Computer Communications 71 (2015) 129–141

Contents lists available at ScienceDirect

Computer Communications

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

Cellular traffic offloading via opportunistic networking with reinforcement learning

Lorenzo Valerio*, Raffaele Bruno, Andrea Passarella

Institute for Informatics and Telematics, National Research Council. Via G. Moruzzi 1, 56214 Pisa, Italy

ARTICLE INFO

Article history:
Received 5 June 2015
Revised 1 September 2015
Accepted 3 September 2015
Available online 11 September 2015

Keywords:
Cellular traffic offloading
Opportunistic networking
Reinforcement learning
Actor–critic
Q-learning

ABSTRACT

The widespread diffusion of mobile phones is triggering an exponential growth of mobile data traffic that is likely to cause, in the near future, considerable traffic overload issues even in last-generation cellular networks. Offloading part of the traffic to other networks is considered a very promising approach and, in particular, in this paper we consider offloading through opportunistic networks of users' devices. However, the performance of this solution strongly depends on the pattern of encounters between mobile nodes, which should therefore be taken into account when designing offloading control algorithms. In this paper we propose an adaptive offloading solution based on the Reinforcement Learning framework and we evaluate and compare the performance of two well known learning algorithms: Actor–Critic and Q-Learning. More precisely, in our solution the controller of the dissemination process, once trained, is able to select a proper number of content replicas to be injected in the opportunistic network to guarantee the timely delivery of contents to all interested users. We show that our system based on Reinforcement Learning is able to automatically learn a very efficient strategy to reduce the traffic on the cellular network, without relying on any additional context information about the opportunistic network. Our solution achieves higher level of offloading with respect to other state-of-the-art approaches, in a range of different mobility settings. Moreover, we show that a more refined learning solution, based on the Actor–Critic algorithm, is significantly more efficient than a simpler solution based on Q-Learning.

1. Introduction

Over the last few years we have witnessed an exponential growth of data traffic in cellular networks. On the one hand this is due to the exponential spreading of mobile devices, such as smartphones and tablets, with multiple heterogeneous wireless interfaces. On the other hand, the diffusion of content-centric services among mobile users, e.g. Netflix, Youtube or Spotify, is triggering a strong increase in the amount of traffic carried by the cellular networks. It has been recently estimated that this huge traffic growth will accelerate in the near future [1]. CISCO forecasts that mobile data traffic will grow at a Compound Annual Growth Rate (CAGR) of 61% between 2014 and 2019, reaching 15.9 Exabytes per month by 2019 [1], resulting in the so called *data tsunami*. The adoption of 4G technologies is giving a significant boost to the cellular capacity with respect to the current demands. However, projections over the next few years show that this additional capacity may be saturated soon, as it is foreseen to grow only by a factor of 1.4 by 2019, which may not be sufficient to cope with the data tsunami effect. As a consequence, bandwidth crunch events similar to those occurred with 3G networks [2] may be expected, unless alternative solutions are devised in time.

A promising solution to reduce the burden on the cellular infrastructure is to divert part of the traffic from the cellular network, through *data offloading* mechanisms [3–6]. Data offloading, indeed, is considered one of the most promising techniques to complement pure cellular networks and cope with the amount of mobile data traffic expected in the next few years [7–9] (see Section 2 for a brief survey on the main offloading approaches presented in the literature).

In this paper we consider content whose delivery is delay tolerant and the offloading is based on a device-to-device (D2D) approach. In this context Opportunistic Networks offer a very powerful alternative to relieve part of the network traffic from the cellular infrastructure. Opportunistic networks [10–13] are self-organising mobile networks where the existence of simultaneous end-to-end paths between nodes is not taken for granted, while disconnections and network partitions are the rule. Opportunistic networks support multi-hop communication by temporarily storing messages at intermediate nodes, until the network reconfigures and better relays (towards the final destinations) become available.

The main research issues in Opportunistic Networks focus on the development of analytical models of data delivery performance

* Corresponding author. Tel.: +39 50 315 3059.
E-mail addresses: lorenzo.valerio@iit.cnr.it, lorenzo.valerio@gmail.com (L. Valerio), raffaele.bruno@iit.cnr.it (R. Bruno), andrea.passarella@iit.cnr.it (A. Passarella).

http://dx.doi.org/10.1016/j.comcom.2015.09.004
0140-3664/

[14–16], routing approaches that consider nodes' aggregation and privacy [17,18], forwarding schemes for hybrid networks (opportunistic networks combined with the infrastructure [19]), real world implementations [20], applications to Vehicular Networks [21,22].

Offloading through opportunistic networks is particularly appealing when the same content is requested by multiple users in a limited geographical area in a non-strictly synchronised way (which rules out the possibility of multicasting on the cellular network). Given the typical Zipf nature of content interest [23], this scenario is very relevant. When offloading is used, content can be sent via the cellular network only to a small fraction of interested users (typically called *seeds*), while the majority of them can be reached via opportunistic networking techniques [14,24–26]. In addition, note that recent results also indicate that opportunistic offloading can work very well in cooperation with cellular multicast in case of synchronised requests [27].

One of the known challenges of pure opportunistic networks is that it is very difficult to guarantee maximum content delivery deadlines, because of the random nature of contacts between nodes (enabled by users' mobility) on which they are based. Even in case of delay-tolerant traffic, it is realistic to assume that content would need to be delivered within a maximum deadline. This is, for example, supported by recent studies of Web traces, which show that the value of requested Web content drastically diminishes if not received within a certain time from the request [23].

