Al-Quraan, M., Khan, A., Mohjazi, L. , Centeno, A. , Zoha, A. and Imran, M. A. (2023) Intelligent beam blockage prediction for seamless connectivity in vision-aided next-generation wireless networks. *IEEE Transactions on Network and Service Management*, 20(2), 1937 - 1948. (doi: [10.1109/TNSM.2022.3216556](https://doi.org/10.1109/TNSM.2022.3216556))

Intelligent Beam Blockage Prediction for Seamless Connectivity in Vision-Aided Next-Generation Wireless Networks.

Mohammad Al-Quraan, *Student Member, IEEE,* Ahsan Khan, Lina Mohjazi, *Senior Member, IEEE,* Anthony Centeno, Ahmed Zoha, *Senior Member, IEEE,* and Muhammad Ali Imran, *Senior Member, IEEE*

Abstract—The upsurge in wireless devices and real-time service demands force the move to a higher frequency spectrum. Millimetre-wave (mmWave) and terahertz (THz) bands combined with the beamforming technology offer significant performance enhancements for future wireless networks. Unfortunately, shrinking cell coverage and severe penetration loss experienced at higher spectrum render mobility management a critical issue in high-frequency wireless networks, especially optimizing beam blockages and frequent handover (HO). Mobility management challenges have become prevalent in city centres and urban areas. To address this, we propose a novel mechanism driven by exploiting wireless signals and on-road surveillance systems to intelligently predict possible blockages in advance and perform timely HO. This paper employs computer vision (CV) to determine obstacles and users' location and speed. In addition, *this study introduces a new HO event, called block event (BLK), defined by the presence of a blocking object and a user moving towards the blocked area*. Moreover, the multivariate regression technique predicts the remaining time until the user reaches the blocked area, hence determining best HO decision. Compared to conventional wireless networks without blockage prediction, simulation results show that our BLK detection and proactive HO algorithm achieves 40% improvement in maintaining user connectivity and the required quality of experience (QoE).

Index Terms—Computer vision, object detection, machine learning, blockage prediction, proactive handover, mobility management, mmWave communications, ultra-dense networks.

I. INTRODUCTION

Millimetre-wave (mmWave) and terahertz (THz) technologies are vital in supporting beyond fifth-generation (B5G) and sixth-generation (6G) networks. The dependence on new high-frequency bands is expected to achieve the global connectivity vision by providing significant enhancements in terms of multi-Gbit/s throughput, supporting a massive number of devices, and delivering ultra-low latency and reliable connections [1]. Moreover, the transition towards higher bands changes the paradigm of future wireless networks to small coverage cells, and thus, forming the concept of ultra-dense networks (UDNs) [2]. High-frequency wireless networks are needed to meet the stringent broadband access demands and realise various revolutionary applications, such as intelligent healthcare, holographic telepresence, and autonomous driving.

M. Al-Quraan, A. Khan, L. Mohjazi, A. Centeno, A. Zoha, and M. A. Imran are with the James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, UK. (e-mail: {m.alquraan.1, a.khan.9}@research.gla.ac.uk, {Lina.Mohjazi, Anthony.Centeno, Ahmed.Zoha, Muhammad.Imran}@glasgow.ac.uk).

Future networks leverage mmWave and THz multiarray antennas that offer beamforming capabilities, which focus the power of radio signal towards the receiving device through line-of-sight (LoS) communication. Beamforming is expected to be extensively used in next-generation networks owing to the provision of remarkable features, like high spatial reuse, increased throughput, boosting capacity, and interference elimination [3]. Despite the outstanding merits offered by such networks, relying on high-frequency beam-based communications is more sensitive to the adverse effects of blockage and penetration losses than those operating at lower frequency bands. For instance, a link budget undergoes a 20 dB or more power loss when the connection is blocked by obstacles, such as human bodies or vehicles [4], [5]. Such a sudden drop in the received power affects the signal quality and degrades the data rate of the communication link, making the network unreliable for time-sensitive applications.

Network densification aims to serve the highly populated urban areas and meet user demands and traffic capacity. However, cell coverage shrinkage and the presence of dynamic users/obstacles means mobility management in UDN is far more complex than in legacy networks, resulting in the recurrence of challenging issues, namely beam blockage and frequent handover (HO) [6]. HO is a fundamental mechanism in any wireless network that transfers the ongoing call or data session from one base station (BS) to another. The 3rd Generation Partnership Project (3GPP) organisation introduced several predefined measurement events; if one occurred, HO must be conducted [7]. Typically, a user equipment (UE)-assisted network for controlling HO receives a measurement report from the user with information about the received signal strength (RSS)/quality of a specific downlink reference signal from the serving BS (S-BS) and other neighbouring BSs. If the condition of a specific event is met, the network will trigger the HO process and negotiations eventuate between the S-BS and the target BS (T-BS) to complete handing the user to the new BS, thus guaranteeing user connectivity.

Unlike in 3G/4G networks, mobility management in B5G networks is accompanied with negative impacts at both user and network levels. Frequent HOs are leading to more data transmission delays and throughput loss in the UE as well as causing increased power consumption and poor network quality of service (QoS), especially if there is rejection from the T-BSs due to full resource occupation. This problem is exacerbated in smart cities due to the highly dynamic

environment and the existence of blocking objects that can shade the serving beam. Selection/reselection of the best beam is a time-consuming task that will result in additional network operating costs. Accordingly, beam blockage and frequent HO are attracting the attention of research bodies in academia and industry to find new solutions that can improve the reliability of UDNs.

A. Related Work

Despite numerous benefits gained when shifting operational frequencies from the lower bands (sub-6GHz) to the higher bands (mmWave and THz), reliance on mmWave and THz technologies introduces critical challenges, such as link blockage and frequent HOs. To this end, many attempts have been made to provide solutions to address the connectivity issue in mmWave networks. In [8], the authors count on the geometry of mmWave channels to predict when and for how much time an LoS link will be blocked by observing the behaviour of the neighbouring non-LoS (NLoS) signals. The main idea is that a connected user will be served by LoS and NLoS links, and detecting a blockage in one of the NLoS connections can be exploited to anticipate when the LoS link will be blocked, thus performing HO proactively. However, such techniques are not efficient in practical scenarios due to assuming slowly moving obstacles, highly scattering environment, and NLoS will be blocked before the LoS links. The work in [9] explores the use of channel state information of the sub-6GHz channels and the effect of frequency-dependent diffraction to form an early warning of possible mmWave signal blockages in hybrid communication systems. Motivated by the fact that the diffraction angle decreases as the frequency increases, the diffracted sub-6GHz signals reach a certain signal strength threshold earlier than the mmWave signals. This work relies on simulations to validate the proposed method, while the authors did not measure the sub-6GHz and mmWave diffraction in reality to identify a significant difference. Similarly, the work in [10] predicts mmWave beam blockages based on diffraction fringe characteristics on mmWave and sub-6GHz signals. When a user approaches a blocking object, the received power fluctuates with the increasing amplitude. This phenomenon is observed earlier on sub-6GHz than mmWave signals. Therefore, mmWave beam blockages can be expected by noting the fluctuating patterns of the sub-6GHz signals. However, employing both frequency bands increases the cost and the complexity of the wireless network.

Liu and Xiao [11] followed a different mechanism to predict beam blockage in heterogeneous mmWave networks by observing the previous beamforming vectors and their fingerprints. The cloud radio access network (C-RAN) is used to maintain a fingerprinting database that can be used to predict possible future blockages and apply the needed countermeasures in advance. The drawback of this mechanism is that it takes a long time to build the fingerprinting database table that also needs to be updated frequently and therefore, it is not suitable for dynamic environments. The study in [12] proposes a joint communication and sensing framework for near real-time extended reality (XR) systems.

