Characteristics of Dynamic Connection and Path Spatial-temporal Evolution in Cluster Flight Spacecraft Network

Yanfeng Shi[1,2], Shengbo Hu[1,2], Jinrong Mo[1,2], Xiaowei Song[1,2], Tingting Yan[1,2]

[1] Institute of Intelligent Information Processing, Guizhou Normal University, China
[2] Department of Education, Center for RFID and WSN Engineering, China
syf@gznu.edu.cn, hsb@nssc.ac.cn, 935472997@qq.com, 935634557@qq.com, 1592624854@qq.com

Abstract

In recent years, spacecraft cluster flight has attracted more and more attention in the field of distributed spacecraft systems. The high-speed flight of the cluster flight spacecraft modules increases the uncertainty of network topology. In order to optimize the orbital design of the cluster flight spacecraft and improve the performance of cluster flight spacecraft network (CFSN), based on the dynamic connection of nodes, this paper studies the characteristics of dynamic connection and path spatial-temporal evolution by the probabilistic connectivity matrix in CFSN. Firstly, based on twin-satellites mode, we establish the mobility model of nodes. And then by adopting empirical statistical method and curve fitting method, the solution of the nodal distance density function in the CFSN is obtained and the threshold range of nodal connection distance are derived under the constraints of CFSN. Finally, using the orbital data generated by STK (Satellite Tool Kit), through the definition of sequential path and a new matrix multiplication, the probabilistic connectivity matrix of sequential path of multiple hops between nodes is obtained. And the characteristics of dynamic connection and path spatial-temporal evolution in a orbital hyper-period is studied. These results can provide theoretical reference for the design and optimization of CFSN.

Keywords: Cluster flight spacecraft network (CFSN), Nodal connection, Probabilistic connectivity matrix, Spatial-temporal evolution

1 Introduction

In recent years, fractionated spacecraft with cluster flight model has become a hot topic in the field of distributed space network, due to its advantages of flexibility, rapid response, low cost, strong scalability and long lifetime. The previous work has made contribution to earth observation and space exploration [1-3]. Fractionated spacecraft distributes the functionality of a traditional large monolithic spacecraft into a number of heterogeneous modules. Each module can be viewed as a node through wireless communication, and the nodes construct the cluster flight spacecraft network (CFSN). Cluster flight spacecrafts require mutual cooperation between nodes to realize information exchange, navigation communication and power sharing. These spacecrafts constitute a virtual satellite platform with information exchange structure [4]. However, because of the high-speed flight of the cluster spacecraft module and the uncertainty of the network topology, the connection between nodes and the path formation change with time and space. There are many challenges to the performance analysis and optimization design of the CFSN [5-6]. In this paper, based on the dynamic connection of cluster flying spacecraft nodes in the early stage [7], the dynamic connection of cluster flying spacecraft network based on probability connection matrix and the spatial-temporal evolution characteristics of the path are studied, in order to further explore the network performance of cluster flying spacecraft.

In fact, for many years, the dynamic connection and path of wireless mobile communication networks, including ad hoc and P2P, have always been research hotspots, and geometric and random graph theory methods are widely used. For example, Wu establishes a linear node movement model and uses geometric aids to study the dynamic connection and path evolution characteristics of nodes in VANET networks under transmission delay constraints [8]. Brook et al. studied the dynamic connection and path evolution characteristics of self-organizing wireless sensor networks by establishing a Bernoulli probability model with node connections and using random graph theory [9]. Padmavathy et al. compared the complexity of dynamic node connection reliability calculation with mobile model and no-movement model for mobile Ad hoc networks [10]. Recently, Mao used random graph theory and probability connection matrix to study the network connection characteristics through the analysis

*Corresponding Author: Shengbo Hu; E-mail: hsb@nssc.ac.cn
DOI: 10.3966/160792642020092105027

of matrix eigenvalues [11].

It should be noted that the above studies are closely related to the distance distribution between nodes. In fact, including signal-to-interference ratio (SIR), node degree, outage probability, link capacity, etc. are also related to the distance distribution between nodes [12-16]. Therefore, the distance distribution between nodes is also a research hotspot. For example, Jone PM uses a stochastic geometric method to give the approximate link distance distribution of the random node moving model, but in some cases the distribution has no closed solution [17]; and KB Baltzis et al. A method for calculating the distance distribution function from node to node is proposed. The proposed method is suitable for overlapping and non-overlapping mobile networks [18]. The idea of using measure in integral geometry [19], Tong Fei et al. a cellular network of a single sub-area, and gives the results of the distance between some nodes [20]; subsequently, they proposed a systematic, arithmetic method to obtain the distance distribution between random nodes of an arbitrary shape wireless network, and obtained The distance distribution function between nodes is usually an implicit [21]. To calculate the distance distribution between nodes, we usually need to know the network shape and the node distribution information. In order to simplify system analysis and reduce computational complexity, square [22] or circular-shaped [23] networks are usually assumed. One of the main disadvantages of the circle is that it can't describe the statistical information of the distance between nodes in a closed form. Therefore, it is still a complicated problem to solve the closed distance distribution between nodes.

Compared with the previous studies, the dynamic connection and path formation of the CFSN is more complicated. In addition to considering the connection and path between nodes at a certain moment, the satellite node has the function of storage and forwarding, and also considers the connection and path between nodes at different times, that is, the dynamic connection of the node and the space-time evolution of the path. This study provides a theoretical support for spacecraft cluster flight including on orbit deployment [24].

The contributions of this paper can be summarized as follows:

1. Based on the spacecraft double-star companion mode, this paper establishes the cluster flight spacecraft node movement model and uses the empirical statistical analysis method to study the distance distribution between the nodes of the spacecraft and the probability connection matrix of the nodes.

2. By defining the sequential path, study the dynamic connection between network nodes and the spatial-temporal evolution characteristics of paths.

