Performance Evaluation of an Energy Efficient Traffic Offloading Protocol for Vehicular Networks

Loreto Pescosolido, Marco Conti, Andrea Passarella
Italian National Research Council - CNR
Institute for Informatics and Telematics - IIT, Pisa, Italy
Email: {loreto.pescosolido, marco.conti, andrea.passarella}@iit.cnr.it

***Abstract*—Offloading data traffic from infrastructure-to-vehicle (I2V) to vehicle-to-vehicle (V2V) communications is a powerful tool for increasing spectral efficiency, as well as energy efficiency, of a vehicular network. Typically, traffic offloading is achieved by exploiting caching capabilities of the mobile nodes (vehicles) and direct V2V communication technologies. In prior network-level studies on traffic offloading, the physical features of wireless channels are taken into account using relatively simplistic models. In this study, we propose an offloading protocol based on V2V communications and evaluate its performance taking into account realistic models of the wireless channel, focusing on energy efficiency as a performance metric. Specifically, we compare the results obtained by using a set of wireless channels and interference models, ranging from simplistic models based on deterministic path loss formulas to a very accurate channel model that has been developed inn the recent years for LTE-A networks. Our goal is twofold: on one hand, we aim to quantify the achieved benefits in terms of energy efficiency; on the other hand, we investigate the impact of using different wireless channel models in the evaluation of the energy efficiency of V2V offloading protocols.**

***Index Terms*—Traffic offloading, caching, VANETs**

I. Introduction

Offloading wireless traffic to device-to-device (D2D) or vehicle-to-vehicle (V2V) communications by caching contents in the nodes has sparkled interest, in the recent years, due to the benefits in terms of reduction of congestion on wireless and/or wired backhaul infrastructures, increased spectral efficiency and reduced energy consumption it may bring about [1]. When an offloading protocol, based on caching contents in the mobile nodes, is in use, a user can obtain a desired content either through an infrastructure-based wireless network, e.g., a cellular network or a WiFi network, or retrieve it from a neighbor that has cached it previously. Furthermore, thanks to the mobility of nodes, even though a certain content may not be available from the user's neighbors at the request instant, it may be still obtained through a D2D link, later on, by a *new* neighbor (which has previously cached the content), i.e., a node that comes close to the requesting user after the initial request instant. To exploit this possibility, in our previous work [2], inspired by [3], we used a *content timeout*, i.e., an interval starting at the content request instant, during which the requesting node tries to obtain the content from devices it encounters. If this is not the case, the content is delivered to the node by the Network Infrastructure (NI), i.e., a Base Station (BS), at the expiration of the timeout. Especially in highly dynamic scenarios, such as vehicular wireless networks, the use of a *content timeout* allows to increase the *offloading efficiency*, i.e., the percentage of contents delivered through D2D or V2V communications. In addition to reducing the traffic carried by the infrastructure, this approach also results in higher *energy efficiency*, defined as the amount of payload bits delivered per energy unit consumed by the overall system, and *spectral efficiency,* defined as the ratio between the amount of delivered payload bits and the amount of physical transmission resources[1,2]. Offloading efficiency, used in [2] and in other existing works (see [1]), is an important high level metric, which captures in a single parameter the effectiveness of the considered offloading protocol in achieving its objective. However, a more in-depth performance evaluation, which can be more easily mapped to the cost incurred by the network operator and the users, needs to target energy efficiency and spectral efficiency as additional performance metrics. In this work, we focus on energy efficiency. We show that, for a reliable evaluation of these performance metrics through simulations, which in turn allows a more effective design of offloading protocols, it is preferable to use channel models that depart from simplistic models based on deterministic

[1]E.g., the Physical Resource Blocks (PRBs) of a LTE-A time-frequency physical layer grid.

[2]Thanks to the lower transmission range, D2D communications require much less energy than conventional Infrastructure-to-Device (I2D) ones, while, at the same time, make it possible to reuse the same transmission resources (frequency, time slots, codes) in closer by regions, since interference ranges of D2D transmissions are smaller than those of I2D ones.

path loss. These models, used for instance to justify the circular coverage and interference regions assumed in the "protocol interference" model, fail to capture the effects of random components affecting the channel attenuation and the effect of frequency selectivity. Instead, it is necessary to use channel models that do account for all these effects, since their impact on system-wise performance is not negligible.

The contribution of this work is two-fold: firstly, we propose a distributed Content Dissemination Management System (CDMS) implementing a content caching and traffic offloading protocol for a set of wireless mobile devices in a given Region of Interest (ROI) with the assistance of a NI (i.e., a set of BS) covering the ROI; secondly, we evaluate the performance of the proposed system in terms of energy efficiency by using physical interference models fed with realistic channel models suited for a urban microcell scenario [4].

