

Constant Rate Ultra Short Probing (CRUSP): Measurements in Live LTE Networks

Vaclav Raida
Technische Universität Wien
Vienna, Austria
vraida@nt.tuwien.ac.at

Philipp Svoboda
Technische Universität Wien
Vienna, Austria
psvoboda@nt.tuwien.ac.at

Martin Kruschke
A1 Telekom Austria
Vienna, Austria
martin.kruschke@A1.at

Markus Rupp
Technische Universität Wien
Vienna, Austria
mrupp@nt.tuwien.ac.at

Abstract—Throughput is one of the most important key performance indicators in cellular mobile networks. In wired networks, there are applications that estimate available throughput based on various low-volume traffic patterns. In wireless networks, however, these techniques have not proven itself to work.

Common throughput measurement tools estimate the available throughput by downloading as much data as possible for a predefined time. If the signal-to-noise ratio is high and the cell load is low, then such approach becomes expensive in terms of the consumed data volume, which users usually have to pay for.

In our previous work, we proposed constant rate ultra short probing (CRUSP); in the controlled LTE test cell, we confirmed that CRUSP can estimate available throughput while consuming less than 2MB. In this paper, we report on several drive tests (vehicular measurements) and analyze the performance of three Austrian operators in order to verify whether CRUSP can also be successfully applied to real live LTE networks.

Due to the highly dynamic nature of the mobile networks, we lack the ground truth when measuring throughput in live cells. To tackle this problem, we scheduled the same number of longer iPerf3 measurements, and then compared the throughput distributions of both tools using the Kolmogorov-Smirnov test.

We find that CRUSP yields the same distribution as iPerf3 if there is no UDP throttling in force. With UDP throttling, CRUSP does not match iPerf3 perfectly, but its distribution is still close. Hence, we can use CRUSP for operator benchmarking. In most cases, CRUSP saves more than 90% of measurement time and more than 90% of data volume.

Index Terms—cellular mobile networks, LTE, measurements, throughput, constant rate ultra short probing, mobile network operators, benchmarking, drive tests, iPerf

I. INTRODUCTION

In the previous paper [1], we proposed the constant rate ultra short probing (CRUSP) as a frugal alternative to traditional throughput measurement applications, which require several seconds and consume tens of megabytes of user data per test.

In this paper, we evaluate the drive-test measurements (nonstatic car measurements) that we conducted in Austria on five different days along five different routes in live LTE networks of three Austrian mobile network operators (MNOs).

We show that CRUSP yields the same throughput distributions as iPerf3 if there is no differentiation between TCP and UDP traffic. Although CRUSP transmits UDP segments, it performs—due to its short test duration—closer to TCP measurements if UDP throttling is in place.

A. Related Work

In [1], we proposed CRUSP as an alternative to conventional, heavy-load measurement applications. We performed a series of tests in our LTE reference cell, which is connected to a fully operational LTE core network. We concluded that CRUSP can measure throughput with the same accuracy as RTR-NetTest [2] while saving significant data volume and requiring only a fraction of measurement duration.

An extensive overview of throughput estimation tools and techniques—PathChirp [3], ASSOLO [4], WBest [5] to name few—is given in [6]. To the best of our knowledge, no one has been able to successfully apply them in cellular mobile networks. (See “Related Work” in [1] and [6] for more details.)

We implemented CRUSP based on FLARP (fast lightweight available rate probing) [6]. Instead of transmitting chirps, we send the data at a constant rate. CRUSP uses the simplest possible traffic pattern. Hence, it can serve as a benchmark to verify whether more advanced patterns bring any advantage.

MONROE platform [7] enables distributed measurements in cellular mobile networks in several European countries. In this paper, we utilize MONROE node to perform measurements on the networks of three different MNOs.

B. Paper Outline

Section II describes our hardware setup. We present settings of iPerf3 and briefly review the main idea behind CRUSP.

Section III introduces the Kolmogorov-Smirnov test, which we apply to compare the throughputs measured by different tools. We use the one-sided Kolmogorov-Smirnov test to benchmark different MNOs.

Section IV summarizes our measurement results. We find that we can use CRUSP instead of iPerf3 and still obtain the same throughput distributions, if no UDP throttling is present. We further discuss the impact and advantages of considering the individual bursts of received data.

