AIAA SciTech Forum
8–12 January 2018, Kissimmee, Florida
2018 AIAA Information Systems-AIAA Infotech @ Aerospace

10.2514/6.2018-1802

Check for updates

A Probabilistic Network Formulation for Satellite Swarm Communications

V. Hunter Adams* and Mason Peck†

This paper presents a scalable method for routing over a dynamic network of spacecraft in orbit. The routing policy is constructed as a series of optimal stopping problems. A spacecraft carrying data destined for a separate node in the network will encounter intermediate spacecraft as it travels along its trajectory. At each of these encounters, it must decide whether to relinquish the data to the intermediate spacecraft or to retain it. This paper presents a method for making that decision in a manner that minimizes the expected time to destination.

Nomenclature

θ	Polar angle from perihelion (rad)
x	State
γ_i	Spacecraft i
t	Time elapsed (sec)
u	Control input
c_T	Terminal cost
T	Time to destination from origin node (sec)
ϕ	Stopping cost
τ	Time to contact with destination node (sec)
V_i^*	Optimal value function
g	Control policy
G	Gravitational constant $\frac{\text{m}^3}{\text{kgsec}^2}$
M	Mass of Earth (kg)
m	Mass of spacecraft (kg)
$p(x)$	Probability density function of variable x
e	Eccentricity
a	Semimajor axis (km)

*PhD Candidate, Mechanical and Aerospace Engineering, Cornell University, 452 Upson Hall, Ithaca NY 14850, AIAA Student Member

†Assistant Professor, Mechanical and Aerospace Engineering, Cornell University, 455 Upson Hall, Ithaca NY 14850, AIAA Full Member

I. Introduction

With the advent of chip-scale spacecraft, Earth-orbiting satellites can now be deployed in unprecedented numbers. The KickSat 1 spacecraft from Cornell University was a 3U CubeSat that carried 104 chip-scale spacecraft into low-Earth orbit in 2014, and its follow-on, Kicksat 2, is expected to launch in 2018 [1]. The maturity of this technology motivates new questions that must be answered for constellations of thousands of spacecraft. Many of the questions this opportunity raises are new to the field of astronautical engineering because, until Kicksat-1, an on-orbit network of this many satellites was inconceivable.

For such constellations, it is likely impractical for the ground to communicate with each node as if it were a traditional satellite. Furthermore, the satellites in such an architecture drift into and out of communication range with one another as they travel along their respective trajectories, introducing subtlety in how cross-linked communications might work. Methods must be developed for designing networks that maintain properties that are important to the user (such as connectedness, ground coverage, and broadcast time). Furthermore, in networks of such size, a certain number of constituent nodes should be expected to fail. The network must be robust to individual failures of nodes, and capable of rerouting around these failures. While many of these questions have been considered in other fields [2][3], there are some decidedly unique features to an on-orbit network. Some of these features include nodes that have dynamics, and network-wide periodicity among edge connections.

This analysis presents a method for routing over these dynamic and uncertain networks. Routing data from any origin node in the network to any destination node in the network can be viewed as a series of decisions. The origin node will encounter a number of intermediate spacecraft as it moves along its trajectory. At each of these encounters, it must decide whether the expected time to destination is minimized by relinquishing or retaining the data that it carries. If the expected time to destination is reduced by passing the data to the intermediate spacecraft, then it reqlinquishes. If not, it retains the data. This decision is repeated for every encounter with an intermediate spacecraft until the data reaches its destination.

This solution can be formulated a series of optimal stopping problems, derived via the standard dynamic programming equations. This paper discusses the solution for a representative example network composed of spacecraft on nested orbits of slightly different altitudes, and then discusses the necessary generalizations for networks of arbitrary, potentially time-varying orbits. The nested orbits example is treated in depth because it leads to a simple formulation of the dynamic programming equations, and it is a realistic configuration for a group of spacecraft randomly deployed from a mothership (like chip-satellites being deployed from a KickSat).

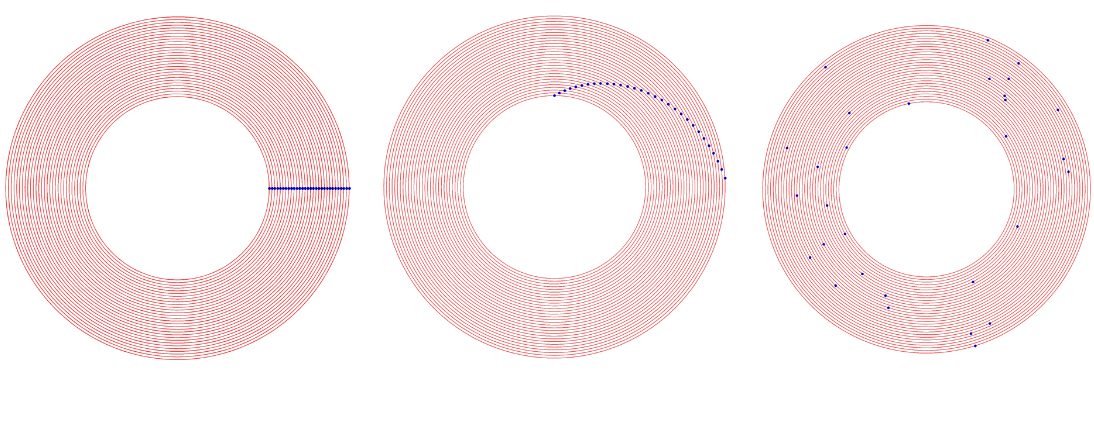

Figure 1: Dispersion of spacecraft with slightly different semimajor axes

II. Problem Statement

Spacecraft that are simultaneously deployed with randomized initial velocities from a common mothership will enter nested orbits of slightly varying eccentricities and semimajor axes. This variation in eccentricity and semimajor axis, along with perturbations from atmospheric drag, solar pressure, and other higher-order effects, leads to rapid dispersion of the spacecraft in the along-track direction and comparatively little dispersion in altitude, as shown in Fig. 1. Thus, it is assumed that any two spacecraft that are at the same angular position in their respective orbits will be near enough to communicate with one another, while those that are not in the same angular position cannot communicate. The probability density function for the position of any particular spacecraft is a function only of its eccentricity and semimajor axis (see appendices). After a sufficient amount of time has passed, the positions of all spacecraft are well approximated by independent distributions (see Fig. 2). The objective is to develop a policy for routing data from an arbitrary origin node to a single destination node (i.e. a ground-station) along a time-optimal path. Each spacecraft is assumed to be capable of measuring time elapsed between encounters with the destination node, from which it can estimate time to contact with the destination.