The proposed system yields a reduction in the system-wise power consumption which depends on the content popularity distribution. Assuming a standard Zipf distribution with parameter α, and with the conservative[3] choice of $\alpha = 1.1$, the achieved power consumption reduction can be as high as 35%.

Our results clearly show the need to use detailed physical channel models to accurately evaluate offloading performance. In addition, they also indicate that cross-layer modifications of the offloading algorithms, which exploit information about the status of the physical channel, could further improve energy efficiency.

The paper is organized as follows: in Section II we describe our system model; in Section III we describe the proposed protocol and the algorithms executed by the involved entities (CDMS and mobile devices); In Section IV we present our results obtained through system level simulations, and in Section V we conclude the paper summarizing our contribution and pointing at future research directions.

II. System model

A. System Geometry and Traffic Model

We consider a ROI populated with mobile users, and assume that the ROI is fully covered by a NI composed of possibly multiple BSs (e.g., a cellular infrastructure). Users enter, roam into, and exit the ROI. In this work we consider mobile devices to be on board of vehicles, either being human hand-held devices or part of the vehicle equipment, assuming simple vehicle mobility models inside the ROI. As users enter the area they become interested in contents according to a given content request process determined by the application layer. This process is characterized by a content-request *arrival* process and a content-interest distribution, which associates to each request a specific content. We assume the content-request arrival processes of different users to be statistically independent. Furthermore, the contents requested by different users or by the same user in different requests are also independent. The CDMS is implemented as software agent that can communicate with the users through the NI, as described in detail in Section III. Time is organized in epochs. In each epoch there are dedicated resources (e.g., a set of Physical Resource Blocks in a time-frequency grid as defined by the LTE-A standards) for allowing the NI to send messages for paging, broadcast, and instructions for individual nodes, the mobile nodes to send messages for implementing neighbor discovery and other control messages required by the offloading protocol (see Section III), and both BSs and mobile nodes to transmit the contents.

B. Channel modeling approach

When considering performance metrics such as power savings and spectral efficiency, the simplistic path loss models underlying the "protocol interference" model, where the signal attenuation is a deterministic function of the distance between transmitter and receiver, without considering shadowing and small scale fading effects, are not sufficiently accurate. In real-life deployments, for a given transmitter-receiver distance, variations of both useful signal and interfering signals strength due to shadowing, small scale fading, and Line Of Sight (LOS) or Non Line Of Sight (NLOS) conditions can be in the range of several tens of dBs. Furthermore, the Large Scale Parameters (LSPs) of the random components of the channel attenuation[4] exhibit spatial correlation[5] and correlation among different parameters of the same link. Additionally, they are also scenario-dependent[6] [4], [6]–[8], and in the same scenario depend on the communi-

[3]The selected value is close to the lower extreme of $\alpha = 1$. With $\alpha \leq 1$, it would be necessary to truncate the Zipf power law to a finite support, in order to obtain a well defined probability density function (PDF) [5]. Higher values of α would result in more skewed PDFs, which would further increase the offloading efficiency.

[4]LSPs include, but are not limited to, first and second order moments for the random generation of LogNormal shadowing, Rician K-Factor, and channel impulse response delay spread.

[5]Spatial correlation means that the LSPs of two different links with a common transmitter (or receiver) are correlated and the correlation coefficient depends on the distance between the two receivers (or transmitters) respectively.

[6]The term "scenario", here, refers to the characteristics of the real-life deployment environment, such as urban macro cell, urban micro cell, rural, etc..

cation being either D2D or I2D. In this work, the target scenario is the Urban Micro cell scenario (UMi) [4].

In this work, we adopt the geometry-based stochastic channel model (GSCM) put forward by the WINNER II European project in [6] and subsequently refined/extended by the ITU [4], the 3GPP [7], and by the METIS Project [8]. In [6], a detailed procedure is described for generating a set of frequency selective channels affected by shadowing and small scale fading. The procedure uses specific, scenario-dependent formulas for computing the path loss and generating the set of Large Scale Parameters (LSPs)[7].

