Efficient Resource Allocation for Video Streaming for 5G Network-to-Vehicle Communications

Farhan Pervez
Technology & Operations
Saudi Telecom Company, Saudi Arabia
Email: fpervez.c@stc.com.sa

Abdulkareem Adinoyi
Technology & Operations
Saudi Telecom Company, Saudi Arabia
Email: aadinoyi.c@stc.com.sa

Halim Yanikomeroglu
Systems & Computer Engineering
Carleton University, Canada
Email: halim@sce.carleton.ca

***Abstract*—Unprecedented bandwidth-intensive, ultra-low latency multi-media services and applications are expected to be delivered through the fifth generation (5G) network. Thus, efficient radio resource management will continue to play an important role as the network service providers strive to provide rich media contents to their users. Therefore, in this work, we consider a cross-layer based framework for efficient allocation of network resources for the transmission of HEVC encoded video streams over 5G-aligned V2X application in mmWave environments. We adopt application and physical layer models, and formulate a multi-vehicle resource allocation problem taking into account the abstracted information of the two corresponding layers. Such a framework is compared with three other standard resource allocation schemes using computed Mean Opinion Score (MOS). Numerical results show that our proposed radio resource management scheme offers a significant improvement of viewer-perceived quality of service compared to the reference approaches.**

Keywords. *HEVC (High Efficiency Video Coding), 5G New Radio, mmWave (millimeter wave), V2X (Vehicle-to-everything), radio resource management (RRM).*

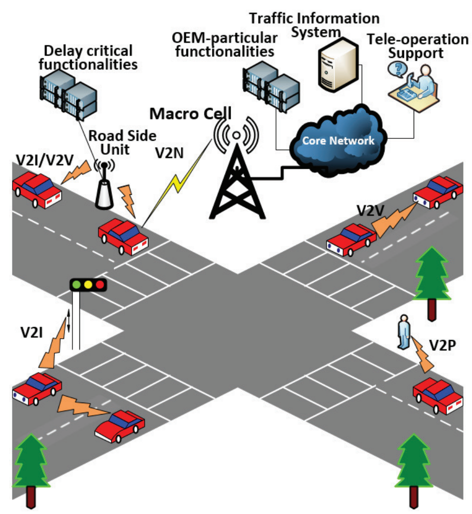

Fig. 1. Example of V2X deployment, transport options and services offering [Note: Here X represents V for vehicle, or N for Network, I for Infrastructure].

I. Introduction

The exciting developments in market tendencies for Internet-of-Things (IoT) and connected vehicles, among other IoT business verticals, create a shift towards more responsive, smart and service-oriented transportation systems. It is expected that Vehicle-to-Everything (V2X) communications, a part of the 5G use cases, will play a significant role in ushering the next thrust of applications and services. Thus, these developments provide interesting scenarios of focus for researchers and other stakeholders to understand the performance of the 5G network, challenges, and opportunities, especially for video streaming and multi-media communications.

Standardization bodies are already doing their part. For instance, V2X is being treated as an important component of 5G networks design in 3GPP Release 15, which is saddled with the responsibility to provide enhancements, not only in terms of the large connectivity requirements, but also for the data rates, latency, and reliability for V2X environment [1]. In numerical terms for example, the goal is to deliver 1000 times more data throughput, serve 10 to 100 times more devices and reduce the minimum latency by a factor of 5 compared to the existing 3GPP LTE networks [2]. To achieve such ambitious performance targets, a large amount of spectrum will be required to facilitate the delivery of ultra-high definition videos. This serves as a motivation in the study for video transmission in the mmWave bands. Fig. 1 demonstrates the proposed design of the future 5G network for providing V2X services that can be supported by different integrated communication subsystems, such as IEEE 802.11p, Vehicular Visible Light Communication (VVLC) [3].

