Arabian Journal for Science and Engineering (2019) 44:9719–9733
https://doi.org/10.1007/s13369-019-03830-w

RESEARCH ARTICLE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

Cluster-Based Architecture Capable for Device-to-Device Millimeter-Wave Communications in 5G Cellular Networks

Ghazal Mesbahi[1] · Akbar Ghaffarpour Rahbar[1]

Received: 3 October 2018 / Accepted: 23 March 2019 / Published online: 3 April 2019

Abstract

The fifth-generation (5G) wireless networks are the newest mobile technologies proposed for supporting high-data-rate traffic and challenges of previous generations such as spectrum crisis and high energy consumption. Millimeter-wave (mmWave) communication is a promising technology for 5G cellular networks aiming at solving microwave spectrum crisis and providing very high data rates for users. Enabling device-to-device (D2D) communications over mmWave networks can improve the efficiency of these networks. In this article, a new cluster-based architecture capable for D2D mmWave communication (CADM) with TDMA-based medium access control structure is proposed to improve the performance of 5G networks. Using clustering for CADM results in reducing energy consumption and prolonging network lifetime. In addition, enabling simultaneous short-distance mmWave connections on the same frequencies in this architecture not only improves data rates, throughput, and spectral efficiency but also reduces end-to-end delay of 5G mobile networks.

Keywords 5G mobile networks · D2D communications · mmWave · Simultaneous communications

1 Introduction

With the impressive success of mobile wireless communications in various generations (1G to 4G), wireless data traffic has grown dramatically over the past few years since many users subscribe to mobile broadband systems every year [1,2]. The explosion of wireless mobile devices and services has caused some challenges such as the spectrum crisis and high energy consumption. The existing techniques can no longer satisfy the needs of users since wireless system designers face the continuously increasing demand for high data rates and mobility required by new wireless applications. Therefore, researches on the fifth-generation (5G) wireless systems started. 5G cellular networks are expected to have much higher network capacity and provide multi-gigabits-per-second data rates for each user to support multimedia applications with stringent quality of service (QoS) requirements. In fact, telecommunication industry has offered this new technology to conquer the defects and limitations of the earlier technologies. In order to achieve these goals, new promising techniques are provided for 5G networks.

Device-to-device (D2D) communication is a promising technique in the next-generation cellular networks for meeting the demand for high data rates [3]. Unlike the traditional communications where traffic has to go through the base station (BS) even if the mobile users are close to each other, D2D communications allow direct exchange of data between two mobile users without using a BS or the core network. D2D functionality is very helpful for offloading traffic from the BS and facilitating resource sharing for users who are spatially close to each other. In addition, D2D communication can be employed in natural calamities. In an earthquake or storm, an urgent communication network can be set up using D2D functionality in a short time to supersede the damaged communication.

The mmWave communication with capability of gigabits-per-second data rates is a promising candidate technology for 5G wireless communications [4]. The 3–300 GHz spectrum with wavelengths range of 1–100 mm is collectively denoted to as mmWave bands. There are several advantages and disadvantages to utilize mmWave high frequencies in future 5G networks. The advantages include: (a) overcome cellular spectrum (below 3 GHz) crisis due to an abundant

✉ Akbar Ghaffarpour Rahbar
ghaffarpour@sut.ac.ir

Ghazal Mesbahi
ghazal.mesbahi@yahoo.com

[1] Computer Networks Research Lab, Sahand University of Technology, Tabriz, Iran

amount of spectrum at mmWave frequencies, (b) use of directional antennas, which result in narrow directive beams, and (c) achieve the inherent security and privacy of mmWave transmission due to the limited transmission range. The disadvantages include [5]: (a) much higher propagation loss compared to the microwave band, (b) some difficulties in diffracting around obstacles due to short wavelengths of mmWave bands, and (c) some difficulties in penetrating through solid materials. These features of mmWave propagation cause challenges to provide seamless coverage, reliable communication, and high-quality services, especially for long-distance communications.

The main motivation for this study is to solve the problems in 5G networks (such as high energy consumption, cellular spectrum shortage, overloaded BSs, high latency, and increased demand for high data rates). Most of the architectures proposed for 5G networks focus only on one of the above challenges, while this article attempts to partly solve most of the above problems and improve the performance of 5G systems by proposing the CADM architecture and utilizing some of 5G promising technologies in this architecture.

Due to the advantages and disadvantages of mmWave communications, we use mmWave bands for short-distance D2D communications to take their advantages in this article. MmWave communications have relatively low multi-user interference due to the directional antenna and high propagation loss, which can support simultaneous communications [6]. By allowing multiple concurrent D2D links, the network capacity can be further improved. Hence, our objective in this paper is to propose a new cluster-based architecture capable for simultaneous D2D mmWave communication (CADM) for 5G mobile networks. Clearly, clustering is one of the techniques to improve energy efficiency in wireless networks by balancing load and energy consumption among wireless devices. Hence, we will use clustering in our proposed architecture to establish multiple concurrent short-distance D2D mmWave connections on the same frequency within clusters, between two cluster normal members (CNMs), and/or between CNM and cluster head (CH), and/or between two adjacent clusters. Also, we propose new cluster formation and CH transfer algorithms in the proposed architecture that can save energy and prolong the network lifetime.

Unlike most architectures, our proposed scheme enables simultaneous connections on the same frequency within a cluster through communicating between a CNM and a CH or between two CNMs by proposing a proper MAC protocol. This scheme increases data rate, throughput, number of concurrent connections, spectral efficiency, while decreasing end-to-end delay. In addition, our proposed clustering algorithm improves the network energy efficiency, lifetime, and overall system performance by choosing the CHs with more suitable conditions and more ideal metrics for long-distance communications.

The remainder of this article is organized as follows. A brief overview of the related work in 5G architectures is explained in Sect. 2. In addition, the network model is discussed in Sect. 3. In Sect. 4, cluster formation and CH selection algorithms are proposed, and intra- and inter-cluster communications are presented for the CADM architecture in Sect. 5. The proposed scheme is simulated using the NS-3 simulator and its performance is evaluated in Sect. 6. Finally, the article is concluded in Sect. 7.

2 Related Works

In this section, we provide a brief overview of the related works on D2D communication, mmWave communication, clustering, and 5G architectures. A survey on D2D communications has been presented in [7]. The work in [8] offers new approaches for underlay D2D communication in spectrum-shared heterogeneous cellular networks. It considers device mode selection, interference mitigation, and resource allocation for underlay D2D communication. In [9], generalized modeling and analysis of the impact of imperfect D2D association are presented for direct D2D communications in spectrum-shared downlink cellular networks. Also, practical interference and energy constraints imposed on the underlying D2D communication network are considered. Resource allocation and interference mitigation have been investigated in [10,11] for heterogeneous networks where the lowest tier consists of D2D cells. The authors of [10,11] address downlink/uplink decoupling user association and its effect on interference management and network-wide D2D performance improvement and suggest an uplink fractional frequency reuse scheme where sub-band bandwidths are adaptively determined.

In [12], a survey of mmWave communications has been presented for 5G mobile networks. The authors of [4] have discussed some key elements to enable mmWave communications in 5G wireless systems such as: mmWave channel characteristics, beamforming technologies, blockage effect in mmWave communications, mmWave transmission in D2D communications, and mmWave transmission in heterogeneous networks. In [13], the motivation for new mmWave cellular systems, methodology, and hardware for measurements has been presented and a variety of measurement results has been offered that show 28 and 38 GHz frequencies can be used when employing steerable directional antennas at BSs and mobile devices. In addition, a survey of clustering schemes has been provided in [14–19].

Some novel architectures for 5G mobile networks has been proposed in [6,20–25]. In [20], an all-IP-based system model has been proposed to design the basic network architecture of 5G mobile systems. In [21], a general architecture for 5G mobile networks has been proposed. This architecture describes the relationships between emerging technologies for 5G networks. A centralized network architecture based on super BS has been proposed in [22]. In [23], a cognitive radio multi-agent architecture has been proposed for 5G wireless networks. An energy-efficient architecture has been proposed in [24] for the next-generation mobile networks using energy-efficient radio technologies. The authors of [6] have introduced an mmWave + 4G system architecture with TDMA-based MAC structure and an effective resource sharing scheme by allowing non-interfering D2D links for 5G cellular networks. In [25], a cluster-based architecture called Hybrid-Band has been proposed for 5G networks that combines ISM band 2.4 GHz spectrum as the Out-Band mode with the cellular spectrum as the In-Band mode for D2D communications. The base scenario of this architecture includes a BS and two clusters. Each cluster has eight devices where one device is chosen as the master (i.e., cluster head (CH)), and the others are the slave devices (i.e., cluster normal member (CNM)). The CH is changed periodically based on the most value of residual charge and signal-to-interference-plus-noise ratio (SINR). All the user equipments (UEs) inside a cluster can directly transmit data to the CH through the Out-Band mode D2D communications. Each cluster has its own CH which establishes the In-Band mode D2D communications. This architecture improves throughput, energy efficiency, BS load, end-to-end delay, and lifetime of 5G networks. Nonetheless, transmission of high data rates and simultaneous connections on the same frequencies in a cluster are not possible in this architecture. This results in relatively low throughput and low spectral efficiency.