We use a simplified version of the channel models in [4], [6]–[8], which we implemented in Matlab. Our implementation involves: (i) the computation of the path loss for all the links whose ends are located at any two points in a 2D rectangular grid with a spatial step of $5\,\mathrm{m}$ in both dimensions, and between each BS and each point in the grid, using the scenario-dependent formulas recommended in [8] for the UMi scenario, also taking into account the different heights of BSs and devices, (ii) the generation of random correlated LSPs for each link, using the procedure described in [6] and with the parameters selected in [8] for the UMi scenario for I2V and V2V communications as appropriate; (iii) the generation of the set of random tap delays and amplitude for each possible link (i.e., the generation of the channel impulse response), according to the procedure described in [6], using the parameters suggested in [8]. We assume omnidirectional antennas and account for different transmit antenna heights of I2V and V2V communications (which do have an effect on the path loss values), assuming an antenna height of $10\,\mathrm{m}$ for BSs and $1.5\,\mathrm{m}$ for mobile terminals. The models in [4], [6]–[8] provide a further level of detail, which we decided to skip for complexity reasons[8].

III. Content Dissemination Management

Each node k has an internal content cache $\mathcal{C}_k$ populated with previously downloaded contents, and a list of neighbors $\mathcal{N}_k$ composed of (j, r_j^k) pairs, where j is the id of any node which is a neighbor of node k and r_j^k is a ranking index of node j as seen by node k. $\mathcal{N}_k$ is updated periodically on the basis of Hello messages containing their id sent periodically by the nodes and received by each other neighbors. The ranking is determined on the basis of a given criterion. In this work, the ranking is based on the channel quality between node j and node k. The channel quality of the link (j, k) is obtained at node k by measuring the received power of the Hello messages sent by node j. Hello messages are transmitted with a predefined transmit power. At any time, the CDMS has a copy of the neighbor list $\mathcal{N}_k$ of each node, and has an index of the contents in each node's cache, although the CDMS does not necessarily hold a copy of the contents itself. The proposed protocol requires the mobile nodes and the CDMS to periodically execute routines implementing a basic neighbor discovery and ranking protocol, and to keep the index of each node's cache at the serving sBS, and its copy of each node's neighbors list, up to date. Algorithms 1-2 describe the actions taken on demand, i.e., as a consequence of content requests, by nodes and CDMS. We briefly introduce the notation required for a correct interpretation of the algorithms: the notation $\biguplus \{\mathcal{C}_j | \text{condition on } j\}$ is used to indicate the union of the caches of nodes satisfying a given condition; the notation $\hat{j}(k, z)$ is used to indicate the node j that has the best ranking r_j^k among the neighbors of node k which have content z in their caches; the notation $j \xrightarrow{z} k$ indicates the transmission of content z from node j to node k. These transmissions are triggered by the CDMS. The remaining notation used in Algorithms 1-2 is self-explanatory and we omit its description for space reasons.

Algorithm 1 describes the actions of a node as it becomes interested in a content. Essentially, it notifies the CDMS that it is interested in that content, and then waits for receiving it either from the NI or from a neighbor. The system guarantees that the content will be delivered within a given *content timeout*. After the reception of the content, the node makes it available for other nodes which will eventually request it, for a limited amount of time determined by a *sharing timeout*. The sharing timeout is introduced to keep the cache degree of occupation stable, i.e., avoid cache overflow, see [2].

Algorithm 2 describes the actions taken by the CDMS to handle a content request. Here, the key point which effectively allows to increase the system energy efficiency is that the CDMS selects the best node for delivering the content, on the basis of channel quality considerations, represented by the ranking of each node's neighbors (steps 5-13). If, however the content cannot be delivered through D2D within the content timeout, the CDMS uses the NI to deliver it (steps 15-20).

The proposed protocol can be implemented in dis-

[7]Namely, LSPs are the delay spread of the channel impulse responses, the Lognormal shadowing, and the Rician K factor (used to parameterize the small scale fading probability distribution.

[8]Particularly, most of the details we skipped become relevant when dealing with multi-element antennas and MIMO communications, which we do not consider in this work.

Algorithm 1 Actions taken by node k to request content z

1: **Upon** request for content z from the application layer
2: **Set** k_content_received = **false**
3: **Send** (k,z)_cont_req to CDMS
4: **while** k_content_received == **false do**
▷ Wait for receiving content z, from NI or from a neighbor
5: **if** content z is received **then**
6: **Set** k_content_received = **true**
7: **Send** (k,z)_ACK to CDMS and/or the sending node
8: **Add** z to $\mathcal{C}_k$
9: **Set** (k,z)_sharing_timeout
10: **break**
11: **end if**
12: **end while**
13: **while** (k,z)_sharing_timeout is not expired **do**
▷ Available for opportunistic sharing of content z
14: **Upon** request from CDMS (step 7 of Algorithm 2)
15: **Send** z to node requesting it
16: **end while**
17: **Remove** content z from $\mathcal{C}_k$
18: **Cancel** (k,z)_sharing_timeout