Furthermore, on the motivation, we emphasize on the critical role of radio resource management for a 5G-aligned connected cars system and propose a cross-layer scheduling algorithm for an efficient utilization of the available radio spectrum. It is known that RRM algorithms are designed to exploit the dynamism in

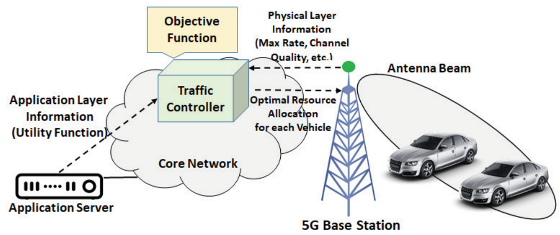

Fig. 2. Typical network-centric scenario and high-level system architecture.

wireless channels by adaptively distributing communication resources towards maximizing or minimizing some important key performance metrics for network performance [4]. In view of the above, this paper considers the application layer modeling of the video streams that are encoded using the bandwidth efficient HEVC mechanism, considering the mmWave channel model and LTE numerology as the key aspects for a realistic evaluation of the performance of V2X in 5G networks. The proposed optimization framework takes information from the application layer, resident at the application server, and the physical layer, transmitted from the 5G base station.

Figure 2 shows an illustration for the V2N scenario. The vehicles are directly served by a 5G network infrastructure. In this case a number of streams are served to the vehicles through a 5G antenna beam using a high-quality HEVC encoded video that is resident at the application server. The resource allocation is embedded in the antenna beam, which is a direct consequence of the massive MIMO/antenna configuration of the 5G. The Traffic Controller (TC) module acts as an optimizer for downlink resource allocation, taking into account the video utility function from the application server, the physical layer information related to the vehicles from the 5G base station, and the objective function. The utility function is either stored at the TC in advance during the session establishment or sent along with the video streams. To optimally distribute resources among the vehicles, the instructions from the TC to the resource allocator (e.g., base station) are transmitted using an enhanced interface in the network system.

To the best of the authors' knowledge, this work is the first to investigate the assignment of optimized resources for 5G-aligned connected cars engaged in video streaming. Previously, resource management techniques have been studied for LTE/LTE-A based V2X systems such as [5], but the assignments are based on a single layer information (i.e., throughput or channel quality). They do not take into account the viewers' perceived quality. In other works, resource allocation has been studied either for massive machine-type communication, such as in [6] or for heterogeneous but non-V2X environment, such as [7], [8], and [9].

The rest of the paper is organized as follows. In the next section, the application layer model including the video quality metric and the utility curves for each quality scalable HEVC encoded video is presented. Insights into the physical layer model are also provided. Section III describes the proposed radio resource management framework over the downlink. Section IV presents the simulation results comparing the proposed scheme with some other standard resource allocation approaches, and finally in Section V, we draw some conclusions from the results presented in the paper.

II. System Model

Below, we discuss the application and physical layer models as adopted in this paper to enable the formulation of the multi-vehicle resource allocation problem for a V2N scenario considering the abstracted information of the two corresponding layers.

A. Application Layer Model

For streaming video applications, especially those in which compression is implemented due to limited bandwidth and challenging 5G air interface, a video quality metric provides a numerical indication of the quality of experience from the viewers perspective, of received media, after compression and/or transmission. Therefore, for our application layer model, we use the concept of utility function and evaluate the viewer perception by using spatio-temporal video quality metric as proposed in [10]. This metric, evaluated on the scale of 0 to 100, is based on peak signal-to-noise ratio (PSNR), frame rate as well as spatial and temporal activity that are obtained from the videos. Furthermore, it is appealing for dynamic optimization and assessing multimedia transmission because it is content-independent.

