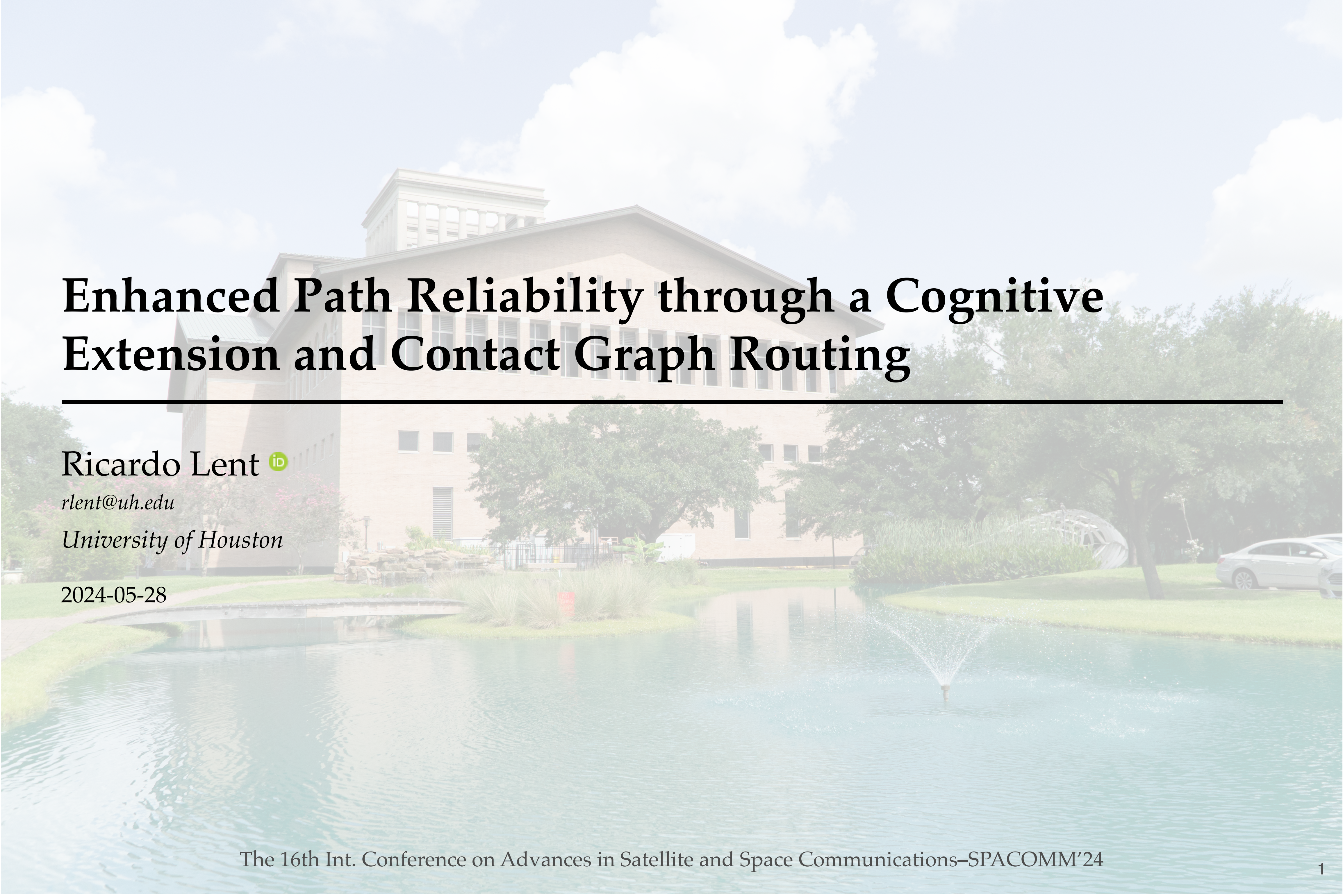


Enhanced Path Reliability through a Cognitive Extension and Contact Graph Routing

---

Ricardo Lent

rlent@uh.edu

University of Houston

2024-05-28

The 16th Int. Conference on Advances in Satellite and Space Communications–SPACOMM’24