Narrow beams blockage is a challenge in such systems, so the inherent sensing capabilities of THz frequencies are exploited to extract environmental sensing parameters and form a high-resolution indoor situational awareness map. This map is then used to localise the users, assess the beam blockages, and determine the availability of LoS links between the users and the serving units. Nevertheless, this solution is suitable for indoor environments that are relatively less complex than outdoor environments.

Subsequently, learning-based approaches have exploited various ML techniques to optimise the operations of high-frequency networks. In the interest of maintaining seamless connectivity and fulfilling the required QoS for mobile users, the work in [13] presents a proactive mobility management scheme. This work comprises a resource reservation and prediction scheme, which uses neural network (NN) to predict the average channel quality and the connected BSs for each real-time user, hence reserving the required resources and guaranteeing the users' QoS requirements. Besides, a proactive HO (PHO) scheme, which aims to avoid frequent HOs and reduce HO latency by relying on the dual connectivity (DC) mechanism. The DC allows each user to connect with more than one BS simultaneously to achieve zero HO interruption time and maintain seamless connectivity. Likewise, the study in [14] exploits the DC and deep learning (DL) algorithm to avoid service interruption during HO decisions. A long short-term memory (LSTM) model is trained to predict the user's future movement trends depending on historical trajectory information to perform efficient HOs. Despite the potential of eliminating the HO intermittency, the DC technique will add more operational complexity to the network as well as to user devices. In addition to incurring more costs, wasting network resources, and increasing energy consumption. Moreover, multiconnectivity will not avoid LoS links blockages and service disconnection. Based on dual band network operation, the authors in [15] exploit the knowledge of the sub-6GHz uplink channel to enhance the reliability of the mmWave downlink channels motivated by the spatial correlation between the two frequency bands. A DL model is trained using a tuple of sub-6GHz channel information and blocking status to determine whether the LoS link is blocked or not. However, this work would not suit realistic scenarios since it only classifies the channel status as blocked/unblocked and cannot avoid blockages in advance.

The emerging research direction of exploiting computer vision (CV) for developing wireless communications and tackling complicated problems in UDNs has gained much interest recently. The consistency between the LoS communication and the direct camera view is envisioned to play a major role in future wireless networks. For instance, the study in [16] leverages camera imagery and DL to tackle the beam blockage problem in mmWave systems. The proposed technique predicts the time series of the mmWave received power to several hundred milliseconds in advance based on the depth images of the served area, allowing for sufficient time to perform HO. However, predicting the received power in advance does not necessarily suit HO decision problems. Furthermore, it requires large quantities of training datasets and computational

resources to prepare a model for accurate prediction. The authors in [17] utilise the visual sensory information collected from the served area to train a DL model and predict beam blockages. RGB images captured by a camera installed on each BS are labelled with the beam blockage status and used to train the ResNet18 model to classify images based on blockage status. However, the proposed technique does not predict in advance, so service disconnection can not be avoided.

B. Motivation and Contribution

As discussed in the previous section, the studies mentioned above have different assumptions that limit their applicability in real practical scenarios. The vision for B5G and 6G networks is to meet the strict requirements of maintaining high levels of QoS/quality of experience (QoE) and ensuring user connectivity to realise real-time services and applications. Therefore, the newly emerging research direction of exploiting CV to enhance the performance of mmWave communication systems is envisioned to assist the operation of such systems, satisfy the stringent demands, and encourage their widespread. Furthermore, next-generation high-frequency wireless networks will be prevalent in smart cities where video surveillance systems are widely available, especially in crowded areas, the main target of deploying UDNs. Given that future BSs will be miniaturised and installed on lampposts, the combination of wireless and vision information is recognised as an attractive approach to optimise the operation of high-frequency wireless networks [17], [18]. However, merging CV in the operation of UDNs is still in its infancy and needs a lot of research devotion to getting the most out of it.

In this paper, we propose a novel CV-assisted PHO mechanism that combines two modes of information, i.e., wireless and imagery information, to predict possible beam blockages in advance and then instructs the network to perform HO in a time that maximises the overall quality of experience (QoE), hence optimising the network performance. An object detection and localisation (ODL) algorithm is adopted to analyse the RGB images from vision sensors, detect obstacles and users, and determine their location and speed. Additionally, a simple NN model is trained using the multivariate regression method to predict the remaining time until the user is being blocked by the obstacle. This work introduces a new HO event, called blocking event (BLK), defined by the existence of a blocking object and a user moving towards the blocked area. Once a BLK event is detected, the proposed algorithm determines the best time to trigger HO and switch the user to another BS, therefore improving the network's reliability. The main contributions of this paper are as follows:

- First, we present a novel solution to the problem of beam blockage and frequent HOs in next-generation wireless networks by utilising CV and NN algorithms. The CV is used to increase the network's awareness of the surrounding environment, and the NN model predicts when a sudden drop in the RSS will happen due to the existence of stationary obstacles, which is a very common challenge in high-frequency networks.
- Furthermore, we introduce a new HO event called BLK, which can be considered in B5G and 6G networks besides
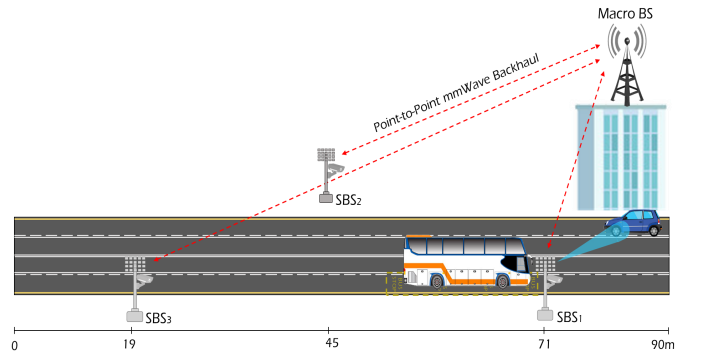

Figure 1: The proposed system model: portion of an UDN including one macro BS and three SBSs each equipped with an RGB camera.

the standardised events defined by the 3GPP [7]. The BLK event is defined by detecting the presence of an obstacle and a user moving toward the blocked area.
- To determine the best point of triggering and completing HO once the BLK event is detected, this study provides an analysis on determining the optimal HO trigger point to keep the user's QoE at a high level.
- Finally, we validate the accuracy of the proposed framework using state-of-the-art simulation tools. The results demonstrate the significance of this solution in maintaining seamless connectivity.

C. Organisation

The rest of this paper is organised as follows. Section II presents the system and channel models adopted to study the beam blockage problem. Section III describes the schematic diagram and the proposed framework in detail. In Section IV, performance evaluation and simulation results are discussed. Finally, Section V gives concluding remarks.

II. SYSTEM AND CHANNEL MODELS

In this section, we offer a detailed description of the system and channel models and discuss the scenario under study.

A. System Model

We consider a wireless communication system consisting of one macro BS and three SBSs[1] covering a 90×15m street, as illustrated in Fig. 1. The system adopts orthogonal frequency division multiplexing (OFDM) with K subcarriers and cyclic prefix of length D and operates at 60 GHz [19]. Each SBS is equipped with a mmWave uniform linear array (ULA) composed of M antenna elements that enable beamforming technology to create LoS beams that can achieve high RSS at a single-antenna user. To reduce cost and power consumption, this study assumes analog beamforming architecture with M phase shifters and a single radio frequency (RF) chain [20]. Further, to simplify the network operation, each SBS adopts a predefined beam codebook $\mathcal{F} = \{\mathbf{f}_i\}_{i=1}^{B}$, where $\mathbf{f}_i \in \mathbb{C}^{M \times 1}$

---

[1]It is important to mention that: (i) extending this proposed framework to include a larger number of SBSs is straightforward, and (ii) the three SBSs model is used for simplicity.

and B is the total number of beams in the codebook. Each beamforming vector $\mathbf{f}_i$ can be expressed as:

$$\mathbf{f}_i = \frac{1}{\sqrt{M}} \left[1 \ e^{j\frac{2\pi}{\lambda} d \sin(\theta_i)} \ \cdots \ e^{j\frac{2\pi}{\lambda}(M-1)d\sin(\theta_i)} \right]^T, \quad (1)$$

where $\frac{1}{\sqrt{M}}$ is the normalisation factor, d is the inter-element distance of the antenna array, λ is the wavelength corresponding to the carrier frequency, and $\theta_i \in \left\{ \frac{2\pi i}{B} \right\}_{i=0}^{B-1}$ is the steering angle.

