**TECHNICAL UNIVERSITY OF CRETE**

**SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING**

**Information & Networks Laboratory**

# *‘A Device-to-Device Caching Approach for the Alleviation of the Fronthaul Link Traffic in Centralized Radio Access Networks’*

***Emmanouil Sofikitis***

Examining Committee:

Professor Michael Paterakis, Supervisor<br>
Professor Athanasios Liavas<br>
Assoc. Professor Aggelos Bletsas

**Diploma Thesis**

**February 2018**

# Acknowledgments

Foremost, I would like to express my sincere gratitude to all the people in the Department of Electrical Computer Engineering, who offered me their help and their knowledge generously throughout my studies.

I am also very grateful to my supervisor, Professor Michael Paterakis, who trusted me and gave me the opportunity to study by his side for a topic that really fascinates me. I am thankful and indebted to him for sharing expertise, sincere and valuable guidance and encouragement extended to me. Besides my supervisor, I would like to thank the rest of my thesis committee: Prof. Athanasios Liavas and Assoc. Prof. Aggelos Bletsas who took the time and effort to examine my diploma thesis.

I wish to thank all my friends for their support, the knowledge that we exchanged and most importantly for the good and the bad times all these years, because they were always there. I take this opportunity to express my gratitude to Machi, for supporting me every day all these years, especially throughout this thesis, and I wish her all the best for her studies!

Last but not least, I am very grateful to my family for the financial support and the unceasing encouragement they offered me all these years. Without them this thesis and the completion of my studies would never be possible.

Emmanouil Sofikitis

# Abstract

The proliferation of modern mobile devices and the ease of access they offer to the Internet, has led to a burst of traffic through the cellular network. This fact leads to the need and creates the motivation for the evolution of Radio Access Networks. Many researchers and organizations suggest the Centralized Radio Access Network architecture as the successor of the current deployment, even though efforts are still ongoing in order to facilitate such architecture, regarding the alleviation of the huge load that would burden the link between the centralized network processors and the antennas deployed throughout the cellular network. This link is referred to as the fronthaul. Considering the ever-growing demand of multimedia services, users' device level caching appears to be a very promising technique to leverage the traffic offloading of the fronthaul link, while Device-to-Device technology offers the services and functions which facilitate a distribution network between cellular devices. In this thesis, we propose a Device-to-Device caching approach integrated into the Centralized Radio Access Network. According to the proposed approach, some users are storing video files in the local memory of their cellular devices in order to serve video file requests of nearby users. We first describe the main characteristics of the Device-to-Device caching network we consider and we present a system model along with its most important specifications and parameters. Then, we present four case models that we design by gradually relaxing our simplifying system model assumptions and progressively employing improved algorithmic techniques. The performance metric adopted is video hit ratio, since it provides a direct measure of the degree of fronthaul link offloading. Finally, we evaluate and discuss the performance of the above four case models obtained via simulation.

# Contents

# List of Tables

# List of Figures

# Chapter 1
# Introduction

Mobile cellular networks for commercial use have emerged in the late 70's with only some thousands of subscribers. An official standard was developed almost one decade later and it was soon after adopted by most countries around the world [1]. Since then, the amazingly rapid advances in the field of information and networks technology gave the opportunity to the telecommunications industry to create an enormous network of interactive users, devices and applications, reaching the spectacular number of 8 billion mobile devices and connections worldwide. According to Cisco, monthly mobile network traffic in 2016 reached 7.2 exabytes, almost twice as high as in 2015 and is expected to increase sevenfold by the end of 2021 [2]. Although this growth was expected, it has nevertheless strained the current technology expertise and infrastructures, opening new challenges to researchers in the field. Several ideas have been proposed aiming to maintain, or even improve, the high quality of services alongside with the ever growing volume of information accessed daily through the mobile cellular network.

## 1.1. Introduction to Centralized Radio Access Networks

According to the literature concerning the evolution of mobile networks, Centralized (or Cloud) Radio Access Network (C-RAN) architecture poses as the most promising way to overcome the limitations of the existing architecture. Utilizing the current infrastructures and radio resources, the bottleneck expressed by Shannon's law of channel capacity is almost reached and deploying more and even smaller cells to cover the increasing number of devices has become unprofitable for the network operators and aggravating for the environment. In order to meet the increasing demand of mobile traffic for network capacity while optimizing cost and energy consumption, network designers are considering a different approach to the basic principles of the architecture which is used to connect the mobile users to the core network. The Centralized Radio Access Network is a novel mobile network architecture that aims to improve the cell deployment of conventional mobile networks used broadly nowadays, where each cell’s functionalities are controlled by its exclusive Base Station in a distributed manner. The main idea behind C-RAN is to gather all baseband processing regarding multiple cells into one entity that is referred to as the Baseband Unit (BBU) pool. That way a balance of loads between Base Stations is accomplished along with faster processing through resource sharing, since the Baseband Units of multiple cells are cooperating under a single processing unit. Furthermore, reduced power consumption, increased flexibility in network upgrades and adaptability to non-uniform traffic is achievable. Most importantly C-RAN is expected to increase the data rate and the Quality of Service (QoS) experienced by the users. However, despite the numerous advantages that this architecture provides, there are some drawbacks that render C-RAN scheme difficult to employ. In the current thesis we focus exclusively on the difficulties arising on the links between the BBU pool and the Remote Radio Heads (RRHs), namely the fronthaul link, and try to mitigate the traffic amount that burdens this component of the Centralized Radio Access Network architecture.

### 1.1.1. Architecture and transition from conventional Radio Access Network

Modern mobile networks, which support up to 4G technologies and services, are deployed in accordance with the Universal Mobile Telecommunications System (UMTS) model. It is a system which was standardized by the Third Generation Partnership Project (3GPP) work in 1998 under the Rel99 release [3]. Its initial design was supporting technologies of the second generation but, through the years and with the releases that followed, it has been modified accordingly to serve the mobile users, delivering to them the requirements and the features of the latest generations of mobile systems specified by major standardization bodies around the world, like ITU, ETSI, etc. However in order to satisfy the rapidly increasing traffic and the strict requirements entailed by the upcoming generations, it is clear that radical changes in almost every layer and every component of the existing mobile network model is a necessity. A major bottleneck in the improvement of the performance of the network is attributed to the component responsible for the data packet transport from the User Equipment (UE) devices to the Core Network (CN) and vice versa,  i.e., the Radio Access Network (RAN). Consequently research and discussions have been made in order to move forward to a new generation of RAN, the C-RAN, which is expected to deliver high data rates, low latency, improved system capacity, inter-cell cooperation, flexible resource allocation and therefore to surpass the restrictions posed by current networks, like the UMTS Terrestrial Radio Access Network (UTRAN) [4, Ch. 2], [5].

UTRAN is the successor of the Global System for Mobile Communications (GSM) Radio Access Network, i.e., the network that was in charge of the radio access of the UEs in the GSM standard. Based on some basic features of the previous edition of the Radio Access Network, the UTRAN evolved gradually to the current form that is described below.

All functions required for the management and the arrangement of the access of the UEs to the Core Network, along with the transport of data packets between them, are supported by the UTRAN. To achieve such complicated and demanding processes, the UTRAN consists of the following basic components:

- **The Remote Radio Head:** The RRH is embedded with some basic radio signal processing units, like RF filter and Power Amplifier.
- **The radio interface:** The wireless link that connects the User Devices with an antenna.
- **The Baseband Unit:** The BBU is a module that contains the digital assets needed to perform all baseband processes necessary, like coding, modulation, Fast Fourier Transform (FFT), etc. In addition it is responsible for data link layer functionalities, like MAC.
- **The link between RRHs and BBU:** In initial designs this link was supported by a coaxial cable, however in later releases it was replaced by a digital fiber optical link.
- **The backhaul link:** This link connects the Radio Access Network with the Core Network which serves as a getaway to external networks.

Several BBUs are controlled by a Radio Network Controller (RNC) which is an essential node connected to the Core Network, implementing network layer functionalities, however in this thesis for simplicity we do not assume the existence of a RNC and we will focus on the architecture described above. More detailed introductions on the fundamentals of UMTS can be found in [3], [4, Ch. 3] and [6].

In an early deployment of the UTRAN, the antenna and the BBU unit were collocated at a tower, called the Base Transceiver Station (BTS) or Base Station (BS) for simplicity. The overall baseband processing and the radio functionalities were operated exclusively by the BBU sitting at the foot of the tower, while at the top of the BS, the antenna served only as a transceiver of the signal processed by the BBU. As a result of the replacement of the coaxial cable connecting the antenna and the BBU with high throughput broadband optical fiber cables, the BBU and the RRH were decoupled and in modern networks they can be separated by up to several kilometers. This modification is significant and induces several advantages, yet it is not sufficient to meet the needs of future mobile networks. Figure 1.1 shows an example of the UTRAN architecture.

The visionaries of the C-RAN architecture are exploiting the abovementioned separation of the radio functionalities and the baseband processing with an even more efficient advancement. The basic idea behind C-RAN is for several BBUs to be gathered in a central station, creating thus the aforementioned BBU pool. A BBU pool is a virtualized cluster which can consist of general purpose processors, installed at the same physical location where the functions of all telecommunication layers (physical, data link and network layer) are gathered, in order to serve several cell sites. A BBU pool can serve up to hundreds of cell sites, by managing and controlling a few thousands of RRHS. This way there is a change from the traditional static dedication of a BBU to certain RRHs to a more dynamic allocation of resources. This approach allows under-provisioned resources and infrastructures to be utilized at their peak capacity, improving the performance of costly and energy insufficient elements in the existing network. There are also major benefits by the C-RAN architecture in terms of network capacity, scalability for future network expansion, energy consumption, carbon footprint, capital and operating expenditures [5]. An initial concept of C-RAN was first introduced by IBM in 2010 [7] and it was further described by China Mobile Research Institute in 2011 [8]. Although it has not been yet standardized by IEEE, China Mobile has already started proprietary C-RAN networks deployment [9]. Figure 1.2 shows an example of a possible C-RAN deployment.

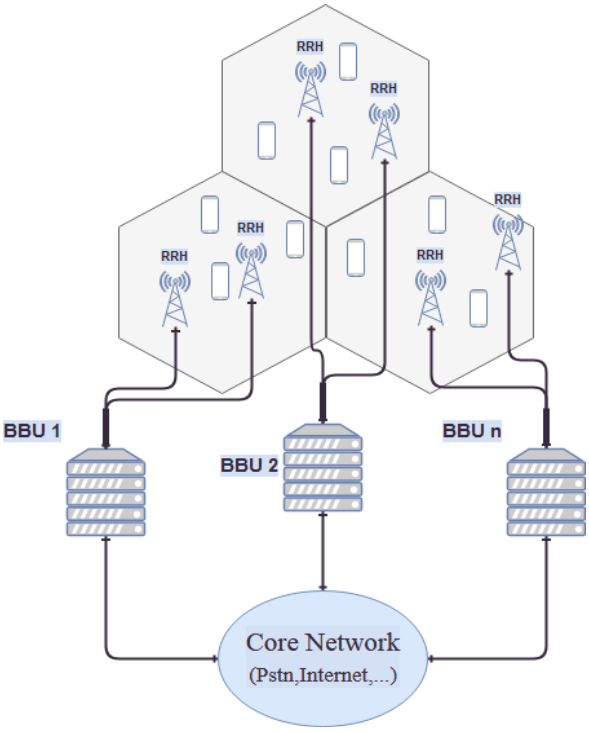

**Figure 1.1 : UTRAN architecture**

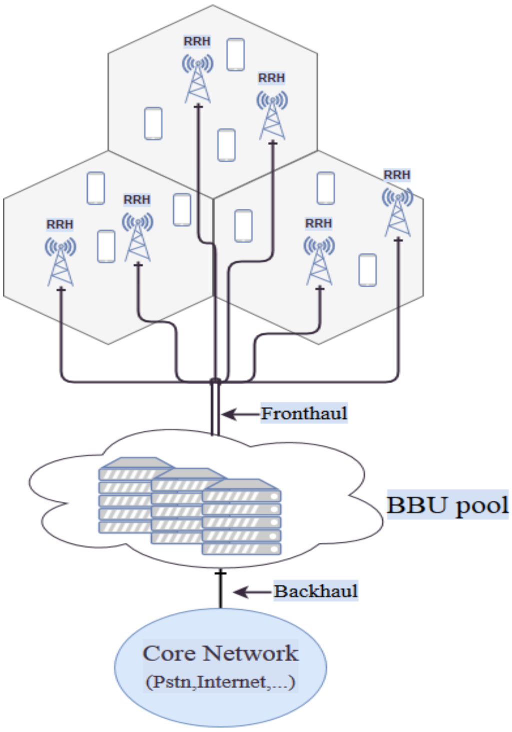

**Figure 1.2 : C-RAN architecture**

### 1.1.2. The fronthaul link challenge

With the advent of the new configuration in the C-RAN architecture where BBUs are centralized and RRHs are distributed at remote cell sites, there is a new transport segment established between them referred to as the fronthaul (FH). The fronthaul link is considered to be the barrier towards the transition of the current architecture to the promising C-RAN architecture and for this reason significant efforts and research are in progress, in order to create solutions that would enable the deployment of the C-RAN architecture.

The fronthaul network transports the digitized signals aggregated from numerous cell sites, which results in a huge amount of data flowing from and to the BBU pool. In addition to the data that a BBU would exchange with a RRH in a traditional UTRAN implementation, the fronthaul link is burdened with the transport of the aggregated data of functions of all telecommunication layers and the vital information necessary to achieve the coordination between multiple cells. Unlike the current installation of RAN, in which the functions of the physical layer and the data link layer are performed in a distributed fashion in each cell by the dedicated Base Station, in a centralized architecture all functions will be gathered in the BBU pool. This approach offers all the benefits discussed earlier in this chapter, but makes fronthaul requirements, in terms of latency and throughput, more stringent [5]. Consequently, although there are already installations of BBUs separated from their dedicated RRHs, the conversion of such installations to a Centralized-RAN will create additional traffic on the link connecting RRHs to the BBU pool.

Furthermore, the requirements for the latency and jitter tolerance and the overall delivered QoS are increasing in the light of new generations of wireless mobile telecommunications technology. In Rel14 [10] the 3GPP committee includes discussions for several use cases of the upcoming 5G mobile networks and in [11] the Next Generation Mobile Networks (NGMN) Alliance presents an outline of its envisioned 5G outlook. The requirements described for various scenarios and enhanced services are highly demanding and are calling for an optimized architecture of the mobile network. Table 1.1 below presents a comparison of 4G and 5G system requirements regarding the data rates provided per mobile user and the end-to-end latency experienced by them.

| **Attribute** | **4G** | **5G** |
|---|---|---|
| Data rate (per user) | Up to 100 Mb/s on average<br>Peaks of 600 Mb/s | • 10× expected on average and peak rates<br>• 100× expected on cell edge |
| End-to-end latency | 10 ms round-trip<br>Up to 50 ms round-trip tolerance | 10 times faster<br>Technology should allow operators to optimize topology to achieve 1 ms end-to-end for ultra-low latency services. |

**Table 1.1 : Comparison of 4G and 5G system requirements**

Aggregated mobile network traffic from multiple cells along with the high QoS standards expected for future generations of cellular networks are the main reasons that cause a huge overhead on the links between RRHs and the BBU pool in C-RAN architecture. Reducing the traffic sent via the FH link is an outright way to facilitate a possible C-RAN deployment. To reduce the fronthaul traffic, a principal method is to push the cached contents further, i.e., caching contents into the “front” of the fronthaul. Therefore, in this thesis we consider caching multimedia content in the devices of cellular network users.

