Ad Hoc Networks 25 (2015) 303–313

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Routing-aware fair contact plan design for predictable delay tolerant networks

J. Fraire *, J.M. Finochietto

Digital Communications Research Lab, Universidad Nacional de Córdoba, CONICET, Argentina

ARTICLE INFO

Article history:
Received 19 February 2014
Received in revised form 15 July 2014
Accepted 23 July 2014
Available online 1 August 2014

Keywords:
Delay tolerant networks
DTN
Contact plan design
Satellite constellations

ABSTRACT

Delay tolerant networks (DTNs) have become a promising solution for extending Internet boundaries to challenged environments such as satellite constellations. In this context, strategies to exploit scarce communication opportunities, while still considering device and application constraints, are still to be investigated to enable the actual deployment of these networks. In particular, the Contact Graph Routing (CGR) scheme has been proposed as it takes advantage of the contact plan, which comprises all future contacts among nodes. However, resource constraints can forbid the totality of these contacts to belong to the contact plan; thus, only those which together meet an overall goal shall be selected. In this article, we consider the problem of designing a contact plan that can provide fairness in link assignment and minimal all-to-all route delay; therefore, achieving equal contact opportunities while favoring end-to-end traffic latency. We formalize this by means of a multi-objective optimization model that can be computationally intractable for large topologies; thus, heuristic algorithms are proposed to compute the contact plan in practice. Finally, we analyze general results from these routines and discuss how they can used to provision valuable contact plans for real networks.

1. Introduction

Delay-Tolerant Networks (DTNs) have received much attention during the last years as they have been proposed for several environments where communications can be challenged by either latency, bandwidth, errors, or stability issues [1]. Even if originally studied to develop a network architecture for the Interplanetary Internet (IPN) [2], DTNs have been recently recognized as an alternative solution for building future satellite applications [3]; in particular, to cope with typical intermittent channels of LEO (Low Earth Orbit) constellation systems [4].

Among the challenges to implement practical DTNs, the definition of a new communication protocol, which does not assume a persistent connectivity between the communication end points, has been addressed by the specification of the Bundle protocol [5], resulting in the availability of several software implementations of the protocol [6,7]. Indeed, traditional protocols like TCP cannot be used for end-to-end communication due to their conversational nature. Neither can traditional routing protocols be employed on environments with intermittent connections, where the network topology is time-varying. To this end, new routing mechanisms have been recently investigated [8–11]. In particular, if topology changes are predictable as in LEO environments [12], the Contact Graph Routing (CGR) [13] scheme is appealing as it takes advantage of the a priori knowledge of the *contact plan* between DTN nodes.

In general, a *contact* can be defined as the opportunity to establish a temporal communication link among two DTN nodes. However, it is possible that a given node may

* Corresponding author. Tel./fax: +54 351 4334147.
E-mail addresses: juanfraire@gmail.com (J. Fraire), jfinochietto@efn.uncor.edu (J.M. Finochietto).

http://dx.doi.org/10.1016/j.adhoc.2014.07.006
1570-8705/

have more than one contact opportunity at a given time but limited or conflicting resources to only make use of one of these opportunities. Furthermore, operational constraints such as power budgets, agency policies, among others might arise. As a result, the contact plan can be thought as a subset of the *contact topology* which comprises all contact opportunities a network of DTN nodes has over a given time window. The design of contact plans has still received little attention, as it is typically assumed that all potential contacts between DTN nodes can belong to the contact plan; in other words, that the contact plan equals the contact topology. Early works [14,15] have focused on contact plans that can enhance the network connectivity for a given topology state, without taking into account the time-evolving nature of the contact topology which enables the transport of data through multiple states. With the advent of DTNs, the design of contact plans that consider time-evolving topologies became much relevant as traffic data can traverse different topology states. Recent works [16,17] have proposed the design of contact plans that can minimize path costs and maximize reliability; however, to the best of our knowledge nor fairness issues neither end-to-end latency have been addressed in the design of contact plans.

