Bitrate Analysis in 5G Networks for Video Streaming Services Using L-Moment Ratio Diagrams

Juan Antonio Ortega Aparicio
Univ. Rey Juan Carlos
Fuenlabrada, Spain
jortegaparicio@gmail.com

David Cortés Polo
Univ. Rey Juan Carlos
Fuenlabrada, Spain
david.cortes.polo@urjc.es

Javier Carmona-Murillo
Univ. de Extremadura
Cáceres, Spain
jcarmur@unex.es

Mihaela I. Chidean
Univ. Rey Juan Carlos
Fuenlabrada, Spain
mihaela.chidean@urjc.es

***Abstract*—Nowadays, the expansion of wireless technologies and the development of new services have arisen new network requirements that completely change the way that network providers supply the resources to these new services. In this context, this work includes a novel analysis based on the L-moments statistical theory oriented to the understanding of the 5G networks, specially in the situation when video streaming applications are used. Our analysis involves computing and representing the L-moment ratio diagram of the bitrate of both uplink and downlink channel for three specific services, i.e. the download of a large file and Netflix and Amazon Prime video streaming. Obtained results show clear differences in the statistical behaviour for all services when uplink and downlink, mainly due the different objectives and characteristics of these links. Results also show significant differences between the Netflix and Amazon Prime video streaming services. However, interesting and surprising similarities have also been detected, specially between the Amazon Prime video streaming service and the download of a large file. Conclusions obtained in this work could be useful input to novel network management and resource allocation algorithms in the next generation networks, being this idea one of our future research lines.**

***Index Terms*—L-moments, Network Traffic, bitrate, 5G, Streaming applications**

I. Introduction

The 5G and beyond (5G and B5G) technologies are developed to address various challenges encountered in the 4th mobile generation. These networks are characterised by complex architectures with high-density deployments and multiple services with stringent requirements. Standards are designed to facilitate a 1000-fold increase in connection density, 3-fold improvement in spectrum efficiency, 10-fold enhancement in data rate, and 10-fold boost in energy efficiency, among others [1]. The achievement of these network capabilities is possible through the integration of advanced technologies such as virtualization, coexistence of traditional radio networks and novel frequency bands with massive MIMO, network slicing, and network software-defined architecture. These paradigms are centred on the capability of the network to facilitate data sharing and information exchange among network entities while also being able to adapt to the specific requirements of each service or user in order to enhance the user experience in terms of network speed, latency, reliability, and the deployment of new services such as autonomous vehicles, the tactile internet, and remote surgery [2].

In this context, the rapid increase in the use of multimedia streaming services like YouTube, Netflix and Mobile TV on smart devices has created new usage models which must be managed by telecommunications companies and service providers. The implementation of 5G and beyond technologies could greatly improve the performance of these services, which demand efficient handling of a high volume of connections with fast data transfer speeds and minimal delay. Moreover, in this challenging context, the usage of Analytical Techniques and Artificial Intelligence for analysing network information has the potential to enhance network performance, the Quality of Service (QoS) offered by the network, and the Quality of Experience (QoE) experienced by users by studying their behaviour [3].

The monitoring of wireless networks presents several difficulties in comparison to the monitoring of wired networks. The conventional techniques utilised for wired networks, such as evaluating parameters once the data has been transmitted, fail to provide an accurate representation of the current status of the wireless network [4]. The analysis process can become even more intricate when the mobile nodes moves from one cell to other, which has a direct impact on the utilisation of network resources. These circumstances complicate the data analysis and network resources administration, two of the most relevant tasks for the network providers.

This work introduces a new analytical methodology based on L-moments [5] to analyse the usage of a 5G wireless network when the mobile node uses video streaming services and to categorise these services based on their attributes. L-moments enable the use of higher-order statistical moments, bypassing limitations related to the necessary amount of data for the estimation process. This advantage, in combination

This work has been partially supported by the European Union NextGenerationEU/PRTR, grant TED2021-131699B-I00 (AEI/FEDER,UE), by the Spanish Ministry of Science and Innovation, grants PID2020-112545RB-C54 and PDC2022-133900-I00 and, by the Univ. Rey Juan Carlos Program for Research Promotion and Development (Ref. F920 and "AYUDA PUENTE 2022, URJC" Ref. F931)