## 1.2. Introduction to Device-to-Device Caching Networks

### 1.2.1. Basic Principles of Device-to-Device communications

Device-to-Device (D2D) communications refer to a set of functions which enables the direct exchange of traffic between two cellular devices without routing through the cellular network. D2D communication has been standardized by the 3GPP under Release 12 [12] and it is referred to as Proximity Services (ProSe). The standard is designed to be integrated to the functions of the Evolved-UTRAN which is the primary Radio Access Network deployed nowadays. The two basic features of ProSe are the D2D discovery and D2D communication functions. D2D discovery refers to the ability of a cellular device user to discover, without relying on the cellular network, other users that are in proximity. D2D communication refers to the direct exchange of data packets between two cellular devices that are in proximity. Here, proximity should be understood in a broader sense than just physical distance (vicinity). It may also be determined based on, for example, channel conditions, signal-to- interference-plus-noise ratio (SINR), throughput and delay.

D2D communication can be divided into two major categories regarding the radio spectrum utilized for the exchange of traffic between two UE devices: inband D2D communication and outband D2D communication. Inband D2D communication occurs on licensed spectrum, meaning that both D2D and cellular links utilize the same frequency spectrum. Inband D2D communications occur on either uplink or downlink radio resources. All D2D communication functions are controlled and supervised by the cellular network, thus creating the opportunity of fully utilizing the licensed spectrum. Depending on the techniques used to share the available licensed spectrum between cellular and D2D links, inband D2D communications are divided into underlay and overlay. However, the approach of inband D2D communication could result in performance degradation of the cellular network because of the interference created between cellular and D2D links. Outband D2D communication exploits unlicensed spectrum. Unlicensed spectrum usually refers to frequency bands that are reserved for Wireless Local Area Networks (WLANs) or for industrial, scientific and medical purposes, also known as ISM bands. With respect to the degree in which the cellular network is involved in the D2D communication functions, outband D2D is further divided into controlled and autonomous communication. In autonomous communication the discovery of nearby devices, the process of setting up a D2D link and the direct communication of two UE devices are entirely performed by the UE devices in a distributed manner, whereas in the case of the controlled communication the UE devices rely on the network for several functions concerning their communication except from the actual traffic exchange. Since in this case D2D and cellular communications occur on different spectrums, there is no interference between them and thus an outband D2D communication setup is a very promising choice. However, some researchers believe that the interference in the unlicensed spectrum will be hard to manage, because the cellular network has limited control on the D2D functions, especially in the autonomous D2D communication case [13], [14].

From a technical perspective, exploiting the natural proximity of communicating devices may provide multiple performance benefits. D2D-enabled UE devices may enjoy high data rates and low end-to-end delay due to the short-range direct communication. In addition, when two UE devices communicate directly with each other rather than routing through the RAN and possibly the CN, there is an obvious saving in network resources. Consequently, this can lead to cellular traffic offloading and congestion alleviation which could prove beneficial to the cellular communication of non-D2D UE devices. Moreover, in inband D2D communication, there is a reuse gain of cellular radio frequency resources throughout an area, since D2D communication occurs on the same spectrum as cellular communication and within small distances [15]. Because of the short-range links used in D2D communication a frequency reuse gain is also implied in outband D2D communication, when D2D pairs use the same channel in an area. In Table 1.2 we show a list of different short-range radio technologies that can be possibly used for D2D communication together with their performance and range capability [14]. FlashLinQ is a recent technology proposed by

Qualcomm, which is designed to perform D2D discovery and link scheduling functions on the licensed spectrum. The interested reader is referred to [16] for further details.

| Technology | Range (m) | Bitrate (Mbps) |
|---|---|---|
| Bluetooth | 20-40 | 2 |
| Wi-Fi | 30-50 | 30 |
| FlashLinQ | 50-500 | 50 |

**Table 1.2 : Comparison of short-range radio technologies for D2D communication**

### 1.2.2. Features of Device-to-Device enabled Caching

According to Cisco [2], video traffic accounted for the 60% of the total mobile traffic in 2016 and it is expected to reach 78% by 2021. In addition, it is observed that video file popularity is very unevenly distributed, meaning that a small portion of the most popular video files accounts for the majority of the video traffic. More specifically in a survey contacted in 2007 that used traffic information collected from YouTube, the results showed that 10% of the top popular video files distributed by YouTube accounted for the 80% of the total views in the same site [17]. Considering these facts caching seems a very promising solution in order to better serve this high demand in video traffic. By exploiting the advent of D2D communication services, the high storing capacity offered in cellular devices and the close proximity of UE devices, especially in areas with high user density, we propose a users' device level caching approach using Device-to-Device proximity services in order to mitigate the limited fronthaul capacity problem by replacing fronthaul capacity with the storage capacity of cellular devices.

UE devices participating in the D2D caching network serve as caching nodes, thus creating a distributed cache in an area. As the number of users increases in an area so is the size of the distributed cache and the number of possible D2D links in the same area. This way the increasing density of users that is straining the current infrastructures is exploited as a benefit in a possible D2D caching implementation. Furthermore such an implementation utilizes existing resources, like the very large storage capacity of cellular devices and enables the reuse of radio resources, which makes D2D caching a very popular candidate for enhancing the performance of the cellular network [18], [19], [20].

## 1.3. Related Work

To our best knowledge, prior work that has proposed an integration of D2D communications into the C-RAN architecture is only presented in [9] and [21]. However both studies do not consider caching at the end users as a means of load alleviation of the cellular network. The authors in [9] propose a C-RAN-based Device-to-Device networking in order to alleviate the fronthaul delay confronted in the C-RAN architecture. A set of protocols and interfaces that would enable the proposed C-RAN-based D2D architecture are presented along with a discussion of the benefits of such a combination in terms of end-to-end delay, throughput and energy consumption of the system. In [21] a D2D underlaid downlink C-RAN system is considered, in order to improve the spectral efficiency in the C-RAN. The main focus of this work is the modeling of the interference caused to downlink cellular users by D2D links and the benefits of the D2D communications in improving the licensed spectrum utilization.

There is a lot of research done regarding a user's device level caching approach along with the integration of D2D into cellular networks in order to improve their performance; in this thesis we refer to [18], [19], [20] regarding this approach. However none of the studies considering such architecture has, to the best of our knowledge, taken into consideration the unique characteristics of the C-RAN.

The authors in [18] propose the collaboration of the cellular network with a controlled outband D2D communication model where each UE device stores one video file in its local memory which afterwards is forwarded to other UE devices if certain conditions are met. The goal of this study is to find a good tradeoff between the number of possible parallel D2D links in a cell and the probability that a UE device can find the requested content in its vicinity. To examine this tradeoff the cell of an assumed cellular network is divided into squared clusters inside which clusters only one active D2D link is permitted. In [19] an inband D2D network underlaying the cellular network with UE devices serving as caching nodes is considered in order to offload traffic from the cellular network. Interestingly, the D2D links are established without the help of the cellular network. Instead, every UE device is equipped with cognitive technology in order for the UE devices to sense the available cellular channels and occupy them for the exchange of traffic. The main goal of this study is to model the priority based cognitive channel access and the availability of content in D2D caching nodes in order to assess the performance of such architecture. Finally, the work in [20] examines a network where some UE devices can store video files in their local memory and distribute them on demand to nearby UE devices. This study emphasizes mainly on the optimization of the caching policy in a way that the D2D network distributing the video files can achieve the maximum possible performance.

## 1.4. Thesis Goal and Contribution

As we have already mentioned, our main goal is to reduce the traffic overhead on the fronthaul link in order to facilitate a possible C-RAN deployment. To accomplish this goal we have designed a novel architecture which combines the upcoming C-RAN architecture with a D2D caching network. Some features and details of the proposed D2D caching network were based on the work in [18], [19] and [20].

The idea of introducing a D2D caching network to enhance the C-RAN architecture was inspired by the work in [22] where the C-RAN architecture is described along with the fronthaul link challenge and several ideas are proposed in order resolve it.

Initially we present a controlled outband D2D communication network, similar to the one in [18], where some UE devices choose to serve as caching nodes by storing one video file in their local memory and distribute it on demand to other UE devices in the same area. Unlike the studies in [18] and [20], we model a network which covers a large area (as large as the coverage of a single BBU pool) inside which there are many UE devices requesting video files and several D2D links can be established in order to serve these requests. Contrary to the work in [20], we employ a very simple admission policy for choosing which video files to cache. More specifically, the cache-enabled UE devices decide which video file to cache according to the video files popularity, meaning that the probability that a specific video file is going to be cached by a cache-enabled UE device is set equal to the probability according to which the same video file is going to be viewed by a user. Finally, in order to evaluate the performance of the proposed network we create four variations of our architecture based on the interference models that we employ, the portion of the video file library we choose to cache and the ability of our architecture to adapt to changes in the popularity distribution of the video files.

Since the architecture we propose is based on the control of some important D2D communication functions by the cellular network, it is important in terms of latency and workload, that these functions are controlled by the same processing entity, i.e., a single BBU pool. By expanding the modeled area of the proposed architecture we are exploiting the main advantage of the C-RAN, which is the simultaneous management by the same entity of multiple cells in a cellular network deployment.

## 1.5. Thesis Outline

The remainder of this work is organized as follows. In Chapter 2 we present and discuss the proposed D2D caching network. Initially in Section 2.2 we describe the main features and specifications of our system. Section 2.3 presents the characteristics and the parameters of our simulation along with four examined case models derived from the initial D2D caching network. In Section 2.4 we present and discuss the results of our simulations.

Finally, in Chapter 3 the conclusions of the work in this Thesis are presented. In Section 3.1 we discuss the overview and the contribution of our work, while in Section 3.2 we present some ideas for future work.

# Chapter 2

# A C-RAN deployment with D2D Caching

## 2.1. Introduction

A successful D2D controlled communication between two UE devices that are associated with different Base Stations requires the cooperation of the two Base Stations for the synchronization of the devices and the D2D transmission scheduling. This fact would make the already stringent requirements of the upcoming 5G standard, regarding the end-to-end latency, even more difficult to attain [15], especially if you take into account the fact that the links between Base Stations are meant for more latency-tolerant functions. Therefore, a possible D2D caching network would benefit from the unique feature of the C-RAN, i.e., the ability of a BBU pool to singlehandedly control all UE devices in a multi-cell environment. Consequently, the proposed architecture is designed accordingly so that it can be extended to a geographic area which is controlled by one BBU pool in a C-RAN deployment, so that all functions required for a D2D link setup, such as synchronization between devices and interference coordination will be carried out by a single processing entity.

## 2.2. System Description

The distance between a RRH and the BBU pool is constrained to 15 km, because of the latency of the link connecting them. According to this constraint on the fronthaul link distance, the theoretical upper limit of the area of coverage of a single BBU pool is taken to be 36 $\mathrm{km}^2$ [23]. In the current study we conduct our simulation over an area of 1 $\mathrm{km}^2$, but, as mentioned above, it can be extended to any area confined to the coverage capability of a single BBU pool.

Over the selected area there are randomly placed users, equipped with UE devices capable of connecting to the Internet through the cellular network as well as connecting to other devices in the area in pairs using interfaces like Bluetooth, Wi-Fi, etc., meaning that all UE devices are D2D enabled. We assume that all users are static. Some users, a certain proportion of the total number of users, decide to cache video files in their local storage and transmit them to other users upon request, thus becoming cache-enabled transmitters. This way a distributed cache is formed throughout the area, capable of serving video requests from UE devices, which requests would otherwise be typically served by the cellular network.

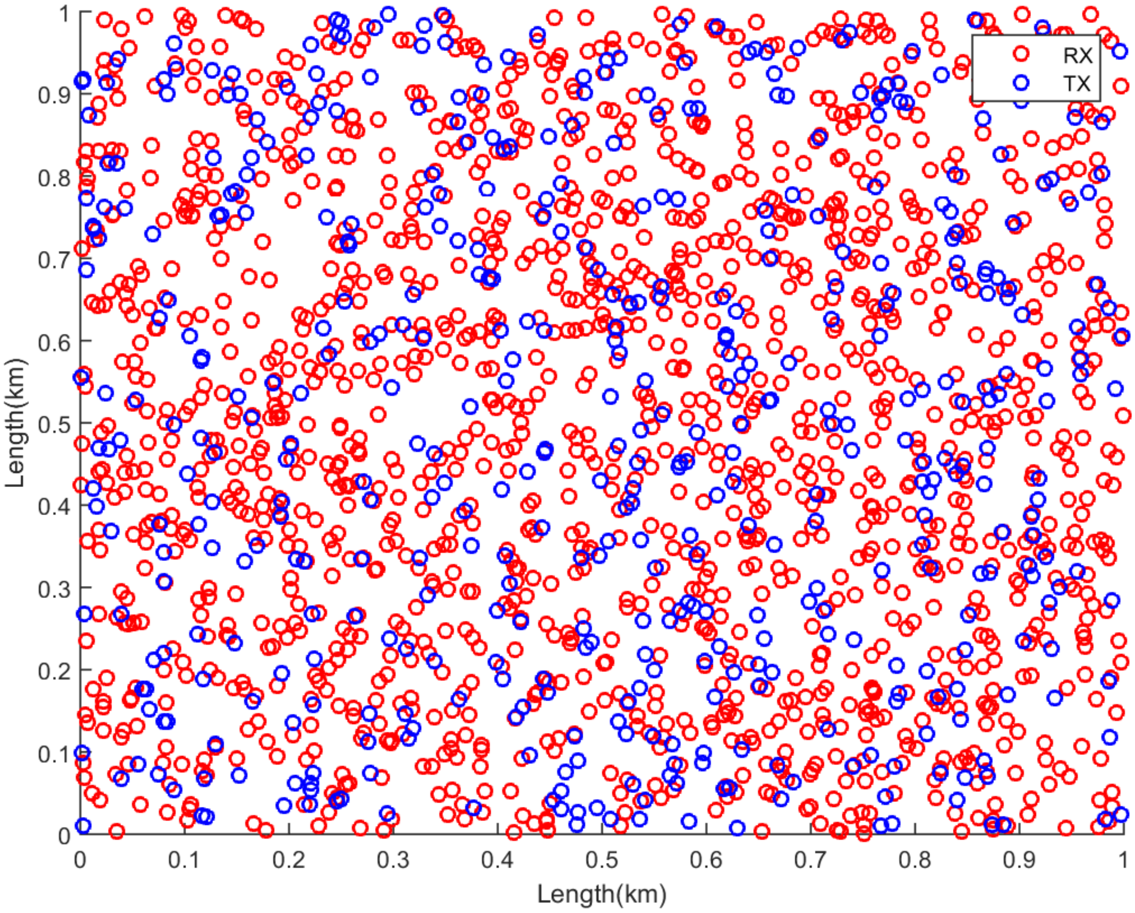

**Figure 2.1 : A possible placement of UE devices**

The distributed cache is populated with files by the applicant transmitters. Caching of a video file may occur when a cache-enabled UE device generates a request for watching a video file or at times with low request traffic in the network. Even though storage capacity in modern cellular devices is the fastest increasing component, capable of storing even hundreds of video files, we assume that every UE device caches only one video file. Nevertheless, this study can be extended to the case in which UE devices can cache multiple video files. All D2D operations are transparent to the users, who simply enter a request for a uniform resource locator (URL) at their UE device, whenever they want to watch a video file. There is a list containing all video files available to users. We will refer to this list as the actual library. A cache-enabled transmitter can also generate requests that will potentially be served by other transmitters or even by its local storage, in the case where the video file cached by a UE device corresponds to the video file requested by its user. In the latter case the request is served immediately without burdening the network or any transmission in which this specific UE device may currently participate. A cache-enabled transmitter is abbreviated by TX and a receiver by RX. Figure 2.1 illustrates a possible placement of UE devices in an area of one squared kilometer.