To overcome the problems of Hybrid-Band architecture, we propose the CADM architecture that can enable concurrent connections on the same frequencies in a cluster and improve energy efficiency and lifetime of 5G networks by proposing a new clustering algorithm.

3 Network Model

In this section, the proposed CADM architecture is detailed and the suggested medium access control (MAC) for this architecture is also described.

3.1 CADM Architecture

Figure 1 illustrates an outline of the proposed heterogeneous cluster-based architecture capable for D2D mmWave communication (CADM) for dense 5G cellular networks, which consists of BSs and mobile devices. The 5G mobile network is expected to support various types of applications. In the proposed architecture, cellular communications are used to support low-data-rate applications such as voice, text, and Web browser with full QoS. In addition, D2D mmWave communications are employed to support the direct exchange of high-volume data at high data rates among users who are spatially close to each other. In cellular 5G networks with CADM architecture, the whole geographical area is partitioned into cells, each of which is covered by one BS placed in the center of the cell. The BSs are able to work in both microwave and mmWave frequencies. Furthermore, each user equipment (UE) has communication modes of both

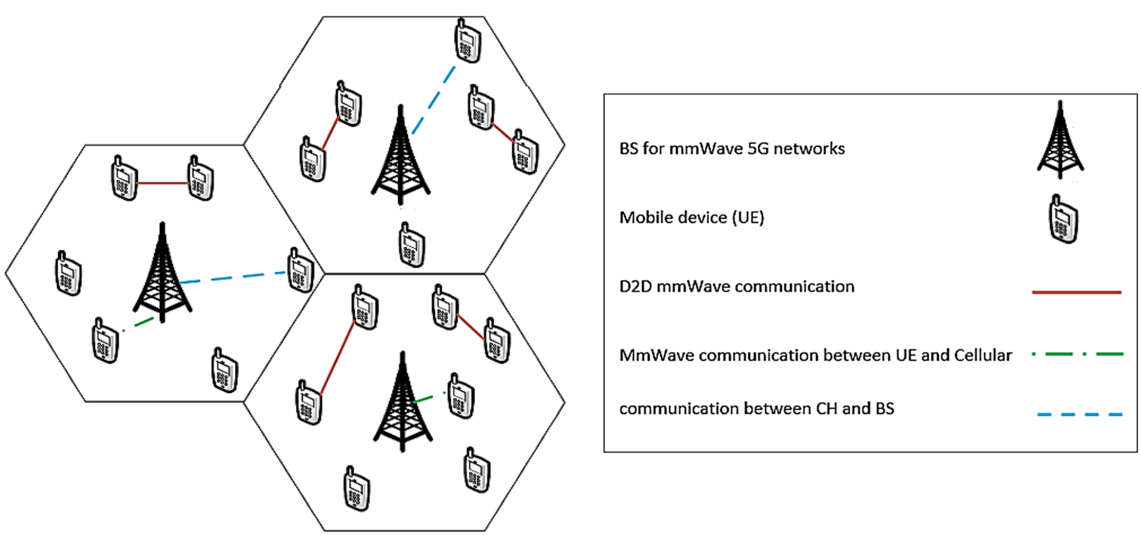
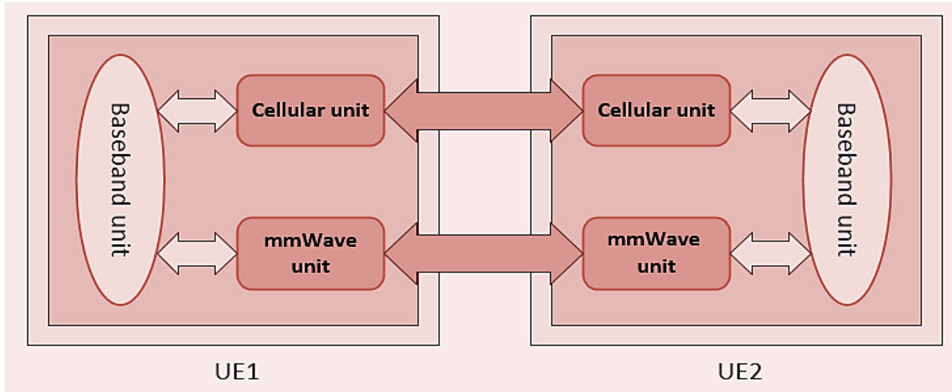


Fig. 1 Heterogeneous cluster-based architecture capable for D2D mmWave communication (CADM) for dense 5G cellular networks

Fig. 2 Wireless operation mode of each node [6]



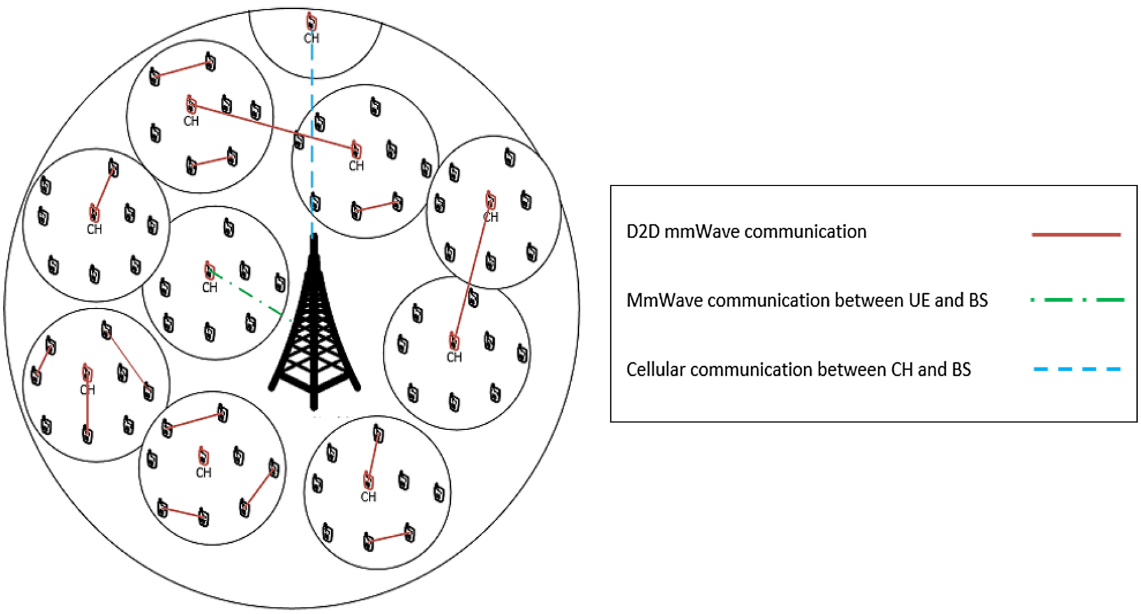


Fig. 3 Proposed architecture for each cell and communications within the cell

cellular and mmWave operations and supports fast mode transition between them, as shown in Fig. 2 [6]. Two UEs can communicate with each other at the same mode. For short-distance and high-data-rate communications, mmWave band is used, and for long-distance and low-data-rate communications, cellular band is used.

BSs and all of wireless devices are equipped with electronically steerable directional antennas with beamforming technologies for mmWave communications and also equipped with omnidirectional antennas for cellular communications. It is supposed that each transceiver can specify the best transmission/reception beam patterns for data transmission by using mmWave beamforming technologies. Figure 3 demonstrates the internal architecture of each cell with the formed clusters and intra-cell communications. As shown in Fig. 3, each cluster has a CH, and each CNM of a cluster can communicate with each other or with its own CH through D2D mmWave communication. The clustering and communications within the proposed architecture will be detailed in Sects. 4 and 5, respectively.

