Experimental Characterization of I2V Wi-Fi Connections in an Urban Testbed

Pedro M. Santos*, Tânia Calçada*, Susana Sargento†, Ana Aguiar*, João Barros*
Instituto de Telecomunicações
*DEEC, Faculdade de Engenharia, Universidade do Porto
†Departamento de Electrónica, Telecomunicações e Informática, Universidade de Aveiro
{pedro.salgueiro,tcalcada}@fe.up.pt, susana@ua.pt, {anaa,jbarros}@fe.up.pt

ABSTRACT

A city-wide platform of sensing devices, or Data Collection Units (DCUs), can harness vehicular networks installed in public transports to gather sensor data at a backend server. Selection of DCU deployment sites must consider connectivity patterns and achievable transfer rates to the vehicles' On-Board Units (OBUs). In this work, we characterize experimentally the connectivity patterns and the achievable data rates and volumes using WiFi between a DCU and passing OBUs of a 400 node vehicular network. Measurements were taken over approximately one month and reveal a strong correlation between connection duration and transferrable data volume. We also evaluate the impact of speed and distance on throughput. Finally, multiple linear regression is used to model attainable data volumes and identify the features that best predict the potential data volume of a tentative deployment site.

Categories and Subject Descriptors: C.2.1 [Network Architecture and Designs]: Store and forward networks; Wireless communication

Keywords: infrastructure-to-vehicle; vehicular data collection; wireless link characterization

1. INTRODUCTION

The feasibility of vehicular data collection from road-side clients is dependent on the performance of the infrastructure-to-vehicle (I2V) opportunistic connections with the mobile nodes. We consider the application scenario of an urban environmental monitoring platform. A number of sensing units distributed accross the city, called Data Collection Units (DCU), produce sensor data that must be gathered at a backend server. Collecting data from a number of disparate locations may be impractical or costly if solutions such as cellular, M2M-dedicated systems or physical cabling to a communication backbone are used. An alternative solution is provided by existing vehicular hotspots with Internet access. Public vehicles, such as buses, are equipped with embedded On-Board Units (OBU) to provide free WiFi access points (APs) and DTN services to passengers and external clients. Although cost-free and providing city-wide coverage, this solution requires road-side clients to handle intermittent service availability. Selection of DCU deployment sites must evaluate the site's ability to support minimum data volume requirements. For this purpose, instantaneous and long-term performance of I2V WiFi links must be characterized and related with the sites' connection opportunities.

In this paper, we present an experimental characterization of the wireless connectivity between a road-side WiFi client and WiFi-enabled public transports. We take advantage of a network of 404 vehicles that operate WiFi APs, V2X communication via DSRC and DTN services. We deployed a WiFi-capable sensing unit, a DCU, atop a traffic light pole of a major street. Over the course of approximately one month, we characterized the duration and bandwidth of the wireless links between DCU and buses' OBUs using UDP streams. Daily connection time reached almost one hour on weekdays, resulting in a transferrable daily volume in the order of 4+ Gigabytes. Transferred data volume per connection was found to be strongly correlated with connection duration. We also measured the buses' speed and position, and found speed to have little impact in throughput. Finally, we performed linear regression over hour-long segments to identify the features that best describe the potential transferrable data volume of a site. The number of connections per hour was the best predictor.

The remainder of this paper is organized as follows. In Section 2, we describe our collection platform and the experiment methodology. Measured data is presented and analyzed in Section 3. In Section 4, we provide an overview on related work. We draw our conclusions in Section 5.

2. EXPERIMENTAL TESTBED

We now present the architecture of our vehicular collection platform. Next, we describe the setting and methodology of our wireless link characterization experiment.

2.1 System Architecture

Our sensing and collection platform encompasses multiple Data Collection Units (DCUs) scattered throughout the city. Their data is forwarded to a backend server by means of a 400-node vehicular network. The architecture of our collection platform, depicted in Figure 1, is composed of three main components: the DCUs, the vehicular network, and the backend server.

CHANTS'15, September 11, 2015, Paris, France.

DOI: http://dx.doi.org/10.1145/2799371.2799386.

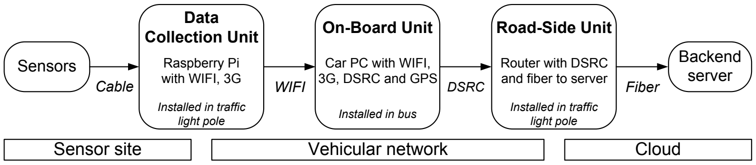


Figure 1: Architecture of vehicular collection platform.

