

Application of Multiple Description Coding for Adaptive QoS Mechanism for Mobile Cloud Computing

Ilan Sadeh*, Stojan Kitanov**, Danco Davcev***

* P.O Box 627 , 56 Hashikma Street, Pardesia Israel 42815

** University for Information Science and Technology "St. Paul the Apostle", Ohrid, Republic of Macedonia

*** Faculty of Computer Science and Engineering, Ss Cyril and Methodius University, Skopje, Republic of Macedonia
{ilansd@yahoo.com, stojan.kitanov@uist.edu.mk, danco.davcev@finki.ukim.mk}

Article Info

Article history:

Received Nov 25th, 2013

Accepted Dec 26th, 2013

Keyword:

Cloud Computing
Mobile Cloud Computing
Multiple Description Coding
Quality of Service
Quality of Fidelity

ABSTRACT

Multimedia transmission over cloud infrastructure is a hot research topic worldwide. It is very strongly related to video streaming, VoIP, mobile networks, and computer networks. The goal is a reliable integration of telephony, video and audio transmission, computing and broadband transmission based on cloud computing. One right approach to pave the way for mobile multimedia and cloud computing is Multiple Description Coding (MDC), i.e. the solution would be: TCP/IP and similar protocols to be used for transmission of text files, and Multiple Description Coding "Send and Forget" algorithm to be used as transmission method for Multimedia over the cloud. Multiple Description Coding would improve the Quality of Service and would provide new service of rate adaptive streaming. This paper presents a new approach for improving the quality of multimedia and other services in the cloud, by using Multiple Description Coding (MDC). Firstly MDC Send and Forget Algorithm is compared with the existing protocols such as TCP/IP, UDP, RTP, etc. Then the Achievable Rate Region for MDC system is evaluated. Finally, a new subset of Quality of Service that considers the blocking in multi-terminal multimedia network and fidelity losses is considered.

Copyright © 2013 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Stojan Kitanov,
Faculty for Information Systems Visualization Animations and Multimedia (ISVMA),
University for Information Science and Technology "St. Paul the Apostle",
Building at ARM, Ul. Partizanska B.B., Ohrid 6000, Republic of Macedonia.
Email: stojan.kitanov@uist.edu.mk

1. INTRODUCTION

Together with the explosive and rapid growth of Internet, mobile networks, mobile applications, and cloud computing, mobile cloud computing is introduced as a potential technology for mobile devices. As mobile network infrastructures continuously improve, their data transmission becomes increasingly available and affordable, and thus they are becoming popular clients to consume any internet web-based applications.

The "cloud" in cloud computing is defined as the set of hardware, networks, storage, services, and interfaces that combine to deliver aspects of computing as a service. Cloud computing provides delivery of services, storage, software and processing capacity over internet, reducing cost, increasing, automating systems, decoupling of service delivery from underlying technology, and providing flexibility and mobility of information. It provides computing utilities and resources that can be leased and released by the customers through the Internet in an on-demand fashion [1], [2].

On the other hand Mobile Cloud Computing (MCC) refers to an infrastructure where both the data storage and the data processing happen outside of the mobile device. Mobile cloud applications move the computing power and data storage away from mobile phones and into the cloud, bringing applications and

mobile computing to not just smartphone users but a much broader range of mobile subscribers. Mobile Cloud Computing (MCC) integrates the cloud computing into the mobile environment and overcomes the obstacles related to the performance (battery life, storage, and bandwidth), environment (heterogeneity, scalability and availability), and security (reliability and privacy) [3], [4].

A General Mobile Cloud Computing Model that is made up of complex network and relationships of and in between Infrastructure Providers, Application/Services Providers, End-Users and Developers all producing and/or consuming applications and/or services on internet is given in [5]. In such a network model different Cloud Computing consumers have different Quality of Service (QoS) requirements of applications against a service provider's capabilities and usage terms [6]. For future Internet, QoS-aware broadband multimedia applications provided by cloud computing and contents delivery services are expected to be the major traffic generator and consumer.

