applied sciences

MDPI

*Article*

# Optimized Two-Tier Caching with Hybrid Millimeter-Wave and Microwave Communications for 6G Networks

**Muhammad Sheraz [1], Teong Chee Chuah [1,*], Mardeni Bin Roslee [1], Manzoor Ahmed [2], Amjad Iqbal [3] and Ala'a Al-Habashna [3,4]**

[1] Centre for Wireless Technology, Faculty of Engineering, Multimedia University, Cyberjaya 63100, Malaysia; engr.msheraz@gmail.com (M.S.); mardeni.roslee@mmu.edu.my (M.B.R.)
[2] School of Computer and Information Science, The Institute for AI Industrial Technology Research, Hubei Engineering University, Xiaogan 432000, China; manzoor.achakzai@gmail.com
[3] Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada; amjad.iqbal68a@gmail.com (A.I.); alaaalhabashna@cmail.carleton.ca (A.A.-H.)
[4] School of Computing and Informatics, Al Hussein Technical University, Amman 11831, Jordan; alaa.alhabashna@htu.edu.jo
* Correspondence: tcchuah@mmu.edu.my

**Abstract:** Data caching is a promising technique to alleviate the data traffic burden from the backhaul and minimize data access delay. However, the cache capacity constraint poses a significant challenge to obtaining content through the cache resource that degrades the caching performance. In this paper, we propose a novel two-tier caching mechanism for data caching on mobile user equipment (UE) and the small base station (SBS) level in ultra-dense 6G heterogeneous networks for reducing data access failure via cache resources. The two-tier caching enables users to retrieve their desired content from cache resources through device-to-device (D2D) communications from neighboring users or the serving SBS. The cache-enabled UE exploits millimeter-wave (mmWave)-based D2D communications, utilizing line-of-sight (LoS) links for high-speed data transmission to content-demanding mobile UE within a limited connection time. In the event of D2D communication failures, a dual-mode hybrid system, combining mmWave and microwave μWave technologies, is utilized to ensure effective data transmission between the SBS and UE to fulfill users' data demands. In the proposed framework. the data transmission speed is optimized through mmWave signals in line-of-sight (LoS) conditions. In non-LoS scenarios, the system switches to μWave mode for obstacle-penetrating signal transmission. Subsequently, we propose a reinforcement learning (RL) approach to optimize cache decisions through the approximation of the Q action-value function. The proposed technique undergoes iterative learning, adapting to dynamic network conditions to enhance the content placement policy and minimize delay. Extensive simulations demonstrate the efficiency of our proposed approach in significantly reducing network delay compared with benchmark schemes.

**Keywords:** caching; delay; D2D; mmWave; microwave; RL; time-varying popularity



**Citation:** Sheraz, M.; Chuah, T.C.; Roslee, M.B.; Ahmed, M.; Iqbal, A.; Al-Habashna, A. Optimized Two-Tier Caching with Hybrid Millimeter-Wave and Microwave Communications for 6G Networks. *Appl. Sci.* **2024**, *14*, 2589. https://doi.org/10.3390/app14062589

Academic Editor: John Xiupu Zhang

Received: 31 January 2024
Revised: 1 March 2024
Accepted: 5 March 2024
Published: 20 March 2024



## 1. Introduction

The proliferation of mobile devices, social networking, high-definition data streaming, and massive device connectivity is straining cellular network capacity. In 6G communication systems, the hyper-dense deployment of small base stations (SBSs) is anticipated to reduce the distance between the serving SBS and user equipment (UE). This approach promises substantial improvement in spectral efficiency and network capacity. Traditionally, microwave (μWave) frequencies have been utilized; however, they frequently experience congestion, imposing limitations on spectrum resources. Therefore, the conventional system will not be able to efficiently deal with the growing demand of the higher data rate. It is known that data rate and data transmission probability are correlated with the throughput [1]. To handle the surging data demands, millimeter-wave (mmWave)

is a de facto technology that is envisaged for future cellular networks [2]. For mmWave communications, high-gain directive antennas are employed to achieve data rates of up to 7 Gbit/s, and this is anticipated to grow up to 40 Gbit/s. Given the abundance of spectra in the mmWave band, which are capable of providing multi-gigabit data services, optimizing communication over mmWave bands has become a prominent research area [3]. To handle an enormous amount of data traffic, the dense deployment of SBSs, exploitation of huge mmWave bandwidth, and edge data caching at SBSs' level and the UE's level are foresighted in 6G networks. The mmWave communications operating at high carrier frequencies experience more path loss than at low carrier frequencies. For instance, at the 60 GHz band, mmWave experiences 28 dB more free space path loss than at 2.4 GHz [4]. To overcome the higher level of channel attenuation, antenna gain is improved by employing beamforming through directive antennas at both the transmitter and receiver ends [5–7]. The beam steering feature of the mmWave communication can greatly reduce path loss and mitigate the interference footprint. mmWave also helps to reduce the antenna size due to its shorter wavelength. Despite mmWave communication providing interference elimination, it is more susceptible to non-line-of-sight (NLoS), penetration loss, and blockages than μWave frequencies [8,9].

One of the noteworthy architectural boosts in future cellular networks will be device-to-device (D2D) communications, which enable UE to establish a direct connection with other UE in proximity without traversing through the SBS [10,11]. As a significant portion of data traffic is generated and consumed locally, D2D can improve spectral efficiency tremendously. D2D communications are a promising approach to lower signaling overhead, improve energy efficiency, enhance the QoS, and reduce backhaul traffic. The UE's data demands trend toward a few popular pieces of content that are requested repeatedly [12]. Therefore, edge caching is a promising technique where SBS and UE storage can function as data cache resources. Thus, when content is requested by a user, the request can be fulfilled through its own cache or a user in proximity via a D2D link [13,14]. Nevertheless, if the desired content is not available in the cache resources of the neighboring UE within the D2D range, then the desired content is obtained from the serving SBS without imposing a data traffic burden on the backhaul link.

The cache resources are limited, which poses a significant challenge to optimizing cache decisions. The caching process consists of two phases. In the first phase, data is cached at the network edge. In the second phase, the UE's data demands can be fulfilled through the cache resources if the requested content is available in the cache. The ability to access data via cache resources can provide robustness and flexibility in data access [15,16]. Moreover, caching content close to the users offers dual benefits. Firstly, the frequently requested content can be obtained through the closely located cache resources. Secondly, caching mitigates the network congestion by eliminating redundant data transmission. This can substantially alleviate the data traffic burden from the backhaul link. Thus, the data caching can be performed during off-peak hours and the UE can fulfill its data demands with less delay during peak hours. This approach reduces network delay by up to 35% because data is accessed through the closely located cache resources rather than the core network [17,18]. In one of the preliminary works [19], the concept of femtocaching is proposed, where data placement at a SBS is performed to minimize network delay. In a realistic urban scenario, UE is densely available. The UE's density can be exploited to enlarge the amount of cached content and boost the data sharing among proximal sets of UE through D2D communications. Hence, two-tier data caching at the SBS and UE levels appears to be a promising approach to significantly reduce network delay.

Moreover, UE is mobile and its position evolves with respect to time. Therefore, UE exhibits intermittent connectivity with the data source. This poses a significant challenge to deliver content within a limited period of time. The UE's transitions from one location to another location are modeled through Markov chain, and the UE's mobility pattern is random, therefore its connectivity follows Poisson distribution [20,21]. However, the UE's mobility can be utilized as an advantage because its mobility can increase the likelihood

of discovering desired content within the D2D communication range. This rise in the likelihood is due to the fact that as the UE's connectivity with other UE varies, there is an increase in the probability of encountering UE containing the desired content. Hence, the inclusion of the UE's mobility features, such as the sojourn time and the UE's contact and inter-contact intervals, is imperative in designing an optimal caching mechanism.