(a) DCU. (b) East-bound. (c) West-bound.

Figure 2: Views of experiment setting.

Data Collection Units are deployed at selected city locations for environmental monitoring. DCUs are composed of an processing unit (Raspberry Pi Model B) equipped with WiFi (IEEE 802.11 b/g) interface and diverse sensor hardware, enclosed in a water-proof casing. A local database stores collected sensor data, which is deleted once acknowledged by the backend server. Figure 2a shows a deployed DCU. For this study, only one DCU is used.

The vehicular network is a large-scale deployment of embedded devices in the city's bus and municipality services fleet that provide WiFi access and vehicle-to-vehicle and vehicle-to-infrastructure communication. Nodes of the vehicular network are of two kinds: mobile nodes named On-Board Units (OBUs), and fixed nodes designated Road-Side Units (RSUs). OBUs are embedded computing devices equipped with GPS receiver, 3G, WiFi (IEEE 802.11 b/g) and DSRC (IEEE 802.11p) interfaces. They offer access point (AP) and DHCP server functionalities by means of Linux's *hostapd* and *udhcpd* services, respectively. A DTN service is offered for non-real-time applications (as it is the case of ours). RSUs are infrastructural routers deployed throughout the city, equipped with DSRC (IEEE 802.11p) antennas and a wired connection to the Internet via fiber.

The final end-point of our architecture, the backend server, collects and stores data from all DCUs deployed in the city.

2.2 Experiment Setting

Our experiments involved a DCU deployed at an street of our city where substantial bus traffic exists. The DCU was placed on a traffic light pole, at a height of approximately 4.5 meters, from where it draws electrical power. This pole is located roughly in the middle of a 600-meter long straight stretch of road, one-way and three lanes wide, with no curves or loss of line-of-sight for at least 300 meters in both directions. Variation in elevation amounts to 9 meters over that 600 meters stretch (less than 1° slope). Figures 2b and 2c shows the street from the DCU point of view.

There are three bus routes going through the street where the DCU is located. This site features multiple stopping opportunities for buses: the traffic light where the DCU is placed, a bus stop at approximately 100 meters towards West and a cross-roads with traffic lights, at 125 meters in the other direction. At the cross-roads, three distinct bus routes pass on the perpendicular street, on both ways. The only stopping opportunity with line-of-sight to the DCU for those buses is at the cross-roads' traffic lights. Night bus routes (0am-6am) only exist on the perpendicular street.

2.3 Experiment Methodology

We now describe our measurement setup and its operation during the experiment. The measurements are managed by an application-level script at the DCU. In our setup, the IP of the DCU is assigned dinamically by the OBUs' access points. The script monitors the DHCP client (*dhclient*) service as it continously waits for an IP assignment. The instant at which a new IP is detected is tagged as the beginning of the connection. To speed up the process of IP acquisition, the DHCP client parameters `initial-interval` and `backoff-cutoff` were set to to the minimum possible, 1 and 2 seconds respectively. The script then initiates a GPS query to the OBU and link quality measurements in a sequential fashion. The end of the connection is identified via timeout, when a query for new GPS data or a session of link quality measurements became unresponsive for 5 seconds. We do not have the information, in real or deferred time, to associate an OBU's MAC/IP address to specific routes.

We perform unidirectional (DCU to OBU) UDP link quality measurements. We use the tool *Iperf* [3] to generate load traffic (i.e., attempts to use the full bandwidth of the link) for a period of one second in each measurement. At the end of the Iperf measurement session, measured throughput, packet loss ratio and jitter for that second is reported. The Iperf client resides on the DCU and the Iperf server at the OBU. The GPS information of the bus is obtained via a query at application-level to the OBU, and contains time, longitude, latitude and speed of the bus. We pair a new set of GPS and Iperf measurements approximately every 2-3 seconds during the period the connection is alive. The MAC address of the associated AP is also stored.

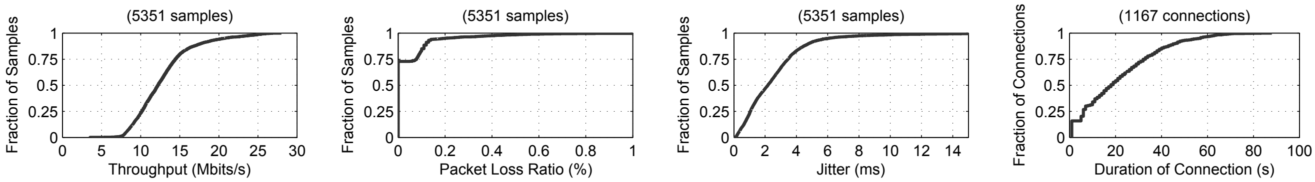
This experiment took place during 25 days, starting in August 19 and ending in September 12 of 2014. During this period, our equipment was the sole user of the OBUs' APs.

