Design of AI-based Digital Twin Network for Multimedia Service Provisioning

Enrico Boffetti, Arcangela Rago, Giuseppe Piro, Gennaro Boggia
Dept. of Electrical and Information Engineering, Politecnico di Bari, Italy
CNIT, Consorzio Nazionale Interuniversitario per le Telecomunicazioni, Italy
Email: e.boffetti@phd.poliba.it, arcangela.rago@poliba.it, giuseppe.piro@poliba.it, gennaro.boggia@poliba.it

***Abstract*—In the era of pervasive connectivity, the effective provisioning of multimedia services surely represents a cornerstone challenge for both infrastructure and service providers. Given the recent evolution of Beyond 5G/6G networks, it is conventionally accepted that a Network Operator (infrastructure provider) shares its resources with tenants (service providers) by giving them the possibility to autonomously and efficiently configure the application services they provide to end users. To this end, this contribution proposes a cutting-edge framework for multimedia service provisioning, that harnesses the power of Artificial Intelligence (AI) in conjunction with Digital Twin Networks (DTNs). Specifically, it considers the real network features collected from the physical real network through Software-Defined Networking Controllers to build the DTN, whose purpose is to provide present and future physical, network, and application statistics through AI for efficient and proactive resource allocation. The DTN exposes to the tenant a concise set of monitoring parameters related to actual network-specific statistics and Quality of Service metrics and also future ones thanks to Deep Learning, without sharing in-depth details of the underlying network through Machine Learning (clustering). This information will be exploited by the service orchestrator for optimal resource allocation and redistribution, which will be deeply investigated in future works.**

***Index Terms*—6G, DTN, AI, QoE, QoS, Clustering**

I. Introduction

The new generation of mobile communications introduces the Digital Twin Network (DTN) as a 6^{th} Generation (6G)-oriented service, which can take advantage of the integration of Artificial Intelligence (AI) [1]. Specifically, the integration of DTNs and AI offers a very powerful approach for modeling, simulating, and optimizing Beyond 5^{th} Generation (5G) and 6G communication systems also in the context of multimedia services, where Quality of Experience (QoE) prediction and effective planning and usage of network resources are needed because of the explosive growth of multimedia traffic [2].

The current scientific literature generally performs Quality of Service (QoS) optimization and QoE estimation. To this end, various scheduling approaches for multimedia services are proposed. They generally adopt QoS metrics to estimate QoE through Mean Opinion Score (MOS) values [3], [4] or they collect MOS values or subjective evaluations directly from users [2], [5]. Moreover, also AI can be adopted for these purposes. To obtain the same MOS value for each user, the transmission power and channels of Access Points (APs) can be dynamically changed for each user thanks to Deep Reinforcement Learning (DRL) [3]. Other works such as [6], [7] calculate Inter-Mean Opinion Score (iMOS) that is the MOS values at the intermediate network nodes, e.g., 5G Voice over Long Term Evolution and Voice over New Radio or in Software-Defined Networking (SDN) with Voice over Internet Protocol flows.

Differently from the current state of the art, this contribution proposes an innovative framework that integrates DTN, the SDN paradigm, and AI for multimedia service provisioning. The adoption of the DTN is of utmost importance in the proposed framework because it represents a revolutionary approach to managing networks. Its significance lies in the capacity to replicate diverse network configurations in real-time, providing a virtual mirror image of the physical network (i.e., Network Operator Infrastructure) [8], [9]. Thanks to SDN facilities and network slicing, network data and statistics are collected for each slice dedicated to users. This information is shared with the DTN, which knows network dynamics, including QoS metrics, of the real physical network [1], [10], [11]. Thus, the DTN simulates and analyzes various setups and pinpoints potential vulnerabilities without impacting the real physical network. The AI integration can boost and enhance the DTN capability for troubleshooting, experimenting with new resource configurations, forecasting the consequences of changes before the actual implementation, and enhancing the overall efficiency of network operations and the network resilience through proactive measures [8], [12]. Specifically, Deep Learning (DL) and Machine Learning (ML) are employed for QoE/QoS prediction and clustering, respectively. The tenant can, then, exploit the outcomes of QoE/QoS prediction based on Long Short-Term Memory (LSTM) and user clustering through the service orchestrator. It enables the redistribution of physical, network, and application resources thanks to AI, and in particular by implementing a DRL approach, in order to guarantee high levels of QoE for all the network users. It is important to note that at the time of writing, and to the best of our knowledge, a first attempt in this direction is presented in [13], where QoS prediction and spatial clustering algorithms have been designed and evaluated. However, it refers to Vehicular-to-Everything services and it does not address QoE estimation and QoE/QoS prediction based on LSTM, which is the state-of-the-art model for online predictions, as stated also there. In addition, it does not sketch a complete framework with new 6G technologies (DTN) and does not deal with the optimal resource allocation and redistribution. To summarize, this paper significantly advances the current state of the art because: i) it designs the overall framework for multimedia service provisioning within the Internet Engineering Task Force (IETF) model for DTN [8]; ii) it integrates DTN and AI as envisioned by the upcoming 6G network architecture; iii) it proposes an innovative AI-based methodology for QoE/QoS prediction and users' clustering, by performing an objective and standardized estimation of QoE and extracting the information of all the users through the simplified and privacy-preserving (without in-depth details of the network) management of clusters for the anticipatory and optimal resource allocation and redistribution; iv) it provides a preliminary discussion on the usage of prediction and clustering outcomes in a realistic scenario.