The rest of this paper organized as follows. In Section 2, First, the earth-centered inertial (ECI) coordinate system is introduced. Then the dynamic topology evolution graph of cluster flying spacecraft network and the sequential path of cluster flying spacecraft network are defined. Finally, the nodal mobility model is established, and the probability connectivity matrix between nodes is defined. Section 3 describes the calculation method of the connection threshold between nodes. Section 4 describes the spatial-temporal evolution graphs between nodes between nodes based on the probability connectivity matrix. Section 5 concludes the paper.

2 System Model

2.1 Earth-Centered Inertial

The orbits of cluster flight spacecrafts can be described using the ECI coordinate system. Earth-centered inertial (ECI) coordinate system is defined in the following standard manner: the fundamental plane is the equatorial plane, the x axis towards the vernal equinox, the z axis points towards the geographic north pole, and $y = z \times x$, as shown in Figure 1.

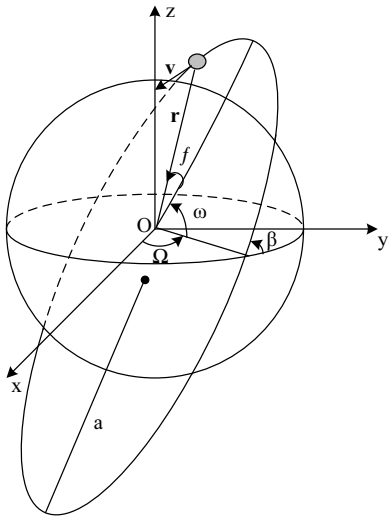


Figure 1. The schematic of the orbital elements in ECI coordinate

The vector α of classical orbital elements in ECI coordinate, which describes natural orbits of a cluster flight spacecraft, is defined as:

$$\alpha \triangleq [a, e, \beta, \omega, f, \Omega]^{\wedge T} \quad \textbf{(1)}$$

Where α is the semi-major axis, e is the eccentricity, $\beta$is the inclination, $\omega$is the argument of perigee, f is the true anomaly, and $\Omega$is the right ascension of ascending node (RAAN).

If the vector $r = [x, y, z]^T$ denotes the position of any satellite in CFSN in ECI coordinate, $v = dr/dt$ is the velocity, $\eta = \sqrt{x^2 + y^2}$ is equatorial projection of

position vector, as well as the maximal and minimal equatorial projections of the position vector of satellite at time *t*, given by $\eta \max_t \eta(t)_{\max}$ and $\eta \min_t \eta(t)_{\min}$, where $\eta(t) = \sqrt{x^2(t) + y^2(t)}$.

2.2 Cluster Flight Spacecraft Network Topology Evolution Graphs and Sequential Path

In a CFSN consisting of N nodes, V denotes the nodes set, and $V \triangleq \{v_1, v_2, \ldots, v_N\}$.

For the sake of analysis, taking frequency division multiple access (FDMA) with subcarrier binary phase shift keying (BPSK) modulation as an example, we establish the link constraints between nodes in CFSN. The system is supposed to be noise limited also.

Generally, given a certain time slot $\delta_k (k = 1, 2, 3\cdots\}$, let vector $r_i(k)$ and $r_j(k)$ denote positions of nodes v_i and v_j, if the cluster spacecraft network node $v_i (i = 1, 2, ..., N\}$ whose position vector is $r_i(k)$ and transmit power P transmits data to the node $v_j (i = 1, 2, ..., N\}$ whose position vector is $r_j(k)$, the other nodes also use the same power P to the node v_j. When the data is transmitted and the signal-to-interference ratio at the node satisfies Equation (2), the node is considered to successfully receive the data sent by the node:

$$\mathrm{SIR} = \frac{Pl\left(\mathbf{r}_i(k) - \mathbf{r}_j(k)\right)}{\sum_{w \neq i,j} Pl\left(\mathbf{r}_w(k) - \mathbf{r}_j(k)\right) S_{\mathrm{BPSK}}\left(\left|f_w - f_j\right|\right)} \geq \Gamma \quad \textbf{(2)}$$

where $S_{BPSK}(|f_w - f_j|)$ is approximately:

$$S_{\mathrm{BPSK}}\left(\left|f_w - f_j\right|\right) \approx \frac{1}{\left(2\pi T_f \left|f_w - f_j\right|\right)^2} \quad \textbf{(3)}$$

Γ is threshold value of SIR, $l(r_i - r_j)$ is path loss factor with respect to channel, $S_{BPSK}(|f_w - f_j|)$ is power spectrum in BPSK modulation, T_f is the symbol duration, f_i is subcarrier frequency of node *j*.

The Equation (2) shows that the construction of network topology depends entirely on the path loss factor $l(r_i - r_j)$, if P and Γ are known. When the value of SIR between nodes v_i and v_j is greater than Γ, there exists a link between them, i.e. an edge of the topological graph. So, the value of SIR directly affects the construction of network topology.

Considering $l((r_i(k) - r_j(k))) = |r_i(k) - r_j(k)|^{-2}$ in free space, we can rewrite (3) as following:

$$\mathrm{SIR} = \frac{\left|\boldsymbol{r}_i(k) - \boldsymbol{r}_j(k)\right|^{-2}}{\sum_{w \neq i,j} \left|\boldsymbol{r}_w(k) - \boldsymbol{r}_j(k)\right|^{-2} S_{\mathrm{BPSK}}\left(\left|f_w - f_j\right|\right)} \quad \textbf{(4)}$$

Let $S_{BPSK}(|f_w - f_j|)$, the distance between the nodes v_i and v_j, $|r_i(k) - r_j(k)| = d$, if d satisfies the following:

$$d^2 \leq \frac{1}{\Gamma S \sum_{w \neq i,j} \left|\boldsymbol{r}_w(k) - \boldsymbol{r}_j(k)\right|^{-2}} \quad \textbf{(5)}$$

There exists an edge between nodes v_i and v_j. In conclusion, the connection problem of the model is transformed into the problem of nodal distance distribution.