II. MEASUREMENT SETUP

In [1], we compared CRUSP with RTR-NetTest (available open source as Open-RMBT [8]). Whereas CRUSP uses UDP protocol, RTR-NetTest generates encrypted TCP traffic. We know that in our reference cell, the TCP and UDP traffic is handled equally. However, for the measurements in live LTE cells of multiple MNOs, we had to check whether there is any

TABLE I
MEASUREMENTS CONDUCTED ON FIVE DIFFERENT DAYS ALONG DIFFERENT ROUTES [ILLUSTRATED IN FIG. 1 (A)]

No.	Time	MNO	Number of tests			CRUSP vs iPerf3 TCP			iPerf3 TCP vs iPerf3 UDP		
			CRUSP	TCP	UDP	H_0 rej.	p -value	$d \cdot 100$	H_0 rej.	p -value	$d \cdot 100$
1	21. 6. 2018 10:00-15:00	A	92	90	87	no	52.35%	11.81	no	66.83%	10.69
		B	90	90	90	no	28.23%	14.44	no	60.95%	11.11
		C	79	74	78	no	25.21%	16.10	yes	0.00%	43.62
2	22. 6. 2018 7:00-14:00	A	172	167	169	no	39.85%	9.59	no	26.53%	10.80
		B	166	167	167	no	25.46%	10.96	no	27.09%	10.78
		C	141	147	153	no	43.66%	10.09	yes	0.00%	27.52
3	25. 6. 2018 8:00-16:00	A	120	119	114	no	37.79%	11.58	no	7.17%	16.62
		B	94	94	94	no	30.71%	13.83	no	51.49%	11.70
		C	77	77	75	no	77.80%	10.39	yes	0.03%	33.11
4	26. 6. 2018 8:00-14:00	A	115	115	111	no	43.19%	11.30	no	84.82%	8.00
		B	132	133	133	no	49.42%	10.05	no	52.78%	9.77
		C	116	117	115	no	10.45%	15.64	yes	0.00%	39.58
5	27. 6. 2018 7:00-14:00	A	113	92	109	no	17.66%	15.18	no	8.01%	17.62
		B	96	81	93	no	62.44%	11.11	no	44.83%	12.82
		C	123	105	125	no	24.59%	13.36	yes	0.00%	31.28
All scenarios 1-5 together		A	612	583	590	no	41.36%	5.07	no	5.48%	7.76
		B	578	565	577	no	39.61%	5.26	no	49.49%	4.87
		C	536	520	546	yes	1.55%	9.52	yes	0.00%	31.28

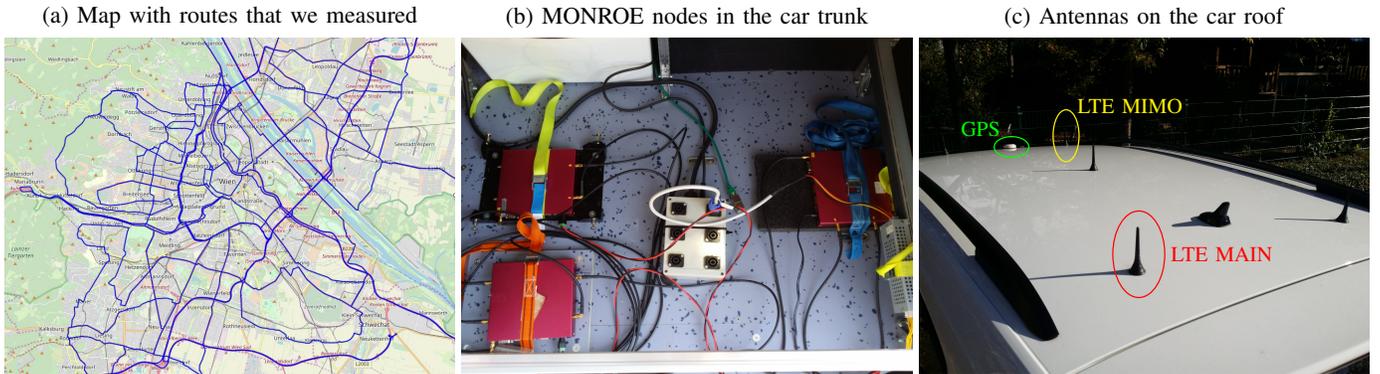


Fig. 1. Measurement setup. (a) During five days, we measured different routes in Vienna (over 500 km). (b) We measured with MONROE nodes, which have in total four LTE modems. Each modem has two antenna ports to support MIMO transmission. (c) Antennas on the car roof. We used four-way splitters so that all modems can share the same two antennas. (The nonmarked antennas were unused in this measurement campaign.)

difference between how TCP and UDP segments are treated. For this reason, we chose iPerf3, which can be executed both in TCP and in UDP modes, to be our reference tool.¹

We wanted to perform the measurements for all three major Austrian operators. Thus, we used a MONROE node (second design, [10]) that—in comparison with NEMO phones [11]—accommodates three LTE modems (Sierra Wireless, LTE Cat. 6). This private development node is not a part of the MONROE scheduling system. We, therefore, scheduled iPerf3 and CRUSP measurements via cron. We performed downlink (DL) measurements only.