III. Dynamic Programming Equations

Supposing that the network is composed of N spacecraft with indices $\gamma_0, \gamma_1, \ldots, \gamma_N$, then the state of the system is fully specified by the index of the spacecraft that presently holds the data being routed to destination γ_d and the time elapsed. Each spacecraft maintains an estimate of its own time to contact with the destination node, which is a function of the total time elapsed $\tau_{\gamma_i \rightarrow \gamma_d}(t)$. For simple orbital configurations, this function can be derived in closed form. For more general cases, the spacecraft may form an estimate for the expected time between contacts with the destination online.

The routing process begins with the data at γ_0. At each encounter, there are two possible control inputs that the spacecraft carrying the data may exert on the system. It may either pass the data to the encountered spacecraft, or it may keep the data for itself. The state definition, feasible control inputs, and state update equations are given by eqns. 1-6.

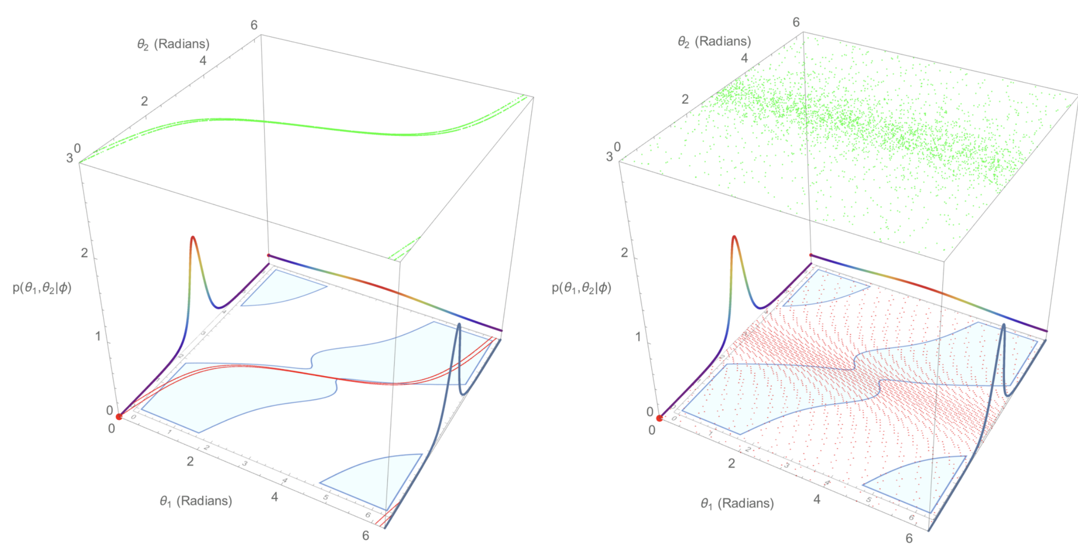

Figure 2: Path of two spacecraft with 0.1-percent different semimajor axes through phase space after two orbits (left) and many orbits (right). Random samples of positions on top, equal time-spaced samples on bottom, blue shaded region shows region of communication.

State :

$$x_i = \begin{bmatrix} \gamma_i \in [\gamma_0, \ldots, \gamma_N] \\ t \end{bmatrix} \tag{1}$$

Initial Condition :

$$x_0 = \begin{bmatrix} \gamma_0 \\ t_0 \end{bmatrix} \tag{2}$$

Control Input :

$$u_i|(x_i \in \gamma_{i \neq d}, t) \in \begin{cases} 1 & \text{Relinquish data to } \gamma_j \\ 0 & \text{Retain data} \end{cases} \tag{3}$$

$$u_i|(x_i \in \gamma_d, t) = \begin{cases} 0 & \text{Retain data} \end{cases} \tag{4}$$

$$\tag{5}$$

State Update Equation :

$$x_{i+1} = f(x_i, u_i) = \begin{bmatrix} \begin{cases} \gamma_0 & x_i = \gamma_0, u_i = 0 \\ \gamma_i & x_i = \gamma_0, u_i = 1 \\ \gamma_d & x_i = \gamma_d \end{cases} \\ t + t_{\gamma_0 \to \gamma_j} \end{bmatrix} \tag{6}$$

From any origin spacecraft, there exists a path (or multiple paths) via intermediate spacecraft that minimizes the amount of time that it takes for the data to get from the origin to the destination ground station. The routing policy is optimal if it is guaranteed to route along one of those paths regardless of the initial configuration of the spacecraft. The optimal time to destination from spacecraft γ_i is given by the variable $\tau^*_{\gamma_i \to \gamma_d}$. This quantity can be used to specify the costs in the dynamic programming equations. The terminal cost is the cost incurred if the origin spacecraft never relinquishes the data to any intermediate spacecraft. The stopping cost is the cost incurred by relinquishing to an intermediate spacecraft. The stage cost is equal to the stopping cost if the spacecraft relinquishes the data, and equal to zero if the spacecraft retains the data.

Terminal Cost :

$$c_T(x_T) = \begin{cases} T - t_0 & x_T = \gamma_0 \\ 0 & x_T \neq \gamma_0 \end{cases} \tag{7}$$

Stopping Cost :

$$\phi(t) = \tau^*_{\gamma_i \to \gamma_d}(t) \tag{8}$$

Stage Cost :

$$c_i(x_i, u_i) = \begin{cases} 0 & u_i = 0 \\ \tau^*_{\gamma_i \to \gamma_d}(t) & u_i = 1 \end{cases} \tag{9}$$

From these costs, one can write down the general expressions for the terminal value function, optimal value functions, and optimal policy in terms of the optimal value function. These general expressions hold for arbitrary network configurations. The explicit expression for the optimal value function will vary depending on the particular network under consideration.

