Enhancing Path Reliability in Contact Graph Routing via Improved Hop Time Estimations

Ricardo Lent
Engineering Technology
University of Houston
Houston, Texas, USA
rlent@uh.edu

***Abstract*—Delay-Tolerant Networking (DTN) enables the forwarding of data bundles over space networks that experience extended link disruptions and path disconnections. Routing in such environments is challenging yet crucial for efficient end-to-end data delivery. Contact Graph Routing (CGR) is the standard routing method adopted for space DTNs. This study enhances CGR by exploring the potential inclusion of a Cognitive Element (CE) that leverages a data-driven approach. The CE is anticipated to use machine learning to estimate average single-hop bundle delivery times based on selected inputs. These estimates then replace the propagation delay that is used as the sole decision metric in CGR's shortest-path algorithm, improving the accuracy of the average one-hop delivery time predictions by allowing consideration of significant factors such as Convergence Layer Adapter (CLA) behavior, configuration parameters, packet drop probabilities, and unreliable contacts. The result is enhanced routing performance. The paper evaluates the CE extension in a simulated Earth-Moon network, assuming implementation independence and examining the effects of contact plan size (defined as a look-ahead window) and potential performance degradation from reduced prediction accuracy due to partial data or model limitations. Insights into the practical benefits of this approach are provided with a focus on realistic contact features and unreliable links.**

Keywords-delay-tolerant networking; routing; reliability; performance evaluation; cognitive networking.

I. Introduction

This paper extends the exploration of the integration of a Cognitive Element (CE) into Contact Graph Routing (CGR) to enhance the routing performance [1] of space Delay-Tolerant Networks (DTNs). A DTN provides crucial services in facilitating the communication among spacecraft, rovers, orbiters, landers, and ground stations in space exploration missions that often times involve significant signal propagation delays because of the long-distance of the communication links and periods of signal disruption due to celestial bodies obstructing line-of-sight communication paths and other factors. DTN plays a key role in LunaNet, NASA's proposed lunar communications and navigation network [2], [3] for the Artemis program, which demands systems capable of efficiently managing the unique challenges of space networking. LunaNet will provide transparent networking services that establish end-to-end data paths through a disconnected and time-varying topology. Routing is a key component of space DTNs that determines the store-carry-and-forward communication path for data bundles. The pre-planned nature of these networks simplifies the routing task, as contact opportunities can be anticipated from the expected positions of nodes as derived from orbital calculations. These calculations not only identify link obstructions but also provide the information required for a link budget analysis. CGR leverages the available future contact information to compute the optimal next-hop for bundles achieving minimum latency to the destination.

However, it is relevant to point out that, despite the deterministic assumption of contacts in scheduled DTNs, variations can still arise due to a multitude of factors. For instance, cloud coverage can bring large signal attenuation at high radio frequencies and in free-space optical links that can disrupt expected contacts between an orbiter and a ground node. Node malfunction and antenna misalignment issues may also occur randomly preventing contact realizations. Moreover, operational priorities may dynamically change resulting in the re-assignment of expected contacts to a different application.

These observations are aligned with the evolution of Opportunistic Contact Graph Routing (OCGR), which explores the potential utilization of non-scheduled contacts associated with a calculated confidence level. OCGR introduces a shift in the path search methodology of CGR, allowing the discovery of the k-shortest paths and the assessment of path reliability. Extending this concept further, it can be assumed that all contacts in a DTN have an opportunistic nature, including scheduled contacts, as they may randomly fail as discussed. Therefore, at least the path searching part of OCGR can be widely applicable to optimize unreliable DTN scenarios, provided each contact can be associated with a confidence level.

One limitation of CGR and OCGR is that the time progression step of each bundle forwarding within the path search algorithm assumes ideal transmission conditions that are determined solely by the link propagation delay or one-way light time. Buffering information is considered unavailable beyond the links leading to neighboring nodes, therefore not fully accounting for transmission and queuing delays. Additionally, protocol dynamics, including the convergence-layer adapter (CLA), particularly concerning the handling of packet losses through retransmissions, are overlooked. These factors contribute to differences between the calculated times within the CGR path search algorithm and the actual bundle

This work was supported by grant #80NSSC22K0259 from the National Aeronautics and Space Administration (NASA).

forwarding performance, which may impact routing optimally specially with network congestion.

Building on the initial exploration of performance gains from integrating a CE into CGR [1], this study extends the evaluation by incorporating additional options. The core concept remains that the CE can generate more accurate estimates of average one-hop bundle delivery times, helping CGR's shortest-time algorithm to identify optimal paths by factoring in predicted network performance metrics. The main contributions of this work include:

1) The concept of using a CE to forecast average single-hop bundle delivery times, which is utilized in the time progression step of CGR, is further developed. The core idea is to introduce a data-driven approach that aids in identifying the best paths by considering factors such as specific CLA behavior, configuration parameters, packet drops, and unreliable contacts. The CE could be trained either offline using an analytical model or historical data, or online with real-time measurements to achieve accurate predictions. This approach eliminates the need for modifications to the CGR algorithm to handle uncertain contacts and random factors, thereby removing the requirement for searching for the k-shortest paths as done in OCGR. Additionally, this study eliminates the constraint that bundle arrivals must coincide within a contact duration by noting the average nature of the delivery time estimations. Network congestion metrics are also used to filter the contact plan before computing the path search.
2) This study evaluates the impact of the contact plan size on routing optimality in CGR and the CE extension. Since CGR relies on building a graph where nodes represent future available contacts, a shorter contact plan can speed up path computation and improve efficiency. However, this also risks insufficient network connectivity for path computation. The study assesses how the size of the contact plan, defined as the look-ahead time window, affects CGR routing performance and the performance of the CE extension.
3) An evaluation of the performance impact of the limitations of the CE in producing accurate average bundle time estimations. The CE provides a function that maps the known network state to forecast the time required for a bundle to reach the next hop. The limitations of the method are therefore related to the accuracy of the network state knowledge, particularly because the required information may not necessarily be available at the nodes. This study provides an implementation-agnostic assessment of the performance advantages and limitations of the CE, identifying the performance bounds of the method across three variations regarding the severity of assumptions involving the network state. In the first scenario, only local state information, normally assumed available in the standard CGR, is assumed. The second and third scenarios require global knowledge, i.e., information external to the node, with differences in how they predict transmission hop times. The evaluation is conducted within the context of an Earth-Moon network [2], employing approximately realistic values for contact features and considering unreliable contacts. The evaluation provides insight into the impact of imperfect CE model predictions on end-to-end bundle routing performance.

The remainder of the paper is structured as follows: Section II reviews related works relevant to this study. Section III provides a detailed explanation of the Cognitive Element (CE) method. Section IV describes the evaluation scenario and simulation assumptions. Section V presents the results demonstrating the CE's effectiveness in optimizing bundle flow across an Earth-Moon network. Lastly, Section VI offers concluding remarks.