Because the physical layer provides a retransmission mechanism, we neglect the impact of packet loss. Therefore, the utility function can be expressed as a function of application data rate, defined as follows:

$$U = f(R_{app}), \tag{1}$$

where R_{app} is the set of possible encoded rates of application, and $U = [0; 100]$ with $U : 0$ reflecting an unacceptable application quality and $U : 100$ referring to the best possible quality experienced by the viewer. This value is mapped onto MOS by using the ITU-adopted expression [11], given as:

$$MOS = \begin{cases} 1, & U < 0; \\ 1 + 0.035U + 7U(U-60) & \\ \times(100-U)10^{-6}, & 0 \leq U \leq 100; \\ 4.5, & U > 100. \end{cases} \tag{2}$$

Fig. 3 shows an example of perceived video quality curves as a function of the encoded bitrate for four different test videos, encoded with the quality scalable HEVC codec [12] at 30 fps for 2K resolution $(2,048 \times 1,080$

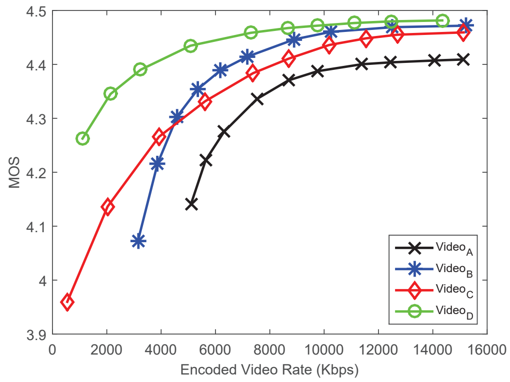

Fig. 3. Four utility curves versus the encoded bit rate.

pixels), comprising of ten layers each that include one base layer and nine enhancement layers. These utility curves are used as an application layer model for formulating the proposed radio resource management.

B. Physical Layer Model

1) Channel Model: An important step in the design of the physical layer is to identify the role of an adequate and appropriate channel model, one that is suitable for the V2X application. To design a 5G V2X system that can serve the challenging V2N use-case, we need to select the channel model carefully. As many of the existing channel models are either developed for specific frequency scenario outside the mmWave bands and for limited vehicle speeds, usually low speeds. In terms of the V2X channel generation process, we can distinguish four main steps. The first is the need to select the link type (e.g., V2V, V2I, V2N, and V2P) and scenario (e.g., highway, urban, rural). The second is to assign propagation condition to the link. The third is calculating the path loss and shadowing; and the last to calculate small scale parameters (such as multi-path delays, arrival and departure angles, etc.). In this work, we have considered V2N, the urban macro cellular and only path loss and shadow fading phenomena since the small-scale part of fading can easily be averaged out.

In view of the above, the model presented in [13] is a suitable one and therefore adopted in this work. It recommends close-in (CI) path loss model obtained through extensive measurement campaign and employing ray-tracing in the frequency band covering 2 GHz to 73.5 GHz. The CI model is physically tied to the transmitter power using a close-in free space reference, and standardizes all measurements around an inherent 1 m free space reference distance. Furthermore, the CI model has a very similar form compared to the existing 3GPP path loss model and it provides high accuracy over a vast range of cm-wave and mmWave frequencies, distances, and scenarios. The CI model is given in [14], as:

$$\begin{aligned} PL^{CI}(f,d)[dB] &= FSPL(f,1m)[dB] \\ &\quad +10n\log_{10}(d)+\mathcal{X}_{\sigma}^{\mathcal{CI}} \end{aligned} \quad (3)$$

where n denotes the single model parameter, the path loss exponent, d is the 3D T-R separation distance, $\mathcal{X}_{\sigma}^{\mathcal{CI}}$ is the shadow fading standard deviation and $FSPL(f,1m)$ denotes the free space path loss in dB at a T-R separation distance of $1m$ at the carrier frequency f:

$$FSPL(f,1m)[dB] = 20\log_{10}\left(\frac{4\pi f}{c}\right) \quad (4)$$

where c denotes the speed of light. The estimated value, as given in [13], for n is 2.7 and for $\mathcal{X}_{\sigma}^{\mathcal{CI}}$ is 10 dB, considering urban macro cellular cm-wave and mmWave environments over a T-R distance of between 45 m to 1450 m.