Introduction

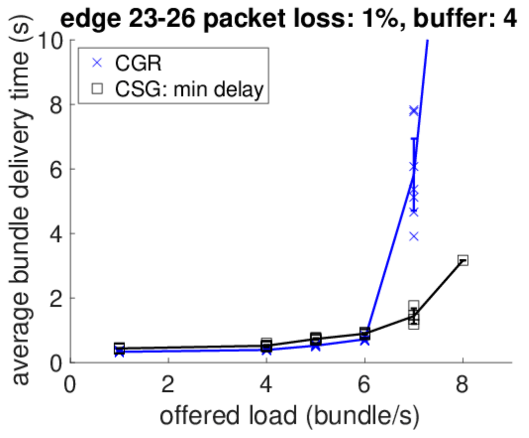
- Space Delay-Tolerant Networks (DTNs) rely on a store-carry-and-forward mechanism to deliver bundles over a dynamic and disconnected network
- The scheduled nature of space DTN allows knowing ahead of time the expected future contacts
- Contact Graph Routing (Schedule-Aware Routing) exploits contact information to discover paths with the shortest bundle delivery times
- Routing “mistakes” involve large costs as link delays are potentially very large

Additional Issues

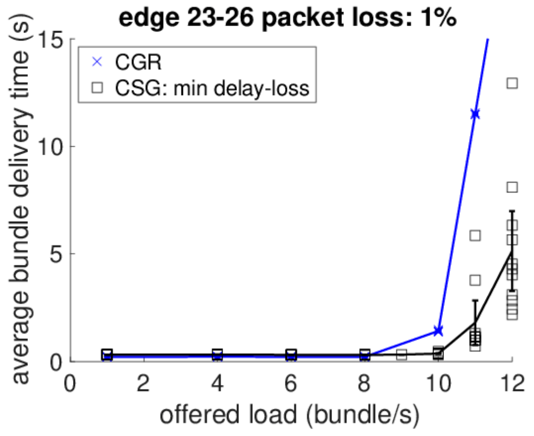
- Bundle transmissions during contacts may be subject to higher signal loss than expected, which brings:
  - Packet losses
  - Buffer overflows
  - Bundle retransmissions (CLA dynamics)
  - Contact overbooking
  - Reduced routing efficacy
- Contacts may not be fully reliable
  - In addition to the effects listed above, the waiting times for the next contact to given neighbor may be longer than predicted

Motivation & Objective

- Prior work demonstrated the advantages of cognitive routing
- Cognitive Space Gateways defines an alternative routing approach to CGR based on reinforcement learning and spiking neural networks
- Key performance enablers are:
  - Accurate bundle delivery time predictions
  - Multi-objective routing
- This work explores whether some of the key aspects of CSG could be implemented in CGR to improve its routing performance



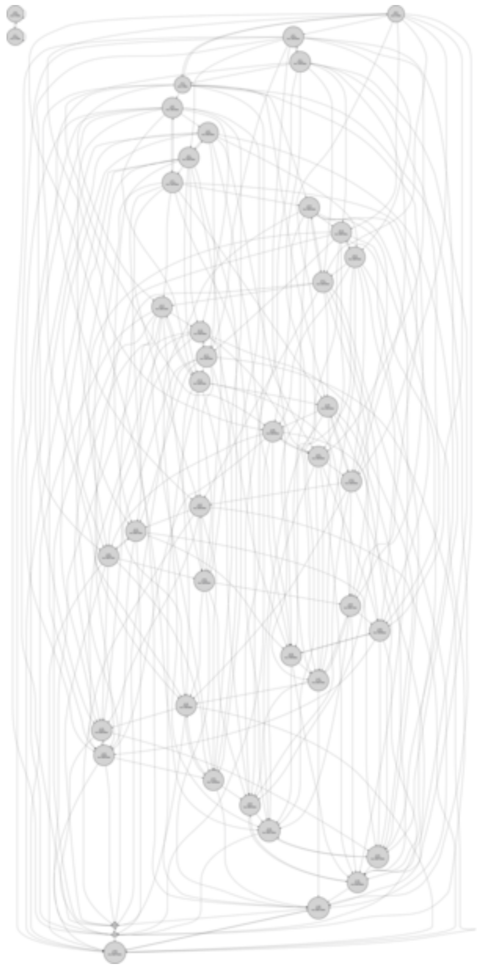
TCP CLA



UDP CLA

Contact Graph Routing

- Relies on a predicted contact plan that is distributed to the network nodes ahead of time
- Contact plan entry: *(source vertex, sink vertex, start time, end time, transmission rate, one-way light time (OWLT), confidence)*
- Builds a contact graph with two auxiliary nodes
- Implements a shortest path graph traversal, usually Dijkstra's with custom link costs that account for the estimated time of bundle arrival
- Extracts the next-hop from the best path