II. Related Works

The reliability of DTN protocols remains a dynamic area of research with application to many ambitious missions [2], [4]. A feature that characterizes space DTNs is the use of scheduled contacts, commonly used jointly with the Bundle Protocol (BP) [5], [6] and Contact Graph Routing (CGR) [7], [8], which begins by constructing a graph, where vertices denote active contacts and links represent logical transitions between contacts—where one contact's endpoint aligns with the next contact's starting point, feasible within a defined time frame. While this process incorporates factors, such as transmission time, propagation delay, and network disruptions, buffering delays are typically overlooked due to the distributed nature of the algorithm, as this information is normally inaccessible. CGR is commonly implemented by adapting Dijkstra or Yen's algorithms, with the latter method raising scalability concerns [9].

The performance of CGR in scenarios involving unreliable links has been explored in various contexts, including satellite constellations [10] and random networks [11]. Reliability has been mainly addressed by BP custody [12] and CLA design via retransmissions, e.g., the Licklider Transmission Protocol (LTP). For experimental results, see for example [13]–[15]. These studies have shed light on CGR's vulnerabilities concerning contact failure rates and random losses. An extension known as Opportunistic CGR (OCGR) [16] investigates the potential integration of nonscheduled contacts—either discovered or predicted—into CGR's standard path search algorithm, assigning them a confidence level. OCGR maintains a record of the contact history of nonscheduled contacts to predict future contacts, alongside their associated properties and confidence levels, calculated based on available contact history [16]. Discovered contacts are assigned a unit confidence [17] and the resulting route is assigned a delivery confidence derived from the product of the confidence levels of the contacts involved. In recent iterations, the implementation of OCGR [18] evaluates path candidates based on their arrival confidence, considering a predefined margin from the highest confidence level.

Additional related methods to this work include CGR extensions to support multigraphs [19], better handling of capacity constraints [20], [21], a variety of networks (and Roaming DTN (RDTN) [22] that integrates roaming nodes with unpredictable motion, Best Routing Under Failures (BRUF) [23], where the routing process is conceptualized as a Markov Decision Process, with certain state transitions becoming probabilistic due to the limited reliability of specific contacts and Routing under Uncertain Contact Plans (RUCoP) [24], [25] that introduces a multiple-copy Markov Decision Process. Also related, is the Cognitive Space Gateway (CSG) [26] where routing decisions are delegated to a Spiking Neural Network which is continually trained after the bundle transmissions using a reinforcement learning approach.

This paper presents an alternative approach to enhance CGR performance, a method known for its computational efficiency and practicality, but limited in handling random factors impacting single-hop bundle transmissions, such as packet losses and contact failures. Unlike previous approaches, this method modifies the conventional one-hop bundle time calculations. Specifically, it introduces the idea of using a cognitive element designed to accurately predict average bundle transmission times. While the implementation of this cognitive element is expected to utilize a neural network or similar structure, this study evaluates its limitations without specifying a particular technology. Instead, it offers widely applicable findings focused on determining performance bounds based on assumptions regarding available network state information used as inputs to the CE.

III. Cognitive Extension and CGR

CGR defines a decentralized approach in which each node calculates the path to the destination node, but using only next-hop information to forward bundles. This method requires access to the contact plan for all future contacts, which is distributed to the DTN nodes in advance.

A. Standard Mechanisms

A contact plan consists of a sequence of entries of the following form: $(\mathcal{I}_i, \mathcal{F}_i, \mathcal{T}_i, \mathcal{S}_i, \mathcal{E}_i, \mathcal{R}_i, \mathcal{O}_i, r_i)$ and that includes a contact identifier $\mathcal{I}_i$, the sending $\mathcal{F}_i$ and receiving node $\mathcal{T}_i$ identifiers, the start $\mathcal{S}_i$ and end $\mathcal{E}_i$ times, the transmission rate $\mathcal{R}_i$ and the propagation delay or one-way light time $\mathcal{O}_i$ that depends on the distance between the nodes. The term r_i, $0 \leq r_i \leq 1$, is the contact confidence as used by O-CGR [16].

To determine routes for each desired destination, CGR builds a contact graph $G = (V, E)$ using each contact entry of the plan as a vertex minus the entries containing excluded nodes (e.g., known failed nodes). A contact graph is a directed acyclic graph where an edge exists when two contacts are logically connected, which happens when the destination node of the first contact matches the sending node of the second contact and the latter expires after the first. The target contact of an edge is called the proximate of the first contact. The contact graph is considered directed as transmissions in the reverse direction of a given contact may not be possible or may occur with different transmission parameters, e.g., different transmission rate, due to transceiver limitations. A start time of the proximate that is later than the current time while using a contact brings forced data buffering due to the corresponding link disruption.

CGR derives the path to the destination node by calculating the shortest path on the contact graph between two auxiliary vertices that are attached to represent the root and terminal contacts. The root is the node executing CGR. These auxiliary contacts involve a zero-cost to the relevant proximates. Starting from the root, a graph traversal iteratively tracks the bundle transmission progress in the network by estimating its arrival time as it is forwarded over contacts. That is, if t_i represents the bundle arrival time calculated at vertex i of the contact graph, the algorithm evaluates the proximate vertices j and greedily chooses the one offering the smallest t_j. Specifically, the evaluation of the proximate vertex j, yields the following arrival time.

$$t_j = \begin{cases} t_i + \mathcal{O}_j & \mathcal{S}_i \leq t_i \\ \mathcal{S}_j + \mathcal{O}_j & \mathcal{S}_i > t_i \end{cases} \quad (1)$$

The calculation does not include transmission times, but that metric is utilized to determine the remaining data volume for transmissions. This additional step enables consideration of whether given contacts are likely to be already fully booked. However, this assessment is restricted to contacts leading to neighboring nodes, as information beyond that scope is unavailable. The output of the algorithm is the path $P = v_0, v_1, \ldots, v_k$, where $v_i \in V$ is a contact and v_0, v_k are the auxiliary contact entries for the source and sink nodes respectively. If t_k is the estimated time to deliver the bundle to the end contact based on (1) for each step, the objective of the algorithm is to minimize t_k among all possible paths from v_0 to v_k in G.

B. Cognitive Element

The central idea of this paper is to enhance the route selection quality in CGR by refining the accuracy of the single-step bundle forwarding time calculation. This involves substituting the computation outlined in (1) with the output of a cognitive element (CE) designed to accurately predict the average time needed to deliver a bundle to the next hop, accounting for the segmentation, transmission and retransmission times of the convergence-layer adapter, buffering delays, and the reliability of contacts, among other factors:

$$t_j = \begin{cases} t_i + y_j & \mathcal{S}_i \leq t_i \\ \mathcal{S}_j + y_j & \mathcal{S}_i > t_i \end{cases} \quad (2)$$

where $y_j = f_\theta(x)$ represents the output of a function f_θ given the specified system state x and the model parameters θ.