### 1.1. Related Works

There are features such as signal propagation and penetration that differentiate mmWave from μWave networks due to hardware and channel characteristics. Since both the mmWave and μWave have their own benefits, mmWave and μWave cellular networks are anticipated to coexist within cellular networks. In [22], a heterogeneous network (HetNet) environment containing both the mmWave and μWave SBSs is considered. The authors maximized data access via cache resources and network success probability considering the data popularity profile and data sizes. In [23], the authors proposed a collaborative edge resource model where dual connectivity is exploited through the deployment of μWave and mmWave base stations empowered by edge capability. The authors focused on joint caching and computation optimization to improve the reliability. Then, a best-first branch-and-bound algorithm was devised to achieve high reliability. The obtained results demonstrate a satisfactory virtual reality delivery.

In [24], the authors focused on cost minimization by integrating data caching and switching between mmWave and μwave signals to improve network coverage. The complexity of cost calculations was reduced by introducing an approximation approach and the results were validated numerically. The obtained results demonstrate a substantial reduction in cost. In the same vein, the authors in [25] considered evolving the data popularity profile and modeled this through the Markov decision process (MDP). The authors deployed a mmWave link for the fronthaul in order to achieve a higher capacity. Moreover, the μWave backhaul link was used to control the network performance degradation due to blockage. The authors proposed a dynamic programming technique to optimize caching in order to minimize the cost. The results demonstrate a substantial improvement in caching and data fetching decisions while adapting to time-varying data popularity. In [26], the authors minimized the average long-term cost by improving caching decisions to enhance content fetching via cache. The authors exploited spectrum switching between μWave and mmWave and introduced a chain point detection-assisted reinforcement learning (RL) approach to observe an evolving environment in order to optimize learning and the caching policy accordingly.

In [27], the authors investigated content pushing in HetNets that are equipped with mmWave hotspots. The content can be cached by UE when it is in contact with the mmWave hotspot. The authors proposed a unified policy to optimize throughput. The obtained results demonstrate a substantial gain in throughput due to the increased content access via the cache resources. In [28], the authors focused on the optimization of cache hits in a D2D environment operating on mmWave signals. The authors proposed a group caching scheme to optimize the objective function. Moreover, the hard-cored-based caching algorithm is designed to avoid redundant caching. The obtained results demonstrate substantial improvement in cache access. In [29], the authors introduced a novel user association-based caching policy where highly popular content is cached on mmWave BSs. Moreover, μWave BSs store a portion of the highly popular content that is stored at mmWave BSs and the content that is not stored by mmWave BSs. The results demonstrate that network performance can be optimized by enhancing the cooperation size of μWave BSs and array size of mmWave BSs. Moreover, the proposed caching approach demonstrates a substantial benefit in data access success probability.

### 1.2. Contributions

Motivated by the potentials of data caching and dual-mode data transmission, we propose a two-tier caching mechanism with hybrid mmWave and μWave communications

to minimize delay. The data caching is executed at both the SBS and UE levels. Moreover, dual-mode transmission is supported between serving SBS and UE. Additionally, mmWave-based D2D communications are utilized to achieve high-speed data services between proximal sets of UE through an LoS link. Furthermore, RL is employed to accurately determine optimal cache decisions while taking into account the limited size of caches, channel conditions, and UE's connectivity patterns. The main contributions of our work are as follows:

- We propose a novel data caching approach for both the SBS and UE levels. We consider a HetNet environment, where SBSs are equipped with dual-mode transmission facilities. In addition, D2D communications are provisioned through mmWave links.
- We model the data placement problem as an MDP. The RL is utilized to perform cache decisions considering the UE's evolving data demanding patterns, cache sizes, and mobility. The network delay assessment takes into account dual-mode data transmission between the SBS and UE. This includes leveraging μWave or mmWave data transmission between the serving SBS and UE. Moreover, to further improve communication capabilities, D2D communication through mmWave transmission is introduced between two sets of UE.
- Experimental results demonstrate a high performance gain in our proposed caching approach in comparison with the state-of-the-art caching approaches under various system parameters.

The remaining paper is organized as follows. Section 2 introduces the system model and problem formulation. In Section 3, an RL-based data caching decision algorithm is proposed to deal with the delay minimization problem. In Section 4, the simulation results are analyzed and a detailed discussion is provided. Finally, in Section 5, the conclusion is provided. Table 1 provides the key notations.

**Table 1.** Notations.

| Symbol | Description |
|---|---|
| $B$ | Bandwidth |
| $I$ | Number of small base stations (SBSs) |
| $D_c$ | Content access delay from core network |
| $\|f_m\|$ | Content size |
| $D_b$ | Delay between SBS and UD |
| $T_D$ | Delay threshold |
| $D'$ | Delay within internet |
| $P_{in}$ | Interference power |
| $N_o$ | Power spectral density |
| $c_i$ | SBS cache capacity |
| $\|h_{i,n}\|^2$ | Second-order statistics of Rayleigh fading channel |
| $\Gamma$ | Signal-to-interference-plus-noise ratio (SINR) |
| $\gamma$ | Skewness of popularity distribution |
| $M$ | Total content |
| $R_T$ | Transmission rate |
| $P_T$ | Transmit power |
| $N$ | Number of sets of UE |

## 2. System Model

We consider a HetNet environment that is illustrated in Figure 1. The network consists of $I$ SBSs that are distributed by utilizing the homogeneous Poisson point process (PPP)

with intensity $\lambda_{SBS}$. A dual-mode transmission system is exploited for SBS transmission. The SBSs are equipped with both μWave and mmWave front-ends to improve data services through the cellular links [30]. Furthermore, the network contains $N$ mobile sets of UE following the independent homogeneous PPP with density $\lambda_u$. The UE sets are represented by the set $\mathcal{U} = \{1, 2, \ldots, N\}$. We consider that UE sets can perform data sharing among themselves via D2D communications. This feature is helpful in alleviating the data traffic burden over the fronthaul. The D2D communications occur on the mmWave spectrum in order to ensure fast data transmission within a limited time period. However, mmWave frequencies are susceptible to path loss and blockages. Therefore, it is assumed that D2D links can be established if any of the UE within D2D range possesses the desired content and there exists a LoS link between the data sender and receiver. We consider a statistical model for both the LoS and NLoS links. The LoS link probability between the content sender and content receiver is $e^{-\varpi r}$, where $r$ denotes the length of the LoS link and $\varpi$ denotes the blockage density. Moreover, the NLoS link probability is $1 - e^{-\varpi r}$ [31].

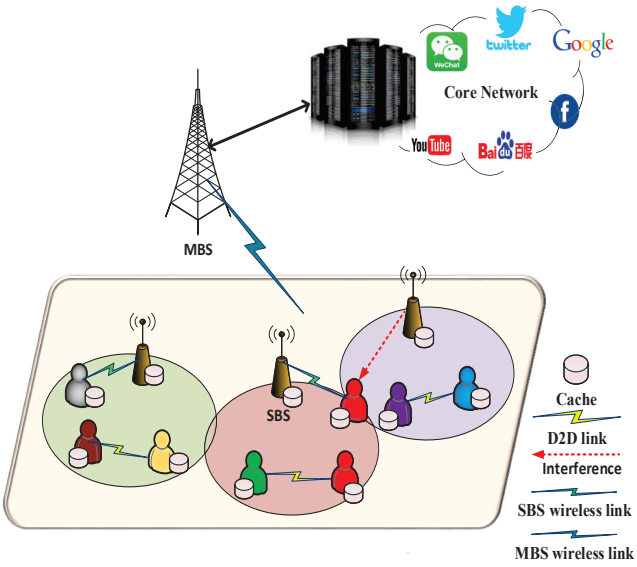


**Figure 1.** Network scenario. We consider SBSs and UE to be equipped with limited cache resources. Data caching at UE enhances data sharing via D2D communications. Moreover, SBS level caching is also performed to facilitate UE that cannot find its desired content in its own cache or the cache of proximal sets of UE. The concurrent caching at UE and SBS levels substantially minimizes network delay.