For the CFSN, the topological graph is denoted by $G(V, E)$, where $V \triangleq \{v_1, v_2, \ldots, v_N\}$ denotes the nodes set and E denotes the edge set. The graph is obtained by connecting each node in the network using a node and converse.

Based on the theory of orbital dynamics, the orbital hyper-period can be divided into $T_0, T_1, T_2, \ldots, T_T$ times for cluster flight spacecrafts [25], so there are T time slots in an orbital hyper-period. The orbital hyper-period is $C = (T_T - T_0)$ [26]. If the topology is supposed to remain static in time slot $\delta_k = [T_{k-1}, T_k]$ $(k = 1, 2, \ldots, T)$, we can define the evolving graph of dynamic topology graph for CFSN as follows:

Definition 1. If the orbital hyper-period can be divided into $T_0, T_0, T_2, \ldots, T_T$ times for CFSN, there are T time slots. If the edge set is E_k at each time slot $\delta_k = [T_{k-1}, T_k]$, the evolving graph of dynamic topology for CFSN in the orbital hyper-period is defined as:

$$G(\mathrm{V}, \mathrm{E}) = \left(\mathrm{V}, \bigcup_{k=1}^{T} \mathrm{E}_k\right).$$

According to the definition 1, the topological graph of the $k-thk-th$ time slot δ_k is denoted by $G(V, E_k)$.

Since the satellite has the function of storage and forwarding, according to the definition 1, if the start node sequential pair, that is, the node sequence pair at time slot δ_1 is $(v_i(1), v_i(1))$, and the sequential pairs at time slot δ_k and time slot δ_{k+1} are $(v_m(k), v_n(k))$ and $(v_m(k+1), v_n(k+1))$ $(i \neq m \neq n)$ respectively. The sequential path of CFSN can be defined:

Definition 2. For a CFSN, a sequential path is a time series of paths in a topological graph, which can be expressed as:

$$< (v_i(1), v_i(1)), (v_i(2), v_i(2)) \cdots, (v_m(1), v_n(1)), (v_n(k+1), v_j(k+1)), \cdots > .$$

In continuous time, data is transmitted from node v_i

to v_j through a sequential path as shown in Figure 2. As can be seen from Figure 2, the node v_i stores data in the first time slot, and the node v_i transmits data to the node v_i in the second time slot, ... and in the k time slot, the node v_m sends the information to the node v_n, and finally in the $k+1$ time slot, the node v_n sends information to the node v_j.

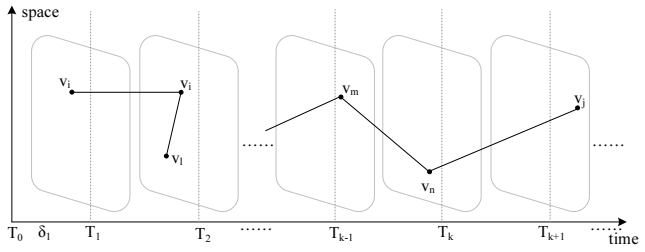


Figure 2. The sequential path

2.3 Nodal Mobility Model and Probabilistic Connectivity Matrix

The nodal mobility model describes all the possible distribution area among nodes that enable resource exchange. To accomplish cluster flight model within bounded distance, for the sake of analysis, we adopt twin-satellites mode to research the mobility model of network nodes. So, we considered that the node position is uniformly distributed on circle within $(M-m)/4$ radius in Figure 3. M is the upper bound of nodal distance in cluster flight, m is the lower bound.

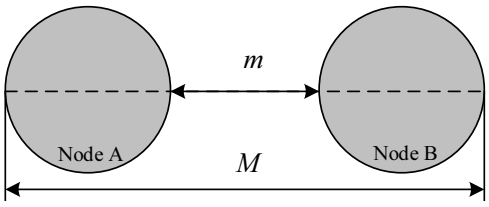


Figure 3. Distribution area of node positions

So the mobility model $M(t)$ for CFSN can be defined as:

Definition 3. In ECI coordinates, if the position sets of N nodes in CFSN is $R(0)=\{r_1(0), r_2(0), ..., r_N(0)\}$ at initial time T_0, the position set is $R(k)=\{r_1(k), r_2(k), ..., r_N(k)\}$, and the positions are uniformly distributed within sphere $B(r_1(0), a)$, $(i=1, 2, ..., N)$ at time T_k, where $r_1(0)$ and $a=(M-m)/4$ are center and radius of the sphere respectively. Moreover, positions among all nodes are mutually independent and independent of all previous locations.

Based on topological graph and the definition of sequential path, in order to better study the connection of nodes, it is necessary to further study the probability connection characteristics between nodes of CFSN. First, the definition of the probabilistic connectivity matrix is given.

Definition 4. For a given CFSN, in each time slot of its orbital hyper-period, its probability connection matrix denoted by $\boldsymbol{Q}$ is a $N\times N$ matrix. q_{ij} is the (i, j) element of $\boldsymbol{Q}$, $q_{ij}=q_{ij}$. q_{ij} is the probability that nodes v_i and v_j are successfully connected. The diagonal entries are all equal to 1.

In definition 4, since the satellite node has the function of storage and forwarding, the diagonal element is 1.

If the topology is supposed to remain static in time slot $\delta_k=[T_{k-1}, T_k]$, the probabilistic connectivity matrix can be denoted by $\boldsymbol{Q}_k$ at time slot σ_k. Then we define the probabilistic connectivity matrix of dynamic topology for CFSN in the following way:

Definition 5. If the orbital hyper-period can be divided into $T_0, T_1, T_2, \ldots, T_T$ times for CFSN, there are T time slots. The probabilistic connectivity matrix of dynamic topology for CFSN in an orbital hyper-period is

$$\boldsymbol{Q}(1,T)=\boldsymbol{Q}_1\cup\cdots\cup\boldsymbol{Q}_T=\bigcup_{k=1}^{T}\boldsymbol{Q}_k$$

Where $\boldsymbol{Q}(1,T)$ is the probabilistic connectivity matrix in an orbital hyper-period with T time slots.