2) Numerology: LTE deploys a rigid frame structure with a fixed Transmission Time Interval (TTI) of 1 ms, contributing to the end-to-end latency. Therefore, to meet the strict latency requirements of V2X, the TTI may be flexibly shortened for 5G as it is being proposed for the 5G New Radio (NR). This can be achieved by increasing the sub-carrier spacing; hence reducing the symbol duration. Furthermore, having a flexible numerology offers an additional degree of freedom to adapt physical layer transmission to various service requirements and channel conditions. In this paper, we considered the 3GPP adapted LTE numerology for the 5G, where a scaling factor of 2^s (where $s = 0, 1, 2, \cdots, N$) is used for implementing the 5G new radio, [15].

As a result, for a scaling factor 4 (i.e., $s = 2$), the sub-carrier spacing of the self-contained frame structure would be tuned to $2^s\times$ 15 KHz, i.e., 60 KHz, the number of physical resource blocks (PRBs) would be $2^s\times$ (Num_PRB_LTE), the PRB size would be $2^s\times$ (PRB_Size_LTE i.e., 180 KHz), a cyclic prefix of $1\mu s$, a guard period of $20.33\mu s$ and the allocated number of symbols would be 12.

3) Spectral Throughput: The system spectral throughput given by the modified Shannon capacity:

$$Thr = \beta\log_2(1 + SINR) \quad (5)$$

indicates the achievable throughput per physical resource block (PRB) in bps/Hz for a given signal-to-interference-plus-noise ratio (SINR). The model approximates the throughput over downlink, after considering link adaptation, hybrid automatic repeat request and coding rate, with a loss β, due to the implementation constraints. The spectral efficiency is taken into account to estimate the maximum achievable rate for each vehicle in the system over the downlink antenna beam.

III. CROSS-LAYER RESOURCE ALLOCATION

The Traffic Controller (TC) as depicted in Fig. 2 is the main entity that decides how to allocate resources across more than one vehicles, requesting video streams within a single antenna beam. The TC finds the most appropriate allocation by taking into account the physical layer information, the application-layer data and the objective function. In this paper, we implement the utility maximization as an example of the objective function, which aims at maximizing the average utility of all K viewers subject to the constrained network resources. The resource-dependent optimization problem for utility maximization is a convex function and defined as follows:

$$\tilde{\alpha}_{opt} = \arg\max_{\alpha} \left(\frac{1}{K} \sum_{k=1}^{K} U_k(\alpha_k) \right) \quad (6)$$

subject to:

$$\sum_{k=1}^{K} \alpha_k \leq 1, \quad (7)$$

where α_k is the fraction of network radio resources given to the vehicle k, and α is all possible sets of vectors of network resources shared for each vehicle. $\tilde{\alpha}_{opt}$ is the optimal resource allocation tuple that achieves the possible maximum of the objective function. $U_k(\alpha_k)$ is the utility function of the given resource share, α_k of vehicle k. The inequality in the defined constraint is due to the limited number of scalable layers we have generated for HEVC encoded video streams. Searching for an optimal resource allocation can be done via a full search algorithm, however, it is computationally expensive and not feasible to appeal to practical and real-time use, especially when there are many connected vehicles in the system. Therefore, we apply the max slope-based resource allocation framework, to solve the optimization problem due to its low computational complexity and short execution time [16]. This is particularly useful for solving discrete resource allocation problems with a linear resource constraint. In principle, as depicted in Fig. 4, when the optimization cycle occurs, the proposed algorithm determines the resource share of every vehicle in the system, by first assigning the resources to all vehicles that are enough to transmit the base layer of the HEVC encoded video stream. Afterwards, in order to send the enhancement layer, the algorithm then assigns the resources to the vehicle whose video provides the maximum gradient on the utility curve. Here, the gradient is defined as the increase in utility with respect to the amount of resources required for that increase. The algorithm runs by assigning resources to the next vehicle that has the maximum gradient on the utility curve. It stops if: a) all resources are utilized, b) all layers are transmitted or c) the remaining resources are not sufficient to transmit the next higher layer of any of the videos.