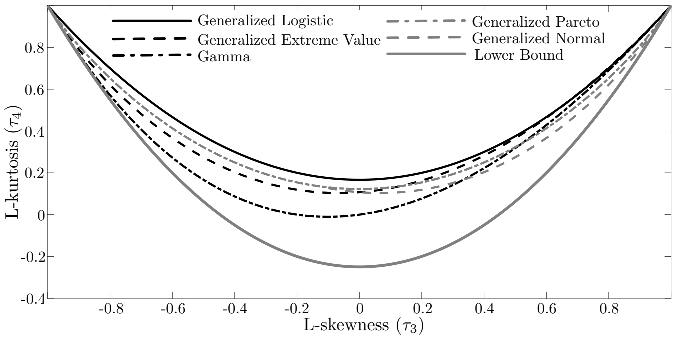

Fig. 1: LmomRD for some common distributions [5].

with the low computational resources required, could even real-time processing of data [6]. In this work, being one of the first approaches to this kind of specific 5G data with this methodology, the study is mainly related to determine similarities and even differences between the video streaming services and discuss possible network and channel causes, at all time in the resource allocation and management context.

The remainder of this paper is organised as follows. Section II describes the methodology used and the techniques applied in this work. The results of the experiments conducted applying the described methodology are introduced in Section III. Section IV discusses some results obtained after the application of the proposed methodology. Finally, Section V contains our concluding remarks.

II. Methodology

This section describes the statistical tools, specifically the L-moments statistical theory and the L-moment ratio diagram (LmomRD), as well as the dataset used in this work.

A. L-moments and LmomRD

One of the common tools used to characterise data and to summarise samples from a given data base are the statistical moments. "Classical" or "conventional" moments, also known in the literature as product moments, are just one of the available alternatives. The L-moment theory [5] is one of the possible and interesting alternatives, mainly due to the direct result interpretation, the unbiased estimators, the outlier robustness and the low sampling variability, all this even for low sample size situations [5], [7]–[10]. Also, L-moments are known to be suitable for data with large skew, large or long tails or outliers [10], [11], characteristics fulfilled by multiple variables measured for a network flows [6], [12].

L-moments are therefore a powerful tool in distributional analysis, specially for data with large range, variation, skewness and outliers. By definition, L-moments are linear combinations of the expected values of order statistics [5]. And one of the most relevant properties for the sake of this work is the fact that L-moments are analogous to product moments in terms of interpretation. That is: the first L-moment *L-location* (λ_1) is the mean of the distribution or average value of the dataset; the second L-moment *L-scale* (λ_2) gives insight into the scale of dispersion; the third L-moment (λ_3) is related to the asymmetry; the fourth L-moment (λ_4) gives insight about the tails of a given distribution, and so on.

This theory also considers ratio moments that are defined as $\tau_r = \lambda_r/\lambda_2$, for $r \geq 3$. *L-skewness* (τ_3) and *L-kurtosis* (τ_4) are specially useful as they give comparable information to the conventional skewness and kurtosis. In addition, L-skewness and L-kurtosis are both lower and upper-bounded by definition, allowing therefore the direct comparison between distributions or data samples, regardless their range. Readers are referred to the primary references like [5] and subsequent works for further details regarding the L-moments theory, such as their formal definition as well as their basic properties, estimators and proofs.

The LmomRD is a extremely useful graphical tool for exploratory analysis as well as for distribution selection tasks [5], [11]. LmomRD plot tuples (usually pairs) of L-moment ratios, being the most common considered pair the $\{\tau_3, \tau_4\}$ one. It is usual that LmomRD also include the theoretical L-moment ratios as points, lines or regions for typical distributions, like the ones included in Fig. 1.

In this work, we calculate sample L-moments and L-moment ratios for a given feature from the considered dataset for different scenarios in order to understand and interpret the similarities and differences. And LmomRD is used for visual result presentation, interpretation and comparison, including in all figures as a visual guideline the theoretical lines from Fig. 1 using also the same legend.