In total, we measured for five days (see first two columns of Tab. I) along different routes in Vienna, Austria [see Fig. 1 (a)]. The MONROE node was placed in the trunk of a car [Fig. 1 (b)].² Each modem of the node had two antenna ports

to support MIMO³ transmission and reception. Each port was connected to an antenna mounted on the car roof [Fig. 1 (c)].

A. iPerf3

We set up our own iPerf3 server at the Institute of Telecommunications, TU Wien to make sure that there were no parallel tests running from other users. In the MONROE node, we executed iPerf3 client in TCP mode and in UDP mode. In both regimes, we run the tests with the default duration (10 s).

In the TCP mode, we applied the default TCP CUBIC congestion control algorithm, and set 10 parallel TCP connections to assure that the throughput ramp-up was sufficiently fast.

In the UDP mode, we used 10 parallel streams. In each, we set the target throughput⁴ to 20 Mbit/s. The resulting sending rate was thus 200 Mbit/s on the server side. Although the LTE Cat. 6 modems can receive up to 300 Mbit/s of DL throughput

¹IETF RFC 6349 “Framework for TCP Throughput Testing” [9] proposes iPerf3 for throughput measurements.

²The two red boxes in the left part of Fig. 1 (b) are the head and tail node, which together form the MONROE node of the second design. The red box on the right is a single node of the first design, which we used for a parallel measurement campaign, and it is not relevant in this paper.

³Multiple input, multiple output.

⁴Called “target bandwidth” in the iPerf3 documentation. We prefer the word “throughput” to avoid confusion with the frequency bandwidth.

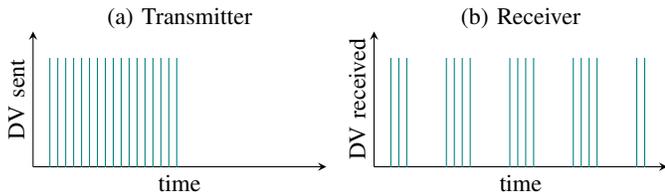


Fig. 2. An illustration of CRUSP framework. Transmitter (a) sends data at constant rate. Receiver (b), however, receives bursts of data. The CRUSP throughput estimation is based on post-processing the received bursts.

if the MNO supports carrier aggregation, we did not expect values higher than 200 Mbit/s in the live LTE networks due to cell load caused by other users.

B. CRUSP

We present in Tab. II the eight different CRUSP configurations we tested. For the rest of the paper, the most relevant configuration for us will be CRUSP1.

CRUSP server transmits a predefined data volume at predefined constant rate [Fig. 2 (a)]. CRUSP client receives bursts of data [Fig. 2 (b)] due to transmission time interval (TTI) or duty polling cycle. (We address this issue in more detail in Section IV-D.) As we discussed in [1], the first and last burst are dropped, and we divide the remaining n bursts by n inter-burst spaces to estimate the throughput.

III. BENCHMARKING METHODOLOGY

In the reference cell [1], we can set an attenuation level to achieve a desired constant signal strength. This leads to repeatable throughput measurements, yielding only small variations in the estimated throughput (ground truth).

In live cells, however, the ground truth is unknown. On the one hand, cell load changes rapidly. On the other hand, in the nonstatic scenario, the signal strength varies as the vehicle moves. Moreover, if we want to compare two different measurement tools, then we cannot measure using both of them at the same time—there is no guarantee that the resources would be split up equally. In a given time slot, only one app can be active to avoid cross-traffic.

A. Comparing Different Measurement Tools

To decide whether two different measurement tools deliver the same throughput estimates, we compare the distributions they generate. The empirical cumulative distribution function (CDF) of random variable (r.v.) Z , denoted as $F_Z(x)$, which is based on n samples z_1, \dots, z_n , is defined as

$$F_Z(x) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(z_i); \quad I_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases} \quad (1)$$

where $I_A(x)$ is an indicator function.