A second observation concerns the interpretation of t_j, which now represents the average time to reach the next hop, rather than the precise definition in CGR. This change is required to properly take into account probabilistic factors, such as transmission errors and contact failures. The idea is that these probabilistic factors will affect the one-hop bundle

delivery time along the path adding uncertainty into the calculation of the final delivery time. With this reinterpretation of t_j, the shortest path algorithm of CGR requires no modification. It continues to identify the route with the smallest average time of arrival t_k (instead of precise time), but now able to accommodate random factors affecting the paths.

In this study, we keep the concept of introducing a CE to CGR separated from its implementation on purpose, recognizing that diverse techniques may be used to define this element. Possible mechanisms encompass a range of neural network architectures, including multi-layer feedforward, convolutional, generative adversarial, recurrent (such as Long Short-Term Memory Networks), autoencoders, graph neural networks, and more. These mechanisms can be implemented using either continuous activation or spiking neurons. Given the potential variations in prediction accuracy resulting among different techniques, our focus is in assessing the performance bounds attained with the introduction of the CE concept and understanding the performance implications of using imperfect inputs for $f_\theta(x)$.

In particular, we focus on studying three variations for $f_\theta(x)$. The first case, which is labeled CE-1, considers $f_\theta(x)$ providing the average one-hop bundle time that aggregates the propagation delay, and additional buffering time required with imperfect contact reliability. The second case, CE-2, improves CE-1 estimation by also aggregating the estimated bundle transmission time. The last case, CE-3, includes in addition to the elements of CE-2 the expected buffering time.

We note that CE-1 and CE-3 use the same expressions as in [1]. However, unlike the previous work, we remove the constraint in the path search algorithm that requires bundle transmissions to be completed within a contact duration. In this study, we consider transmissions as simple approximations of the average time required to deliver a bundle with reliability constraints. We note that this change tends to improve throughput as observed in the evaluation scenario.

Regarding the training of the models, CE-1 and CE-2 are comparatively the easiest to train since they require only local state information, which is available in the standard contact plan, i.e., transmission rate, propagation delay, and confidence level. The latter parameter can be understood as an estimate of contact reliability. CE-3, however, requires predicting the global state, as buffer occupancy levels dynamically change. CE-2 and CE-3 could be further improved if the channel bit-error-rate value could be estimated as this value determines the extended times required for retransmissions. To maintain the study's focus on evaluating the effectiveness of the CE concept rather than discussing specific approaches, we omit further details of the training phase for these models.

C. Model Approximation

CE-based predictions address the effect of contact reliability and are calculated using the following models [27], which provides the average time required to deliver a packet over contact j:

CE-1:

$$y_j = \mathcal{O}_j + \frac{1-r}{r}C \tag{3}$$

CE-2:

$$y_j = \frac{L}{R_j} + \mathcal{O}_j + \frac{1-r_j}{r_j}C \tag{4}$$

CE-3:

$$y_j = \frac{L+B}{R_j} + \mathcal{O}_j + \frac{1-r_j}{r_j}C \tag{5}$$

where L is the average bundle size, R_j the link rate, P_j the propagation delay, C the average time between contacts, B the buffer occupancy and r_j the contact's reliability. Note that the subscript j emphasizes that the parameters is per-contact and that both B and C refer to values associated with the next-hop node of proximate j. The value y_j is then used to advance time in the iteration of the shortest path calculation as used in 2. Assuming that the contact plan includes r_j as done with O-CGR, all of the inputs can be directly extracted from the contact plan except for L. However, L can be iteratively estimated from bundle arrivals through exponential averaging, that is, each arriving bundle of size l allows updating L as follows: $L \leftarrow \alpha l + (1-\alpha)L$, where α, $0 \leq \alpha \leq 1$, is a hyperparameter.

IV. Evaluation Scenario

Consider a communication scenario where a node located on the lunar surface regularly emits messages to a terrestrial sink. This scenario corresponds to a typical space exploration communication model, such as a rover collecting scientific data that is then sent for analysis or a hub aggregating data from various sources before transmitting it to Earth. The time required to deliver the data, i.e., the response time, and the risk of data loss provide sensible assessment of the benefits of introducing the CE extension.

To implement the scenario two simulators were developed. The first simulator generates the contact plan by estimating the locations of nodes from orbital calculations that accounts for both Earth's and Moon's rotation and translation, which helps to determine transmission opportunities based on the line-of-sight between nodes. The starting separation distance between the nodes is used to define the one-way light time that appears in the contact plan for each entry. The second simulator evaluates routing performance by using the generated contact plan and implementing an event-driven simulation of the bundle transmission dynamics, including buffering and drops, while considering potential contact failures that prolong buffering times.

The traffic originates from a Lunar node positioned on the southern far side of the Moon at -19.94, -200.07 in the selenographic coordinate system. As such, the node is not visible from Earth, but three orbital relays are available to forward the data: LO1, LO2, and LO3. For simplicity, Keplerian assumptions were used to model their orbits. The orbits are characterized by inclinations of 10, 40, and -40 degrees, and Right Ascension of the Ascending Node (RAAN) values of 4.462, 90, and 40 respectively. It is relevant to note