The BBU pool is in charge of scheduling every transmission by pairing a requesting UE device with a suitable cache-enabled transmitter. For this purpose, the scheduler stores the location of every UE device and a list of the video files cached by each TX. Furthermore, there is a record of all active D2D links which is updated every time a D2D communication is initiated or when a transmission between two UE devices is completed. When a request for a video file is generated from a UE device, it is received by the BBU pool. Upon receiving a request, the BBU pool examines the possibility that the request can be served by the distributed cache via the D2D network. If there is a TX with the requested file stored in its local cache, in the vicinity of the requesting UE device (defined by a radius r around the UE device), the BBU pool marks this pair as a potential link. Provided that this condition is met the BBU pool examines the subsequent condition, regarding the interference between active D2D links. If no other D2D communication is active in this area around the requesting UE device, the potential link is qualified to become active. The D2D communication session is set up by the BBU pool, by directing the requesting RX at the qualified TX. The D2D link remains active until the transmission is over. In case more than one TX is qualified to serve the request, the closest to the requesting UE device is chosen. If the conditions mentioned above are not fulfilled, the request is served via the cellular network. In the following section we will further describe these conditions, which must be satisfied in order to establish a successful transmission.

We assume that the time and the processing resources required for the management of the D2D pairing by the BBU pool are negligible, since the BBU pool is equipped with powerful processors. We further assume that the size of the data sent over the fronthaul link, for the coordination of the D2D communications, is also negligible relatively to the size of the video files.

We assume outband controlled D2D communication over the same channel (single channel network) for every transmission between D2D enabled UE devices. Our model emphasizes functions concerning the media access and the network layer of the telecommunication system; therefore we are interested in studying the interference between links in a simplified manner. Particularly, we use an approach to model interference by characterizing its effects, i.e., whether an active D2D link interferes with another D2D link or not. For this reason we use the Protocol Interference Model formalized by Gupta and Kumar in [24]. The Protocol Interference Model assumes a simple deterministic path loss model and that all terminals transmit at the same power, conditions that are suitable for our study. According to this model, all UE devices employ a common *range* $r$ for their transmissions. When UE device $TX_i$transmits to UE device $RX_i$over a specific channel, this transmission is successfully received by $RX_i$ if

i) the distance between $\mathrm{TX_i}$ and $\mathrm{RX_i}$ is no more than r, i.e.,

$$d(TX_i, RX_i) \leq r \text{ , and} \tag{1}$$

ii) for every other node $\mathrm{TX_k}$ simultaneously transmitting over the same channel

$$d(TX_k, RX_i) \geq (1 + \Delta)r \text{ ,} \tag{2}$$

where $d(TX_i, RX_i)$ and $d(TX_k, RX_i)$ represent the distance between the UE devices $TX_i$ and $RX_i$ and between the UE devices $TX_k$ and $RX_i$, respectively, and $\Delta > 0$ represents the spatial protection margin, which defines an extra guard zone around receivers to guarantee successful transmissions [25].

Figure 2.2 illustrates a case where a potential link is qualified and the D2D communication is initiated. Four TXs, with the requested video file in their local storage, are located in the vicinity of the requesting UE device and the closest is chosen to serve the request. It is important to notice that all other active TXs are not located inside the interference guard zone of the requesting RX. On the contrary, Figure 2.3

illustrates a case where a potential D2D link is not qualified because the requirements of the Protocol Interference Model are not fulfilled. Even though, there is a potential TX in the vicinity of the requesting RX, the D2D communication is rejected by the BBU pool, because an active TX is located closer than the tolerable distance defined by the Interference Protocol Model.

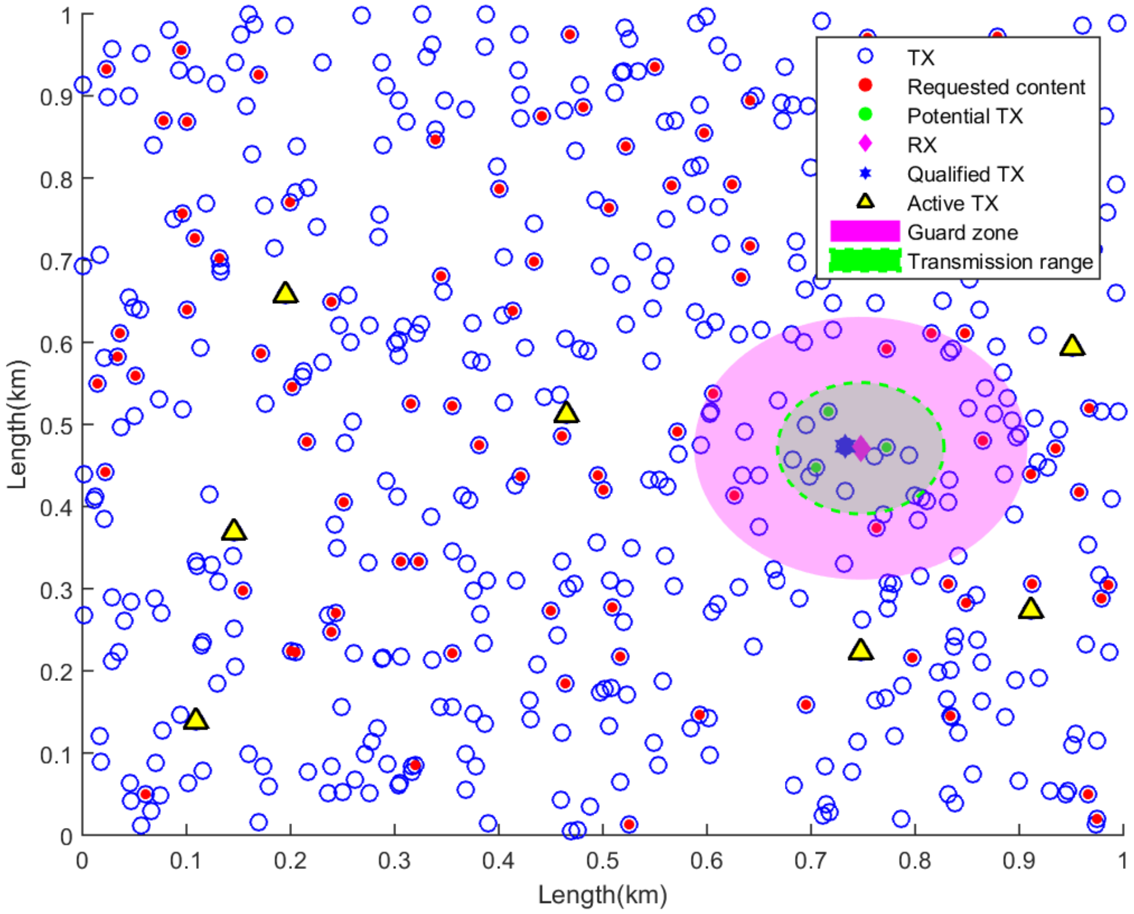

**Figure 2.2 : Qualified D2D transmission**

The value of the transmission range, employed by every UE device, is selected accordingly to achieve high network performance. Smaller radius implies that every transmission has a smaller interference zone; therefore more links can be simultaneously active inside a specific area. However, as the transmission range of a UE device is getting smaller, the probability to find a TX, with the requested file stored, in its vicinity is decreasing. Therefore, it is important to find a good tradeoff between the number of possible parallel links in the area and the probability of finding the requested content within the vicinity of a requesting UE device. For brevity we denote this tradeoff by interference-AoC (Area of Coverage) tradeoff. Through our study, we observed that the transmission range that needs to be employed in order to achieve higher network performance depends on the popularity of the video files and more specifically on the skew coefficient of the Zipf distribution, used to model the popularity distribution of the video files. We assume that the exact parameters of the Zipf distribution are known to the BBU pool (or that they can be estimated by the BBU pool) and thus the desired range for every UE device can be chosen accordingly. In such case, the BBU pool informs the UE devices to employ the suitable transmission range by broadcasting the corresponding value.

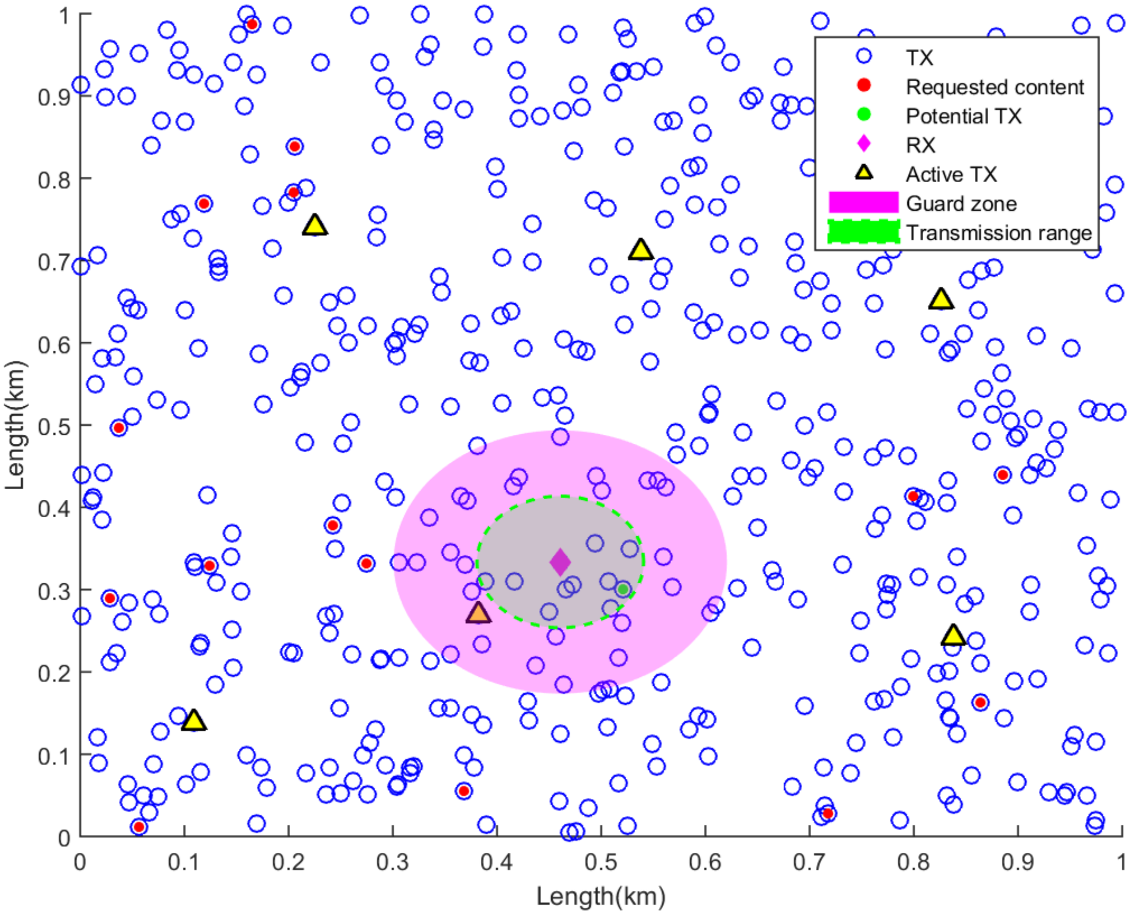

**Figure 2.3 : Disqualified D2D transmission**

## 2.3. Performance Evaluation

### 2.3.1. Performance metric

The main goal of the D2D caching network that we study is to efficiently distribute video files to requesting users in order to alleviate the traffic load burdening the fronthaul link of a C-RAN deployment. Video hit ratio is the primary metric that provides a direct measure of the degree in which this offloading is achieved, and it is defined as the percentage of the total number of requests generated by all users that can be served by the D2D caching network.

### 2.3.2. Characteristics of our simulation

We conducted event-driven simulations to evaluate the performance of our models. The parameters, the distributions and the processes listed and described are common for every model. The simulations have been performed using the MATLAB® platform created by MathWorks® and the random number generator used was seeded based on the current time of the simulation to avoid repeating the same random number arrays.

We simulate the requests of a number of users distributed over a two-dimensional plane. The area of the plane is denoted by W, with a default value equal to 1km$^2$. We use a spatial Poisson Point Process (PPP) for the placement of the UE devices (nodes) over the area. According to this distribution, the probability that there are n UE devices in an area of specific dimensions is given by a Poisson distribution and thus equal to $\frac{(\kappa W)^n}{n!}e^{-\kappa W}$, where κ denotes the average density of the points inside a square unit. The default value for κ is taken to be 5000 $\frac{users}{km^2}$. This number falls into the medium device density use case suggested by the 5G-PPP in [26]. Each UE device $i$ is then placed over the area with uniformly distributed random coordinates $X_i$ and $Y_i$. There are not cases where two or more UE devices are collocated. According to [27] PPP has been by far the most popular spatial model in simulations where transmitters and receivers are located randomly over a large area. A proportion (a) of the total number of users decides to become TXs. The default value for a is 25%.

The requests for video files are generated by users via their UE device. We assume that user requests for video files arrive according to a Poisson process, therefore the inter-arrival times are exponentially distributed with mean λ, with a default value for $\lambda = 2$ seconds. The mean rate ${}^{1}/_{\lambda}$ of the Poisson request arrival Process results cumulatively from all users and it is independent of the total number of users. Video files are selected from a total of M distinct videos. The size of the video files is not defined; instead we are interested in the duration of a successful video file transmission. We assume that the duration of a video file transmission follows an exponential distribution with mean μ. The default value for μ is taken to be 120 seconds.

The popularity of each of the M videos is assumed to follow a Zipf-like distribution Zipf $(s, M)$, where s corresponds to the degree of the skew coefficient and M to the total number of videos. Based on numerous studies, Zipf distributions have been established as good models to the measured popularity of video files [17], [28]. Every video is defined by a rank k. The first video with rank $k = 1$ is the most popular, i.e., it is requested by users with the highest frequency, and the last video with rank $k = M$ is the least popular, i.e., it is the least frequently requested by the users. According to Zipf's law, the request frequency of an element of rank k is equal to $\frac{1/k^s}{\sum_{n=1}^{M} 1/n^s}$. For $s = 0$ the distribution is uniform with no skew; on the contrary as the value of s is increasing the distribution is becoming more skewed, corresponding to higher content reuse, i.e., the first few popular video files account for the majority of video requests. In our simulation the value of s ranges from 0.5 to 1.8.