The remainder of the paper is as follows. Section II describes the proposal and provides some technical details on the adopted AI approaches. Section III presents the preliminary investigation and the early results. Section IV draws future research activities and highlights related optimization opportunities and, finally, Section V concludes the paper.

II. Proposal

The cutting-edge framework presented herein aims to provide multimedia services by exploiting AI-based DTN. Specifically, the tenant, i.e., the service provider, provides multimedia services through the physical network of the Network Operator, i.e., the infrastructure provider. As shown in Fig. 1, the Network Operator Infrastructure, which includes the various 5G/6G network elements (e.g., mobile users, base stations, routers, switches, computers, servers, satellites) manages and accepts resource requests issued by tenants [1]. Through the Southbound Interface of SDN Controllers, the Network Operator Infrastructure is connected to them, which collects information and enables network control (e.g., OpenFlow, OpFlex, NetConf) [10], [11]. Thus, through the Northbound Interfaces of SDN Controllers, the instance of the DTN knows the network infrastructure state and what resources are available. At this point, through its Southbound Interface, the DTN requests information on the Network Operator Infrastructure and stores it in its Data Repository. The DTN generates network models and, through the Functional Models, can calculate key network parameters. In the first phase, the QoE Estimator module receives from the Data Repository the QoS metrics of the network and estimates the QoE of the users (i.e., MOS). Then, QoS parameters and MOS values are given to the QoE/QoS Predictor, which is a module that contains the DL Model responsible for predicting users' future QoS and QoE values (i.e., QoS metrics and QoE values). The Clustering module through a ML model gets the clusters of users for the instant of execution t and subsequent ones (i.e., $t+1$, $t+2$, and so on) because the tenant can use this information for any requests of resource reallocation to increase user satisfaction. Note that the number of clusters can change over time. For this reason, in Fig. 1 the DTN shares with the tenant predicted information on different numbers of clusters (i.e., N, M, P). The tenant receives this information and, thanks to the service orchestrator and the help of AI, computes the bandwidth needed for optimal resource allocation and the increase of QoE, avoiding over-provisioning and, if necessary, by sending to the DTN the requests for more bandwidth and resources. These requests arrive at the Configuration Planner, which creates various possible configurations to test via the Network Emulator in a parallel manner. The Network Emulator, after testing all possible configurations, chooses the best one and sends it to the Configuration Deployer, whose function is to send the configuration to the interested SDN Controllers.

In this way, the DTN aims to proactively optimize network resources thanks to AI. It performs QoE estimation through the perfect knowledge of monitoring parameters and QoS metrics and consequently performs QoE/QoS prediction and user clustering thanks to DL and ML. Then, only the processed network statistics are exposed to the tenant, without sharing all the in-depth details of the underlying network, by safeguarding the privacy of the Network Operator [14]. Finally, the tenant has a service orchestrator, which is able to ask for the redistribution of physical, network, and application resources to guarantee high levels of QoE (i.e., MOS values) for all the network users. To this end, it is supported by AI and specifically it can implement a DRL algorithm. The main functionalities covered by the proposal, with a focus on the DTN, are introduced below. Only a high-level description is provided in this position paper, specifically for the resource allocation and redistribution, and the complete design and analysis are delayed for future research activities.