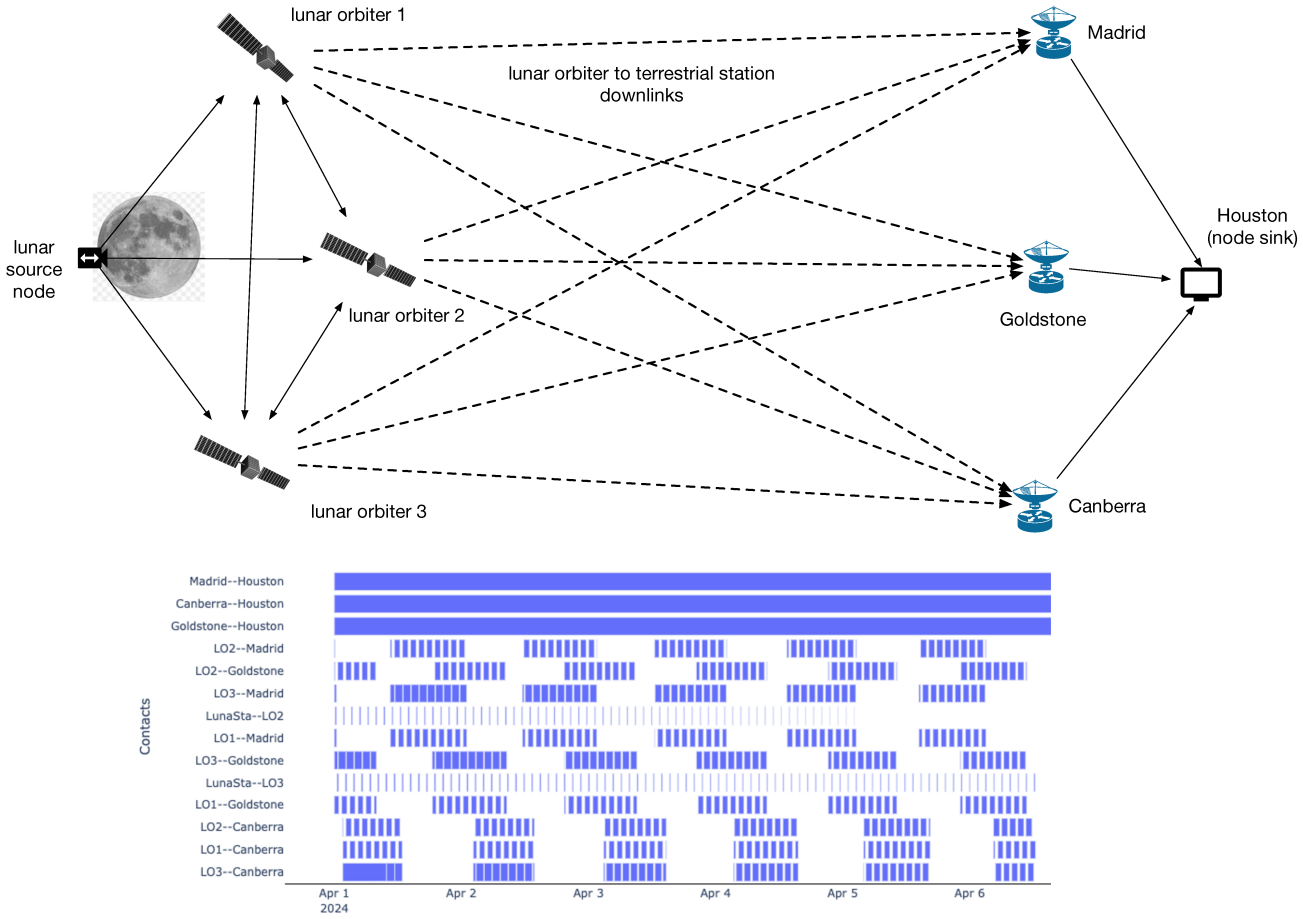


Figure 1. The evaluation scenario involves a source located on the surface of the Moon that generates traffic addressed to a terrestrial sink. The traffic can be routed through contacts provided by three lunar orbiters and three ground stations, as depicted in the figure. The dashed lines represent links that may be affected by unreliable contacts. The lower part of the figure shows the contact pattern between the nodes.

that these orbits were not modeled after existing lunar satellites but were defined to facilitate the establishment of contacts of varying durations with the rover and Earth stations. However, we consider the scenario to be close enough to that of a real mission.

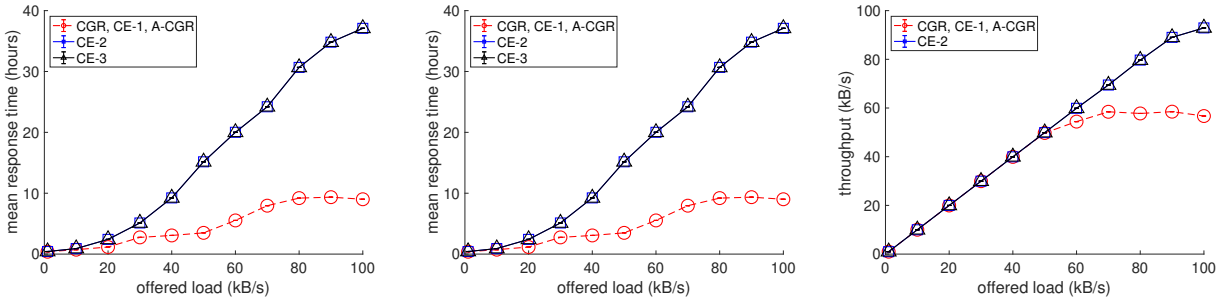
The terrestrial ground stations are modeled to match with the locations of the Deep-Space Communication (DSN) complexes in Canberra, Madrid, and Goldstone. The sink is assumed to be situated in Houston, with a permanent link established from each DSN location to Houston. In reality, these connections would traverse one or more networks. To account for this, the propagation delays for the terrestrial connections were calculated based on the distance between the involved nodes as the direct as-the-crow-flies distance with an additional 20% margin to account for network overhead.

The critical part of the system lies in the communication between the lunar node and the lunar orbiters, as well as between the lunar orbiters and the terrestrial stations, where the links provide on-off capacity. The evaluation topology is depicted in Figure 1 and Table I summarizes the features of these contacts, including the average and standard deviation of the contact duration and period (i.e., the inter-contact time). The contact pattern extracted from the orbital model is shown in the bottom chart of Figure 1. Note that the topology allows for sending orbital crosslink traffic if required by the routing protocol and provided contacts are available based on the orbital model.

TABLE I. AVERAGE (μ) AND STANDARD DEVIATION (σ) OF CONTACT DURATIONS AND PERIOD LENGTHS (TIME BETWEEN CONSECUTIVE CONTACTS) FOR THE EARTH-MOON EVALUATION NETWORK.

Contact type	**Duration μ**	**Duration σ**	**Period μ**	**Period σ**
Rover to LO1	13.0	1.9	91.4	3.4
Rover to LO2	10.4	3.8	90.6	10.3
Rover to LO3	13.3	2.5	91.1	6.5
LO1 to Madrid	55.0	10.8	149.7	186.7
LO1 to Canberra	53.7	12.8	160.9	217.7
LO1 to Goldstone	55.3	9.5	147.1	185.2
LO2 to Madrid	56.7	19.0	147.5	181.6
LO2 to Canberra	61.2	12.0	170.4	229.9
LO2 to Goldstone	58.9	16.6	147.9	189.4
LO3 to Madrid	69.8	15.1	152.3	196.4
LO3 to Canberra	75.7	67.4	178.6	239.4
LO3 to Goldstone	68.8	16.7	146.6	194.4

The transmission rate for the terrestrial (wired) links was



(a) Mean bundle response time. (b) Bundle loss ratio. (c) Throughput.

Figure 2. Flow performance metrics in a network with reliable contacts.

fixed at 2 Mbps, while all wireless transmissions were set to 100 Kbps. In all cases, the links are assumed to have negligible bit error rates (BER). It is also assumed that all contacts are reliable, possibly except for the ones between the orbiters and ground stations, due to the long distances involved. To this end, we evaluate three scenarios: (1) where none of the downlinks are affected, (2) where all downlinks from the three lunar orbiters are affected with reliability factors of 0.95, 0.85, and 0.5, and (3) where the downlinks from a single lunar orbiter have a limited reliability of 0.5 while the others remain fully reliable. While these values have been chosen arbitrarily, they are intended to illustrate a good range of possible scenarios that could occur in a real system.

In this context, the CGR agent running in the lunar node decides which orbiter will handle the bundle forwarding to Earth. The selected orbiter then determines which terrestrial station will receive the bundle before forwarding it to the sink. Also, the bundles are not associated with a finite deadline and the node buffers are assumed to be large enough to ignore the impact of buffer overflows, so bundles that miss any given contact simply continue waiting in the buffer for future service.