As a result automated negotiation is needed to accommodate different consumer's QoS requirements. The result of such a negotiation is a Service Level Agreement (SLA), an electronic contract that establishes all relevant aspects of the service. The SLA contains the QoS requirements of Applications hosted by a cloud based computing platform, such as timeliness, scalability, response-time, throughput, failure probability, and dependability (availability, security, safety, reliability, etc.). The SLA guarantees the corresponding QoS during the applications execution, especially important for sensitive and critical applications. The QoS aspects in a network are influenced by a large number of factors such as: Channel conditions, resource allocation policies, available resources, delay sensitive/insensitive traffic. Therefore an adaptive QoS resource reservation management algorithm is needed to exist in the network.

One possible solution to serve all the users that have different bandwidth requirements for video streaming over multicast heterogeneous networks is the server to transcode the stream to each of the user bandwidths [7], [8]. However this is neither efficient nor feasible solution. Another approach is to encode the video into a number of streams of different bit rates that users join to best match their bandwidth requirements [7], [8]. However, because of the wide range of bandwidth requirements, and limited number of video streams, this would result many receivers to get a stream substantially lower than their bandwidth requirements.

A much better approach to address the heterogeneity issue is to use Multiple Description Coding (MDC), where the video source is encoded into multiple independent descriptions with different bandwidths [9] – [12]. A receiver, depending on its available bandwidth joins different descriptions to meet its bandwidth requirements. Additionally when the source data is splitted into several descriptors, and like that the encoded streams are being transmitted over several channels, each channel would encounter almost flat fading and the probability of receiving some information from at least one channel is greatly increased with the price of increased overhead [13], [14]

MDC is an adequate approach for providing rate-adaptive streaming, where the content providers can send all the descriptions of a stream without having to think about the receiver limitations. Receivers that don't have the possibility to sustain a certain data rate, they can only subscribe to a subset of the multiple descriptions. This relieves the sender from sending streams with different rates to different receivers.

Multiple Description Coding (MDC) system is an efficient and reliable communication system aiming to improve QoS (Quality of Service) by breaking a message into more distorted replicas, which together contain enough information to reconstruct the original message. A low fidelity reproduction can be obtained if only one channel is intact.

Some of the possible applications can be in the field of mobile multimedia, cellular networks and cloud computing in general. Particularly for cellular and mobile networks MDC will grant different number of descriptors to mobile devices that have different device grades (battery power, and maximum data rate being supported) [13]. More descriptors will be granted to the mobile devices that have higher device grades or capabilities. The quality of received packets increases as the number of granted descriptors increases. The ultimate goal is a reliable integration between telephony, video and audio transmission computing and broadband transmission based on mobile cloud computing.

This paper proposes how the Quality of Service for Mobile Cloud Computing can be improved by using Multiple Description Coding. This paper determines the bound for the achievable rate region for Multiple Description Coding systems and presents a new approach for improving the quality of multimedia and other services in the cloud. Furthermore, it introduces a new subset of Quality of Service that takes into account the issues of blocking in multi terminal multimedia network and fidelity losses.

This paper is organized as follows. Section 2 presents the related work. Section 3 explains the advantages of Multiple Description Coding "Send and Forget" Method versus existing classical techniques TCP/IP, UDP and RTP. Section 4 focuses on Achievable Rate Region (ARR) for MDC Systems. Section 5 defines a potentially new approach of Quality of Service that can be used in Cloud Computing. Finally Section 6 concludes the paper, and presents the future work.

2. RELATED WORK

Recently, with the popularization and promotion of the Internet, mobile smartphone devices cloud computing and mobile cloud computing, there is an inevitably high demand to transmit image and video at real time in packet-switching networks and narrowband networks. The low computing and processing power of mobile and wireless devices, and the increasingly serious congestion problem in wireless communication networks and the Internet, along with the growing complexity of heterogeneity in networks, and the introduction of Cloud Computing and Mobile Cloud Computing have brought many challenges to the traditional video image coding.

