National Aeronautics and Space Administration

Development of a Cognitive Delay Tolerant Network Node

Principal Investigator: Marc Sanchez Net (332); Co-Investigators: Pau Buzzi (Texas A&M), Daniel Selva (Texas A&M), Scott Burleigh (JPL, retired)

Program: FY21 SURP **Strategic Focus Area: Disruption tolerant networks**

Objectives

The goal of this SURP is to explore the use of Artificial Intelligence (AI) to automate and eventually optimize monitoring and management of Delay Tolerant Network (DTN) nodes [1]. To achieve this goal, we have identified three specific aims:

1. Formulate the problem of managing a DTN node memory using Reinforcement Learning (RL) and demonstrate that an RL agent can be trained to autonomously adjust input and output data rates, as well as the types of bundles accepted, to avoid memory overflows in the presence of non-stationary traffic conditions;
2. Test the performance of the RL agent using DtnSim and Openai Gym [2] in a realistic cis-lunar network scenario and compare it to the performance of alternative agents based on static rule-based expert systems, and control theory.

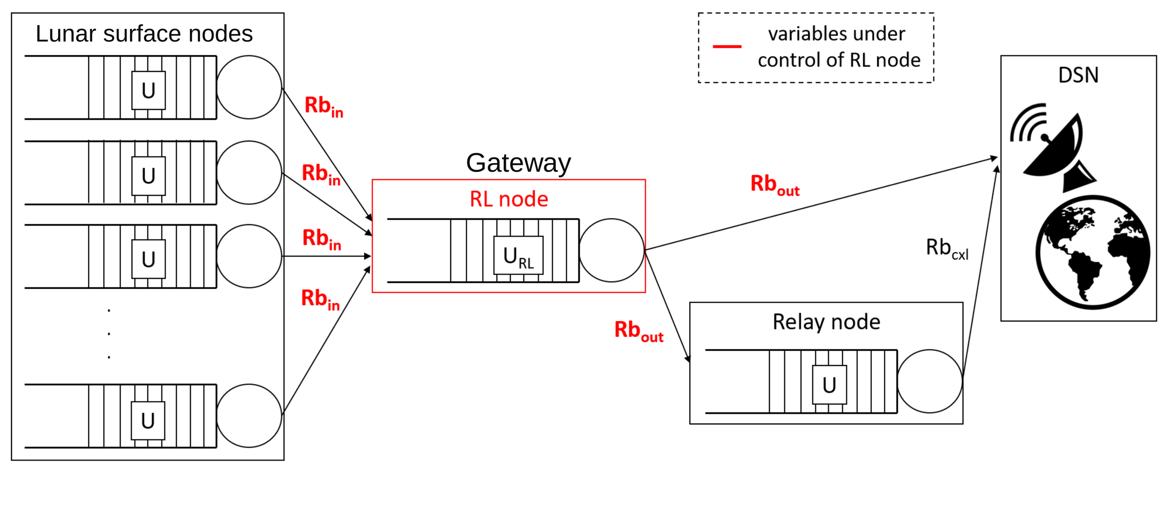

Figure 1. Illustration of the Delay Tolerant Network of the Lunar mission used as training/test scenario, consisting of Moon-to-Earth network with 46 nodes and two activities that generate more than 2,500 traffic flows, most of which originate at the lunar surface and need to be delivered to Earth. . The RL node is highlighted in red. Variables controlled by the RL node are in red bold font.

3. Integrate the control agent with ION, the reference flight implementation of DTN being deployed on the DSN and in spacecraft such as KPLO and Lunar IceCube.
4. Test the developed RL agent in an emulated environment using Raspberry Pis communicating over an ad-hoc mesh network (i.e., direct and unstable peer-to-peer communication links) using the Interplanetary Overlay Network (ION) software (downscoped during FY21 due to COVID pandemic).