3. RESULTS AND DISCUSSION

We describe the pre-processing and analysis performed to our measurements, followed by a discussion on site selection.

3.1 Trace Pre-processing

The obtained raw samples were pre-processed to remove invalid or corrupt measurements. We observed that Iperf provides some anomalously large throughput values. Our measurement setup records throughput only at the application level and, even if the expectable drop in throughput due to IP/UDP overhead is disregarded, in no circumstance values higher than the physical layer's maximum nominal bit rate (55 Mbits/s for IEEE 802.11g) can be expected, which was the case. We filtered the measured samples depending on whether their value exceeded the physical layer's nominal bit rate. On the set of samples below 55 Mbits/s, we observed by histogram analysis that frequency of samples above 30 Mbits/s is null. This accounts for the expectable overhead and corroborates the validity of the selected samples. Regarding analysis of connections, we use the stored MAC addresses to identify the beginning and end instants, and quantify the connection duration. If two consecutive sets of samples associated to the same MAC address are apart by more than 60 seconds, we consider those to be two independent connections. Total transferred data volume for a connection is estimated by multiplying the total connection time and the average throughput for that connection.



(a) Throughput. **(b) Packet Loss Ratio.** **(c) Jitter.** **(d) Connection duration.**

Figure 3: Empirical CDFs over all measured samples and identified connections.

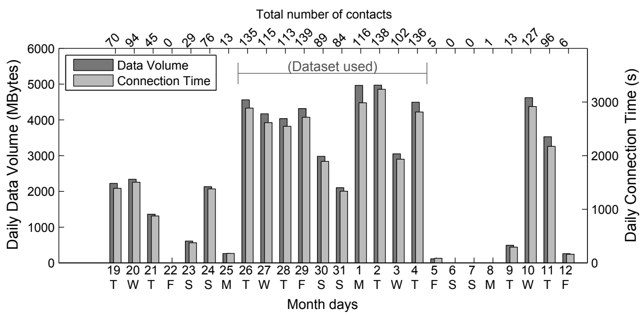


Figure 4: Average data volume transferred per day, over all days. Length bar indicates period of stable operation.

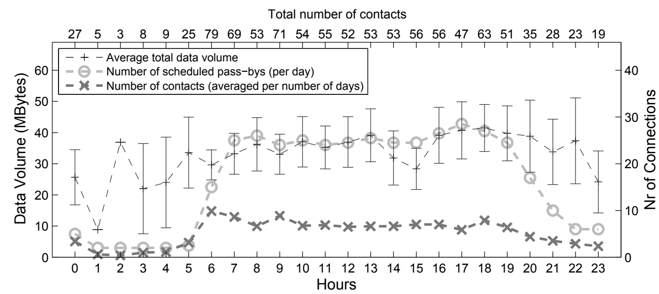


Figure 5: Average data volume, number of connections and number of scheduled bus pass-bys per hour for week-days of stable period. Confidence intervals of 95%[1].

Additionally, estimated transferred data volumes per day were not consistent throughout the experiment period. The operation of our setup was interrupted on day 21 for an experiment on delay-tolerant communication, and due to malfunction on 22 and 23 of August, and from 5 to 9 of September. Other interruption periods, namely day 25 of August and 3 of September, may be associated to external problems such as disruptions in the WiFi service operation, in the public authority vehicles operation and/or in road traffic. We restricted our dataset to a period of stable operation, between August 26 and September 4 (indicated in Figure 4 by the grey length bar). For this period we obtained 5351 samples and identified 1167 connections to buses.

3.2 Results and Analysis

An analysis of the measured data after pre-processing is now presented. The empirical cumulative distribution of throughput is shown in Figure 3a. The mean throughput considering all samples is 12.82 Mbits/s and the respective standard deviation is 3.78 Mbits/s. The cumulative distribution of measured jitter is presented in Figure 3c. The mean is 2.627 ms and the standard deviation is 2.628ms. Figure 3b depicts the ECDF of the packet loss rate. Packet loss was zero in 72.82% of the samples. The ECDF of connection duration is shown in Figure 3d. The mean connection time is 21.38 seconds and the standard deviation is 17.39 seconds.