Terminal Value Function :

$$V_T^*(x_T) = \begin{cases} T - t_0 & x_T = \gamma_0 \\ 0 & x_T \neq \gamma_0 \end{cases} \tag{10}$$

Optimal Value Functions :

$$\begin{aligned} V_i^*(\gamma_d) &= \min_{u \in \{0\}} E\left[c_i(x_i = \gamma_d, u_i = 0) + V_{i+1}^*(x_{i+1} = \gamma_d)\right] \\ &= 0 + 0 = 0 \end{aligned} \tag{11}$$

$$\begin{aligned} V_i^*(\gamma_i) &= \min_{u \in \{0,1\}} E\left[c_i(x_i = \gamma_i, u_t = 0, 1) + V_{i+1}^*(x_{i+1})\right] \\ &= \min\left[\tau^*_{\gamma_i \to \gamma_d}(t), \quad V_{i+1}^*(x_{i+1})\right] \end{aligned} \tag{12}$$

Optimal Policy :

$$g_i^* = \begin{cases} 1 & \tau^*_{\gamma_i \to \gamma_d}(t) < V_{i+1}^*(x_{i+1}) \\ 0 & \tau^*_{\gamma_i \to \gamma_d}(t) \geq V_{i+1}^*(x_{i+1}) \end{cases} \tag{13}$$

If the optimal time-to-go for the intermediate spacecraft is greater than or equal to optimal value function at the $(i+1)^{th}$ step, then the spacecraft retains the data. If the optimal time-to-go for the intermediate spacecraft is less than the optimal value function at the $(i+1)^{th}$ step, then the spacecraft relinquishes the data.

IV. Optimal Value Function for Nested Orbits

For the particular case where any two spacecraft that are at the same angular position are capable of communicating, then it can be shown that the fastest route to the destination is never via an intermediate spacecraft with a longer time to destination than the origin spacecraft, and that there is never cost incurred by relinquishing data to a faster spacecraft (even if the origin spacecraft encounters an even-faster spacecraft after handoff has occurred). Thus, for spacecraft on nested orbits, the optimal value function is given by eqn. 14.

$$V_{i+1}^*(x_{i+1}) = \tau_{\gamma_0 \to \gamma_d}(t) \tag{14}$$

IV.A. Proof

Suppose that spacecraft A encounters spacecraft B, which has a shorter time to destination D than A. Then, assume that spacecraft A encounters spacecraft C before B has made contact, which has a shorter time to destination D than spacecraft B and A. It can be shown that spacecraft B encounters spacecraft C before spacecraft C reaches the destination D. All possible cases for this series of interactions are enumerated below. For each case, it will be shown that that B must encounter C before C encounters D.

1. A overtakes B then A overtakes C
2. A overtakes B then C overtakes A
3. B overtakes A then A overtakes C
4. B overtakes A then C overtakes A

IV.A.1. Case 1

Let the present positions of spacecraft A, B, and C be given by θ_A, θ_B, θ_C. Let the position of the destination be given by θ_D When spacecraft A overtakes B, let $\theta_A = \theta_B$. Because all spacecraft are orbiting in the same direction, the angular position of each will increase at a constant rate. For this particular case, the rate at which the destination spacecraft's angular position increases is greater than that of spacecraft B, which is greater than that of spacecraft A. At the moment of AB alignment, spacecraft C is at a position $\theta_C > \theta_A = \theta_B$, since it has yet to be overtaken by A.

$$\theta_D < (\theta_B = \theta_A) < \theta_C \tag{15}$$
$$\theta_D < \theta_B < \theta_A < \theta_C \tag{16}$$
$$\theta_D < \theta_B < (\theta_A = \theta_C) \tag{17}$$
$$\theta_D < \theta_B < \theta_C < \theta_A \tag{18}$$
$$(\theta_D = \theta_C) < \theta_B < \theta_A \tag{19}$$

In order for the system to go from states 18-19, it must pass through the intermediate state given by eqn. 20. This proves that, for the case where A is overtakes B with time to destination less than A, and then overtakes C with time to destination less than A and B, then B must overtake C before being overtaken by D, allowing the two to exchange information. This indicates that there is no time penalty incurred for passing data to B rather than waiting for C.

$$\theta_D < (\theta_B = \theta_C) < \theta_A \tag{20}$$

IV.A.2. Case 2

Consider the angular positions of the spacecraft and destination at the moment of AB alignment, and then step through each configuration until CD alignment (as in case 1).

$$\theta_C < \theta_D < (\theta_B = \theta_A) \tag{21}$$
$$\theta_C < \theta_D < \theta_B < \theta_A \tag{22}$$
$$\theta_D < \theta_B < (\theta_A = \theta_C) \tag{23}$$

In order to get from states 22-23, the system must have passed through intermediate state given by eqn. 24. This proves that, for the case where A overtakes B with time to destination less than A, and then is overtaken by C with time to destination less than A and B, then C must overtake B before overtaking D, allowing the two to exchange information. This indicates that there is no time penalty incurred for passing data to B rather than waiting for C.

$$\theta_D < (\theta_B = \theta_C) < \theta_A \tag{24}$$

IV.A.3. Case 3

Consider the angular positions of the spacecraft and destination at the moment of AB alignment, and then step through each configuration until CD alignment (as in case 1).

$$\theta_D < (\theta_B = \theta_A) < \theta_C \tag{25}$$
$$\theta_D < \theta_A < \theta_B < \theta_C \tag{26}$$
$$\theta_D < (\theta_A = \theta_C) < \theta_B \tag{27}$$

In order to get from states 26-27, the system must have passed through intermediate state given by eqn. 28. This proves that, for the case where B overtakes A with time to destination less than A, and then A overtakes C with time to destination less than A and B, then B must overtake C before C is overtaken by D, allowing the two to exchange information. This indicates that there is no time penalty incurred for passing data to B rather than waiting for C.

$$\theta_D < \theta_A < (\theta_B = \theta_C) \tag{28}$$

IV.A.4. Case 4

Consider the angular positions of the spacecraft and destination at the moment of AB alignment, and then step through each configuration until CD alignment (as in case 1).

$$\theta_C < (\theta_B = \theta_A) < \theta_D \tag{29}$$

$$(\theta_C = \theta_A) < \theta_B < \theta_D \tag{30}$$

$$\theta_A < \theta_C < \theta_B < \theta_D \tag{31}$$

$$\theta_A < \theta_B < \left(\tilde{\theta_D} = \tilde{\theta_C}\right) \tag{32}$$