A. Physical network monitoring

The Network Operator Infrastructure manages resource requests issued by tenants and interacts with the SDN Controller, which implements monitoring functionalities and retrieves network dynamics, including QoS metrics. Network data and statistics are collected for each slice dedicated to users and are exposed to the DTN through Northbound Interfaces (e.g., by using RESTful API [10]). This information is transmitted from the SDN Controllers to the DTN [1]. The interaction between the SDN Controller and the Network Operator Infrastructure/DTN is implemented through conventional protocols (i.e., OpenFlow, RestConf, etc.) and interfaces [10], [11], [14].

B. QoE/QoS estimation, prediction, and clustering

The Network Operator can generate instances of the DTN, making precise copies either of specific network segments or related to relevant features for the tenants. In the proposed framework, the Network Operator utilizes the DTN to replicate and instantiate specific features. Going into detail, the usage of physical resources and QoS metrics like bandwidth, latency, and Packet Loss Rate (PLR) are collected and exposed to the DTN for QoE estimation. In particular, MOS can be adopted [3]–[5]. It is a standard subjective metric that measures the level of user satisfaction on a scale of 1 (bad) to 5 (excellent). Through the International Telecommunication Union (ITU) standardized E-model [15], the DTN estimates the MOS value of users.

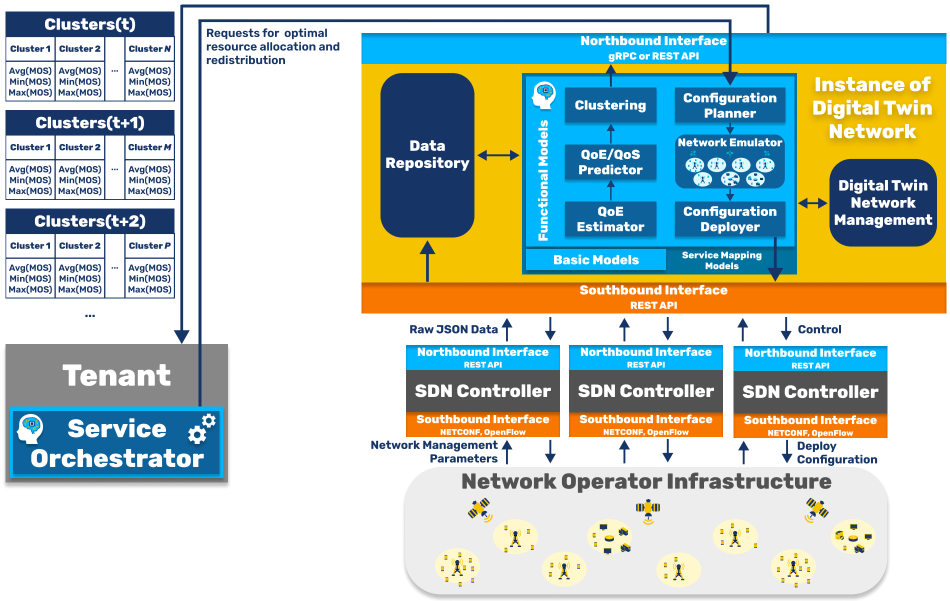


Fig. 1. General architecture with the proposed AI-based DTN.

Below there are the details on the focus of the proposal, i.e., DTN and AI.

1) DTN: Standard Development Organizations, such as IETF or ITU, have initiated efforts to formulate a definition for a DTN [8]. IETF also shared a reference architecture of DTN that is divided into physical network, instance of DTN, and application. According to the reference architecture, the instance of DTN collects network data and control messages from the real physical network (i.e., raw JSON data) and stores them in the Data Repository. Thus, the collected information is temporally stored in the Data Repository (e.g., in JSON files) so that it is available to the Functional and Basic Models.