Very Simplified CGR Algorithm

```
function Dijkstra(G, s_node, d_node, t):
    # contact graph: G, current time: t

    s = G.add_contact_vertex(s_node)
    d = G.add_contact_vertex(d_node)
    for vertex in G:
        eta[vertex] = inf
        prev[vertex] = undef

    arrival[s]=0, Q={s}
    while Q is not empty:
        u = vertex in Q with min eta[u]
        remove u from Q

        for v in neighbor(u):
            if cost(u, v) < eta[v]:
                eta[v] = cost(u, v, t)
                prev[v] = u

    return dist, prev
```

```
function cost(u, v, t):
    # contacts: u, v, current time: t

    if S(u) < t:
        return t + owlt(u, v)
    else:
        return S(u) + owlt(u, v)
```

OCGR Variation

- Introduces opportunistic contacts which are associated with a confidence level based on contact history
- In fact, all contacts can be seen as opportunistic
- It finds the k-shortest paths and their confidence
- Paths are considered if within certain confidence bounds
- The time progression aspect used in Dijkstra's remains the OWLT

Cognitive Extension

- The core ideas are to:
  - Implement the `cost` function of CGR/OCGR as a predictive neural network (NN) model or related model
  - Re-interpret the estimated time of arrival as an average value
- The model is essentially a regressor $cost = f_\theta(x)$
  - x is the system state
  - θ is the set of model parameters

Cognitive Extension (cont'd)

- The system state x is defined by:
  - Local information (static and available at the nodes), e.g., CLA type and transmission rates, OWLT, from the contact plan
  - Global information (dynamic), e.g., network buffer occupancy
- Different methods may be used to define the model
- Conventional neural networks based on continuous-activation functions
  - MLP
  - LSTM/GRU
  - Graph neural networks
- Spiking neural networks

Concept Evaluation

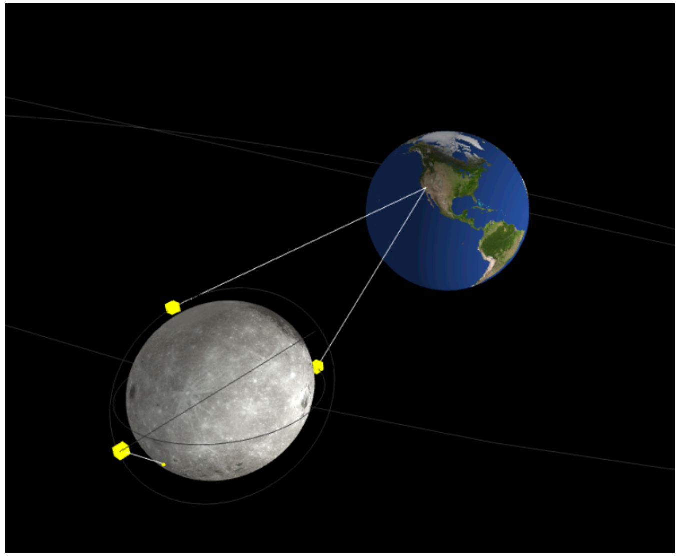
- The choice of the model and definition of the system state will impact the routing performance of the enhanced CGR method
- We keep the model implementation specifics separated from the proposed idea:
  - Ideal predictive performance is obtained from an analytical model
  - Separated local and global availability assumptions for x (CE-A, CE-B)
  - Errors introduced through a probabilistic deviation of the model predictions:

$$cost = \max\{y_{min}, y \times \mathcal{N}(1, \sigma_e)\}$$

where $\mathcal{N}$ is a sample from a normal distribution with unit mean and standard deviation σ_e