In this work, we significantly extend our recent work [18] in order to investigate the design of routing-aware fair contact plans, where all DTN nodes shall have the best possible access to available links among them while optimizing all-to-all route delays. Fairness can be achieved over time by considering the time-evolving nature of the network topology as well as previous link assignments, while route delays can be optimized in a subsequent heuristic stage. Our goal is to compute proper contact plans that can provide equal opportunities to all DTN nodes for the purpose of exchanging data traffic in both direct and routed neighbors. To this end, we formalize this problem, and then propose and evaluate computational efficient algorithms that can be used to design these contact plans.

The paper is structured as follows. Section 2 describes the contact topology and the contact plan models on which fairness and routing criteria metrics are based. We formalize the problem by means of multi-objective optimization problem in Section 3, for which novel algorithmic alternatives are proposed in Section 4. Section 5 evaluates the performance of these strategies in terms of fairness and routing metrics in a general and particular case study respectively. Finally, Section 6 concludes the work and discusses future directions.

2. System model

Nodes from a satellite constellation exhibit a concrete physical trajectory (i.e., orbits), where communication links become both sporadic and foreseeable. In other words, since their position and attitude can be accurately forecasted by precise analytic or numerical models, the same remains true for communication contacts between them. As a result, such constellations can be considered predictable DTNs, where traffic flows in a store, carry, and forward fashion before reaching its destination.

2.1. Contact topology

The aforementioned contact forecast is rendered in the *contact topology*. In order to illustrate the latter, we consider the case of a satellite network with 4 nodes as shown in Fig. 1(a). Here, all nodes are equipped with transponders and antennas enabling cross-links near the pole, while only contacts between nodes 1 and 2, and nodes 3 and 4, are feasible on the equator area. The time-evolving nature of these links can be captured by means of graphs [19], capable of symbolizing links availability over time. This representation can be thought as a finite state machine (FSM) in which each state is characterized by a graph whose arcs, in turn, represent a feasible communication (i.e., contact) between nodes during a period of time. Each state can be identified by $k = 1, 2, \ldots K$ conforming K graphs that comprises the same set of nodes but different arcs among them. Particularly, in the suggested scenario, 3 states can describe the contact topology, representing the communication link evolution during half an orbit time frame.

In particular, a contact topology consists of $p_{k,i,j}$ links between node i and j at state k, where $p_{k,i,j}$ may adopt an integer identifier related to the communication interface. If no contact is feasible, then $p_{k,i,j} = 0$, and $p_{k,i,j} = 1$ if the contact among i and j is possible. For instance, we could state that $p_{1,1,2} = p_{1,2,1} = 1$ and $p_{2,2,3} = p_{2,3,2} = 1$, implying the feasibility of contacts via the pole and equator communication interfaces. Besides, at state $k = 1,\ p_{1,2,3} = p_{1,3,2} = 0$ since no physical link exist between nodes 2 and 3. In general, the contact topology can be defined by a three dimensional *physical* adjacency matrix $[P]_{k,i,j}$.

Finally, the contact topology can be configured to satisfy strict system limitations such as forbidden transmission over a specific geographical area, or specific time interval where interference to geostationary can be provoked, or even operational issues like specific agency policies among others. In the other hand, resource-dependent constraints with impact in the maximum number of interfaces are the focus of research of this article and defined in Section 2.2.

2.2. Contact plan

The contact topology assumes no resource limitation on nodes as it is based only on the feasibility of physical communication. However, nodes may only be able to make use of limited contacts at a given time, even if more are potentially available. This can be due to power budget limitations, hardware designs, interference requirements, among others. As a result, the maximum number of simultaneously active interfaces can be thought as a restriction to the contact topology $[P]_{k,i,j}$. The problem then lies in selecting the most appropriate set of possible contacts in a time-evolving network with a certain criteria subject to restrictions or limitations. Consequently, we define an integer *contact* matrix $[C]_i$ constituted by the maximum c_i contacts a node i can simultaneously maintain. For example, if we