To determine the best beam vector that can achieve maximum received power, the mmWave user will send a pilot message to the SBS that can be used to train the B beams in order to find the optimal beam $\mathbf{f}^\star$. Once $\mathbf{f}^\star$ is determined, the received downlink signal at the user's receiver at the kth subcarrier is given as follows:

$$y_k = \mathbf{h}_k^T \mathbf{f}^\star s_k + n_k, \quad (2)$$

where $\mathbf{h} \in \mathbb{C}^{M \times K}$ represents the mmWave channel between the SBS and the user, s is the transmitted symbol, and $n \sim \mathcal{N}\left(0, \sigma^2\right)$ is the additive white Gaussian noise (AWGN).

In addition to the ULA, each SBS has a vision sensor (an RGB camera with a standard definition (SD) resolution) that captures visual information from the covered area to assist the operation of the UDN. The vision information is transmitted to a central server located at the macro BS through 10Gbps point-to-point mmWave backhaul links [21]. The role of the central server is to collect, process, and use the visual information to train an ML model that can proactively predict possible beam blockages. This paper assumes a simple scenario including a moving user (vehicle) and a stationary blocking object (bus) that blocks the LoS communication between the SBS and the user.

B. Channel Model

This study adopts the geometric mmWave channel model. The reasons for selecting this model are: i) it captures the physical properties of the signal propagation, ii) it enables the direct use of accurate channel simulation tools, such as ray tracing, the tool selected to generate the wireless data used to validate this work. The mmWave channel model at the kth subcarrier can be written as [20]:

$$\mathbf{h}_k = \sum_{d=0}^{D-1} \sum_{\ell=1}^{L} \alpha_\ell e^{-j\frac{2\pi k}{K}d} p\left(dT_s - \tau_\ell\right) \mathbf{a}\left(\theta_\ell, \phi_\ell\right), \quad (3)$$

where L denotes the number of channel paths, α_ℓ, τ_ℓ, θ_ℓ, ϕ_ℓ are the gain, delay, azimuth and elevation angles of the arrival of path ℓ, respectively. Also, T_s represents the sampling time.

Optimal Beam and RSS: The maximum RSS value at any x location is associated with finding the optimal beamforming vector $\mathbf{f}_x^\star$ that can achieve this value. In other words, determining $\mathbf{f}_x^\star$ means obtaining the maximum RSS and vice versa. As a result, these can be mathematically expressed as:

$$\mathbf{f}_x^\star = \underset{\mathbf{f}\in\mathcal{F}}{\mathrm{argmax}} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\left[\left\| (\mathbf{h}_{k,x})^T \mathbf{f} \right\|_2^2 \right], \quad (4)$$

and

$$RSS_x = \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\left[\left\| (\mathbf{h}_{k,x})^T \mathbf{f}_x^\star \right\|_2^2 \right], \quad (5)$$

where $\mathbf{h}_{k,x}$ is the kth subcarrier's mmWave channel between the SBS and the user at the location x.

III. PROPOSED CV-ASSISTED PHO FRAMEWORK

The key idea of this work is to anticipate future beam blockage using CV and NN to perform timely PHO. Beam blockage prediction is a very challenging task because it depends on finding the location of a moving user and its possible sources of blockage in a realistic wireless scenario. In CV, ODL is used to identify an object's class and location coordinates; however, object detection alone is not enough to determine future blockages that necessitate: first, an efficient system that can detect the moving users (wireless users) and the potential source of blockage. The second is extracting augmented information, including speed, time, and distance from the blocked area. Guided by the above notions, the beam blockage prediction is divided into two sub-tasks, (i) ODL, which is used to determine the types of objects and their location to calculate their speed, (ii) Using the multivariate regression model to predict the remaining time until the user reaches the blocked area based on the information extracted from the RGB images. Before delving into a detailed discussion of the various components of the proposed framework, we highlight the assumptions made in this study as follows:

1) We assume the availability of vision sensors integrated with SBSs. Moreover, the sensors can provide flat RGB images showing that the width of the street's upper and lower pixels are the same.
2) The vision sensors are not affected by the time of the day (day/night), as well as, the weather conditions.
3) The macro BS has local processing units, where optimisation and local decision-making occur. In addition, the ODL consistently provides high performance in predicting the objects and determining bounding boxes within acceptable precision.
4) The wireless user is identified in the image, and the network correctly performs the proactive HO, if necessary, for that user.

In the worst-case scenario, if the assumptions are not met, the wireless network will lose the vision assistance in solving the beam blockage problem and revert to a simple wireless network without proactive blockage prediction.

A. Schematic Diagram of the Proposed Framework

This study aims to achieve a blockage prediction mechanism, such that the network can proactively HO the user well before it reaches the blocked area. *Once a BLK event is spotted out in the camera's field of view, the algorithm's*

main task is to predict the time needed by the user to reach the shadowed area, denoted as $\boldsymbol{T_{toBLK}}$. This time allows us to determine the best instant to perform HO before the user reaches that area and undergoes service interruption. Fig. 2 demonstrates the schematic diagram of the proposed technique. We rely on a heuristic approach since wireless networks generally serve dynamic and unpredictable environments like urban areas and smart cities. Moreover, increasing the network complexity usually leads to NP-hard optimisation problems. Such problems demand a very high computation time, which is intolerable and challenges latency-sensitive applications, such as intelligent transportation systems. Therefore, the nature of the problem we are looking at motivates the use of heuristic models [22], [23]. Multivariate regression is used to predict the T_{toBLK} by modelling and training a two-hidden layer NN. Initially, the server will build a complete view of the covered area (the street) with the exact coordinates and locations of each SBS. The RGB cameras continuously[2] capture images from the covered area, then every SBS adds its identification number and timestamp to each image before sending it to the central server through the mmWave backhaul link. Once received, the server performs the following tasks:

- First, it uses the ODL algorithm to detect blockages/users and updates its view. If a BLK event is detected, the server will move to the next step; otherwise, it will return to the detection phase.
- Then, the server identifies the user's exact location and updates its view, the user's location information and the timestamp difference between two consecutive images are used to determine the user's speed.
- After that, the location and speed information are stored for model training/retraining and used to predict the T_{toBLK}.
- Finally, if the T_{toBLK} is greater than the execution time of the proposed algorithm (T_{exec}), the server will wait for a specific time and then send a HO trigger event to the network in order to HO the user to another SBS. Otherwise, it will return to the detection phase.