Users that choose to become cache-enabled transmitters are requesting a video file, of the total of M videos, to cache in their local memory. The probability that a user chooses a certain video is also given by the same Zipf-like distribution, as the one described above. In addition we examined some cases where the TXs select video files from a smaller library, consisting of the V-most popular elements of the M total videos. The default value of V is taken to be 10. We select this value for the parameter V as the one achieving best network performance, after comparing results for various values of V.

The duration of the simulations we conducted for every case model is defined by the simDuration parameter. The default value of simDuration is taken to be 200 hours. Through our simulations we observed that this value can provide steady state results with insignificant deviations. All the parameter default values are shown in Table 2.1.

| Notation | Definition | Default value |
|---|---|---|
| W | The area of the plane in squared kilometers. | 1 $km^2$ |
| a | The proportion of the users that decides to serve as TXs. | 25% |
| κ | The density of users in the unit area. | 5,000 $^{users}/_{km^2}$ |
| radius | Transmission range of every UE device. | Variable<br>$10 \le radius \le 170$<br>(m) |
| M | The total number of distinct videos. | 1000 |
| s | Exponent of the Zipf distribution corresponding to the skew of the popularity distribution of the video files. | Variable<br>$0.5 \le s \le 1.8$ |
| simDuration | The duration of the simulation. | 200 hours |
| V | The size of the library from which TXs choose a video file to cache. This parameter is used only in some cases. The value of **V** is always smaller than **M**. | 10 |
| λ | Mean request inter-arrival time. | 2 seconds |
| μ | Mean duration of a video file transmission. | 120 seconds |
| batch | Length of the sampling rate in Case Model 4 | 200 requests |
| window | Length of the sampling window in Case Model 4 | 2,000 requests |

**Table 2.1 : Simulation Parameters and their default value**

### 2.3.3. Simulated Case Models and Techniques

In this section we discuss the specifications of three different case models describing the D2D network with caching we examined. These models present the progression of our assumptions, in an effort to create a rather realistic model, and the techniques we progressively employed in order to achieve better network performance. We use this approach in order to understand and measure the impact of each improvement. The performance evaluation of each model will be discussed later in Section 2.4.

Although there are differences in the assumptions that we make for each case model we have employed some common characteristics. The area of the simulated plane is the same for every case model and all UE devices are assumed static. For every case model we use the spatial Poisson Point Process for the placement of the UE devices. In addition, all UE devices are D2D enabled and a certain proportion of the total number of UE devices decide to serve as TXs. We assume that all UE devices generate requests for video files. Moreover, every TX is assumed able of caching one video file. The caching policy, according to which a video file is chosen to be cached, is also common for every case model. Lastly, we assume outband controlled D2D communication in every case model and that all UE devices employ a common transmission range. All assumptions made for a case model apply to each subsequent case model, unless specified otherwise.

#### 2.3.3.1. Case Model 1

The first case model that we examined is a simple and basic model which can offer a first glance at the capacity that can be achieved by the D2D caching network described in this thesis. In this case model when the requested video file is found in the vicinity of the requesting UE device, the BBU pool initiates the potential D2D transmission. The interference between simultaneous transmissions is neglected. It is obvious that this assumption leads to optimistic results, due to the fact that UE devices can employ the highest possible transmission power in order to find and deliver the requesting video file throughout the simulated area, ignoring adjacent transmissions. However the results of this case model, when compared with the results of the case models which employ the Interference Protocol Model, can provide a better understanding of the interference-AoC tradeoff. Additionally, we assume that RXs and TXs choose from the actual library of videos which video files to watch and cache, respectively.

#### 2.3.3.2. Case Model 2

By applying the Interference Protocol Model for every D2D transmission in Case Model 1 we obtain a case model which can provide more realistic results than the one before. In this case model we try to balance the interference-AoC tradeoff by finding a suitable transmission range in order to achieve high network performance.

We examine the results of this case model for different values of the Zipf's distribution skew coefficient. We have noticed that the highest network performance achieved in our simulation, as well as the transmission range we have to select in order to achieve this performance, are strongly dependent on the value of the skew coefficient. Therefore, according to the results of this case model, we assign the most suitable transmission range to every value of the skew coefficient, thus creating a correspondence of values of the skew coefficient and values of the transmission range. Using this relation of the skew coefficient and the transmission range we have created a reference table according to which the BBU pool selects and broadcasts the desired transmission range value for a given value of the skew coefficient to all UE devices.

#### 2.3.3.3. Case Model 3

In this case model we assume that the cache-enabled UE devices and the requesting UE devices choose from different libraries which video files to cache and watch, respectively. More specifically, we create a library for the cache-enabled UE devices which is formed from a few most popular video files of the actual library. We will refer to this library as the caching library. The same proportion (as in the previous case models) of UE devices choose to serve as TXs, however in this case model they are offered a much smaller range of video files to select from; therefore we have higher repetition of the distinct video files throughout the distributed cache. The correspondence of values of the skew coefficient and values of the transmission range that was established in the previous case model is different here, because in this case it is affected by the change of the distribution of the cached video files throughout the designated area. Therefore, in this case model we have created a new reference table according to which our network adapts the transmission range to every value of the skew coefficient. This case model represents the base model of our network and we use the results derived for this case model as a benchmark.

### 2.3.3.4. Case Model 4

#### 2.3.3.4.1. The cumulative distribution function of the Zipf distribution

The techniques and the methods that we employed for this case model are based in a key attribute of the Zipf distribution. It is obvious that for different sets of data values drawn from a Zipf distribution with given parameters, the cumulative distribution function (cdf) based on the different data sets tends to converge as the number of samples in each data set is increasing. In our case, if we calculate the percentage of requests generated for the two hundred most popular video files in the library (20% of M), we observe that it strongly depends on the value of the skew coefficient. In Figure 2.4 we show the cdf of the Zipf distribution versus the size of the data set drawn from this distribution for different values of the skew coefficient and in Figure 2.5 we demonstrate the cdf of data sets of different sizes drawn from a Zipf distribution with given skew coefficient, $s = 1$.

#### 2.3.3.4.2. Case model description

For this case model we have considered a more realistic scenario in which the parameter of the distribution describing the popularity of the video files varies over time. More specifically, we assume that the skew coefficient of the Zipf distribution takes up different values over fixed time intervals throughout the duration of the simulation. This variation can be attributed, for example, to the occurrence of a significant and popular event in a region, which may cause users or organizations to upload videos regarding this event. For a period of time these videos will attract the interest of users, a fact which will consequently create a burst of request traffic for a narrow range of the most popular items in the library of video files. This burst of request traffic leads to an increase of the skew coefficient. Gradually the users will lose interest in the particular event, the number of requests generated for the video files regarding this event will decrease and the value of the skew coefficient will decrease.

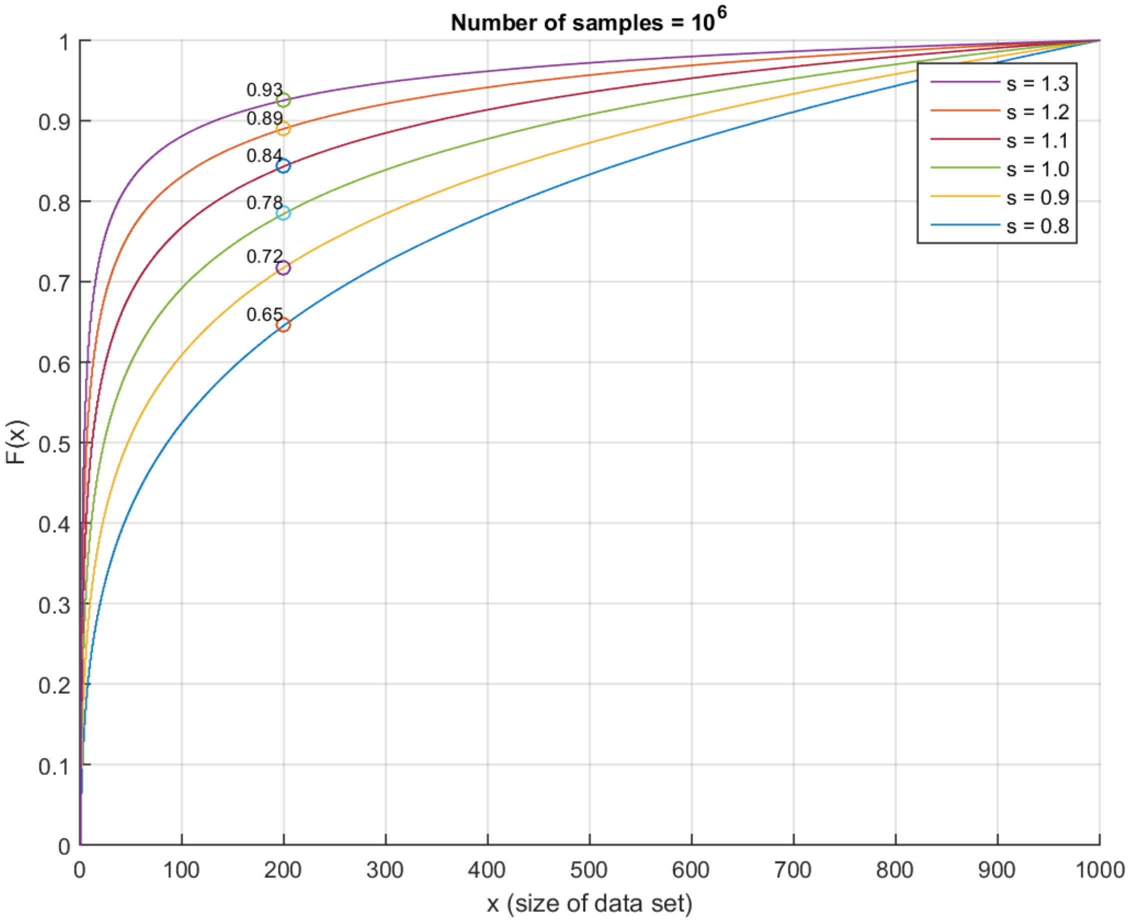

**Figure 2.4 : Cdf of the Zipf distribution for different values of the skew coefficient**

As we mentioned earlier in previous sections, there is a strong dependence of the transmission range employed by the UE devices on the skew coefficient of the Zipf popularity distribution. It is, therefore, important to detect the fluctuations of the skew coefficient value, when they occur, in order to maintain the high network performance achieved by Case Model 3 described in 2.4.1. For the purpose of estimating the skew coefficient of the Zipf distribution from which a set of video file ids are drawn from, we created a reference table in which we matched the cumulative distribution function of the two hundred most popular video files from various sets of video files to the value of the skew coefficient according to which each set was generated. More specifically, we generated random integers drawn from a Zipf distribution for different values of the skew coefficient (a million for each value) and we calculated the cumulative distribution function for each sample. This approach is used in [29] to observe the behavior of power laws and it is considered an accurate method. This way, if we take a random sample of subsequent requests without any knowledge of the exponent parameter (skew coefficient) characterizing the Zipf distribution, according to which these requests were generated, and we calculate the cdf of the two hundred most frequent distinct values, we can match the obtained result to a value of the skew coefficient, by using the above reference table. As the size of the sample of subsequent requests is increasing, so is the probability of an accurate estimation of the skew coefficient.

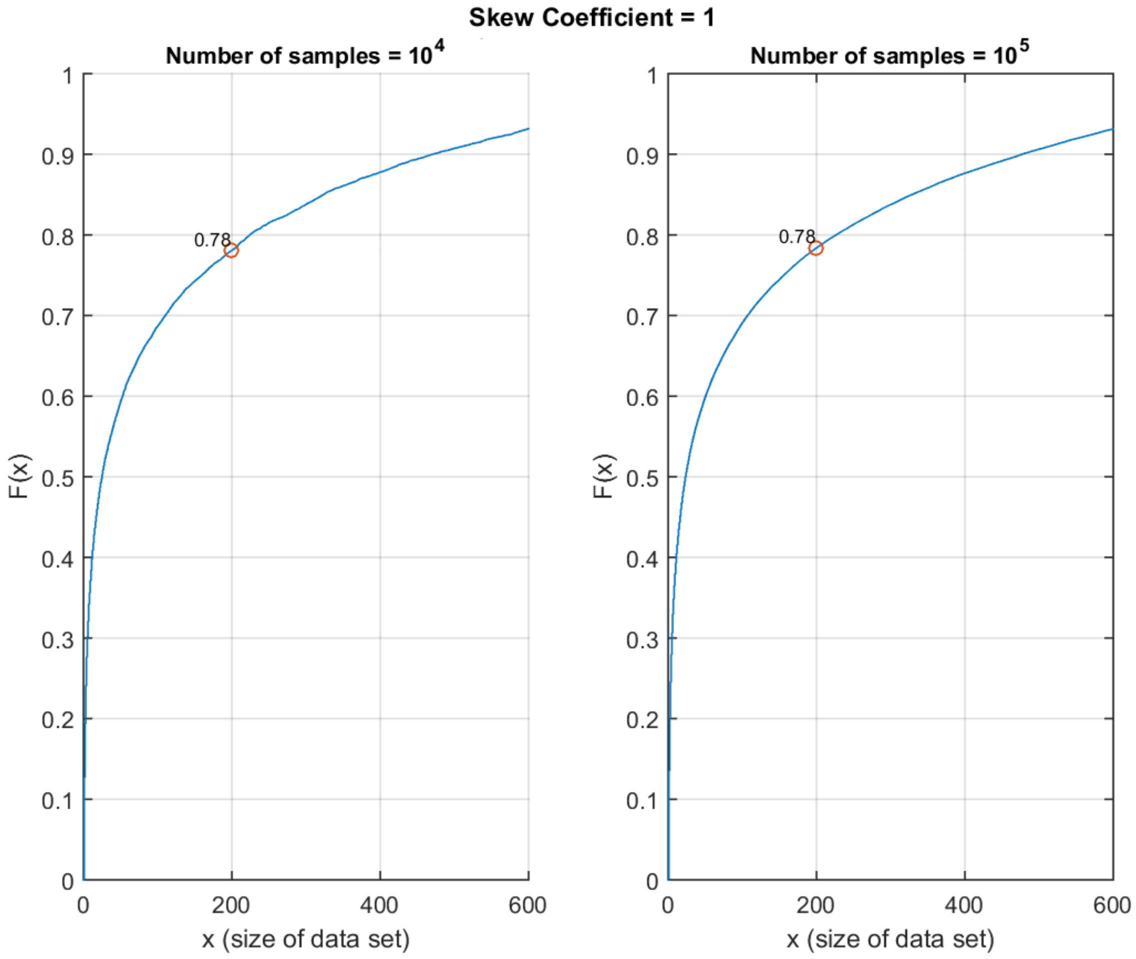

**Figure 2.5 : Cdf of different data sets drawn from the Zipf distribution**

### 2.3.3.4.3. Estimating the CDF

In order to achieve a reliable estimation of the cdf of the generated requests drawn from the Zipf popularity distribution, through which we can accurately detect the variation of the skew coefficient, we have employed the following scheme. We create a vector and we fill it with the requested video file ids as the simulation progresses. We use a window of fixed length, which slides forward as the simulation proceeds, for sampling these ids. Indicatively, the window can typically sample at most 2,000 requests; while in every simulation on average 360,000 requests are generated. The sampling occurs cumulatively in batches beginning from the left bound of the sampling window; the size of each batch is typically one tenth of the size of the sampling window. When the end of the sampling window is reached, at which point we have an aggregated sample equal to the size of the sampling window, we shift the sampling window to the right by a distance equal to half the size of the sampling window. This way in our sampling process we discard the ids which corresponded to the first half of the sampling window before shifting it to the right. By employing this sampling method, after the first 1,000 requests of our simulation, our sample always consists of at least 1,000 requests. We repeat this process until the id of the video file corresponding to the last request is included in the calculation estimating the cdf.