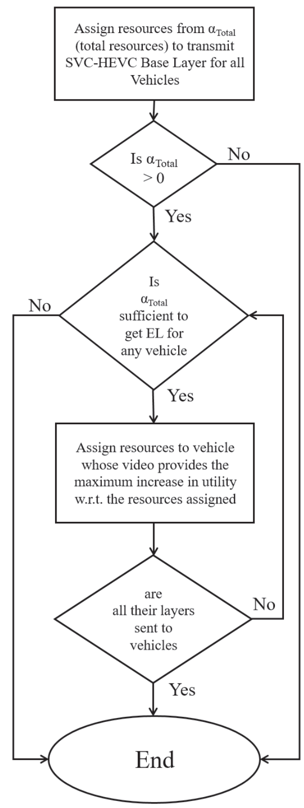

Fig. 4. The flowchart for assigning resources for base layer and enhancement layers (EL).

We consider a radio layer model with resource allocation period of one second, not taking into account short-term channel variations. Rather, we adopt an average channel quality for each vehicle over the resource allocation period. The current data rate R_k for vehicle k is calculated depending on its resource share α_k (obtained from $\tilde{\alpha}_{opt}$), and the maximum achievable data rate $R_{max,k}$, calculated by using the spectral throughput for this vehicle and for this resource allocation period if all resources would be allocated to the vehicle k, is given below as:

$$R_k = \alpha_k R_{max,k}, \quad 0 \leq \alpha_k \leq 1 \quad \forall_k \in N. \quad (8)$$

This achievable data rate R_k for vehicle k is used to estimate the MOS using (1), as defined in Section II.

IV. NUMERICAL RESULTS

The results presented in this section of the paper are based on the system model described in Fig. 2. Since Network Function Virtualization (NFV) will be a key enabler of 5G network, we assume that the core network will not constitute congestion challenges for the network. The main bottle-neck will be provided by the air-interface constraints. Therefore, we have assumed in

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency	38 GHz
5G NR scaling factor	4
System bandwidth	320 MHz
Number of PRBs	400
PRB size	12 subcarriers
Sub-carrier spacing	60 KHz
Bandwidth per PRB	720 KHz
Sub-frame interval	0.25 ms
Number of vehicles	63
Vehicle speed	36 Kmph
T-R distance	50 m - 1450 m
Channel model	Urban macro-cell

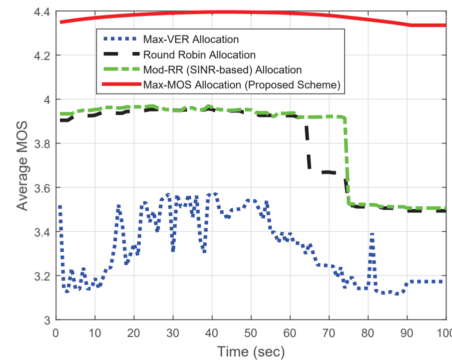

Fig. 5. The average MOS performance of the video streamers using the proposed scheme and other reference resource allocation schemes.

this investigation that the packets suffer no congestion over the 5G core network. Furthermore, we consider a single antenna beam and one array of elements serving multiple vehicles. Along with that, the inter-beam handover case is also left out. There are four different HEVC encoded test videos being streamed by the vehicles as an infotainment service. Table I provides a summary of the simulation parameters employed in the performance evaluation of the proposed set-up.

Here, we evaluate the proposed scheme based on maximizing the MOS (Max-MOS) and compare it with some well-known resource allocation schemes, such as the classical round robin (RR), the modified round robin that takes SINR into account (Mod-RR) and the maximizing video encoding rate (Max-VER). The number of optimization cycles is 100 and the total simulation time is 100 seconds. For mmWave environments this simulation time is enough to evaluate the performance of the resource allocation schemes, as the vehicle is expected to be handed over to another radio beam or base station given the high speed that it will be moving. Furthermore, to confirm our results for the vehicles that remain connected to the same beam, we implement multiple simulation scenarios. The objective of maximizing VER resource allocation is to determine the resource share α_k of each streaming vehicle k in such a way to maximize the overall video encoding rate of the antenna beam. Conversely, the modified RR allocation assigns resources starting from the vehicle with maximum SINR moving towards that with minimum SINR in a round robin fashion. Next, we present the numerical results.