2) AI: Functional Models are used for network analysis, emulation, prediction, etc., while Basic Models are the models referring to the network element models, topology, and all other information to fully characterize in real-time the real physical network [8]. By focusing the attention on the Functional Models, the integration of AI in the DTN is crucial: the QoS metrics are adopted not only to calculate the MOS values for each user (i.e., QoE Estimator module) but also for QoE/QoS prediction and user clustering. The QoE/QoS Predictor module implements a LSTM network to predict the next QoS metrics and MOS values for each user by analyzing them in the previous time instants. This kind of DL solution can extract temporal correlations of data through LSTM memory cells [13], [16]. Then, the Clustering module thanks to ML groups the users based on QoS metrics and allows reducing the size of exchanged data with the tenant.

C. Towards an optimal multimedia service provisioning

The DTN exposes to the tenant, through its interfaces [11], bandwidth and the average, the minimum, and the maximum MOS values for each cluster and shares their predicted future values. Thus, thanks to AI, the DTN can provide in advance high-level indicators on QoE for all the users. They enable simplified network management, without sharing QoS metrics by reducing the amount of exchanged data and energy consumption and masking in-depth details of the Network Operator. By knowing the future predicted values of MOS, i.e., future QoE, the tenant through AI, and specifically DRL, can decide if more resources from the Network Operator are needed to guarantee high levels of QoE for all the users, minimizing the bandwidth to be allocated in order to avoid resource over-provisioning and deliver a reliable streaming service meeting Service Level Agreement (SLA).

III. Preliminary Investigation

The preliminary results discussed in this paper refer to some AI functionalities presented in Section II-B. The Web Real-Time Communications (WebRTC) use case is considered as an example to preliminarily test the proposed framework and network statistics, including QoS metrics. Realistic tests are conducted in the time domain.

A. Dataset

This paper considers the dataset presented in [17], which reports various tests conducted in various countries around the world using WebRTC technology in the context of the Measuring Mobile Broadband Networks in Europe (MONROE) project. The analyzed scenario includes a group of 150 nodes, consisting of both mobile devices (e.g., within delivery trucks, trains, or buses) and stationary ones (e.g., volunteers who host nodes in their residences). Among the captured features, latency and PLR per user are used. Specifically, the PLR is obtained through the received packets and lost packets per user. In the proposed framework (Fig. 1), the SDN Controller exposes these QoS metrics to the instance of DTN, which calculates the MOS values (i.e., QoE estimation).

TABLE I
CONFIGURATION PARAMETERS AND PERFORMANCE OF THE LSTM NETWORKS.

hidden size	*number of layers*	number of trainable parameters [#]	MSE_{train} $[\cdot 10^{-3}]$	MSE_{val} $[\cdot 10^{-3}]$	MAE_{train} $[\cdot 10^{-3}]$	MAE_{val} $[\cdot 10^{-3}]$
5	1	166	1	2.8	16.2	17.5
40	2	20041	0.6	2.8	17.7	19.7
100	2	122101	0.4	2.7	10.7	16.9
150	2	273151	0.5	2.7	14.3	16.8
200	**2**	**484201**	**0.4**	**2.7**	**10.2**	**13.5**
200	1	162601	0.5	2.7	15.5	17.2
200	3	805801	0.4	2.7	15.8	18.5
300	2	1086301	0.4	2.8	15.8	15.9

B. Evaluation setup for QoE prediction and clustering

The adopted AI algorithms for QoE prediction and clustering, i.e., DL and ML algorithms, have been implemented in Python, as detailed below.

1) QoE prediction: The prediction architecture is implemented using *torch.nn*, a Python-based API for neural networks built on PyTorch [18]. It utilizes an observation window T of 20s. Adam optimization with a learning rate of 0.001 is employed for weight updates, with 100 *epochs* and a *batch size* of 64. Training hyperparameters include an *input size* of 1 (representing the MOS value per user), various *hidden sizes* (number of features/units in the hidden state), and different numbers of LSTM layers. The *output size* is set to 1 for predicting the next MOS value. Preliminary evaluation of QoE prediction covers time instants τ ranging from $0s$ to $150s$.

2) Clustering: The proposed clustering solutions have been implemented with PyClustering and Scikit-learn libraries, using the K-Means, X-Means, and Clustering Using REpresentatives (CURE) algorithm [19], [20]. The users are grouped over time according to QoS metrics. In order to evaluate the performance of clustering for each time instant τ, the dataset has been properly pre-processed. Firstly, the Min-Max Scaler has been applied to normalize the data. To select the number of clusters, the *silhouette* method has been adopted in K-Means and CURE, while X-Means is the extended version of K-means with efficient estimation of the number of clusters [21].