3.2 Medium Access Control

In our proposed architecture, time-division multiple access (TDMA) is adopted for channel access in millimeter-wave wireless 5G networks with the superframes shown in Fig. 4. In TDMA method, time is partitioned into superframes, each of which includes Z time slots. Transmission requests and signaling information are transferred to the BS through the control channel, and the BS properly allocates channels and time slots to each UE.

The usage of directional antennas in mmWave transmission results in creating narrow directive beams, which brings a new feature to mmWave systems [4]. This feature allows multiple D2D mmWave connections to be established simultaneously using the same frequency as long as there is no

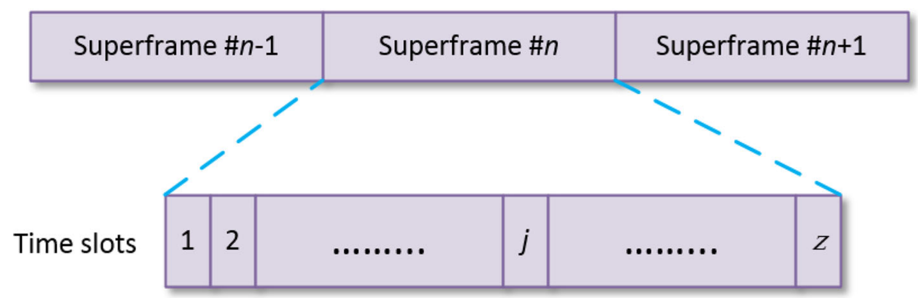


Fig. 4 Superframe-based scheduling structure for each cell

collision between the connections. This improves spectral efficiency and decreases end-to-end delay of 5G networks. Therefore, each time slot is actually a TDMA slot assigned by the BS for multiple communication links on the same frequency.

The BS checks the requests of all nodes which have data to transmit by employing the centralized polling method [6]. Therefore, all of transmission requests are collected by BSs at the beginning of each superframe. Each transmission request indicates the number of needed time slots. The BS constantly checks the possibility of using the requested link simultaneously with other users using the same link without any collisions. If the link does not interfere with existing links during the requested time, the requested link will be active at that time. After reviewing all links, the set of active links for the current time slot is obtained. If the required number of time slots of a link is satisfied, it means that the link has sent all of its data and does not have any data to send in time slots. Therefore, it should be set inactive, and it is not necessary to check this link in the following time slots. The above procedure is continued until all the time slots have been traversed. If the request of a link is not satisfied in the current superframe, it will be retransmitted in the next superframe to share the resources with other links. Note that two links which have the same node cannot work simultaneously due to half-duplex constraint of wireless communications. In addition, there should be a maximum of one connection between the same UE and BS in each time slot.

This method reduces the collisions between concurrent same-frequency links, but there is still possibility of collision due to the mobility of mobile nodes. After assigning time slots to a UE, it can move in different directions. Therefore, when an UE is allowed to transmit data in specified time slots, it is probable that the UE is located in a place that interferes with other connections using the same frequency. To support mobility of UEs by the medium access control protocol and minimize the collision among connections of the UEs requesting to transmit data on the same mmWave frequencies, the CSMA/CA protocol is run by the UEs at the beginning of the assigned time slots. If there is no collision probability for the mmWave connections of the set of active links according to CSMA/CA, the connections will be established in the allocated slots. But if there is a collision probability for some mmWave connections, the connections with collision probability cannot be established in the allocated time slots, the CSMA/CA protocol ends and the UEs which could not send their data in the assigned time slots due to collision probability send a request for allocating new time slots to their BS.

4 Clustering

Since mobile devices usually have limited battery power, it is essential to provide an effective energy management scheme that reduces power consumption. To manage the energy of UEs, we use the weighted clustering algorithms based on combined metrics [14] to balance energy consumption among UEs. In clustering, since a UE chosen as a CH has the most activity, it will have the most battery consumption. Therefore, CHs need to be changed periodically to prolong the network lifetime.

In our proposed clustering algorithm, we consider the metrics taken in Table 1 for CH selection.

The selection of a UE with less R_N causes the UE with the highest remaining energy level to be selected as a CH. Besides, in order to prevent the selection of UEs with very poor channel conditions as a CH that causes lower throughput, we consider the S_N metric. The selection of a UE with less S_N increases the network throughput. Since high mobility (especially in the mmWave bands where directional antennas are used) decreases data rates and throughput, as

Table 1 Clustering metrics and variables

Metric	Variable
The normalized value of consumed battery power (battery charge) of the candidate UE to be CH till current time T	R_N
The normalized value of 1 / SINR of the candidate CH	S_N
The normalized value of average speed of the candidate CH till current time T	V_N
The normalized value of the total distances between the UEs and the candidate CH in a cluster	D_N

well as increase in energy consumption (due to the need for retransmissions), the selection of a UE with less V_N as a CH is more suitable. In addition, longer-distance results in more transmission power and more energy consumption. By considering the D_N metric, the UE that has the smallest sum of distances to all other UEs in a cell is chosen as CH. This decreases total transmission power and energy consumption.

In the proposed clustering algorithm, we define $IV_{initial}$ as the parameter for the initial selection of a CH during the clusters formation. This parameter determines the selection or non-selection of a candidate UE as CH based on the above metrics. In addition, we define $IV_{periodic}$ as the parameter of the CH updated after forming the clusters. This parameter is a criterion for updating the CHs in each period and determines the selection or non-selection of a candidate UE as CH according to the above metrics. Whatever the value of these two parameters are lower for a UE, the UE is more suitable to be a CH. Due to the density and congestion of the network, in order to prevent excessive increase in load on the CH and to ensure the efficiency of the system, it is assumed that each CH can ideally support δ CNMs.

In this algorithm, in order to avoid random CH selection, R_N and S_N metrics are taken into account when forming the cluster (CH initial selection). For choosing the CHs in the next periods (CH transfer), in addition to R_N and S_N metrics, D_N and V_N metrics are also considered. Note that the value of V_N is calculated from Eq. (1) [26]:

$$V_N = \frac{1}{T}\sum_{t=1}^{T}\sqrt{(X_t - X_{t-1})^2 + (Y_t - Y_{t-1})^2} \quad (1)$$

where (X_t, Y_t) and (X_{t-1}, Y_{t-1}) are the coordinates of the candidate CH at time t and $(t-1)$, respectively.

4.1 Cluster Formation

In this section, we describe the proposed algorithm for creating clusters. In this algorithm, it is assumed that $T1$ is the time interval that a UE waits for receiving a HELLO message from a CH, and if it is unable to receive HELLO packets from any CH during $T1$, the UE begins to form its own cluster as a CH. A HELLO packet is a message that each UE places its information in it and periodically broadcasts it to its neighbors. Additionally, it is assumed that $T2$ is the time interval that the UE will wait to resend its request to the CH after receiving a reject response from the CH to join the cluster.

The proposed algorithm for cluster formation and selection of a CH during clusters formation in a cell is as follows. Each mobile UE is in the initial state as soon as it enters the cell or it is switched on. The UE, to specify its situation, forms a HELLO packet, calculates its $IV_{initial}$ and places it in the HELLO packet. The UE broadcasts the HELLO packets to its neighbors and then starts listening to the broadcast messages from the neighboring UEs in order to receive a message from a CH for joining to a cluster.

If the UE is unable to receive HELLO packets from any CH during $T1$ time span, it means that there is no cluster in the available range. Hence, the UE begins to form its own cluster as the CH and broadcasts the HELLO message periodically. However, if it receives a HELLO packet from a CH during $T1$ time interval, it indicates that there is a cluster nearby to join it. In this case, the UE sends a request to the CH to join the cluster. The CH sends a response to the UE after receiving the request.

If the UE is allowed to join the cluster according to the received response, the UE compares its $IV_{initial}$ with $IV_{initial}$ of the CH, which is located in the received HELLO message from the CH. If $IV_{initial}$ of the UE is larger than $IV_{initial}$ of the CH, it means that the conditions of current CH are more ideal than the UE to be CH. Thus, it remains as the CH, the UE joins the cluster as a CNM and the CH adds the information of the UE as a new CNM to its list. Otherwise, if $IV_{initial}$ of the UE is less than or equal to $IV_{initial}$ of the CH, it means that the UE has better conditions to be CH. Hence, the UE changes its status from initial to CH, places its information into the HELLO message, and periodically broadcasts the HELLO message. Moreover, the previous CH changes its status from CH to a normal member of that cluster.