Fig. 5 represents the comparison between four different resource management techniques in terms of average MOS achieved for a single scenario. The red curve shows the performance of the proposed scheme whereas the green curve shows the average MOS achieved by using Mod-RR allocation. The black curve shows the average MOS of the viewers for RR scheduling and the blue one shows the average MOS of the Max-VER algorithm.

The results show the advantage in terms of viewers' experience for our proposed framework in comparison to the other three radio resource management schemes. The figure shows that as the vehicles come to a halt after 90 seconds the average MOS remains constant there on. We can also observe that the RR allocation outperforms the Max-VER framework that is due to the large number of resources (PRBs) available to be fairly distributed between the vehicles. This is because in RR allocation, all the streamers get a high number of layers therefore having a higher average MOS. Contrarily, in Max-VER, to maximize the video encoding rate the algorithm may provide a higher number of layers for some viewers and no layers for some resulting in the minimum MOS value of 1 for such viewers. In this case, the resource distribution strategy is unfair and the average MOS is severely degraded. Such impact can be clearly seen in Fig. 6, in which we investigate the number of vehicles perceiving minimal quality of MOS 1 over the simulation time. Here, we see that there are many viewers with MOS 1 for the Max-VER scheme. We expect the Max-VER to outperform RR if there are much fewer resources available or very high video encoded rates.

To gain further insight into the system performance, we extend the simulation to cover different scenarios (e.g., by changing the position and the mobility track of each vehicle) and computing the Cumulative Distribution Function (CDF) for the two system performance metrics, namely: the average MOS of all streamers and the number of streamers with MOS equal to 1, as depicted in Fig. 7 and Fig. 8. Fig. 7 shows that the proposed scheme provides a superior average MOS performance over the reference schemes. An improvement of 0.5 or more in average MOS is obtained compared to the other approaches for 80% of the time. Moreover, in terms of the number of streamers with MOS of 1

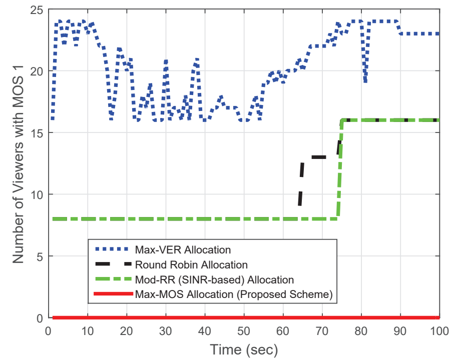

Fig. 6. The number of viewers experiencing minimal quality of MOS 1 for the proposed scheme and the other reference resource allocation schemes.

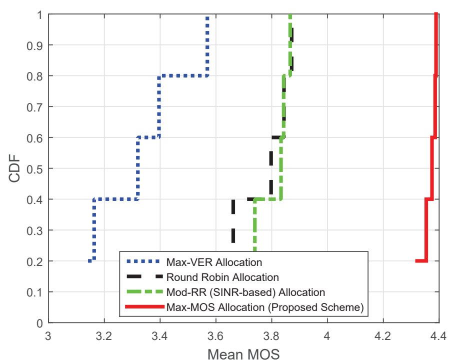

Fig. 7. The CDF of the average MOS of all streamers.

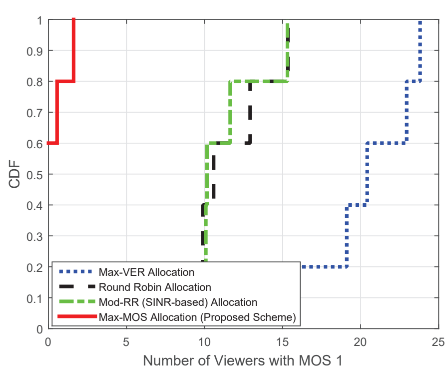

Fig. 8. The CDF of the number of viewers experiencing minimal quality for the different resource allocation schemes.

(worst video quality), the proposed scheme also shows a considerably superior performance advantage over the reference schemes as observed in Fig. 8.