In order to get from states 31-32, the system must have passed through intermediate state given by eqn. 33. This proves that, for the case where A is overtaken by B with time to destination less than A, and then overtaken by C with time to destination less than A and B, then C must overtake B before overtaking D, allowing the two to exchange information. This indicates that there is no time penalty incurred for passing data to B rather than waiting for C.

$$\theta_A < (\theta_C = \theta_B) < \theta_D \tag{33}$$

V. Nested Orbits Results

The above four cases prove that, if $\tau_{\gamma_j \to \gamma_d} < \tau_{\gamma_i \to \gamma_d}$, then $\tau^*_{\gamma_j \to \gamma_d} < \tau^*_{\gamma_i \to \gamma_d}$ $\forall i, j, d$. Thus, the total time to the destination is minimized by optimizing over the immediate timestep. For the nested orbits, the myopic policy is optimal. The optimal policy may be rewritten as shown in eqn. 34.

$$g_i^* = \begin{cases} 1 & \tau_{\gamma_i \to \gamma_d}(t) < \tau_{\gamma_0 \to \gamma_d}(t) \\ 0 & \tau_{\gamma_i \to \gamma_d}(t) \geq \tau_{\gamma_0 \to \gamma_d}(t) \end{cases} \tag{34}$$

Note that the policy given by eqn. 34 is time-optimal in the case that spacecraft at the same angular position on separate orbits are close enough to communicate. A suboptimal routing policy for the case where this assumption does not hold can be derived by generalizing the above derivation. The above policy is scalable in that each spacecraft must only maintain knowledge of its own time to contact with the destination ground station. Global knowledge is not required for time-optimal routing over these configurations of orbits. For the limiting case where each spacecraft travels on a circular orbit, $\tau_{\gamma_i \to \gamma_d}(t)$ is given by eqns. 35-37. Eqn. 35 gives the time to next contact with the destination node as a function of total time elapsed, eqn. 36 gives the time between contacts, and eqn. 37 gives the time to first alignment, given initial phase offsets θ_i and θ_d.

$$\tau_{\gamma_i \to \gamma_d} = \text{Mod}\left[t_{p,i \to d} - (t - t_{1,i \to d}), t_{p,i \to d}\right] \tag{35}$$

$$t_{p,i \to d} = 2\pi \left[\frac{G\left(M+m\right)}{a_i^3}\right]^{-\frac{1}{2}} \left[1 - \left(\frac{a_i}{a_d}\right)^{\frac{3}{2}}\right]^{-1} \tag{36}$$

$$t_{1,i \to d} = \left(2\pi - \text{Mod}\left((\theta_i - \theta_d) + 2\pi, 2\pi\right]\right) \left[\frac{G\left(M+m\right)}{a_i^3}\right]^{-\frac{1}{2}} \left[1 - \left(\frac{a_i}{a_d}\right)^{\frac{3}{2}}\right]^{-1} \tag{37}$$

VI. Future Work: Methods for Generalization

In the general case, it is not necessarily true that all spacecraft are capable of communication with all other spacecraft and with the destination. Furthermore, it is not neccesarily the case that all spacecraft are orbiting in the same direction. In this situation, the myopic policy is no longer optimal. The shortest route to the destination may route through an intermediate spacecraft with a longer (or infinite) time to destination, since the orbits are no longer nested. With global knowledge of the network, one could find an optimal route from an arbitrary origin to an arbitrary destination. However, for computationally limited spacecraft like the chip-satellites, this method is entirely unscalable. Instead, one can derive a suboptimal routing policy based only on local knowledge that scales to networks of arbitrary numbers of nodes.

As in the nested orbits case, the routing problem can be setup as a series of decisions. The state, initial conditions, control inputs, and state update equations the same as before, augmented with t_h, the total time that the spacecraft with the data has been holding the data.

$$\text{State}:$$

$$x_i = \begin{bmatrix} \gamma_i \in [\gamma_0, \ldots, \gamma_N] \\ t \\ t_h \end{bmatrix} \tag{38}$$

$$\text{Initial Condition}:$$

$$x_0 = \begin{bmatrix} \gamma_0 \\ t_0 \\ 0 \end{bmatrix} \tag{39}$$

$$\text{Control Input}:$$

$$u_i|(x_i \in \gamma_{i \neq d}, t, t_h) \in \begin{cases} 1 & \text{Relinquish data to } \gamma_j \\ 0 & \text{Retain data} \end{cases} \tag{40}$$

$$u_i|(x_i \in \gamma_d, t, t_h) = \begin{cases} 0 & \text{Retain data} \end{cases} \tag{41}$$

$$\tag{42}$$

$$\text{State Update Equation}:$$

$$x_{i+1} = f(x_i, u_i) = \begin{bmatrix} \begin{cases} \gamma_0 & x_i = \gamma_0, u_i = 0 \\ \gamma_i & x_i = \gamma_0, u_i = 1 \\ \gamma_d & x_i = \gamma_d \end{cases} \\ t + t_{\gamma_0 \to \gamma_j} \\ t_h + t_{\gamma_0 \to \gamma_j} \end{bmatrix} \tag{43}$$

In the nested case, where every spacecraft will eventually come into communication range with every other spacecraft and with the destination, the cost is simply the time to destination. This is no longer true in the general case. Instead, the cost is a function not only of the time to destination (which may be infinite, and is itself a function of the total time that the spacecraft has held the data), but also of the number of unique encounters that a spacecraft is expected to have N_{γ_i}, and the expected number of additional spacecraft that the origin node will encounter N_r. In the nested case, T represented the total time to encounter of the origin spacecraft with the destination. In this case T is a tunable parameter that represents the propensity of a spacecraft to covet or to share a particular data packet.

$$\text{Terminal Cost}:$$

$$c_T(x_T) = \begin{cases} \min\left(C, \tau_{\gamma_0 \to \gamma_d}(t_0)\right) & x_T = \gamma_0 \\ 0 & x_T \neq \gamma_0 \end{cases} \tag{44}$$

$$\text{Stopping Cost}:$$

$$\phi(\gamma_i, t, t_h) = f\left(\tau_{\gamma_0 \to \gamma_d}(t), \tau_{\gamma_i \to \gamma_d}(t), N_{\gamma_i}, N_r(t_h)\right) \tag{45}$$

$$\text{Stage Cost}:$$

$$c_i(x_i, u_i) = \begin{cases} 0 & u_i = 0 \\ \phi(\gamma_i, t, t_h) & u_i = 1 \end{cases} \tag{46}$$