If the UE receives a rejection from the CH, it means that the number of the normal members of the cluster is equal to δ, and therefore, the UE is not allowed to join the cluster due to the capacity of the cluster. Therefore, the UE will resend its request after $T2$ interval so that if the cluster will have a capacity during $T2$ (due to the exiting of some UEs from the cluster), the UE can join the cluster. If the UE is not allowed to join the cluster after sending maximum three requests, the UE begins to form its own cluster as the CH and broadcasts the HELLO message periodically. Figure 5 shows the procedure of the cluster formation.

Parameter $IV_{initial}$ is calculated according to Eq. (2):

$$\begin{aligned} &IV_{initial} = w_1 S_N + w_2 R_N \\ &w_1, w_2 \quad \in [0, 1], w_1 + w_2 = 1, S_N, R_N \quad \in [0, 1] \end{aligned} \quad (2)$$

where w_1 and w_2 are weight factors. As Eq. (2) shows, the weight factor w_1 is the coefficient of S_N and the weight factor w_2 is the coefficient of R_N, which their values can vary between 0 and 1 depending on the network conditions while satisfying the conditions of Eq. (2). Thus, if w_1 has a value close to 0, the S_N metric will not have much effect on the selection of CH, and instead, the effect of the R_N metric will be significant; and if w_1 has a value close to 1, the effect of the S_N metric on the CH selection will be significant and instead the R_N metric will not have much effect. For example, if

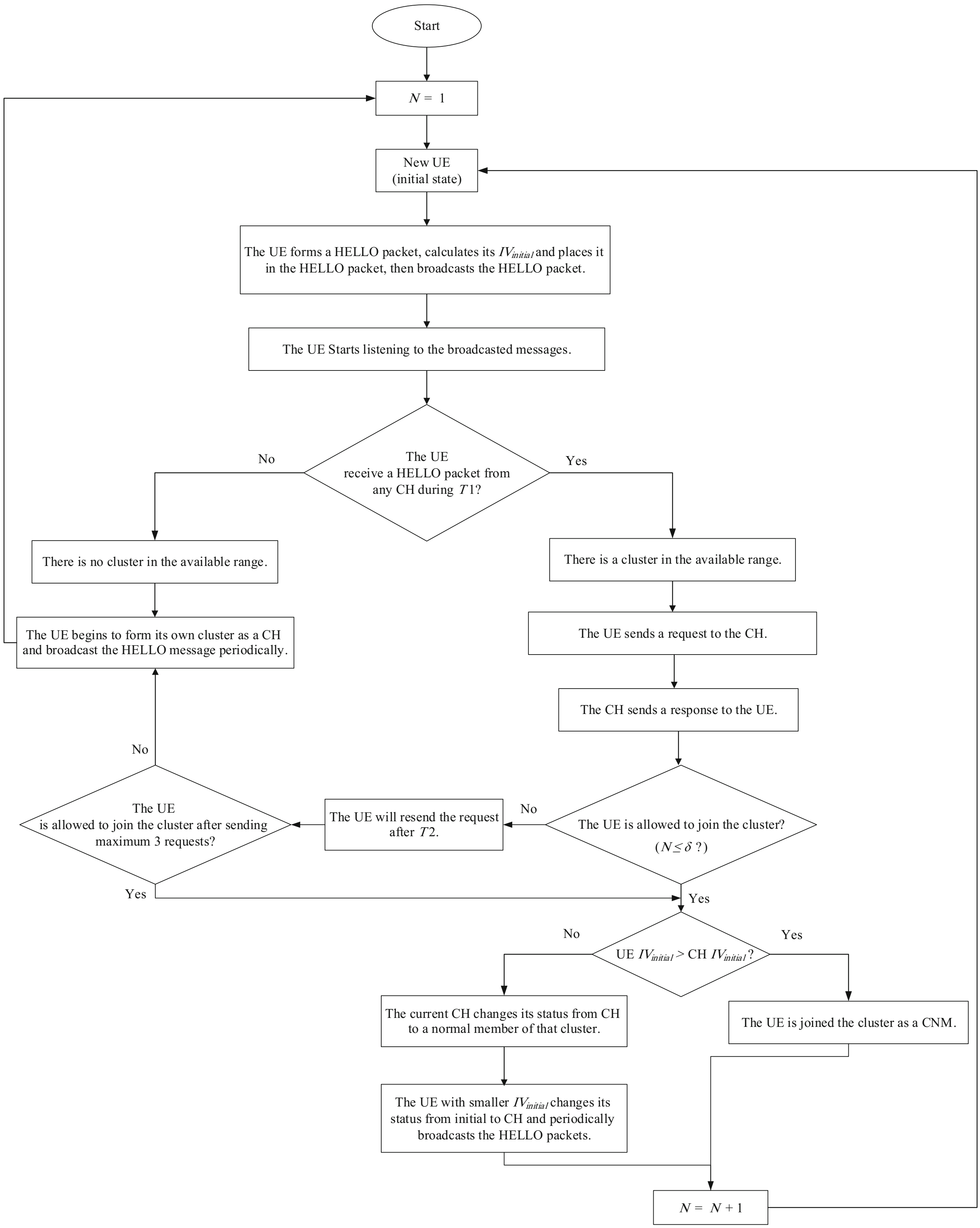


Fig. 5 Cluster formation algorithm

$w_1 = 0$, a number of CNMs with fully charged, very weak channel conditions will be selected as CHs so that the cluster will have the highest lifetime. However, if $w_1 = 1$, then the CNM with the highest SINR and the best channel conditions will be selected as the CH, but the amount of battery charge will not be affected in the CH selection, thus reducing the lifetime of the network. Therefore, it is necessary to balance the value of weight factors and set them according to the network conditions.

It is worth noting that the computational cost of this algorithm is equal to O(K), where K is the maximum number of cell members.

4.2 CH Transfer

Since a CH consumes more battery than a typical node, in order to increase cluster lifetime, all CHs in the network should be changed periodically. This may also happen whenever an alternation occurs in the network (such as CH shutdown or CH movement from one cluster to another one).

As shown in the flowchart of Fig. 6, the proposed CH transfer algorithm for each cluster is as follows. Each member of the cluster calculates its $IV_{periodic}$ and sends $IV_{periodic}$ into a packet in the given time slot to its CH. The CH compares the smallest received $IV_{periodic}$ with its own $IV_{periodic}$ denoted as IV_{CH}. If the smallest $IV_{periodic}$ is greater than IV_{CH}, this means that the current CH has ideal parameters for being CH, so it remains as the CH. Otherwise, CNM with the smallest $IV_{periodic}$ will be selected as the new CH.

Parameter $IV_{periodic}$ is calculated according to Eq. (3):

$$\begin{aligned} &IV_{periodic} = w_3 S_N + w_4 R_N + w_5 V_N + w_6 D_N \\ &w_3, w_4, w_5, w_6 \in [0, 1], w_3 + w_4 + w_5 + w_6 = 1, \\ &\quad S_N, R_N, V_N, D_N \quad \in [0, 1] \end{aligned} \tag{3}$$

where w_3, w_4, w_5, and w_6 are weight factors, in which their values can vary between 0 and 1 depending on the network conditions while satisfying the conditions of Eq. (3). Whatever the weight factor of a metric is closer to 1, the metric will have a greater effect on CH selection, and whatever the weight factor of a metric is closer to 0, the metric will have a lower effect on CH selection. Therefore, it is necessary to balance the value of weight factors and set them according to the network conditions. For example, if in a network the quality of channels and therefore the network throughput are low, we can consider the value of the weight factor w_3 more than other weight factors to improve network throughput. Or if in a network the energy efficiency is important, we can consider the value of the weight factor w_4 more than other weight factors. The proposed scheme improves energy efficiency, lifetime and throughput of 5G networks due to considering S_N, R_N, V_N, D_N metrics as explained in Sect. 4.

It is worth noting that the computational cost of this algorithm is equal to O(δ), where δ is the maximum number of cluster members.

5 Communications Within the CADM Architecture

In this section, we describe the communications within CADM architecture including intra-cluster communications and inter-cluster communications. In intra-cluster communications, communication of simultaneous mmWave connections between the UEs within a cluster is described. In inter-cluster communications, the communications between adjacent clusters and communications between the UEs and the BS are explained.