B. Dataset

The dataset considered in this work is a 5G dataset collected from a major Irish mobile operator [13]. This dataset includes multiple features covering from throughput to channel and context information for 5G networks for two mobility patterns and applications. The two mobility patterns are static and driving a car, including therefore two of the most common mobility situations for 5G network users. Also, the applications considered for the dataset are the download of a large file (>200MB) and Netflix and Amazon Prime video streaming services, all three being standard and common 5G applications with high quality and network resource requirements. Readers are referred to the primary references like [13] and subsequent works for further details regarding the dataset generation and specific characteristics.

As stated in [13], data were collected using the version 18.7 of the G-NetTrack Pro application installed on a Samsung S10 5G Android device. The compete dataset includes 83 experiments corresponding to more than 3100 minutes. For each experiment, i.e. covering all possibilities of mobility pattern and application combinations, a total of 80GB of data was consumed. This value is determined by the mobile plan offered by the operator when the experiments were performed. The complete dataset includes not only information about the packets sent to the network, but also the information exchanged by the mobile node which it is connected to a base station of the network operator using the standard configuration (physical and MAC configuration) provided by the network provider like channel-related metrics, context-related metrics, cell-related metrics and throughput information (uplink and downlink).

The metrics considered in this work are downlink and uplink bitrates measured at the application layer, dataset features `DL_bitrate` and `UL_bitrate`, respectively. The selection of these metrics is based on their high impact in resource allocation procedures, as well as on the fact that they directly determine the quality of a service as perceived by users. Therefore, considering the two mobility pattern, the three applications and the two metric possible combinations, in this work a total of 12 scenarios are analysed. The minimum sample size for these scenarios is 11.5×10^3 samples, a sufficient data amount to properly perform the L-moments estimation.

III. RESULTS

This section includes the detailed description of the obtained results, while the result discussion and interpretation is included in the next section.

Figure 2 includes the LmomRD obtained for the 12 considered scenarios, dividing them in four different categories to facilitate the visualisation of the figures and result comparison. Specifically, the criteria used to criteria used to categorise are the mobility pattern (static vs. driving a car, one for each sub-figure column) and the link (uplink vs. downlink, one for each sub-figure row). Among the sub-figures, a different marker is used for each application: blue triangle for the download of a large file; red square of the Netflix video streaming service and; orange circle for the Amazon Prime video streaming service. The different lines from these sub-figures are included out of habit in state-of-the-art literature mainly because they proved to be convenient for result comparison; these lines follow the legend of Fig. 1.

The LmomRD represents the τ_3 and τ_4, i.e. the L-skewness and L-kurtosis of the data, for each scenario. These figures are obtained using $n = 500$ non-overlapping and temporarily ordered data samples for the estimation of each L-moment. The considered n fulfils the minimum data samples to guarantee a proper L-moment estimation with low estimation error [5], [7]. At the same time, with this n, multiple L-moment estimations can be shown for each scenario, allowing a proper observation of the statistical behaviour for each case [12].

Regarding the L-skewness, all 12 scenarios reveal positive skewness, i.e. both mode and median are lower than the average and the tail of the distribution is located at the right. However, depending on the scenario, specific values give more insight regarding the specific statistical behaviour. For example, the Netflix video streaming service consistently has $\tau_3 > 0.65$ for the four cases, revealing highly positive skewed data, meaning that the measured `bitrate` is most of the time on the lower part or the range. On the other hand, both the download of a large file and the Amazon Prime video streaming service show differences between the four cases, revealing different statistic behaviour for the bitrate feature. The discussion section includes several possible motivations and interpretations of these differences, as well as potential usages of this differences from the network management point of view.

On the other hand, the L-kurtosis is positive for 11 scenarios, achieving some negative values for the Amazon Prime video streaming service for the downlink in the static scenario. Nevertheless, rather than the exact values it is more useful to focus on the range of estimated τ_4 as it reveals the statistical behaviour and the stationary characteristic of the bitrate. Therefore, scenarios with an estimated τ_4 close to 0 reveal that the tails of the bitrate distribution are leaner, i.e. there is a lower probability for outliers and the data is more concentrated around the mean of the distribution. And scenarios with heavy-tail distributions, i.e. high values for the estimated τ_4, show higher outlier probabilities. It can be observed that for Netflix video streaming service the L-kurtosis is bounded by $0.4 < \tau_4 < 0.75$ for the four cases, showing consistency in the statistical bitrate behaviour for both downlink and uplink for the two mobility patterns.