Before going into the details of data popularity profile, beamforming, and communication link modeling, the data access procedure is summarized as follows:

1. Self data offloading: First, the user checks its own cache for the desired content. The request is fulfilled with immediate effect through self data offloading subject to the availability of required content in its own cache.
2. D2D data offloading: If the desired content is not cached locally, then neighboring UE sets within proximity or D2D communication range ($R_{D2D}$) are approached. If the requested content is cached by at least one UE, a D2D communication link is established, and the requested content is obtained.
3. SBS data offloading: In case of data access failure through UE caches, the data is demanded from serving SBS. Then, the serving SBS will provide the content to the content demanding UE via mmWave or μWave transmission depending on the link probability from its own cache or by fetching from the core network.

### 2.1. Data Popularity Profile Modeling

We consider a data library comprising $M$ content represented by the set $\mathcal{M} = \{1, 2, \ldots, M\}$. For homogeneity, we consider each content size is equal to $f_m$. The cache capacity of SBS and a UE is considered to be equal to $c_i < M$ and $c_n < M$, respectively. Since the cache resources of serving SBS and UE are limited, it is necessary to determine the popularity of content because data demands are trended accordingly [32]. The data popularity profile follows a Zipf distribution, where a content $m \in \mathcal{M}$ request frequency is inversely proportional to its rank that is evaluated as follows:

$$p_m = \frac{m^{-\gamma}}{\sum_{m=1}^{M} m^{-\gamma}}, \tag{1}$$

where $\gamma$ represents the skewness of popularity distribution that controls the content popularity. If $\gamma$ is large, less content will have high popularity, leading to the concentration of popularity weights on only a few pieces of content.

### 2.2. Beamforming Model

The mmWave spectrum provides substantial bandwidth, which can facilitate fast data transmission. Nevertheless, mmWave signals are characterized by shorter wavelengths that are highly prone to path loss and attenuation. To address this issue, we leverage beamforming at both the transmitter and receiver ends to enhance signal quality and mitigate noise, path loss, and interference. Considering mmWave path loss effects, highly directive antenna beams are considered to enhance data transmission rate and coverage [33]. In this work, we adopt the IEEE 802.15.3c reference antenna model, characterized by a main lobe of Gaussian form and constant level side lobes [34]. The directional antenna gain is expressed by

$$G(\theta) = \begin{cases} G_m - 3.01 \cdot \left(\frac{2\theta}{\omega_m}\right)^2, & 0^\circ \leq |\theta| \leq \frac{\theta_m}{2}, \\ G_s, & \frac{\theta_m}{2} \leq \theta \leq 180^\circ, \end{cases} \tag{2}$$

where $G_m$ and $G_s$ represent the maximum antenna and side lobe gains, respectively. $\theta$ is an arbitrary angle of the main lobe center having a range of [0°,180°], $\omega_m$ represents the angle of half-power beamwidth, and $\theta_m$ denotes the angle of fist null beamwidth, $\theta_m = 2.6 \cdot \omega_m$, $G_m = 10 \log \left( \frac{1.6162}{\sin\left(\frac{\omega_m}{2}\right)} \right)^2$, and $G_s = -0.4111 \cdot \ln(\omega_m) - 10.579$.

### 2.3. Communication Model

mmWave signals are highly susceptible to blockages in a dense urban environment [35–37]. Therefore, for mmWave communications, we consider only LoS data transmission link for both the SBS and D2D modes [38]. Conversely, for µWave communication, we leverage an omni-directional antenna for data transmission because blockages have negligible effect over µWave frequency due to low penetration losses. For consistency in data transmission links (i.e., mmWave cellular, µWave cellular, and mmWave D2D), we assume that the content-requesting user always prioritizes the lowest path loss link. This ensures the establishment of a connection with the content transmitter, whether it is an SBS or proximity UE within D2D range.

#### 2.3.1. Data Rate Characterization of the Cellular Access Link

At 28 GHz, mmWave signal propagation exhibits minimal multipath effects owing to directive antennas. The communication link closely matches an additive white Gaussian noise (AWGN) channel [39]. Consequently, we do not consider multipath effect in mmWave communications. Furthermore, the interference in case of mmWave data transmission link occurs due to transmitters having LoS paths to the data-receiving UE. Thus, in case of downlink, interference is exhibited by the neighboring cells. We derive the signal-to-

interference-plus-noise-ratio (SINR) at UE $n$ for mmWave-based cellular access mode in the following expression:

$$\Gamma_{mm} = \frac{P_i G_t(i,n) G_r(i,n) d_{i,n}^{-\delta}}{\sigma^2 + \underbrace{\sum_{j \neq i} P_j G_t(j,n) G_r(j,n) d_{j,n}^{-\delta}}_{I}}, \tag{3}$$

where $P_i$ represents the transmission power of SBS $i$. Both $G_t(i,n)$ and $G_r(i,n)$ represent the directional antenna gains of data-transmitting SBS $i$ and data-receiving UE $n$, respectively. $d_{i,n}$ represents the distance from SBS $i$ to UE $n$ and $\delta$ represents the path loss exponent. $\sigma^2$ denotes the AWGN power density, and $I$ denotes the inter-cell interference from all the other SBSs that are not associated with the UE $n$. Similarly, both $G_t(j,n)$ and $G_r(j,n)$ represent the directional antenna gain of the interfering SBS $j$ and data-receiving UE $n$, respectively. The transmission rate provided by SBS $i$ operating in mmWave mode is computed as

$$R_{mm} = W \log_2(1 + \Gamma_{mm}). \tag{4}$$

SBS and UE are assumed to be equipped with steerable directive antennas and perfect training is performed to accomplish successful data transmission. However, mmWave provides high bandwidth, so it is vulnerable to high propagation loss in case of NLoS communication. In situations where mmWave access connectivity is unavailable, data transmission is performed through μWave spectrum, leveraging its negligible blockage and penetration losses. When UE $n$ is served by μWave band, the SINR is determined by

$$\Gamma_{\mu} = \frac{P_i |h_{i,n}|^2 d_{i,n}^{-\delta}}{\sigma^2 + \sum_{j \neq i} P_j |h_{j,n}|^2 d_{j,n}^{-\delta}}. \tag{5}$$

where $|h_{i,n}|^2$ denotes the channel fading coefficient from serving SBS $i$ to receiving UE $n$. $|h_{j,n}|^2$ represents the channel fading coefficient from interfering SBS $j$ to UE $n$. The transmission rate is derived as

$$R_{\mu} = W \log_2(1 + \Gamma_{\mu}). \tag{6}$$

#### 2.3.2. Data Rate Characterization of the D2D Link

We also perform data caching at the UE level to increase data traffic offloading through the cache resources. This enables UE to obtain the desired content from proximal UE through D2D communications, given the availability of the requested content at the proximal UE and the availability of LoS link. Moreover, for D2D connectivity, both sets of UE should have LoS for mmWave data transmission. During D2D communications, the data-receiving UE experience interference from other D2D transmitters and cellular uplink UE that are operating at the same frequency. Thus, the transmission rate is evaluated as

$$R_{v,n} = W \log_2 \left( 1 + \frac{P_v G_t(v,n) G_r(v,n) d_{v,n}^{-\delta}}{\sigma^2 + \sum_{k \in \mathcal{U} \setminus \{v\}} P_k G_t(k,n) G_r(k,n) d_{k,n}^{-\delta}} \right), \tag{7}$$

where $P_v$ is the transmission power of data-sending UE $v$. $G_t(v,n)$ and $G_r(v,n)$ denote the directive antenna gains of the transmitter and receiver, respectively. $d_{v,n}$ represents the distance between data-sending UE $v$ and data-receiving UE $n$, $\delta$ represents the path loss exponent, and $\sigma^2$ denotes the white Gaussian noise. $G_t(k,n)$ and $G_r(k,n)$ designate the directional antenna gains of interfering UE $k$ and data-receiving UE $n$, respectively. $d_{k,n}$ represents the distance between interfering UE $k$ and data-receiving UE $n$, respectively.

