Performance Evaluation 115 (2017) 108–131

Contents lists available at ScienceDirect

Performance Evaluation

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

A theoretical analysis of buffer occupancy for Intermittently-Connected Networks

L. Boero [c], M. Cello [a,*,1], G. Gnecco [b], M. Marchese [c], F. Patrone [c], M. Sanguineti [d]

[a] Nokia Bell Labs, Blanchardstown Business & Technology Park, Snugborough Road, Dublin 15, Ireland
[b] IMT-School for Advanced Studies, Piazza S. Francesco, 19, 55100 Lucca, Italy
[c] University of Genoa, DITEN Department, Via all'Opera Pia 13, 16145, Genova, Italy
[d] University of Genoa, DIBRIS Department, Via all'Opera Pia 13, 16145, Genova, Italy

ARTICLE INFO

Article history:
Received 19 October 2016
Received in revised form 1 June 2017
Accepted 7 August 2017
Available online 30 August 2017

Keywords:
Intermittently-Connected Networks
Congestion control
Ad-hoc networks
Markov chains
Epidemic routing

ABSTRACT

Network congestion is a well-known problem that may heavily affect the overall network performance. Congestion control approaches in Intermittently-Connected Networks (ICNs) differ from those used in classical networks, since the assumptions of "universal connectivity" of the nodes and "global information" about the network do not hold. In this paper, an analytical framework is proposed to investigate node buffer occupancy in ICNs through bulk-arrivals/bulk-services queuing models. A relation in the z-domain between the discrete probability densities of the buffer state occupancies and of the sizes of the arriving bulks is exploited to analyze two classes of forwarding strategies for ICNs. The infinite- and finite-buffer cases are investigated, simulated, and compared in terms of the concept of stochastic order, which is also used to compare models obtained for different parameter choices. The results can be exploited for buffer dimensioning and for deriving estimates of performance metrics such as average buffer occupancy, average delivery delay, and buffer overflow probability. The theoretical analysis is complemented by numerical outcomes from a network simulator and from real mobility traces.

1. Introduction

Various applications recently emerged, which operate under network conditions where the assumptions of "universal connectivity" of the nodes (i.e., the constant presence of an end-to-end path between any two nodes) and "global information" about the network (i.e., the knowledge of the state of the whole network) do not hold. Examples are sensor networks [1], social networks and pocket switched networks [2,3], smart environments, and vehicular ad-hoc networks [4]. A common denomination is *Intermittently-Connected Networks* (*ICNs*) [5]. As in such contexts the networks may be disconnected most of the time and it may even happen that an end-to-end path between source and destination is never available, classical routing and data delivery-approaches (see, e.g., [6]) may fail [7]. Indeed, node mobility in ICNs causes frequent disconnections, then reducing the average path duration significantly [8]. Consequently, in many cases, the time needed to repair a broken path is close to or even larger than the average path duration, and the expected throughput of reactive protocols (i.e., protocols that compute a route only when it is needed) is very small [9]. Other approaches to routing in ICNs involve the use of additional communication resources (e.g., satellite and UAV) forced to follow a given trajectory between disconnected parts

---

[*] Corresponding author.

E-mail addresses: luca.boero@edu.unige.it (L. Boero), marco.cello@nokia-bell-labs.com (M. Cello), giorgio.gnecco@imtlucca.it (G. Gnecco), mario.marchese@unige.it (M. Marchese), f.patrone@edu.unige.it (F. Patrone), marcello.sanguineti@unige.it (M. Sanguineti).

[1] The work has been performed while M. Cello was employed at the University of Genoa.

http://dx.doi.org/10.1016/j.peva.2017.08.001
0166-5316/

of the network, in order to bridge the gaps (DataMule, Message Ferries, etc.) [10,11]. In other cases, such as inter-planetary networks [12], intermittent connectivity is often predictable, hence classical routing algorithms may be adapted to compute shortest delivery delay paths by taking into account future connectivity [13].