L-moment ratio diagram are quite useful also when the "cloud" of points is also observed. Cases when the points are less scattered, like the ones obtained for the download of large file scenario for example, reveal a higher stationarity as the estimated τ_3 and τ_4 are more similar regardless of the $n = 500$ data samples used for the estimation. Moreover, as the data is ordered on a temporary basis, this little scattering also shows that the statistics of the considered feature are similar among all the considered temporal range. On the other hand, cases when the points are more scattered, like the ones obtained for the Netflix video streaming service for example, reveal less continuity in the statistical behaviour among the performed experiments, at least in one of the considered L-moments.

IV. DISCUSSION

First of all, by using this analysis based on the L-moment theory, the analysis is independent on the bitrate range achieved for each scenario, allowing an in depth insight of the statistical behaviour of this network feature. It has already been shown that the download of a large file requires significant higher bandwidth than the video streaming services [13]. And by abstracting from the specific bitrate values we are able to go beyond with the result interpretation, understand the network behaviour and be able to propose more efficient and customised network management techniques.

Starting with the download of a large file scenario, it is clear that the LmomRD reveals significant statistical behaviour between the uplink and downlink. These differences are expected, as both links are different in objective, characteristics and resource utilisation: the uplink is mainly used for network access and control, while the downlink is mainly used for the actual data transmission. From Figure 2 we can identify that the uplink bitrate can be modelled, with little inaccuracies, using the Gumbel ($\tau_3 = 0.1699$; $\tau_4 = 0.1504$) or Normal ($\tau_3 = 0$; $\tau_4 = 0.1226$) distributions for the static and with the Exponential ($\tau_3 = 1/3$; $\tau_4 = 1/6$) distribution for driving a car mobility patterns, respectively [5]. Besides validating