The transferred data volume and total connection time per day is shown in Figure 4. We highlight the apparent correlation between the daily data volume and the daily total contact time visible in the figure. The Pearson's correlation test provided a correlation coefficient of 0.965 between contact time and transferred data volume for all individual connections, and of 0.99 for the daily totals. The total daily data volume can reach close to 5 Gigabytes (see days 1 and 2 of September). Daily total data volumes on the weekend of August 30-31 are smaller than in week days. The average measured number of connections per day was 86.5 on weekend and 124.3 on week days, and the average connection time was 18.7 and 21.9 seconds respectively. A smaller number of connections on weekends is consistent with less scheduled pass-bys, whereas faster connections may be caused by less traffic and/or less passengers.

The hourly average data volume per hour during weekdays of the stable period, the number of contacts averaged per number of days and the number of scheduled bus pass-bys, are shown in Figure 5. An apparent trend is that data volumes transferred during the night period (11pm-7am) are lower than those of peak hours (7am-10am and 4pm-8pm). We hypothesize that this is related to the corresponding number of measured connections/scheduled pass-bys. Despite correlation values with hourly data volume not being very strong (Pearson coefficient of 0.562 for the number of connections and 0.663 for the number of scheduled pass-bys), the curves exhibit similar shapes. Regarding the relation between the number of measured connections and of scheduled pass-bys, recall that, as described in Section 2.2, there are two sets of routes passing within range of the DCU with different stopping opportunities. The number of scheduled pass-bys shown in Figure 5 include both sets. Pearson coefficient between both metrics is 0.829.

Impact of distance and speed on throughput is presented in Figures 6 and 7. Measured throughput decays as the distance between terminals increases, which indicates lower signal-to-noise ratios caused by a decay in received power (as modeled by the line-of-sight path loss model). Samples at distances of 110 and 130 meters have usually velocity zero; these distances are coincidental with the nearby bus stop and cross-roads. Regarding speed, throughput showed little variation over the whole range of measured velocities, which is consistent with the existing results [5, 4, 2]. Note that higher speeds are typically recorded at closer distances.

3.3 Discussion on Site Selection

Our goal is to understand the characteristics of opportunistic connectivity that impact the most the total transferable data volume at a potential DCU deployment site. Coloquially, we ask if larger hourly data volume transfers are achieved at sites where few but long connections occur

[1] C.I.s for hours 1 and 2 omitted due to lack of statistical significance.

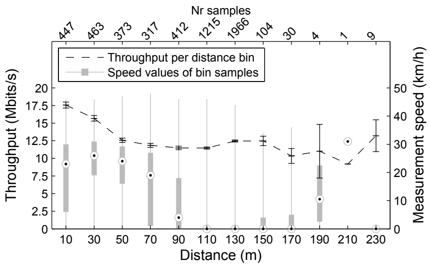


Figure 6: C.I.s (of 95%) of throughput binned by distance, and boxplots of each bin samples' respective speeds.

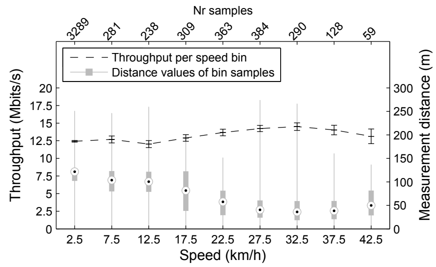


Figure 7: C.I.s (of 95%) of throughput binned by speed, and boxplots of each bin samples' respective distances.

Method	thr_avg	dur_avg	nrconn
Pearson	0.288	0.577	0.828
Kendall	0.205	0.463	0.683
Spearman	0.306	0.635	0.846

Figure 8: Correlation of `tdv` and selected features.

Crit.	thr_avg	dur_avg	nrconn
AIC	0.305	1.527e-49	1.437e-88

Figure 9: *p*-values for selected predictors. Inclusion/exclusion criteria was Akaike Information Criterion.

(near a bus stop), or in sites where connections are numerous but short (a major road without traffic lights or bus stops).

The presented results provided substantial information for this purpose. Transferred data volumes are highly correlated with connection duration, per instance and per day. This is supported by the large Pearson correlation coefficient presented. Therefore, we conclude that locations with the largest daily total connection times should be favoured. Also, throughput measurements present negligible influence of speed and exhibited a small variance over the whole range of velocities recorded. We conclude that connections with buses moving at urban speeds support similar throughput as with stopped buses, indicating that locations with few stopping opportunities may sustain high daily data volumes.