However, bundle drops are still possible under certain conditions. For computational efficiency, the length of the contact plan used by the CGR for route determination is limited to a pre-selected look-ahead window. That is, the contact plan is filtered to contain unexpired contacts whose starting time is not later than the current time plus the look-ahead window. If it is not possible to determine a path for the bundle using the information in the contact plan–for example, if all contacts are expected to be busy–the bundle is considered lost for the purposes of this study. This is measured by the "routing miss" metric. With the use of varied reliability factors and look-ahead window sizes, it is possible to identify the level of impact and assess the robustness of the routing scheme under various conditions.

V. Results and Discussion

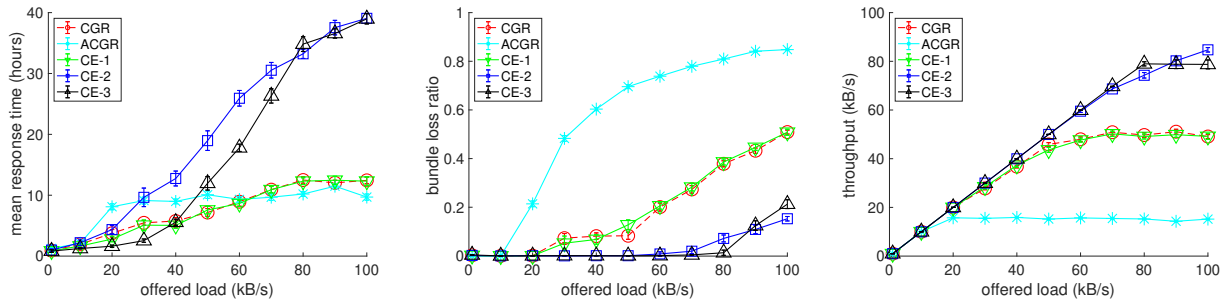
The routing performance is evaluated through observations of the response time, bundle loss ratio, and throughput of a test flow, and studied under simulation conditions where buffer capacities are uncapped, bit error rates (BER) are negligible, and no deadlines are imposed on bundle delivery times. Throughout is calculated as the product of the offered load and the delivery ratio (one minus the bundle loss ratio), both of which can be directly measured in the simulator. In the tests, bundles are generated at a constant rate of one every 100 seconds, while the bundle size is varied as an experimental parameter to adjust the offered load of the flow. The results reported in the next sections show in all cases the 95% confidence interval of the acquired samples for each experimental factor.

A. Reliable Contacts

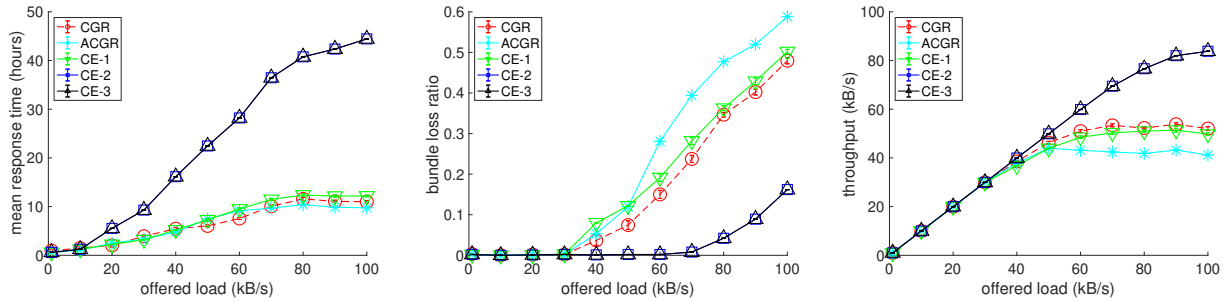
It is initially assumed that all contacts are reliable. Figure 2 (a) shows the average bundle response time that was observed with such conditions as a function of traffic load. The response time metric includes both transmission and buffering times, with the latter determined by the time required for bundles to reach the head of the transmission buffers. Buffering time is influenced by factors such as traffic load, transmission rates, and the waiting time for contacts along the selected path. The response time of a bundle is measured as the difference between its arrival time at the sink and its generation time by the simulator.

It can be observed that the average time required to transmit small files is around one hour or less. This duration is primarily determined by the waiting times for the next contact opportunities, as transmission times are short and, under light traffic conditions, buffers tend to remain empty. With increasing file sizes, there is a corresponding rise in both storage and transmission demands, leading to an increase in the average response time. The results provided by CGR and CE-1 are identical given the reliable contacts assumption. Also A-CGR, which will be discussed in the next section, yields identical results. CE-2 and CE-3 produced about the same response times in this scenario.

The simulations were run with a look-ahead time window of 6 hours. It was observed that the standard CGR and CE-1 started to have difficulties in determining the path for bundles with loads above 50 kB/s using the contacts contained within that window duration. This was not the case with CE-2 and CE-3 as can be observed in Figure 2 (b). Lower bundle losses



(a) Case 1: Mean response time. (b) Case 1: Bundle loss ratio. (c) Case 1: Throughput.



(d) Case 2: Mean response time. (e) Case 2: Bundle loss ratio. (f) Case 2: Throughput.

Figure 3. Flow metrics obtained in a network with unreliable contacts in the lunar orbiter to terrestrial station downlinks: (a), (b) and (c) with all downlinks unreliable; (d), (e) and (f) downlinks from one lunar orbiter unreliable.

benefits throughput as shown in Figure 2 (c) and increase the average response time as a result given that a large number of bundles are kept in the network, which explains the larger response time of CE-2 and CE-3 compared to the other methods.

B. Unreliable Contacts

We consider two cases as representative of the broad spectrum of possibilities involving unreliable contacts affecting the downlinks from the lunar orbiters to a terrestrial station. Case 1 assumes that the downlinks originated from each of the three lunar orbiters are affected by reliability factors of 0.95, 0.85, and 0.5. Case 2 assumes that all contacts are reliable except for the downlinks from one lunar orbiter, which have a contact reliability factor of 0.5. Figure 3 depicts the results.

In addition to CGR and the three cognitive extensions, we introduce A-CGR as a basic routing method that attempts to improve the route computation by enforcing the use of contacts with a reliability factor that is above a predefined threshold. This threshold was set to 0.9 in the experiments. This basic logic makes A-CGR functionally related to O-CGR, which evaluates the reliability of paths when making routing decisions. Although not identical to O-CGR, A-CGR provides reasonable baseline performance.

The charts on the left part of Figure 3, i.e., the ones labeled (a), (c) and (d), correspond to Case 1. It can be observed that A-CGR produced the lowest throughput given that the action of removing contacts despite offering limited reliability, also removes network capacity leading to a higher level of bundle drops. The response time of A-CGR, CE-1 and the standard CGR was observed to be very close. CE-2 and CE-3 achieved the highest throughput of all methods as both consider the impact of buffering delays when building the contact graph, with CE-3 producing lower response times than CE-2 by also considering buffering delays within the shortest path calculation. The results for Case 2 in the right part of Figure 3 follow a similar pattern although less penalized as only one link was affected by unreliable contacts. Also, the results with CE-2 and CE-3 were very close.