5.1 Communications Inside a Cluster

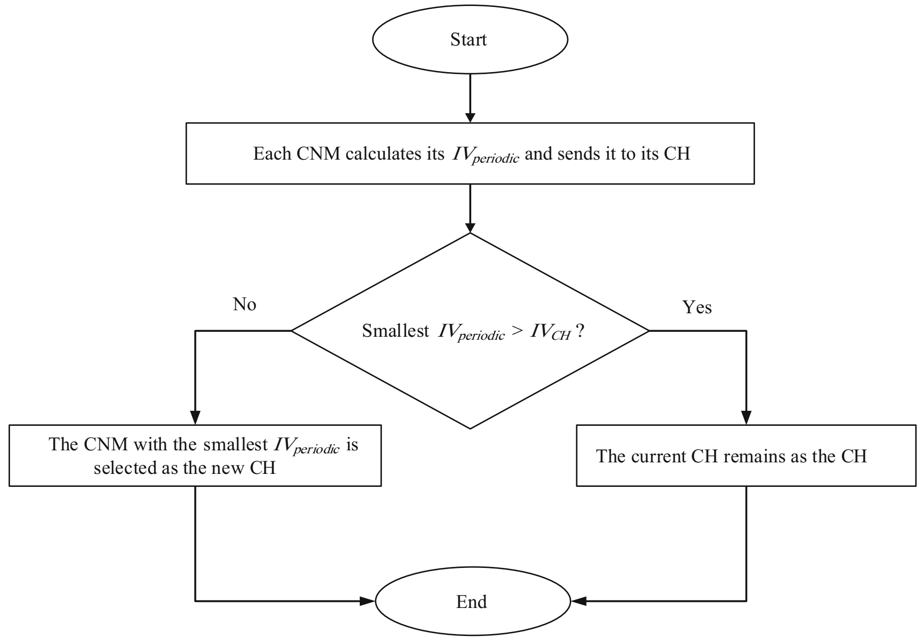
Typically, communications within a cluster, even if two devices are adjacent, are performed via CH and there can be only one communication in each time slot. Thus, establishing the D2D communications causes high collision at low frequencies, due to the use of omnidirectional antennas and small areas of the cluster. This results in higher energy consumption, lower cluster lifetime, and higher end-to-end latency.

As pointed out in Sect. 3.2, the use of directional antennas in the mmWave transmission produces narrow directive beams. This brings a new feature for mmWave systems, which can reduce fading, multi-path and interference [4]. We can use this feature of mmWave frequencies in short-distance D2D communications that allows a UE to communicate directly and without additional hops with its neighboring UE with minimal interference. Using this feature, within each cluster and in each time slot, we can have a D2D mmWave communication between CH and CNM, and several D2D mmWave communications between CNMs on the same frequency. This can increase network capacity, reduce latency, and prevent excessive consumption of the CH energy. Therefore, the BS properly allocates resources to devices asking for data transmission to transmit in the same time slots if simultaneous transmission of them is possible. Figure 7 shows the simultaneous D2D mmWave communications inside a cluster in a same time slot on the same frequency.

5.2 Communications Outside a Cluster

The communications between two adjacent clusters in dense networks are established through the D2D mmWave communications due to small distance between two adjacent CHs. Also, the communications between the BS and the clusters adjacent to the BS are established using the mmWave connec-

Fig. 6 CH transfer algorithm



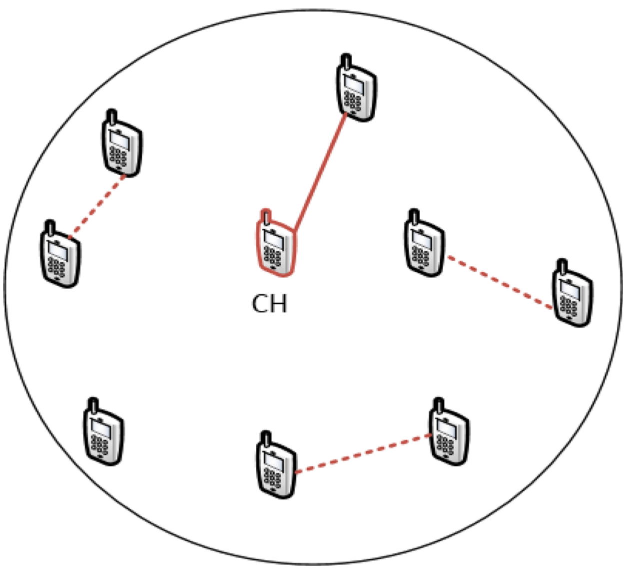


Fig. 7 Simultaneous D2D mmWave communications inside a cluster at the same time slot and on the same frequency

tions. In addition, the communications between the BS and other clusters as well as the communications between the two non-adjacent clusters are performed using cell bands. The connections between cells are also established through traditional cellular communications. Due to the density of the network, it is supposed that communicating between two adjacent CHs is possible.

Details of the communications are as follows. If the source UE intends to transmit to a destination UE within the same cluster, the connection between the two devices is directly established by the D2D mmWave communication without requiring to intervention of the CH. If the source UE intends to send to the destination UE located in one of the adjacent clusters, it sends its data using the D2D mmWave communication to its own CH, and the CH sends the data to the destination CH using the D2D mmWave connection. Finally, the data are delivered to the destination UE. If the source UE intends to send to the destination UE that is not in adjacent clusters, we will no longer use D2D mmWave communications. Because using of D2D connections for long-distance communications results in multi-hop transmission; this causes that a large number of UEs interact in the communication, so that their power consumption is increased and their battery power is decreased. Also, the use of mmWave communications is not suitable for long-distance communications due to the characteristics of transmission at such high frequencies, including signal attenuation because of free space propagation, atmospheric gases, and rain. Hence, traditional cellular communications are used to communicate between two UEs that are not in the same cluster or adjacent clusters. Here, the source UE transmits the data to its own CH and the CH sends the data using the cellular communication through the base station to the destination CH, and then, the destination CH delivers it to the destination UE. Due to the role of CH in communications with the longer distance, we have tried in our proposed clustering algorithm to select the UE with ideal conditions (i.e., the smallest $IV_{initial}/IV_{periodic}$) as a CH.

Note that the UEs within the clusters adjacent to the BS can also communicate with the BS using the mmWave band. Due to the directive and narrow mmWave beams, the connection between a UE and the BS can be established simultaneously with the same-frequency D2D mmWave connections at the

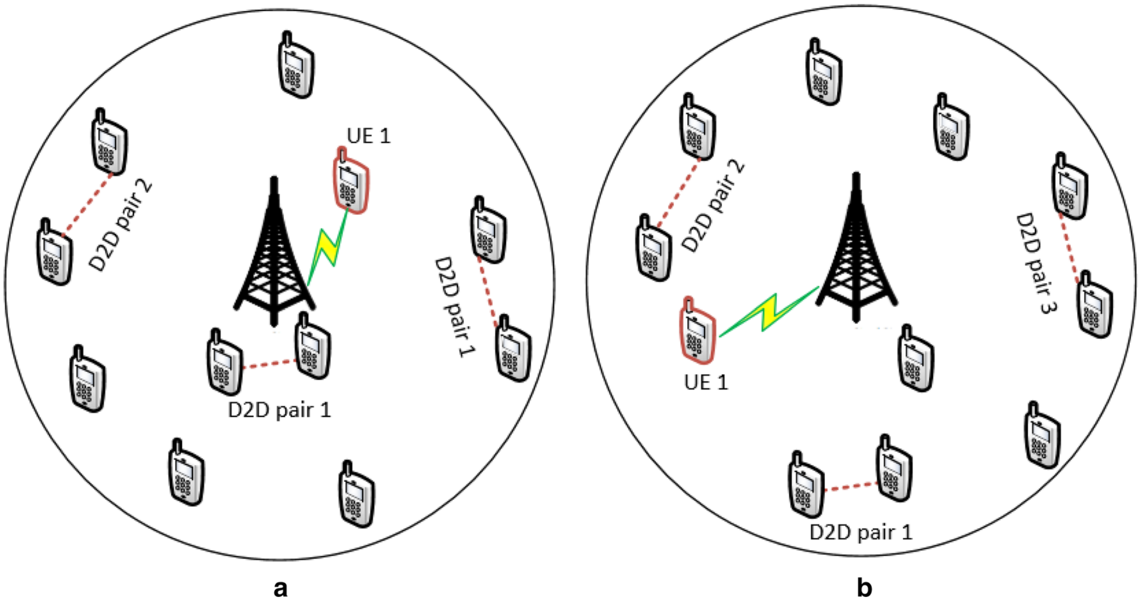


Fig. 8 Simultaneous mmWave communications between two UEs and between UE and BS

same time slots as long as there is no probability of collision between them. For example, in the downlink mmWave transmission in Fig. 8a, the D2D pair 1 located close to the BS can maintain their communication until their connection beam avoids the main lobe of the BS transmission beam [4]. In the uplink mmWave transmission of Fig. 8b, UE 1 can still maintain its connection even if the co-channel D2D pair 2 is in communication because transmission beams of the D2D pair are not directed toward UE 1.