Additional insight was obtained from analysis over hour-long segments. We selected the hourly number of contacts (`nrconn`), the hourly average throughput (`thr_avg`) and the hourly average of connection duration (`dur_avg`) as potential features to provide indication about the hourly total transferred data volume (`tdv`). We computed `nrconn`, `thr_avg` and `dur_avg` from the data of the stable operation period, totalling 223 observations, and calculated their correlation with `tdv`. Results are shown in Table 8. The number of connections of a site, `nrconn`, exhibits the highest correlation to `tdv`. This indicates that locations with a large number of connections to buses are preferential.

Multiple regression analysis further supported this conclusion. We input `tdv` as the variable to be modelled and, as predictors, `nrconn`, `thr_avg` and `dur_avg`. We used stepwise regression and constrained it to finding a linear model without predictor interactions within 20 rounds. The p-values at the last round are presented in Table 9. The best predictor is the hourly number of connections, `nrconn`. This result strengthens the conclusion that a large number of connections is preferential over a large average connection duration.

4. RELATED WORK

We provide an overview of the literature on vehicle/infrastructure (V2I and I2V) connectivity. Message ferrying for data delivery through vehicular networks was first proposed in [6]. Regarding V2I-I2V communication based on the IEEE 802.11a/b/g standard, the authors of [5] report a UDP throughput of 35 Mbits/s for payload sizes of 1250 bytes at a velocity of 120km/h between a mobile client and an AP AP in the middle of a two-kilometer highway stretch. Results for other velocities are similar, showing that speed has little impact on throughput. The work described in [2] uses a fixed AP in a radiation-free zone and evaluate packet losses and throughput of UDP, TCP and web traffic connections, for a range of velocities. A data volume of 6.5 MBytes is reported at 75 miles per hour. It is concluded that the main factor that limits transferred data volume is the connection lifetime. The authors of [4] report a 200 meter window around a static node in which link throughput from a mobile node can reach nominal values and PLR can be almost to zero. Transmitted data volume was 9 MBytes. In [1], V2V and V2I communication is tested using IEEE 802.11b. Packet loss rates, jitter and retransmissions at the MAC layer remain close to zero, at speeds up to 130 km/h.

5. CONCLUSION AND FUTURE WORK

In this work, we evaluate the performance of wireless links between a fixed road-side WiFi client and a platform of WiFi APs deployed in a bus fleet. Measurements were taken at a single location over the course of almost one month. We report almost one hour of total daily contact time and 4+ Gigabytes of daily transferrable data volume in work days. Transferred data volume and total connection time are strongly correlated. Our goal is to identify relevant criteria to perform site selection. We found that the hourly number of connections is the best predictor of hourly total data volume. Future work will target city-scale site selection using these results. If bus frequency and mobility patterns for a given site are known, connection time can be predicted.

Acknowledgments

This work was supported by projects and grants SenseBusNet (PEst-OE/EEI/LA0008/2013), I-City for Future Mobility (NORTE-07-0124-FEDER-000064), Vital Responder 2 (PTDC/EEI-ELC/2760/2012), FP7-Future Cities (FP7-REGPOT-2012-2013-1, 316296), FCT Ph.D. grant (SFRH/BD/67178/2009), IT (UID/EEA/50008/2013) and by financial framework of R&D Unit 50008 (FCT/MEC through national funds and co-funded by FEDER-PT2020 partnership agreement).

6. REFERENCES

[1] P. Bergamo, M. Cesana, D. Maniezzo, G. Pau, K. Yao, D. Whiteman, and M. Gerla. Ieee 802.11 wireless network under aggressive mobility scenario. In *Proc. International Teletraffic Congress (ITC), Las Vegas, NV*, 2003.

[2] R. Gass, J. Scott, and C. Diot. Measurements of in-motion 802.11 networking. In *7th IEEE Workshop on Mobile Computing Systems and Applications*, pages 69–74, Aug 2006.

[3] NLANR/DAST. Iperf. https://iperf.fr/.

[4] J. Ott and D. Kutscher. Drive-thru internet: Ieee 802.11b for "automobile" users. In *23rd Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2004*, volume 1, pages –373, March 2004.

[5] M. Wellens, B. Westphal, and P. Mahonen. Performance evaluation of ieee 802.11-based wlans in vehicular scenarios. In *65th IEEE Vehicular Technology Conference*, pages 1167–71, 2007.

[6] W. Zhao, M. Ammar, and E. Zegura. A message ferrying approach for data delivery in sparse mobile ad hoc networks. In *Proceedings of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 187–198. ACM, 2004.