### 2.4. Problem Formulation

In this subsection, we formulate our delay objective function. Whenever a user needs content, it first checks its own cache resource. If the desired content is available, the request is fulfilled without delay. If the desired content is not available in user's own cache, then it requests the desired content from the proximal UE within D2D range. If the desired content is available in the cache of a neighboring UE, it can be provided to the content demanding UE by leveraging mmWave-based D2D communication link. The number of sets of UE lying within the D2D communication range of content demanding UE $n$ is represented by $\mathrm{Y}_n$, which is expressed as

$$\mathrm{Y}_n = \{v \in \mathcal{U} : \|n - v\| \leq R_{D2D}\}, \tag{8}$$

where $\|\cdot\|$ denotes the Euclidean distance between content sender UE $v$ and content receiver UE $n$, and $R_{D2D}$ denotes the D2D communication range. But, if the desired content is also not available in the cache of neighboring UE, then the UE $n$ sends its request to the serving SBS. If the demanded content is cached by the serving SBS, the request is fulfilled immediately; otherwise, the content is fetched from the core network by the serving SBS and forwarded to the data-requesting UE. In this work, we focus on a two-tier caching mechanism. We define a binary decision variable $e_n^m$. If UE $n$ caches content $m$, $e_n^m = 1$; otherwise, it equals 0. We define another decision variable $y_n^m = \cup_{v \in \mathrm{Y}_n} e_v^m$ to determine whether the desired content $m$ of UE $n$ is available or not in the neighborhood or adjacent content set. The data request probability of UE $n$ is defined by $p(m|n)$, which follows $p_m$ as determined in Equation (1). Hence, the probability of UE $n$ obtaining content $m$ via a D2D link is determined by

$$\chi_m = \frac{1}{N} \sum_{n=1}^{N} (1 - y_n^m) p_m. \tag{9}$$

The delay of D2D communication is determined as follows:

$$D_{D2D} = \sum_{m=1}^{M} (p_m - \chi_m) d_{D2D}, \tag{10}$$

where $d_{D2D}$ denotes the data delivery delay between two sets of UE exploiting D2D communication.

If the content cannot be obtained within a certain delay threshold $T_D$, the request is sent to the serving SBS. When the demanded content is cached by the serving SBS, UE can then download the content via mmWave- or μWave-based cellular access link rates denoted by $R_{mm}$ or $R_{\mu}$, respectively. The total time slots required to successfully receive content $m$ is evaluated as

$$T = \arg\min \left\{ T : R_T \geq \frac{|f_m|}{\Delta t} \right\}, \tag{11}$$

where $\Delta t$ represents the duration of one complete time slot. For ease of comprehension, the transmission rate through mmWave or μWave transmission link is denoted by $R_T$. Thus, the serving SBS containing the desired content incurs a certain delay while transmitting data to the UE, which is evaluated as

$$D_i = \frac{|f_m|}{\mathbb{E}[R_T]}, \tag{12}$$

where $R_T$ denotes the transmission rate (either in mmWave or μWave mode) between data transmitter and receiver. In order to obtain the delay of obtaining content $m$ from the core network, we designate it with $D_c$, which is derived as

$$D_c = D' + \frac{|f_m|}{\mathbb{E}[R_T]}, \tag{13}$$

where $D'$ represents the content delivery delay incurred within the internet, and its value depends on the network congestion.

Subsequently, a decision variable $x_{im}$ is defined to determine availability of content $m$ in the cache of SBS $i$. If $x_{im} = 1$, it means the desired content $m$ is available in the cache of SBS $i$; otherwise, it equals 0. Hence, the data delivery delay is evaluated as

$$\bar{D} = D_{D2D} + \sum_{m=1}^{M} \tau_m (x_{im} D_i + (1 - x_{im}) D_c), \tag{14}$$

Our main objective is to minimize the network delay subject to the cache capacity constraints of both the serving SBS and UE. Hence, our problem is formulated as follows:

$$\min_{\{x_{im}\},\{e_n^m\}} \quad \bar{D} \tag{15a}$$

$$s.t. \sum_{m=1}^{M} |f_m| x_{im} \leq c_i, \quad \forall i \in [1, I], \tag{15b}$$

$$\sum_{m=1}^{M} |f_m| e_n^m \leq c_n, \quad \forall n \in [1, N], \tag{15c}$$

$$x_{im} \in [0, 1], \quad \forall i \in [1, I], m \in [1, M], \tag{15d}$$

$$e_n^m \in [0, 1], \quad \forall n \in [1, N], m \in [1, M], \tag{15e}$$

where the amount of content cached in serving SBS $i$ and UE $n$ is constrained by Equation (15b) and (15c), respectively. Our problem is the mixed-integer programming problem, which is NP-hard [19]. It is a cumbersome task to find an optimal caching strategy to address the problem. Therefore, we utilize an RL approach to optimize the objective function.

## 3. Proposed Algorithm

In this section, we propose an RL algorithm to optimize the objective function considering the evolving network dynamics and the complexities of multidimensional action. First, the data placement problem is modeled through an MDP, then Q-learning, a model-free RL approach, is exploited to determine the action selection policy for optimizing caching decisions. The Q-learning approach optimizes caching without any prior knowledge about the underlying dynamics of the system. This makes Q-learning suitable to optimize caching through learning and adjusting strategies based on the obtained rewards in the dynamic environment. Therefore, we define a tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{R}(s, a)\}$.

- State: The data requests for content $m \in \mathcal{M}$ at any given time $t$ are represented by $v_{it}^m$ and the availability of content $m \in \mathcal{M}$ in the cache is represented by $x_{cm}$. Thus, the system state at any given time $t$ is expressed by $s_{ct}^m = (v_t^m, x_{ct}^m)$. $v_t^m$ and $x_{ct}^m$ are binary variables that can take a value of either 0 or 1. If $v_t^m = 1$, it means the request for content $m$ is generated at time $t$; otherwise, its value is equal to 0. Moreover, if $x_{ct}^m = 1$, then the requested content $m$ is available in the cache; otherwise, its value equals 0.
- Action: The data transmission of content $m$ by cache $c$ is represented by $r_{ct}^m$, and the caching of content $m$ in cache $c$ is represented by $h_{ct}^m$. Both $r_{ct}^m$ and $h_{ct}^m$ are binary variables that are either 0 or 1. If $r_{ct}^m = 1$, it means the requested content $m$ is sent to the demanding UE by cache $c$; otherwise, it equals 0. Similarly, $h_{ct}^m = 1$ if content $m$ is placed in the cache $c$ at time $t$, else it will be equal to 0. Thus, the action space at any given time $t$ is expressed as $a_{ct}^m = (r_{ct}^m, h_{ct}^m)$.
- Reward: A reward function $R_{ct}^m = R(s_{ct}^m, a_{ct}^m)$ is defined that is based on performing each action $a_{ct}^m \in \mathcal{A}$ in state $s_{ct}^m$.

Our objective is to increase the rewards associated with optimizing network performance through optimal data placement. We propose a mechanism aimed to improve the rewards in terms of minimizing delay in the following manner:

- When a cache $c$ receives requests for content $m$, the cache state $s_{it}^{m} \in \mathcal{S}$ is checked to determine whether the content $m$ is cached or not during any given time $t$.
- After observing the cache state, an action $a_{ct}^{m}$ is executed, leading to a transition to the next state $s_{c(t+1)}^{m} \in \mathcal{S}$. A reward $R_{ct}^{m}$ is also achieved.
- This complete procedure is repeated.