In the event that the origin node will never come into contact with the destination, and that none of its neighbors come into contact with the destination either, then the router should preferentially choose spacecraft with more unique connections than spacecraft with less unique connections. In doing so, it maximizes the liklihood of finding a spacecraft with a connection to the destination. The terminal value function, optimal value functions, and optimal stopping rule can be expressed in terms of these generalized expressions.

$$\text{Terminal Value Function :}$$
$$V_T^*(x_T) = \begin{cases} \min\left(C, \tau_{\gamma_0 \to \gamma_d}(t_0)\right) & x_T = \gamma_0 \\ 0 & x_T \neq \gamma_0 \end{cases} \tag{47}$$

$$\text{Optimal Value Functions :}$$
$$\begin{aligned} V_i^*(\gamma_d) &= \min_{u \in \{0\}} E\left[c_i(x_i = \gamma_d, u_i = 0) + V_{i+1}^*(x_{i+1} = \gamma_d)\right] \\ &= 0 + 0 = 0 \end{aligned} \tag{48}$$
$$\begin{aligned} V_i^*(\gamma_i) &= \min_{u \in \{0,1\}} E\left[c_i(x_i = \gamma_i, u_t = 0, 1) + V_{i+1}^*(x_{i+1})\right] \\ &= \min\left[\phi(\gamma_i, t, t_h), \quad V_{i+1}^*(x_{i+1})\right] \end{aligned} \tag{49}$$

$$\text{Optimal Policy :}$$
$$g_i^* = \begin{cases} 1 & \phi(\gamma_i, t, t_h) < V_{i+1}^*(x_{i+1}) \\ 0 & \phi(\gamma_i, t, t_h) \geq V_{i+1}^*(x_{i+1}) \end{cases} \tag{50}$$

Future work will involve developing a closed-form expression for $\phi(\gamma_i, t, t_h)$ and, as a consequence, $V_{i+1}^*(x_{i+1})$.

Appendix

The appendices provide additional details on the dynamic assumptions of the system and prove some of the properties of this system that are exploited in the network analysis.

Swept Area from Angular Position

Consider an elliptical orbit with semimajor axis a and semiminor axis b, as shown in Fig. 3. Earth sits at

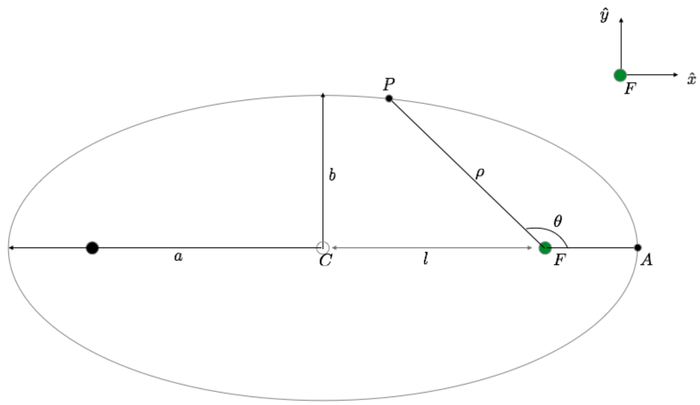

Figure 3: Swept area geometry, origin at focus

focal point F, with perhelion at A. As the spacecraft traverses the orbit, it sweeps out the area AFP. The position of the spacecraft is specified by its distance from the Earth (ρ) and the angle from perihelion (θ). Some of the geometric relationships among the above quantities are given by (not derived here):

$$e = \text{eccentricity} = \sqrt{1 - \frac{b^2}{a^2}} \tag{51}$$

$$l = \text{linear eccentricity} = ae \tag{52}$$

Additionally, as shown above:

$$\rho(\theta) = \frac{a(1-e^2)}{1 + e\cos\theta} \tag{53}$$

$$\tag{54}$$

The location of the spacecraft (in Cartesian coordinates, with the origin on the Earth) can be parametrized as:

$$^F\mathbf{P} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} \tag{55}$$

$$= \begin{bmatrix} \frac{a(1-e^2)}{1+e\cos\theta}\cos\theta \\ \frac{a(1-e^2)}{1+e\cos\theta}\sin\theta \end{bmatrix} \tag{56}$$

Consider scaling the y-coordinate such that the trajectory is a circle of radius a (the semimajor axis), is shown in Fig. 4. This can be done by scaling the y-axis by $\frac{a}{b} = \frac{1}{\sqrt{1-e^2}}$. The point P gets mapped to Q, and the new position can be parametrized as:

$$^F\mathbf{Q} = \begin{bmatrix} r\cos\theta \\ \frac{1}{\sqrt{1-e^2}} r\sin\theta \end{bmatrix} \tag{57}$$

$$= \frac{a}{1+e\cos\theta}\begin{bmatrix} (1-e^2)\cos\theta \\ \sqrt{(1-e^2)}\sin\theta \end{bmatrix} \tag{58}$$

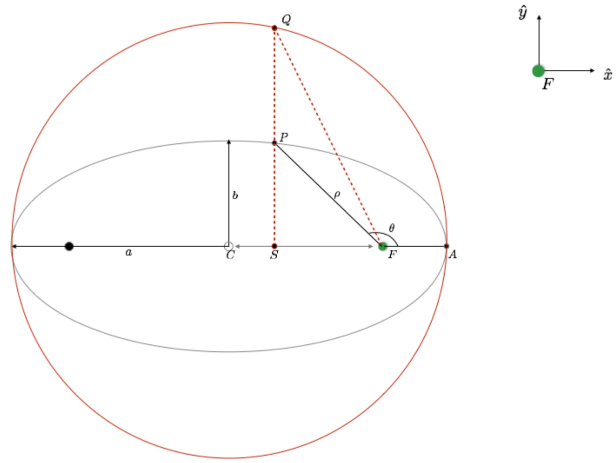

Figure 4: Swept area geometry, scaled to circle

Move the origin to the center of the circle, as shown in Fig. 5. The position of Q is now given by:

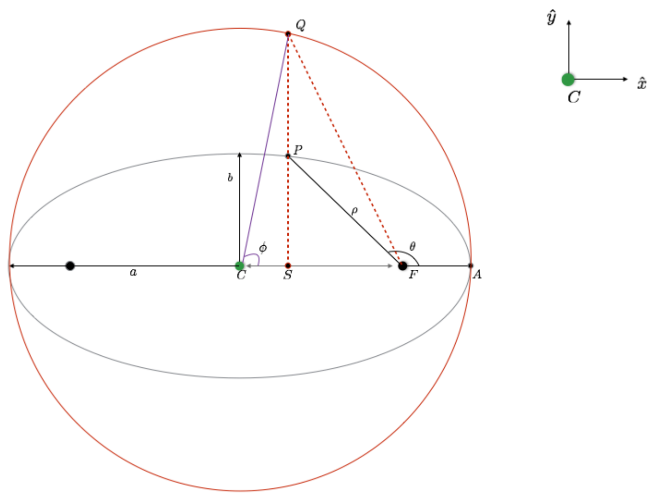

Figure 5: Swept area geometry, scaled to circle, origin at center

$$
{}^{C}\mathbf{Q} = \begin{bmatrix} r\cos\theta + ae \\ \frac{1}{\sqrt{1-e^2}} r\sin\theta \end{bmatrix} \tag{59}
$$

$$
= \frac{a}{1+e\cos\theta} \begin{bmatrix} (1-e^2)\cos\theta + e(1+e\cos\theta) \\ \sqrt{(1-e^2)}\sin\theta \end{bmatrix} \tag{60}
$$

The angle ϕ is the eccentric anomaly. From the geometry shown above, we have:

$$
\tan\phi = \frac{\sqrt{1-e^2}\sin\theta}{(1-e^2)\cos\theta + e + e^2\cos\theta} \tag{61}
$$

Using a Wiererstrauss substitution, this can be simplified to:

$$
\tan\frac{\phi}{2} = \sqrt{\frac{1-e}{1+e}}\tan\frac{\theta}{2} \tag{62}
$$

Thus:

$$\phi = \text{atan2}\left(\sqrt{1-e^2}\sin\theta, \quad (1-e^2)\cos\theta + e + e^2\cos\theta\right) \tag{63}$$

$$= 2\tan^{-1}\left[\sqrt{\frac{1-e}{1+e}}\tan\frac{\theta}{2}\right] \tag{64}$$

And the reverse is of course also true:

$$\theta = 2\tan^{-1}\left[\sqrt{\frac{1+e}{1-e}}\tan\frac{\phi}{2}\right] \tag{65}$$

With this angle ϕ, it is easy to calculate the area of sector ACQ. This sector includes the region in which we're interested.

$$A_{ACQ} = \frac{1}{2}a^2\phi \tag{66}$$

To obtain the area of the region of interest, subtract and scale back the above result. To begin, subtract off the area of triangle FCQ, leaving the region AFQ remaining:

$$A_{AFQ} = A_{ACQ} - A_{FCQ} \tag{67}$$

$$= \frac{1}{2}a^2\phi - \frac{1}{2}a^2 e\sin\phi \tag{68}$$

$$= \frac{1}{2}a^2\left[\phi - e\sin\phi\right] \tag{69}$$

This is still the area for a circle. To get back to the area for the ellipse, undo the initial scaling by multiplying by $\frac{b}{a}$ (since scaling in the y-direction scales the area by the same factor):

$$A_{AFP} = \frac{1}{2}ab\left[\phi - e\sin\phi\right] \tag{70}$$

Probability Density from Swept Area

Because the spacecraft sweeps equal areas in equal times (shown above), the liklihood of finding the spacecraft in this angular region is given by:

$$P(AFP) = \frac{A_{AFP}}{A_{total}} \tag{71}$$

$$= \frac{\frac{1}{2}ab\left[\phi - e\sin\phi\right]}{\pi ab} \tag{72}$$

$$\tag{73}$$

This leads directly to the probability distribution function in ϕ.

$$P(\phi) = \frac{\phi - e\sin\phi}{2\pi} \tag{74}$$

Substituting the expression for θ yields the probability distribution in θ:

$$P(\theta) = \frac{2\tan^{-1}\left(\sqrt{\frac{1-e}{e+1}}\tan\left(\frac{\theta}{2}\right)\right) - e\sin\left(2\tan^{-1}\left(\sqrt{\frac{1-e}{e+1}}\tan\left(\frac{\theta}{2}\right)\right)\right)}{2\pi} \tag{75}$$

The probability density function is obtained by taking the derivative of the distribution function with respect to ϕ (or θ in the case of $P(\theta)$):

$$p(\phi) = \frac{\partial P(\phi)}{\partial\phi} = \frac{1 - e\cos(\phi)}{2\pi} \tag{76}$$

$$p(\theta) = \frac{\partial P(\theta)}{\partial\theta} = \frac{(1-e)^{3/2}}{2\pi\left(\frac{1}{e+1}\right)^{3/2}(e\cos(\theta)+1)^2} \tag{77}$$

Conditional Probability Density

Consider two orbits, Orbit A and Orbit B, each containing a single spacecraft. Before measuring the system, the prior distributions for the positions of spacecraft A and B are given by Eqn. 103 (with the proper eccentricity substituted for each orbit). These are the prior distributions for A and B. However, upon sampling the position of A, the probability density function for A collapses to a delta function at the angular position for A. The probability density function for B collapses to a train of delta functions, the location for each corresponding to one of the possible positions for B.

Prior

$$p(\theta_A) = \frac{(1-e_A)^{3/2}}{2\pi\left(\frac{1}{e_A+1}\right)^{3/2}(e_A\cos(\theta)+1)^2} \tag{78}$$

$$p(\theta_B) = \frac{(1-e_B)^{3/2}}{2\pi\left(\frac{1}{e_B+1}\right)^{3/2}(e_B\cos(\theta)+1)^2} \tag{79}$$

Posterior

$$p_A(\theta) = \delta(\theta-\theta_A) \tag{80}$$

$$p_B(\theta) = \sum_{i=0}^{N}\frac{1}{N}\cdot\delta\left(\theta-\theta_i\right)+\lambda \tag{81}$$

N, in this case, it the total number of possible locations for B. This is obtained by taking the ratio of the orbital periods of A and B. Represented as a ratio of integers, the denominator of that fraction represents the number of possible locations for B given a particular location for A. If the orbital period for B were half that of A, for example, then there would be two possible locations for B for each location of A. In some cases, the ratio of orbital periods is an irrational (or close to irrational) number. In such situations, the position of A provides no information on the position of B, and the posterior distribution for B is the same as the prior distribution. λ is the phase offset, which depends on the initial configuration of A and B.