Often, neither additional resources with controlled behavior nor predictable trajectories and connectivity are available. In such cases, one of the most common approaches is *epidemic routing* [14], which is based on the replication and transmission of messages to newly-discovered contacts that do not already possess a copy of the message. Existing epidemic protocols try to avoid congestion by limiting, either in a deterministic or in a non-deterministic way, the number of copies of a message inside a network. Concerning the first case, we mention *spraying algorithms* (Spray and Wait, Binary Spray and Wait, and Spray and Focus) [15], in which each generated message can be replicated at most an a-priori fixed number L of times. As regards nondeterministic approaches, the replication of a copy with some probability [16–18] and intelligent filtering replication strategies using history-based or utility-based routing [1,5,19–23] have been investigated. In these approaches, every node maintains a utility value for each other node in the network. Such a value contains indirect information about, e.g., relative node locations, likelihood to encounter the destination node, etc.

In general, message replication performed by epidemic routing paradigms imposes a high storage overhead on wireless nodes [24] and very likely node buffers run out of capacity. So, an analytical framework for congestion control management is highly recommendable. Besides an investigation of performance metrics related to the behavior of the node buffers (e.g., average buffer occupancy and its standard deviation, buffer overflow probability), it is important for such an analytical framework to model also their trade-off with respect to other performance metrics. Among these metrics we mention the average time to destination of the first copy of a message reaching the destination node (average delivery delay), under different forwarding strategies, considering also the effect of changes in their parameters, which can be useful for their optimization.

Providing the analytical framework described above is the main subject of the present work. Buffer management strategies were proposed and experimentally evaluated in [25,26] (e.g., Drop-Random, Drop-Oldest, Evict shortest life time first, etc.) and others. In [27], two buffer management strategies, called HBD (History Based Drop) and FBD (Flood Based Drop), were proposed and investigated both theoretically and numerically. A peculiar feature of these two strategies is that they do not require a global knowledge of the state of the network. A batch of works derives performance bounds for epidemic routing and its variations from an analytical point of view. In [28], the authors investigate the average delivery delay for flooding by exploiting the Susceptible–Infectious–Recovered (SIR) model [29]. In [18], a model is developed to analyze the average delivery delay and its relative trade-offs with energy consumption and buffer requirements in the so-called (p, q)-*epidemic routing*, where p and q represent, respectively, the probability that a node accepts a packet copy from another node when none of the two nodes is the source and the probability that a node accepts a packet copy from the source node of the packet. Deactivation-controlled and lifetime-controlled epidemic routing are presented, respectively, in [30] and in [31]. In [32], an analytical framework is proposed to analyze the expected delay for some routing schemes in opportunistic networks under contention. In [33], the authors develop a mathematical framework based on a Markov chain to get insights into the global congestion behavior (through the computation of parameters, such as number of drops, number of replications, etc.) and propose a local congestion control strategy. Buffer dimensioning issues are investigated in [34] and [35], considering forwarding strategies different from the ones of this paper.

In this paper, we focus on single nodes modeled through Markov chains whose states are the number of packets in node buffers, and on the behavior of the nodes when the size of the incoming data bulk (i.e., its number of packets) varies. Differently, most literature studies the underlying Markov chain that models the number of copies of a single packet in the network. An advantage of our approach is that it provides information not only on the average buffer occupancy, but also on its standard deviation and on the probability that the buffer occupancy is above a given threshold. Using the concept of stochastic order, our model makes feasible an extension to the finite-buffer case of several results obtained for infinite buffers, providing for the former bounds on the average buffer occupancy and its standard deviation, on the buffer overflow probability, and on the average delivery delay. In the latter two cases our model takes into account the size of the bulk, too. We also exploit stochastic order arguments to relate models obtained for different choices of the parameters. All such properties derive from focusing on the behavior of node buffers, instead than on single packets. We apply the framework to two specific well-known forwarding protocols (i.e., q-forwarding and two-hop forwarding), showing how one can optimize their parameters taking into account the trade-off between different performance metrics. This work extends substantially our papers [36] and [37], providing additional contributions in

- examining a larger variety of performance metrics;
- relating (in terms of the stochastic order) the infinite- and finite-buffer cases, and models obtained for different choices of their parameters;
- optimizing the trade-off between different performance metrics (this issue is particularly important, as it allows one to choose the best parameters for the forwarding strategies investigated);
- comparing the theoretical results with numerical outcomes from a network simulator and from real mobility traces, validating the results of our analysis (particularly, some upper bounds on performance metrics).

The paper is organized as follows. Section 2 presents an analytical framework, based on bulk-arrivals/ bulk-services queues, to model ICN nodes behavior and to capture the main feature of ICNs, i.e., their intermittent connectivity. In Section 3, we express the z-transform of the stationary discrete probability density of the state occupancy of each ICN node buffer,