3 The Nodal Distance Distribution

3.1 Problem Description

According to the nodal mobility model in Section 2.2, we establish the two-dimensional cartesian coordinate system. Let's assume two nodes A and B are randomly located. In this case, the coordinates of the nodes are $(r_1\cos\varphi_1, r_1\sin\varphi_1)$ and $(D+r_2\cos\varphi_2, r_2\sin\varphi_2)$ with $r_{1,2}\in[0, R]$ and $\varphi_{1,2}\in[0, 2\pi]$, D with the value $(M+m)/2$ is the distance between the centers of the two-dimensional model, see Figure 4.

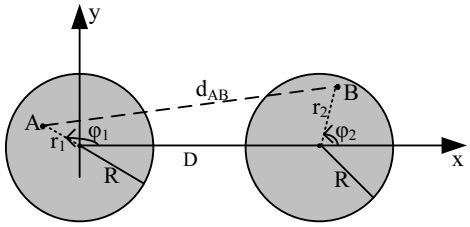


Figure 4. The two-dimensional cartesian coordinate system of nodal distribution

$r_{1,2}$ and $\varphi_{1,2}$ denote r_1, r_2, ϕ_1 and ϕ_2, respectively. r_1 and r_2 respectively represent the node-to-center distances in the respective distribution regions, and the ϕ_1 and ϕ_2 respectively represent the angles with the x-axis as the starting edge and r_1 and r_2 as the final edges. Therefore, the probability density functions of

the random variables r_1, r_2, ϕ_1 and ϕ_2 are expressed as

$$f_{r_{1,2}}\left(r_{1,2}\right)=\begin{cases}\dfrac{2r_{1,2}}{R^2}, & 0\le r_{1,2}\le R\\ 0, & else\end{cases} \tag{6}$$

$$f_{\varphi_{1,2}}\left(\varphi_{1,2}\right)=U\left(0,2\pi\right) \tag{7}$$

Where $U(0,2\pi)$ represents a uniform distribution over the range $[0,2\pi)$. According to the established cartesian coordinate system, the distance d_{AB} between the two nodes is a random variable calculated from

$$d_{AB}=\sqrt{\left(r_1\cos\varphi_1-\left(D+r_2\cos\varphi_2\right)\right)^2+\left(r_1\sin\varphi_1-r_2\sin\varphi_2\right)^2} \tag{8}$$

Despite the simplicity of above equation, the derivation of the distance density cannot be given in closed-form in non-overlapping circular-shaped networks.

3.2 Empirical Probability Density Function of Distance Between Nodes

According to the foregoing, we suggest an approach that is based on curve fitting, i.e. we approximate the distance density with polynomial.

Step 1: According to the model proposed in Section 3.1, $D=(M+m)/2$ is normalized, so that the empirical probability density function of the distance between nodes can be constructed as:

$$\hat{f}_h\left(h\right)\approx\begin{cases}\sum_{i=0}^{c}p_i h^i, & m/\mathrm{D}\le h\le M/D\\ 0, & else\end{cases} \tag{9}$$

Where p_i is the polynomial coefficient, c is polynomial order, and h is the distance uniformly distributed between any two nodes in the bounded range.

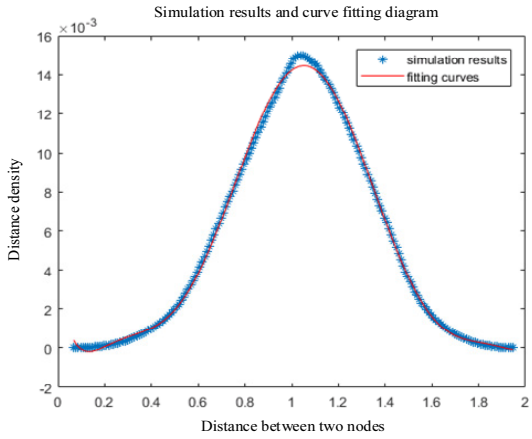
Step 2: Determine the number of samples. According to the relationship between the number of samples and the convergence and error [27-28], we assume that the error is greater than 10^{-3}, that is, the probability of $\varepsilon=10^{-3}$ is 10^{-6}, and the corresponding number of samples is at least 3×10^{-6}.

Step 3: Determine the polynomial order. The polynomial order affects the residual. After analysis, it is determined that the eighth-order polynomial approximation is used.

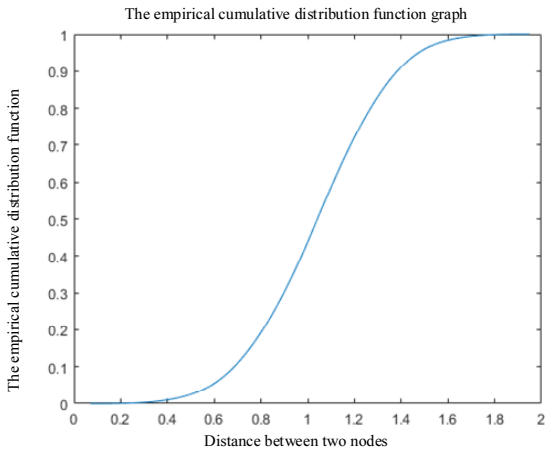
Step 4: Determine the polynomial coefficient p_i. Under the MATLAB environment, according to the node movement model shown in Figure 3, when, $m=2.5$km and $M=90.1km$, the random number is generated, that is, the possible spatial positions of the nodes. The distance between nodes is calculated by Equation (8). The experimental simulation parameters $(R^*,R^*)=(0.473,0.473)$. (R^*,R^*) represents the radius of the node distribution area in the two-dimensional model, and R^* is normalized by D. According to the definition of empirical cumulative distribution function in [27] and the minimum mean square error rule, the polynomial coefficients of the empirical probability density function are listed in Table 1. The approximate probability density function and the empirical cumulative distribution function of the distance between the nodes in the cluster flying spacecraft network are obtained and shown in Figure 5.