At the moment, the vast majority of state-of-the art codecs uses Single Description (SD) video coding [15] – [22]. However these SD video coding relies on the traditional TCP/IP protocol, where the system quality seriously declines when the network packet loss is serious. Additionally, excessive Automatic Repeat-reQuests (ARQs) will cause excessive delay, and strong Forward Error Correction (FEC) will also bring additional delay because of its complexity, seriously affecting the real-time play of the video [16].

Therefore, the traditional TCP/IP protocol and the existing coding methods are not efficient for transmission of image and multimedia streaming applications. In order to provide high-quality video services to the users in wireless mobile terminals, it is necessary to overcome the low operational ability in the terminal and the problem caused by unreliable transmission in the existing network. Therefore it is necessary to design video coding that has low coding complexity and strong error resilient ability. One possible solution of this issue could be a full separation of the traffic: TCP/IP, or similar protocol for transmission of text files, and Multiple Description Coding (MDC) for Multimedia communication.

Multiple description coding (MDC) is a Source Coding technique that fragments a single media stream into a few sub-streams referred to as descriptions [9] – [12]. The packets of each description are routed over multiple, disjoint paths. In order to decode the media stream, any description can be used. However, the quality improves with the number of descriptions received in parallel. The idea of MDC is to provide “Send and Forget” method, improve error resilience to multimedia streams and improve Quality of Service.

Besides increased fault tolerance, MDC allows for rate-adaptive streaming: Content providers send all descriptions of a stream without paying attention to the download limitations of clients. Receivers that can't sustain the data rate only subscribe to a subset of these streams, thus freeing the content provider from sending additional streams at lower data rates.

Historically, the theory of lossy source coding emerged from the theory of lossless source coding based on Asymptotic Equipartition Property (AEP) or Shannon McMillan Breiman Theorem [23] – [26]. This theorem states the general property of the output samples of a stochastic source and it is the basis of Information Theory. In 1990 Ornstein and Shields [27] proposed an algorithm that converges to Shannon capacity known as Rate Distortion Function. However this algorithm assumes apriori knowledge of source distribution.

Later Sadeh proposed a generalized universal coding algorithm (Sadeh Algorithm) based on Wiener Ziv and Ornstein and Shield's Algorithms. Sadeh Algorithm does not require knowledge of source distribution [28] – [31]. He has also shown the asymptotic expansion of the Rate Distortion Bound, with a large tolerable error probability, by using Large Deviation Theory. This methods can be applied in Multiple Description Coding to determine the asymptotic expansion of the bounds of real algorithms.

Many research papers studied MDC in the context of source coding [32] – [37]. An inner bound on the achievable-rate region was developed in [33] and tightness for some special cases in [36]. The present work goes into this direction.

To this purpose, the work of Sadeh established the Achievable Rate Region of Multiple Description Coding [29]. The exact relations between rates and distortions in all cases are given in that work. Before Sadeh work there were known only inner and outer bounds for the Achievable Rate Region.

Despite the aforementioned advantages of MDC, SD codecs for multimedia communications still dominate in industry. However the cloud computing and next generation of mobile networks are the technologies provide a very good area for application of Real-Time Multiple Description Coding, since it is clear that the old protocols such as TCP/IP do not fit. Mobile multimedia and cloud computing provide a very good area for application of Real-Time Multiple description coding.

3. MULTIPLE DESCRIPTION CODING “SEND AND FORGET” METHOD VERSUS CLASSICAL TECHNIQUES

One of the important and innovative techniques in data transmission is Multiple Description Coding or MDC [9] – [12]. It is a source coding technique that partitions the stream into n substreams ($n \geq 2$). These substreams are often referred to as descriptions. The packets of each description are routed over multiple,

(partially) disjoint paths. In order to decode the media stream, any description can be used, however, the quality improves with the number of descriptions received in parallel. Using this approach, certain flexibility can be obtained in recovering from possible errors in transmission.

Since an arbitrary subset of descriptions can be used to decode the original stream, network congestion or packet loss which are common in best-effort networks such as the Internet will not interrupt the stream but only cause a (temporary) loss of quality.