To maximize the reward, it is imperative to pick the optimal action from $\mathcal{A}$, which is a function $\varpi^{*} : \mathcal{S} \rightarrow \mathcal{A}$ defined to map a state $s \in \mathcal{S}$ to an action $a \in \mathcal{A}$. Hence, $\varpi^{*}(\cdot)$ determines which action $a_{ct}^{m} \in \mathcal{A}$ is beneficial to be executed during the state $s_{ct}^{m}$ in order to increase the reward $R_{ct}^{m}$. A state-value function $V$ is defined to evaluate the data placement efficiency as follows:

$$V = \sum_{t=0}^{\infty} \alpha^{t} R_{ct}^{m}, \tag{16}$$

where $\alpha \in (0, 1)$ denotes the discount parameter, which controls the long-term rewards resulting from the current actions to capture the system dynamics. We exploit Q-learning to determine an optimal $\varpi^{*}(\cdot)$ to improve the accuracy of $V_c$, which is determined as follows:

$$V_c = \sum_{t=0}^{\infty} \alpha^{t} R_{ct}^{m}. \tag{17}$$

To optimize the $V_c$ estimation accuracy, the Q-value for each state-action pair of a SBS $c$ is determined by

$$\begin{aligned} Q_c(s_{ct}^{m}, a_{ct}^{m}) = \{R(s_{ct}^{m}, a_{ct}^{m})\} + \alpha \sum_{s_{c(t+1)}^{m} \in \mathcal{S}} P\left(s_{c(t+1)}^{m} | s_{ct}^{m}, a_{ct}^{m}\right) \\ \max_{a_{c(t+1)}^{m} \in \mathcal{A}} Q\left(s_{c(t+1)}^{m}, a_{b(t+1)}^{m}\right), \end{aligned} \tag{18}$$

where $s_{ct}^{m}$ and $s_{c(t+1)}^{m}$ denote the current and next cache state, respectively, in response to the execution of action $a_{ct}^{m}$ during state $s_{ct}^{m}$. Moreover, the probability of state transition is determined by $P\left(s_{c(t+1)}^{m} | s_{ct}^{m}, a_{ct}^{m}\right) = P\left(v_{c(t+1)}^{m}, x_{c(t+1)}^{m} | v_{ct}^{m}, x_{ct}^{m}, a_{ct}^{m}\right)$.

We exploit the potential of the $\epsilon$ greedy policy to optimize action selection in order to maximize the rewards. The $\epsilon$ greedy policy performs a trade-off between the exploration and exploitation of actions. During the exploitation phase, those actions are selected that can improve the long-term rewards. Since this approach may halt the system's performance, it is necessary to also check the impact of new actions that can further improve long-term rewards. Consequently, exploration is performed to find new actions that can significantly enhance the achievable rewards in terms of network delay minimization. Hence, the $\epsilon$ greedy policy provides a balance between exploitation and exploration by a factor of $\epsilon \in (0, 1)$ [40]. The $V_i$ is evaluated as

$$V_i = \max_{a_{it}^{m} \in \mathcal{A}} Q_i(s_{it}^{m}, a_{it}^{m}). \tag{19}$$

After determining the optimal Q-value for all state-action pairs, the optimal policy is determined as

$$\varpi^{*}(s) = \arg\max_{a_{ct}^{m} \in \mathcal{A}} Q_c(s_{it}^{m}, a_{ct}^{m}). \tag{20}$$

The Q-learning algorithm is utilized to iteratively update Q-values using the following update rule [41].

$$Q_c^{t+1}(s_{ct}^m, a_{ct}^m) = (1-\psi)Q_c^t(s_{ct}^m, a_{ct}^m) + \psi\left(R_{ct}^m(s_{ct}^m, a_{ct}^m) + \alpha {V_c}^t(s_{ct}^m + a_{ct}^m)\right), \tag{21}$$

where $\psi$ represents the learning rate, and $(s_{ct}^m + a_{ct}^m)$ represents the transition to a new state from the previous state $s_{ct}^m$ due to performing an action $a_{ct}^m$ at any given time $t$.

### *3.1. Reward Function*

The reward function is determined by taking the difference between the new caching gain and the loss of removing content. The gain of the system is determined as follows:

$$\zeta_{ct}^{m^+} = \sum_{m=1}^{M} r_{ct}^{m^+}, \tag{22}$$

where $r_{ct}^{m^+}$ denotes the rise in cache hits due to caching new content $m$ in cache $c$. If the gain is large, then the delay is low. Furthermore, the loss of removing content is determined as

$$\zeta_{ct}^{m^-} = \sum_{m=1}^{M} r_{ct}^{m^-}, \tag{23}$$

where $r_{ct}^{m^-}$ denotes the decrease in the cache hits as a result of replacing the previous content with new content. This shows that the replaced content generated more requests. Hence, an increase in loss value demonstrates a rise in data access delay. After calculating both the gain and loss values, the reward function is evaluated as follows:

$$R(s_{ct}^m, a_{ct}^m) = \zeta_{ct}^{m^+} - \zeta_{ct}^{m^-}, \tag{24}$$

where $\zeta_{ct}^{m^+}$ and $\zeta_{ct}^{m^-}$ represent the gain and loss values, respectively.

### *3.2. Proposed Algorithm*

In this subsection, we discuss the proposed Algorithm 1. First, Q-values for every state-action pair are initialized. From steps 2 to 8, state-action pairs are evaluated iteratively to select actions that can maximize the Q-function. During steps 4 to 7, stochastic decisions are performed that are influenced by the random probability $p$ to exploit the exploration and exploitation of actions depending on $\epsilon$. In this manner, optimal caching actions are performed. Furthermore, in steps 9 and 10, the rewards are observed based on the executed action $a_{ct}^m$ and system transition to the next state $s_{c(t+1)}^m$, respectively. The system's long-term rewards are improved according to Equation (19). Finally, in Step 11, the Q-table is updated based on the reward values.

**Algorithm 1** RL-based dynamic data caching decision algorithm.

1: Initializing $Q_c(s_{ct}^m, a_{ct}^m)$ for every state-action pair.
2: **for** $\mathcal{T} = 1, 2, \ldots, T$ **do**
3: Select random probability $p$.
4: if $p \geq \epsilon$
5: $a_{ct}^m = \arg\max_{a_{ct}^m \in \mathcal{A}} Q_c(s_{ct}^m, a_{ct}^m)$,
6: otherwise,
7: randomly select an action $a_{ct}^m$.
8: Execute action $a_{ct}^m$ in the system.
9: Obtain reward $R_{ct}^m$.
10: $s_{ct}^m$ shift to the new state $s_{c(t+1)}^m$ and the subsequent interval.
11: Update,
$Q_c^{t+1}(s_{ct}^m, a_{ct}^m) = (1-\psi)Q_c^t(s_{it}^m, a_{ct}^m) + \psi(R_{ct}^m(s_{ct}^m, a_{ct}^m) + \alpha V_c(s_{ct}^m + a_{ct}^m))$,
12: **end for**

## 4. Simulations

In this section, after analytical modeling, we evaluate the performance of our proposed algorithm. The simulation parameters are provided in Table 2. For the performance evaluation, we utilize a real human mobility trajectory, which is collected in Korea Advanced Institute of Science & Technology (KAIST), Korea [42,43]. For evaluation purposes, we set $\epsilon = 0.5$ in order to ensure that all actions are explored sufficiently before selecting a particular action for reward maximization. In addition, we set $\psi$ and $\alpha$ to 0.6 and 0.7, respectively. The experiments were performed using MATLAB R2017a on an X64-based processor system that is equipped with an Intel(R) Core(TM) i7-7500 CPU running at 2.90 GHz and 16 GB RAM. Simulations were executed for 500 time periods and 800 iterations were performed for each simulation setup.

