








# TCP Performance over Current Cellular Access: A Comprehensive Analysis

Eneko Atxutegi<sup>1</sup><sup>(✉)</sup>, Åke Arvidsson<sup>2</sup>, Fidel Liberal<sup>1</sup>,  
Karl-Johan Grinnemo<sup>3</sup>, and Anna Brunstrom<sup>3</sup>

<sup>1</sup> University of the Basque Country (UPV/EHU), Bilbao, Spain  
{[eneko.atxutegi](mailto:eneko.atxutegi@ehu.eus),[fidel.liberal](mailto:fidel.liberal@ehu.eus)}@ehu.eus

<sup>2</sup> Kristianstad University, Kristianstad, Sweden  
[ake.arvidsson@hkr.se](mailto:ake.arvidsson@hkr.se)

<sup>3</sup> Karlstad University, Karlstad, Sweden  
{[karl-johan.grinnemo](mailto:karl-johan.grinnemo@kau.se),[anna.brunstrom](mailto:anna.brunstrom@kau.se)}@kau.se

**Abstract.** Mobile Internet usage has increased significantly over the last decade and it is expected to grow to almost 4 billion users by 2020. Even after the great effort dedicated to improving the performance, there still exist unresolved questions and problems regarding the interaction between TCP and mobile broadband technologies such as LTE. This chapter presents a thorough investigation of the behavior of distinct TCP implementation under various network conditions in different LTE deployments including to which extent TCP is capable of adapting to the rapid variability of mobile networks under different network loads, with distinct flow types, during start-up phase and in mobile scenarios at different speeds. Loss-based algorithms tend to completely fill the queue, creating huge standing queues and inducing packet losses both under stillness and mobility circumstances. On the other side delay-based variants are capable of limiting the standing queue size and decreasing the amount of packets that are dropped in the eNodeB, but under some circumstances they are not able to reach the maximum capacity. Similarly, under mobility in which the radio conditions are more challenging for TCP, the loss-based TCP implementations offer better throughput and are able to better utilize available resources than the delay-based variants do. Finally, CUBIC under highly variable circumstances usually enters congestion avoidance phase prematurely, provoking a slower and longer start-up phase due to the use of Hybrid Slow-Start mechanism. Therefore, CUBIC is unable to efficiently utilize radio resources during shorter transmission sessions.

**Keywords:** TCP adaptability · LTE · Flow size · Slow-Start  
Mobility

## 1 Introduction

Mobile Internet usage has increased significantly over the last decade, growing almost 18-fold over the past 5 years and more than half a million new mobile devices and connections in 2016 [1]. The following years are expected to be

equally promising with 4G traffic reaching quotas of more than three-quarters of the total mobile traffic by 2021. The growth expectation is not only related to the traffic volume itself but also to the average speed. To continue this growth and to meet user expectations, all the involved stakeholders have a common interest in fast downloads, quick responses, high utilization and few packet losses.

Since a large part of mobile Internet comprises TCP flows, the performance of TCP over cellular networks has become an important research topic. Even though in the last three decades many different TCP implementations have been developed [2] each of them targeting a different Congestion Control Algorithm (CCA), there still exists room for improvement in terms of achieved throughput and resulting delay over highly variable mobile networks.

Previous studies and proposals have reported their results regarding the interaction effects between mobile networks and TCP [3–5] and tried to define suitable CCAs for mobile networks [6]. However, none of them have extensively study the implication of a wide range of TCP implementations in a variety of static and moving scenarios. This chapter complements and extends previous works on mobile networks by studying and evaluating the behavior of a selection of TCP variants with different packet sizes, network loads, during start-up and mobility with different speeds, i.e. scenarios that are considered challenging for TCP. In order to appropriately study the different sources capable of impacting the final performance, the chapter suggests a bottom-up scenario with respect to complexity starting with static conditions so as to understand the responsiveness of TCP under distinct network status and load combinations and finishing with a variety of mobility scenarios.

The chapter is organized as follows. Section 2 covers related work. In Sect. 3, a brief overview of the studied TCP variants is provided and the LTE testbeds are described. Next, in Sect. 4, we explain the methodology regarding the performed measurements and the studied scenarios. The findings and results from our work are presented in Sect. 5. Finally, Sect. 6 concludes the chapter with a summary and a discussion of future work.

## 2 Related Work

TCP and LTE cellular access have been deeply studied throughout the last years. Most of the studies have either research the TCP side or mobile network side. However, a significant amount of researchers have been attracted by the interaction between TCP and LTE.

One of the first basis of such interaction is the impact that radio retransmissions have into the delay increment and how they therefore degrade the achieved goodput [7, 8]. It has been proven that the number of simultaneously active User Equipments (UEs) towards a common eNodeB has a huge impact on the effective available bandwidth due to radio resources being shared. Thus, the work [9] found that sudden increases in background traffic load have an important effect in the Round-Trip Time (RTT) increment. This cross-traffic effect severely influences the network playground for TCP, provoking sudden changes in the network

conditions and making TCP struggle while following the fluctuations in the available capacity. This chapter compiles a more detailed treatment of the effects of buffering in the radio access part of LTE by also considering the performance of high-speed/long-delay variants of TCP in these kinds of networks.

The so-called bufferbloat effect has also a huge impact into the performance of TCP over LTE [10]. The bufferbloat effect is possible due to the configuration of long queues both in the end-nodes and intermediate nodes, which can accumulate a great number of packets without any drop. However, that excessive packet buffering in a single queue in the end-to-end network path, causes a great latency increase and therefore, throughput degradation. Our work does not merely focus on bufferbloat, but considers the implications of different TCP variants in queue build-up under certain network conditions.

Considering that many flows in Internet are short, it is important to verify the efficiency of TCP to carry out such transmissions over cellular networks, it has been demonstrated [11] that under some network conditions TCP fails to correctly utilize the available capacity and therefore, the flows last longer than necessary. The current work complements such works and analyzes the impact that different flow sizes have in the performance outcome of different TCP flavors. To this end, our work not only focuses on the stationary phases of TCP but also on its behavior during start-up due to its significant impact in short flows performance. In particular, we study the Hybrid Slow-Start scheme [12], and evaluate how it operates in LTE networks in comparison with the Standard Slow-Start scheme.

Other studies have measured TCP over live LTE networks. Apart from the classic metrics of TCP throughput and RTT in [4] they also measured the delay caused by mobile devices going from idle to connected state. In [13], measurement trials were carried out over the cellular access of four Swedish operators and the diurnal variation of TCP throughput and delay were analyzed. [14,15] studies did similar TCP measurements, however, they did not consider daily variations. None of these live measurements took into account the impact of speed in the performance of TCP, or the behavior on different types of CCAs. So, to the best of our knowledge, our work both complements and extends these works through the study and evaluation of the behavior of common TCP variants in LTE networks under mobility with different speeds.

There are only a few works that have considered the impact of different speeds on the performance of TCP over LTE networks. Even though some works [5] have studied different speeds, the primary metrics were more related to the radio part with spectral efficiency and share of resource blocks among the UEs. Even though the utilization of such radio resources was studied, one or two simple variants of TCP were utilized in a multi-user resource share, leading to TCP micro-effects masking. Also, in [3] the impact of speed on TCP in LTE was studied. The work focused on uplink and downlink throughput, RTTs and also considered time-of-day variations. Still, they did not consider how the CCA factor into the TCP performance at different velocities. Our work serves to cover all the options and extends the previous studies with multiple mobility patterns, different speeds and a wide range of CCAs.

### 3 Research Environment

In order to compare the behavior of TCP in LTE networks, we first choose the TCP variants AND identify the LTE working parameters. This section first describes the most important features of the selected CCAs and later presents the LTE setup.

#### 3.1 TCP Variants

TCP variants fall into three categories according to the CCA mechanism used: loss-based, delay-based and combined loss- and delay-based. Along this chapter, the analysis starts with five CCAs and, with every measurement phase, we will reduce the group, avoiding the repetitive usage of TCP solutions that do not work well in mobile networks. A brief overview of the TCP variants is given below together with the classification of CCAs in Table 1.

**Table 1.** Selected TCP CCAs and their category

CCA category	Selected TCP CCA
Loss-based	TCP NewReno
	TCP CUBIC
Delay-based	TCP CDG
Hybrid with bandwidth estimation	Westwood+
Hybrid without bandwidth estimation	Illinois

- (i) TCP NewReno [16] employs the well-known additive increase multiplicative decrease (AIMD) mechanism that is common to most CCAs. During the Slow-Start period the *cwnd* increases by one packet per acknowledgment (ACK) reception until it reaches the value of *ssthresh*. Afterwards, the *cwnd* enters the congestion avoidance phase, with an increment of one packet per RTT period (standard synchronization with RTT or RTT-synchronized). If a 3-duplicate ACKs (3DUPACK) are received or a time-out occurs, the CCA deduces that some link is congested. After 3DUPACK, NewReno establishes the *cwnd* to the half (basic back-off) and the new *ssthresh* to previous *cwnd*. However, if a time-out occurs the *cwnd* will be decreased to one packet. NewReno is essential in the measurements since it represents the base TCP behavior.
- (ii) TCP CUBIC [17] employs a different mechanism compared with AIMD based on a cubical function. After a decrease of the *cwnd*, the *cwnd* ramps up in a concave shape, until it achieves the value that the *cwnd* had before the reduction. Afterwards, CUBIC increases its growth rate and ramps-up in a convex shape. CUBIC uses Hybrid Slow-Start [12] mechanism in the

sender instead of the Standard Slow-Start phase. Hybrid Slow-Start aims at finding the proper exit point for standard Slow-Start in order to avoid massive packet losses. The detection of such an exit point is based on the measurements of ACK trains and RTT delay samples. The TCP CUBIC implementation has been selected for the analysis due to its widespread use due to the fact that it currently is the default CCA in Linux servers, whose market share comprises the 67% of world-wide servers (as stated by W3Techs [18]).

- (iii) TCP CAIA delay gradient (CDG) [19] modifies the TCP sender to use RTT gradients as a congestion indicator. CDG also calculates the state of the bottleneck queue so that packet losses are treated as congestion signals only when the queue is full. Finally, CDG also uses Hybrid Slow-Start but with a more strict configuration than CUBIC. The selection of TCP CDG has been based on its novel use of delay gradients in the AIMD mechanism and to evaluate the actual usefulness of such a different feature in mobile networks.
- (iv) TCP Westwood+ [20] is capable of estimating the available bandwidth and minimum RTT (RTT<sub>min</sub>) by measuring ACK inter-arrival times. The estimations are used to decide the new *cwnd* after a congestion episode of 3DUPACK. With timeouts the *ssthresh* is calculated in accordance to the estimations and the *cwnd* is set to 1 segment. TCP Westwood+ has been selected in this study for its hybrid behavior using loss-based mechanisms together with delay-awareness.
- (v) TCP Illinois [21] controls the AIMD mechanism by the estimated queuing delay and buffer size. In a normal situation when no queuing delay is detected, the *cwnd* is increased by 10 packets per RTT. If estimated delay starts increasing, the increment of *cwnd* will be gradually lowering until the minimum value of 0.3 packets per RTT is reached. When the RTT is considered as high as compared to the baseline RTT, the loss is considered as buffer overflow, whereas in low RTT the loss counts as packet corruption. Developed to perform efficiently within high speed networks, its loss-based and delay-awareness make a perfect candidate for our study.

### 3.2 LTE Setup

In order to evaluate the performance of LTE three different environments have been used: simulation, emulation and controlled deployment. Most of the work described in this chapter has been carried out over the simulated environment and for comparison purposes the findings and results have been correlated with the behavior in the other two deployments. Since the configuration and explanation of the simulated environment is comprised of many parameters and in order to help the reader understand the setup, Table 2 gathers the most important information about the simulation environment regarding the configuration parameters and experiment-related conditions.

As the simulated environment, ns-3 simulator with the LTE capabilities of LENA module is used. This module also allows to create standard-based fading

traces that can be applied to the channel between the UE and the eNodeB. Since ns-3 does not exactly use the available TCP implementations in the Linux kernel, we used Direct Code Execution (DCE) Cradle [22] to be able to run real TCP implementations in ns-3. In order to simulate the distance to the server, the propagation delay between the fixed remote host and the Packet Data Network Gateway (PGW) was set to 40 ms. In the Radio Link Control (RLC) layer we selected the Acknowledged Mode (AM) in order to resemble the most commonly deployed configuration in real-world. We modified the mechanism to be able to support a limitation in terms of packets, establishing in our setup a common packet buffer size in the eNodeB of 750 packets. Regarding the radio resources, the eNodeB was configured to have a standard value of 100 available physical resource blocks (PRB). We simulated the frequency band 7 (2600 MHz), one of the most commonly used commercial LTE frequency bands (in Europe).

Background flows are used to load the network with multiple short TCP connections, similar to the behavior of real networks. The same TCP variant is used for both background and foreground traffic in order not to be affected by issues of TCP friendliness. The amount of data transferred in a background connection as well as the inter-arrival time between two connections were drawn from uniform random distributions.

The controlled testbed aims at providing a measurement platform with the ability to measure TCP in more realistic radio conditions in order to confirm or reject the findings and assumptions made in simulated environment in relation to the behavior of TCP over LTE. We have used the iMinds'/iMEC's LTE facility (LTE w-iLab.t [23]) in Zwijnaarde, Ghent. Apart from the provisioning of all the agents involved in LTE, the deployment allows ad-hoc mobility patterns while

**Table 2.** Simulation parameters

Simulation environment	
Simulator	ns-3 LENA LTE model
Linux Kernel	4.3 (DCE)
CCA	NewReno/CUBIC/Illinois/CDG/Westwood+
Parameter	Value
One-way delay PGW-Server	40 ms
MAC scheduler	Proportional fair
AMC model	MiError
Number of PRBs	100
LTE band	7 (2600 MHz)
RLC mode	AM
RLC transmission queue	750 PDUs
Pathloss model	FriisPropagationLossModel
Fading models	EVA60/EVA200

experimenting. It is important to underline that in this environment, the LTE transmissions are done over the air, thus allowing a proper study of TCP and LTE events. Even though the movement is real, the space limitation could limit the employed speed.

The emulated testbed targets the validation of simulated results of TCP under mobility circumstances. To this end, a LTE emulator or LTE-in-a-box (Aeroflex 7100) has been used. This emulator is capable of creating the LTE radio signal and all the necessary LTE protocol events to support the attachment and registration of any LTE device through a radiofrequency cable or over the air. The tests have been completed with an smartphone, a couple of servers and a controller to synchronize the experiments and all the equipment involved during the assessments. Since the UE in the emulated testbed is not able to physically move, the controller would continuously manipulate the baseline Signal-to-interference-plus-noise ratio (SINR) levels and Aeroflex would apply the corresponding fading pattern so as to model actual movement.

## 4 Methodology Description

The intrinsic operation mode of LTE (i.e. resource sharing, scheduling, HARQ mechanisms) results in a constant change in the available capacity. Even considering single-UE scenarios, different positions and fadings would lead to have a different SINR and it would therefore report a distinct Channel Quality Indicator (CQI) to the eNodeB. Thus, the eNodeB would assign a different available capacity for the channel of the UE through the Modulation and Coding Scheme (MCS) and transport block size (tbSize). Due such fluctuations in the radio side, the *cwnd* will be continuously evolving in order to obtain a resulting goodput as close as possible to the available capacity. The relative progressions of both parameters (available capacity and achieved capacity) play a fundamental role in the final performance.

In this section, the applied methodology will be presented. Figure 1 shows the different scenarios that have been used in the analysis of the effects between TCP's different CCAs and LTE. The methodology and reasoning of each scenario is explained below.

(I) **Implication of cross-traffic and responsiveness of TCP:** The static scenario aims at providing insights of the evolution and responsiveness of TCP under different background traffics (I point in Fig. 1). There are three main goals with this scenario: the comparative study of TCP behavior with and without a loaded cell, the analysis of TCP focusing on short flows and the responsiveness comparison of TCP variants with a sudden capacity increase and decrease. Several metrics are gathered at different nodes along the path. At the source, TCP state information such as *cwnd* and *ssthresh* is saved. At the eNodeB, the transmission buffer length, the drop count and the Packet Data Convergence Protocol (PDCP) delay (i.e., the time it takes for a PDCP Protocol Data Unit -PDU- to go from the eNodeB to the UE), are logged. Finally, in the UE, the goodput is recorded. Since it is measured at the application level, packet losses and/or reordering may result in goodput spikes.

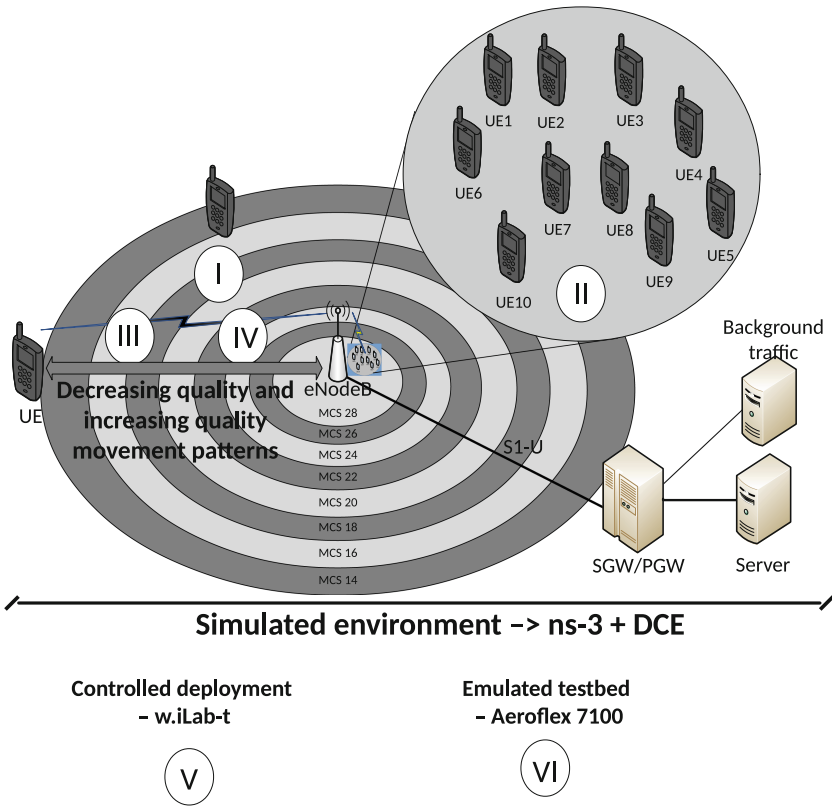


Fig. 1. Scenarios in use.

(II) Start-up performance: The aim of this scenario is to analyze the impact of different CCAs’ Slow-Start phases, such as the above mentioned standard Slow-Start and Hybrid Slow-Start, and determine their adequacy or inadequacy in broadband mobile networks. To that purpose, we deployed 10 static and scattered UEs (II point in Fig. 1) in good radio conditions (CQI 15) so as to study the start-up performance in a simplified multi-user scenario and set some basis for the understanding of the following scenarios. In the server, the *cwnd*, RTT, outstanding data and goodput has been collected.

(III & IV) Cell outwards/inwards movement resulting on decreasing/increasing available capacity: The decreasing quality movement scenario evaluates the behavior of TCP with a constantly worsening channel quality on average (III point in Fig. 1). The idea is to assess the CCA’s adaptability in a continuous capacity reduction (on average) environment and the impact of UE’s speed on the final performance. To help simulate different speeds, two Extended Vehicular A Model (EVA) fading patterns are applies: one for the velocity of 60 km/h (common limitation in rural roads) and one for 200 km/h (common maximum speed



in high-speed trains). Apart from the usual metrics in the evaluation of CCAs, the main metric for simulated mobility-based scenarios is the relation between the available capacity (extracted from the `tbSize`) and the achieved goodput. On the other hand, the increasing quality movement represents the behavior of TCP on a constantly improving channel quality (IV point in Fig. 1). Therefore, these simulations aim at evaluating the CCA's adaptability under different UE's speeds in a continuous increasing capacity (on average) environment.

(V) Correlation of TCP behavior in deployments as similar as possible to live commercial LTE networks: The scenario (V point in Fig. 1) aims at providing a measurement platform with the ability to measure TCP in more realistic radio conditions in order to confirm or reject the findings and assumptions made in simulated environment in relation to the behavior of TCP over LTE. Since the equipment in the scenario is fully real (see description in [23]), the scheduling, queuing and the rest of the features that could have an impact on delay are realistic and represent more clearly what would happen in live scenarios, helping in the verification of findings.

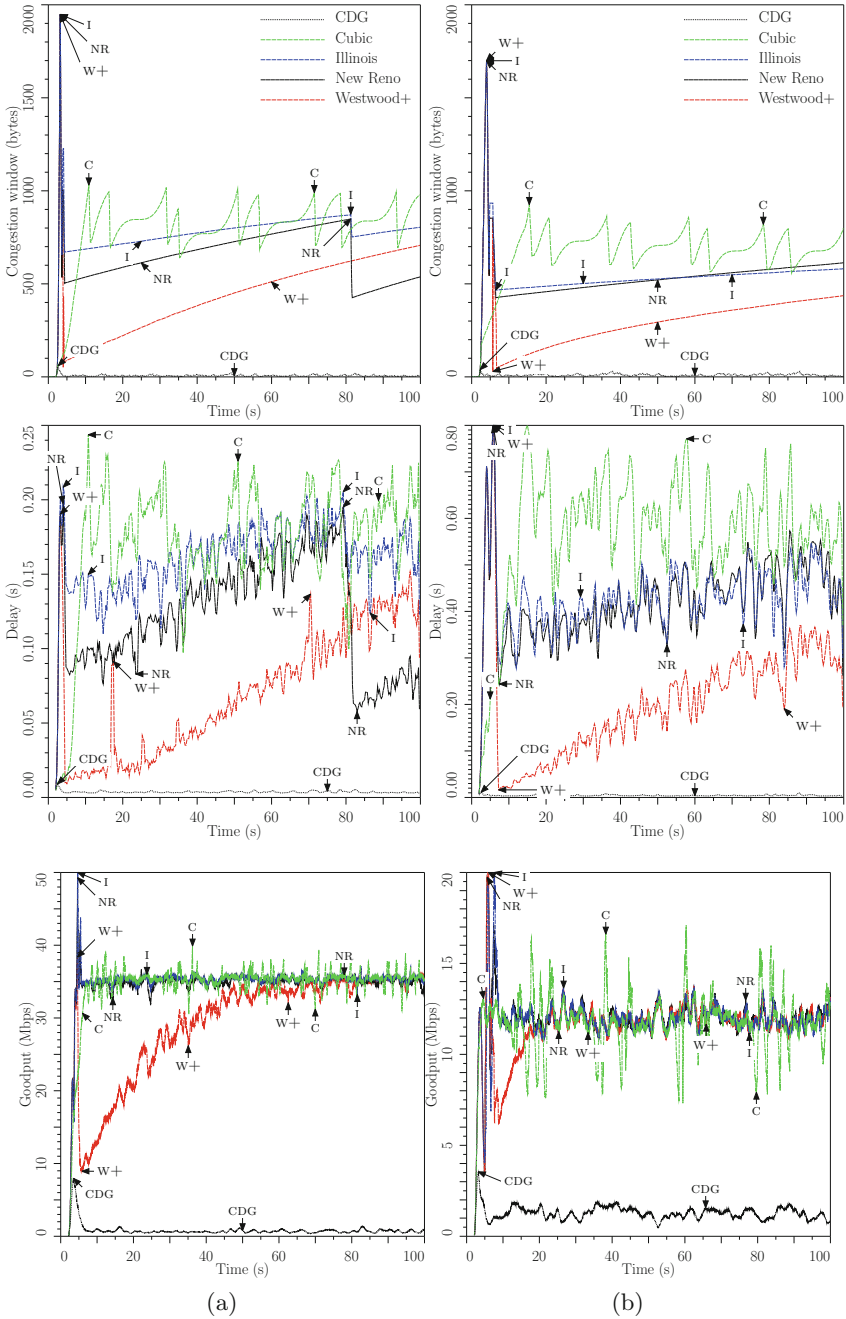
(VI) Emulated support to correlate mobility-based scenarios: Since the previous scenario is limited in terms of speed, the emulated testbed (VI point in Fig. 1) due to the utilization of real UEs and the ability to emulate movement, is capable of confirming and clarifying performance trade-offs that in simulated environment could be blurry. In order to better understand the evolution of different performance-related parameters, in the server, the `cwnd`, RTT, outstanding data and goodput have been collected.

## 5 Analysis of the Interactions Observed in Different Scenarios

This section is divided in five main parts: implication of cross-traffic and responsiveness of TCP (with scenario I), the start-up performance (with scenario II), both decreasing quality and increasing quality movement scenarios (with scenario III and IV), the correlation of findings in the controlled deployment (with scenario V) and finally, the correlation of findings regarding mobility scenarios' over emulated testbed (with scenario VI).

### 5.1 Cross-Traffic Impact and Responsiveness of TCP

This subsection is responsible for covering different kind of traffic loads and behaviors while the UE is static. The location of the UE among the different measurements is the same and thus, the results are comparable. The subsection is divided in three main experiments: the comparison between a single-UE without cross-traffic and a loaded network, the impact of short flows and finally, sudden increase and decrease of the available capacity.



**Fig. 2.** Performance comparison of the selected CCAs: (a) Base single flow behavior; (b) Single flow behavior over loaded network.

### Base Behavior and Behavior in a Loaded Network

According to the selected position, the UE has a maximum throughput around the half of the total maximum (35 Mbps). Different experimental trials are carried out with and without background traffic to study the responsiveness of TCP and infer whether the background traffic has the same impact among the CCAs or not. In order to make easier the reading, Table 3 gathers the most important point of the following explanation.

The three subfigures on the left of Fig. 2 depict the results regarding the scenario with no background traffic. The differences between the loss-based TCP variants and delay-based ones are remarkable even in such a simplified scenario. Loss-based implementations manage to achieve the maximum capacity and create a long standing queue delay (up to 250 ms), whereas delay-based variants, such as CDG, keep the delay controlled but fail while trying to reach full resource utilization. In the case of Westwood+, it is clear that the applied back-off after Slow-Start is very drastic and due to this, it takes longer to ramp-up. Illinois minimally reduces the *cwnd*, causing huge standing queue delay comparing with more conservative implementations like NewReno. In the case of CUBIC, it suffers for the deficient behavior of Hybrid Slow-Start. The mechanism exits to the congestion avoidance phase in an early stage and therefore reduces its growth pace far from the maximum achievable capacity, severely impacting in the time it takes to converge.

The three subfigures on the right of Fig. 2 show the outcome for the same scenario but with background traffic. The total target load of the background

**Table 3.** Findings wrap-up in base behavior and behavior in a loaded network

CCA	Conditions	Behavior
CUBIC	Base behavior	Slightly suffers for the deficient behavior of Hybrid Slow-Start
	Loaded network	No impact of Hybrid Slow-Start
NewReno	Base behavior	Easily achieves maximum capacity
	Loaded network	Similar behavior but with higher delay and more unstable goodput
Illinois	Base behavior	Easily achieves maximum capacity. However, it creates a huge standing queue
	Loaded network	Very similar to NewReno but with slightly higher delay
CDG	Base behavior	Keeps the delay controlled but fails while trying to reach full resource utilization
	Loaded network	The differences with loss-based CCAs are reduced
Westwood+	Base behavior	Very aggressive back-off that impacts the time needed to ramp-up
	Loaded network	The impact of the back-off application is minimized

traffic is set to the 50% of the link capacity. The capacity reduction minimizes the performance gap between loss-based and delay-based variants and still, the more capacity a CCA gets, the harder impact it inflicts in terms of queuing delay (Illinois as an example). Big differences appear comparing with the base example without background traffic, mostly related to a significant increment in the queuing delay and the reduction of the gap in terms of capacity to reflect the differences amongst the CCAs. RTT-clocked CCAs suffer due to a lengthen of the time between implementation decisions. In contrast, CUBIC behaves better because it does not suffer for RTT increase. The scenario itself due to its reduction in the available capacity cushions the underperformance of Hybrid Slow-Start.

### Short Flows Study

Live measurements have shown that many flows over Internet are small (90% of downstreams carry no more than 35.9 KB of data [4]). Therefore it is important to assess the impact that such load distribution has in final performance. In order to do so, the previous foreground TCP flow must be replaced by a succession of short flows following an exponential distribution regarding their amount of data.

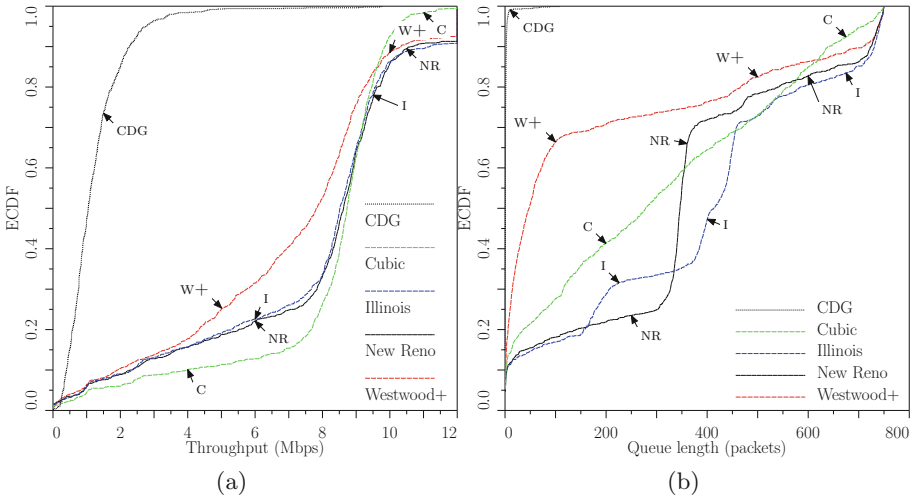
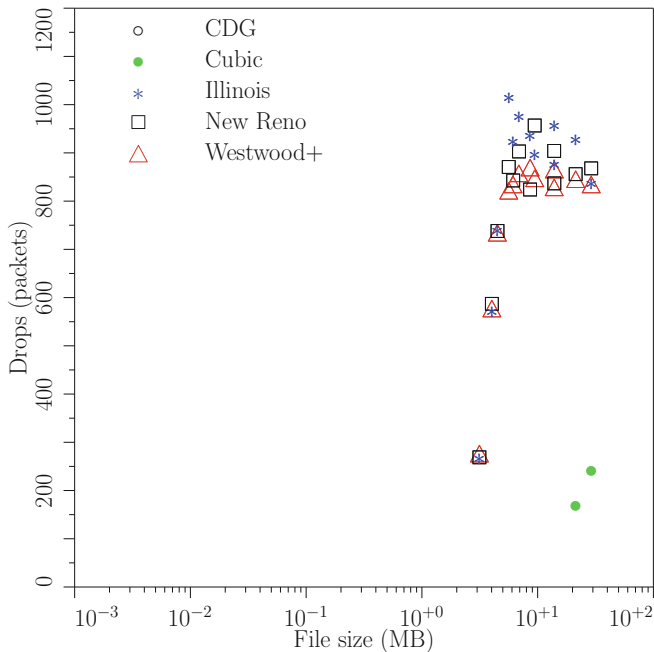


Fig. 3. Throughput and queue size ECDF at 700 m

The Fig. 3 represent as an Empirical Cumulative Distribution Function (ECDF) the results obtained regarding the achieved throughput and standing queue size. In Fig. 4 the size of the flows and the number of induced drops are correlated.

In Fig. 3a, it is clear that the achieved throughput is very similar among most the CCAs and their differences really appear regarding the amount of enqueued packets in Fig. 3b. The delay-based variant, CDG, successfully limits



**Fig. 4.** eNodeB drops at 700 m

the enqueued packets while loss-based implementations overshoot causing a great standing queue. Due to the detected behavior of Westwood+ in the beginning of the transmissions and the short duration of the flows, it prompts little ability to inject packets in the eNodeB. In contrast, NewReno, Illinois and CUBIC happen to be the average solutions. If we compare two deficient solutions such as CUBIC and Westwood+, we clearly see that even with short flows and therefore quick transmission duration, the premature exit from Slow-Start for the former performs better than the excessive back-off of the latter.

Considering the reported findings, Fig. 4 shows the number of packets that have exceeded the queue size with each flow size, being therefore dropped. It is clear that the more aggressive the CCA is, the more packet are dropped by the eNodeB. Illinois for instance has a more aggressive behavior in congestion avoidance phase. It enqueues more packets and gets more packets dropped. As a result Illinois suffers on average 100 more dropped packets than any other TCP candidate. Once again, the behavior of Hybrid Slow-Start is clearly shown. If we avoid the fact that the transmissions with Hybrid Slow-Start take slightly more time to be completed, it only suffers drops with longer transmissions and when it has congestion events, the number of losses are very few. With loss-based AIMD mechanisms, the drop packets metric appears to be directly related to the aggressiveness and back-off strategy. NewReno and Westwood+ have quite similar results (Table 4).

**Table 4.** Findings wrap-up in short flows study

CCA	Behavior
CUBIC	Thanks to the underperformance of Hybrid Slow-Start, it only suffers drops with longer transmissions and when it has congestion events, the number of losses are very few
NewReno	The average solution
Illinois	Very aggressive behavior that results in 100 more dropped packets on average
CDG	Successfully limits the enqueued packets
Westwood+	In the beginning of the transmissions and the short duration of the flows, it prompts little ability to inject packets due to the aggressive back-off

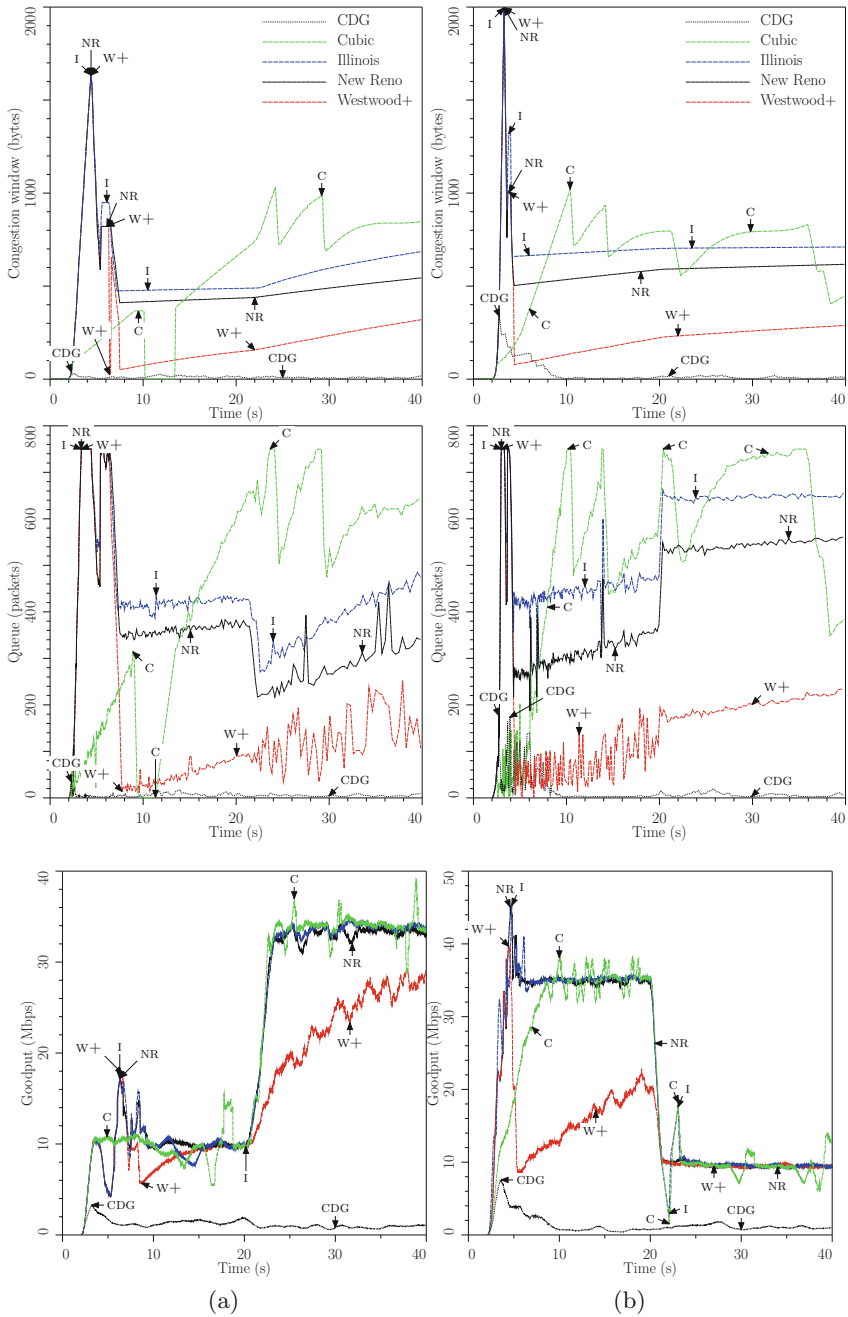
### Sudden Increase and Decrease of the Available Capacity

Once the main features of the CCAs have been detected in loaded scenarios in comparison with the base behavior as well as the impact of different short flows on the drop rate, it is important to study the responsiveness of CCAs in big and sudden capacity changes. To this end, two type of simulations are carried out: with the background traffic being stopped at 20s of the test and with the background traffic being started at 20s of the test. In order to make easier the reading, Table 5 gathers the most important point of the following explanation.

On the one hand, the left part of Fig. 5 shows the results regarding the scenario with a sudden capacity increase. In general, as soon as the capacity increases, the queue size is lowered due to a release of previously enqueued packets. It is clear that loss-based CCAs quickly respond to an additional bandwidth assignment. However, Westwood+ still suffers from the excessive reduction of the *cwnd* after the Slow-Start phase. During the congestion avoidance phase, its AIMD mechanism is very conservative and the enqueued packets tend to be almost 0, therefore with a new and greater achievable capacity, the adaptation ability of the CCA is very weak. In the case of delay-based variants, since they mainly focus on reducing the delay over path, they usually fail to increase their pace and thus, the new available capacity is wasted.

On the other hand, the right part of Fig. 5 depicts the case in which the background traffic is activated at 20s. Due to the sudden reduction of available capacity, the queue size suffer an instant increment because of the relation between the same number of incoming packets to the eNodeB and the drastic reduction of outgoing ones. The Fig. 5 clearly shows that all CCAs but CDG are able to successfully react to the capacity reduction. However, in some cases such as CUBIC, the CCA takes more time to stabilize to the new pace.

These simulations reflect that most CCAs, even delay-based implementations, are capable of reducing their throughput when sudden available capacity decreases happen but delay-based variants struggle to adapt their pace to bandwidth increases.



**Fig. 5.** Performance comparison of the selected CCAs: (a) Sudden capacity increase; (b) Sudden capacity decrease.

**Table 5.** Findings wrap-up in sudden increase and decrease of the available capacity

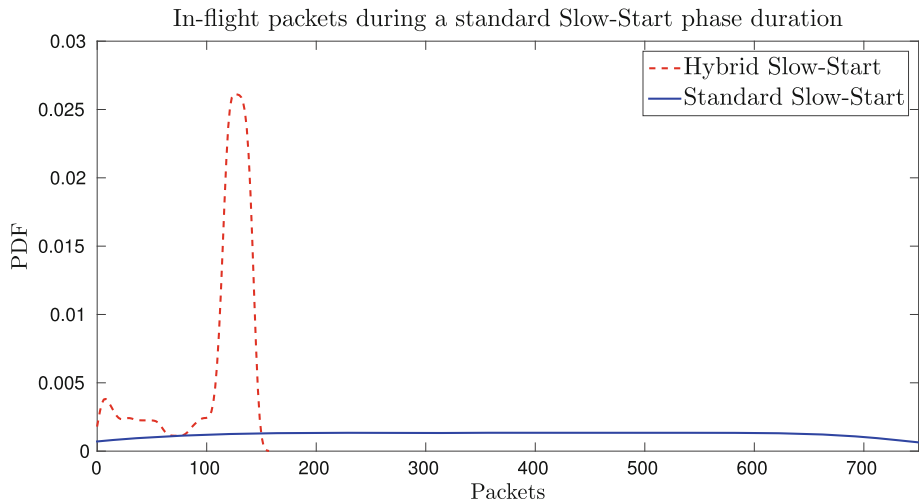
CCA	Conditions	Behavior
CUBIC	Inc. available cap.	Good performance without the impact of Hybrid Slow-Start due to the low available capacity at the beginning of the transmission
	Dec. available cap.	Impact of Hybrid Slow-Start in the beginning. Aggressive behavior in congestion avoidance phase that leads to an instant huge increment of queue size while reducing the available capacity
NewReno	Inc. available cap.	Good responsiveness and average delay impact
	Dec. available cap.	Average loss-based solution that suffers and instant standing queue increase while reducing the available capacity
Illinois	Inc. available cap.	Good responsiveness and greater induced delay than NewReno
	Dec. available cap.	Its aggressiveness is harmful in this scenario and takes some time to stabilize the goodput
CDG	Inc. available cap.	Fails to increase its pace and thus, the new available capacity is wasted
	Dec. available cap.	Bad performance in terms of goodput but full control of the delay that is always close to the baseline delay
Westwood+	Inc. available cap.	Its AIMD mechanism is very conservative and the enqueued packets tend to be very few, being not capable of responding to a sudden greater capacity assignment
	Dec. available cap.	The combination of its dynamics (with a slow ramp-up ability) and the available capacity reduction happen to get the best performance due to the achievement of the maximum goodput and the lowest impact in terms of delay

## 5.2 Start-Up Performance

In very simplified scenarios, we have seen that the behavior of Standard Slow-Start and Hybrid Slow-Start differs leading in some occasions to a successful avoidance of massive losses with Hybrid Slow-Start. However, the LTE cells are usually more crowded and therefore the UEs could inflict more delay as cross-traffic that could impact Hybrid Slow-Start. The target is to assess whether the internal mechanisms of Hybrid Slow-Start could provoke an early exit from the standard ramp-up, following to a slow increment of the *cwnd* and therefore a significant underutilization of radio resources or not.

We first measured the convergence behavior of the Standard Slow-Start, recording the packets that are in-flight at every moment. Later, we assessed





**Fig. 6.** Hybrid Slow-Start impact in mobile networks: injected packets during a standard Slow-Start period.

the same for Hybrid Slow-Start. Figure 6 shows the probability density function (PDF) of the number of injected packets for both mechanisms in the time Standard Slow-Start takes to converge. This is, we would compare in the fastest convergence period of time, the ability of both methods to put packets in-flight.

Figure 6 shows the behavior of Standard Slow-Start has a equal distribution of packets in flight, whereas Hybrid has an imbalanced distribution presumably formed by the period of time in which Hybrid Slow-Start has ramped-up as Standard Slow-Start and the period after detecting a delay variation and behaving under the incremental pace of congestion avoidance phase. The distribution represents the huge difference between both methods regarding the ability to inject packets which leads to a extrapolation of the time needed to converge or achieve the maximum capacity from the beginning of the transmission.

It is clear that not only in simplified scenarios, but also in multi-UE measurements, Hybrid Slow-Start suffers due to the detection of delay increment and the early trigger of exit condition from fast ramp-up. So, under some delay variability circumstances Hybrid Slow-Start slows-down the ramp-up of TCP. In some situations, this effect could underutilize the available radio resources and lengthens the time needed to converge, directly impacting on the quality experienced by users (QoE).

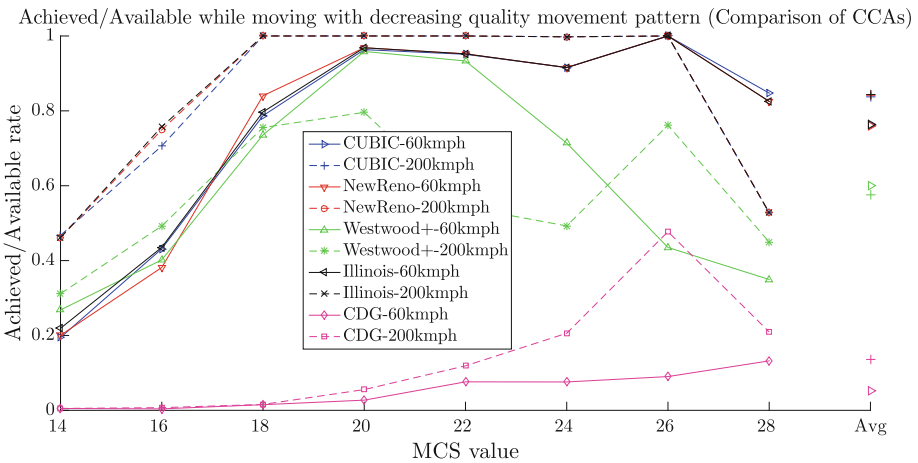
### 5.3 Mobility Performance

This subsection covers the analysis of decreasing quality and increasing quality movement for the selection of CCAs. Even though it has been proven in Subsect. 5.1 that some CCAs fail in mobile networks (Westwood+ and CDG), they have been kept for comparison and confirmation purposes. In order to measure the ability

or inability of distinct TCP implementation to take advantage of radio resources, the results will be presented as the portion of *tbSize* that has been actually utilized every Transmission Time Interval (TTI). In other words, since the TTI is commonly configured in 1 ms, the portion of *tbSize* will show how many bits are used for the UE every millisecond. Considering that different MCS values lead to have distinct available capacity and therefore a different achievable throughput, the analysis is divided in MCS ranges.

**Decreasing Quality Movement**

The decreasing quality movement scenario stands for the continuous movement evolution of a certain UE from the eNodeB to a further location. In other words, on average the obtained SINR due to the distance from the UE to the eNodeB and the fading will have a tendency to be worse. So will be the reported CQI and the assigned MCS (instead of worse, it is a tendency to become a more robust modulation). In such a transition, the CCA will need to adapt to the different available capacities. Figure 7 shows the difference between the available capacity and the achieved MCS for different CCAs under distinct speeds, all classified by average MCS.



**Fig. 7.** Achieved/Available capacity at different speeds for different TCP variants (decreasing quality movement).

Figure 7 clearly depicts three main areas:

Slow-Start phase: Located in the coverage zone associated to MCS 28, during the transmission establishment and first ramp-up, the *cwnd* is not great enough to take full advantage of available radio resources. Considering that Standard Slow-Start converges very fast, the MCS 28 area also takes the first back-off application. For that reason Westwood+ or NewReno among others do not report the same result. Since the distance associated with a MCS is covered a lot faster

at 200 km/h, the *cwnd* has no time to grow quickly enough and therefore the impact of ramp-up is more significant for the scenario at 200 km/h, prompting a lower value of achieved/available for this speed.

“Bufferbloat” area: While in the area between MCS 26 to 18-20, the CCAs are able to take advantage of already enqueued packets in the eNodeB (*bufferbloat* effect). However the effect itself has a drawback in relation to the inflicted delay. This feature is more present in the examples at 60 km/h. The time spent in each MCS area makes it possible to the TCP variant to inject packets throughout a longer time, getting loss packets and requiring to recover from them under high-delay conditions and therefore, not allowing the CCA to achieve maximum capacity.

Queue draining zone: Regardless the speed, it is an area in which the radio conditions are not good enough to maintain a full utilization of resources. Even though the average MCS values are between 18 to 14, fading conditions force the eNodeB to operate with very low MCS values (achieving sometimes MCS 4 and 6) in some drastic fades. With each sudden fade, it is easier to receive more robust modulations, leading the packets to need stronger segmentation. As a side effect, both the queue size of the eNodeB and the delay increase. The recovery of losses in such network conditions is also a harmful process for TCP that leads to queue starvation events. When it comes to faster UE scenario, the eNodeB is able to lengthen the utilization of previously enqueued packet to further positions, therefore, the draining effect is slower or at least happens in further positions.

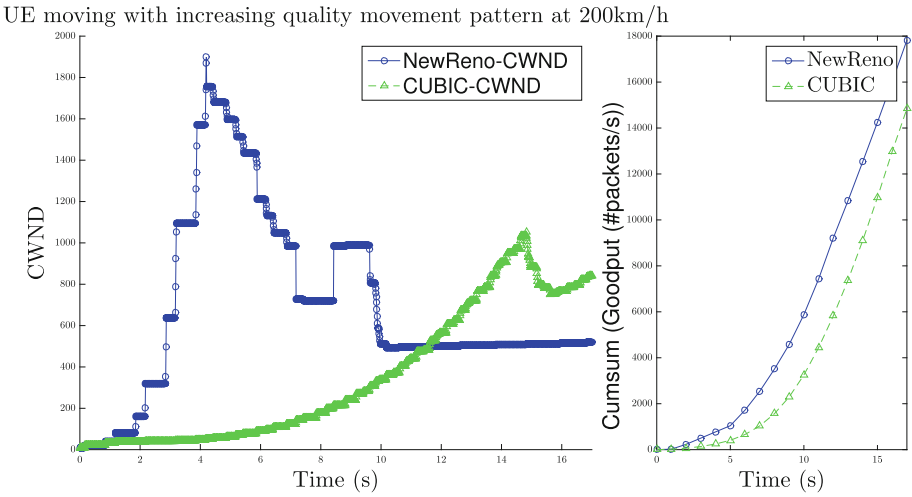
Figure 7 shows that in decreasing quality movement the differences in loss-based CCAs are minimum, getting more credit of aggressiveness at 60 km/h and RTT-synchronization at 200 km/h (NewReno and Illinois over CUBIC). Once again and even in a scenario that moves towards worse radio position, the delay-based variants have demonstrated to be unable to cope with the delay variability of LTE. CDG maintains a RTT close to the baseline RTT but underutilizes most of the assigned bandwidth. In the case of Westwood+, even though it is a scenario that helps get the maximum capacity to the weak AIMD mechanisms due to its continuous achievable capacity reduce, it takes very long time to achieve such a task at 60 km/h and it is not capable of doing so at 200 km/h.

### Increasing Quality Movement

Once analyzed the decreasing quality movement and the behavior of different CCAs under distinct speeds, it is necessary to study the increasing quality movement in a constant evolution of the channel quality to better positions. Considering the findings in decreasing quality movement, it is important to determine whether the different methods of Slow-Start equally struggle under challenging radio condition or not and analyze whether the aggressiveness of TCP overshoots sufficient packets to serve a continuous greater capacity or not.

Trying to better explain the effects of this scenario in the beginning of the transmission and the relation between *cwnd* evolution and achieved goodput, Fig. 8 represents the relation between them. The graphs have been split for the

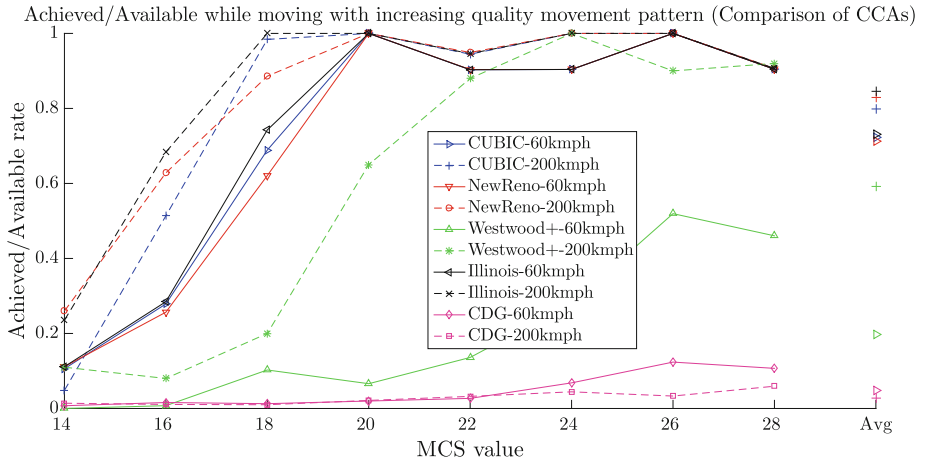
better understanding in two blocks: the result in relation to the *cwnd* evolution is on the left and goodput's cumulative sum on the right. Figure 8 depicts the behavior difference between CUBIC with Hybrid Slow-Start and NewReno with Standard Slow-Start. It is clear that the network conditions are challenging because even in Standard Slow-Start the shape of the *cwnd* is very stepped. In such conditions in which the delay variability is also a hard drawback to tackle, the Hybrid Slow-Start mechanism detects an increment in the delay that is considered enough to trigger an early exit to congestion avoidance phase. The resultant cumulative goodput of both CCAs is represented on the right where the graphs shows a big outcome gap between both methods. Once again the underperformance of Hybrid Slow-Start is shown. Besides, in this case the early exit of fast ramp-up is provoked in a single-UE scenario in which the movement and fading are the only sources that vary the delay.



**Fig. 8.** NewReno vs. CUBIC in increasing quality scenarios at 200 km/h.

Considering the explained effect regarding how Hybrid Slow-Start could affect the performance, we will now proceed to study the performance differences under different speeds between NewReno, CUBIC, Westwood+, Illinois and CDG, classified by average MCS levels (see Fig. 9). At a first glance, the figure looks very similar to Fig. 7, but some differences are present. The behavior of such scenario is divided in two areas.

Ramp-up phase: The hardest radio conditions for the channel are present from MCS 14 to 18. In such a challenging conditions the CCAs initialize the transmission and employ the selected Slow-Start method in a try to ramp-up and convergence as fast as possible without inducing a bursty loss event. As seen beforehand, at 200 km/h the performances of Standard Sow-Start and Hybrid



**Fig. 9.** Achieved/Available capacity at different speeds for different TCP variants (increasing quality movement).

Slow-Start are completely different, leading to a better utilization of the network resources in the case of Standard Slow-Start. Besides, in MCS 14 in some occasions it is not only present the Slow-Start phase but part of the first back-off and application of congestion phase as well. The growth limitation is comparatively very similar for 60 km/h and 200 km/h during this phase and establishes an undodgeable boundary for loss recovery. However, in faster scenarios the time spent in weakest radio conditions is less and the impact of such challenging conditions is less significant in the final outcome. Apart from that, in the case of Standard Slow-Start, at 200 km/h the first loss event will happen in better radio conditions than for 60 km/h and therefore, the ability to recover the lost packets is greater at 200 km/h.

Stationary area: Throughout MCS 20 to 28, TCP is able to take close to full advantage of available capacity. However, it has to be mentioned that, due to that transition speed and applied back-offs while recovering from losses, the CCAs are not able to rise sufficiently the *cwnd*, causing some channel underutilization.

Even though, in general, the CCAs follow the identified phases, there are some differences among the CCAs that result in a distinct outcome for the same network conditions (see the wrap-up Table 6).

Different Slow-Start methods affect the availability to take full advantage of radio resources, dividing the performance in two major groups. (1) Among the CCAs with same Slow-Start phase, some differences appear in MCS 14 due to the different AIMD policy applied when a loss is detected. As stated before, the higher speed, the comparatively longer Slow-Start phase and therefore, a decrease of the loss recovery effect due to the recovery taking place in better radio conditions. (2) For Hybrid Slow-Start mechanism, a difference between CUBIC and CDG appear regarding the delay sensitivity to quit fast ramp-up

**Table 6.** Findings wrap-up in mobility scenarios

CCA	Behavior
CUBIC	It only suffers the impact of Hybrid Slow-Start in the very beginning of the transmission in increasing quality movement pattern under high speed (see Fig. 8)
NewReno	Very good performance in terms of achieved available capacity in simplified single-user mobility scenarios
Illinois	The best results due to the combination of the delay-awareness and the aggressiveness
CDG	It has demonstrated very weak performance over cellular access under mobility in terms of bandwidth utilization
Westwood+	Only able to reach full utilization after a long ramp-up period in decreasing quality movement at 60 km/h and in increasing quality movement at 200 km/h

(as stated in Subsect. 3.1). In relation to the effect of speed, the faster the UE moves, the higher delay variability and therefore quicker skip to a slow increase phase, suffering more wasted bandwidth utilization at 200 km/h.

Westwood+ and CDG have been proven to be not adequate for mobile networks. The former is able to reach full utilization after a long ramp-up period in decreasing quality movement at 60 km/h and in increasing quality movement at 200 km/h. The time spent is due to a poor available bandwidth estimation and consequent drastic back-off policy. At 200 km/h in decreasing quality scenario the CCA does not allow sufficient time so as to achieve the maximum bandwidth. On the contrary, in increasing quality movement, the fastest scenario allows the CCA get the maximum capacity. The latter has demonstrated very weak performance over cellular access. It has to be underlined that the main objective of the CCA regarding the control of end-to-end delay is fulfilled, however, regardless the speed and scenario, the CCA has not been able to rise to the 10% of the available capacity, consequently leading to a 90% of resource underutilization. Therefore CDG is not suitable for mobile network as is configured now.

The group formed by NewReno, CUBIC and Illinois have shown a very successful performance regardless the speed and movement pattern. As stated in previous explanation, the shortening of the challenging periods could make a difference in terms of greater achieved capacity. On average (see average values on the right) at 60 km/h 3 CCAs are very similar and it is only under 200 km/h speed circumstances when CUBIC performs poorly due to Hybrid Slow-Start and Illinois get a slight advantage of its delay-awareness to make the most of using available resources.

All the gathered results are consistent with the findings regarding decreasing quality movement (in Subsubsect. 5.3), the performance of different Slow-Start methods (in Subsect. 5.2) and the preliminary analysis in regards to the impact of different cross-traffic in the performance of CCAs. However, since the results

have been obtained in a single LTE deployment, it is important to determine to which extent our findings could be extrapolated as a general-purpose behavior of CCAs and whether the results are biased towards the simulated/emulated testbed or not.

### 5.4 Correlation of TCP Behavior over w-iLab.t LTE Testbed

The current subsection aims at representing and explaining the behavior of a selection of CCAs over the controlled LTE testbed called w-iLab.t. Since the deployment is formed with completely real equipment (i.e. UEs, eNodeBs, femtocells, servers), the internal mechanisms of LTE and the interaction with TCP are closer to real-world behavior and therefore the variability is presumably higher comparing with simulated environment. Thus, such testbed allows carrying out experiments that represent the performance of the reality in a smaller scale. CDG was removed from the comparison set for its incompatibility with mobile networks. Westwood+ is kept in the selection of CCAs to confirm or deny the underperformance under more variable circumstances.

We configured three different paths to be followed by the robots with decreasing quality and increasing quality movements. The location of those movement patterns were located in different places of the femtocell, having a pattern close to the eNodeB, another one close to the spacial limits of the testbed and a third one in the middle of the previous two. After ten experiments over the different configurations/patterns, we gathered the following average throughput values for CUBIC, NewReno, Westwood+ and Illinois.

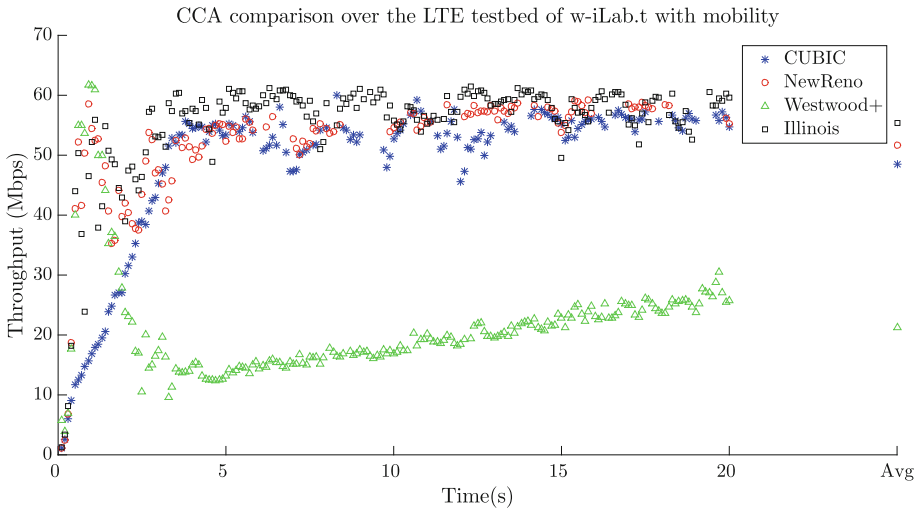


Fig. 10. CCA comparison over w-iLab-t under mobility circumstances.

Figure 10 shows that the previous findings in ns-3 were accurate enough to explain the possible effect of CCAs in other LTE deployments. In fact, some deficiencies such as the ones regarding Hybrid Slow-Start and Westwood+ are more harmful than in simulation environment, causing a greater gap between the available capacity and the achieved one.

In general three are the most important features to be underlined. First, the drastic back-off application of Westwood+ leads the CCA to be incapable of achieving the maximum capacity even within 20s of transmission. Looking at the growth tendency, the CCA may well take around 1 min to convergence which is an unacceptable value in order to provide a good service to the UEs. Second, the underperformance of Hybrid Slow-Start is more remarkable in this testbed and the results prompt a convergence time around 4.5 s. The performance difference with Standard Slow-Start (present in NewReno and Illinois) could be cushioned if the transmission is long enough (average value of CUBIC is close to NewReno or Illinois in 20 s transmission). However, the impact in short-lived flows would be more notable. Third, the performance of NewReno and Illinois are very similar and the only distinction appear due to the greater aggressiveness of Illinois for its delay-awareness. Nevertheless, the utilized femtocells give a very good channel quality regardless the mobility pattern, movement patterns or speed. Thus, the “signal quality rings” that are present in real-world could not be represented. Therefore, in order to better understand the performance tradeoff of CUBIC, NewReno and Illinois in congestion avoidance phase during mobility circumstances, an additional analysis was demanded.

### 5.5 Performance Tradeoff of Selected TCP Variants Under Mobility in Emulated Testbed

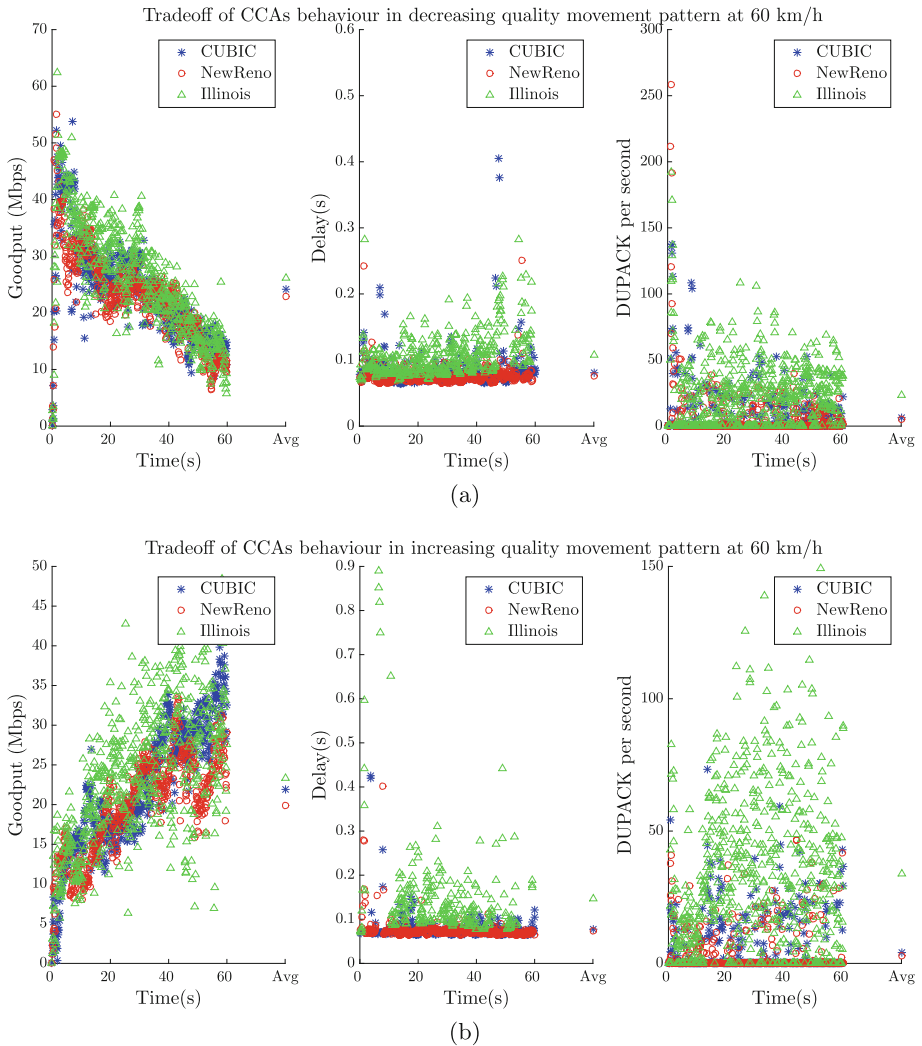
Once the previous findings regarding the behavior of CCAs have been demonstrated in a controlled testbed, this subsection covers the comparison analysis of most adequate TCP flavors (CUBIC, NewReno and Illinois) over emulated testbed with mobile scenarios of decreasing quality and increasing quality movement. The previous scenarios have shown that CUBIC, NewReno and Illinois have a very close outcome. Therefore, this subsection will serve not only as a confirmation step of the findings in another testbed but to also carry out experiments in mobility circumstances with a realistic representation of “signal quality rings”.

The testbed itself is not able to emulate movement due to the fixed position of the UE attached to a radio cable. Nonetheless, a computer that plays the role of a experiment controlled, is capable of establishing the baseline SINR at any moment. Besides, the lte-in-a-box called Aeroflex 7100 applies a fading pattern to such a variable baseline SINR, modelling this way the effect of movement with a static UE.

To help decide the best timing for different baseline SINR values, averaged SINR traces obtained from ns-3 with a UE moving in decreasing quality and increasing quality movement patterns at 60 km/h are used. In order to give more realism to the experiments, the EVA60 fading model in Aeroflex 7100 are applied.



We have decided to only use the scenarios at 60 km/h due to the result equality in ns-3. At 200 km/h the differences among CCAs were noticeable. Therefore, these experiments add additional information to the previous inconclusive outcomes and gives more insight regarding the differences among the selected CC. Figure 11 depicts the average goodput, end-to-end delay and duplicated ACK (DUPACK) events per second as a sign of congestion for decreasing quality movement at the top and for increasing quality movement at the bottom.



**Fig. 11.** Performance tradeoff of CCAs in the emulated testbed at 60 km/h: (a) Decreasing quality movement; (b) Increasing quality movement.

In this case, the differences among the CCAs are noticeable for both movement patterns. The goodput results do not prompt any new feature and classify the performance of the selected TCP implementations from better to worse as Illinois, CUBIC and NewReno. This outcome equally applies for decreasing quality and increasing quality movement, getting slightly more difference in increasing quality movement due to the continuous capacity increase and the availability to cushion overshoots. The simplest way to proceed would be to say that Illinois is the best amongst the CCAs. However, depending on the performance objective, the decision could be another one. The reasons are manifold. First, even though the goodput performance is better for Illinois, it induces delay and the consequent packet losses are a way larger than in the examples of CUBIC and NewReno. Second, if we compare the overall performance of CUBIC and NewReno, we see that in spite of the delay and DUPACK events being very similar, CUBIC makes the most in terms of goodput. Therefore, trying to avoid massive packet losses and delay infliction, the selection of CUBIC would be more desirable in this simple comparison. Third, for comparison purposes, since the objective of this scenario was the understanding of congestion avoidance phases and the adaptability to mobile LTE scenarios, the Hybrid Slow-Start mechanism was disabled. Taking into account this detail and depending on the requirements of the application, the selection of NewReno could not be discarded. To conclude this tradeoff study, it is clear that Illinois, CUBIC and NewReno have very similar results, but it cannot be easily decided whether one is better than the other because each of them has its “bright side” and drawback.

## 6 Conclusion

This chapter has tried to shed some light in the explanation of CCAs adaptability to different mobile network situation including the implication of different type of cross-traffics, the start-up phase and mobile UEs with increasing quality and decreasing quality movement patterns. The chapter has also included different LTE deployments so as to confirm and clarify the obtained result in the simulated environment. Table 7 wrap-ups the detected findings and confirmations of CCAs behavior under distinct circumstances.

Simple static experimentation with different background traffic profiles and behaviors has demonstrated that loss-based TCP mechanisms reach the maximum capacity quicker than delay-based variants. The former achieves greater throughputs but fails limiting the standing queue size and therefore inflict severe delays. The latter is able to keep the end-to-end delay close to the baseline delay value but struggle to ramp-up or speed up its injection pace, wasting this way a great amount of the available radio resources.

Different scenarios have shown the huge impact that Hybrid Slow-Start mechanism has under some delay variability circumstances. Having in mind that the delay’s instability is one of the main features in mobile networks, Hybrid Slow-Start is capable of slowing down the start-up phase leading to a bad resource utilization. Taking into account the widespread usage of CUBIC due to the

**Table 7.** Findings wrap-up

CCA	Simulated env.	Controlled testbed	Emulated testbed
CUBIC	It suffers from its delay sensitivity in Hybrid Slow-Start phase, being very harmful and provoking mobile network capabilities underutilization	Confirmed behavior of Hybrid Slow-Start with even greater impact. Long transmission would suffer such effect but it would be more significant in short-lived ones	In simplified mobility scenarios, the cubical congestion avoidance phase allows a good available capacity utilization while the delay is lower than with Illinois (closest CCA in terms of goodput)
NewReno	I has responded very positively to different network situation, showing that it is still a good TCP candidate to be utilized in certain situations. Its speed weaknesses in fixed networks could result in a valuable feature in mobile networks	Confirmation of the good performance	Some precise mobility circumstances have shown a deficient performance of NewReno leading to resource underutilization and may well indicate which mobile network circumstances are not suitable for the protocol
Illinois	Very similar to the performance of NewReno with bigger impact in delay due to its greater aggressiveness. Such aggressiveness allows performing slightly better in scenarios that require rapid adaptability (under mobility)	Overall performance of Illinois has been demonstrated, showing in close-to-the reality scenarios better performance than NewReno in terms of achieved throughput	Under mobility circumstances, a slight gap increment in the outcome of Illinois and NewReno has been found. The results may indicate that under more realistic conditions the breach will be even greater
CDG	It has demonstrated very weak performance with all scenarios over mobile networks in terms of bandwidth utilization. However it has shown a good control of the delay keeping it close to the baseline delay	–	–
Westwood+	Found a problem with a drastic back-off application that is capable of provoking underutilization of the radio resources under certain network situations	Confirmation of the findings noticing even greater impact of the deficiency. The closer to real-world, the poorer assessment of the available capacity and therefore, the more deficient the application of the back-off policy	–

presence of it by default in most Web servers, the problem is even worse. Even though, long transmissions suffer the impact of the underperformance of Hybrid Slow-Start, the effect is greater in the case of short-lived flows.

Regarding the mobility scenarios, two have been studied. In decreasing quality movement most CCAs are able to achieve the maximum capacity during good radio conditions and they lengthen the utilization of previously enqueued packets while running towards worse channel qualities. At higher speeds, the already enqueued packets are driven to further positions comparing with lower speeds, improving the average capacity utilization. In increasing quality movement, regardless the speed, the transmission initialization and first ramp-up happens in very challenging radio conditions, requiring CCAs availability to scale, recover from losses and AIMD mechanisms' suitability to make the most of available capacity.

In relation to the specific features of each CCAs' adaptability, several findings have to be mentioned: (1) CUBIC suffers from its delay sensitivity in Hybrid Slow-Start phase, being very harmful and provoking mobile network capabilities underutilization. (2) CDG keeps the delay close to the baseline delay value but is incapable of growing its pace in order to utilize greater capacities. In its current state is not suitable for mobile networks and it could more suitable for wired networks where the delay variation is not that abrupt. However, the delay boundaries of the protocol may well be adapted to cellular networks' constraints. (3) Westwood+ has shown to be incapable to properly estimate the available bandwidth, leading to big *cwnd* reductions and the necessity to grow-up from very low values and very weak AIMD incremental pace. The adaptation of the estimation is required in order to make it suitable for mobile networks. (4) NewReno and Illinois have demonstrated to beat the other CCAs (apart from CUBIC in some situations) under different loads, traffic patterns, mobility and speed contexts. Even though in simulated environment the only detected difference has appeared in increasing quality scenario in which the delay-awareness and greater aggressiveness has given to Illinois the best performance regarding the use of available capacity, in emulated testbed the differences have been also present in decreasing quality movement. Since the emulated testbed has shown a slight gap increment in the outcome of Illinois and NewReno, the results may indicate that under more realistic conditions the breach will be even greater.

The feature-based findings have been confirmed over the LTE deployment of w-iLab.t and the performance tradeoff of the best CCAs has been explained under mobility circumstances in order to give insights regarding the appropriate selection depending on the application requirements. This chapter has given an overview of the behavior of the different TCP mechanisms in a LTE network under different circumstances. This work might be of value as a validation of the performance of different CCAs and as an indication of fruitful directions for the improvement of TCP congestion control over cellular networks.

Some knowledge from the network would help TCP decide the best strategy in accordance to the network conditions. The envisioned scenario is aligned with the main features of mobile edge computing (MEC) management that would

allow removing as much end-to-end TCP variant dependency as possible. In the same way, other initiatives such as QUIC [24] that propose transport services in the user-space of the operating system with TCP-alike CCAs on top of UDP (UDP as a substrate) could take advantage of this comprehensive analysis in order to select the most appropriate TCP candidate (i.e. depending on multi-criteria that considers both network state and application requirements) in each network conditions and enable such TCP-alike implementation.

**Acknowledgments.** This chapter has been possible thanks to the Cost Action IC1304 through the STSMs entitled “Evaluation of QoE-optimized transport protocols on cellular access” and “Evaluation of modern transport protocols over iMinds LTE facilities”. The work has been partially funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 644399 (MONROE) through the open call project MARI-L and by the Spanish Ministerio de Economía y Competitividad (MINECO) under grant TEC2016-80090-C2-2-R (5RANVIR). The views expressed are solely those of the author(s). The authors would like to thank Rémi Robert for the useful discussions on the experiments and his work and dedication in the simulated environment.

## References

1. Cisco: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021. Technical report, Cisco (2017)
2. Callegari, C., et al.: Behavior analysis of TCP linux variants. *Comput. Netw.* **56**(1), 462–476 (2012)
3. Huang, J., et al.: A close examination of performance and power characteristics of 4G LTE networks. In: *MobiSys 2012*, pp. 225–238. ACM, New York (2012)
4. Huang, J., et al.: An in-depth study of LTE: effect of network protocol and application behavior on performance. *SIGCOMM Comput. Commun. Rev.* **43**(4), 363–374 (2013)
5. Merz, R., et al.: Performance of LTE in a high-velocity environment: a measurement study. In: *AllThingsCellular 2014*, pp. 47–52. ACM, New York (2014)
6. Johansson, I.: Congestion control for 4G and 5G access. Internet-Draft draft-johansson-cc-for-4g-5g-02. IETF Secretariat, July 2016. <http://www.ietf.org/internet-drafts/draft-johansson-cc-for-4g-5g-02.txt>
7. Alfredsson, S., et al.: Cross-layer analysis of TCP performance in a 4G system. In: *SoftCOM*, pp. 1–6, September 2007
8. Park, H.S., et al.: TCP performance issues in LTE networks. In: *ICTC 2011*, pp. 493–496 (2011)
9. Nguyen, B., et al.: Towards understanding TCP performance on LTE/EPC mobile networks. In: *Proceedings of AllThingsCellular 2014*, pp. 41–46. ACM, New York (2014)
10. Alfredsson, S., et al.: Impact of TCP congestion control on bufferbloat in cellular networks. In: *WoWMoM 2013*, June 2013 (2013)
11. Garcia, J., et al.: A measurement based study of TCP protocol efficiency in cellular networks. In: *Proceedings of WiOpt 2014*, pp. 131–136 (2014)
12. Ha, S., Rhee, I.: Taming the elephants: new TCP slow start. *Comput. Netw.* **55**(9), 2092–2110 (2011)

13. Garcia, J., et al.: Examining TCP short flow performance in cellular networks through active and passive measurements. In: *AllThingsCellular 2015*, pp. 7–12. ACM, New York (2015)
14. Chen, Y.C., et al.: Measuring cellular networks: characterizing 3G, 4G, and path diversity. In: *Annual Conference of International Technology Alliance*, pp. 1–6, December 2010
15. Wylie-Green, M.P., et al.: Throughput, capacity, handover and latency performance in a 3GPP LTE FDD field trial. In: *GLOBECOM 2010*, pp. 1–6 (2010)
16. Henderson, T., et al.: The NewReno Modification to TCP’s Fast Recovery Algorithm. RFC 6582 (Proposed Standard), April 2012
17. Ha, S., et al.: CUBIC: a new TCP-friendly high-speed TCP variant. *SIGOPS Oper. Syst. Rev.* **42**(5), 64–74 (2008)
18. WeTechs: Usage Statistics and Market Share of Operating Systems for Websites (2017)
19. Hayes, D.A., Armitage, G.: Revisiting TCP congestion control using delay gradients. In: Domingo-Pascual, J., Manzoni, P., Palazzo, S., Pont, A., Scoglio, C. (eds.) *NETWORKING 2011*. LNCS, vol. 6641, pp. 328–341. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-20798-3\\_25](https://doi.org/10.1007/978-3-642-20798-3_25)
20. Mascolo, S., et al.: TCP westwood: bandwidth estimation for enhanced transport over wireless links. In: *MobiCom 2001*, pp. 287–297. ACM, New York (2001)
21. Liu, S., et al.: TCP-Illinois: a loss and delay-based congestion control algorithm for high-speed networks. In: *valuetools 2006*. ACM, New York (2006)
22. Tazaki, H., et al.: Direct code execution: revisiting library OS architecture for reproducible network experiments. In: *CoNEXT 2013*, pp. 217–228. ACM, New York (2013)
23. Bouckaert, S., Vandenberghe, W., Jooris, B., Moerman, I., Demeester, P.: The w-iLab.t testbed. In: Magedanz, T., Gavras, A., Thanh, N.H., Chase, J.S. (eds.) *TridentCom 2010*. LNCS, vol. 46, pp. 145–154. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-17851-1\\_11](https://doi.org/10.1007/978-3-642-17851-1_11)
24. Langley, A., Riddoch, A., Wilk, A., Vicente, A., Krasic, C., Zhang, D., Yang, F., Kouranov, F., Swett, I., Iyengar, J., et al.: The QUIC transport protocol: design and internet-scale deployment. In: *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, pp. 183–196. ACM (2017)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