A schematic presentation of the estimation algorithm we employed is shown in Figure 2.6. Each vector represents the vector we use to store the ids of the video files requested by the users and corresponds to a different moment during the simulation. The pink-shaded cells of the vectors represent the video request ids that are currently included in the sample used for the estimation of the cdf. The last request id stored in the vector is shown by the pointer *Current request*. In the first vector the left bound of the sampling window corresponds to the first request id stored in the vector. Although the last request id stored is the 1,747th, the sample used for the estimation of the cdf contains only the first 1,600 elements of the vector, because of the sampling method we employ. The second vector shown in Figure 2.6 represents a later instance in the simulation in which the last request id stored is the 2,445th. As described earlier, when the end of the sampling window is reached, we shift the sampling window to the right by a distance equal to half the length of the sampling window, discarding thus the first 1,000 elements of the vector. In this case, and after the shift of the sampling window, the left bound of the sampling window corresponds to the 1,001st request id and the right bound to the 3,000th request id.

### 2.3.3.4.4. Updating the transmission range

Every time we sample a batch of requests, we estimate the cdf of the video file ids from the beginning of the sampling window. If we detect a change in the estimated values, derived from the batches of samples, we assume that the BBU pool matches the new estimated value of the skew coefficient with a value of the transmission range according to the reference table created for Case Model 3 (Table 2.2 (b) in Section

2.4.1) and broadcasts the new transmission range value to all UE devices. Broadcasting a value to all UE devices is assumed to have no significant negative effect on the overall network performance. Therefore, we make sure that the transmission range is frequently updated, in order for the network to better adapt to possible changes of the popularity of the video files. In Section 2.4.4 we further examine the network performance gain offered by promptly adapting the transmission range to the changes of the skew coefficient characterizing the Zipf popularity distribution.

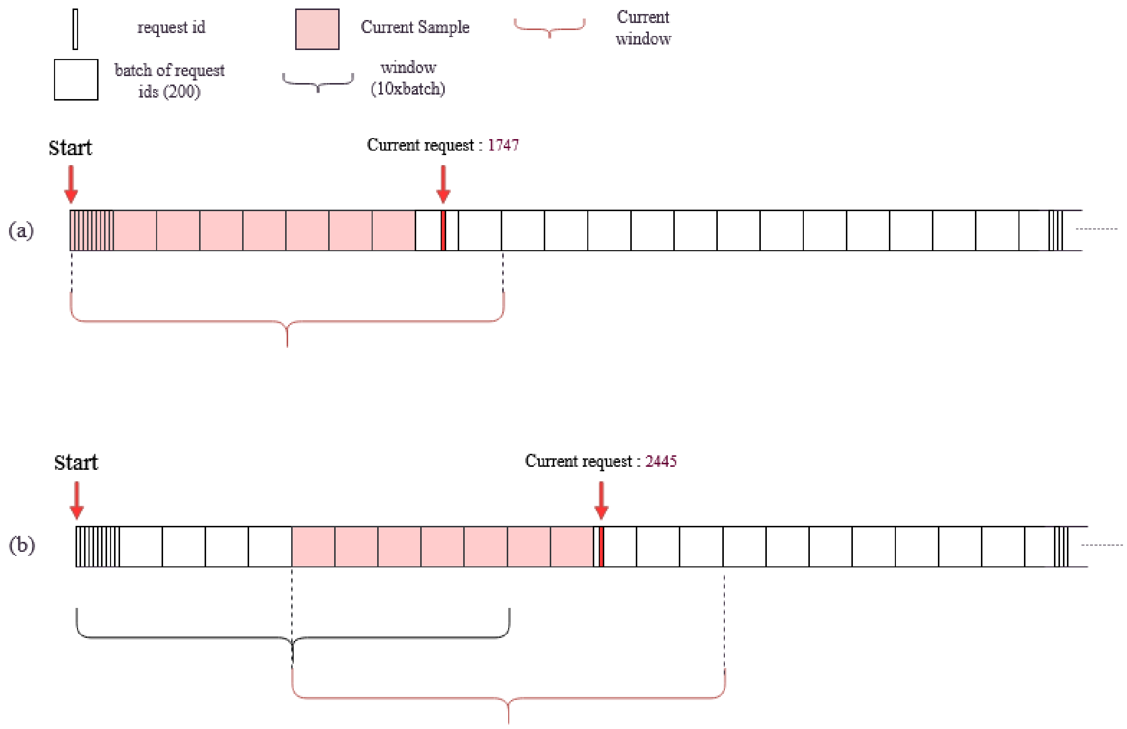

**Figure 2.6 : Schematic presentation of the employed estimation scheme**

### 2.3.3.4.5. Changing the Caching according to the new skew coefficient

Whenever a cache-enabled UE device applies to store a video file, it draws a random video file id from a Zipf distribution and proceeds to cache the corresponding video file, which file would be typically stored throughout the duration of the simulation in the previous case models. This distribution, for simplicity we will refer to it as the caching distribution, has the same parameters as the Zipf popularity distribution according to which users generate requests for video files. In the scenario in which the skew coefficient of the Zipf popularity distribution changes throughout the simulation, after the first change, there is no correlation between the caching distribution and the popularity distribution. Because of this inconsistency, the performance of the network is degraded. In order to maintain high performance in the network it is essential that the caching adapts to the changes in the popularity of the video files. To facilitate result comparisons with the previous case models, when a change in the cache is considered necessary, we assume that all TXs discard the currently cached content and proceed to cache video files according to the new estimation of the skew coefficient of the Zipf popularity distribution. Considering the workload caused in the network by such a change in the caching, we employ a scheme which will help avoiding unnecessary cache refreshes. An unnecessary cache refresh could occur if the estimation of the skew coefficient is not accurate or if the skew coefficient changes for a very short period of time after which it returns to the initial value. In our simulation we assume that the skew coefficient changes every simulated day (approx. after on average 43,000 requests).

We keep the value of the skew coefficient, characterizing the caching distribution currently employed, into a variable, named cc. In the beginning of the simulation the skew coefficient of the caching distribution employed has a value equal to the initial value of the skew coefficient of the Zipf popularity distribution. We assign this value to facilitate result comparisons with the network performance achieved by Case Model 3. When the end of the window, used to sample the ids of the requests, is reached, at which point we have an aggregated sample equal to the size of the sampling window, we assign the last estimation of the skew coefficient to a vector. Every time a new value of the estimated skew coefficient is assigned to this vector we compare it with the previous value stored.

a) If the two values are not equal we compare the current estimated value with the next estimated value that is going to be stored in the vector.
b) If the last two values are equal we then compare their common value to the value stored in the variable cc.
c) Finally, if these compared values are not equal, we initiate a cache refresh and we replace the content of the variable cc with the last estimation of the skew coefficient.

If any of the above conditions are not met we continue the sampling process without any change in the caching and we repeat the above comparison method for the next estimated value assigned in the vector.

## 2.4. Presentation and Discussion of Representative Simulation Results

In this section we present the results derived from our simulations. In Section 2.4.1we show the network performance of Case Models 2 and 3. We emphasize on the superiority of the network performance of Case Model 3, compared to the network performance of Case Model 2, and the techniques we employed in order to achieve it. Section 2.4.2 presents the behavior of Case Model 4 in two scenarios in which the skew coefficient of the Zipf popularity distribution varies over time. Next, in Section 2.4.3, we present the impact of three primary simulation parameters on the network performance of each simulated case model. Finally, in Section 2.4.4 we examine the impact of adapting the suitable transmission range in the case in which the video popularity distribution and the caching distribution of the network are inconsistent.

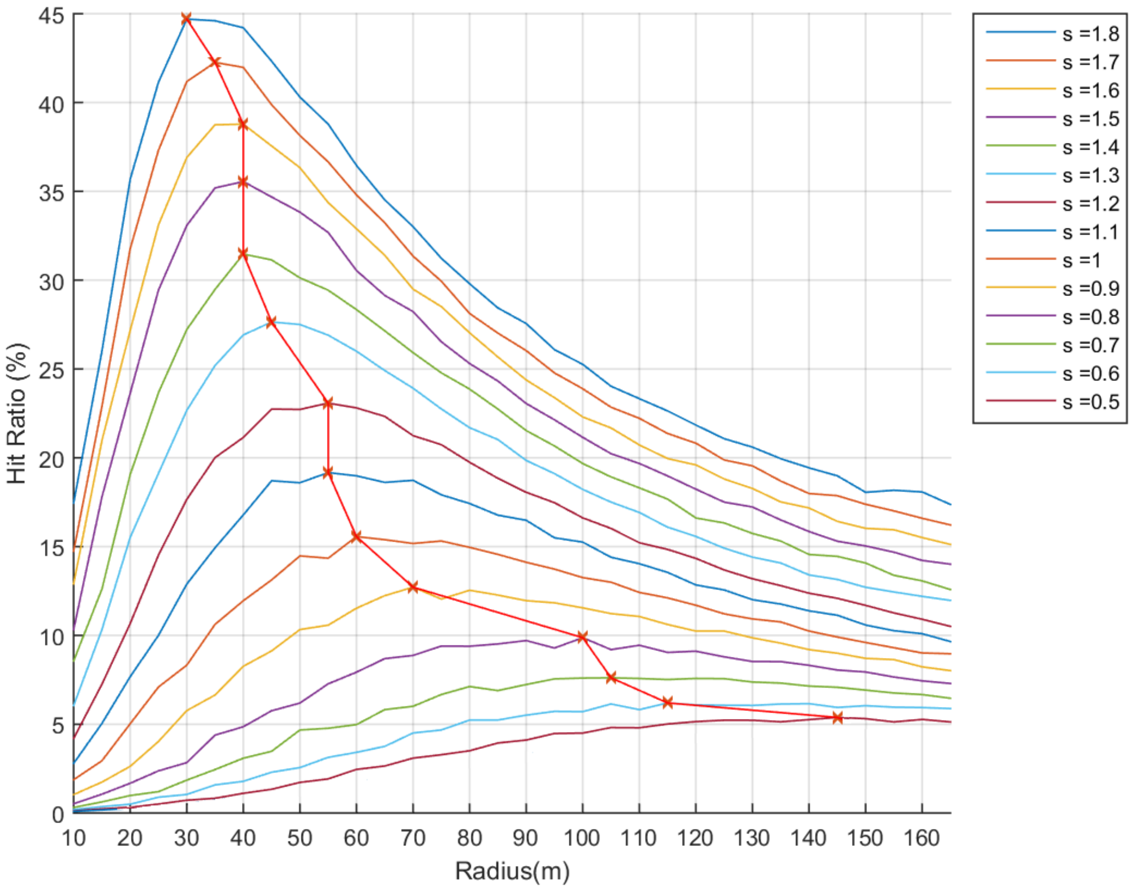

**Figure 2.7 : Case Model 2 network performance for every combination of values of transmission range and skew coefficient**

### 2.4.1. Network performance of Case Model 2 and Case Model 3

For Case Model 2 we have created a reference table filled with values of the transmission range and the corresponding values of the skew coefficient in order to find the suitable transmission range which sufficiently balances the interference-AoC tradeoff described in Section 2.2. For this purpose we conducted simulations for every combination of skew coefficient values and transmission range values. More specifically, the transmission range takes up values ranging from 10 meters up to 170 meters with 5 meters step (a total of 33 distinct values), whereas the values of the skew coefficient range from 0.5 up to 1.8 with step 0.1 (a total of 14 distinct values). The results shown in Figure 2.7 indicate that the best network performance, with respect to the hit ratio, is achieved for high values of the skew coefficient and short transmission ranges. Meaning that we can achieve higher spatial reuse when the popularity of video files is highly skewed, i.e., a few most popular video files account for the majority of the views. In addition we observe that transmission range has a greater impact on the network performance for higher values of the skew coefficient. The red symbols “x” in Figure 2.7 mark the best hit ratio achieved for each value of the skew coefficient.

Similarly, we created a table for Case Model 3. We recall that in this case model the caching library consists only of the 10 most popular video files of the actual library. Our simulations show that, by narrowing the range of items from which TXs choose which video file to cache, we can achieve better hit ratio for every value of the skew coefficient while using even shorter transmission range for each of these values. We came up with this modification by observing that, in Case Model 2, the first ten most popular video files account for almost all successful transmissions performed by the D2D network (in some cases they account for more than 99% of the successful transmission). This can be attributed to the following facts. The percentage of requests corresponding to video files other than the ten most popular is very small, considering that these video files comprise the 99% of the library. In addition, the probability that a UE device is able to find a requested video file, belonging to the 990 least popular video files of the video library, in its vicinity, is insignificant. Therefore, by caching only the ten most popular video files of the actual library, the probability of successfully finding and transmitting a video file increases. In Table 2.2 we show the contents of the reference tables created in Case Model 2 and Case Model 3 along with the peak network performance achieved by employing the corresponding suitable transmission range for every value of the skew coefficient.

| s | radius (m) | Hit Ratio (%) |
|---|---|---|
| 0.5 | 145 | 5.4 |
| 0.6 | 115 | 6.2 |
| 0.7 | 105 | 7.6 |
| 0.8 | 100 | 9.9 |
| 0.9 | 70 | 13 |
| 1.0 | 60 | 16 |
| 1.1 | 55 | 19 |
| 1.2 | 55 | 23 |
| 1.3 | 45 | 28 |
| 1.4 | 40 | 31 |
| 1.5 | 40 | 36 |
| 1.6 | 40 | 39 |
| 1.7 | 35 | 42 |
| 1.8 | 30 | 45 |

**(a) Case Model 2**

| s | radius (m) | Hit Ratio (%) |
|---|---|---|
| 0.5 | 75 | 5.7 |
| 0.6 | 70 | 7.7 |
| 0.7 | 65 | 10 |
| 0.8 | 60 | 13 |
| 0.9 | 55 | 16 |
| 1.0 | 50 | 20 |
| 1.1 | 45 | 24 |
| 1.2 | 45 | 27 |
| 1.3 | 40 | 31 |
| 1.4 | 40 | 35 |
| 1.5 | 40 | 38 |
| 1.6 | 35 | 41 |
| 1.7 | 35 | 44 |
| 1.8 | 35 | 46 |

**(b) Case Model 3**

**Table 2.2 : Suitable transmission range for each value of the skew coefficient and the corresponding hit ratio**

### 2.4.2. Adaptability of Case Model 4 and network performance

In order to examine the network performance of Case Model 4 and the sensitivity of the proposed adaptive algorithm to the changes of the skew coefficient of the Zipf popularity distribution, we assume a scenario in which the skew coefficient changes values every simulated day until it returns to its initial value, based on the scenario described in Section 2.3.3.4.2. More specifically, we assign an initial value to the skew coefficient equal to 0.7 and we assume that after 43,000 requests (approximately equal to one simulated day) the value of the skew coefficient rises and becomes equal to 1.5. This value remains for another 43,000 requests; after that the value decreases gradually, every 43,000 requests with step equal to 0.1, until it reaches the initial value, which is equal to 0.7.