C. QoE prediction analysis

Different configurations of the LSTM network are evaluated for QoE prediction. Table I reports the configuration parameters and the prediction performance obtained with different values of *hidden size* and *number of layers*, as anticipated in Section III-B. Also, the number of trainable parameters is reported for the complexity analysis: the higher the number of parameters, the higher the complexity level. By making a trade-off between complexity and prediction performance, the best configuration is highlighted in bold. It achieves the lowest value of Mean Square Error (MSE) and Mean Absolute Error (MAE) for training and validation (30% of data) phases. Fig. 2 shows the prediction loss (i.e., MSE) as a function of the number of epochs considered during the training phase in order to analyze the convergence: stable loss values are reached in a short time.

To provide further insight, Fig. 3 shows the trend and the related prediction over time of MOS values for a generic test and unknown user: they almost overlap. Specifically, the black line represents the trend of the true MOS value observed in the sliding window T=$20s$ and the red dots are the predicted MOS values in the next time instant. Obtained results confirm the ability of LSTM to suitably process time series. The effectiveness of the proposed prediction approach is also demonstrated for a different number of users N_U, i.e. 10 and 20, over time, as shown in Fig. 4 for 10 example time instants. The figures highlight the median value (i.e., the red line), the 25^{th} and the 75^{th} percentile (i.e., the bottom line and the top line of the blue rectangle), as well as the minimum and the maximum MAE value (i.e., the edges of the vertical black line) of MOS predictions. It can be noted that the median error value for N_U=10 and N_U=20 does not exceed 0.02 and 0.05, respectively.

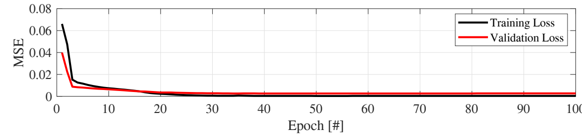


Fig. 2. MSE vs number of epochs for the best LSTM configuration.

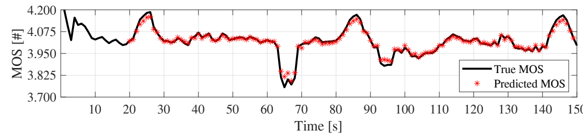
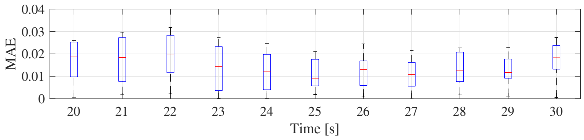


Fig. 3. MOS prediction for a generic unknown user.

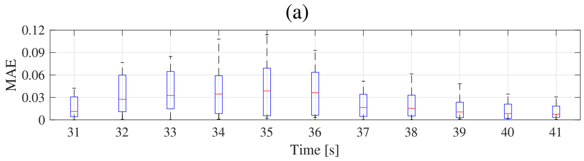
D. Clustering analysis

The clustering algorithms K-Means, X-Means, and CURE are applied to users at different time observation instants τ to analyze the time trend. The clustering performance for two representative time instants (τ=$60s$ and τ=$70s$) is illustrated in scatter-plots (Fig. 5). These plots show how considering delay and PLR, crucial QoS metrics for MOS calculation, aids in user clustering. User mobility is evident: for instance, the user with ID 568 (a train user in Norway) exhibits varying delay and PLR values between τ=$60s$ and τ=$70s$. At τ=$60s$, all algorithms produce two clusters (Fig. 5a). However, at τ=$70s$, K-Means and CURE generate three clusters, while X-Means two (Fig. 5b). Table II summarizes the clustering results with the average MOS values for each cluster ($\overline{MOS_x}$). Notably, at τ=$60s$, K-Means and X-Means have identical average MOS values, whereas CURE discerns differences among users in clusters 1 and 2. At $\tau = 70s$, K-Means and CURE exhibit similar behaviors, while X-Means fails to distinguish users with low MOS values in cluster 2, as observed through the other algorithms.