The T_{exec} is defined as the time required by the proposed algorithm to be completed, starting from when the RGB images are captured until the HO process is completed. It is expressed in (6) and can be evaluated by the summation of four sub times: (i) The time required to send two consecutive RGB images to the central server (T_{RGB}), (ii) the time needed to perform ODL on the two images (T_{ODL}), (iii) regression model inference time (T_{inf}), and (iv) HO implementation time (T_{HO}).

$$T_{exec} = T_{RGB} + T_{ODL} + T_{inf} + T_{HO}. \tag{6}$$

In addition, we define a new time parameter notated as (T_w), which is the waiting time after completing the regression inference and before triggering HO. The value of T_w can vary based on defining the optimal trigger region, as will be discussed in Section III-D; however, the maximum value of

[2]Note that some cameras have a motion detection feature that can be activated to reduce the amount of vision information sent to the server [24].
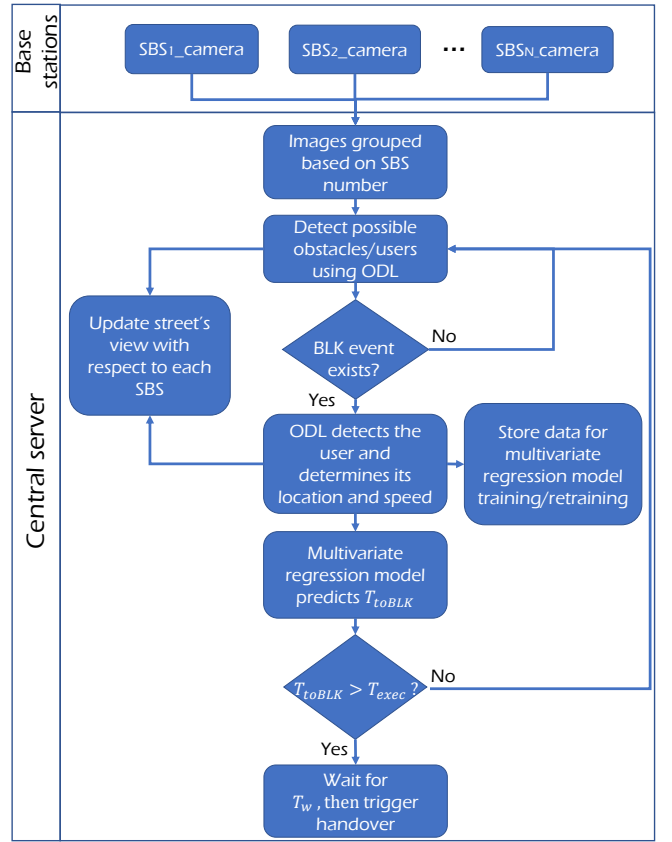

Figure 2: Schematic diagram of our proposed framework.

T_w is given as follows:

$$T_w^{max} = T_{toBLK} - T_{exec}. \tag{7}$$

It is noteworthy that the values of all parameters specified in (6) are fixed, whereas the value of T_w varies based on the user's location and speed. Determining the values of T_{RGB}, T_{ODL}, T_{inf}, and T_{HO} depends on the capacity of the mmWave backhaul links, the type of model used, and the specifications of the central server. While choosing the value of T_w is related to defining the optimal HO trigger region where the algorithm will find out the optimal time to perform HO.

B. Object Detection and Localisation (ODL)

The central server needs to receive at least two consecutive images (frames) from each SBS to detect the objects' presence, position, and speed. Table I lists the most common camera resolutions for surveillance applications and the required image transmission time over 10 Gbps mmWave backhaul links. It can be noticed from Table I that higher camera resolutions incur longer transmission times due to the production of larger image sizes. Since this study considers an SD camera resolution, T_{RGB} is equal to the transmission time of two SD images plus 38.5 ms, which is the time difference between capturing two consecutive RGB images, assuming that each camera records vision information at 26 fps [25]. Hence, the time required to transmit these images from the SBS to the server, T_{RGB}, would be about 40 ms.
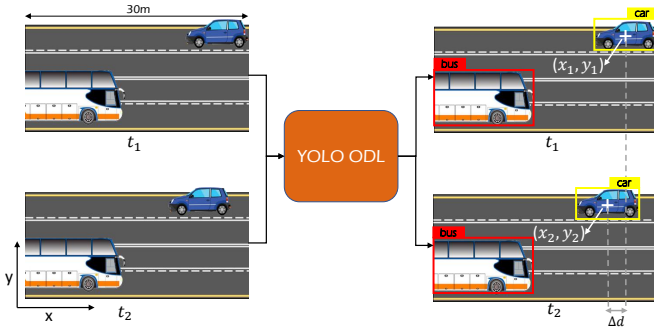
Figure 3: Using ODL to detect objects and determine their locations. This information is used to determine the speed of the moving object.

Table I: The most common camera resolutions with associated image transmission times over 10 Gbps links.

Term	Resolution	Transmission time (ms)
2CIF	704×240	0.4
SD	640×480	0.7
4CIF	704×480	0.8
HD	1280×720	2.2
FHD	1920×1080	5

Once the server receives the visual information, the first step is to process that information to obtain the location of the objects. In this work, we adopt a state-of-the-art detection model, you only look once (YOLO) version 3, which can provide fast and accurate real-time object detection [26]. Instead of developing and training an object detection model from scratch, YOLO models can be used directly without the need for any modification. Furthermore, the main objective of ODL is to identify if a BLK event exists and determine the objects' locations in the pixel scale, then this information will be converted to the meter scale to find the speed of the objects. The server will feed the RGB images to the object detection algorithm, which in its turn detects the objects within the image by drawing bounding boxes around them and adding tags showing their categories, as illustrated in Fig. 3. Moreover, this algorithm will provide the location information of the objects by determining the coordinates of the upper left and lower right corners of the bounding boxes. We use this information to determine the centre of the moving user, as shown in Fig. 3.

The location information obtained from the YOLOv3 model is in pixel scale, and to determine the user's speed, we need to convert this information into a metric scale. The cameras' field of view is set to 100 degrees to ensure that the entire street is covered with minimal overlapping. Since the cameras capture images in two dimensions only, we assume that the image is flat and its width is equal to 30m distance[3]. Therefore, the following formula can be used to determine the user's

---

[3]Determining image width and height in meter scale can be easily performed in real-world systems by capturing an image from the camera and measuring the actual image dimensions based on image corners.

displacement in metres:

$$Travelled\_distance = \frac{W_m}{W_p} \times \Delta d, \tag{8}$$

where W_m is the image width in meters, W_p is the image width in pixels, and $\Delta d = |x_1 - x_2|$ is the x-axis user displacement in the pixel scale, assuming the user is moving in a straight trajectory, as demonstrated in Fig. 3. For example, if a camera produces images with a resolution of 640×480, then every 21 pixels are approximately equal to a distance of one meter. After evaluating the travelled distance, it is necessary to determine the associated travel time, which can be easily measured by exploiting the timestamp information of each image. Since the camera's frame rate is 26 fps, the time difference between two consecutive images will be about 38.5 ms. Now, the following speed formula can be used to determine the speed of the moving user:

$$Speed = \frac{Distance}{Time}. \tag{9}$$

Since the proposed solution depends on the T_{exec} to determine if we have sufficient time to perform PHO, as demonstrated in Fig. 2, we need to find out the detection time of the YOLOv3 model (T_{ODL}). Benefiting from the high performance of multi-access edge computing (MEC) servers, the T_{ODL} can be reduced significantly to tens of ms when using the MEC server as a central server. Based on the analysis presented in [26], we assume that the T_{ODL} requires 102 ms for detecting objects in two images.

C. Multivariate Regression: Learning and Prediction

In this section, we will discuss the reasons for selecting ML techniques to predict the value of T_{toBLK} and the training and inferencing processes for the given model. Using analytical methods can fulfil the task of determining the T_{toBLK} in general. However, the main objective of this work is to propose a flexible, scalable, and transferable framework that can be employed in complex and dynamic environments. Therefore, we used a NN model for the following reasons:

1) Analytical solutions generally perform very well in static systems, where the model's assumptions do not change. This study targets high-frequency wireless networks, which are highly dynamic and unpredictable. For instance, predicting T_{toBLK} greatly depends on the cameras resolution, which varies in practical wireless systems and provides diverse data. Machine learning models learn from problem-specific data to automate the process of an analytical model and solve the associated tasks [27]. Therefore, we use neural networks to provide a more generalised and scalable solution.