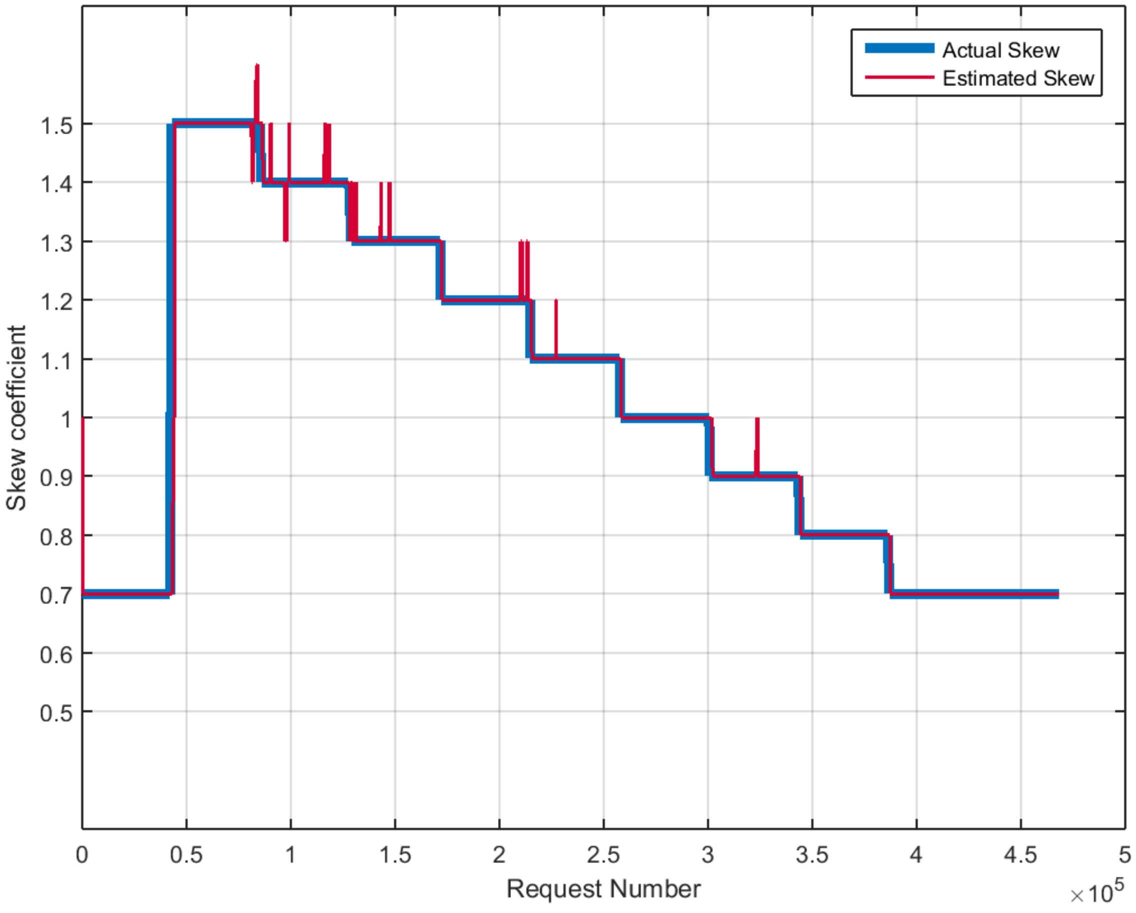

**Figure 2.8 : The changes of the values of the skew coefficient in the first scenario and the estimated values of our algorithm after every batch of requests**

The results of our simulation show that, in this scenario, our algorithm accurately estimates, after every batch of requests, the value of the skew coefficient with a success rate equal to 95.55%. In addition we observe that our system initiates cache refreshes successfully, i.e., only when the skew coefficient of the Zipf popularity distribution changes values, with a relatively small delay, which is approximately equal to three simulated hours (approx. three times the size of the sampling window). We present the changes of the skew coefficient of the Zipf popularity distribution compared to the estimated values of the algorithm of Case Model 4 in Figure 2.8 and Figure 2.9. In Figure 2.8 we show 2,344 estimated values; each of them was calculated after every batch of requests (every 200 requests). In Figure 2.9 we show 235 estimated values; each of them was calculated every time our simulation reached the end of the sampling window (every 2,000 requests). In addition, in Figure 2.9 we present the exact moments a cache refresh is initiated. In both figures the horizontal axis represents the progression of the simulation and it is expressed in generated so far number of requests.

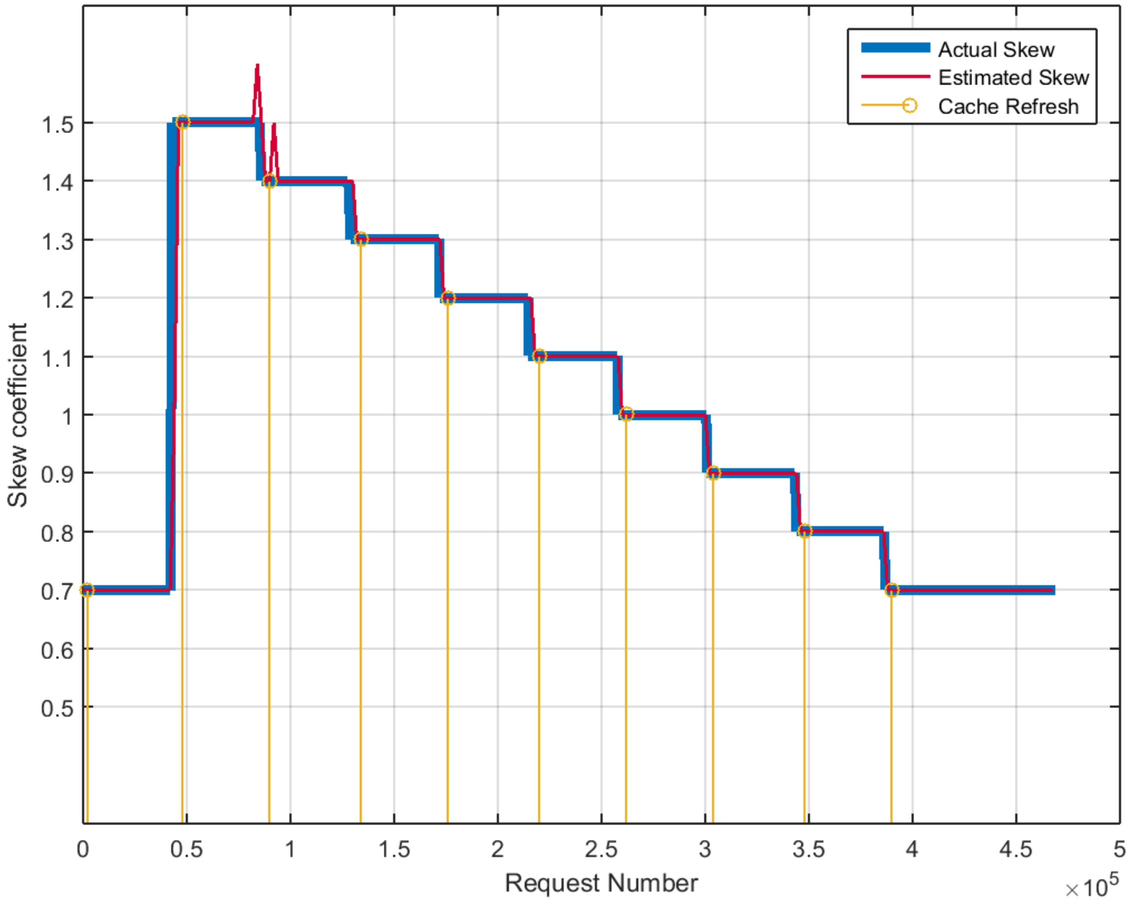

**Figure 2.9 : The changes of the values of the skew coefficient in the first scenario and the estimated values of our algorithm for every shift of the sampling window**

In order to further demonstrate the ability of Case Model 4 to adapt to the changes of the popularity distribution of the video files, we examine a second scenario in which we generate a random pattern of variations of the skew coefficient and we present in Figure 2.10 the behavior of Case Model 4 in this scenario. We generate uniformly distributed random sizes of intervals, in the range of 4,000 to 86,000 requests, over which the skew coefficient takes up uniformly distributed random values, in the range of 0.5 to 1.8. From the results shown in Figure 2.10 we observe that our algorithm initiates only one unnecessary cache refresh right after the arrival of the 432,000th request. This can be attributed to the fact that for such a high value (1.5) of the skew coefficient the difference between the cdf of elements drawn from the Zipf distribution with this value of the skew coefficient and the cdf of elements drawn from the same distribution with adjacent skew coefficient values (1.6) is small and thus our algorithm is more likely to produce a false estimation, whereas for lower values of the skew coefficient this difference is greater. Nevertheless the network performance achieved is close to the maximum hit ratio and superior to the network performance achieved by Case Model 3.

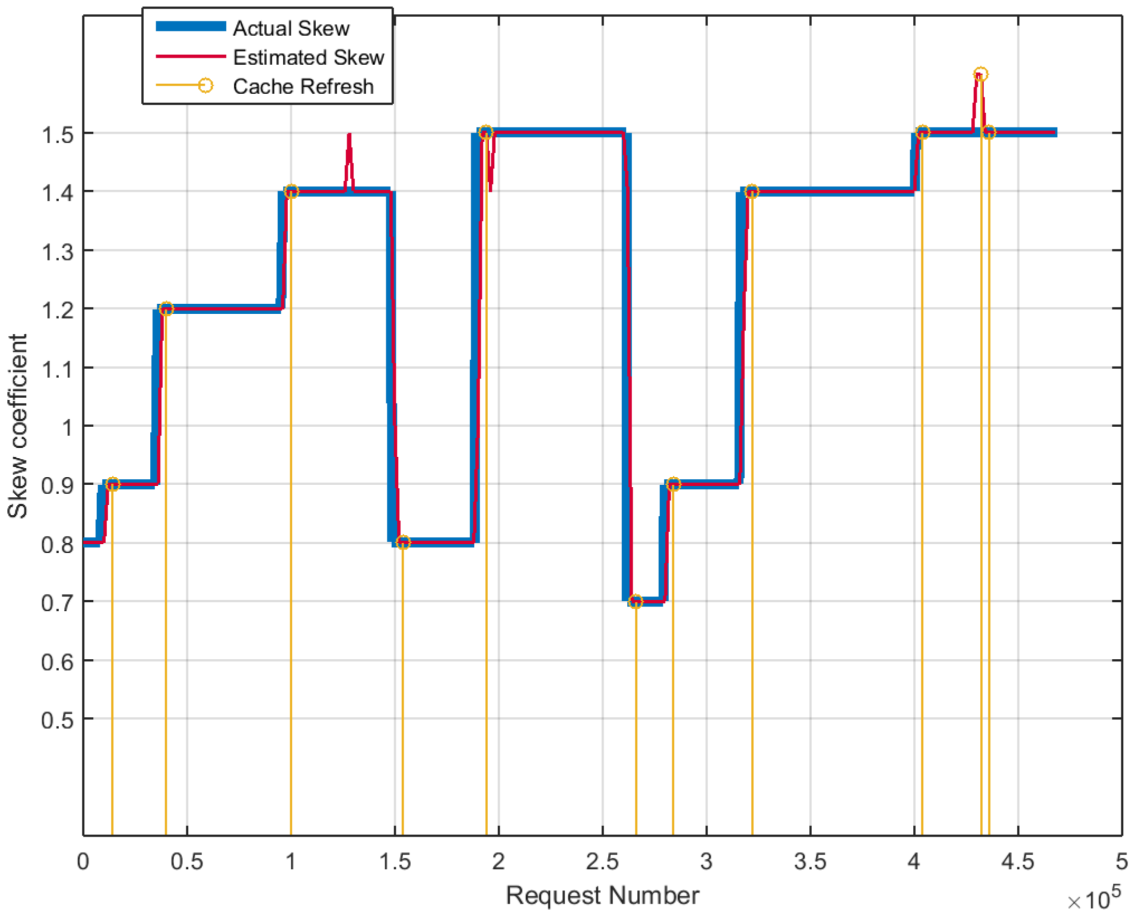

**Figure 2.10 : The changes of the values of the skew coefficient in the second scenario and the estimated values of our algorithm for every shift of the sampling window**

When we compare the hit ratio achieved by Case Model 3 in both scenarios, the simulation results show a significant gain in the network performance achieved by Case Model 4. We mention here that, as described in Section 2.3.3.3, Case Model 3 maintains the initial caching which was formed according to the first value of the skew coefficient of the Zipf popularity distribution at the beginning of the simulation. In the first scenario the first value of the skew coefficient is equal to 0.7, whereas in the second scenario the first value is equal to 0.8. According to the reference table created for Case Model 3, the value of the transmission range, corresponding to the initial value of the skew coefficient for each scenario, is equal to 65 meters in the first scenario and 60 meters in the second scenario. These values of the transmission range are also assumed constant throughout the simulations. The results of the network performance for Case Model 4 compared to the network performance for Case Model 3 are shown in Table 2.3. In addition, in order to evaluate the potential of the performance of Case Model 4, we show in the same table the hit ratio achieved by Case Model 3, when the exact variations of the skew coefficient of the Zipf popularity distribution are assumed known from the beginning of the simulation and the suitable caching distribution and transmission range are employed accordingly. The network performance achieved by Case Model 3 in this case is considered the upper limit of the performance of our algorithm. We observe that Case Model 4 achieves hit ratios close to the maximum hit ratio for both scenarios.

| Scenario id | Case Model 3 (%) | Case Model 4 (%) | Maximum (%) |
|---|---|---|---|
| 1 | 18.24 | 21.34 | 21.36 |
| 2 | 25.38 | 29.11 | 33.2 |

**Table 2.3 : Network performance achieved by Case Model 3 and Case Model 4 compared to the maximum achievable hit ratio**

### 2.4.3. The impact of transmission range, percentage of TXs and skew coefficient on network performance

We examined the performance of all case models for different values of the transmission range employed by the UE devices, the percentage of the users that decide to serve as TXs and the skew coefficient characterizing the Zipf popularity distribution. The simulation results show that the performance of all case models strongly depends on the transmission range, except from case model 4, for which the performance is not affected at all. This is expected, due to the fact that in case model 4 the BBU pool chooses the transmission range employed by the UE devices according to the reference table created for Case Model 3. On the contrary, when we examine the impact of the percentage of users that decide to serve as TXs, the results show that this parameter affects the performance of every case model in the same manner, i.e., the hit ratio of every case model is increasing for higher values of the parameter a. Lastly we examined the impact of the skew coefficient of the Zipf popularity distribution on the network performance for each case model. We observe that the increase of the skew coefficient has a beneficial effect to the performance of every case model. It is interesting that the performances of Case Model 2 and Case Model 4 represent a lower and upper bound of the performance of Case Model 3, respectively. The results of our simulations show that this behavior holds for every fixed value of the transmission range that we tested.