By analyzing the clustering performance at the different time instants, the results of K-Means, X-Means, and CURE algorithms are quite similar, especially for K-Means and CURE that adopt the silhouette method for the number of clusters. Since CURE has higher complexity with respect to K-Means [22], K-Means can be selected as the most suitable clustering

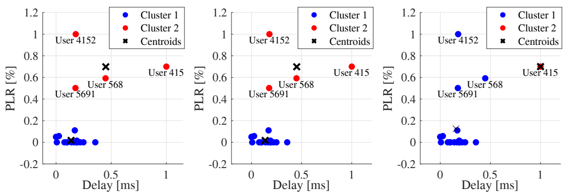


(a)

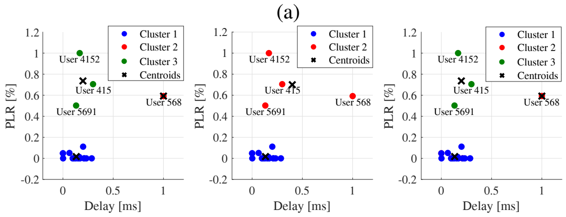


(b)

Fig. 4. MAE of MOS predictions over time for (a) $N_U = 10$ and (b) $N_U = 20$.



(a)



(b)

Fig. 5. K-Means, X-Means, and CURE at (a) $\tau = 60s$ and (b) $\tau = 70s$.

algorithm in the analyzed scenario for its high implementation efficiency.

As an example of the data (and therefore also storage and energy) reduction, the amount of information to be exchanged between the instance of DTN and the tenant with and without clustering (by using K-means) is shown below in the case of different numbers of users N_U and clusters N_C. For example, with N_U=20, the total number of the values of latency and PLR per user is 40, whose size is 2880 bytes for the baseline approach without clustering as shown in Fig. 6. In the same case but with the clustering approach, there are N_C=2 with 4 and 17 users in cluster 1 and cluster 2, respectively, and only 6 values, i.e., three values (average, minimum, and maximum MOS) per cluster, whose size is 160 bytes, are transmitted. Therefore, in this example case, the data saving between the clustering and baseline approach corresponds to 94.44%. To conclude, the proposed approach leads to a saving of exchanged data, becoming more prevalent with the increase in the number of users N_U, which will be grouped in a small number of clusters $N_C < N_U$.

IV. Optimization Opportunities and Future Directions

A preliminary investigation on resource management and redistribution has been conducted, neglecting mobility impairments, by assuming that the maximum allocation permitted by

TABLE II
Average MOS values for each Cluster obtained by different Clustering Algorithms.

	K-Means	*X-Means*	*CURE*
$\tau = 60s$	$\overline{MOS_1}$ = 2.971 $\overline{MOS_2}$ = 3.999	$\overline{MOS_1}$ = 2.971 $\overline{MOS_2}$ = 3.999	$\overline{MOS_1}$ = 1 $\overline{MOS_2}$ = 3.938
$\tau = 70s$	$\overline{MOS_1}$ = 4.031 $\overline{MOS_2}$ = 1.132 $\overline{MOS_3}$ = 3.919	$\overline{MOS_1}$ = 3.807 $\overline{MOS_2}$ = 3.222	$\overline{MOS_1}$ = 4.031 $\overline{MOS_2}$ = 1.132 $\overline{MOS_3}$ = 3.919

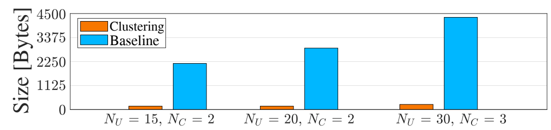


Fig. 6. Exchanged data size for Clustering and Baseline approaches in the case of different numbers of users N_U and clusters N_C.

the Network Operator is 5 Mbps [23]. As part of this analysis, a specific time observation instance as an example wherein users exhibit low and high MOS values has been examined. The idea is to reduce bandwidth for clusters with high MOS values, allocating more bandwidth to clusters with low MOS values to improve their QoE. For example, at τ=110s, there are two clusters with average MOS values $\overline{MOS_1}$=1.518 (poor quality) and $\overline{MOS_2}$=4.105 (good quality), respectively, as shown in Fig. 7. Subsequently, at τ=140s, increasing available bandwidth by 4.76% with respect to the maximum allocable for Cluster 1 and reducing it by 1.80% for Cluster 2, the average MOS value of Cluster 1 improved to $\overline{MOS_1}$=3.802, while the average MOS value of Cluster 2 is only reduced to $\overline{MOS_2}$=4.021, obtaining similar values per cluster.