We apply the two-sided two-sample Kolmogorov–Smirnov test to assess the following hypothesis:

\mathcal{H}_0 : Throughput measurement samples collected by tool 1 (r.v. A) come from the same distribution as the samples collected by tool 2 (r.v. B).

\mathcal{H}_1 : Measurements conducted by tool 1 and tool 2 come from different distributions.

The Kolmogorov–Smirnov statistic d is the maximum absolute difference between two empirical CDFs $F_A(x)$, $F_B(x)$:

$$d = \max_x \{|F_A(x) - F_B(x)|\}. \quad (2)$$

The p -value is the conditional probability that the statistic D (r.v.) is larger than the observed value d , given that both r.v. A and B come from the same distribution (\mathcal{H}_0 is in force):

$$p = \text{Probability}(D > d \mid \mathcal{H}_0). \quad (3)$$

The asymptotic p -value can be approximated by [12, (14.3.18)]. In the rest of the paper, we set the significance level to $\alpha = 5\%$. The null hypothesis \mathcal{H}_0 is rejected if $p < \alpha$.

B. Benchmarking MNOs

We usually see from the CDFs, which MNO achieves higher throughput. However, to quantify the difference between two CDFs, we apply the one-sided two-sample Kolmogorov–Smirnov test. The null hypothesis \mathcal{H}_0 stays the same, but we modify the alternative hypothesis \mathcal{H}_1 :

\mathcal{H}_0 : Samples of MNO 1 (r.v. A) come from the same distribution as the samples of MNO 2 (r.v. B).

\mathcal{H}_1 : CDF of MNO 1 is larger than the CDF of MNO 2. (For the sake of brevity we write MNO 1 > MNO 2.)⁵

In the one-sided test, the statistic d is defined as

$$d = \max_x \{F_A(x) - F_B(x)\}, \quad (4)$$

and also the p -value is calculated slightly differently, see [13, Tab. 1, 2 (first row)] and [14].

IV. MEASUREMENT RESULTS AND DISCUSSION

We plot all CDFs with 95% lower and upper confidence bounds, which we obtain by Greenwood’s formula that approximates the variance of a Kaplan–Meier estimator [15].

A. CRUSP vs iPerf3 TCP vs iPerf3 UDP

Fig. 3 depicts the CDFs of the throughputs measured by CRUSP1, iPerf3 in TCP mode, and iPerf3 in UDP mode for all three MNOs during all five days. Tab. I compares the CDFs by the two-sided Kolmogorov–Smirnov test (Section III-A) for each day separately as well as for all five days together.

In the case of MNO A and B, we cannot reject the null hypothesis \mathcal{H}_0 that CRUSP and iPerf3 TCP measurements come from the same distribution. Similarly, we cannot disprove that iPerf3 TCP and UDP have the same distributions. We conclude that MNO A and B handle the TCP and UDP traffic equally. Furthermore, we can replace iPerf3 with CRUSP and still obtain comparable throughput characteristic.

We assume that MNO C throttles UDP traffic because the CDF of UDP is significantly higher than the CDF of TCP, with $d > 0.27$ and $p < 10^{-4}$ in each of the five

⁵ $F_A(x) > F_B(x)$ means that the realizations of r.v. A tend to be smaller than the realizations of r.v. B . Our notation MNO 1 > MNO 2, thus, means that MNO 2 performs better than MNO 1.

TABLE II
EIGHT DIFFERENT CRUSP CONFIGURATIONS

Configuration:	CRUSP1	CRUSP2	CRUSP3	CRUSP4	CRUSP5	CRUSP6	CRUSP7	CRUSP8
Volume:	1.86 MB	930 kB	310 kB	155 kB	74.4 kB	37.2 kB	10.0 kB	5.0 kB
Transmit rate:	300 Mbit/s	150 Mbit/s						