For performance validation, we compare the performance of our proposed caching approach with several benchmark approaches, including most popular caching (MPC), random caching (RC), least frequently used (LFU) caching, and no-D2D. The outlined benchmark approaches are discussed below:

- Most popular caching (MPC): MPC only caches popular content instead of caching all content. This provides cache resource conservation [44].
- Random caching (RC): In this scheme, random data caching decisions are made, resulting in a reduction in complexity without compromising data service quality [45].
- Least frequently used (LFU): This scheme focuses on data demand frequency, where data demands follows data popularity. This scheme incorporates data popularity and data demand frequency to optimize data caching [46].
- no-D2D: As its name implies, no D2D communication is used in this scheme. Moreover, data caching is performed only at the SBS level by exploiting Q-learning.

**Table 2.** System parameters.

| Parameter | Notation | Value |
|---|---|---|
| Small cell size | $A_{sc}$ | 250 m × 250 m |
| Millimeter-wave bandwidth | $W_{mm}$ | 2.16 GHz |
| Microwave bandwidth | $W_{\mu}$ | 20 MHz |
| SBS transmission power | $P_i$ | 30 dBm |
| UE transmission power | $P_v$ | 23 dBm |
| mmWave noise spectral density | $\sigma^2_{mm}$ | −134 dBm/MHz |
| Microwave noise spectral density | $\sigma^2_{\mu}$ | −174 dBm/Hz |
| Discount factor | $\alpha$ | 0.7 |
| Learning rate | $\psi$ | 0.8 |
| Half-power beamwidth | $\omega_m$ | 30° |
| Popularity skewness | $\gamma$ | 1 |
| Content size | $f_m$ | 100 Mbit |

Figure 2 illustrates the impact of varying amounts of content on network delay. The obtained results demonstrate that the delay increases with increasing amounts of content. The increase in the delay shows the unavailability of the requested content within the cache resources. The absence of requested content in the cache incurs an additional delay due to fetching requested content from the core network, substantially increasing data access delay. The obtained results demonstrate that when the amount of content is equal to 600, our proposed caching scheme attains 29%, 30.7%, 16.6%, and 24.8% lower delay than MPC, RC, LFU, and no-D2D schemes, respectively. These results highlight that our proposed caching approach can flexibly perform data caching and improve cache access.

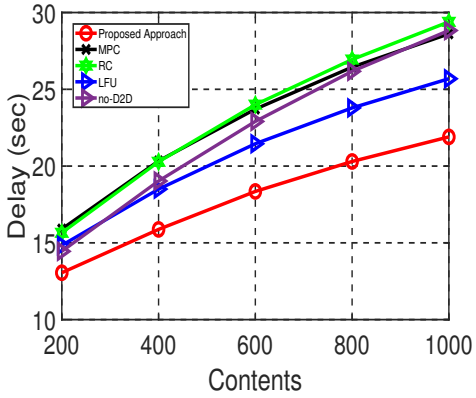


**Figure 2.** Impact of varying content quantity on network delay, where $\gamma$ is set to 1.

Figure 3 illustrates the influence of the UE's density on the delay. There is a direct relation between the number of sets of UE and data requests. When the amount of UE increases within a small cell, it constitutes more data traffic. However, our proposed approach exploits the UE's density to reduce data access by leveraging D2D communications. This enables the UE to fulfill the data demands through D2D communication. For instance, when the number of sets of UE is set to 30, our proposed caching scheme achieves 37.4%, 35%, 26.3%, and 46.3% lower delay than MPC, RC, LFU, and no-D2D schemes, respectively. The performance gain underscores the efficiency of D2D communications in exerting a positive influence on the network performance in terms of minimizing data access delay.

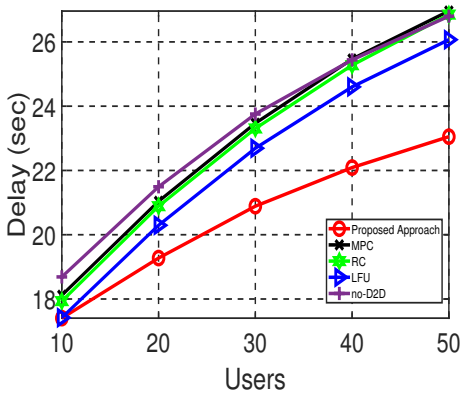


**Figure 3.** Impact of number of users on network delay, when $\gamma$ =1, $M$ = 1000, $c_n$ = 20, and $c_i$ = 50.

In Figure 4, the popularity skewness ($\gamma$) is varied, and its effect on the network performance is observed. A large $\gamma$ leads to less content being responsible for generating most of the data traffic. Our findings indicate that when $\gamma$ increases, the delay decreases. The results show that beyond $\gamma$ = 2.5, all the schemes show stable performance. The obtained results demonstrate that the popularity skewness $\gamma$ imposes a substantial impact on network performance, and it is imperative to accurately determine which content should be cached to reduce network delay. Our proposed scheme exhibits better performance than other baseline schemes, manifesting the efficiency of our proposed scheme in accurately capturing the data popularity profile to improve data caching.

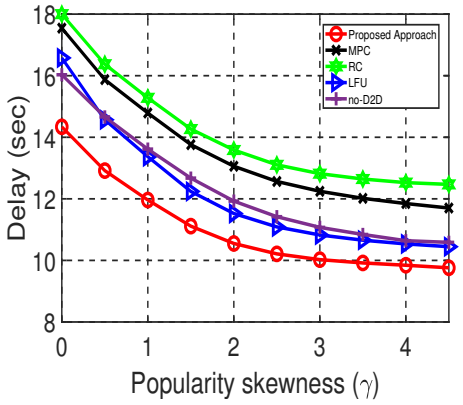


**Figure 4.** Network delay vs. $\gamma$ under various caching schemes when $M = 400$, $N = 50$, $c_n = 20$, and $c_i = 50$.

In Figure 5, we analyze the impact of the data popularity profile on the performance of our proposed scheme. We varied the amount of content while varying $\gamma$ values. The results illustrate that when $\gamma$ is low, the delay is high because the data popularity is scattered. Consequently, more content is responsible for generating data traffic that cannot be cached. This leads to more data fetching from the core network, imposing a data traffic burden on the backhaul. Conversely, when the $\gamma$ value rises, there is a profound reduction in the delay as the amount of popular content that can be located in the cache resources shrinks to meet demands. The influence of popularity skewness is evident because in the case of high popularity skewness, the delay is low. This finding underscores the importance of the data popularity profile during the design of a caching mechanism.

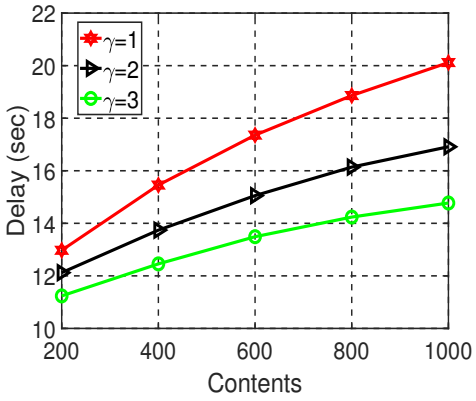


**Figure 5.** Influence of popularity skewness ($\gamma$) on network delay under varying amounts of content in the proposed scheme.

When the cache capacity is large, a large amount of content can be cached. However, both the SBS and UE are equipped with limited cache spaces, imposing a constraint on the maximum amount of content that can be cached. Consequently, uncached content results in a data traffic burden on the backhaul link. In Figure 6, we observe the impact of increasing the cache capacity of the SBS on the network delay. There is a monotonic decay in the performance across all the approaches. Nevertheless, our proposed caching scheme demonstrates the best performance compared with the benchmark approaches. The results demonstrate that when the cache capacity can accommodate 150 pieces of content, our proposed caching approach shows 25.2%, 23.3%, 11.3%, and 29.5% lower delay compared with the MPC, RC, LFU, and no-D2D approaches, respectively. Therefore, in order to maximize the advantages of the caching mechanism, it is crucial to cache only that content that generates the majority of data requests.