Future research directions include the usage of physical resource blocks and their correlation with application-level statistics in order to formulate a DRL algorithm, which jointly exploits the QoE/QoS prediction and MOS-based clustering for resource optimization and redistribution. In this way, high levels of QoE will be achieved and energy efficiency will be improved. Specifically, the DRL module of the tenant's service orchestrator will facilitate resource redistribution in cases where the average MOS value of clusters indicates low QoE. This will lead to requests for an additional percentage of bandwidth from the DTN instance operated by the Network Operator. Such proactive adjustments will be driven by predicted statistics and cluster behaviors. The specific percentage increase in bandwidth allocation will be determined on a case-by-case basis by considering network constraints to avoid resource over-provisioning while optimizing the network configuration. After the complete design and evaluation of the AI (DL, ML, DRL) modules, the approach can be extended for multiple tenants. Moreover, the development of the remaining components will characterize future research activities to fully implement the designed framework. Also, in the context of Zero-touch network and Service Management (ZSM), DTN is a key enabler for proactive network management according to user needs and demands. Specifically, ZSM includes AI techniques to achieve higher levels of automation and efficiency. Thus, the proposed AI-based DTN could be investigated as a reference solution

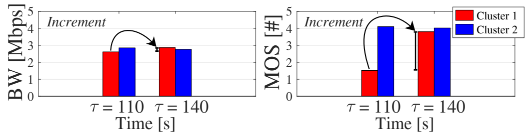


Fig. 7. Improvement in $\overline{MOS}$ (right) with an increase in the allocated bandwidth (left) between two example time instances.

for the creation of Zero-Touch Networks (ZTNs). Lastly, the whole framework could be also placed in the 6G-oriented scenario with integrated Terrestrial/Non-Terrestrial Networks, according to the latest 3rd Generation Partnership Project (3GPP) specifications.

V. Conclusions

This paper preliminarily designed a novel framework to effectively manage multimedia service provisioning in Beyond 5G and 6G communication systems through the integration of Artificial Intelligence and Digital Twin Networks, aligning with ongoing Internet-Draft efforts by the IETF. Its components and functionalities have been sketched, with a focus on Quality of Experience (QoE) estimation through the standardized Mean Opinion Score (MOS), QoE/Quality of Service (QoS) prediction, and user clustering. Different configurations of Long Short-Term Memory (LSTM) schemes have been efficiently employed to predict the temporal dynamics of MOS values. Moreover, three different clustering algorithms (i.e., K-Means, X-Means, and CURE) have been analyzed by considering real QoS metrics. The MOS values allow for effective representation of user clusters, reducing exchanged data size. Future research activities include the implementation of a Deep Reinforcement Learning approach for optimal resource allocation and redistribution.

Acknowledgment

This work was supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, in the context of partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”, CUP: D93C22000910001), national center on “Sustainable Mobility” (CN00000023 - program “MOST”, CUP: D93C22000410001), and partnership on “Cybersecurity” (PE00000007 - program “SERICS”, CUP: D33C22001300002, project ISP5G+). It was also supported by the PRIN 2022 projects INSPIRE (grant no. 2022BEXMXN_01) and HORUS (grant no. 2022P44KA8) funded by the Italian MUR, by the HORIZON MSCA project BRIDGITISE (grant no. 101119554) and the HORIZON JU SNS project 6G-GOALS (grant no. 101139232), and by “The house of emerging technologies of Matera (CTEMT)” project funded by the Italian MIMIT.

References

[1] A. Masaracchia, V. Sharma, B. Canberk *et al.*, “Digital Twin for 6G: Taxonomy, Research Challenges, and the Road Ahead,” *IEEE Open Journal of the Communications Society*, vol. 3, pp. 2137–2150, 2022.

[2] Z. Zhang, Y. Zhou, L. Teng *et al.*, “Quality-of-Experience Evaluation for Digital Twins in 6G Network Environments,” *IEEE Transactions on Broadcasting*, pp. 1–13, 2024.

[3] H. D. Moura, J. M. Oliveira, D. Soares *et al.*, “Improved Video QoE in Wireless Networks Using Deep Reinforcement Learning,” in *2023 19th International Conference on Network and Service Management (CNSM)*, 2023, pp. 1–7.