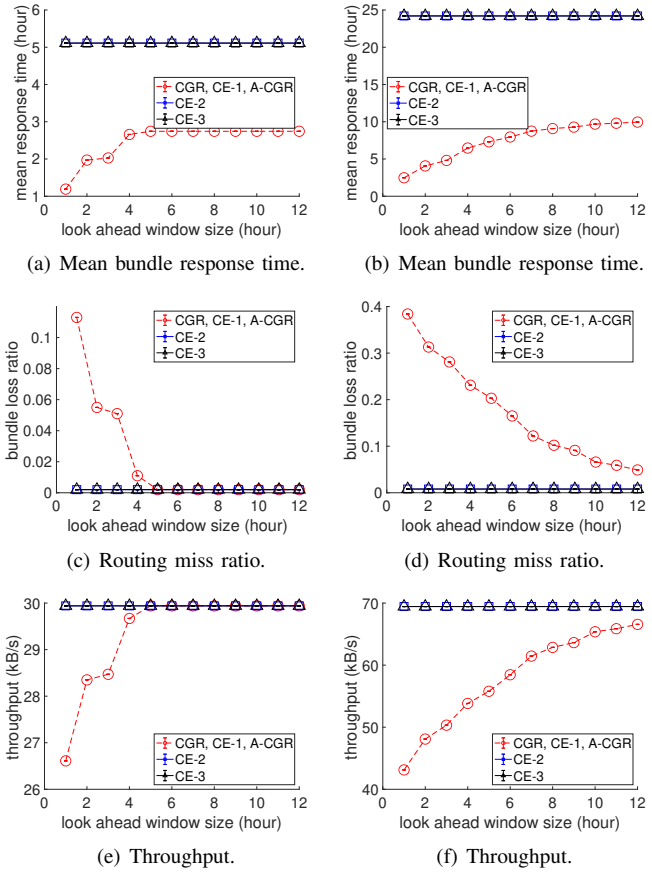
C. Impact of the Look-Ahead Window Size

The previous results were obtained with the look-ahead window size set at 6 hours. The next set of results evaluates the impact of selecting different look-ahead window sizes on routing performance. Observations were collected for two traffic load points: low at 30 kB/s and high at 70 kB/s. Based on prior results, the response time difference at high load is approximately twice as much as at low load, allowing us to observe performance differences between these two cases.

Figure 4 presents the observations collected in a network with reliable contacts. The figures in the left column show the results for low load, while those in the right column show results for high load. In each case, the flow metrics are presented in terms of average response time, bundle loss ratio, and bundle throughput. The results indicate that, for both cases, CE-2 and CE-3 yield similar outcomes across the range of look-ahead window sizes tested. However, CGR, CE-1, and A-CGR showed sensitivity to the look-ahead window size. In a

reliable network, these three methods produce identical results. At low traffic load, the CGR, CE-1, and A-CGR methods showed throughput improvement up to a look-ahead window size of approximately five hours, though with higher mean response times. At high traffic load, they follow a similar trend but over a much larger span.

Figures 5 and 6 report the same metrics for both traffic loads but with the tests running on a network with unreliable contacts. As before, two cases were observed: the first with all downlinks from the lunar orbiters affected by limited contact reliability at 0.95, 0.85, and 0.5 per orbiter (Figure 5), and another affecting a single orbiter at 0.5 (Figures 6). It can be observed that the trend initially noted in the reliable network continues in both scenarios with unreliable contacts. As the look-ahead window increases, CGR, CE-1, and A-CGR show larger mean response times, lower loss, and higher throughput. Performance differences between these methods are evident, with A-CGR tending to yield higher response times and lower throughput than the other methods in the first scenario.

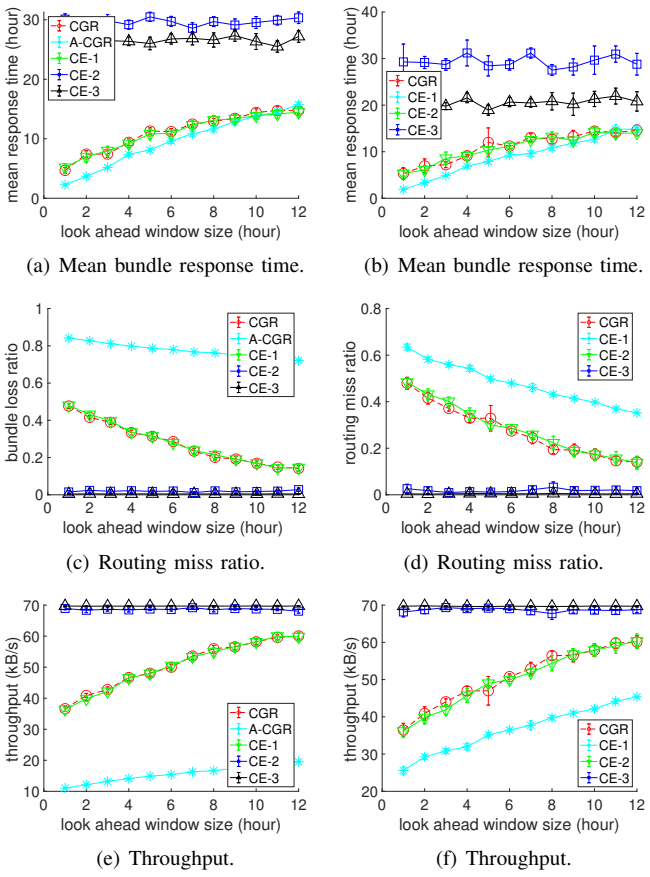


(a) Mean bundle response time. (b) Mean bundle response time.

(c) Routing miss ratio. (d) Routing miss ratio.

(e) Throughput. (f) Throughput.

Figure 4. Impact of the look ahead window size with reliable contacts for traffic loads of 30 kB/s (left column) and 70 kB/s (right column)



(a) Mean bundle response time. (b) Mean bundle response time.

(c) Routing miss ratio. (d) Routing miss ratio.

(e) Throughput. (f) Throughput.