Multiple Description Coding [9] – [12] is an adequate approach for providing rate-adaptive streaming: content providers can send all the descriptions of a stream without having to think about the receiver limitations. Receivers that don't have the possibility to sustain a certain data rate, can only subscribe to a subset of the multiple descriptions. This relieves the sender from sending streams with different rates to different receivers.

Multiple Description Coding is much superior than the existing network transmission technologies, such as TCP/IP, UDP and RTP, with regard to: packet loss, delay, jitter, blocking probability or consistence, bit/packet error rate probability, costs, network congestion.

In the TCP/IP protocol, widely used in computer networks, packet loss is a normal event [38]. When a packet loss occurs, the TCP/IP protocol sends extra packets that repeat the information lost, multiplying the data rate sent. In the best case, this effort causes an important delay in the transmission, while very often it pushes the entire network into a 'congestion collapse', where an important part of the packets is lost. This causes a degradation of the network throughput.

With the solution proposed in this work, the delay caused by packet re-transmission is reduced to zero. "Send and forget" method over two parallel channels performs better than TCP/IP, with regard to packet loss, network congestion and delay. Though, the TCP/IP protocol is considered improper for real-time applications such as VoIP, streaming, video-conferencing, etc. For such applications, protocols like the Real-time Transport Protocol (RTP) running over the User Datagram Protocol (UDP) are usually recommended instead [25].

UDP provides a procedure for application programs to send messages to other programs with a limited number of protocol mechanisms. The protocol is transaction oriented, and delivery and duplicate protection are not guaranteed. It has no handshaking dialogues, and thus exposes any unreliability of the underlying network protocol to the user's program. As it assumes IP as the underlying protocol, there is no guarantee of delivery, ordering or duplicate protection [38].

In situations in which error checking and correction is either not necessary or performed in the application, UDP can be a good choice, because it avoids the overhead of processing at network interface level. Time-sensitive applications often use UDP because dropping packets is preferable to waiting for delayed packets, which may not be an option in a real-time system.

Real-time multimedia streaming applications require timely delivery of information and can tolerate some packet loss to achieve this goal. For example, loss of a packet in audio application may result in loss of a fraction of a second of audio data, which can be made unnoticeable with suitable error concealment algorithms.

An important protocol that defines a standardized format for sending video and audio over IP networks is the Real time Transport Protocol (RTP). The majority of the RTP implementations are built on the User Datagram Protocol [38]. RTP is designed for end-to-end, real-time, transfer of stream data. The protocol provides facility for jitter compensation and detection of out of sequence arrival in data, that are common during transmissions on an IP network. RTP is regarded as the primary standard for audio/video transport in IP networks.

However, the User Datagram Protocol, and as a consequence the RTP protocol, suffer from possible blocking of the transmission, packet loss, bit/packet errors. In addition, there might be a significant variation in delay or jitter, which is the most frustrating parameter for certain devices. With the proposed approach, the average delay will be reduced and, as a consequence, relevant reduction of jitter is expected. These are closely related to QoS.

Sending the message through two (possibly) independent channels is an application of the general idea of Multiple Description Coding (MDC) [9] – [12]. Pushing a signal through two parallel channels can only reduce the bit/packet-error rate probability. Even if one of the channels emits a faulty bit/packet, the receiver can still use the bit/packet emitted by the other channel to reconstruct the correct signal. In this case an error detection code can be used to detect the errors, if any, and eventually switch to the correctly transmitted bit/packet.

The approach here proposed in this paper reduces the blocking probability in communication network, and under certain conditions it can be very low. Transmitting the signal through a channel pair, significantly reduces the costs of the physical mean of communication, because now the signal transmission can be with channels of lower capacity.

For all these reasons, this approach can be considered a very promising technique for use in information networks that support real-time applications such as video streaming, VoIP, mobile telephony, and particularly Cloud Computing, and Mobile Cloud Computing.

4. ACHIEVABLE RATE REGION (ARR) FOR MDC SYSTEMS

In order to determine the multirate distortion function for MDC systems, firstly it is necessary to prove that optimal pair encoder-decoder is a quasi deterministic system. Then it is necessary to prove that optimal MDC system is quasi-deterministic.