Radial Distance Distribution

From the above probability distribution for the polar position of the spacecraft, one can find the probability distribution for distance from Earth. It has been shown that the distance from the Earth is given by the below equation:

$$\rho(\theta) = \frac{a(1-e^2)}{1+e\cos\theta} = g(\theta) \tag{82}$$

In the range $0 \le \theta \le 2\pi$, the above function for ρ is not monotonic. Thus, one can solve for the inverse function $g^{-1}(\rho)$:

$$g^{-1}(\rho) = \pm\cos^{-1}\left(\frac{-ae^2+a-\rho}{e\rho}\right) \tag{83}$$

The distribution function for ρ is given by:

$$p(\rho) = \sum_{k=1}^{n(\rho)}\left|\frac{d}{d\rho}\left(g_k^{-1}(\rho)\right)\right|\cdot p_\theta(g_k^{-1}(\rho)) \tag{84}$$

$$= \left|\frac{a-ae^2}{e\rho^2\sqrt{1-\frac{(a(e^2-1)+\rho)^2}{e^2\rho^2}}}\right|\left(-\frac{\sqrt{\frac{2}{e+1}-1}\rho^2}{2\pi a^2(e-1)}\right)+\left|\frac{a\left(e^2-1\right)}{e\rho^2\sqrt{1-\frac{(a(e^2-1)+\rho)^2}{e^2\rho^2}}}\right|\left(-\frac{\sqrt{\frac{2}{e+1}-1}\rho^2}{2\pi a^2(e-1)}\right) \tag{85}$$

$$= -\frac{\sqrt{\frac{2}{e+1}-1}\rho^2\left|\frac{a\sqrt{e^2-1}}{\rho\sqrt{-(ea+a-\rho)(a(e-1)+\rho)}}\right|}{\pi a^2(e-1)} \tag{86}$$

This probability density function, valid in the range $a(1-e) \leq \rho \leq a(1+e)$, is verified by Monte Carlo, as shown in Fig. 6.

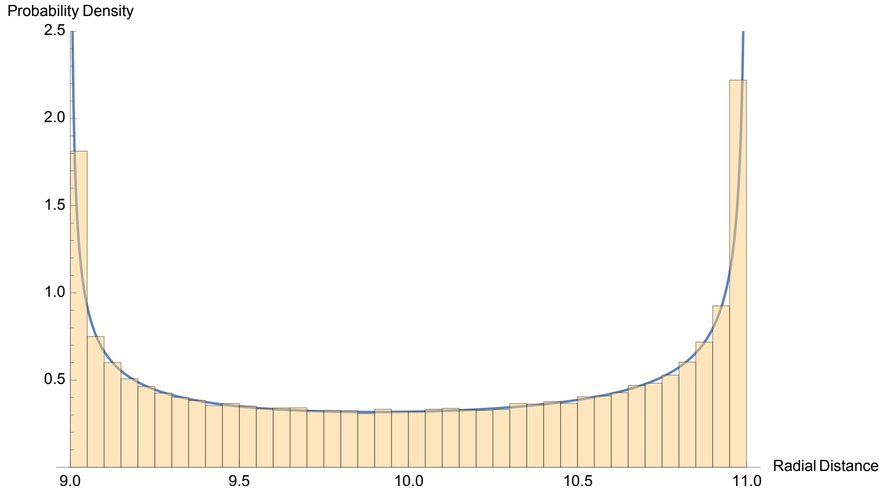

Figure 6: Monte Carlo verification of probability density function for $a = 1$, $e = 0.1$

Pairwise Distance Scalar Function

The position of each spacecraft may be represented in a Cartesian system

$$x_1 = \rho_1 \cos\theta_1 \qquad x_2 = \rho_2 \cos\theta_2 \tag{87}$$

$$= \frac{a_1(1-e_1^2)}{1+e_1\cos\theta_1}\cos\theta_1 \qquad = \frac{a_2(1-e_2^2)}{1+e_2\cos\theta_2}\cos\theta_2 \tag{88}$$

$$y_1 = \rho_1 \sin\theta_1 \qquad y_2 = \rho_2 \sin\theta_2 \tag{89}$$

$$= \frac{a_1(1-e_1^2)}{1+e_1\cos\theta_1}\sin\theta_1 \qquad = \frac{a_2(1-e_2^2)}{1+e_2\cos\theta_2}\sin\theta_2 \tag{90}$$

$$z_1 = 0 \qquad z_2 = 0 \tag{91}$$

Orbit 2 is rotated relative to orbit 1. This rotation can be represented by a quaternion, from which one can form the rotation matrix shown below.

$$A(\mathbf{q}) = \begin{bmatrix} 1-2q_2^2-2q_3^2 & 2q_1q_2-2q_3q_4 & 2q_1q_3+2q_2q_4 \\ 2q_1q_2+2q_3q_4 & 1-2q_1^2-2q_3^2 & 2q_2q_3-2q_1q_4 \\ 2q_1q_3-2q_2q_4 & 2q_1q_4+2q_2q_3 & 1-2q_1^2-2q_2^2 \end{bmatrix} \tag{92}$$

The rotated position for a spacecraft on orbit 2 is given by:

$$\begin{bmatrix} x_2' \\ y_2' \\ z_2' \end{bmatrix} = \begin{bmatrix} 1-2q_2^2-2q_3^2 & 2q_1q_2-2q_3q_4 & 2q_1q_3+2q_2q_4 \\ 2q_1q_2+2q_3q_4 & 1-2q_1^2-2q_3^2 & 2q_2q_3-2q_1q_4 \\ 2q_1q_3-2q_2q_4 & 2q_1q_4+2q_2q_3 & 1-2q_1^2-2q_2^2 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \tag{93}$$

$$= \begin{bmatrix} \left(-2q_3^3-2q_2^2+1\right)x_2 + (2q_1q_2-2q_3q_4)\,y_2 + (2q_1q_3+2q_2q_4)\,z_2 \\ (2q_1q_2+2q_3q_4)\,x_2 + \left(-2q_1^2-2q_3^2+1\right)y_2 + (2q_2q_3-2q_1q_4)\,z_2 \\ (2q_1q_3-2q_2q_4)\,x_2 + (2q_2q_3+2q_1q_4)\,y_2 + \left(-2q_1^2-2q_2^2+1\right)z_2 \end{bmatrix} \tag{94}$$