The proposed scheme is highly effective for nowadays dense and high-traffic Internet networks because whenever a new popular video or an updated software is placed on the Internet, all the Internet users tend to buy and download it in real time with a high speed and low latency. Therefore, if one of the UEs has downloaded one or more files such as several high-volume videos or softwares from the BS, it is probable that some UEs within the same cluster or adjacent clusters will also require the same information; these data can be transmitted using D2D mmWave communications in a very high speed and a short time without needing to any BS. This improves performance of the system and reduces the BS load and latency of the network.

6 Performance Evaluation

In this section, we evaluate the proposed architecture (CADM) in terms of data rates, throughput, energy consumption, network lifetime, number of simultaneous connections, spectral efficiency, and end-to-end delay of the system, through the simulations based on the NS-3.25 simulator. The evaluations are performed by comparing CADM with the basic architecture of 5G networks called Basic architecture [20], in which none of the presented technologies for 5G networks have been used, and the Hybrid-Band architecture [25], in which the D2D and clustering technologies have been used. Assumptions for simulations are as follows.

1. There are eight cellular channels with 400 MHz bandwidth and 5 mmWave channels with 2 GHz bandwidth for data transmission, and a control channel for sending signaling information.
2. Each mmWave channel can support up to 10 concurrent connections.
3. Each UE has one omnidirectional antenna and three directional antennas.
4. The initial battery charge and SINR of all UEs follow the uniform random distribution.
5. The channel condition keeps no change for a long time; that is, the SINR is constant for each CNM during the whole survival time.
6. The direction and angle motion of each UE are random, and the speed of each one follows the uniform random distribution with values between 0 and 25 m/s.
7. To balance the value of weight factors, we consider equal values for weights (i.e., $w_1 = w_2 = 0.5$ and $w_3 = w_4 = w_5 = w_6 = 0.25$) in our simulations because of not considering any certain condition. In other words, all four criteria have the same importance for us.
8. To calculate the energy consumption of UEs, the energy model presented in [27] is employed in the simulations.

Table 2 Simulation parameters

Parameter	Value
Number of cells	3
Number of BSs in each cell	1
Cell size	500 m
Cellular carrier frequency	2 GHz
Number of cell channels	8
Cell channel bandwidth	400 MHz
MmWave carrier frequency	30 GHz
Number of mmWave channels	5
MmWave channel bandwidth	2 GHz
Superframe length	32 time slots
Time slot length	100 µs
Packet size	1500 bytes
Max. number of cluster members (δ)	8
Periodic time of CH transfer	30 min
Weight factors in $IV_{initial}$	$w_1 = w_2 = 0.5$
Weight factors in $IV_{periodic}$	$w_3 = w_4 = w_5 = w_6 = 0.25$
Time interval $T1$	100 ms
Time interval $T2$	500 ms
Max power of BS transmission	43 dBm
BS receiver sensitivity	−123.4 dBm
Max power of UE transmission	24 dBm
UE receiver sensitivity	−107.5 dBm
Max energy of UE	100 J
Shadowing	Normal distribution with standard deviation of 8 dB

9. Path loss model is based on Eq. (4) [28]:

$$P_{\text{LOS}} = \begin{cases} 1, & d \leq 1 \text{ m} \\ \exp\left(-\left(d-1\right)/9.4\right), & d > 1 \text{ m} \end{cases} \quad (4)$$

   where d is the distance from the BS in meters.
10. For drawing each diagram, the simulation is carried out 15 times; hence, each point of every diagram is an average of 15 simulations with confidence interval of 95% and maximum error of 5%. The used values for the simulations are summarized in Table 2.

The simulation results are shown in Figs. 9, 10, 11, 12, 13, 14, and 15 and compared for 5G networks with three architectures as the Basic architecture, Hybrid-Band architecture and CADM architecture. Figure 9 shows the average data rate for these three architectures during simulation time with a fixed number of 80 UEs. As shown in Fig. 9, 5G with the Basic architecture has the lowest data rate; the Hybrid-Band architecture, which uses the 2.4 GHz band and D2D communications, has a relatively higher data rate, and the CADM architecture has the highest average data rates among the three schemes, where its highest value reaches to 847 Mbps according to the diagram. This is because of using the high-bandwidth short-distance mmWave communications along with the cellular communications. The reason for the fluctuation of the CADM architecture between the two values of 592 and 847 is the simultaneous use of cellular communications and mmWave communications with two different bandwidths. When the number of mmWave connections with the bandwidth of 2 GHz is higher, the average data rates will be higher, and when the number of cellular connections with the bandwidth of 400 MHz is higher, the average data rates will be lower.

Figure 10 shows the throughput percentage of 5G networks with the three discussed architectures for 20 to 100 UEs, where throughput denotes the successful delivery rates of bits in a given time period. As shown in Fig. 10, all of the three diagrams have the least amount of throughput when there are 20 UEs in the network. This is because of the low demands for establishing connections and not using some of the channels, thus leading to the successful delivery of fewer bits. As Fig. 10 shows, the 5G network with the proposed architecture has the highest amount of throughput percentage, indicating that the proposed scheme is significantly better than the other two schemes. The reasons for this improvement are the ability of high-bandwidth and short-distance D2D mmWave communication, providing a proper medium access control model, and considering the SINR and the speed in selection of cluster heads. This causes a cluster member with the highest quality channel and the lowest speed to perform the highest amount of data transmission and to communicate with the BS.

Figure 11 shows the energy consumption percentage of UEs for a fixed number of 80 UEs for 5G networks with the three referred architectures. As Fig. 11 illustrates, the highest energy consumption of UEs at any time of the simulation is related to the 5G network with Basic architecture; after that, the 5G network with the Hybrid-Band architecture has the highest power consumption and 5G with the CADM architecture has the lowest energy consumption and therefore the highest energy efficiency. This is because of the used metrics (such as the consumed battery power till current time, the mobility, and the sum of distances of the UEs within the cluster to the candidate UE interested in being CH) considered in the proposed clustering algorithm, which dramatically reduces the energy loss of the UEs and increases network energy efficiency compared to the two other schemes.

Figure 12 shows the normalized lifetime of 5G networks. The network lifetime denotes the time span that a network can operate and perform its tasks. As shown in Fig. 12, the 5G networks with the CADM architecture, Hybrid-Band architecture and Basic architecture have the highest lifetime from left to right, respectively. This is directly related to the energy consumption of UEs. In fact, a 5G network with the Basic