2) Given the dynamicity and complexity of the targeted environment, we need a model that can scale well under these conditions. NNs are flexible and adapt to the dynamics of the problem by learning from diverse data. Moreover, the transfer learning feature allows these models to be transferable, providing our solution a new dimension to perform distributed learning.
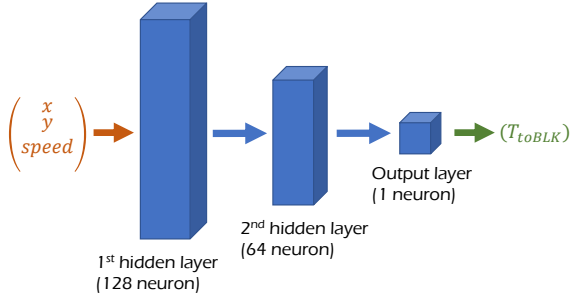
Figure 4: A two-hidden layer neural network to perform regression.

Table II: Sample of the training dataset.

x (m)	y (m)	Speed (m/s)	T_{toBLK} (ms)
68.47	9	1.5	60.00
74.21	10	5	1166.00
81.22	9	10	1284.00
71.88	10.5	12	291.67
75.92	9.5	15	502.67
81.22	10	18	713.33
86.91	11	20	926.50

Figure 5: Multivariate regression model training and validation loss versus number of epochs.

DL algorithms have achieved breakthroughs in various areas but at the expense of high computing and energy consumption. Combining CV with DL to assist the operation of UDNs will increase the models' computational complexity, rendering this fusion inefficient [28]. Since this study depends mainly on vision information, we carefully select the ML model that does not require much training time and achieves the expected results. In the previous section, we discussed using the pretrained YOLOv3 ODL model, which is off-the-shelf and can be used directly; thus, no further training is required. Moreover, to predict the T_{toBLK}, we select a simple NN model to conduct multivariate regression, as shown in Fig. 4. The multivariate regression technique is a statistical approach that measures the relationship between dependent variables (i.e., T_{toBLK}) with more than one independent variable (i.e., x, y, speed). Besides, instead of using information-rich RGB images in model training/inferencing, the proposed technique only requires extracting the user's location and speed to be fed into the NN model. This yields to significant savings in time.

Training Phase: The operations of the proposed regression model include model training and inference; both require the availability of data samples. For training the initial model, training datasets are readily generated using random (x, y) locations confined to our system model's street dimensions and using several speed values that reflect the expected vehicle speed in urban areas. The dependent variable T_{toBLK} is calculated by fixing the location of the obstacle (bus) at an arbitrary location, for example at $x = 68.38$m as illustrated in Fig. 1, and using the speed formula given in (9). Table II shows a small sample of nearly ten thousand generated data samples that are divided into 70% training,
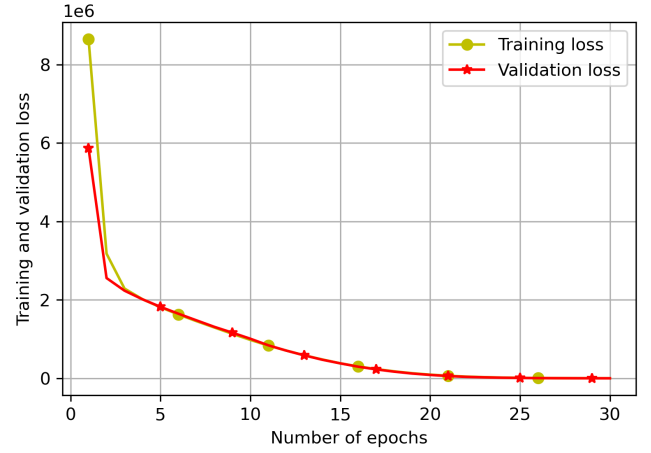
20% validation, and 10% testing. Adaptive moment estimation (Adam) is used as an optimiser in the training process. Moreover, other model hyperparameters, such as the number of epochs, batch size, metric, and activation function, are set to E=50, B=20, mean square error (MSE), and linear activation function, respectively. Fig. 5 shows the training and validation loss for the model in each epoch. This figure shows the effectiveness of this model since the training and validation loss approaches zero as the number of epochs increases. It is worth mentioning that the number of epochs depends on the number of data samples used for model training; a small dataset and a large number of epochs will lead to model overfitting, which is an unwanted behaviour for predictive modelling. The coefficient of determination (R squared) metric is adopted for model performance evaluation to measure the linear correlation between the predicted and actual values using the test dataset. The values of R squared range between 0 and 1; 1 is the optimal value that can be achieved. Using the test dataset, our model achieves 0.9998, which indicates the superb performance of this model. Finally, based on the generated dataset and the above mentioned hyperparameters, the proposed regression model took about 20 seconds to be trained using typical personal computer resources. However, using the MEC server, the training time will be less than one second [29].

Inference Phase: Using the trained regression model, our framework is now ready to predict T_{toBLK} for the user detected in images received from the cameras. After completing the street view and updating the blockage status, whenever the server receives RGB images from the SBSs, it will use the ODL to extract the location and speed of the moving user as discussed in Sec. III-B. The user's information is now ready to be fed to the input layer of the regression model to infer the remaining time until the user reaches the blocked area. Furthermore, by using the same MEC server resources, the value of T_{inf} is around 1 ms.
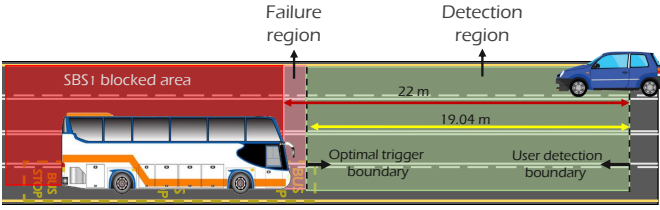
Figure 6: Optimal trigger distance for a user with a speed of 30 mph.

Table III: Optimal trigger distance based on different user speeds.

$Speed$ (mph)	T_{toBLK} (sec)	T_w (sec)	D_{opt} (m)
5	9.85	9.63	21.52
10	4.92	4.70	21.01
15	3.28	3.06	20.52
20	2.46	2.24	20.03
25	1.97	1.75	19.56
30	1.64	1.42	19.04
35	1.40	1.18	18.46

D. Optimal HO Trigger Distance

Once all the parameters like user location, speed, and T_{toBLK} are known, the final stage is to determine the optimal HO trigger location to maintain the user's QoS/QoE as high as possible. This is the optimal distance at which the central server initiates the HO request after detecting the BLK event, and performs the PHO with minimum performance degradation. In this work, we use a threshold distance-based setup, where the central server determines the optimal trigger distance using the following equation:

$$D_{opt} = S_u \times T_w^{max}, \tag{10}$$

where S_u is the speed of the user, which is already known from ODL. From equations (7) and (10), the variable trigger distance D can be inferred as:

$$D \leq S_u(T_{toBLK} - T_{exec}), \tag{11}$$

where T_{exec} is the sum of the four sub times i.e., T_{RGB} is 40 ms, T_{ODL} 102 ms and T_{inf} is 1 ms whereas the T_{HO} is 80 ms as will be discussed in the next section. The formula in (11) can be used to differentiate between early and optimal HO decisions.

To determine D_{opt}, we first need to identify the detection region, which can be defined as the region in which the proposed algorithm will monitor and detect if a wireless user is spotted in that zone. The detection region is confined within two boundaries, i.e. the user detection boundary and the optimal trigger boundary. This region is followed by the failure region located between the optimal trigger boundary and the blocked area, as shown in Fig. 6. The user detection boundary is where the proposed algorithm begins the detection process for any BLK event. On the other hand, the optimal trigger boundary, denoted by D_{opt}, is the minimum distance from the blocked area where a successful optimal HO is performed. For instance, if HO is performed before the optimal trigger boundary, the user will experience undesirable performance degradation due to the wireless channel's path loss. Whereas, if the HO is performed beyond the optimal trigger boundary, the user's LoS link will be blocked by the blocking object due to insufficient time to execute the PHO algorithm.