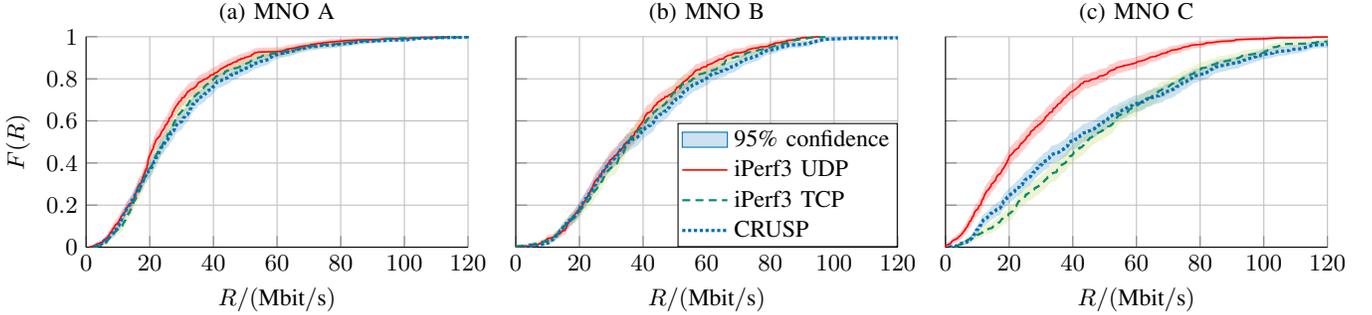


Fig. 3. Empirical CDFs of throughput measured by CRUSP1, iPerf3 in TCP mode, and iPerf3 in UDP mode for all three MNOs.

TABLE III
ONE-SIDED KOLMOGOROV-SMIRNOV TEST FOR MNO BENCHMARKING

Alternative hypothesis	CRUSP			iPerf3 TCP		
	H_0 rej.	p -value	$d \cdot 100$	H_0 rej.	p -value	$d \cdot 100$
A > B	yes	0.00%	22.12	yes	0.00%	27.97
B > C	yes	0.00%	13.97	yes	0.00%	19.18
A > C	yes	0.00%	27.52	yes	0.00%	37.31

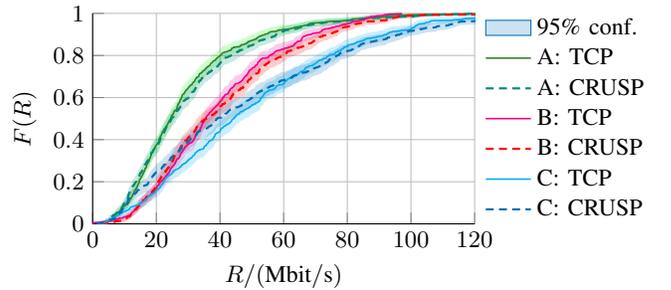


Fig. 4. CDFs of CRUSP1 and iPerf3 TCP show that we can effectively use CRUSP for MNO benchmarking and obtain results comparable to iPerf3.

days. Although CRUSP uses UDP traffic for throughput estimation, its estimates are closer to the results of the TCP tests. We suspect that the UDP throttling takes effect only after a certain time interval passes or after a certain data volume is transmitted (similar to traffic shaped by leaky bucket algorithm in [16]). CRUSP's estimates are thus closer to the unthrottled rate due to the short test duration and low data volume.⁶ Nevertheless, if we consider all measurements together, then the CRUSP1 and iPerf3 TCP are assessed as different distributions.

B. Benchmarking MNOs

Interestingly, although we measured along different routes on each day, we always obtained the same ranking: MNO A > MNO B > MNO C (i.e., C performs better than B, which performs better than A). We thus present only the total ranking based on measurements from all five days.

Fig. 4 shows the CDFs of CRUSP1 and iPerf3 TCP. We summarize the numerical values calculated by the one-sided Kolmogorov-Smirnov test (Section III-B) in Tab. III. Although we evaluated the CRUSP and iPerf3 TCP distributions as different in the case of MNO C, we can still effectively use CRUSP and achieve the same ranking as with iPerf3 TCP.

⁶An alternative explanation is that the UDP traffic congests the network because there is no congestion control algorithm. This would not happen with CRUSP due to its short duration and low data volume.

C. Data Consumption and Test Duration

We have seen that CRUSP can generate the same distribution as iPerf3 TCP. However, CRUSP is more efficient than iPerf3 and other similar tools in terms of measurement duration and consumed data volume. The duration of iPerf3 (default is 10 s) can be set to an arbitrary value, but other applications like RTR-NetTest [2], Speedtest [17], Netradar [18], OpenSignal [19], do not allow adjusting the test duration.

Fig. 5 (a) displays the CDF of data volume consumed per test. Whereas CRUSP1 is designed to transmit 1.86 MB per test, the consumption of iPerf3 TCP depends on the current channel capacity (the higher the available throughput, the higher the data volume consumption per fixed duration).

In Fig. 5 (b), we show the CDFs of the test duration that CRUSP1 needs to transmit the whole data volume. The higher the available throughput is, the shorter measurement time CRUSP1 requires. Both plots are based on the tests conducted during the five-day period.

D. Lower Volumes and Advantage of Considering Bursts

For a given maximum throughput, it is possible to test multiple configurations and find out which volume

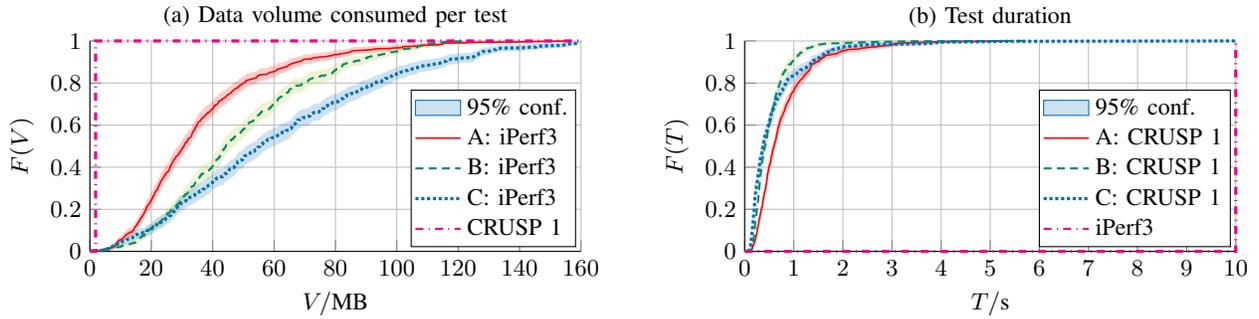


Fig. 5. (a) CRUSP transmits a predefined constant volume and is more efficient than iPerf3 TCP. (b) iPerf3 transmits at full rate for a predefined duration. Most of the CRUSP tests require less than 1 s.

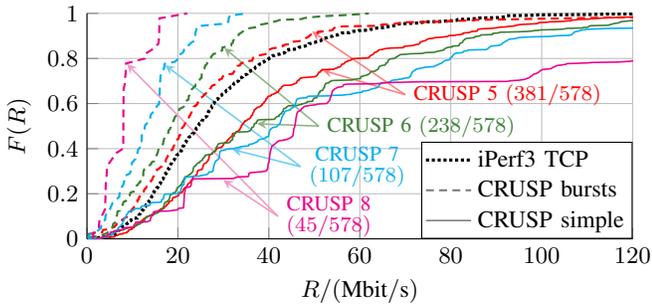


Fig. 6. Different CRUSP configurations, MNO B. The lower the transmitted volume is (higher configuration number), the more “CRUSP simple” overestimates and “CRUSP bursts” underestimates the actual throughput. The first number indicates how many tests were successful when considering bursts. For “CRUSP simple,” all tests (578) are always successful.

is sufficient to estimate that throughput (as we did in [1]). In live cells, however, we cannot easily evaluate the performance of CRUSP configurations with lower data volumes (unlike in our reference cell) because we do not know the ground truth (the current maximum throughput can be anything up to 300 Mbit/s). We cannot measure with two configurations simultaneously and if we conduct the measurements successively, then the available throughput may also change.

Therefore, the lower the CRUSP data volume is, the more cases CRUSP underestimates the throughput. The CDF is larger than the CDFs corresponding to higher volumes (see “CRUSP bursts” in Fig. 6).

If we do not consider bursts (“CRUSP simple”), then we face the opposite problem—overestimation. Correctly, we should divide the volume of n bursts by the duration of n inter-burst spaces, which correspond to end-to-end one-way delay. When ignoring bursts, we divide the total volume (n bursts) by the total duration ($n - 1$ inter-burst spaces). This effect becomes stronger for lower volumes (less bursts). In extreme cases, the overestimate exceeds the physical limit.

Considering bursts has the advantage such that if the data volume is too low, then we receive less than three bursts. The test thus fails (see [1, Section III-B]). Fig. 6 contains the number of successful tests for each “CRUSP bursts”

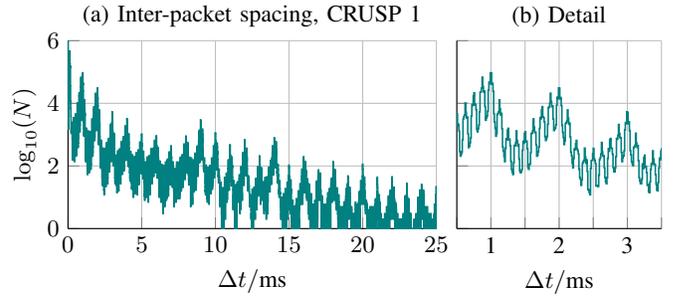


Fig. 7. Inter-packet spacing. In (a), we observe a strong peak every 1 ms (probably due to TTI). In (b), we recognize smaller peaks every 125 μ s (probably due to duty polling cycle of the receiver).

configuration (for “CRUSP simple,” no test can fail). This means that “CRUSP bursts” can be made adaptive—if we observe unsuccessful tests, then we can increase the volume. In the case of “CRUSP simple,” we discover the problem only after plotting the distribution of many tests.

CRUSP configurations 2–4 are not shown here. We tested them to see whether a lower volume would be sufficient. However, the Kolmogorov-Smirnov test did not confirm this.

E. Inter-Burst Spacing

The reason why the overestimated CDFs of “CRUSP simple” in Fig. 6 resemble stepwise constant functions is that the packet spacing is not uniformly distributed. Histogram in Fig. 7 shows the distribution of the packet spaces of all CRUSP1 tests (the spacing of other CRUSP configurations does not look differently).⁷

Almost 10^6 spaces are smaller than 1 ms, and these correspond to intra-burst spacing. In Fig. 7 (a) we see high peaks at integer multiples of 1 ms. This is probably caused by the TTI; in LTE, TTI = 1 ms. If we take a closer look [Fig. 7 (b)], we can also see smaller peaks. There are eight peaks per millisecond; most of the spaces are integer multiples of 125 μ s. This could correspond to duty polling cycle (proposed in [20]).

⁷In [1], most of the spaces cumulated near 6 ms. The difference may be due to the fact that we measured in live cells and used a different hardware.

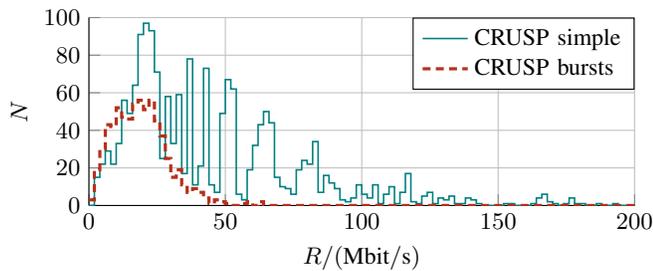


Fig. 8. Throughput measured with CRUSP6, MNO B. Throughput estimates of “CRUSP simple” exhibit quantization. With “CRUSP bursts,” we get rid of overestimates that are based on less than three bursts.

The quantization of inter-burst spacing leads to the quantization of throughput estimates. This becomes visible, especially if there are just few bursts per test (Fig. 8, “CRUSP simple”). If we consider bursts, then such tests are evaluated as invalid, and the throughput quantization is suppressed (“CRUSP bursts”).

V. CONCLUSION

To verify whether the constant rate ultra short probing (CRUSP) is usable not only in a well-controlled reference cell but also “in the wild” in live LTE cells, we performed a series of drive tests conducted for five days along different routes.

For MNOs, which treat the UDP and TCP traffic equally, the Kolmogorov-Smirnov test has revealed that CRUSP delivers results that are comparable to the iPerf3 in TCP mode. We cannot reject the hypothesis that the CRUSP and iPerf3 TCP throughput estimates come from the same distribution. In one MNO, we found evidence of UDP throttling. In that case, the CRUSP estimates tend to be closer to the TCP measurements.

We have demonstrated that CRUSP can be used for operator benchmarking. It generates a similar benchmark but consumes only a fraction of data volume (1.86 MB per test was enough) compared with traditional applications. CRUSP also requires shorter test duration—75% of all CRUSP tests took less than 1 s; 50% of all CRUSP tests took less than 0.6 s.

Furthermore, we investigated what happens if we decrease the volume that CRUSP uses and the impact of considering or not considering received bursts of data.

A. Outlook and Challenges

We have seen that by considering bursts, the throughput estimate may fail if we use too little data volume. This opens a possibility that we can make the CRUSP scheme adaptive—if we observe too many failed estimates, we can increase the data volume. Conversely, if no tests fail, then we can try to decrease the volume to achieve even shorter test duration.