Definition 1. A system is "quasi deterministic" if there is almost no randomness in the assignment of a time process v at the output of a system to an input time process u . Even if there is some randomness (for example: initial random) in the assignment process, we impose that its contribution to the conditional entropy rate vanishes in the limit. Such system can be mathematically represented as:

$$H(v|u) = \lim_{n \rightarrow \infty} \left(\frac{1}{n} H(v_1^n | u_1^n) \right) \quad (1)$$

where $H(v|u)$ and $H(v_1^n | u_1^n)$ are the conditional entropies, n is the size of the block, and u_1^n, v_1^n represent the range from 1 to n for the input and output processes u and v , respectively.

Claim 1. An optimal pair of encoder decoder is a quasi deterministic system.

Proof: If the system is not quasi deterministic then there is randomness in the assignment of the time process v at the output of the system to the n input time process u . Therefore

$$H(v|u) > 0 \quad (2)$$

The reproduction entropy of the output process v is

$$H(v) = I(u; v) + H(v|u) \geq I(u; v) \quad (3)$$

where $I(u; v)$ is the mutual information for the input and output processes u and v . By applying quasideterministic assignment $H(v|u) = 0$ of an output process the reproduction entropy can be reduces in order to obtain obtain better encoder decoder pair:

$$H(v) = I(u; v) \quad (4)$$

This result agrees with the remarkable implication of the sliding block coding theorem which asserts that there exists a joint measure with the property $H(v|u) = 0$ that yields distortion and rate arbitrary close to the infimum [39].

Definition 2. A Multiple Description Coding system is an efficient and reliable communication system trying to secure QoS (Quality of Service) by breaking a message into two (or more) distorted replicas, which together contain enough information to reconstruct the original message. A low fidelity reproduction is obtained if only one channel is intact.

Let's assume that the message is fragmented into two distorted replicas, although the solution can be extended to more than two. Obviously if the probability of total failure (Blocking Event) of a channel is denoted by p , then if there are two available channels, and assuming independence of the events of blocking a channel, then the probability that both channels are blocked is p^2 much lower than the single channel case. Even if the events of blocking the two channels simultaneously, are dependent (and probably they are) still there will be a great improvement comparing to the single channel case. Therefore, sharp reduction of Probability of Blocking is a good reason to adopt Multiple Description Coding system.

Possible applications of MDC can be in telephony, internet and in particular on cellular channels. A digitized voice or video signal could be divided in half (or other fraction) and sent over two routes. If either route is blocked or disconnected, a reduced fidelity reproduction is still available to the receiver. We apply the process assignment concept to these multiterminal systems and present a bound for the achievable rate region for Multiple Description Coding systems as required in mobile systems.

Claim 2. An optimal Multiple Description Coding system is quasi deterministic.

Proof: Every optimal "Multiple Description Coding System" is "quasi deterministic" since such system can be represented as a set of pairs of encoder - decoder. The structure is visualized in Figure 1. The problem can be formulated as follows [28, 32]: There are four alphabets, the source alphabet U and three reproducing alphabets, possibly of different sizes V , W , and Y .

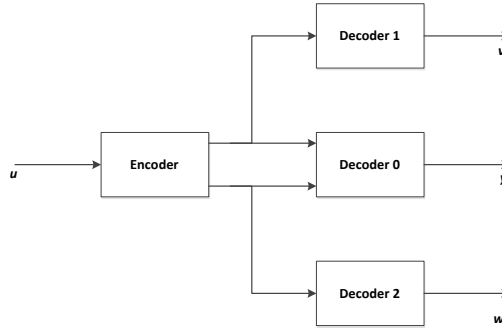


Figure 1. Multiple Description Coding System

Let d_0 , d_1 , and d_2 represent the distortion measure functions that describe the distortion suffered each time the source produces a letter $u \in U$ and the user receives letters $v \in V$, $w \in W$ and $y \in Y$, respectively:

$$d_0 : |U \times Y| \rightarrow R^+ \quad (5)$$

$$d_1 : |U \times V| \rightarrow R^+ \quad (6)$$

$$d_2 : |U \times W| \rightarrow R^+ \quad (7)$$

where R^+ non negative real number. Each distortion measure function is always a non negative real number – (like cost function, penalty, loss, etc.)