Figure 2.11 shows the results of the simulations for different values of the transmission range. The transmission range varies from 10 meters to 170 meters with an increment step equal to 5 meters. The results from the simulations of Case Model 2 and Case Model 3 show that even though Case Model 3 outperforms Case Model 2, they respond similarly to the changes of the transmission range, i.e., they achieve peak performance for a specific value of the transmission range; though this value for is lower for Case Model 3. We observe this behavior for every fixed value of the skew coefficient that we tested. This is due to the fact that Case Models 2 and 3 differ only in the size of the caching library, where the size is equal to the actual library and equal to ten, respectively. The network performance of Case Model 1 is increasing almost proportionally with the increase of the transmission range. This is attributed to the fact that in Case Model 1 the interference between active D2D transmissions is not taken into consideration. As a result, the BBU pool can increase the transmission range employed by the UE devices in order to find the requested video files throughout the area without rejecting potential D2D transmissions. On the other hand, the performance of Case Model 4 is stable despite the changes in transmission range and is equal to the highest hit ratio, achieved only by Case Model 3 when the value of the transmission range tested is in accordance to  the reference table created to balance the interference-AoC tradeoff in Case Model 3. This is also expected, because in Case Model 4 the transmission range is not given as a parameter, rather the system matches the suitable transmission range for the given value of the skew coefficient according to the reference table (Table 2.2 (b)). The skew coefficient has a fixed value equal to 1 and 25% of the total number of users are assumed cache-enabled transmitters. We tested the impact of the transmission range on the network performance for different values of the system parameters s and a, and we observed the same behavior as above for all case models through the changes of the transmission range.

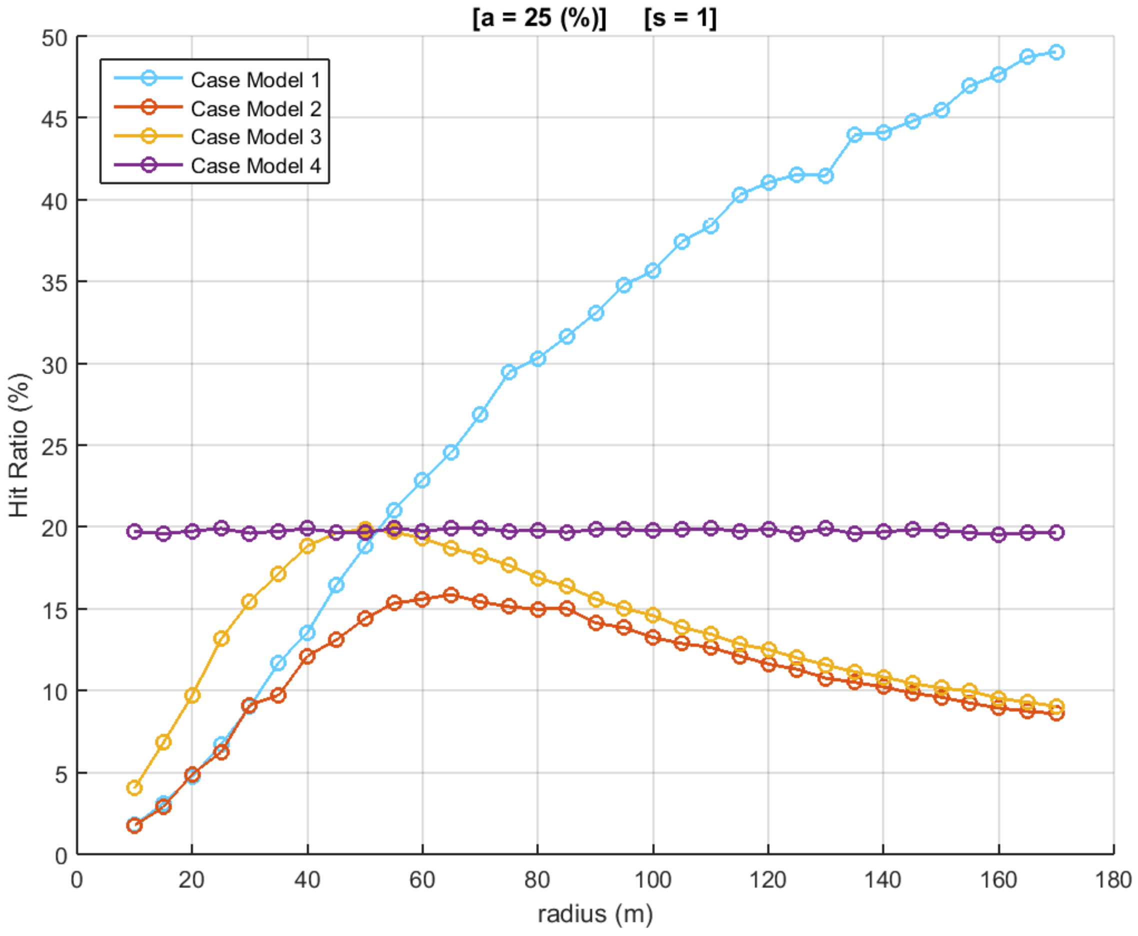

**Figure 2.11 : The impact of the transmission range on the network performance**

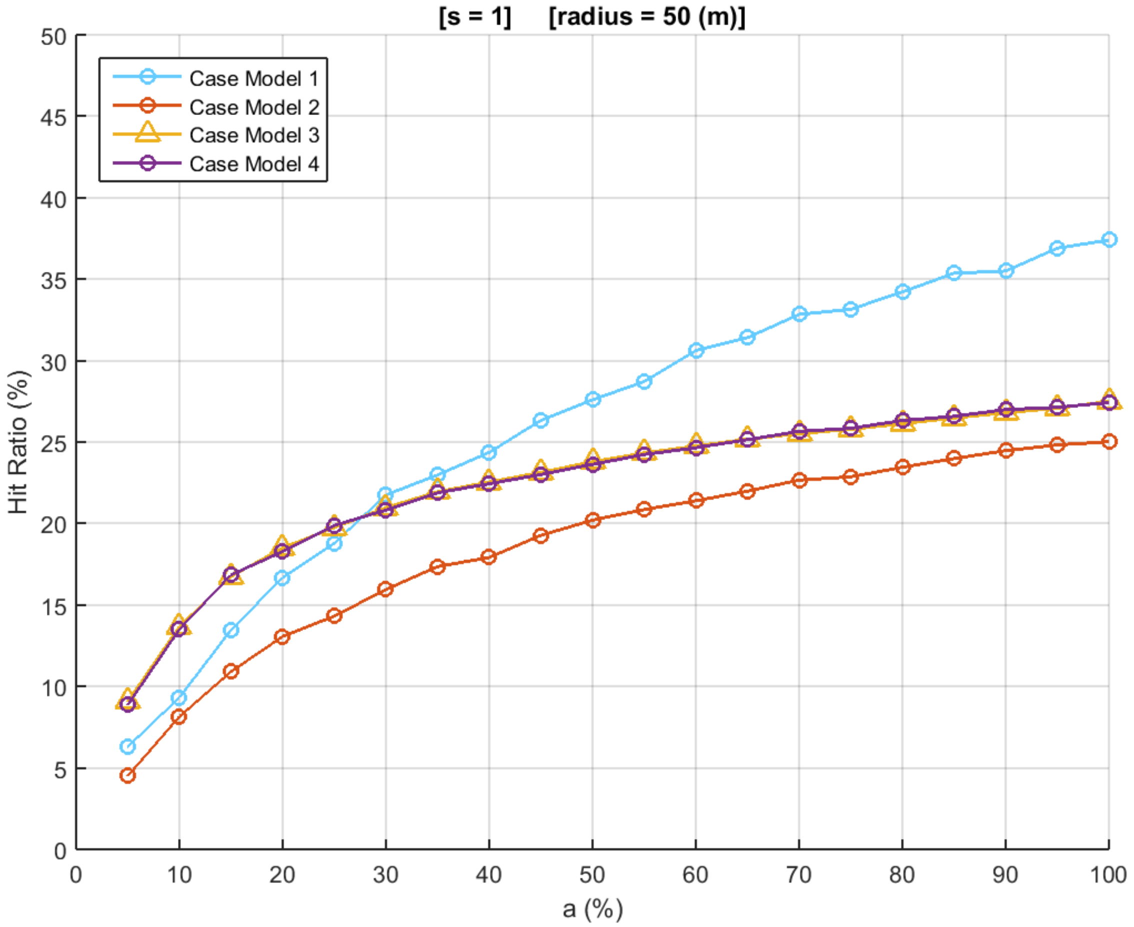

**Figure 2.12 : The impact of the density of cache-enabled UE devices on the network performance**

The network performance achieved by all case models for different percentage of users serving as TXs (i.e., different values of the parameter a) is shown in Figure 2.12. It is obvious that when more users decide to become TXs, it is more likely for UE devices to find the requested video file in their vicinity. Therefore the performance of every case model is improving for higher values of the parameter a. Nevertheless, as the percentage of users that are serving as TXs is increasing, the performance of all case models, except Case Model 1, tends to reach a limit value. If the density of TXs in an area is very high, the chance that a potential transmission is going to be disqualified because of the interference caused by active transmissions is increasing. Due to this fact the rate of increase of the performance achieved by case models that comply with the Interference Protocol Model is decreasing as the density of TXs is getting higher. Even though the performance of these case models can be improved with a proper adjustment of the transmission range, in this thesis we assume the percentage of UE devices that decide to become TXs constant, therefore we do not examine the correlation between the values of transmission range and the variations of this percentage. The network performance of Case Model 1 is increasing almost proportionally to the increase of the values of parameter a, as it is not affected by the interference between D2D transmissions. The skew coefficient and the transmission range have fixed values equal to 1 and 50 meters, respectively. We observe that the performance of Case Models 3 and 4 are equal. This is expected since the transmission range employed in this simulation is 50 meters and as it can be seen from Table 2.2 (b) this is the most suitable transmission range for Case Model 3. The same transmission range is also employed by Case Model 4.

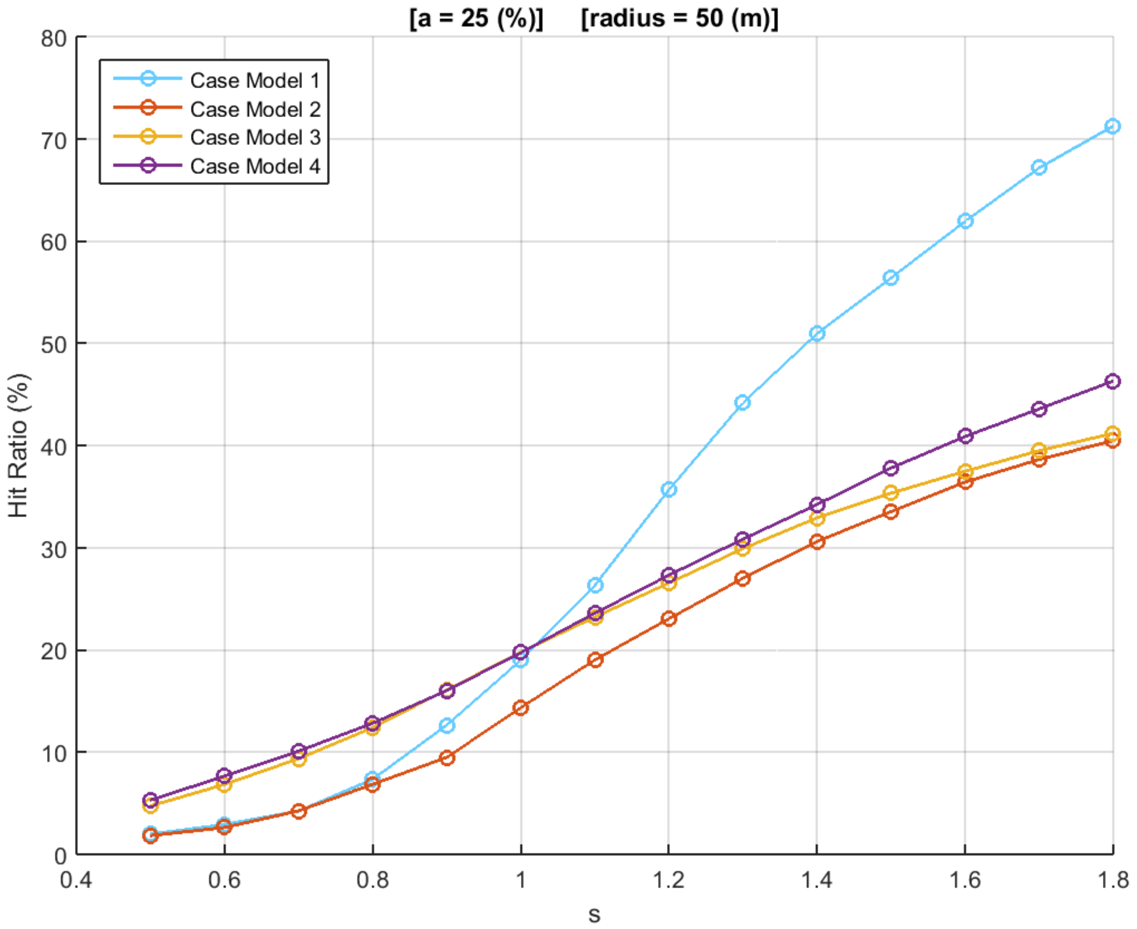

**Figure 2.13 : The impact of the skew coefficient on the network performance**

Finally, the impact of the skew coefficient characterizing the Zipf popularity distribution of the generated requests on network performance is shown in Figure 2.13. The parameter a has a fixed value equal to 25% and the value of the transmission range is set to be equal to 50 meters. We observe that the network performance achieved by Case Model 1 is increasing almost exponentially with the increase of the value of the skew coefficient; this is because its performance is not affected by the constraints of the Interference Protocol Model. On the contrary, the network performances achieved by Case Models 2, 3 and 4 are increasing almost proportionally with the increase of the value of the skew coefficient. We also observe that the network performance of Case Model 3 is always lower than the performance of Case Model 4 and higher than the performance of case Model 2. This can be attributed to the fact that Case Model 3 is an improved version of Case Model 2 and that Case Model 4 is expected to achieve the highest hit ratio among the Case Models, examined in this thesis, which employ the Interference Protocol Model. This is due to the fact that Case Model 4 adapts the transmission range employed by the UE devices according to the reference table shown in Table 2.2 (b) for every distinct value of the skew coefficient.

### 2.4.4. The importance of promptly adapting the transmission range

We examined the results of four simulated cases of Case Model 3 in order to demonstrate the importance of an accurate adaptation of the transmission range employed by the UE devices to the changes of the skew coefficient of the Zipf popularity distribution, according to the results of the reference table shown in Table 2.2 (b) in Section 2.4.1. Our results show that even when the caching is not consistent to the current value of the skew coefficient of the Zipf distribution modeling the popularity of the video files, there is a significant performance gain when our system adapts the transmission range accordingly. Table 2.4 shows the results of these four simulations. In the first two simulations we assume that the Zipf popularity distribution is characterized by a skew coefficient value equal to 1.4, whereas the skew coefficient of the Zipf distribution according to which the caching of video files is performed, which for brevity we denote by cs, has a value equal to 0.7. In the first simulation we assume that the value of the transmission range adapts to the skew coefficient of the Zipf popularity distribution and therefore has a value equal to 40 meters (this case corresponds to the Case Model 4), whereas in the second simulation we assign a value equal to 65 meters to the transmission range, meaning that the transmission range is **not** chosen accordingly (this case corresponds to the Case Model 3). For the following two simulations we employ the same logic with inverted values. More specifically, in simulations 3 and 4 the Zipf popularity distribution is characterized by a skew coefficient value equal to 0.7, whereas the value of cs is equal to 1.4. The values of the transmission range are 65 and 40 meters, respectively.