Figure 5. Impact of the look ahead window size with with all downlinks affected by unreliable contacts (0.95, 0.85, 0.5 per orbiter) for traffic loads of 30 kB/s (left column) and 70 kB/s (right column)

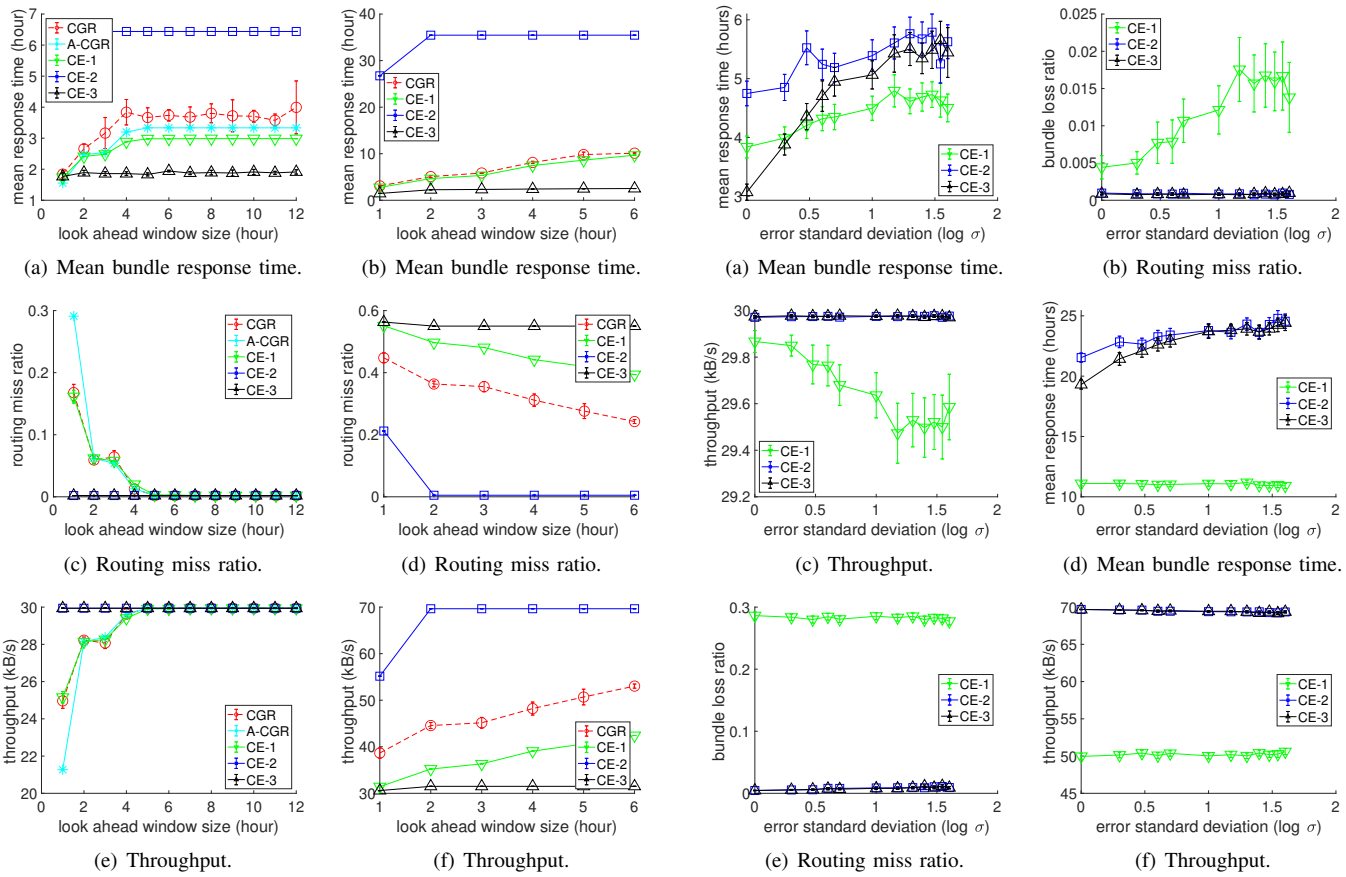
D. Impact of Prediction Errors in the Reliability Factor

The exact mechanism to determine reliability factors is left unspecified in this study, but it is interesting to observe how errors in this estimation would affect the performance metrics of the traffic flow. To achieve control over the error level, deviations where introduced to the model approximation y_j given by 3, 4, or 5, as follows:

$$y'_j = \max\{y_{min}, y_j \times \mathcal{N}(1, \sigma_e)\} \quad (6)$$

where y_{min} is a selected lower bound (0.1 in the tests) and $\mathcal{N}(1, \sigma_e)$ is a sample from a normal distribution with unit mean and standard deviation σ_e. The value y'_j is used in place of y_j when evaluating the effect of contact reliability in expression 2.

Observations were collected for two reference traffic load points at 30 kB/s and 70 kB/s as before for one of the scenarios to illustrate the impact of the estimation error given by parameter σ_e. Figure 7 shows the mean response time and throughput as a function of factor σ_e in the network with unreliable contacts. As before, assuming reliability factors of $0.95, 0.85, 0.5$ for the downlinks originating from each lunar orbiter. The top row containing figures (a), (b), (c) depicts the resulting flow metrics for low traffic, whereas the bottom row with figures (d), (e), (f) depicts the same metrics under high traffic. For low traffic loads (30 kB/s), all three extensions experience an increase in the response time with larger values of σ_e which is expected as the accuracy of the information



(a) Mean bundle response time. (b) Mean bundle response time.

(c) Routing miss ratio. (d) Routing miss ratio.

(e) Throughput. (f) Throughput.

Figure 6. Impact of the look ahead window size with one downlink affected by unreliable contacts r = 0.5 for traffic loads of 30 kB/s (left column) and 70 kB/s (right column)

(a) Mean bundle response time. (b) Routing miss ratio.

(c) Throughput. (d) Mean bundle response time.

(e) Routing miss ratio. (f) Throughput.

Figure 7. Impact of the prediction error σ_e with all downlinks originating from the three lunar orbiters affected by unreliable contacts $(0.95, 0.85, 0.5)$ and for traffic loads of 30 kB/s (top row) and 70 kB/s (bottom row).

available for routing becomes corrupted. However, both CE-2 and CE-3 demonstrated good resiliency in terms of throughout unlike CE-1 given the low bundle loss levels. With high traffic (70 kB/s), CE-2 and CE-3 show an increase in the response time with σ_e, but not CE-1 with the loss and throughput metrics of all methods unchanged. These results indicate that at the selected traffic level, the buffers of all downlinks reach a level of saturate that makes little difference choosing one downlink over another.

VI. Conclusion

In conclusion, this study assesses the integration of a Cognitive Element into CGR and its impact on routing performance. With the use of a data-driven methodology, the CE is expected to predict average single-hop bundle delivery times, accounting for latency-related factors such as CLA protocol behavior (e.g., retransmission dynamics), configuration parameters, and random variables like packet drops and unreliable contacts. This paper builds on prior work by removing the constraint in the path search algorithm that required bundle arrivals to fall within the bounds of a given contact. The use of average values from CE estimations allows better handling of the effects of unreliable contacts in the path search. Additionally, the study explores filtering contacts from the contact plan based on their predicted availability after factoring in expected buffer occupancies in the network. This approach results in shorter graphs, faster computation, and enhanced routing performance.

Comprehensive simulations conducted within an Earth-Moon network simulated context, assuming realistic contact features and accounting for unreliable contacts, show significant improvements in routing performance with the inclusion of a CE compared to the conventional CGR approach. This was evident when considering both regular network information available at a DTN node, i.e., the information contained in the contact plan, and extending this information to include network-wide buffer occupancies, i.e., global information. Unsurprisingly, the latter assumption yielded significant throughput improvements, particularly for traffic loads exceeding 50 kB/s in the tests, i.e., under congestion. For those cases, the simulations also showed lower requirements for the length of the contact plan.