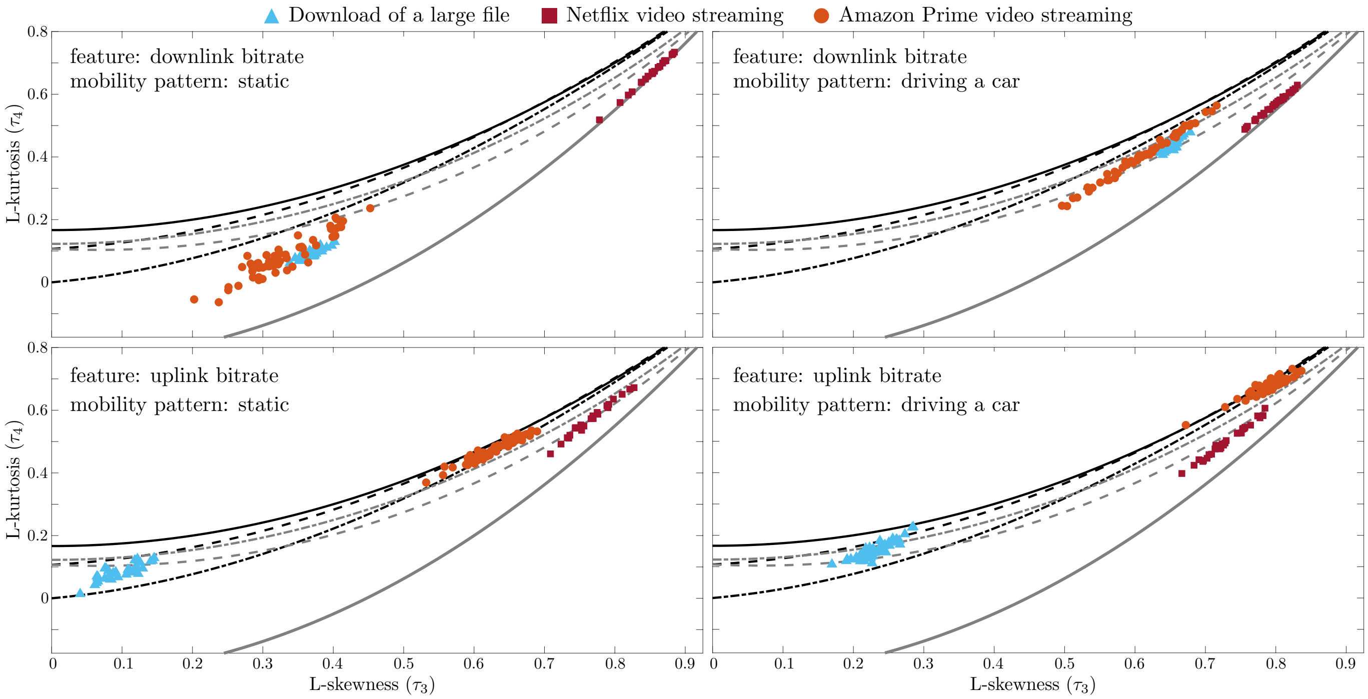

Fig. 2: LmomRD. Downlink and uplink bitrate features are represented in the top and bottom row, respectively. The static and driving a car mobility patterns are represented in the left and right column, respectively. The download of a large file and Netflix and Amazon Prime video streaming services are represented using triangles, squares and circles, respectively.

this analysis method for this scenario, this information could be highly useful for the modelling of the network access and resource management algorithms for similar applications.

Next, regarding the driving a car mobility pattern for the download of a large file scenario, both τ_3 and τ_4 increase, although in the downlink case it is more apparent, revealing a more challenging wireless channel. This difference between the static and driving a car mobility patterns is also apparent for the video streaming applications, specially for Amazon Prime. There are clear differences between these mobility patterns from channel propagation points of view [14], and higher levels of fading, multi-path or Doppler effect are expected at higher relative velocity between the mobile device and the base station. And, as it can be seen, the LmomRD are able to capture these different scenarios, therefore it can be a useful tool for the network resource management algorithms for any of the considered or similar services. Moreover, the few computational resources that are required for the method based in the L-moment theory used in this work is a further advantage for any network management procedure [6].

Results for the Netflix video streaming service also leads to interesting conclusions, as it can be observed that the obtained LmomRD for all four cases are quite similar, with even higher similarities comparing results from the same link. Following the above reasoning, this results reveal highly efficient resource management in terms of consistency with a proper channel adaptation for this service, with independence on the channel quality.

One of the most interesting conclusions obtained from the results is the fact that, contrary to what one might think at first, Amazon Prime video streaming service has more statistical similarities for the downlink scenarios with the download of large file service than with the Netflix video streaming service. As analysed in [13], Netflix requires a significantly higher bandwidth than Amazon prime video for both mobility patterns because of the Open Connect mechanism implemented by Netflix. This mechanism avoids the service single point of failures and increases the QoS of the multimedia transmission by installing Content Delivery Networks in the Internet Services Providers to improve the QoE of the view. With this mechanism the clients can receive the multimedia content with higher throughput compared to other services located in centralised datacenters. Furthermore, the streaming clients usually implements a buffering mechanism which must be filled prior to playback. The buffer mechanism employed by the Netflix client is designed to store a substantial amount of information prior to playback, in contrast to the continuous buffer filling technique utilised by Amazon Prime Video. These differing buffering techniques require distinct network management, as each service has its own requirements.

Finally, the analysis presented in this work enables discrimination between the network usage for both streaming services and enables effective network management (both Core and Radio) to meet the analysed requirements. Netflix requires a burst channel to fill the buffer and play the content, whereas Amazon Prime Video requires sustained network resources throughout the multimedia session. These results are consistent with previous works like [15] where strategy and implementation differences between streaming services are pointed out and analysed.

V. Conclusions

The advancement of next-generation networks, such as 5G and B5G, is facilitated by the utilisation of information extracted from existing networks. This information allows for the deployment of virtual functions, orchestration of physical and virtual Radio Access Networks, modelling of spectrum usage, and creation of flexible services within the network. Within this context, in this work we analysed the bitrate for the uplink and downlink in an actual 5G network with methods from the L-moments statistical theory. A total of three different services and two mobility patterns were considered, i.e. the download of a large file and Netflix and Amazon Prime video streaming services for the static and driving a car settings. Results reveal evident statistical differences between the uplink and downlink scenarios, clearly justified by the inherent attributes of each link. Moreover, results show high statistical similarities for the Netflix video streaming service with independence of the link and the mobility pattern, revealing efficient resource allocation for the wide variety of channel states expected for the considered scenarios. On the other hand, differences between the two video streaming services considered were observed, while the Amazon Prime service has similarities with the download of a large file service. These results and conclusions could lead to the optimisation of network management and resource allocation algorithms, that take into account the specific network requirements for each considered service.