Table 1. Coefficients of the fitting polynomials

p_i	Value
p_0	0.0034
p_1	-0.0764
p_2	0.5919
p_3	-2.1499
p_4	4.1737
p_5	-4.3913
p_6	2.5135
p_7	-0.7380
p_8	0.0872



(a) Distance density



(b) The empirical cumulative distribution function

Figure 5. The diagrams of distance distribution function

Thus, the approximate probability density function $\hat{f}_h(h)$ of the distance h between nodes is obtained.

3.3 Calculation of Nodal Connective Distance Threshold

Under the constraints of CFSN, the threshold of connection between nodes is calculated. The relevant parameters are set as shown in Table 2.

Table 2. Parameters setting

Parameter	Value
Γ	12dB
P	0.15w
N	5
f_1	5kHz
f_2	10kHz
f_3	15kHz
f_4	20kHz
f_5	25kHz
m	2.5km
M	90.1km
S	$1/(5\pi)^2$

In a noise-limited system environment, from Equation (5), we can get the distance threshold of any two nodes v_i and v_j connected successfully:

$$\begin{cases} d_t^2 = \dfrac{1}{\Gamma S \sum_{w \neq i,j} \left| \mathbf{r}_w - \mathbf{r}_j \right|^{-2}} \\ m \le d_t \le M \end{cases} \tag{10}$$

That is, if $|r_i - r_j| \le d_t$, nodes v_i and v_j are connected to each other.

For ease of calculation, the distance thresholds for the minimum and maximum interference cases are considered separately.

Case 1: When the connected nodes are subjected to the least interference, that is, $|r_w - r_j| = M$. According to (10), the threshold range of node connection is $2.5\text{km} \le d_M \le 90.1\text{km}$.

Case 2: When the connected nodes are subjected to the greatest interference, that is, $|r_w - r_j| = m$. According to (10), the threshold range of node connection is $2.5\text{km} \le d_M \le 6.5\text{km}$.

4 Spatial-temporal Evolution Characteristics of the Connectivity Matrix

4.1 The Characteristics of Probabilistic Connectivity Matrix

Definition 5 of Section 2.3 indicates that the state of nodes connection can be represented by a probabilistic connectivity matrix at any time slot.

According to Equation (5), since there may be multiple paths between node pairs v_i and v_j or some of the path may share common edges, the calculation of $\boldsymbol{Q}_k$ is not simple. Under the condition that there is an edge between the pair nodes v_i and v_j, the conditional probability can be used to calculate the connectivity probability of the pair nodes with distance d_{ij}. Considering different threshold ranges, the probability can be expressed as:

$$q_{ij} = \Pr\left\{ d = d_{ij} \middle| d \in d_M \quad or \quad d_m \right\} \tag{11}$$

Where d_M and d_m represent the thresholds of nodes connective distances under the minimum and maximum interference respectively.

Using the orbital data of CFSN consisting of five satellites in [6], each orbital hyper-period is approximately 6464 seconds, and each time slot $\delta_k (k = 1, 2, ..., 108)$ has a width of 60 seconds, and $T_0 = 0s$.

The probability of connection between nodes under connection threshold can be calculated by the definition of the probabilistic connectivity matrix and Equation (11).

Therefore, the probabilistic connectivity matrix $\boldsymbol{Q}_k$ of nodes at any time slot can be obtained, and $\boldsymbol{Q}_k$ is a square matrix of 5×5. According to the previous analysis, under the condition of the threshold $d_M = 90\text{km}$, the probabilistic connectivity matrix $\boldsymbol{Q}_1$ of the first time slot δ_1 is given as follows:

$$\boldsymbol{Q}_1 = \begin{bmatrix} 1 & 0 & 0.0004 & 0.0142 & 0.0138 \\ 0 & 1 & 0.0004 & 0.0127 & 0.0146 \\ 0.0004 & 0.0004 & 1 & 0.007 & 0.0114 \\ 0.0142 & 0.0127 & 0.007 & 1 & 0.0001 \\ 0.0138 & 0.0146 & 0.0114 & 0.0001 & 1 \end{bmatrix}.$$

We have illustrated how to construct probabilistic connectivity matrix for CFSN at any time slot. We now discuss the likelihood of connections of multi-hops between nodes. To do so, the following theorem is given:

Theorem 1. The probability a path of two hops exists between nodes v_i and v_j in two different consecutive time slots is:

$$1 - \prod_{l \neq i,j} \left(1 - (i,l)(l,j) \right) \tag{12}$$

where, (i,l) and (i,j) are elements corresponding to the probabilistic connectivity matrix in two consecutive time slots, (i,l) is an element of the

previous time slot, and (i, j) is the latter time slot.

Proof. Each element (i, j) of the probabilistic connectivity matrix $\boldsymbol{Q}$ is the probability the nodes v_i and v_j connects successfully.

In two consecutive time slots, since self-loops are not considered, a path of length two between nodes v_i and v_j must pass through an intermediate node $v_l(l \neq i, j)$ and can be expressed as $<(v_i(k), v_l(k)), (v_l(k+1), v_j(k+1)), (v_j(k+2), v_j(k+2))>$.

The value of two hops between nodes v_i and v_j in two consecutive time slots is the probability of the union of a set of events defined by the likelihood of paths through all intermediate nodes.

Generally, for s independent events in the probability space, according to inclusion-exclusion principle, the probability of s independent events $A_1, ...A_s$ is:

$$\Pr\left(\bigcup_{i=1}^{s}\mathbf{A}_i\right)=\sum_{k=1}^{s}(-1)^{k-1}\sum_{\substack{1\subset\{1,\ldots,s\}\\|1|=k}}\Pr(\mathbf{A}_1) \tag{13}$$

Equation (13) shows that as the number of events increases the number of factors involved also increases, making this calculation awkward for large number of events.