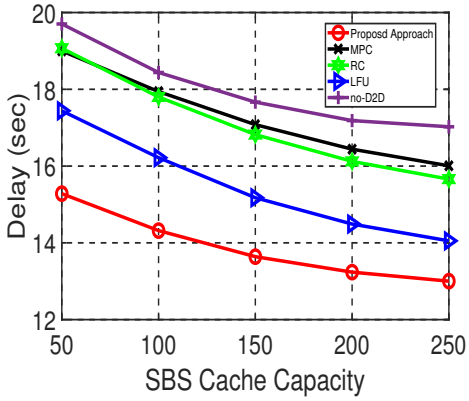


**Figure 6.** Impact of increasing SBS cache capacity on network delay.

In Figure 7, we observe the impact of increasing the amount of content on the percentage of backhaul consumption. When the amount of content increases, it is not possible to cache all the content in the limited cache resources. Furthermore, the UE's data demands exhibit diversity, leading to more requests for uncached content. For instance, when there are 300 pieces of content, our proposed caching scheme demonstrates a reduction of 37%, 33%, 22%, and 44.6% in data access via backhaul compared with MPC, RC, LFU, and no-D2D, respectively. The obtained results illustrate that our proposed caching scheme can substantially alleviate the traffic burden on the backhaul link through effective data caching.

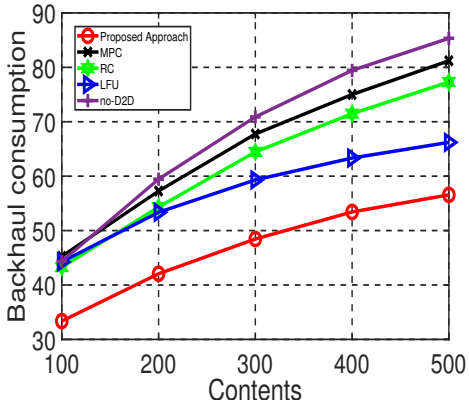


**Figure 7.** Impact of increasing the amount of content on backhaul consumption, when $\gamma = 1$, $N = 50$, $c_i = 50$, and $c_n = 10$.

## 5. Conclusions

In this paper, we introduced a novel two-tier caching approach and exploited dual-model data transmission through mmWave and μWave bands between the UE and serving SBS. Additionally, D2D communications through an mmWave link are incorporated, enabling the UE to obtain the desired content through mmWave-based D2D communications. Furthermore, an RL approach is utilized in the proposed caching scheme to optimize two-tier caching decisions, while taking into account evolving network conditions, cache capacity constraints, and diverse users demands. In an iterative manner, the caching policy is optimized by selecting actions that can improve the rewards and minimize delays. The obtained results demonstrate that our proposed approach substantially minimized delays compared with the benchmark approaches.

In future work, we aim to utilize mobile unmanned aerial vehicles (UAVs) as cache-enabled aerial base-stations to facilitate high-speed data transmission via LoS links to UE, ensuring seamless connectivity to handoff UE and mitigating communication disruptions.

**Author Contributions:** Conceptualization and methodology, M.S. and M.A.; Supervision, M.A.; software, M.S., M.A. and A.I.; writing, M.S., M.A. and T.C.C.; review and editing, T.C.C., M.B.R. and A.A.-H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Multimedia University research grant MMUI/230013.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Zheng, K.; Liu, X.; Wang, B.; Zheng, H.; Chi, K.; Yao, Y. Throughput Maximization of Wireless-Powered Communication Networks: An Energy Threshold Approach. *IEEE Trans. Veh. Technol.* **2021**, *70*, 1292–1306. [CrossRef]
2. Deng, B.; Nan, J.; Cao, W.; Wang, W. A Survey on Integration of Network Communication into Vehicle Real-Time Motion Control. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 2755–2790. [CrossRef]
3. Rasilainen, K.; Phan, T.D.; Berg, M.; Pärssinen, A.; Soh, P.J. Hardware Aspects of Sub-THz Antennas and Reconfigurable Intelligent Surfaces for 6G Communications. *IEEE J. Sel. Areas Commun.* **2023**, *41*, 2530–2546. [CrossRef]
4. Singh, S.; Mudumbai, R.; Madhow, U. Interference Analysis for Highly Directional 60-GHz Mesh Networks: The Case for Rethinking Medium Access Control. *IEEE/ACM Trans. Netw.* **2011**, *19*, 1513–1527. [CrossRef]
5. Ning, B.; Tian, Z.; Mei, W.; Chen, Z.; Han, C.; Li, S.; Yuan, J.; Zhang, R. Beamforming Technologies for Ultra-Massive MIMO in Terahertz Communications. *IEEE Open J. Commun. Soc.* **2023**, *4*, 614–658. [CrossRef]
6. Ge, Y.; Fan, J.; Li, G.Y.; Wang, L.C. Intelligent Reflecting Surface-Enhanced UAV Communications: Advances, Challenges, and Prospects. *IEEE Wirel. Commun.* **2023**, *30*, 119–126. [CrossRef]
7. Lyu, Y.; Yuan, Z.; Zhang, F.; Kyösti, P.; Fan, W. Virtual Antenna Array for W-Band Channel Sounding: Design, Implementation, and Experimental Validation. *IEEE J. Sel. Top. Signal Process.* **2023**, *17*, 729–744. [CrossRef]
8. Zhang, J.; Xi, R.; He, Y.; Sun, Y.; Guo, X.; Wang, W.; Na, X.; Liu, Y.; Shi, Z.; Gu, T. A Survey of mmWave-Based Human Sensing: Technology, Platforms and Applications. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 2052–2087. [CrossRef]
9. Ozkaptan, C.D.; Zhu, H.; Ekici, E.; Altintas, O. A mmWave MIMO Joint Radar-Communication Testbed with Radar-assisted Precoding. *IEEE Trans. Wirel. Commun.* **2023**. [CrossRef]
10. Sheraz, M.; Shafique, S.; Imran, S.; Asif, M.; Ullah, R.; Ibrar, M.; Bartoszewicz, A.; Mobayen, S. Mobility-Aware Data Caching to Improve D2D Communications in Heterogeneous Networks. *Electronics* **2022**, *11*, 3434. [CrossRef]
11. Al-Habashna, A.; Fernandes, S.; Wainer, G. DASH-based peer-to-peer video streaming in cellular networks. In Proceedings of the 2016 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS), Montreal, QC, Canada, 24–27 July 2016; pp. 1–8. [CrossRef]
12. Zhang, Z.; Lung, C.H.; Wei, X.; Chen, M.; Chatterjee, S.; Zhang, Z. In-Network Caching for ICN-Based IoT (ICN-IoT): A Comprehensive Survey. *IEEE Internet Things J.* **2023**, *10*, 14595–14620. [CrossRef]
13. Al-Habashna, A.; Wainer, G. Improving video transmission in cellular networks with cached and segmented video download algorithms. *Mob. Netw. Appl.* **2018**, *23*, 543–559. [CrossRef]
14. Al-Habashna, A.; Wainer, G. QoE awareness in progressive caching and DASH-based D2D video streaming in cellular networks. *Wirel. Netw.* **2020**, *26*, 2051–2073. [CrossRef]
15. Bepari, D.; Mondal, S.; Chandra, A.; Shukla, R.; Liu, Y.; Guizani, M.; Nallanathan, A. A Survey on Applications of Cache-Aided NOMA. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 1571–1603. [CrossRef]
16. Zhou, M.; Li, J.; Yuan, J.; Xie, M.; Tan, W.; Yin, R.; Yang, L. An Architecture for AoI and Cache Hybrid Multicast/Unicast/D2D With Cell-Free Massive MIMO Systems. *IEEE Access* **2023**, *11*, 43080–43088. [CrossRef]
17. Xiao, H.; Zhuang, Y.; Xu, C.; Wang, W.; Zhang, H.; Ding, R.; Cao, T.; Zhong, L.; Muntean, G.M. Transcoding-Enabled Cloud–Edge–Terminal Collaborative Video Caching in Heterogeneous IoT Networks: An Online Learning Approach with Time-Varying Information. *IEEE Internet Things J.* **2024**, *11*, 296–310. [CrossRef]
18. Gu, H.; Zhao, L.; Han, Z.; Zheng, G.; Song, S. AI-Enhanced Cloud-Edge-Terminal Collaborative Network: Survey, Applications, and Future Directions. *IEEE Commun. Surv. Tutor.* **2023**. [CrossRef]
19. Shanmugam, K.; Golrezaei, N.; Dimakis, A.G.; Molisch, A.F.; Caire, G. FemtoCaching: Wireless Content Delivery Through Distributed Caching Helpers. *IEEE Trans. Inf. Theory* **2013**, *59*, 8402–8413. [CrossRef]
20. Chiputa, M.; Zhang, M.; Chong, P.H.J. Pattern Based Mobility Management in 5G Networks With a Game Theoretic-Jump Markov Linear System Approach. *IEEE Access* **2023**, *11*, 116410–116422. [CrossRef]
21. Dou, J.; Xie, G.; Tian, Z.; Cui, L.; Yu, S. Modeling and Analyzing the Spatial-Temporal Propagation of Malware in Mobile Wearable IoT Networks. *IEEE Internet Things J.* **2023**, *11*, 2438–2452. [CrossRef]