Background

- DTN is a new set of communication protocols developed with the intent of extending the Internet to users that experience long delays and/or unexpected disruptions in their link service, which is common in deep space operations. While core protocols of DTN are quite mature (e.g., Bundle Protocol), how to configure and manage a DTN remains an area of active research.
- We propose to extend the capabilities of DTN by automating its monitor and control system via an intelligent agent that interacts with a DTN <u>without</u> human intervention: a **Cognitive Delay Tolerant Node (CDTN).** This intelligent agent is capable of perceiving changes in the node's and/or network's state (e.g., processor load, memory usage) and trigger actions that optimize its performance or protect it against anticipated failures.
- Expect to reduce the need for constant network monitoring and operation in JPL-maintained infrastructure such as the Deep Space Network or Mars Relay Network.

Approach

1) Reinforcement Learning (centralized controller over cislunar network, see Figure 1)

<u>ACTION SPACE</u>: 7 types of actions: (1) Drop packets, (2/3) increase/decrease peer's radio data rate (4) send packets to crosslink, (5) increase DSN/crosslink downlink data rate, (6) decrease DSN/crosslink downlink data rate, and (7) do nothing.

<u>STATE SPACE</u>: Defined by (1) the memory state of the RL node its network peers, (2) the data rate of all proximity links (assumed equal across all surface users for simplicity), (3) the data rate to the crosslink, and (4) the data rate of the DTE link with the DSNs:

$$\vec{s} = [\mathrm{u}, \max_{i=1,\dots,N} u_i, Rb_{in}, Rb_{out}]$$

<u>REWARD FUNCTION</u>: Defined to maximize the number of bits that go through the Gateway and arrive at their intended destination, as well as minimizing the cost of all links controlled by the gateway. The function f(s) aims to minimize buffer utilization to avoid memory overflow by driving the reward to 0 when any node in the network gets congested.

$$R(s,a) = f(u) \cdot f\left(\max_{i=1,\dots,N} u_i\right) \cdot \frac{\# \, bits_{gwy \to dest}}{cost_{X \to gwy} + cost_{gwy \to DSN} + cost_{gwy \to cxl}}$$

$$E_{\text{link}} = \int P(t)\mathrm{d}t = \int E_b \cdot R_b(t)\mathrm{d}t \propto \int R_b(t)\mathrm{d}t = \text{cost}_{\text{link}}$$

2) Control Theory (decentralized controller)

A feedback control loop was designed to monitor and control the contact plan data rate of a single link to avoid memory overflows. Unlike the RL model, which uses a trained neural network to encode the control policy, this control loop can be studied analytically and provides guaranteed memory overflow protection.

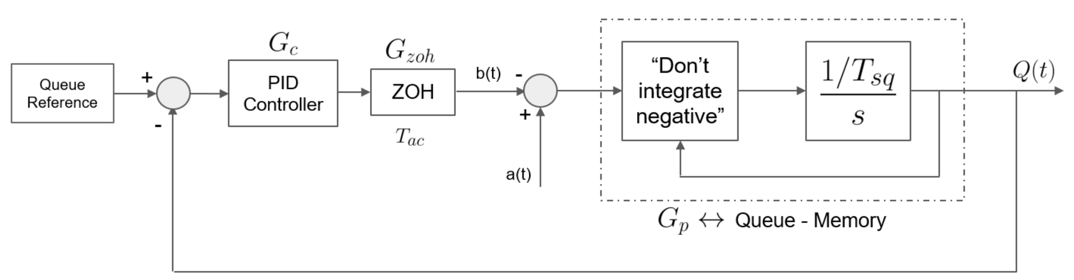

Figure 2. Formulation of a PID controller to control the memory state of a DTN node

Results

Tests were conducted using a simulated environment modeling a cislunar DTN with more than 1,500 traffic flows of different type (telemetry, video, voice, biomedical data). We showed that:

1) An RL agent successfully manages the resource allocation and buffer utilization of a cislunar DTN network, dynamically adapting to varying traffic flows and minimizing buffer overflows. When overflow was inevitable, bundles of low priority were discarded.