Algorithm 2 Actions taken by CDMS for handling content request (k,z)

1: **Upon** receiving (k,z)_cont_req
2: **Set** (k,z)_served = **false**
3: **Set** (k,z)_content_timeout
4: **while** (k,z)_content_timeout is not expired **do**
5: **if** $z \in \biguplus\{\mathcal{C}_j | j \in \mathcal{N}_k\}$ **then**
6: **Identify** $\hat{j}(k,z)$
7: **Trigger** transmission $\hat{j}(k,z) \xrightarrow{z} k$
8: **Wait** for (k,z)_ACK
9: **Upon** (k,z)_ACK reception
10: **Set** (k,z)_served = **true**
11: **Remove** (k,z) from $\mathcal{L}_{\text{req}}$
12: **break**
13: **end if**
14: **end while**
15: **if** (k,z)_served == **false**
16: **Send** z to k
17: **Wait** for ACK_(k,z)
18: **Upon** reception of ACK_(k,z)
19: **Set** (k,z)_served = **true**
20: **end if**
21: **Cancel** (k,z)_content_timeout

tributed way when multiple BSs are involved. In this case, suitable handover strategy is assumed to be in place (e.g., through a NFV component allocated appropriately in the infrastructure topology) to handle, e.g., the case of nodes waiting for a content moving outside of the region covered by the BS which has received the content request, and other situations that may occur due to user mobility.

IV. Performance evaluation

In this work, due to space reasons, we focus on energy efficiency as the target performance metric, leaving the analysis of spectral efficiency to a future work. We provide a comparison among the results obtained by using increasingly complex channel models for I2V and V2V communications. Particularly, we consider: (i) a simplistic channel model which accounts only for a deterministic path loss (PL) determined by the Friirs equation $PL_{\text{dB}} = 20\log_{10} d_{\text{m}} - 27,5582 + 20\log_{10} f_{\text{MHz}}$; (ii) a similar model[9], but with path loss exponent equal to 3; (iii) a deterministic path loss model which uses the path loss formulas 5-4, 5-5, and 5.6 in [8] (indicated with "METIS PL"); (iv) a model which includes path loss and correlated Lognormal shadowing, according to the parameters provided by [8] ("METIS PL+SH"); (v) a model which also includes small scale fading and frequency selectivity, with parameters selected from [8] as appropriate ("METIS PL+SH+FD"). In this work, we assume omnidirectional antennas and radiation patterns. We consider a street of length 1 km and width 20 m. Two BSs are deployed along the street at coordinates [300,10] and [700,10] (meters). Vehicles enter the street from both ends and travel through it at a constant speed. The speed of each vehicle is randomly selected from a uniform distribution in the range $[3,8]$ m/s. The content requests arrival process of each user is modeled as a Poisson arrival process with interarrival rate $\lambda = 1/120\,\text{s}^{-1}$ (one request every 2 minutes for each vehicle), and the content interest process is modeled by Zipf distribution with parameter $\alpha = 1.1$, with independent requests. The content size is 1 MByte. The *content timeout* is 20 s, and the *sharing timeout* is 10 minutes. The system bandwidth is 10 MHz. For simplicity, we assume that a Shannon capacity achieving physical layer is implemented and that an interference coordination mechanism is in place. In addition, we assume an ideal TDMA MAC in which users are allocated the entire system bandwidth during their transmission. The power used to successfully transmit a content (which is used as input to the computation of the energy efficiency of the system) equals the power necessary to achieve a capacity of 100 Mbps on the selected link exploiting the full system bandwidth. The required power is computed using the specific channel realization among the two points where transmitter and receiver are positioned at the time of transmission.

We performed a set of 10 independent simulations, each lasting 1 hour, using our custom Matlab simulator (see Section II). Each simulation corresponds to a different realization of the entire channel set, which

[9]Particularly, for the second model we used

$$PL_{\text{dB}} = \begin{cases} 20\log_{10} d_{\text{m}} - 27,5582 + 20\log_{10} f_{\text{MHz}} & d \le d_0 \\ 10\eta\log_{10} d_{\text{m}} + K_{\text{dB}} + 20\log_{10} f_{\text{MHz}} & d \ge d_0 \end{cases},$$

where: $K_{\text{dB}} = (20 - 10\eta)\log_{10} d_0 - 27,5582$, $d_0 = 10\,\text{m}$, $\eta = 3$.