22. Ochia, O.E.; Fapojuwo, A.O. Popularity and Size-Aware Caching With Cooperative Transmission in Hybrid Microwave/Millimeter Wave Heterogeneous Networks. *IEEE Trans. Commun.* **2021**, *69*, 4599–4614. [CrossRef]
23. Gu, Z.; Lu, H.; Hong, P.; Zhang, Y. Reliability Enhancement for VR Delivery in Mobile-Edge Empowered Dual-Connectivity Sub-6 GHz and mmWave HetNets. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 2210–2226. [CrossRef]
24. Rostampoor, J.; Adve, R. Dynamic Caching in a Hybrid Millimeter-wave/Microwave C-RAN. In Proceedings of the 2022 IEEE International Conference on Communications Workshops (ICC Workshops), Seoul, Republic of Korea, 16–20 May 2022; pp. 1–6.
25. Rostampoor, J.; Adve, R.S. Optimizing Caching in a C-RAN With a Hybrid Millimeter-Wave/Microwave Fronthaul Link via Dynamic Programming. *IEEE Trans. Commun.* **2023**, *71*, 923–934. [CrossRef]
26. Rostampoor, J.; Adve, R.; Afana, A.; Ahmed, Y. CPRL: Change Point Detection and Reinforcement Learning to Optimize Cache Placement Strategies. *IEEE Trans. Commun.* **2023**. [CrossRef]
27. Xie, Z.; Chen, W.; Poor, H.V. A Unified Framework for Pushing in Two-Tier Heterogeneous Networks with mmWave Hotspots. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 19–31. [CrossRef]
28. Lin, Z.; Fang, Y.; Chen, P.; Chen, F.; Zhang, G. Modeling and Analysis of Edge Caching for 6G mmWave Vehicular Networks. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 7422–7434. [CrossRef]
29. Lin, H.; Zhang, C.; Huang, Y.; Zhao, R.; Yang, L. Performance Analysis of Cache-Enabled User Association for Hybrid Heterogeneous Cellular Networks. *IEEE Trans. Commun.* **2022**, *70*, 2518–2531. [CrossRef]
30. Ravanshid, A.; Rost, P.; Michalopoulos, D.S.; Phan, V.V.; Bakker, H.; Aziz, D.; Tayade, S.; Schotten, H.D.; Wong, S.; Holland, O. Multi-connectivity functional architectures in 5G. In Proceedings of the 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 23–27 May 2016, pp. 187–192.
31. Bai, T.; Vaze, R.; Heath, R.W. Analysis of Blockage Effects on Urban Cellular Networks. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 5070–5083. [CrossRef]
32. Kiskani, M.K.; Sadjadpour, H.R. Throughput Analysis of Decentralized Coded Content Caching in Cellular Networks. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 663–672. [CrossRef]
33. Niu, Y.; Liu, Y.; Li, Y.; Chen, X.; Zhong, Z.; Han, Z. Device-to-Device Communications Enabled Energy Efficient Multicast Scheduling in mmWave Small Cells. *IEEE Trans. Commun.* **2018**, *66*, 1093–1109. [CrossRef]
34. Liu, Y.; Chen, X.; Niu, Y.; Ai, B.; Li, Y.; Jin, D. Mobility-Aware Transmission Scheduling Scheme for Millimeter-Wave Cells. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 5991–6004. [CrossRef]
35. Geng, S.; Kivinen, J.; Zhao, X.; Vainikainen, P. Millimeter-Wave Propagation Channel Characterization for Short-Range Wireless Communications. *IEEE Trans. Veh. Technol.* **2009**, *58*, 3–13. [CrossRef]
36. Hao Xu.; Kukshya, V.; Rappaport, T.S. Spatial and temporal characteristics of 60-GHz indoor channels. *IEEE J. Sel. Areas Commun.* **2002**, *20*, 620–630. [CrossRef]
37. Daniels, R.C.; Heath, R.W. 60 GHz wireless communications: Emerging requirements and design recommendations. *IEEE Veh. Technol. Mag.* **2007**, *2*, 41–50. [CrossRef]
38. Singh, S.; Ziliotto, F.; Madhow, U.; Belding, E.; Rodwell, M. Blockage and directivity in 60 GHz wireless personal area networks: From cross-layer model to multihop MAC design. *IEEE J. Sel. Areas Commun.* **2009**, *27*, 1400–1413. [CrossRef]
39. Niu, Y.; Li, Y.; Jin, D.; Su, L.; Vasilakos, A.V. A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges. *Wirel. Netw.* **2015**, *21*, 2657–2676. [CrossRef]
40. Sutton, R.S.; Barto, A.G. *Introduction to Reinforcement Learning*; University of Alberta: Edmonton, AL, Canada, 1998.
41. Even-Dar, E.; Mansour, Y. Convergence of Optimistic and Incremental Q-Learning. In Proceedings of the NIPS, Vancouver, BC, Canada, 3–7 December 2001.
42. Dai, Z.; Liu, C.H.; Ye, Y.; Han, R.; Yuan, Y.; Wang, G.; Tang, J. AoI-minimal UAV Crowdsensing by Model-based Graph Convolutional Reinforcement Learning. In Proceedings of the IEEE INFOCOM 2022—IEEE Conference on Computer Communications, London, UK, 2–5 May 2022; pp. 1029–1038.
43. Rana, R.; Yang, M.; Wark, T.; Chou, C.T.; Hu, W. SimpleTrack: Adaptive Trajectory Compression With Deterministic Projection Matrix for Mobile Sensor Networks. *IEEE Sens. J.* **2015**, *15*, 365–373. [CrossRef]
44. Chen, Z.; Lee, J.; Quek, T.Q.S.; Kountouris, M. Cooperative Caching and Transmission Design in Cluster-Centric Small Cell Networks. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 3401–3415. [CrossRef]
45. Psounis, K.; Prabhakar, B. A randomized Web-cache replacement scheme. In Proceedings of the IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213), Anchorage, AK, USA, 22–26 April 2001; Volume 3, pp. 1407–1415.
46. Chatzieleftheriou, L.E.; Karaliopoulos, M.; Koutsopoulos, I. Jointly Optimizing Content Caching and Recommendations in Small Cell Networks. *IEEE Trans. Mob. Comput.* **2019**, *18*, 125–138. [CrossRef]