A stream of symbols from the source alphabet enters the encoder, and two bit streams leave the encoder at the rates R_1 bits per input symbol and R_2 bits per input symbol, respectively. Both bit streams enter the central decoder (Decoder 0). Only one bit stream enters each side decoder (Decoder 1 and Decoder 2). Side Decoder 1 is required to describe the source data using the alphabet V with average distortion over all the pair of sequences limited by D_1 under the distortion measure d_1 . Side Decoder 2 is required to describe the source data using the alphabet W with average distortion D_2 under the distortion measure d_2 . The central Decoder (Decoder 0) is required to describe the source data using the alphabet Y with average distortion D_0 under the distortion measure d_0 . The average per letter distortions for the output processes y , v and w for large values of block length n are not exceeding D_0 , D_1 and D_2 , respectively, i.e.:

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n d_0(u_k, y_k) \right) \leq D_0 \quad (8)$$

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n d_1(u_k, v_k) \right) \leq D_1 \quad (9)$$

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n d_2(u_k, w_k) \right) \leq D_2 \quad (10)$$

In order to obtain optimal system, each pair of encoder-decoder must preserve "quasi determinism" property and therefore the whole system is quasi-deterministic:

$$H(v|u) = H(w|u) = H(y|u) = 0 \quad (11)$$

Now the achievable rate region for a Multiple Description Coding system can be determined.

Definition 3. The Achievable Rate Region (ARR) for a source u with an a-priori known probability distribution P , is the region of all rate pairs (R_1, R_2) on the channels such that the deterministic mapping from the input process u to the output processes y , v , and w yields average per letter distortions not exceeding D_0 , D_1 and D_2 , respectively.

This means that for each value of distortion vector D_0 , D_1 and D_2 , there is an Achievable Rate Region (ARR) defined as the set of all rate pairs (R_1, R_2) for which there exist codes of distortion at most the distortion vector D_0 , D_1 and D_2 .

The next claim is about the Multi Rate Distortion Theorem, that states the Achievable Rate Region for a Multiple Description Coding System [28].

Claim 3. Let's assume a source u with an a-priori known probability distribution P . The Achievable Rate Region (ARR) is the region of all rate pairs (R_1, R_2) on the channels such that the deterministic mapping from the input process u to the output processes y , v , and w yields average per letter distortions not exceeding D_0 , D_1 and D_2 , respectively. The region ARR is given by all pairs of (R_1, R_2) defined by the convex hull bounded by the following inequalities:

$$R_1 + R_2 \geq I(u; y) \quad (12)$$

$$R_1 \geq I(u; y) \quad (13)$$

$$R_2 \geq I(u; y) \quad (14)$$

For all the deterministic mappings from the input process u to the output processes y , v , and w represented by all the possible conditional per letter probability distribution $Q = Q_{vwy|u}$, such that the average per letter distortions not exceeding D_0 , D_1 and D_2 , respectively. That is,

$$E\{d_0(u, y)\} \leq D_0 \quad (15)$$

$$E\{d_1(u, v)\} \leq D_1 \quad (16)$$

$$E\{d_2(u, w)\} \leq D_2 \quad (17)$$

Proof: All pairs of rates (R_1, R_2) in the region ARR must satisfy the limit of the output entropy. That is:

$$R_1 + R_2 \geq H(y) \quad (18)$$

$$R_1 \geq H(v) \quad (19)$$

$$R_2 \geq H(w) \quad (20)$$

and the average per letter distortions not exceeding D_0 , D_1 and D_2 given with the equations (15), (16) and (17). On the other hand, for all pairs of rates (R_1, R_2) not in the region ARR there is no machine with the above property.