Therefore, the algorithm has to select the best T_w value so the user can be handed over to another SBS at the optimal point and avoid service disconnection. To successfully perform HO, the T_{exec} is found to be equals 223 ms as discussed earlier. The formula in (11) can be used to determine the best point for performing optimal HO and study the impact of performing early HOs on system performance. For extensive analysis, optimal trigger distances based on different speeds

are given in Table III. For example, if a car is detected at $x = 90$m and is moving at 30 mph, the optimal trigger distance to perform PHO is 19.04 m, as shown in Fig. 6. Once a BLK event is detected, the optimal trigger distance is calculated using (11). S_u is already known, T_{toBKL} is obtained using regression analysis, and T_{exec} is also known. It is possible to perform PHO within the detection region. However, this early HO will significantly degrade the RSS value, which is undesirable. Therefore, in our proposed model, the central server waits for T_w^{max} until the user reaches the optimal distance to complete the PHO request.

E. Proactive Handover Mechanism

T_{HO} is one of the key parameters of our algorithm to determine whether there is enough time to perform HO and avoid radio links failures. Conventionally, if the user's link is disconnected, a sequence of steps precedes reconnecting the user to the same or another SBS. The steps include beam failure detection, beam failure recovery, cell search, and contention-based/free random access [30]. Supposing the network employs proactive blockage prediction, the first two steps can be avoided, whereas the cell search can be performed while the user is still connected to the serving SBS. Therefore, T_{HO} boils down to the latency accompanied by performing contention-based or contention-free random access. This study considers contention-based random access, which requires 80 ms according to the 3GPP specifications [30], [31]. Since the values of all parameters in (6) are determined, the central server is aware of the time needed to execute the proposed algorithm, T_{exec} equals 223 ms. If the central server detects a BLK event in the received RGB images, the server will predict the time needed until the user reaches the blocked area (T_{toBLK}). If T_{toBLK} is greater than T_{exec}, then our algorithm has a high probability of successfully triggering and completing HO. Whereas when T_{toBLK} is less than T_{exec}, the time needed to complete the HO process and avoid radio link failure is insufficient, which means that the user undergoes a service interruption.

IV. PERFORMANCE EVALUATION AND RESULTS

To investigate the effectiveness of the proposed CV-based PHO framework, we utilise a publicly available dataset called vision wireless (ViWi) [32]. The ViWi dataset combines visual and wireless information generated using Wireless InSite ray-tracing software and 3D game modelling for mmWave wireless
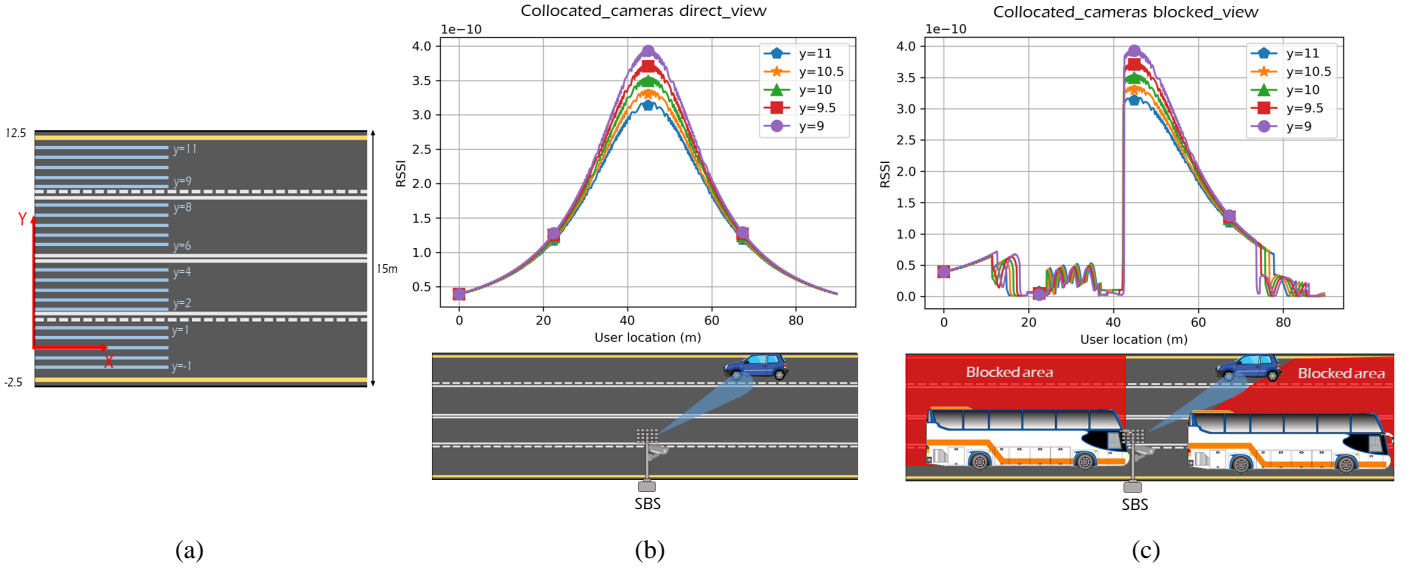
Figure 7: Analysing ViWi information to construct the system model. (a) Locating the origin of the Cartesian coordinates, (b) Using ViWi information from colocated cameras direct view scenario to model SBS_2 in our system model, and (c) ViWi information from colocated cameras blocked view scenario shows a similar RSS pattern when there is no blockage.

systems. It considers four different scenarios based on the camera location (collocated and distributed) and the view (direct and blocked). Furthermore, each data sample represents 4-tuple of user location, RGB image, depth image, and the wireless channel. The distinctive feature of ViWi is that it is a parametric, systematic, and scalable data generation framework that can be used to produce data based on different scenario requirements. In the performance evaluation, the focus is to track the strength of the signal received by the moving user and test the value of the proposed algorithm in maintaining a good received signal during movement compared to traditional mmWave systems (i.e., without PHO). It should be noted that, although the considered system is simple and does not necessarily reflect a practical scenario, the results obtained from our proposed framework (see Section IV-B) demonstrate the potential of exploiting vision information to enhance the reliability of future wireless networks. A more complex scenario will be considered in our future work.

A. Simulation Setup

In our simulation, we consider a simple environment containing a blocking object located near the SBS_1 and a single user moving at a speed of 30 mph, as illustrated in Fig. 1. The user is moving from left to right and is served from SBS_1 since the RSS from that BS is higher than other SBSs. Moreover, the obstacle is static and blocks the LoS communication between the SBS_1 and the user when it reaches the obstacle's blocked region. Our system model is different from any of the scenarios introduced with the ViWi dataset; however, we were able to produce both visual and wireless data for our model by merging the two scenarios of collocated cameras, direct and blocked view. Fig. 7 shows the analysis performed to form our system model. Initially, we analyse the information provided in the ViWi dataset to determine the place of origin of the Cartesian coordinate system, as it is not mentioned in the
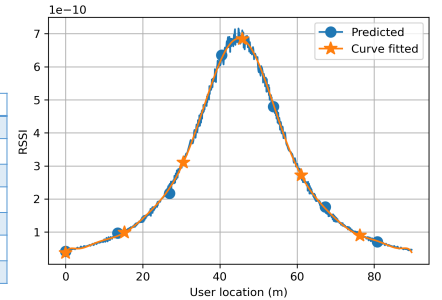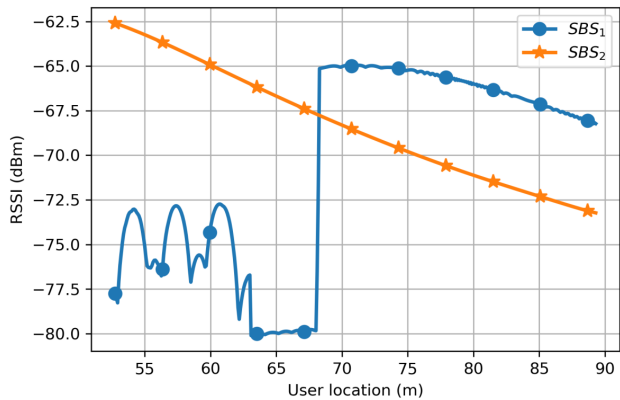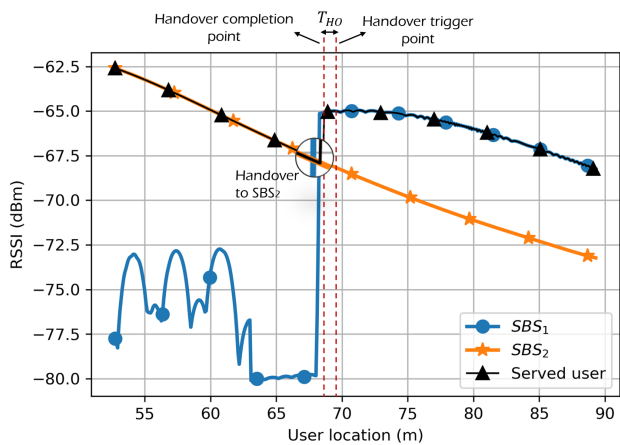