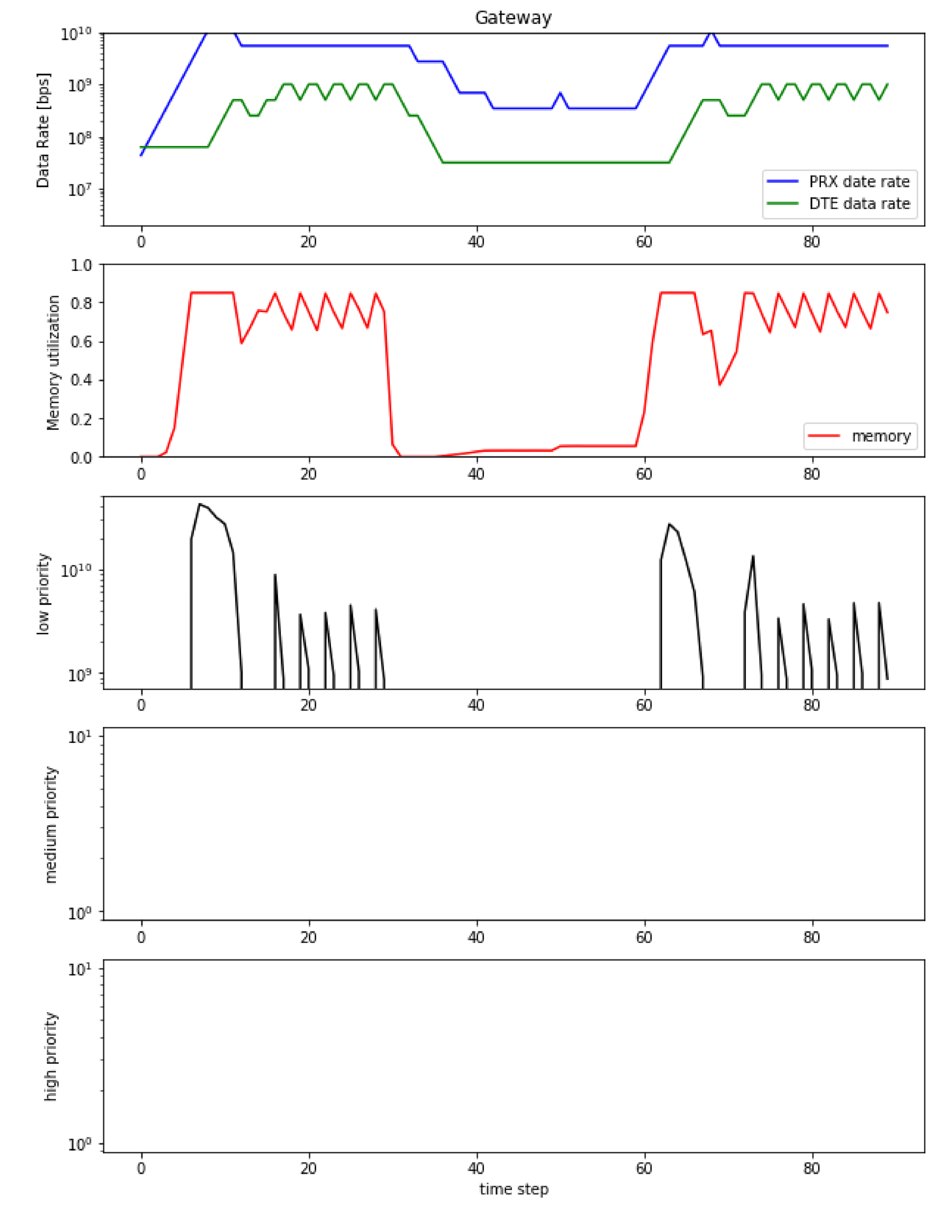

Figure 3. Progression over time of the data rate of all links transmitting bundles to the Gateway [bps] and the data rate of the downlinks with the DSNs [bps] (first), the memory utilization over time of the Gateway (second), and the number of bits dropped for three different levels of priority (third, fourth and fifth)

2) A PID controller can successfully manage a single DTN node under non-stationary traffic conditions. We can observe that changes in the input traffic in the DTN node trigger appropriate data rate changes to avoid memory overflows and control the transients.