[4] I.-S. Comșa, A. Molnar, I. Tal *et al.*, “Improved Quality of Online Education Using Prioritized Multi-Agent Reinforcement Learning for Video Traffic Scheduling,” *IEEE Transactions on Broadcasting*, vol. 69, no. 2, pp. 436–454, 2023.

[5] A. del Río, J. Serrano, D. Jimenez *et al.*, “A Deep Reinforcement Learning Quality Optimization Framework for Multimedia Streaming over 5G Networks,” *Applied Sciences*, vol. 12, no. 20, 2022.

[6] J. J. A. Arnez and C. B. B. De Souza, “Analysis of inter-Mean-Opinion-Score (iMOS) and QoS metrics for 5G Voice over LTE (VoLTE) and Voice over New Radio (VoNR) services over IP-Multimedia-Subsystem (IMS) in 5G mobile networks,” in *2023 6th International Conference on Advanced Communication Technologies and Networking*, 2023.

[7] C. Thorpe, A. Hava, J. Langlois *et al.*, “iMOS: Enabling VoIP QoS Monitoring at Intermediate Nodes in an OpenFlow SDN,” in *2016 IEEE International Conference on Cloud Engineering Workshop (IC2EW)*, 2016, pp. 76–81.

[8] C. Zhou, H. Yang, X. Duan *et al.*, “Network Digital Twin: Concepts and Reference Architecture,” Internet Engineering Task Force, Internet-Draft draft-irtf-nmrg-network-digital-twin-arch-05, Mar. 2024, work in Progress. [Online]. Available: https://datatracker.ietf.org/doc/draft-irtf-nmrg-network-digital-twin-arch/05/

[9] M. El Rajab, L. Yang, and A. Shami, “Zero-touch networks: Towards next-generation network automation,” *Computer Networks*, p. 110294, 2024.

[10] Z. Latif, K. Sharif, F. Li *et al.*, “A comprehensive survey of interface protocols for Software Defined Networks,” *Journal of Network and Computer Applications*, vol. 156, p. 102563, 2020.

[11] A. Lombardo, G. Morabito, S. Quattropani, and C. Ricci, “Design, implementation, and testing of a microservices-based Digital Twins framework for network management and control,” in *2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2022, pp. 590–595.

[12] X. Lin, L. Kundu, C. Dick *et al.*, “6G Digital Twin Networks: From Theory to Practice,” *IEEE Communications Magazine*, vol. 61, no. 11, pp. 72–78, 2023.

[13] S. Barmpounakis, N. Maroulis, N. Koursioumpas *et al.*, “AI-driven, QoS prediction for V2X communications in beyond 5G systems,” *Computer Networks*, vol. 217, p. 109341, 2022.

[14] A. A. Barakabitze, N. Barman, A. Ahmad *et al.*, “QoE Management of Multimedia Streaming Services in Future Networks: A Tutorial and Survey,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 526–565, 2020.

[15] Recommendation ITU-T G.107, “The EModel: a computational model for use in transmission planning,” ITU-T, Tech. Rep., 2015.

[16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, http://www.deeplearningbook.org.

[17] C. Midoglu, M. Moulay, V. Mancuso *et al.*, “Open video datasets over operational mobile networks with MONROE,” in *Proceedings of the 9th ACM Multimedia Systems Conference (MMSys ’18)*, 2018, p. 426–431.

[18] A. Paszke, S. Gross, F. Massa *et al.*, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Advances in Neural Information Processing Systems 32*, H. Wallach *et al.*, Eds. Curran Associates, Inc., 2019, pp. 8024–8035.

[19] F. Pedregosa, G. Varoquaux, A. Gramfort *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[20] A. V. Novikov, “Pyclustering: Data mining library,” *Journal of Open Source Software*, vol. 4, no. 36, p. 1230, 2019.

[21] D. Pelleg and A. W. Moore, “X-means: Extending k-means with efficient estimation of the number of clusters.” in *Icml*, vol. 1. Stanford, CA, 2000, pp. 727–734.

[22] S. Jian, G. Pang, L. Cao *et al.*, “CURE: Flexible Categorical Data Representation by Hierarchical Coupling Learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 5, pp. 853–866, 2019.

[23] V. Vasilev, J. Leguay, S. Paris, L. Maggi, and M. Debbah, “Predicting QoE factors with machine learning,” in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.