We note that this study provides an implementation-agnostic assessment of the proposed approach using an analytical definition of the CE and its prediction errors. In practice, the CE is expected to be implemented using a neural network or related mechanism, with its structure, training algorithm,

and data quality affecting its prediction accuracy. The study addressed the potential performance degradation due to prediction errors. In particular, the results indicate that the CE method shows sensitivity to small errors, which can lead to delays of up to twice as much, although throughput remains largely unaffected. These findings highlight the advantages of using a cognitive networking approach to optimize space DTN performance and point to the importance of designing an accurate CE. Future research will focus on developing practical applications of this concept.

REFERENCES

[1] R. Lent, "Enhanced path reliability of Contact Graph Routing through a cognitive extension," in *The Sixteenth International Conference on Advances in Satellite and Space Communications (SPACOMM 2024)*, June 2024.

[2] D. J. Israel, K. D. Mauldin, C. J. Roberts, J. W. Mitchell, A. A. Pulkkinen, L. V. D. Cooper, M. A. Johnson, S. D. Christe, and C. J. Gramling, "Lunanet: a flexible and extensible lunar exploration communications and navigation infrastructure," in *2020 IEEE Aerospace Conference*, 2020, pp. 1–14.

[3] R. Dudukovich, D. Gormley, S. Kancharla, K. Wagner, R. Short, D. Brooks, J. Fantl, S. Janardhanan, and A. Fung, "Toward the development of a multi-agent cognitive networking system for the lunar environment," *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 269–283, 2022.

[4] C. Caini and R. Firrincieli, "Application of Contact Graph Routing to LEO satellite DTN communications," in *2012 IEEE International Conference on Communications (ICC)*, 2012, pp. 3301–3305.

[5] K. Scott and S. C. Burleigh, "Bundle Protocol Specification," RFC 5050, Nov. 2007. [Online]. Available: https://www.rfc-editor.org/info/rfc5050

[6] S. Burleigh, K. Fall, and E. J. Birrane, "Bundle Protocol Version 7," RFC 9171, Jan. 2022. [Online]. Available: https://www.rfc-editor.org/info/rfc9171

[7] S. Burleigh, C. Caini, J. J. Messina, and M. Rodolfi, "Toward a unified routing framework for delay-tolerant networking," in *2016 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, 2016, pp. 82–86.

[8] J. Segui, E. Jennings, and S. Burleigh, "Enhancing Contact Graph Routing for delay tolerant space networking," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, 2011, pp. 1–6.

[9] O. De Jonckère, J. A. Fraire, and S. Burleigh, "On the tractability of Yen's algorithm and contact graph modeling in Contact Graph Routing," in *2023 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, 2023, pp. 80–86.

[10] J. A. Fraire, P. Madoery, S. Burleigh, M. Feldmann, J. Finochietto, A. Charif, N. Zergainoh, R. Velazco, and S. Céspedes, "Assessing Contact Graph Routing performance and reliability in distributed satellite constellations," *J. Comput. Netw. Commun.*, vol. 2017, jan 2017.

[11] P. G. Madoery, F. D. Raverta, J. A. Fraire, and J. M. Finochietto, "Routing in space delay tolerant networks under uncertain contact plans," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.

[12] K. Zhao, R. Wang, S. C. Burleigh, A. Sabbagh, W. Wu, and M. De Sanctis, "Performance of bundle protocol for deep-space communications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2347–2361, 2016.

[13] R. Wang, X. Liu, L. Yang, Y. Xi, M. De Sanctis, K. Zhao, H. Yang, and S. C. Burleigh, "A study of DTN for reliable data delivery from space station to ground station," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 5, pp. 1344–1358, 2024.

[14] C. Caini, T. de Cola, A. Shrestha, and A. Zappacosta, "Ltp performance on near-earth optical links," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 6, pp. 9501–9511, 2023.

[15] J. Liang, X. Liu, R. Wang, L. Yang, X. Li, C. Tang, and K. Zhao, "Ltp for reliable data delivery from space station to ground station in the presence of link disruption," *IEEE Aerospace and Electronic Systems Magazine*, vol. 38, no. 9, pp. 24–33, 2023.

[16] M. S. Net and S. Burleigh, "Evaluation of opportunistic Contact Graph Routing in random mobility environments," in *2018 6th IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*. IEEE, dec 2018.

[17] A. Berlati, S. Burleigh, C. Caini, F. Fiorini, J. Messina, S. Pozza, M. Rodolfi, and G. Tempesta, "Implementation of (o-)cgr in the one," in *2017 6th International Conference on Space Mission Challenges for Information Technology (SMC-IT)*, 2017, pp. 132–135.

[18] The Interplanetary Overlay Network (ION) software distribution, "ION-DTN," https://sourceforge.net/projects/ion-dtn. Retrieved: 04-01-2024.

[19] M. Moy, R. Kassouf-Short, N. Kortas, J. Cleveland, B. Tomko, D. Conricode, Y. Kirkpatrick, R. Cardona, B. Heller, and J. Curry, "Contact multigraph routing: Overview and implementation," in *2023 IEEE Aerospace Conference*, 2023, pp. 1–9.

[20] H. Liang, X. Xu, Y. Li, and Y. Yao, "An optimized Contact Graph Routing algorithm in deep space communication," in *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, 2020, pp. 147–151.

[21] S. Dhara, C. Goel, R. Datta, and S. Ghose, "CGR-SPI: A new enhanced Contact Graph Routing for multi-source data communication in deep space network," in *2019 IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT)*, 2019, pp. 33–40.

[22] D. Ta, R. Menon, J. Taggart, A. Tettamanti, S. Feaser, P. Torrado, and J. Smith, "Roaming DTN: Integrating unscheduled nodes into contact plan based DTN networks," in *2023 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW)*, 2023, pp. 1–9.

[23] F. D. Raverta, R. Demasi, P. G. Madoery, J. A. Fraire, J. M. Finochietto, and P. R. D'Argenio, "A Markov decision process for routing in space DTNs with uncertain contact plans," in *2018 6th IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, 2018, pp. 189–194.

[24] F. D. Raverta, J. A. Fraire, P. G. Madoery, R. A. Demasi, J. M. Finochietto, and P. R. D'Argenio, "Routing in delay-tolerant networks under uncertain contact plans," *CoRR*, vol. abs/2108.07092, 2021.

[25] P. R. D'Argenio, J. Fraire, A. Hartmanns, and F. Raverta, "Comparing statistical, analytical, and learning-based routing approaches for delay-tolerant networks," in *ACM Transactions on Modeling and Computer Simulation*. New York, NY, USA: Association for Computing Machinery, May 2024. [Online]. Available: https://doi.org/10.1145/3665927

[26] R. Lent, "Implementing a cognitive routing method for high-rate delay tolerant networking," in *2023 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW)*, 2023, pp. 1–6.

[27] R. Lent, "Assessing DTN routing performance in the presence of unreliable contacts," in *GLOBECOM 2024–2024 IEEE Global Communications Conference*, December 2024.