Figure 8: Determining the RSS from SBS_2 at trajectory $y=9$ using the curve fitting tool.

ViWi dataset. We used the ViWi trajectories ($y=9$ to 11) to identify the location of the origin and other trajectories, as shown in Fig. 7(a). Then, for each trajectory considered in ViWi, we plot the RSS against the vehicle location for the two scenarios, as illustrated in Fig. 7 (b) and (c). From these figures, we can conclude that each SBS has the same signal pattern (i.e. bell shape) in any trajectory with the highest signal strength at the same x-location of the SBS. To model the RSS of SBS_2, we used the information in Fig. 7(b) and applied curve fitting to generate the mathematical formula. First, we predict the RSS value for each other trajectory ($y=-1$ to $y=8$) at each user location, and then we used the curve fitting process to create the mathematical formula that gives the RSS versus vehicle location (x). Fig. 8 illustrates the generated mathematical formula and RSS from SBS_2 at $y=9$, which is the trajectory used in the evaluation. Finally, Python programs are used to conduct the simulations.

(a)

(b)

Figure 9: Performance evaluation of the proposed framework. (a) RSS from SBS_1 and SBS_2, (b) Using the CV-assisted PHO algorithm to detect BLK event and trigger PHO.

B. Simulation Results

In what follows, we examine the usefulness of the proposed algorithm in maintaining physical link connectivity and ensuring a timely and seamless transition from one SBS to another. The RSS indicator (RSSI) is used as a metric to measure the quality of the received signals from the nearby SBSs. Fig. 9(a) plots the RSSI received from SBS_1 and SBS_2 at the $y=9$ trajectory along the street. It is noticed that the signal drops from SBS_1 when the user reaches the area behind the blocking object since the beam undergoes severe attenuation. In contrast, the signal received from SBS_2 does not experience any interruptions because the LoS path is clear between the user and the SBS_2. In a traditional wireless network (i.e., does not employ the PHO algorithm), the vehicle experiences a connectivity disruption when entering the blocking area. Such interruption may lead to service drop and the need to initiate a new connection, which means more delay and poor QoE, inconsistent with the vision of 5G/6G wireless networks of providing ultra reliable and low latency communications. Fig. 9(b) demonstrates the capability of the proposed PHO algorithm in proactively predicting beam blockages. This fig-
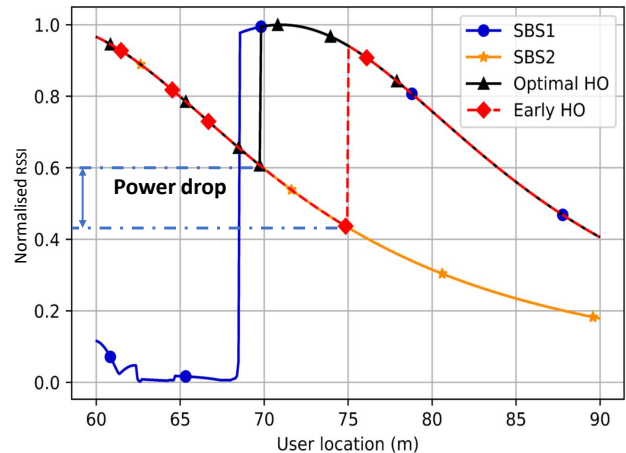

Figure 10: The normalised RSSI as function of the user location when the user speed is fixed at 30 mph.

ure reveals the efficiency of the PHO algorithm in identifying the BLK event in advance and triggering a timely HO. The merit of determining the optimal point of triggering HO in our algorithm is vital to maintain the QoE at high levels and avoid early HO, which could lead to bad system performance. The points of triggering HO and HO completion are also shown in the figure, which illustrates how our algorithm is also QoE-aware.

Fig. 10 shows the results of the early/optimal HO trigger distance for the user moving at the speed of 30 mph. The optimal trigger boundary is the minimum distance where the central server can perform the successful HO with minimum performance degradation measured in the percentage drop in normalised RSSI. In our case, the optimal distance for the user moving at a speed of 30 mph is found to be around 70 m from origin. The central server performs HO once a BLK event is detected, as a result, the resources of the user shift from SBS1 to SBS2. During the HO process, the user will experience a drop in RSSI due to path loss. For instance, if the central server performs an early HO i.e., 5 m before the optimal trigger boundary, there is a power drop of approximately 20 % as shown in Fig. 10. therefore, the optimal trigger distance provides the trade-off between the PHO success rate and the drop in RSSI to maintain the seamless connectivity.

Next, we investigate the effectiveness of the PHO algorithm in improving the reliability of high-frequency wireless networks by considering a real-time application sensitive to service interruption and network latency. We assume a moving user that is running a video call and consider the mean opinion score (MOS) as a metric, which is a measurement of the QoE. MOS is a measure of the media's overall perceived service quality based on human judgement and ranges from 1 to 5 (1-bad, 2-poor, 3-fair, 4-good, 5-excellent) [33]. Fig. 11 shows the MOS value versus the user location, which is experiencing different RSS based on the distance from the SBS and the presence of obstacles. To translate the values of RSS to the corresponding values of MOS, we adopt the mapping table in
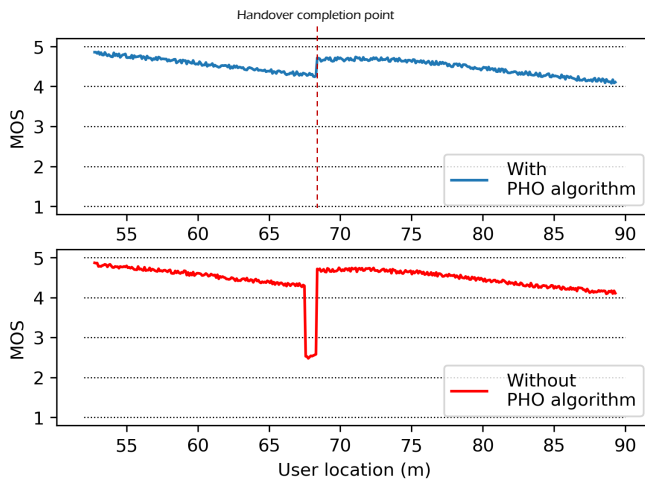
Figure 11: Measuring the QoE with and without PHO.

[34]. From Fig. 11 we notice that without the PHO algorithm, the user will experience a service interruption when it reaches the blocked area, and the call quality is dropped by 40% to the poor MOS region. Call quality will remain poor until the user drops the serving SBS link and connects to a new SBS. Whereas, using the PHO technique, our algorithm intelligently detects the existence of a blockage and countermeasures the possible signal blockage by triggering HO in advance. Therefore, maintaining the perceived MOS at the excellent region. This work enhances the reliability of UDNs that will facilitate the realisation of future latency-sensitive applications. It is worth noting that the interest in this research direction is increasing, and the recent paper [35] has considered a multi-user scenario, which will also be considered in our subsequent study.

V. CONCLUSIONS

This paper proposed a novel CV-based PHO framework to address the challenge of frequent HO and beam blockage in next-generation wireless networks. The key idea is to raise the network's awareness of the surrounding environment by leveraging visual information and using the CV to predict BLK events and perform PHO. A pretrained object detection model, in addition to a multivariate regression model, are used to predict the obstacle/user location and the time remaining before the user arrives at the blocked area. Moreover, this framework is QoE-aware, where we presented an analysis of the optimal location/time to perform HO while minimising the drop in QoE. The evaluation results demonstrated that this framework is able to avoid 40% service reduction and maintain a high level of perceived QoE. Accordingly, this work improves the performance of the UDN by making it more dynamic in interaction with the environment, which is inline with the vision of achieving low-latency and time-sensitive applications in B5G and 6G networks.

REFERENCES

[1] X. You *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inform. Sci.*, vol. 64, no. 1, pp. 1–74, Jan. 2021.

[2] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surv. Tuts.*, vol. 18, no. 4, pp. 2522–2545, May 2016.

[3] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE commun. surv. tuts.*, vol. 18, no. 2, pp. 949–973, Dec. 2015.

[4] S. Collonge, G. Zaharia, and G. E. Zein, "Influence of the human activity on wide-band characteristics of the 60 GHz indoor radio channel," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 2396–2406, Nov. 2004.

[5] A. Yamamoto *et al.*, "Path-loss prediction models for intervehicle communication at 60 GHz," *IEEE Trans. veh. technol.*, vol. 57, no. 1, pp. 65–78, Jan. 2008.

[6] G. Yu *et al.*, "A hierarchical SDN architecture for ultra-dense millimeter-wave cellular networks," *IEEE Commun. Maga.*, vol. 56, no. 6, pp. 79–85, June 2018.

[7] 3rd Generation Partnership Project (3GPP) Technical Specification Group Radio Access Network NR, "Radio Resource Control (RRC) protocol specification," *in 3GPP TS 38.331 version 16.6.0 Release 16*, Sept. 2021.

[8] J. Bao, T. Shu, and H. Li, "Handover prediction based on geometry method in mmWave communications-a sensing approach," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops), Kansas City, MO, USA*, May 2018, pp. 1–6.

[9] Z. Ali, A. Duel-Hallen, and H. Hallen, "Early warning of mmWave signal blockage and AoA transition using Sub-6 GHz observations," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 207–211, Nov. 2019.

[10] L. Yu *et al.*, "Long-range blockage prediction based on diffraction fringe characteristics for mmWave communications," *IEEE Commun. Lett.*, Apr. 2022.

[11] C.-J. Liu and L. Xiao, "Interference and blockage prediction in mmWave-enabled HetNets," in *Proc. IEEE 26th Int. Symp. Model. Anal. Simul. Comput. Telecom. Syst. (MASCOTS), Milwaukee, WI, USA*, Sept. 2018, pp. 201–208.

[12] C. Chaccour *et al.*, "Joint sensing and communication for situational awareness in wireless THz systems," *arXiv preprint arXiv:2111.14044*, Nov. 2021. [Online]. Available: http://arxiv.org/abs/2111.14044

[13] K. Qi *et al.*, "Dual connectivity-aided proactive handover and resource reservation for mobile users," *IEEE Access*, vol. 9, pp. 36 100–36 113, Feb. 2021.

[14] C. Wang *et al.*, "Deep learning-based intelligent dual connectivity for mobility management in dense network," in *Proc. IEEE 88th Vehic. Technol. Conf. (VTC-Fall), Chicago, IL, USA*, Aug. 2018, pp. 1–5.

[15] M. Alrabeiah and A. Alkhateeb, "Deep learning for mmWave beam and blockage prediction using Sub-6 GHz channels," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5504–5518, June 2020.

[16] T. Nishio *et al.*, "Proactive received power prediction using machine learning and depth images for mmWave networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2413–2427, Aug. 2019.

[17] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *Proc. IEEE 91st Vehic. Technol. Conf. (VTC2020-Spring), Antwerp, Belgium*, May 2020, pp. 1–5.

[18] A. R. Khan *et al.*, "When federated learning meets vision: An outlook on opportunities and challenges," in *Proc. EAI Int. Conf. Body Area Netw., Glasgow, UK.* Springer, Dec. 2021, pp. 308–319.

[19] X. Cheng, M. Wang, and S. Li, "Compressive sensing-based beamforming for millimeter-wave OFDM systems," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 371–386, Oct. 2016.

[20] R. W. Heath *et al.*, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE j. sel. topics sig. process.*, vol. 10, no. 3, pp. 436–453, Feb. 2016.

[21] M. M. Ahamed and S. Faruque, "5G backhaul: requirements, challenges, and emerging technologies," *Broadband Communications Networks: Recent Advances and Lessons from Practice*, vol. 43, Nov. 2018.

[22] Z. L. Fazliu *et al.*, "Mmwave beam management in urban vehicular networks," *IEEE Syst. J.*, vol. 15, no. 2, pp. 2798–2809, Jan. 2020.

[23] Y. Jian *et al.*, "WiMove: Toward infrastructure mobility in mmWave WiFi," in *Proc. 18th ACM Symp. Mobility Manage. Wireless Access*, Nov. 2020, pp. 11–20.

[24] S. Wan, S. Ding, and C. Chen, "Edge computing enabled video segmentation for real-time traffic monitoring in internet of vehicles," *Pattern Recognit.*, vol. 121, p. 108146, Jan. 2022.

[25] R. Nawaratne *et al.*, "Spatiotemporal anomaly detection using deep learning for real-time video surveillance," *IEEE Trans. Ind. Inform.*, vol. 16, no. 1, pp. 393–402, Aug. 2019.

[26] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, Apr. 2018. [Online]. Available: http://arxiv.org/abs/1804.02767

[27] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, Sept. 2021.

[28] M. Al-Quraan *et al.*, "A hybrid data manipulation approach for energy and latency-efficient vision-aided UDNs," in *Proc. IEEE Eighth Int. Conf. Softw. Defined Syst. (SDS), Gandia, Spain*, Dec. 2021, pp. 1–7.

[29] Y. Huang *et al.*, "When deep learning meets edge computing," in *Proc. IEEE 25th international conference on network protocols (ICNP), Toronto, ON, Canada*, Oct. 2017, pp. 1–2.

[30] 3rd Generation Partnership Project (3GPP) Technical Specification Group Radio Access Network NR, "UE conformance specification; Part 1: Protocol," *in 3GPP TS 38.523-1 version 15.2.0 Release 15*, Apr. 2019.

[31] J. Thota and A. Aijaz, "On performance evaluation of random access enhancements for 5G uRLLC," in *Proc. Wireless Commun. Netw. Conf. (WCNC), Marrakesh, Morocco*. IEEE, Apr. 2019, pp. 1–7.

[32] M. Alrabeiah *et al.*, "ViWi: A deep learning dataset framework for vision-aided wireless communications," in *Proc. IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium*, May 2020, pp. 1–5.

[33] G. Cermak, M. Pinson, and S. Wolf, "The relationship among video quality, screen resolution, and bit rate," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 258–262, Mar. 2011.

[34] I.-H. Mkwawa, E. Jammeh, and L. Sun, "Mapping of received signal strength indicator to QoE in VOIP applications over WLAN," in *Proc. 2012 Fourth Int. Workshop on Qual. Multimedia Experience*, July 2012, pp. 156–157.

[35] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-aided 6G wireless communications: Blockage prediction and proactive handoff," *IEEE Trans. Vehic. Technol.*, vol. 70, no. 10, pp. 10193–10208, Aug. 2021.