This behavior justifies our selection of the size of the batches of requests, which represents the sampling rate that we use for the estimation of the cdf of the generated video request ids in Case Model 4. Even though the probability to accurately estimate the cdf of the generated requests using a sample of 200 video file ids (which is equal to the size of every batch) is ranging, according to our simulation results, between 47%-68% depending on the value of the skew coefficient, we prefer to use this sampling rate because it is more sensitive to the changes of the skew coefficient. It is worth mentioning here that we estimate the cdf using a sample of 200 requests only at the beginning of the simulation, i.e., when we sample the first batch of video requests. As the simulation progresses, the ids of the video requests are sampled cumulatively from the beginning of every sampling window and due to the shifting method that we employ, the sample always consists of at least 1,000 requests. As a result the probability of an accurate estimation of the cdf turns out to be significantly higher than 68% (at least 80%), yet the skew coefficient is estimated every 200 requests, thus ensuring a frequent and accurate adaptation of the transmission range. Consequently, we achieve a better hit ratio compared to the case in which the sampling rate is equal to the size of the sampling window (which is equal to 2,000 requests), because in the former case the transmission range promptly adapts to a possible change of the skew coefficient of the Zipf popularity distribution. As described in Section 2.4.2, every time that the network detects a change in the skew coefficient of the Zipf popularity distribution there is a delay of approximately three hours (corresponding to approx. 6,000 simulated requests) before the caching distribution adapts to this change. Consequently, with the employed method, i.e., sampling with batches of size equal to 200 requests, we manage to partially make up for the degradation of the network performance caused by the inconsistency between the popularity distribution and the caching distribution.

In Figure 2.14 we present the probability to accurately estimate the skew coefficient characterizing the Zipf distribution of requested video file ids for different sizes of samples, according to our simulation. In order to produce these results we generate samples according to the Zipf distribution with given values of the skew coefficient equal to 0.6, 1 and 1.5; after that we use our estimation algorithm to estimate the skew coefficient of the samples and finally we check if the result of our estimation algorithm is equal to the corresponding values of the skew coefficient. The sizes of the samples are varying from 200 elements to 2,000 elements with an increment step of 200. We repeat this process 100,000 times for each sample size and each value of the skew coefficient. The significant difference between the success rates of our algorithm for different values of the skew coefficient of the Zipf distribution can be attributed to the exponential nature of the Zipf distribution, according to which the frequency of a request for a video file of specific rank drawn from this distribution depends exponentially on the value of the skew coefficient. The reader is referred to the remark we make in the discussion of the results presented in Figure 2.10 in Section 2.4.2, regarding the accuracy of our skew estimation method when the skew coefficient is high (e.g., approaches the higher value examined). The reader is also referred to [29] for further details.

| Simulation id | s | cs | radius (meters) | Hit Ratio (%) |
|---|---|---|---|---|
| 1 | 1.4 | 0.7 | 40 | 32.12 |
| 2 | 1.4 | 0.7 | 65 | 28.19 |
| 3 | 0.7 | 1.4 | 65 | 9.24 |
| 4 | 0.7 | 1.4 | 40 | 7.65 |

**Table 2.4 : The impact of successful adaptation of the transmission range**

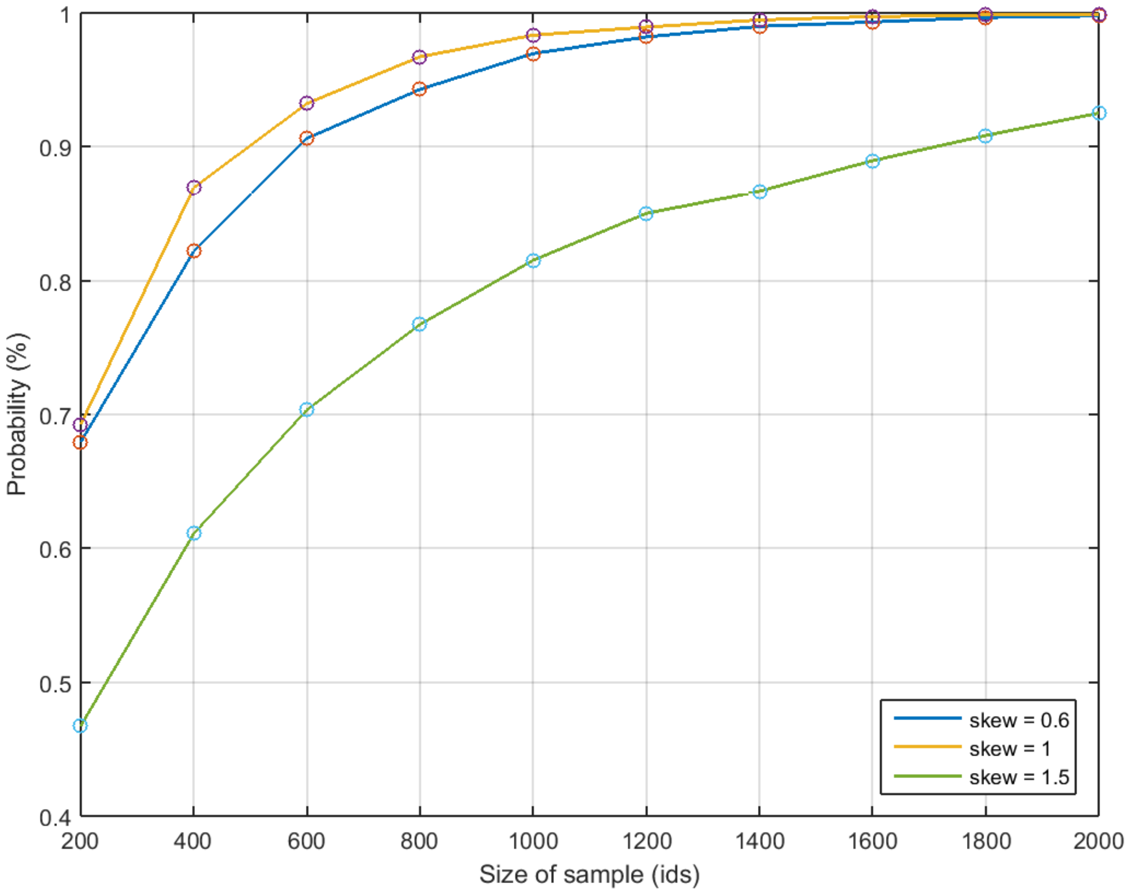

**Figure 2.14 : Success rate of the skew coefficient estimation algorithm**

# Chapter 3
# Conclusions

## 3.1. Overview of Work and Contribution

In this thesis we focus on the fronthaul link challenge imposed in the C-RAN architecture, which is considered by many the enabler of the future generations of mobile services [5], [7], [8], [9], [22]. The main advantage of C-RAN, i.e., the centralization of all functions into a single entity (BBU pool), creates the need of high throughput, low latency links between the BBU pool and the associated RRHs. In order to relax these requirements, we take advantage of the high demand in multimedia content which accounts for the majority of the cellular traffic, and the fact that video files popularity is unevenly distributed. We create a D2D caching network to serve a portion of the video files traffic "in front" of the fronthaul link. More specifically, the proposed architecture is a D2D outband controlled communication network with UE devices serving as caching nodes in collaboration with the C-RAN architecture. The BBU pool is in charge of the D2D links setup and the interference avoidance, and after receiving video files requests from users it matches the requesting UE device with a qualified caching node, which forwards the video file to the requesting UE device. We have created four variations of the proposed architecture in order to evaluate the network performance of the D2D caching network under different conditions. Overall, our simulations show that such an approach is beneficial to the cellular network and can alleviate the traffic load of the fronthaul link. The network performance achieved is measured in terms of the achieved video hit ratio and varies from 5.7% to 46%, depending on the characteristics of the modeled video popularity distribution.

Contrary to related work in the literature concerning the use of D2D technology along with caching into the UE devices, our model design takes into consideration the centralized nature of the C-RAN architecture and the powerful processing capability, which allows a single BBU pool to gather the functions of all telecommunication layers in order to serve multiple cells in a cellular network. Specifically, our model is not limited to a single cell and can simulate the requests of multiple users placed over a large geographical area. In addition, since the proposed communication scheme is designed to utilize unlicensed spectrum, the functions of the D2D network do not interfere with the cellular communications, thus making the proposed architecture scalable and easy to manage, in terms of interference. In conclusion we believe that, the suggested integration of a D2D caching network into the C-RAN can offer significant gains in network performance and can leverage the employment of the C-RAN architecture.

## 3.2. Ideas for Future Work

Even though our model of the proposed D2D caching network integrated into C-RAN is able to capture the fundamental benefits to the cellular network of such collaboration, it is essential to evaluate its performance in even more realistic wireless scenarios. Cases like medium or high user mobility and different types of traffic, e.g., bursty traffic, could affect the performance of the network and are probably calling for a different approach in the design of the architecture.

In addition, the main characteristic of the proposed architecture, which is the direct video file exchange between UE devices without relying on the cellular network, creates conflicting motivations for both mobile Internet service providers (ISPs) and mobile Internet users. From the scope of the mobile ISP, which supports the D2D network along with the cellular network, when a user request is served via the D2D network there is a benefit in load alleviation, but the UE devices are not charged because they do not make use of the cellular resources. However, since the D2D communications are partially controlled by the cellular network, there is an operating expense which burdens the mobile ISP. On the other hand, a D2D network is obviously sustained by the participation of the users to it; for this reason the users need to be urged to serve as caching nodes. Although it is probable that a user would prefer to be served from a nearby UE device, because of the high data rate and the low cost provided by a D2D communication (lower than using the cellular network), there is not a clear motivation for a user to decide to serve video requests, because in such case the user’s UE device would consume additional battery power. Consequently, a charging plan that would ensure a viable solution for network operators along with benefits for the D2D participants is necessary.

The D2D network that we describe in this thesis is based on a single channel priority model. This model has a spatial constraint regarding the interference between adjacent D2D links. It would be interesting to study a multi-channel D2D network and evaluate its performance. Such a network would require the design of a more complicated interference avoidance mechanism than the one employed in our work.

# References

[1] A. Molisch, “Applications and Requirements of Wireless Services,” in *Wireless communications*, 2nd ed. Chichester, UK: John Wiley & Sons, 2011,ch. 1, sec. 1.1, pp.4-8.

[2] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021,” Cisco, San Jose, CA, USA, Mar. 2017.

[3] 3GPP, “3rd Generation Partnership Project; Technical Specifications and Technical Reports for a UTRAN-based 3GPP system (Release 1999),” TR 21.101, 1999

[4] Xi Li, *Radio Access Network Dimensioning for 3G UMTS*, 1st ed. Wiesbaden, Germany: Vieweg+Teubner Verlag / Springer Fachmedien GmbH, 2011, pp. 15-60.

[5] M. Vaezi and Y. Zhang, “Cloud RAN,” in *Cloud Mobile Networks*. Cham, Switzerland: Springer, 2017, ch. 7.

[6] H. Kaaranen *et al.*, *UMTS networks,* 1st ed. Chichester: John Wiley & Sons, 2005.

[7] Y. Lin, L. Shao, Z. Zhu, Q. Wang and R. Sabhikhi, "Wireless network cloud: Architecture and system requirements", *IBM Journal of Research and Development*, vol. 54, no. 1, pp. 4:1-4:12, 2010.

[8] “C-RAN The Road Towards Green RAN,” China Mobile Research Institute, Tech. Rep., October 2011.

[9] K. M. S. Huq, S. Mumtaz, J. Rodriguez, P. Marques, B. Okyere and V. Frascolla, "Enhanced C-RAN Using D2D Network," in *IEEE Communications Magazine*, vol. 55, no. 3, pp. 100-107, March 2017.

[10] 3GPP, “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Feasibility Study on New Services and

Markets Technology Enablers; Stage 1 (Release 14)” TR 22.891, in Release 14, 2016

[11] “5G White Paper,” NGMN, v1.0, 2015

[12] 3GPP, “3rd Generation Partnership Project; Technical Specification Group SA; Feasibility Study for Proximity Services (ProSe) (Release 12),” TR 22.803 V1.0.0, Aug. 2012.

[13] S. Mumtaz, K. M. Saidul Huq and J. Rodriguez, "Direct mobile-to-mobile communication: Paradigm for 5G," in *IEEE Wireless Communications*, vol. 21, no. 5, pp. 14-23, October 2014.

[14] A. Pyattaev et al., “Network-Assisted D2D Over WiFi Direct” in *Smart Device to Smart Device Communication*, S. Mumtaz and J. Rodriguez. Cham, Switzerland: Springer, 2014, pp. 165-218

[15] X. Lin, J. G. Andrews, A. Ghosh and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," in *IEEE Communications Magazine*, vol. 52, no. 4, pp. 40-48, April 2014.

[16] X. Wu *et al.*, "FlashLinQ: A Synchronous Distributed Scheduler for Peer-to-Peer Ad Hoc Networks," in *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1215-1228, Aug. 2013.

[17] M. Cha, H. Kwak, , P. Rodriguez, , Y. Ahn, and S. Moon, “I tube, you tube, everybody tubes,” in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement - IMC '07*, pp.1-14, 2007.

[18] N. Golrezaei, P. Mansourifard, A. F. Molisch and A. G. Dimakis, "Base-Station Assisted Device-to-Device Communications for High-Throughput Wireless Video Networks," in *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3665-3676, July 2014.

[19] C. Yang, X. Zhao, Y. Yao and B. Xia, "Modeling and Analysis for Cache-Enabled Cognitive D2D Communications in Cellular Networks," *2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-6.

[20] H. J. Kang, K. Y. Park, K. Cho and C. G. Kang, "Mobile caching policies for device-to-device (D2D) content delivery networking," *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Toronto, ON, 2014, pp. 299-304.

[21] J. Liu, M. Sheng, T. Q. S. Quek and J. Li, "D2D enhanced cloud radio access networks with coordinated multi-point," *2016 IEEE International Conference on Communications (ICC),* Kuala Lumpur, 2016, pp. 1-6.

[22] J. Tang and T. Q. S. Quek, "The role of cloud computing in content-centric mobile networking," in *IEEE Communications Magazine*, vol. 54, no. 8, pp. 52-59, August 2016.

[23] K. Murphy, "Centralized ran and fronthaul", *Ericsson Tech. Rep.*, May 2015.

[24] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *in IEEE Transactions on Information Theory,* vol. 46, no. 2, pp. 388-404, Mar 2000.

[25] P. Cardieri, "Modeling Interference in Wireless Ad Hoc Networks," in *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 551-572, Fourth Quarter 2010.

[26] 5G-PPP, "5G PPP use cases and performance evaluation models," V. 1.0, 2016.

[27] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal and S. Weber, "A primer on spatial modeling and analysis in wireless networks," in *IEEE Communications Magazine*, vol. 48, no. 11, pp. 156-163, Nov 2010.

[28] Y. Zhou, L. Chen, C. Yang and D. M. Chiu, "Video Popularity Dynamics and Its Implication for Replication," in *IEEE Transactions on Multimedia*, vol. 17, no. 8, pp. 1273-1285, Aug. 2015.

[29] M. E. J. Newman, “Power laws, Pareto distributions and Zipf's law,” in *Contemporary Physics*, vol. 46, no.5, pp.323-351, Sept 2005