Fig. 9 Average data rate

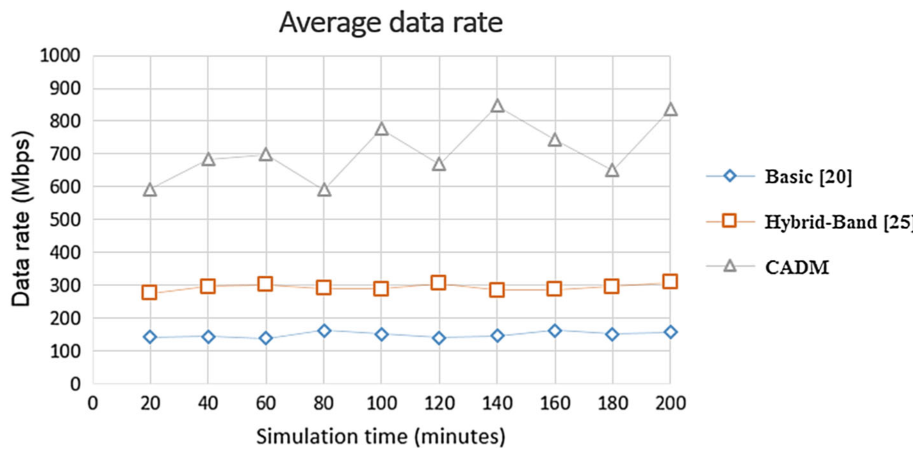


Fig. 10 Throughput percentage

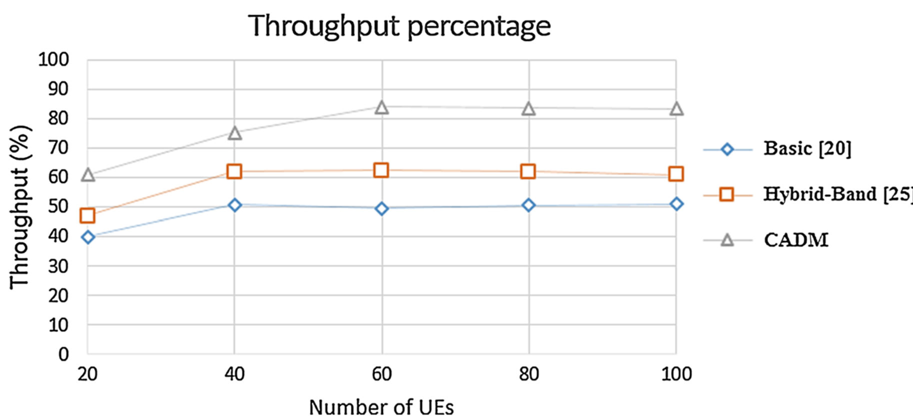
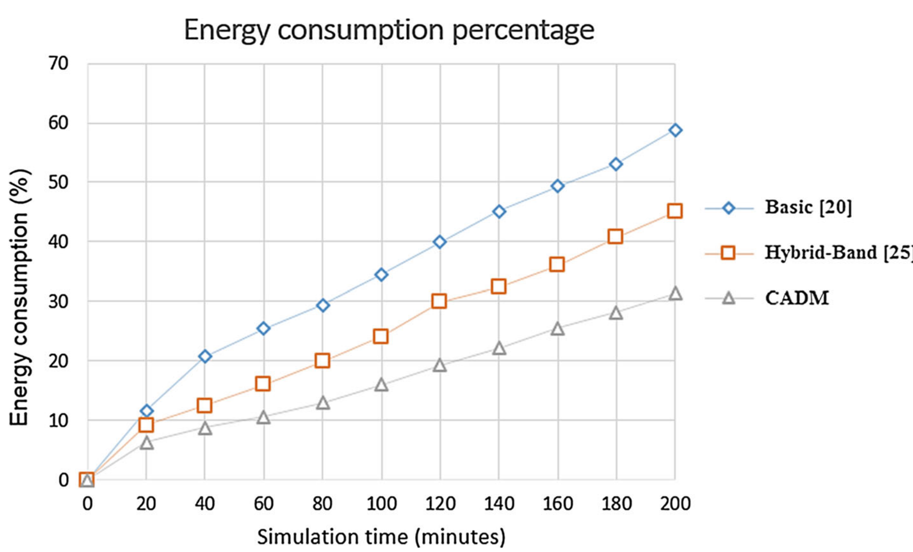


Fig. 11 Energy consumption percentage



architecture, in which the clustering are not used, has the shortest lifetime, and the proposed scheme with the proposed clustering algorithm has the longest lifetime.

Figure 13 shows the number of concurrent connections, assuming that there are 13 distinct channels in the network for the three architectures. As shown in Fig. 13, when there are 20 UEs in the network, all the three architectures have the smallest number of simultaneous connections due to the low traffic demand. In 5G networks with Basic and Hybrid-Band architectures, the number of connections does not exceed from 13 (the number of existing channels on the network) due to the use of cellular bands without the ability to communicate simultaneously. However, under the CADM architecture, when there are 100 UEs in the network,

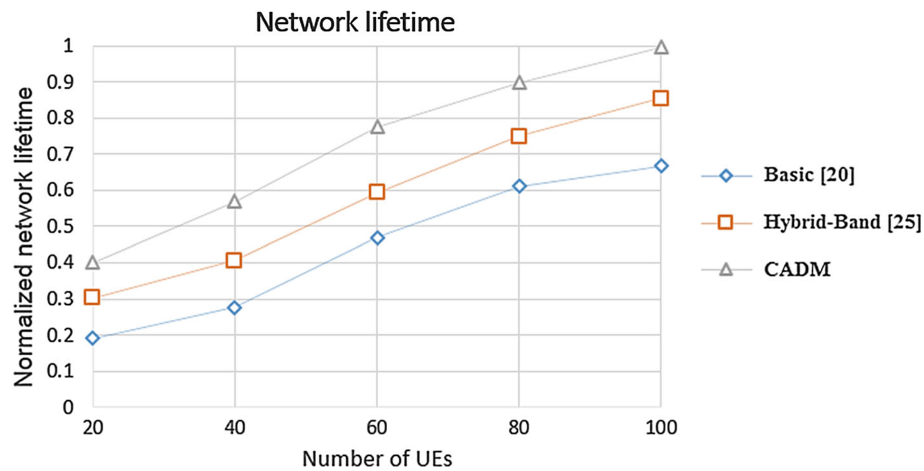


Fig. 12 Normalized values of network lifetime

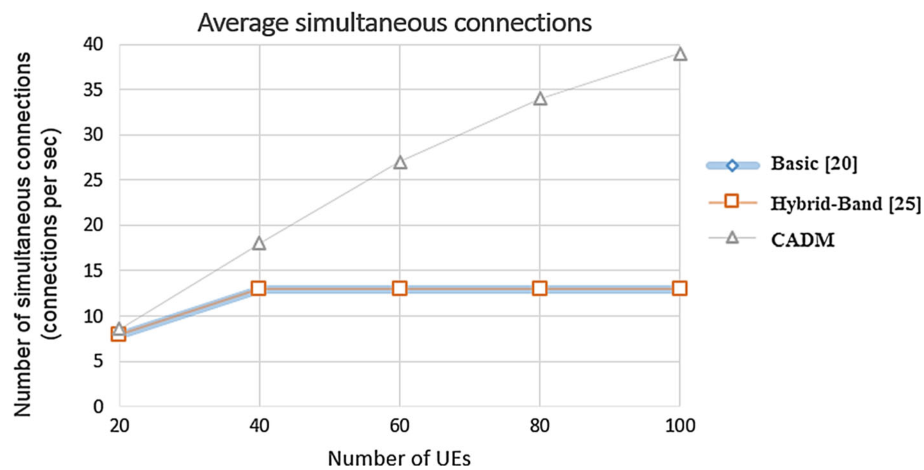


Fig. 13 Average number of simultaneous connections

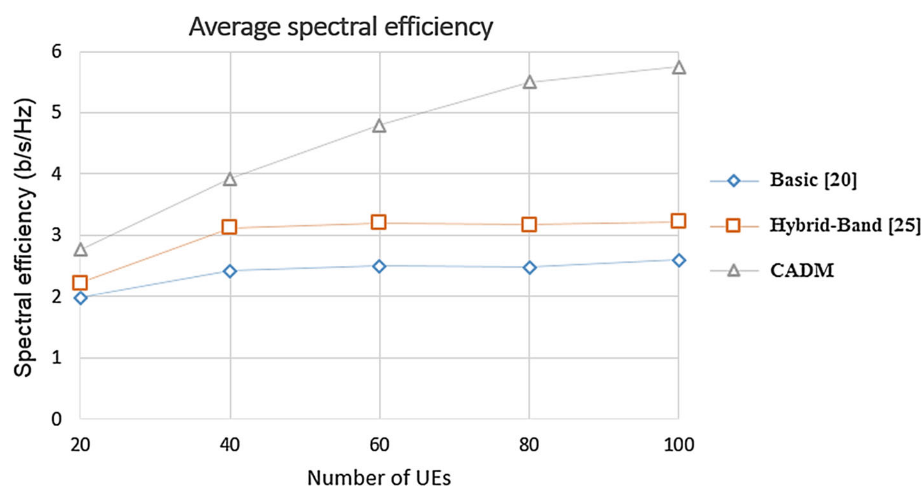


Fig. 14 Average spectral efficiency

the number of simultaneous connections reaches to about 40 connections per second, due to utilizing 5 mmWave channels with the ability of supporting 10 simultaneous connections for each mmWave channel.