So far, we have tested CRUSP only in LTE networks. However, the question that still remains is: How well would CRUSP perform in older UMTS networks? We assume that due to slower reactivity of UMTS, we may need to exclude more than just the first and last burst.

Currently, we are focusing on polishing the CRUSP source code. Our aim is to make CRUSP an open source so that other researchers can also evaluate its applicability.

ACKNOWLEDGMENT

This work has been funded by the ITC, TU Wien. The financial support of the Austrian BMFWF and the National Foundation for Research, Technology, and Development is gratefully acknowledged. The research has been cofinanced by the Czech GA CR (Project No. 17-18675S and No. 13-38735S) and by the Czech Ministry of Education in the frame of the National Sustainability Program under grant LO1401, and supported by the Austrian FFG, Bridge Project No. 871261. We thank A1 Telekom Austria AG for their support.

REFERENCES

- [1] V. Raida, P. Svoboda, and M. Rupp, “Constant rate ultra short probing (CRUSP) measurements in LTE networks,” in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Chicago, USA, Aug. 2018.
- [2] Rundfunk und Telekom Regulierungs-GmbH. (2018) RTR – NetTest. [Online]. Available: <https://www.netztest.at/en/>
- [3] V. J. Ribeiro, R. H. Riedi, R. G. Baraniuk, J. Navratil, and L. Cottrell, “pathChirp: Efficient available bandwidth estimation for network paths,” in *In Passive and Active Measurement Workshop*, 2003.
- [4] E. Goldoni, G. Rossi, and A. Torelli, “Assolo, a new method for available bandwidth estimation,” in *2009 Fourth International Conference on Internet Monitoring and Protection*, May 2009, pp. 130–136.
- [5] M. Li, M. Claypool, and R. Kinicki, “WBTest: A bandwidth estimation tool for IEEE 802.11 wireless networks,” in *2008 33rd IEEE Conference on Local Computer Networks (LCN)*, Oct. 2008, pp. 374–381.
- [6] M. Rindler, P. Svoboda, and M. Rupp, “FLARP, fast lightweight available rate probing: Benchmarking mobile broadband networks,” in *2017 IEEE International Conference on Communications*, May 2017.
- [7] Ö. Alay *et al.*, “Measuring and assessing mobile broadband networks with MONROE,” in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks*, Jun. 2016.
- [8] GitHub, Inc. (2018) RTR-NetTest open source repository. [Online]. Available: <https://github.com/rtr-nettest/open-rmbt>
- [9] B. Constantine, G. Forget, R. Geib, and R. Schrage, “Framework for TCP throughput testing,” Internet Requests for Comments, RFC 6349, Aug. 2011. [Online]. Available: <https://www.rfc-editor.org/rfc/rfc6349.txt>
- [10] M. Peón-Quirós *et al.*, “MONROE platform user manual,” Report, Rev. 1.0, Oct. 2018. [Online]. Available: https://github.com/MONROE-PROJECT/UserManual/user_manual.pdf
- [11] Keysight Technologies. (2018) Nemo handy handheld measurement solution. [Online]. Available: <https://www.keysight.com/en/pd-2767485-pn-NTH00000A/nemo-handy>
- [12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York, NY, USA: Cambridge University Press, 2007.
- [13] M. A. Stephens, “Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 32, no. 1, 1970.
- [14] The MathWorks, Inc. (2018) Two-sample Kolmogorov-Smirnov test – MATLAB kstest2. [Online]. Available: <https://www.mathworks.com/help/stats/kstest2.html>
- [15] —. (2018) Empirical cumulative distribution function – MATLAB ecdf. [Online]. Available: <https://www.mathworks.com/help/stats/ecdf.html>
- [16] V. Raida, P. Svoboda, and M. Rupp, “Lightweight detection of tariff limits in cellular mobile networks,” in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2018.
- [17] Ookla, LLC. (2018) Speedtest by Ookla. [Online]. Available: <http://www.speedtest.net>
- [18] Netradar Mobile Analytics. (2018) Consumers – Netradar. [Online]. Available: <http://www.netradar.com/consumers/>
- [19] OpenSignal. (2018) 3G and 4G LTE cell coverage map. [Online]. Available: <https://opensignal.com/>
- [20] Y. Xu *et al.*, “An end-to-end measurement study of modern cellular data networks,” in *Proceedings of the 15th International Conference on Passive and Active Measurement*, 2014, pp. 34–45.