Results and conclusions obtained from this work lead to multiple future works. For example, the insight obtained regarding the similarities and differences between the considered scenarios should allow the work towards the proposal of novel and more efficient network management and resource allocation algorithms, even considering some level of service customisation. For this purpose, we consider that also other network flow and network state parameters should be considered in the analysis. We also consider that this analysis could be extended to additional 5G services with different requirements and specific conclusions could also contribute to the improvement of the aforementioned network management and resource allocation tasks. In addition, in order to approach a similar analysis from the multimedia data and streaming services point of view a even more detailed database would be required, including details such as the specific video resolution and quality.

References

[1] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, “6g wireless networks: Vision, requirements, architecture, and key technologies,” *IEEE Vehicular Technology Magazine*, vol. 14, pp. 28–41, 9 2019.

[2] F. Paolucci, F. Cugini, P. Castoldi, and T. Osinski, “Enhancing 5g sdn/nfv edge with p4 data plane programmability,” *IEEE Network*, vol. 35, pp. 154–160, 5 2021.

[3] D. Naboulsi, M. Fiore, S. Ribot, and R. Stanica, “Large-scale mobile traffic analysis: A survey,” *IEEE Communications Surveys and Tutorials*, vol. 18, pp. 124–161, 1 2016.

[4] D. S. Medeiros, H. N. Cunha Neto, M. A. Lopez, L. C. S Magalhães, N. C. Fernandes, A. B. Vieira, E. F. Silva, and D. M. F Mattos, “A survey on data analysis on large-scale wireless networks: online stream processing, trends, and challenges,” *Journal of Internet Services and Applications*, vol. 11, no. 1, pp. 1–48, 2020.

[5] J. R. Hosking, “L-moments: Analysis and estimation of distributions using linear combinations of order statistics,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 52, no. 1, pp. 105–124, 1990.

[6] J. Galeano-Brajones, M. I. Chidean, F. Luna, and J. Carmona-Murillo, “A novel approach for flow analysis in software-based networks using l-moments theory,” *Computer Communications*, vol. 201, pp. 116–122, 2023.

[7] J. R. Stedinger, “Frequency analysis of extreme events.” *in Handbook of Hydrology.*, 1993.

[8] H. A. David and H. N. Nagaraja, *Order statistics*. John Wiley & Sons, 2004.

[9] M. Jones, “On some expressions for variance, covariance, skewness and l-moments,” *Journal of Statistical Planning and Inference*, vol. 126, no. 1, pp. 97–106, 2004.

[10] W. H. Asquith, “Univariate Distributional Analysis with L-moment Statistics using R,” Ph.D. dissertation, Texas Tech University, 2011.

[11] R. M. Vogel and N. M. Fennessey, “L moment diagrams should replace product moment diagrams,” *Water resources research*, vol. 29, no. 6, pp. 1745–1752, 1993.

[12] M. I. Chidean, J. Carmona-Murillo, R. H. Jacobsen, and Q. Zhang, “Network Traffic Characterization Using L-moment Ratio Diagrams,” in *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*. IEEE, 2019, pp. 555–560.

[13] D. Raca, D. Leahy, C. J. Sreenan, and J. J. Quinlan, “Beyond throughput, the next generation: a 5g dataset with channel and context metrics,” in *Proceedings of the 11th ACM multimedia systems conference*, 2020, pp. 303–308.

[14] A. F. Molisch, *Wireless communications*. John Wiley & Sons, 2012.

[15] A. Rao, A. Legout, Y.-s. Lim, D. Towsley, C. Barakat, and W. Dabbous, “Network characteristics of video streaming traffic,” in *Proceedings of the seventh conference on emerging networking experiments and technologies*, 2011, pp. 1–12.