Equation (11) is equivalent to

$$1-\prod_{l\neq i,j}\left(1-q_{il}(k)q_{lj}(k+1)\right) \tag{14}$$

Equation (14) computes the complement of the intersection of all the complements of the elementary event, which is logically equivalent to the sum of a set of events.

Compared Equation (13) with Equation (14), Equation (14) is more efficient to compute.

Since the matrix elements represent the connection probability between nodes, Equation (12) is the probability a path of two hops exists between nodes v_i and v_j in two different consecutive time slots. QED.

In order to express the probability of multi-hops connection between nodes of CFSN in continuous time slots more intuitively, a new definition of matrix multiplication is introduced:

Definition 6. Since the probabilistic connectivity matrix is a square matrix, the product of the specified matrix $\boldsymbol{Q}_k$ and $\boldsymbol{Q}_{k+1}$ is a new matrix $\boldsymbol{H}=(h_{ij})$, where

$$h_{ij}=1-\prod_{l\neq i,j}\left(1-q_{il}(k)q_{lj}(k+1)\right)\quad(i\neq j) \tag{15}$$

where $q_{il}(k)$ and $q_{ij}(k+1)$ are the elements of matrixes $\boldsymbol{Q}_k$ and $\boldsymbol{Q}_{k+1}$ respectively.

In definition 6, the probability that a walk of length two exists between any two nodes is the product of two matrixes at consecutive time slots. Here, after L hops between nodes in consecutive time slots, the probabilistic connectivity matrix is denoted as $\boldsymbol{H}^L=\boldsymbol{Q}_k\boldsymbol{Q}_{k+1}\cdots\boldsymbol{Q}_{k+L-1}$.

For example, for CFSN, the probabilistic connectivity matrixes of five nodes in the second and third time slots are expressed as follows:

$$\boldsymbol{Q}_2=\begin{bmatrix}1 & 0 & 0.0005 & 0.0148 & 0.0126\\ 0 & 1 & 0.0004 & 0.0136 & 0.0137\\ 0.0005 & 0.0004 & 1 & 0.0075 & 0.0122\\ 0.0148 & 0.0136 & 0.0075 & 1 & 0.0001\\ 0.0126 & 0.0137 & 0.0122 & 0.0001 & 1\end{bmatrix},$$

$$\boldsymbol{Q}_3=\begin{bmatrix}1 & 0 & 0.0006 & 0.0149 & 0.0115\\ 0 & 1 & 0.0005 & 0.0143 & 0.0126\\ 0.0006 & 0.0005 & 1 & 0.0078 & 0.0126\\ 0.0149 & 0.0143 & 0.0078 & 1 & 0.0002\\ 0.0115 & 0.0126 & 0.0126 & 0.0002 & 1\end{bmatrix},$$

The element q_{ij} in the matrix $\boldsymbol{Q}$ represents the connectivity probability between nodes v_i and v_j in the second time slot. It is important to note that $q_{ij}=0$, indicating that in the second time slot, nodes v_i and v_j are not connected. Similarly, in the third time slot, nodes v_i and v_j are also not connected.

Then, after these two time slots, the probabilistic connectivity matrix after two hops is

$$\mathbf{H}^2=\mathbf{Q}_2\times\mathbf{Q}_3=\begin{bmatrix}1 & 0.0004 & 0.0003 & 0 & 0\\ 0.0004 & 1 & 0.0003 & 0 & 0\\ 0.0003 & 0.0003 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0.0004\\ 0 & 0 & 0 & 0.0004 & 1\end{bmatrix}.$$

The element h_{ij} in the matrix $\boldsymbol{H}^2$ represents the connectivity probability of between nodes v_i and v_j through two hops. It can be seen that after two hops, nodes v_i and v_j can establish a connection through all possible sequential paths

$$<(v_1(1), v_1(1)), (v_1(2), v_3(2)), (v_3(3), v_2(3))>,$$
$$<(v_1(1), v_1(1)), (v_1(2), v_4(2)), (v_4(3), v_2(3))>,$$

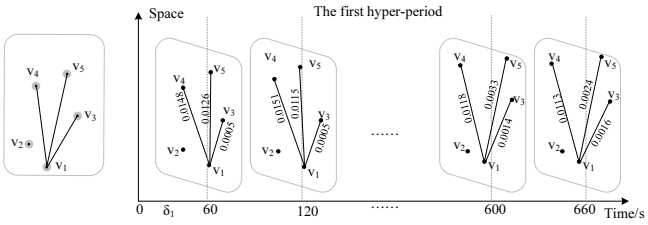
or

$$<(v_1(1), v_1(1)), (v_1(2), v_5(2)), (v_5(3), v_2(3))>.$$

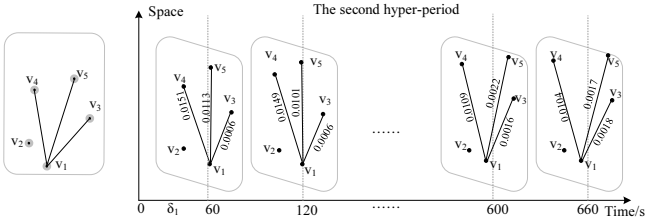
4.2 Spatial-temporal Evolution Graphs Between Nodes

Taking different thresholds range into consideration, the topology evolution diagram of CFSN is described as follows:

(1) When the distance between interference nodes is the largest, the threshold of the connective distance between nodes is $d_M = 90\text{km}$. Figure 6 shows the spatial-temporal evolution of the two orbital hyper-period of the CFSN using star plot [29], which reflects the change of the connection probability between the nodes. Where (a) shows the change of node 1 connected with other four nodes in 11 time slots of the first hyper-period, and (b) shows the change of the second hyper-period.