Evaluation Network

- Two-phase simulation
- We first extract the contact patterns of an Earth-Moon network for about 4 days:
    - Source located on the far side of the Moon
    - Three lunar orbiters
    - Three ground stations: Goldstone, Madrid, Canberra
    - Sink located in Houston
- Event-driven simulation of packet dynamics provide accurate statistics of bundle arrivals



Note: 10% of the actual distances shown for visualization purposes

Contact Features

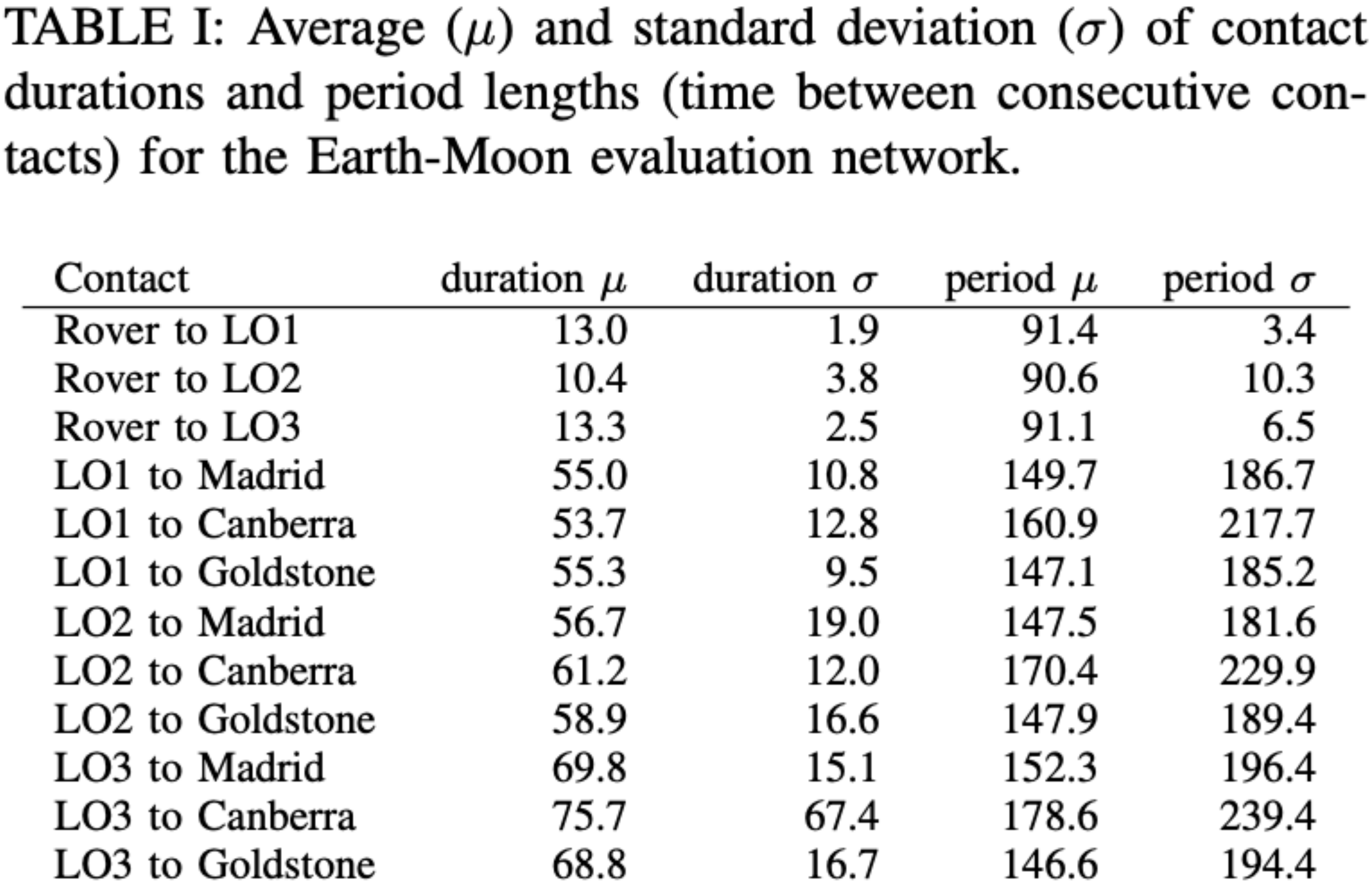


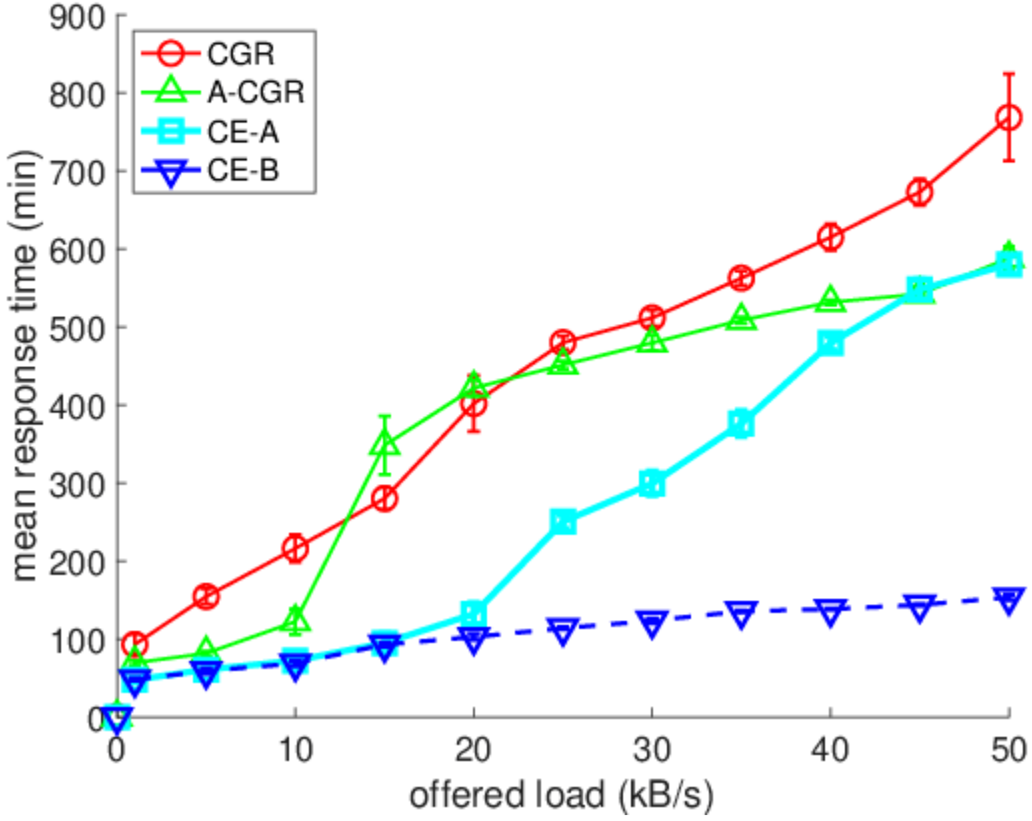
TABLE I: Average (μ) and standard deviation (σ) of contact durations and period lengths (time between consecutive contacts) for the Earth-Moon evaluation network.