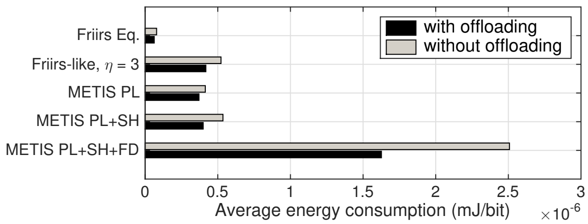

Figure 1. Energy consumption with and without offloading

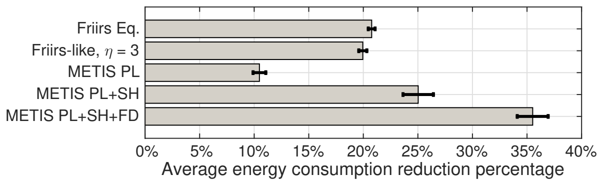

Figure 2. Reduction percentage of the average energy consumption.

includes the channels between any two grid points, and the channels between any grid point and each BS.

Fig. 1 displays the average transmit power, obtained when content requests are fulfilled only by the BSs (grey bars), and when the offloading mechanism is in operation, i.e., contents can be delivered *either* by neighbors *or* BSs (black bars). Fig. 2 represent the average and 95% confidence intervals (computed over the set of simulations) of the percentage by which power consumption is reduced when the offloading mechanism is implemented. In terms of absolute performance, it can be seen that there is a clear advantage in using the offloading mechanism, which allows for an up to 35% reduction of energy consumption. While this is an important result, it is worth mentioning that the exact percentage is specific of the system configuration considered in the preliminary result set presented in this work[10]. The result we aim to emphasize here, and that can be considered general (not related to the specific system configuration), is that using more realistic channel models shows higher energy savings. The main reason behind this behavior is that realistic channel models account for random shadowing and fading: consider a vehicle attached to a BS with an unlucky channel, e.g., a channel affected by strong shadowing, and having, say, K neighbors with a desired content, with channels affected by different shadowing/fading realizations. In this situation, to obtain the content, the system selects the neighbor with the best channel. Since the "best of K" channel among a set of channels is, in a statistical sense, better than a single channel (the I2V channel, in our example), the required power with offloading will be, on average, lower than with I2V communications. This type of diversity gain cannot be captured by channel models only including deterministic path loss, since they do not differentiate, statistically, I2V channels from "best of K" V2V ones. For this reason, a performance evaluation based on simplistic models tends to underestimate the advantage of using offloading.

[10]System parameters that affect the performance are related to the content request process, vehicle traffic, content and sharing timeout selection, and density of BSs. These parameters have an impact on the offloading efficiency [2] and hence on the energy efficiency. Here, suffice it to say that power savings of up to 80% may be achieved with higher offloading ratios.

V. Conclusion

In this work, we have proposed a traffic offloading protocol based on content caching at the mobile nodes and evaluated its performance in terms of energy efficiency. We have showed that the proposed protocol can achieve a considerable improvement in terms of power consumption compared to a scheme that only uses I2V communications. Furthermore, the system-level performance gain depends critically on characteristics of the radio channels that are not captured by simplistic channel models not accounting for random fluctuations around the mean path loss, namely shadowing and small scale fading. Particularly, the results obtained with the realistic models show greater benefits of offloading. This result suggests that physical channel parameters play a significant role in determining energy efficiency of offloading. This observation could be exploited to design further optimised offloading algorithms.

Acknowledgement

This work is partially funded by the EC under the H2020 REPLICATE (691735), SoBigData (654024) and AUTOWARE (723909) projects.

References

[1] F. Rebecchi et al., "Data Offloading Techniques in Cellular Networks: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 580–603, 2015.

[2] R. Bruno, A. Masaracchia, and A. Passarella, "Offloading through opportunistic networks with dynamic content requests," in *Proc. IEEE MASS '14*, Oct. 2014, pp. 586–593.

[3] J. Whitbeck et al., "Push-and-track: Saving infrastructure bandwidth through opportunistic forwarding," in *Pervasive and Mobile Computing*, vol. 8, no. 5, 2012, pp. 682–697.

[4] ITU-R M.2135-1, "Guidelines for evaluation of radio interface technologies for IMT-Advanced," International Telecommunication Union (ITU), Report M.2135-1, Dec. 2009.

[5] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary physics*, vol. 46, no. 5, pp. 323–351, 2005.

[6] IST WINNER II Project Deliverable 1.1.2, "WINNER II Channel Models, Part I," Tech. Rep., 2007.

[7] 3GPP, "Study on 3D channel model for LTE (Release 12)," 3rd Generation Partnership Project; Technical Specification Group Radio Access Network, T.R. 36.873, V12.4.0, 03 2017.

[8] ICT METIS Project Deliverable 1.4, "METIS Channel Models," Tech. Rep., 2015.