V. Conclusion

This paper presents a cross-layer based framework for efficient allocation of network resources for the transmission of HEVC encoded video streams for connected-cars in 5G networks operated in the mmWave regime. Using a suitable application and physical layer models, we have formulated a multi-vehicle resource allocation problem, which has been solved using the max-slope algorithm. We have compared the obtained numerical results with three other standard resource allocation schemes using computed mean opinion score, as a measure of the level of viewers' experience and satisfaction. We have obtained an improvement that is greater than 0.5 in terms of average MOS over all other reference approaches. Furthermore, in terms of the number of streamers with worst possible video quality, the proposed scheme also demonstrates a substantial performance advantage over the reference schemes.

References

[1] 3GPP TR 22.886 V1.1.0, "Study on enhancement of 3GPP Support for 5G V2X Services (Release 15), 3GPP Std," *3GPP,* Nov. 2016.

[2] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, et al., "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Communications Magazine,* 52(5), pp. 44-51, 2014.

[3] M. Boban, K. Manolakis, M. Ibrahim, S. Bazzi, and W. Xu. "Design aspects for 5G V2X physical layer," *IEEE Conference Communications and Networking,* pp. 1-7, 2016.

[4] M. Salem, A. Adinoyi, M. Rahman, H. Yanikomeroglu, D. Falconer, Y. Kim, et al, "An overview of radio resource management in relay-enhanced OFDMA-based networks," *IEEE Communications Surveys and Tutorials,* 12(3), pp. 422-438, 2010.

[5] P. Luoto, M. Bennis, P. Pirinen, S. Samarakoon, K. Horneman, and M. Latva-aho, "System Level Performance Evaluation of LTE-V2X Network," *European Wireless Conference,* pp. 1-5, 2016.

[6] N. H. Mahmood, M. Lauridsen, G. Berardinelli, D. Catania, and P. Mogensen, "Radio resource management techniques for eMBB and mMTC services in 5G dense small cell scenarios," *IEEE VTC-Fall,* pp. 1-6, 2016.

[7] M. Hasan, and E. Hossain, "Distributed resource allocation in 5G cellular networks," *Towards 5G: Applications, Requirements and Candidate Technologies,* pp. 129-161, 2014.

[8] N. Wang, E. Hossain, and V. K. Bhargava, "Backhauling 5G small cells: A radio resource management perspective," *IEEE Wireless Communications,* 22(5), pp. 41-49, 2015.

[9] N. Vo, T. Q. Duong, and M. Guizani, "QoE-Oriented Resource Efficiency for 5G Two-Tier Cellular Networks: A FemtoCaching Framework," *IEEE Global Communications Conf. (GLOBECOM),* pp. 1-6, 2016.

[10] Y. Peng, and E. Steinbach, "A novel full-reference video quality metric and its application to wireless video transmission." *IEEE International Conf. in Image Processing,* pp. 2517-2520, 2011.

[11] ITU-T, "The E-model: a computational model for use in transmission planning," *ITU-T G.107,* June 2015.

[12] T. Hinz, P. Helle, H. Lakshman, M. Siekmann, J. Stegemann, H. Schwarz, et al., "An HEVC extension for spatial and quality scalable video coding," *Visual Information Processing and Communication (IV13),* 2013.

[13] S. Sun, T. S. Rappaport, S. Rangan, T. A. Thomas, A. Ghosh, I. Z. Kovacs, et al, "Propagation path loss models for 5G urban micro-and macro-cellular scenarios," *IEEE Vehicular Technology Conf. (VTC Spring),* pp. 1-6, 2016.

[14] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. on Communications,* 63(9), pp. 3029-3056, 2015

[15] 3GPP, "Report of 3GPP TSG RAN WG1 86 v0.2.0, 3GPP Std.," *3GPP* August 2016

[16] F. Pervez, M. S. Raheel, J. Gill, and K. Iqbal, "Uplink Resource Allocation of Multiple DASH Streams for QoE-based Real and Non-real Time Viewing over LTE," *IEEE International Conference on Next Generation Mobile Applications, Security and Technologies,* pp. 12-17, 2016.