, 12–16 April 2010; pp. 1–4.

15. Pang, T.K.; Trachtman, E. Inter-Satellite Data Relay System (IDRS) for LEO Satellites Using a Commercially Available GEO Satellite System. In Proceedings of the Small Satellite Conference, Virtual, 1–6 August 2020.

16. Kopp, B.; Harris, J.; Lauand, C. Utilizing Existing Commercial Geostationary Earth Orbit Fixed Satellite Services for Low Earth Orbit Satellite Communication Relays with Earth. *New Space* 2019, 7, 19–30. [CrossRef]

17. Toyoshima, M. Technological trends and future prospects of satellite communications for mega-constellations with small satellites. In [International Communications Satellite Systems Conference]. *IET Conf. Proc.* 2019, 29, 28–29.

18. Xiao, T.; Wei, D. Design of multibeam antenna with different beam size. In Proceedings of the 2019 International Applied Computational Electromagnetics Society Symposium—China (ACES), Nanjing, China, 8–11 August 2019; p. 9060596.

19. Ding, W.; Tao, X. Ka-band multibeam antenna for Chinasat-16. In Proceedings of the 2019 International Applied Computational Electromagnetics Society Symposium—China (ACES), Nanjing, China, 8–11 August 2019; p. 9060682.

20. Farserotu, J.; Prasad, R. A survey of future broadband multimedia satellite systems, issues and trends. *IEEE Commun. Mag.* 2000, 38, 128–133. [CrossRef]

21. Camino, J.; Mourguès, S.; Artigues, C.; Houssin, L. A greedy approach combined with graph coloring for non-uniform beam layouts under antenna constraints in multibeam satellite systems. In Proceedings of the 2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC), Livorno, Italy, 8–10 September 2014; pp. 374–381.

22. Camino, J.-T.; Artigues, C.; Houssin, L.; Mourguès, S. Linearization of Euclidean norm dependent inequalities applied to multibeam satellites design. *Comput. Optim. Appl.* 2019, 73, 679–705. [CrossRef]

23. Camino, J.; Artigues, C.; Houssin, L.; Mourguès, S. Mixed-integer linear programming for multibeam satellite systems design: Application to the beam layout optimization. In Proceedings of the 2016 Annual IEEE Systems Conference (SysCon), Orlando, FL, USA, 18–21 April 2016; pp. 1–6.

24. Tsiropoulou, E.E.; Paruchuri, S.T.; Baras, J.S. Interest, energy and physical-aware coalition formation and resource allocation in smart IoT applications. In Proceedings of the 2017 51st Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD, USA, 22–24 March 2017; pp. 1–6.

25. Sarakis, L.; Kormentzas, G. Handovers for Ubiquitous and Optimal Broadband Connectivity among Cooperative Networking Environments. In *Mobile Networks and Management*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 3–13.
26. Xu, F.; Li, Y.; Wang, H.; Zhang, P.; Jin, D. Understanding Mobile Traffic Patterns of Large Scale Cellular Towers in Urban Environment. *IEEE ACM Trans. Netw.* 2017, 25, 1147–1161. [CrossRef]

27. Likas, A.; Vlassis, N.; Verbeek, J. The global k-means clustering algorithm. *Pattern Recognit.* 2003, 36, 451–461. [CrossRef]

28. Kang, S. k-Nearest Neighbor Learning with Graph Neural Networks. *Mathematics* 2021, 9, 830. [CrossRef]

29. Matkowski, J.; Pasteczka, P. Invariant Means, Complementary Averages of Means, and a Characterization of the Beta-Type Means. *Mathematics* 2020, 8, 1753. [CrossRef]

30. Yi, F.; Moon, I. Extended K-Means Algorithm. In Proceedings of the 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, Washington, DC, USA, 26–27 August 2013; pp. 263–266.

31. Moon, I. K-means based Clustering Method with a Fixed Number of Cluster Members. *J. Korea Multimed. Soc.* 2014, 17, 1160–1170. [CrossRef]

32. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc. Ser. C* 1979, 28, 100–108. [CrossRef]