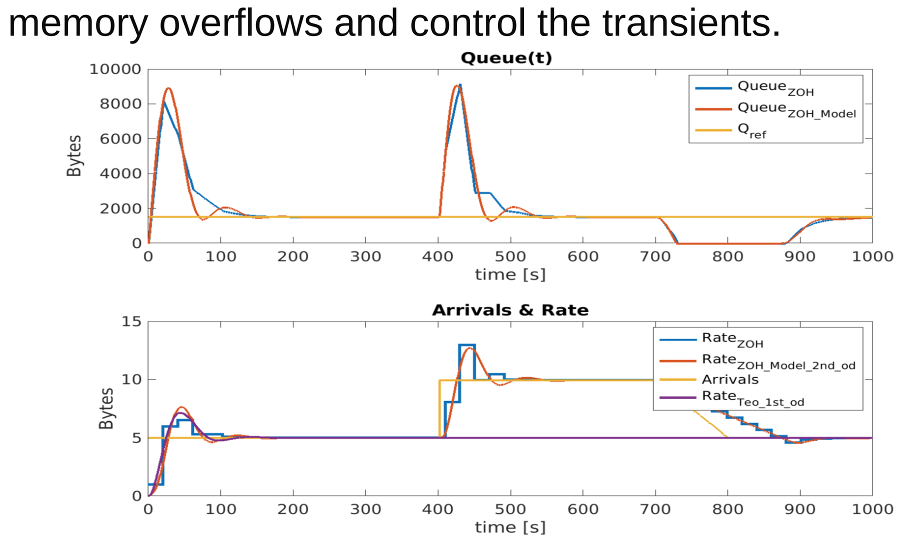

Figure 4. Evolution over time of the memory and output data rate in a DTN node controlled by a PID

A new version of pyion (interface between ION and Python) was released, which includes experimental features that allow network management primitives to be executed in ION from our RL agent or control-theoretical agent.

Significance/Benefits to JPL and NASA

1. Deep space flight autonomy, which is an area of active research at JPL given that our spacecraft can rarely have human-in-the-loop intervention;
2. DSN and MRN network operations, which are infrastructure-like projects which are already embracing automation and autonomy (e.g., follow-the-Sun DSN operations) to improve efficiency.

Publications

A. Garcia Buzzi, P., Selva, D., and Sanchez-Net, M., "Autonomous Delay Tolerant Network Management Using Reinforcement Learning," 2020 AIAA ASCEND, AIAA, 2020.
B. Garcia Buzzi, Pau, Daniel Selva, and Marc Sanchez Net. "Autonomous Delay Tolerant Network Management Using Reinforcement Learning." Journal of Aerospace Information Systems.
C. Garcia Buzzi, P., Selva, D., and Sanchez-Net, M., "Deep Q-learning for Delay Tolerant Network Management in a Reactive Imaging Earth Observing Constellation", 2021 AIAA ASCEND, AIAA, 2021.

References

1. Brockman, Greg, et al. "Openai gym." arXiv preprint arXiv:1606.01540 (2016).
2. Nag, Sreeja, et al. "Autonomous scheduling of agile spacecraft constellations with delay tolerant networking for reactive imaging." arXiv preprint arXiv:2010.09940 (2020).

PI/Task Mgr Contact
Email: marc.sanchez.net@jpl.nasa.gov

Poster No.
Clearance No.
JPLTask#SP20031

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov