

# Learning congestion over millimeter-wave channels

Luis Diez, Ramón Agüero, Alfonso Fernández  
Communications Engineering Department  
University of Cantabria  
{ldiez, ramon}@tlmat.unican.es,  
alfonso.fernandezgu@alumnos.unican.es

Yasir Zaki, Muhammad Khan  
Communication Networks Lab  
New York University Abu Dhabi  
{yz48, mk7406}@nyu.edu

**Abstract**—This paper studies how learning techniques can be used by the congestion control algorithms employed by transport protocols over 5G wireless channels, in particular millimeter waves. We show how metrics measured at the transport layer might be valuable to ascertain the congestion level. In situations characterized by a high correlation between such parameters and the actual congestion, it is observed that the performance of unsupervised learning methods is comparable to supervised learning approaches. Exploiting the ns-3 platform to perform an in-depth, realistic assessment, allows us to study the impact of various layers of the protocol stack. We also consider different scheduling policies to discriminate whether the allocation of radio resources impacts the performance of the proposed scheme.

**Index Terms**—5G, millimeter waves, machine learning, congestion control, network simulation

## I. INTRODUCTION

Millimeter wave (mmWave) is believed to be one of the key radio technologies to cope with the capacity requirements of 5G communications [1], [2]. However, due to its high frequency working range, it poses additional challenges [3]. Among others, the mmWave physical channel exhibits high oxygen absorption, diffraction and penetration losses, which lead to highly varying physical capacity. At the transport layer, this capacity variability might be perceived as congestion, so mechanisms could be triggered to reduce the traffic rate. This way, transport layer protocols, such as TCP, cannot fully harness the communication capacity.

Legacy loss-based congestion control algorithms, such as Cubic, misbehave when bottleneck buffers are large, leading to buffer-bloat situations, and performance hindering, when buffers are small [4]. In addition, these algorithms do not appropriately deal with throughput reduction over wireless links, which cause an under-utilization of the corresponding resources (wireless channel).

In the latest years, new delay-based congestion control algorithms have been proposed to overcome these limitations [5]–[7], offering a better performance over cellular networks. However, they also proved ineffective [8] in certain 5G scenarios. In the particular case of mmWave, there exist studies [9]–[11] that show some of the issues considering the interplay of congestion control algorithms with such wireless technology, and compare the performance of existing algorithms.

In this work, we present the initial steps towards an appropriate and holistic design of congestion control algorithms,

which would be able to learn congestion state over mmWave channels.

In particular, we consider one of the distinctive features of this technology. It is known that channel fading leads to long delay situations, but once it becomes reliable again, communication capacity increases very quickly. On the other hand, over the transport (wired) network, delay due to network congestion takes place at a different pace. Having this difference in mind, it would be appealing to differentiate both phenomena and adapt the congestion control algorithm accordingly. We also assume that losses are very rare in the wireless segment, due to FEC and ARQ techniques used at different levels in the protocol stack of cellular networks. Altogether, in this paper we focus on the characterization of the delay over mmWave access networks, and we analyze the potential to predict congestion (channel fading) at the transport layer, by exploiting machine learning techniques.

The congestion prediction aims to mimic the behavior of Explicit Congestion Notification (ECN) scheme [12], [13], where packets are marked by routers when buffer occupancy exceeds a certain threshold. This feedback eventually reaches the sender, which accordingly modifies the sending rate, as in [14]. Nevertheless, this function is not enabled in most routers [15]. Machine learning in congestion control solutions has been exploited in the past [16] and congestion learning has been particularly analyzed [8], [17]. In [17] congestion learnability is analyzed in general, but the authors do not consider the particular features of mmWave channels. On the other hand, in [8], learnability of congestion in mmWave networks is addressed, but the network emulation does not consider the relationship of the congestion and other protocol stack elements of the cellular access network.

All in all, the contributions of this paper are:

- We consider a realistic full stack implementation, so that the results take into account how specific network configurations may impact the congestion state.
- Exploiting the aforementioned protocol stack, we analyze the relationship of buffer length increase in the wireless access network, and we evaluate metrics that could be used to predict congestion at the transport layer.
- We showcase the influence of different network configurations over the congestion state and such metrics, considering in particular scheduling policies.

- We analyze the potential of using unsupervised learning for congestion prediction, and we compare them with supervised solutions.
- The tools and code used for this work are available in a public code repository <sup>1</sup>.

The paper is structured as follows: in Section II we present the network setup and the information obtained for the congestion learning. Afterwards, we analyze the relationship of the network state with the congestion metrics, under different scenarios and network configurations in Section III. Then, congestion learning is studied in Section IV, where supervised and unsupervised algorithms are compared. Finally, we conclude the paper in Section V, summarizing the outcomes of the research presented herewith and providing an outlook of our future work.

## II. SCENARIOS AND MEASUREMENTS

In this section we describe the scenario that we consider and the procedure that we follow to obtain the data that is afterwards analyzed. In our previous work [8], the whole end-to-end network was emulated by using the MahiMahi [18] link emulator, which provides cellular-channel traces. Although this type of setup enables the analysis of the overall system behavior, it does not allow us to isolate the contribution of the different network segments to the end-to-end behavior. On the other hand, in this work we rely on the ns-3 [19] platform and, in particular, on the mmWave extension described in [20], [21]. This way, we are able to obtain information from different layers of the protocol stack, and to better understand how network configuration may impact the potential prediction of the congestion state.

### A. Scenario setup

The evaluation scenario consists of one user and one mmWave base station, as depicted in Figure 1. We acknowledge that we would need to consider the influence of other users, since their traffic might actually impact the perceived congestion. However, considering that we only aim at studying the congestion over the wireless access segment, and that it has a very high capacity, we can safely state that congestion would be little influenced by other users. In addition, this setup eases the analysis of the different metrics used for the congestion prediction.

Although there exist different channel models [22], [23], in this work we focus on the one defined by the 3GPP [24]. In particular, we apply the urban macro channel model, where both line-of-sight and non-line-of-sight situations may happen, depending on the distance, as described in [24]. For each scenario configuration the user remains static at a certain distance from the base station, and the mmWave channel matrix is updated every millisecond. For simplicity, both UE and base station are equipped with a single antenna, and we assume the beam is perfectly oriented towards the direction of the receiver node. In addition, the scenario is configured

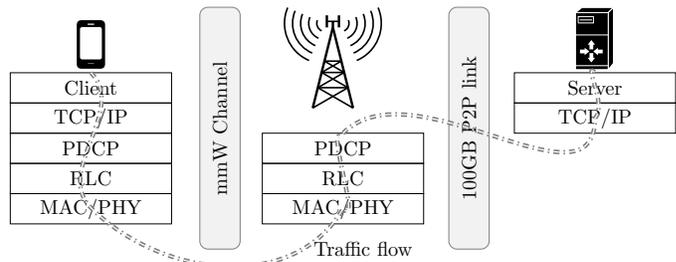


Fig. 1: Simplified scenario representation

with 4 self blocking obstacles with the use equipment staying in portrait mode. Extending this analysis to more complex scenarios, considering a larger number of users, other channel models, or more antenna elements, will be part of our future work. However, the methodology would not change, and the extension to consider such elements would be rather straightforward.

Due to the relevance of the RLC protocol in our study, it is interesting to briefly describe its purpose and possible configurations. Among other tasks, RLC protocol is responsible for segmentation and reassembly of RLC-SDUs coming from the PDCP protocol, see Figure 1. In addition, it ensures duplication detection and deletion, as well as in-order delivery to the upper layer. When configured in acknowledged mode (AM), this protocol also implements ARQ, to provide reliable link-level communications, and the connection is closed if the maximum number of re-transmissions is exceeded.

As for the MAC and link layers, we enable hybrid automatic repeat request (HARQ) in our setup, and we configure the RLC protocol in the AM mode, since it is the most common configuration. It is worth noting that, with this configuration, losses over the access segment are very unlikely, and so are neglected in our analysis. Hence, congestion control mechanisms would be mainly initiated due to delay (i.e. RTO timer expiration in TCP), provided large buffers exist in the rest of the network, thus ensuring no losses.

Finally, we use a constant bit rate traffic pattern, using UDP as transport protocol, so congestion control does not occur. The reason behind using UDP is to analyze the relationship between congestion in the access segment and the metrics that may be gathered at transport level. On the other hand, if we had used TCP the observed relationship would have been influenced by the particular congestion control algorithm, thus hiding the relationship of the analyzed metrics.

### B. Metrics

In order to analyze the relationship between congestion in the radio access segment and its impact on the transport layer, we have defined a set of metrics to be gathered from the different levels of the protocol stack. It is worth recalling that, as shown in Figure 1, we assume infinite capacity in the transport network, implemented with a 100GB/s point-to-point link.

<sup>1</sup>[https://github.com/ldziez/mmw\\_cong\\_ns3.git](https://github.com/ldziez/mmw_cong_ns3.git)

- Physical layer: although the mmWave physical channel is seen as a black-box to the upper layers, we measure its evolution to understand how its capacity variation translates to upper-layers metrics. Particularly, we monitor the amount of data sent to the physical channel as the transport-block-size actually used in each transmission opportunity. This parameter is relevant to understand how some network configurations, such as the scheduler policy, may impact the whole system behavior, including congestion-related metrics at transport layer.
- MAC layer: in this layer we specifically focus on the RLC protocol. Congestion in the access segment due to physical channel fluctuations translates into a higher buffer occupancy at the access-stratum protocol stack. Although other buffers exist at the physical layer to implement the HARQ, the main storing takes place at the RLC level. We monitor the evolution of the RLC buffer length, as an indicator of the level of congestion in the access network. It is calculated as the difference between bytes received at the protocol instance and those that were sent and acknowledged, so that it considers the actual occupancy of transmission and re-transmission buffers. In our scenario configuration, RLC buffers are considered to have an infinite capacity, so that RLC SDUs are not dropped.
- Transport layer: considering that, with the aforementioned setting, losses are rather unlikely, congestion is viewed at the transport layer by a delay variation. For that reason, we registered the packet delay as well as the inter-arrival-time (IaT). While these metrics would not be directly accessible from the congestion control algorithm, both can be easily estimated by means of the round-trip-time (RTT) and the time between ACK segments, respectively.

### III. CHANNEL-CONGESTION RELATIONSHIP

After describing the scenario setup, we analyze the relationship of the above metrics under various scenario configurations, characterized by the distance between user and base station, and so the channel quality, as well as the user traffic rate. Besides, we apply different schedulers, to analyze whether this impacts the transport layer metrics. Hereinafter, we will use correlation to refer to the Pearson's correlation coefficient, which is defined, for two random variables  $X$  and  $Y$ , as  $\rho_{XY} = \frac{cov_{XY}}{\sigma_X \sigma_Y}$ , where  $cov_{XY}$  is the corresponding covariance, and  $\sigma_X$  and  $\sigma_Y$  represent the standard deviations of  $X$  and  $Y$ , respectively.

#### A. Scheduling policy impact

First, we show the temporary evolution of the parameters described in Section II to illustrate the impact of the scheduler policy on the transport level metrics. In particular, we have used two different schedulers. The first one is a greedy scheduler that allocates radio resources according to the RLC buffer status and the selected modulation and coding scheme

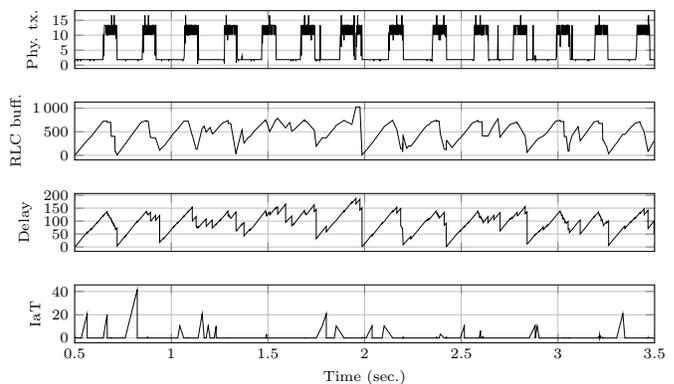


Fig. 2: System parameters with distance and traffic rate of 30 meters and 20Mbps, respectively, using the **greedy** scheduler

(MCS), while the second one is the well-known max-weight algorithm [25].

As an example, Figures 2 and 3 depict the instantaneous evolution of the different metrics during 3 seconds, in a scenario where the user is 30 meters away from the base station and the traffic rate is set to 50 Mbps. We show the bytes sent to the physical channel in kB, the RLC buffers' occupancy in kB, as well as the packet delay and IaT in milliseconds.

As can be observed in Figure 2, the traffic is sent to the physical channel in a bursty manner, leveraging those periods with good channel conditions (i.e. LoS). This yields a clear variation of the RLC buffer occupancy, which increases when the transmitted traffic is low, and gets empty periodically, concurring with longer physical transmissions. In turn, the results also evince that the delay observed at the transport layer follows a similar pattern to that of the RLC buffer. Opposed to that, the IaT exhibits a different behavior.

Then, Figure 3 shows that the max-weight scheduler renders a different behavior. Again, the traffic is sent in bursts, but with different periodicity and burst duration. As can be seen, it leads to different patterns of the RLC buffer length, as well as of the parameters measured at the transport layer. Nevertheless, in both cases the delay seems to be a good metric for the buffer occupancy.

#### B. Correlation analysis

After observing the network behavior at the different levels, we have broadened the analysis, considering different scenarios, to better understand the influence of the network configuration (i.e. scheduler) and the particular characteristics of the scenario setup. Figures 4 and 5 illustrate the linear correlation of the RLC buffer occupancy with both the delay and the IaT for the greedy and max-weight schedulers, respectively. This analysis has been carried out over scenarios with different traffic rates and distances between user and base station (as shown in the x-axis), for connections lasting 30 seconds. Traffic rates are chosen in the range [10, 1000] Mbps, and the distances are set to ensure that communication was not dropped due to timer expiration.

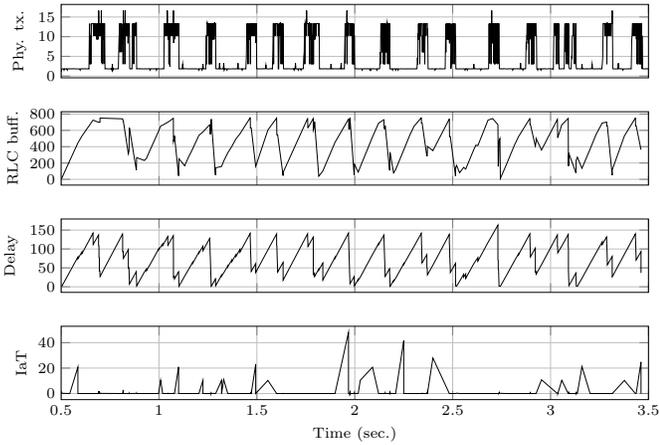


Fig. 3: System parameters with distance and traffic rate of 30 meters and 50Mbps, respectively, using **max weight** scheduler

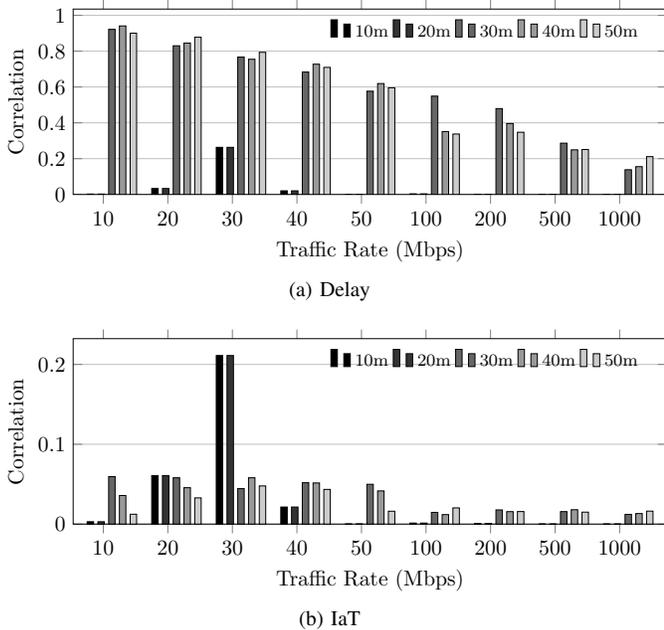


Fig. 4: Linear correlation between the RLC buffer and transport level metrics using the **greedy** scheduler

Figure 4 yields interesting results. First, we can observe that the delay correlation is much higher than the one seen for the IaT, thus making it a better metric to predict the congestion. In addition, we can see that, for small distances (10 and 20 meters), the delay correlation is almost zero, and it just slightly increases (still being rather low) for the configuration using 30 Mbps. The reason is that, under those circumstances, the good channel conditions are very good, thus ensuring LoS almost all the time. Therefore, the variations in the different parameters do not influence each other. Finally, the figure also shows that the delay correlation decreases as the traffic rate increases, while it is less dependent of the corresponding distance between user and base station.

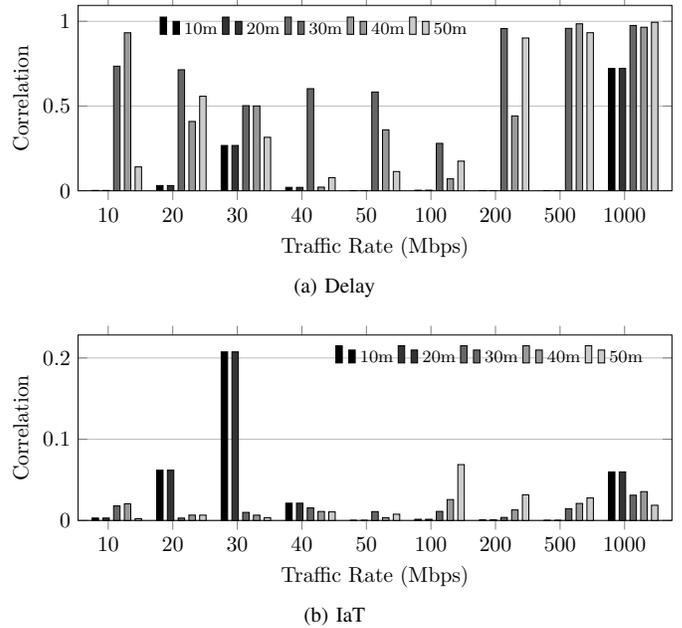


Fig. 5: Linear correlation between the RLC buffer and transport level metrics using the **max-weight** scheduler

On the other hand, Figure 5 yields slightly different trends. While the delay correlation is remarkably higher than that of the IaT, there is not a clear trend as we increase the traffic rate. Interestingly, the higher delay correlation with this configuration is seen with the largest data rate (1 Gbps).

These results indicate that, as expected, there exists a relationship between the channel quality and the delay measured at the transport layer. However, we also observe that such correlation strongly depends on the network configuration (i.e. scheduling policy), which is unknown at the transport layer. Then, it would be necessary to exploit additional metrics or other relationships to identify congestion in all scenarios.

#### IV. UNSUPERVISED CONGESTION LEARNING

After analyzing the evolution and relationship of the different parameters, we discuss in this section the potential to predict congestion events, based on the delay-related metrics. This way, it would be possible to mimic the behavior of the ECN scheme used in congestion control algorithms like DCTCP, without requiring neither the intermediate routers nor the receiver to enable this feature.

The congestion prediction is modeled as a mapping from delay and IaT quantities to 0 or 1 labels, which represent congestion and no-congestion events. This is a 2-class classification problem, for which supervised learning algorithms, such as Support Vector Machine (SVM), can be applied. However, from a practical perspective, obtaining the training information required for the supervised learning may be expensive in terms of time and memory. For that reason, it seems sensible exploring unsupervised learning algorithms, and study their potential to predict congestion. Having this in mind, in this

section we compare the prediction performance of an unsupervised learning algorithm, K-means, with that drawn by SVM, which is thus used as a benchmark. Taking into account the results obtained in the previous section, learnability analysis is performed over scenarios where correlation is not negligible, and so we ensure that prediction is indeed feasible.

Concretely, from the set of scenarios described and analyzed in Section III we focus on those with a distance of 30 meters between user and base station, and analyze the algorithms' performance as we increase the traffic rate. Due to space limitations we only use the traces obtained with the greedy scheduler. In addition, we define the congestion events by setting a threshold to the RLC buffer of 5 MB.

Before applying the algorithms, the dataset was scaled so that delay and IaT samples become comparable. Each feature was scaled to be in the  $[0, 1]$  range. Although different weights can be given to each parameter, in this work both features are given the same relative importance. For both algorithms and the different traffic rates, we have obtained the following parameters:

- Precision: it is the ratio of correctly predicted positive observations to the overall positive observations. It therefore shows the probability that a predicted congestion event corresponds to an actual congestion event; this is, how trustworthy the positive prediction is.
- Recall (sensitivity): it is defined as the ratio of correctly predicted positive observations to all actual positive observations. This metric tells us the probability of detection of a congestion event.
- Accuracy: it is the ratio of correct predictions to the overall number of observations. While this metric is valid, it can yield misleading results when the classification problem is unbalanced.

Figure 6 depicts the prediction metric of both algorithms. As can be observed when the data rate is low and medium (up to 50 Mbps), both algorithms are able to provide a good prediction. These rates correspond to correlation values between buffer occupancy and delay of  $\geq 0.6$ , as was depicted in Figure 4. It is worth noting that the high accuracy shows that the predicted congestion is an actual congestion event with high probability. Additionally, the high value of the recall metric is an indication that most of the congestion events are successfully identified by both algorithms.

On the other hand, when traffic rate is high, the prediction metrics worsen. In particular, the precision and recall deterioration with K-means indicate that congestion prediction could not be trusted with high rates. In addition, we also see that the accuracy is less impacted by the traffic rate, as a consequence of the imbalance between the number of congestion and no-congestion observations.

Furthermore, for the scenarios we consider, we have seen that supervised and unsupervised learning, SVM and K-means respectively, yield comparable performances for low and medium traffic rates. Indeed, unsupervised learning exhibits slightly higher precision, at the expense of lower recall. Finally, it is worth noting that, for both algorithms, prediction

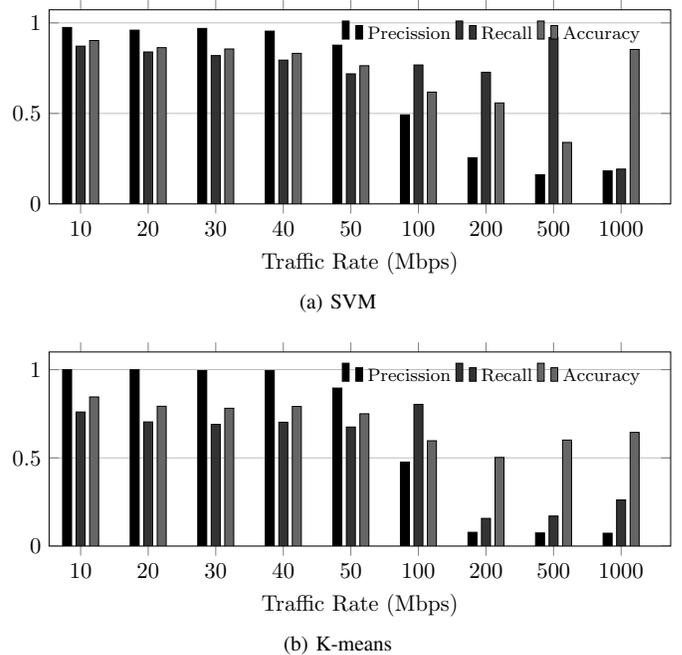


Fig. 6: Congestion prediction metrics of the SVM and K-means algorithms. Greedy scheduler is used and distance is set to 30 meters

is mostly based on the delay metrics, while the IaT provides little information to separate the samples, being a direct consequence of the low correlation between buffer occupancy and IaT, as was observed in Figure 4.

## V. CONCLUSION

In this paper we have discussed how prediction techniques, based on observable metrics, can be exploited by congestion control algorithms of transport protocols when used over 5G wireless channels. This brings interesting possibilities for the appropriate design of transport protocols, since legacy schemes behave rather poorly over prominent 5G wireless channels (for instance, mmWave). The results have shed some light regarding some of the particular metrics that should be considered and the circumstances where they can yield some benefits.

In particular, we have seen that congestion over the wireless access segment does not only depend on the physical channel characteristics, but also on how radio resources are exploited, i.e. scheduling policies. Furthermore, we ascertained that the correlation between congestion and transport protocol measurable parameters (delay and inter-arrival-time) depends on the particular scenario setup and network configuration. In this sense, we could see that some situations are characterized by having a rather strong correlation between the analyzed parameters. Hence, the proposed approach brings interesting benefits in these circumstances.

Furthermore, the analysis that was carried out shows that, under such beneficial situations, the performance of unsupervised learning techniques was comparable to that exhibited by

supervised ones. On the other hand, we would need to promote methods to improve congestion learning in circumstances where the correlation is lower.

In our future work, we plan to broaden this analysis, by increasing the complexity of the scenario: considering more users, different mobility patterns, as well as network deployments (including more complex physical layer configurations). Also, we will consider different learning techniques, both by improving the input parameters (for instance, filtering the data with moving averages), or exploiting other algorithms (e.g. Hidden Markov Models or Gaussian Mixture Models). We will also thoroughly look at congestion events, comparing the patterns caused by congestion at the wireless segment with those coming from the transport (edge-to-edge) network. This would allow us to better adapt the learning algorithms to the phenomenon that caused the congestion. Finally, we also plan to extend the classification techniques, including various wireless-based congestion situations.

#### ACKNOWLEDGMENT

This work has been funded by the Spanish Government (Ministerio de Economía y Competitividad, Fondo Europeo de Desarrollo Regional, MINECO-FEDER) by means of the project FIERCE: Future Internet Enabled Resilient smart CitiEs (RTI2018-093475-AI00).

#### REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What Will 5G Be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [2] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," *IEEE communications magazine*, vol. 55, no. 6, pp. 64–71, 2017.
- [3] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.
- [4] N. Cardwell, Y. Cheng, C. S. Gunn, S. H. Yeganeh, and V. Jacobson, "BBR: Congestion-Based Congestion Control," *Queue*, vol. 14, no. 5, pp. 50:20–50:53, Oct. 2016. [Online]. Available: <http://doi.acm.org/10.1145/3012426.3022184>
- [5] Y. Zaki, T. Pötsch, J. Chen, L. Subramanian, and C. Görg, "Adaptive congestion control for unpredictable cellular networks," in *Proceedings of the ACM SIGCOMM 2015 Conference*, London, UK, 2015, pp. 509–522.
- [6] K. Winstein, A. Sivaraman, H. Balakrishnan *et al.*, "Stochastic Forecasts Achieve High Throughput and Low Delay over Cellular Networks," in *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (NSDI 13)*, Lombard, IL, 2013.
- [7] M. Dong, T. Meng, D. Zarchy, E. Arslan, Y. Gilad, B. Godfrey, and M. Schapira, "{PCC} Vivace: Online-Learning Congestion Control," in *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*, 2018, pp. 343–356.
- [8] T. Ahmad, S. Iyer, L. Diez, Y. Zaki, R. Agüero, and L. Subramanian, "Learning Congestion State For mmWave Channels," in *Proceedings of the 3rd ACM Workshop on Millimeter Wave Networks and Sensing Systems (mmNets '19)*, Los Cabos, Mexico. ACM, 2019.
- [9] M. Zhang, M. Mezzavilla, R. Ford, S. Rangan, S. Panwar, E. Mellios, D. Kong, A. Nix, and M. Zorzi, "Transport layer performance in 5G mmWave cellular," in *2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2016, pp. 730–735.
- [10] M. Zhang, M. Polese, M. Mezzavilla, J. Zhu, S. Rangan, S. Panwar, and a. M. Zorzi, "Will TCP Work in mmWave 5G Cellular Networks?" *IEEE Communications Magazine*, vol. 57, no. 1, pp. 65–71, January 2019.
- [11] H. D. Le, C. T. Nguyen, V. V. Mai, and A. T. Pham, "On the Throughput Performance of TCP Cubic in Millimeter-Wave Cellular Networks," *IEEE Access*, vol. 7, pp. 178 618–178 630, 2019.
- [12] M. Caesar, D. Caldwell, N. Feamster, J. Rexford, A. Shaikh, and J. van der Merwe, "Design and Implementation of a Routing Control Platform," in *Proceedings of the 2Nd Conference on Symposium on Networked Systems Design & Implementation - Volume 2*, ser. NSDI'05. Berkeley, CA, USA: USENIX Association, 2005, pp. 15–28. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1251203.1251205>
- [13] D. Katabi, M. Handley, and C. Rohrs, "Congestion control for high bandwidth-delay product networks," *ACM SIGCOMM computer communication review*, vol. 32, no. 4, pp. 89–102, 2002.
- [14] M. Alizadeh, A. Greenberg, D. A. Maltz, J. Padhye, P. Patel, B. Prabhakar, S. Sengupta, and M. Sridharan, "Data Center TCP (DCTCP)," in *Proceedings of the ACM SIGCOMM 2010 Conference*, ser. SIGCOMM '10. New York, NY, USA: ACM, 2010, pp. 63–74.
- [15] B. Trammell, M. Kühlewind, D. Boppart, I. Learmonth, G. Fairhurst, and R. Scheffenegger, "Enabling internet-wide deployment of explicit congestion notification," in *International Conference on Passive and Active Network Measurement*. Springer, 2015, pp. 193–205.
- [16] P. Geurts, I. E. Khayat, and G. Leduc, "A Machine Learning Approach to Improve Congestion Control over Wireless Computer Networks," in *Proceedings of the Fourth IEEE International Conference on Data Mining*, ser. ICDM '04. Washington, DC, USA: IEEE Computer Society, 2004, pp. 383–386. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1032649.1033486>
- [17] A. Sivaraman, K. Winstein, P. Thaker, and H. Balakrishnan, "An experimental study of the learnability of congestion control," in *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4. ACM, 2014, pp. 479–490.
- [18] R. Netravali, A. Sivaraman, S. Das, A. Goyal, K. Winstein, J. Mickens, and H. Balakrishnan, "Mahimahi: Accurate Record-and-Replay for HTTP," in *2015 USENIX Annual Technical Conference (USENIX ATC 15)*. Santa Clara, CA: USENIX Association, Jul. 2015, pp. 417–429. [Online]. Available: <https://www.usenix.org/conference/atc15/technical-session/presentation/netravali>
- [19] G. F. Riley and T. R. Henderson, *The ns-3 Network Simulator*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 15–34. [Online]. Available: [https://doi.org/10.1007/978-3-642-12331-3\\_2](https://doi.org/10.1007/978-3-642-12331-3_2)
- [20] R. Ford, M. Zhang, S. Dutta, M. Mezzavilla, S. Rangan, and M. Zorzi, "A Framework for End-to-End Evaluation of 5G mmWave Cellular Networks in Ns-3," in *Proceedings of the Workshop on Ns-3*, ser. WNS3 '16. New York, NY, USA: ACM, 2016, pp. 85–92. [Online]. Available: <http://doi.acm.org/10.1145/2915371.2915380>
- [21] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, and S. Rangan, "5G MmWave Module for the Ns-3 Network Simulator," in *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, ser. MSWiM '15. New York, NY, USA: ACM, 2015, pp. 283–290. [Online]. Available: <http://doi.acm.org/10.1145/2811587.2811619>
- [22] M. K. Samimi and T. S. Rappaport, "3-D Millimeter-Wave Statistical Channel Model for 5G Wireless System Design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2207–2225, July 2016.
- [23] T. S. Rappaport, S. Sun, and M. Shafi, "Investigation and Comparison of 3GPP and NYUSIM Channel Models for 5G Wireless Communications," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Sep. 2017, pp. 1–5.
- [24] 3GPP, "Study on channel model for frequency spectrum above 6 GHz," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 38.900 V14.2.0, June 2017.
- [25] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992.