Figure 14 demonstrates the average spectral efficiency of 5G networks with the three discussed architectures. The spectral efficiency specifies the efficiency of using a radio frequency spectrum. When there are more than 40 UEs in the network, both Basic and Hybrid-Band architectures have a near-zero slope, due to the constant average data rate and average number of simultaneous connections. But in the CADM architecture, with increasing the number of UEs, the number of concurrent connections and also spectral efficiency increases. As Fig. 14 depicts, the highest

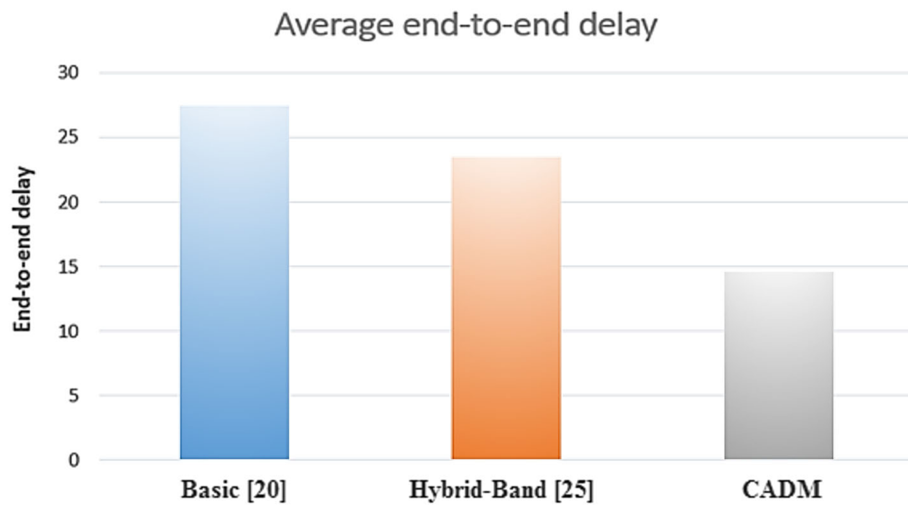


Fig. 15 Average end-to-end delay

spectral efficiency for all UEs is related to the CADM proposal, which has the most efficient use of bandwidth. This is because of possibility of establishing mmWave communications along with cellular communications and providing a medium access control model that allows simultaneous mmWave communications.

Figure 15 displays the end-to-end delay of 5G networks with three Basic, Hybrid-Band, and CADM architectures for a fixed number of 80 UEs. As illustrated in Fig. 15, the most end-to-end delay is related to the Basic architecture, in which cellular communication is used and D2D capability is not used. The lowest end-to-end delay, with a significant improvement compared to the other schemes, is for the proposed CADM architecture, due to the use of the simultaneous and short-distance mmWave communications with high data rates.

7 Conclusion

In order to improve the performance of the fifth-generation technology, a new cluster-based architecture capable for D2D mmWave communication (CADM) has been proposed for 5G mobile networks utilizing mmWave communications along with cellular communications. As the simulation diagrams show, the proposed algorithms for clustering and CH selection in the suggested architecture result in energy savings and increase in network lifetime. Furthermore, providing simultaneous D2D mmWave communications on a frequency in CADM as well as a suitable medium access control protocol can reduce the BS load and end-to-end latency, and increase data rates, throughput, the number of concurrent connections, and spectral efficiency of 5G networks compared to the previous schemes.

References

1. Ramnarayan; Kumar, V.; Kumar, V.: A new generation wireless mobile network-5G. Int. J. Comput. Appl. **70**, 26–29 (2013)
2. Wang, C.-X.; Haider, F.; Gao, X.; You, X.-H.; Yang, Y.; Yuan, D.; et al.: Cellular architecture and key technologies for 5G wireless communication networks. IEEE Commun. Mag. **52**, 122–130 (2014)
3. Tehrani, M.N.; Uysal, M.; Yanikomeroglu, H.: Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions. IEEE Commun. Mag. **52**, 86–92 (2014)
4. Wei, L.; Hu, R.Q.; Qian, Y.; Wu, G.: Key elements to enable millimeter wave communications for 5G wireless systems. IEEE Wirel. Commun. **21**, 136–143 (2014)
5. Alsharif, M.H.; Nordin, R.: Evolution towards fifth generation (5G) wireless networks: current trends and challenges in the deployment of millimetre wave, massive MIMO, and small cells. Springer Telecommun. Syst. **64**, 617–637 (2016)
6. Qiao, J.; Shen, X.S.; Mark, J.W.; Shen, Q.; He, Y.; Lei, L.: Enabling device-to-device communications in millimeter-wave 5G cellular networks. IEEE Commun. Mag. **53**, 209–215 (2015)
7. Gandotra, P.; Jha, R.K.; Jain, S.: A survey on device-to-device (D2D) communication. J. Netw. Comput. Appl. **78**, 9–29 (2017)
8. Radaydeh, R.M.; Al-Qahtani, F.S.; Celik, A.; Alouini, M.-S.: Dynamic downlink spectrum access for D2D-enabled heterogeneous networks. In: GLOBECOM 2017–2017 IEEE Global Communications Conference, pp. 1–7. Singapore (2017)
9. Radaydeh, R.M.; Al-Qahtani, F.; Celik, A.; Qaraqe, K.A.; Alouini, M.-S.: Imperfect D2D association in spectrum-shared cellular networks under interference and transmit power constraints. In: 2018 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–6. Kansas City (2018)
10. Celik, A.; Radaydeh, R.M.; Al-Qahtani, F.S.; Alouini, M.-S.: Resource allocation and interference management for D2D-enabled DL/UL decoupled Het-Nets. IEEE Access **5**, 22735–22749 (2017)
11. Celik, A.; Radaydeh, R.M.; Al-Qahtani, F.S.; Alouini, M.-S.: Joint interference management and resource allocation for device-to-device (D2D) communications underlying downlink/uplink decoupled (DUDe) heterogeneous networks. In: 2017 IEEE International Conference on Communications (ICC), pp. 1–6. Paris (2017)
12. 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. **21**, 2657–2676 (2015)

13. Rappaport, T.S.; Sun, S.; Mayzus, R.; Zhao, H.; Azar, Y.; Wang, K.; et al.: Millimeter wave mobile communications for 5G cellular: it will work!. IEEE Access **1**, 335–349 (2013)
14. Yu, J.Y.; Chong, P.H.J.: A survey of clustering schemes for mobile ad hoc networks. IEEE Commun. Surv. Tutor. **7**, 32–48 (2005)
15. Berkhin, P.: A survey of clustering data mining techniques. In: Kogan, J., Nicholas, C., Teboulle, M. (eds.) Grouping Multidimensional Data: Recent Advances in Clustering, pp. 25–71. Springer, Berlin (2006)
16. Correa, B.A.; Ospina, L.; Hincapié, R.C.: Survey of clustering techniques for mobile ad hoc networks. Rev. Fac. de Ing. Univ. de Antioq. **41**, 145–161 (2007)
17. Abbasi, A.A.; Younis, M.: A survey on clustering algorithms for wireless sensor networks. Comput. Commun. **30**, 2826–2841 (2007)
18. Wunsch, D.; Xu, R.: Survey of clustering algorithms. IEEE Trans. Neural Netw. **16**, 645–678 (2005)
19. Rai, P.; Singh, S.: A survey of clustering techniques. Int. J. Comput. Appl. **7**, 1–5 (2010)
20. Singh, S.; Singh, P.: Key concepts and network architecture for 5G mobile technology. Int. J. Sci. Res. Eng. Technol. (IJSRET) **1**, 165–170 (2012)
21. Gupta, A.; Jha, R.K.: A survey of 5G network: architecture and emerging technologies. IEEE Access **3**, 1206–1232 (2015)
22. Qian, M.; Wang, Y.; Zhou, Y.; Tian, L.; Shi, J.: A super base station based centralized network architecture for 5G mobile communication systems. Elsevier Digit. Commun. Netw. **1**, 152–159 (2015)
23. Zhang, Z.; Zhang, W.; Zeadally, S.; Wang, Y.; Liu, Y.: Cognitive radio spectrum sensing framework based on multi-agent architecture for 5G networks. IEEE Wirel. Commun. **22**, 34–39 (2015)
24. Abrol, A.; Jha, R.K.: Power optimization in 5G networks: a step towards GrEEn communication. IEEE Access **4**, 1355–1374 (2016)
25. Lin, Z.; Gao, Z.; Huang, L.; Chen, C.-Y.; Chao, H.-C.: Hybrid architecture performance analysis for device-to-device communication in 5G cellular network. Springer Mob. Netw. Appl. **20**, 713–724 (2015)
26. Chatterjee, M.; Das, S.K.; Turgut, D.: WCA: a weighted clustering algorithm for mobile ad hoc networks. Cluster Comput. **5**, 193–204 (2002)
27. Pasca, S.T.V.; Akilesh, B.; Anand, A.V.; Tamma, B.R.: A NS-3 module for LTE UE energy consumption. In: 2016 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), pp. 1–6. Bangalore (2016)
28. Haneda, K.; Tian, L.; Asplund, H.; Li, J.; Wang, Y.; Steer, D.; et al.: Indoor 5G 3GPP-like channel models for office and shopping mall environments. In: 2016 IEEE International Conference on Communications Workshops (ICC), pp. 694–699. Kuala Lumpur (2016)