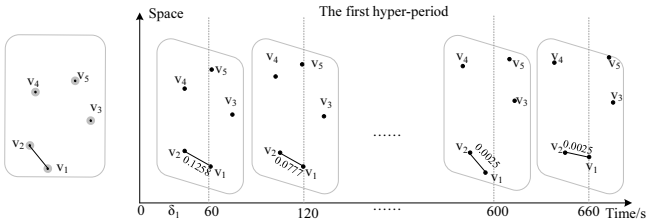
(a) The change of connection probability between nodes in the first hyper-period



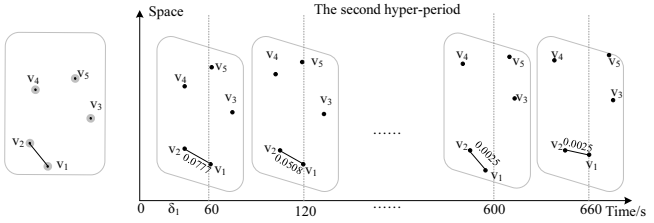
(b) The change of connection probability between nodes in the second hyper-period

Figure 6. Spatial-temporal evolution diagram in CFSN with $d_M = 90\text{km}$

(2) When the distance between interference nodes is the smallest, the threshold of connective distance between nodes is $d_M = 6.5\text{km}$. Because some distance can not satisfy the threshold, the nodes are not connected, i.e. there is no edge between nodes. In this case, Figure 7 shows the spatial-temporal evolution of the two orbital hyper-period of the CFSN. Where (a) shows the change of node 1 connected with other four nodes in 11 time slots of the first hyper-period, and (b) shows the change of the second hyper-period.



(a) The change of connection probability between nodes in the first hyper-period



(b) The change of connection probability between nodes in the second hyper-period

Figure 7. Spatial-temporal evolution diagram in CFSN with $d_M = 6.5\text{km}$

Figure 6 and Figure 7 describe the evolution of the connection and location of nodes over time. The connectivity and its probability between two nodes are determined by nodal distance. It is observed that the probability of connection between nodes has a small change in different time slots of any hyper-period, and the nodal distance changes, as shown in Figure 6(a) or Figure 6(b) and Figure 7(a) or Figure 7(b). Compared (a) with (b) in Figure 6 or Figure 7, the probability of connection between nodes also has a slight change in corresponding time slots of different hyper-periods. Due to the influence of the stochastic of node movement, there is a slight difference. Compared Figure 6 with Figure 7, it is shown that different thresholds greatly affect the connection between nodes. When the connective nodes are the most disturbed, the probability is small. And when the connective nodes are the least disturbed, the probability is large.

5 Conclusion

Due to the high-speed flight of the cluster flight spacecraft module, the uncertainty of the network topology, the connection between nodes and the path formation change with time and space. Therefore, based on twin-satellites mode and empirical statistical method, we establish the mobility model of nodes and study the nodal distance distribution and the probabilistic connectivity matrix for CFSN. We also study the characteristics of dynamic connection and path spatial-temporal evolution through the definition of sequential path. The following conclusions are obtained:

(1) Using the empirical statistics method, the distance probability density function of the CFSN is helpful to the orbit design and analysis.

(2) Through the definition of sequential path, the connection probability between nodes under the minimum and maximum interference cases and the multi-hops probabilistic connectivity matrix of nodes are studied. The results are beneficial to optimize and design the network performance of the CFSN.

(3) The results of the topology evolution graphs show that the connection probability between nodes is almost periodical, which is helpful to study the orbit control of the cluster spacecraft.

Acknowledgments

This research is a project partially supported by the National Natural Science Foundation of China (Grant No. 61362004) and Guizhou Province Education Department Projects of China (Grant No. KY [2017] 031 and KY [2020] 007).

Reference

[1] L. Mazal, P. Gurfil, Cluster Flight Algorithms for Disaggregated Satellites, *Journal of Guidance, Control and dynamic*, Vol. 36, No. 1, pp. 124-135, January, 2013.

[2] S. Nag, C. K. Gatebe, O. Weck, Observing System Simulations for Small Satellite Formations Estimating Bidirectional Reflectance, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 43, No. 2, pp. 102-118, December, 2015.

[3] J. Chu, J. Guo, E. Gill, Distributed Asynchronous Planning and Task Allocation Algorithm for Autonomous Cluster Flight of Fractionated Spacecraft, *International Journal of Space Science and Engineering*, Vol. 2, No. 2, pp. 205-223, April, 2014.

[4] I. F. Akyildiz, J. M. Jornet, S. Nie, A New CubeSat Design with Reconfigurable Multi-band Radios for Dynamic Spectrum Satellite Communication Networks, *Ad Hoc Networks*, Vol. 86, pp. 166-178, April, 2019.

[5] D. Selva, A. Golkar, O. Korobova, I. L. Cruz, P. Collopy, O. L. Weck, Distributed Earth Satellite Systems: What Is Needed to Move Forward, *Journal of Aerospace Information Systems*, Vol. 14, No. 8, pp. 412-438, August, 2017.

[6] R. I. Ansari, S. A. Hassan, S. Ali, C. Chrysostomou, M. Lestas, On the Outage Analysis of a D2D Network with Uniform Node Distribution in a Circular Region, *Physical Communication*, Vol. 25, Part. 1, pp. 277-283, December, 2017.

[7] T. Yan, S. Hu, J. Mo, Path Formation Time in the Noise-Limited Fractionated Spacecraft Network with FDMA, *International Journal of Aerospace Engineering*, Vol. 2018, Article ID 9124132, October, 2018.

[8] J. Wu, Connectivity of Mobile Linear Networks with Dynamic Node Population and Delay Constraint, *IEEE Journal on Selected Areas in Communications*, Vol. 27, No. 7, pp. 1218-1225, September, 2009.

[9] R. R. Brooks, B. Pillai, S. Racunas, S. Rai, Mobile Network Analysis Using Probabilistic Connectivity Matrices, *IEEE transactions on Systems Man and Cybernetics-Part C: Applications and Reviews*, Vol. 37, No. 4, pp. 694-702, July, 2007.