The squared distance between these two spacecraft is given by:

$$d^2(x_1, y_1, z_1, x_2', y_2', z_2') = (x_1 - x_2')^2 + (y_1 - y_2')^2 + (z_1 - z_2')^2 \tag{95}$$

Written as a vector expression:

$$d^2(\theta_1, \theta_2) = \left[\begin{bmatrix} \rho_1(\theta_1, a_1, e_1)\cos\theta_1 \\ \rho_1(\theta_1, a_1, e_1)\sin\theta_1 \\ 0 \end{bmatrix} - A(\mathbf{q})\begin{bmatrix} \rho_2(\theta_2, a_2, e_2)\cos\theta_2 \\ \rho_2(\theta_2, a_2, e_2)\sin\theta_2 \\ 0 \end{bmatrix}\right]^T \left[\begin{bmatrix} \rho_1(\theta_1, a_1, e_1)\cos\theta_1 \\ \rho_1(\theta_1, a_1, e_1)\sin\theta_1 \\ 0 \end{bmatrix} - A(\mathbf{q})\begin{bmatrix} \rho_2(\theta_2, a_2, e_2)\cos\theta_2 \\ \rho_2(\theta_2, a_2, e_2)\sin\theta_2 \\ 0 \end{bmatrix}\right] \tag{96}$$

Substituting:

$$\begin{aligned} d^2(\theta_1, \theta_2) = &\left(\frac{a_1\left(1-e_1^2\right)\sin\left(\theta_1\right)}{e_1\cos\left(\theta_1\right)+1} - \frac{a_2\left(1-e_2^2\right)\left(2q_1q_2+2q_3q_4\right)\cos\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1} - \frac{a_2\left(1-e_2^2\right)\left(-2q_1^2-2q_3^2+1\right)\sin\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1}\right)^2 \\ &+\left(-\frac{a_2\left(1-e_2^2\right)\left(2q_1q_3-2q_2q_4\right)\cos\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1} - \frac{a_2\left(1-e_2^2\right)\left(2q_2q_3+2q_1q_4\right)\sin\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1}\right)^2 \\ &+\left(\frac{a_1\left(1-e_1^2\right)\cos\left(\theta_1\right)}{e_1\cos\left(\theta_1\right)+1} - \frac{a_2\left(1-e_2^2\right)\left(-2q_3^3-2q_2^2+1\right)\cos\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1} - \frac{a_2\left(1-e_2^2\right)\left(2q_1q_2-2q_3q_4\right)\sin\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1}\right)^2 \end{aligned} \tag{97}$$

In the case that the two orbits are coplanar ($q = [0, 0, 0, 1]^T$), the above expression reduces to:

$$d^2(\theta_1, \theta_2) = \left(\frac{a_1\left(1-e_1^2\right)\cos\left(\theta_1\right)}{e_1\cos\left(\theta_1\right)+1} - \frac{a_2\left(1-e_2^2\right)\cos\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1}\right)^2 + \left(\frac{a_1\left(1-e_1^2\right)\sin\left(\theta_1\right)}{e_1\cos\left(\theta_1\right)+1} - \frac{a_2\left(1-e_2^2\right)\sin\left(\theta_2\right)}{e_2\cos\left(\theta_2\right)+1}\right)^2 \tag{98}$$

With this scalar function for separating distance, one can find contours of constant distance as a function of the polar angles θ_1 and θ_2, as shown in Fig. 7.

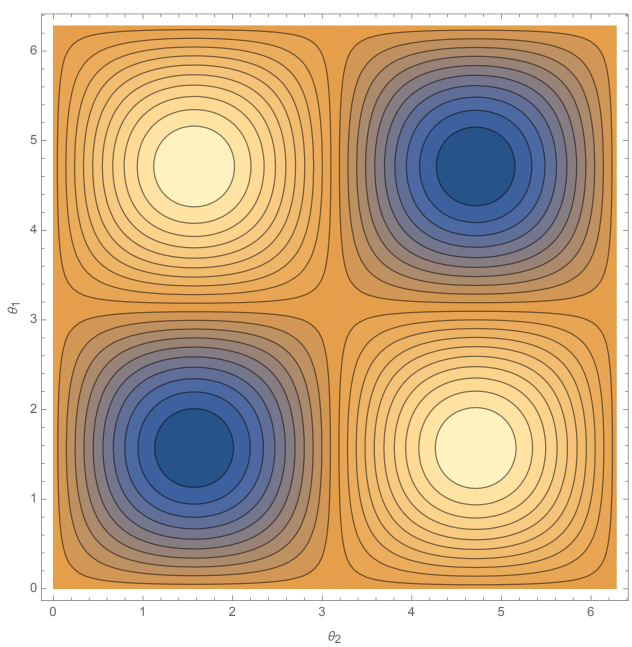

Figure 7: Constant distance contours for equal semimajor axes, zero eccentricity, and 90-degree relative rotation of orbits

References

[1]Manchester, Zachary, Mason Peck, and Andrew Filo. "Kicksat: a crowd-funded mission to demonstrate the worlds smallest spacecraft." (2013).

[2]Broch, Josh, et al. "A performance comparison of multi-hop wireless ad hoc network routing protocols." Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking. ACM, 1998.

[3]Parker, Lynne E. "Current state of the art in distributed autonomous mobile robotics." Distributed Autonomous Robotic Systems 4. Springer Japan, 2000. 3-12.

[4]Szebehely, Victor, ed. Celestial mechanics and astrodynamics. Elsevier, 2014.

[5]Kepler, Johannes, et al. Astronomia nova. CH Beck'sche Verlagsbuchhandlung, 1937.

[6]Kepler, Johannes. Harmonices mundi. Culture et civilisation, 1968.