Contact	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

Additional Simulation Parameters

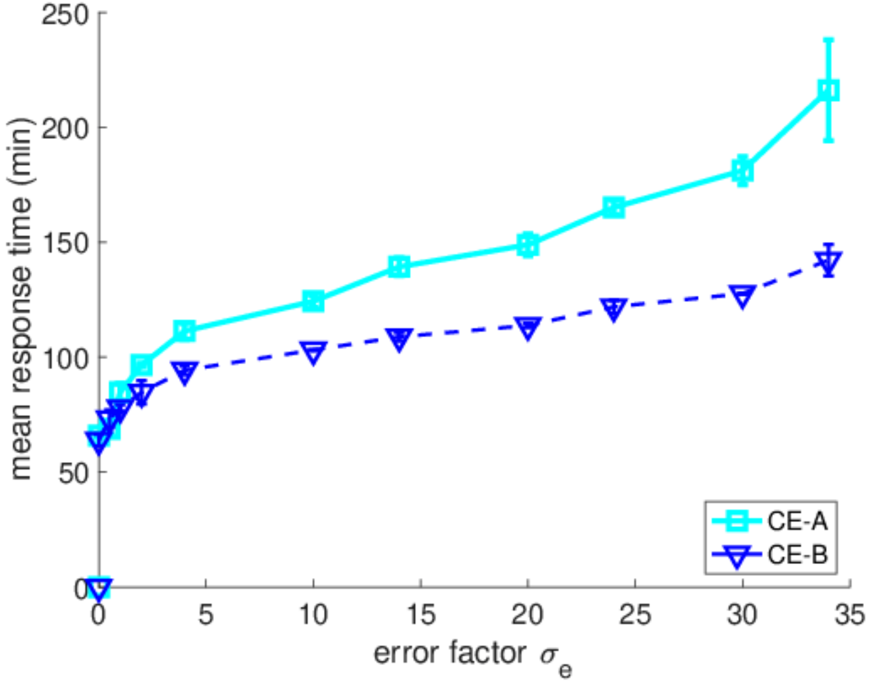
- Wired transmission rates: 2 Mbps
- Wireless transmission rates: 100 Kbps
- Negligible bit-error rates
- OWLT given by the orbital simulation, for wired links, it is given by the as-the-crow-flies distance plus 20%
- All contacts are reliable except for the ones from the orbiters to ground: 0.95, 0.85, 0.5
- The transmission rate was fixed to 0.01 bundle/s but the bundle size was varied to make the offered load an experimental factor

Routing Performance vs. Traffic Load

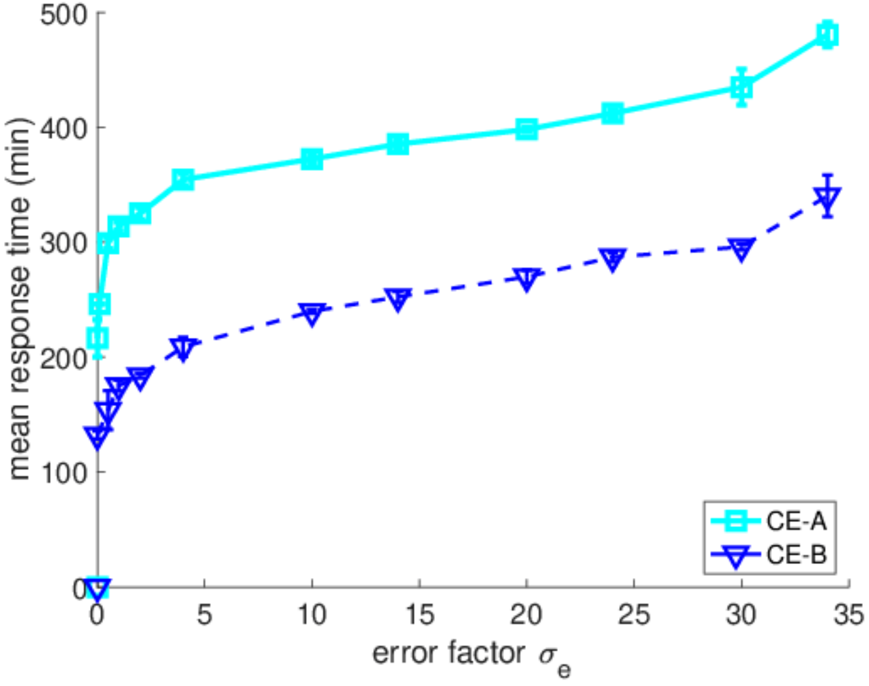


- Independent variable shows the traffic load
- Dependent variable is the average end-to-end bundle delivery time (response time)
- Performance reference given by conventional CGR and A-CGR (CGR excluding low-reliable contacts–below 0.9)
- CE-A: local information
- CE-B: global information

Impact of Link Time Prediction Error



offered load: 10 kB/s



offered load: 30 kB/s

Conclusion

- In this paper, the efficacy of a cognitive extension (CE) for CGR was discussed
- Significant performance improvements in response times were observed of 50% or lower compared to CGR with the use of the CE
- The results emphasized on the role of:
    - The system state, i.e., local vs. global information
    - The prediction accuracy, with 2 to 3 times response time increases with diminishing accuracy
- Future work will address the evaluation of specific prediction models and the experimental evaluation of the concept