[10] N. Padmavathy, S. K. Chaturvedi, Reliability Evaluation of Mobile Ad Hoc Network: with and without Mobility Consideration, *Procedia Computer Science*, Vol. 46, pp. 1126-1139, April, 2015.

[11] S. Dasgupta, G. Mao, B. Anderson, A New Measure of Wireless Network Connectivity, *IEEE Transactions on Mobile Computing*, Vol. 14, No. 9, pp. 1765-1779, September, 2015.

[12] V. Naghshin, A. M. Rabiei, N. C. Beaulieu, B. Maham, Accurate Statistical Analysis of a Single Interference in Random Networks with Uniformly Distributed Nodes, *IEEE Communications Letters*, Vol. 18, No. 2, pp. 197-200, February, 2014.

[13] Y. Zhuang, Y. Luo, L. Cai, J. Pan, A Geometric Probability Model for Capacity Analysis and Interference Estimation in Wireless Mobile Cellular Systems, *2011 IEEE Global Telecommunications Conference*, Houston, TX, 2011, pp. 1-6.

[14] K. B. Baltzis, Analytical and Closed-form Expressions for the Distribution of Path Loss in Hexagonal Cellular Networks, *Wireless Personal Communications*, Vol. 60, No. 4, pp. 599-610, October, 2011.

[15] P. Fan, G. Li, K. Cai, K. B. Letaief, On the Geometrical Characteristic of Wireless Ad-Hoc Networks and Its Application in Network Performance Analysis, *IEEE Transactions on Wireless Communications*, Vol. 6, No. 4, pp. 1256-1265, April, 2007.

[16] S. Weber, J. G. Andrews, N. Jindal, An Overview of the Transmission Capacity of Wireless Networks, *IEEE Transactions on Communications*, Vol. 58, No. 12, pp. 3593-3604, December, 2010.

[17] J. P. Mullen, Robust Approximations to the Distribution of Link Distances in a Wireless Network Occupying a Rectangular Region, *ACM SIGMOBILE Mobile Computing and Communications Review*, Vol. 7, No. 2, pp. 80-91, April, 2003.

[18] K. B. Baltzis, A Geometric Method for Computing the Nodal Distance Distribution in Mobile Networks, *Progress in Electromagnetics Research*, Vol. 114, pp. 159-175, February, 2011.

[19] L. A. Santaló, *Integral Geometry and Geometric Probability*, Cambridge University Press, 2004.

[20] F. Tong, Y. Wan, L. Zheng, J. Pan, L. Cai, A Probabilistic Distance-Based Modeling and Analysis for Cellular Networks With Underlaying Device-to-Device Communications, *IEEE Transactions on Wireless Communications*, Vol. 16, No. 1, pp. 451-463, January, 2017.

[21] F. Tong, J. Pan, Random-to-Random Nodal Distance Distributions in Finite Wireless Networks, *IEEE Transactions on Vehicular Technology*, Vol. 66, No. 11, pp. 10070-10083,

November, 2017.

[22] M. Ahmadi, F. Tong, L. Zheng, J. Pan, Performance Analysis for Two-tier Cellular Systems Based on Probabilistic Distance Models, *2015 IEEE Conference on Computer Communications (INFOCOM)*, Kowloon, Hong Kong, 2015, pp. 352-360.

[23] S. J. Nawaz, S. Wyne, K. B. Baltzis, S. M. Gulfam, K. Cumanan, A Tunable 3-D statistical Channel Model for Spatio-temporal Characteristics of Wireless Communication Networks, *Transactions on Emerging Telecommunications Technologies*, Vol. 28, No. 12, pp. e3213, December, 2017.

[24] P. Liu, X. Chen, Y. Zhao, Safe Deployment of Cluster-flying Nano-satellites Using Relative E/I Vector Separation, *Advances in Space Research*, Vol. 64, No. 4, pp. 964-981, August 2019.

[25] J. Huang, Y. Su, L. Huang, W. Liu, F. Wang, An Optimized Snapshot Division Strategy for Satellite Network in GNSS, *IEEE Communications Letters*, Vol. 20, No. 12, pp. 2406-2409, December, 2016.

[26] A. Kandhalu, R. Rajkumar, QoS-Based Resource Allocation for Next-Generation Spacecraft Networks, *2012 IEEE 33rd Real-Time Systems Symposium*, San Juan, Puerto Rico, 2012, pp. 163-172.

[27] A. DasGupta, *Probability for Statistics and Machine Learning*, Springer, 2011.

[28] A. W. Vaart, *Asymptotic Statistics*, Cambridge University Press, 1998.

[29] J. Park, A. Barabási, Distribution of Node Characteristics in Complex Networks, *Proceedings of the National Academy of Sciences*, Vol. 104, No. 46, pp. 17916-17920, November, 2007.

Biographies

Yanfeng Shi received the B.S. degree in Communication Engineering from Yanshan University in 2005, and the M.S. degree in Signal Processing from Guizhou Normal University in 2018. He has been a lecturer at Guizhou Normal University since 2005. His main research interests include wireless communication and radar signal processing.

Shengbo Hu received his Radio Technology B.Sc. degree in 1985 from Southeast University, received his M.Sc. degree in 1992 from Electronic Engineering Institute of PLA, received his Ph.D. degree in 2006 from Chongqing University, now he is a professor in Guizhou Normal University. His main research interests include intelligent sensing and wireless communication technologies.

Jinrong Mo received her B.Sc. degree and M.Sc. degree in 2013 and 2016 from Guizhou Normal University, and now is studying for her Ph.D. degree in Guizhou Normal University. Her main research interests is applied mathematics.

Xiaowei Song received her M.Sc. degree in 2015 Guizhou Normal University Since 2015, she has been with the Institute of Intelligent Information Processing, Guizhou Normal University, Guiyang, China. Her main research interests include microwave filter and antenna.

Tingting Yan received her B.Sc. degree and M.Sc. degree in 2014 and 2017 from Guizhou Normal University, and now is studying for her Ph.D. degree in Guizhou Normal University. Her main research interests is applied mathematics.