In this paper (as in our previous work [28] and in several previous papers [29,30]), we consider an operator-assisted offloading system. Specifically a central dissemination controller decides dynamically over time to which nodes content must be sent through the cellular network, based on the current status of the dissemination process. This is tracked through ACKs sent by nodes (over the cellular network) upon receiving content (over the opportunistic network). Injections of content to specific nodes through the cellular network help to boost the dissemination process over the opportunistic network, and can also be used as a last-resort option to send content to nodes that have not yet received it through the opportunistic network by the delivery deadline. As we show in the paper, using such an approach guarantees maximum delivery deadlines, and reduces significantly the traffic carried by the cellular network at the same time.

In this paper we design a new solution to control the content dissemination process described before. Precisely, the controller has to cope with two problems: (i) identifying how many nodes should be delegated to disseminate a piece of content and (ii) which of them are the most useful to speed up the dissemination process. As far as point (i) is concerned, due to high dynamicity of the opportunistic dissemination process, it would be desirable that the central dissemination controller would be able to cope with this problem autonomously, in order to reduce as much as possible planning and fine tuning of algorithms' parameters, or previous knowledge about the behaviour of the network nodes. As explained in more detail in the following, we address the first problem through a Reinforcement Learning (RL) approach. As far as point (ii) is concerned, we adopt a heuristic mechanism that permits to identify, online, what nodes are the most useful to spread contents.

We cast the offloading problem as a RL problem, in a very general way such that we can apply and try different RL learning algorithms. Specifically, in our set-up a learning agent (the controller) interacts with an unknown environment (the opportunistic network). For each content, at fixed time steps, the agent observes the state of the *system* (i.e. the current diffusion of the content), takes an *action* (i.e. decides whether to inject new content replicas) and receives a *reward*, i.e. a feedback to evaluate the action taken. Its goal is to learn a *policy* (i.e. a mapping from states to actions) that maximises the long term reward.

To demonstrate the generality of our approach we also present two instances of our RL framework by using two well-known and very robust solutions: the Actor–Critic and the Q-Learning algorithms [31,32]. These two approaches represent the trade-off between convergence speed and precision. In fact, we find that the Actor–Critic learning algorithm, although slightly slower in the learning phase, has better performance than the Q-Learning based approach. Conversely, Q-Learning can be very responsive in the learning phase but its final accuracy (after the learning phase is over) is worse.

Through a comprehensive set of simulative experiments we will show that solutions based on RL are in general more efficient than other state-of-the-art approaches. Specifically, with respect to the state-of-the-art benchmarks, our solution achieves an increased offloading gain (the percentage of offloaded traffic) of about 20%. Moreover, the comparison between the Actor–Critic and the Q-Learning algorithm shows that the former can achieve higher offloading gain of about 17%. Finally, our results show that a solution based on RL does not require pre-existing knowledge about the nodes behaviour (e.g., degree on the contact graph), but is able to automatically learn the most appropriate control strategy for the offloading task.

This paper extends our previous work in [28]. With respect to [28] we show that the formulation of our offloading procedure is general enough to be independent from the specific RL algorithm applied. We also evaluate the performance of two learning approaches (Actor–Critic and Q-Learning) applied to the offloading problem and, as in our previous work, we compare their performance to Droid [29], a state of art solution in cellular offloading. Moreover, we analyse in more detail, with respect to [28], the behaviour of each considered approach in order to better understand the reasons behind their different behaviour.

The rest of this paper is organised as follows. In Section 2 we review other offloading approaches and we present to what extent they differ from our offloading solution. In Section 3 we formalise the offloading problem as a RL problem. In Section 4 we provide a brief background on the RL algorithm we use in this paper and in Section 5 we show how they are used in the offloading problem. Section 6 presents the performance of all the considered offloading approaches while Section 7 concludes the paper.

2. Related work

Over the last years, much effort has been devoted to find good solutions to cope with the problem of data offloading in cellular networks. In this section we briefly review the main literature on offloading, and present the main results regarding data offloading solutions based on opportunistic networks for delay tolerant content. A more complete and exhaustive survey on state of the art of data offloading techniques in cellular networks can be found in [33].

Offloading traffic from cellular networks can be done in several ways, depending on the type of constraints that should be met for the content delivery [33] and the specific technique applied. From a technological point of view, a first distinction is between *inbound* and *outboud* techniques [34]. In the former case, offloading is performed exploiting the same spectrum of the cellular network, while the latter exploits network technologies that use other parts of the spectrum, typically in unlicensed bands. Without loss of generality, in the rest of the paper we refer to outbound techniques that use WiFi networks to offload traffic, although the proposed offloading control mechanism can be used also with inbound offloading (most notably, with LTE-D2D, one of the reference offloading solutions defined by the LTE standard [35]).

An orthogonal classification is between AP-based and device-to-device (D2D) offloading [33]. In AP-based offloading (using WiFi technologies), end-users located inside a hot-spot coverage area might use WiFi connection as an alternative to the cellular network when they need to exchange data. In some schemes data transfers are even deferred when no WiFi connectivity is available, waiting for the next time the user will move close to a WiFi AP [36]. Due to the