Let Q be a transition matrix which simulate a multi-channel where the input process u is transformed into the output processes y ; v ; w defined on alphabets U , V , W , Y . The matrix Q approximates the data coding at the level of a single letter, even though, at the whole block level, the transformation is quasi-deterministic. At the level of a single letter the matrix Q represents the uncertainty of the transition of a single letter of alphabet U to the triplet defined on the alphabets V , W , Y . The set of all machines (encoder-decoder) that transform the input process u to the output processes v , w , y subject to the above fidelity constraints is described by the set

$$Q = \{Q_{vwy|u} \mid E\{d_0(u, y)\} \leq D_0, E\{d_1(u, v)\} \leq D_1, E\{d_2(u, w)\} \leq D_2, H(v|u) = H(w|u) = H(y, u) = 0, \} \quad (21)$$

Because $H(v|u) = H(w|u) = H(y|u) = 0$ it follows that the output entropies are equal to the mutual information terms:

$$H(v) = I(u; v) + H(v|u) = I(u; v) \quad (22)$$

$$H(w) = I(u; w) + H(w|u) = I(u; w) \quad (23)$$

$$H(y) = I(u; y) + H(y|u) = I(u; y) \quad (24)$$

Now the definition of the region described by equations (18), (19) and (20) can be written as:

$$R_1 + R_2 \geq I(u; y) \quad (25)$$

$$R_1 \geq I(u; v) \quad (26)$$

$$R_2 \geq I(u; w) \quad (27)$$

and the set of all rational conditional probability distribution Q described by

$$Q = \{Q_{vwy|u} \mid E\{d_0(u, y)\} \leq D_0, E\{d_1(u, v)\} \leq D_1, E\{d_2(u, w)\} \leq D_2\} \quad (28)$$

This represents the region of rates of distortion at most the distortion vector D_0 , D_1 and D_2 .

5. NEW APPROACH TO QUALITY OF SERVICE

The Achievable Rate Region for MDC Systems can be applied to define a new Quality of Service Approach. Quality of Service was defined by the ITU in 1994 [6] to include basic criteria on the characteristic parameters of a connection, such as service response time, loss, signal-to-noise ratio, cross-talk, echo, interrupts, frequency response, loudness levels, and so on. As a subset of telephony QoS is Grade of Service (GoS) criteria to include aspects of a connection in relation to the capacity and coverage of a network. An example is guaranteed maximum blocking probability and outage probability [40].

When considering the computer networks or other kinds of networks for data communication, some kind of resource reservation can be applied. With this approach to Quality of Service, the most important to consider is the possibility to differentiate the available resources for the end user. The level of performance depends on the type of subscription or on the type of user. In some cases, a certain bit rate, delay, jitter, blocking probability, packet dropping probability and/or bit error rate may be guaranteed. When the network capacity is not sufficient, guaranteeing good parameters for the Quality of service can be very important.

This is especially true in real-time streaming audio/video applications, such as VoIP, online gaming and IP-TV, as these often require fixed bit rate and are delay sensitive. The same reasoning applies for other networks with a limited capacity such as the mobile communication networks.

A network or protocol supporting QoS may agree on a traffic contract with the application software and reserve capacity in the network nodes, for example during a particular session in the establishment phase. During the session it may monitor the achieved performance level, for example the data rate and delay, and dynamically control the scheduling priorities in the network nodes. It could potentially release the reserved capacity during a tear down phase. In a best effort network or service usually Quality of Service is not supported. An alternative approach to control mechanisms is to provide a high quality communication over a best-effort network. This is done by over-provisioning the capacity so that it is sufficient for the expected peak traffic load. In such case network congestion is avoided and there is no need for QoS mechanisms.

Another definition for Quality of Service is related to services based on the application level. As an example of such services are telephony networks or video and audio streaming. When considering these kind of services the Quality of Service is used as a mean for the evaluation of the user satisfaction. From this point of view, QoS is the acceptable cumulative effect on subscriber satisfaction of all imperfections affecting the service. Other metrics known as Quality of Experience (QoE) is based on the user perception of the quality of multimedia content [41] and the required degree of satisfaction of the user or the targeted number of happy customers. Examples of measures and measurement methods are Mean Opinion Score (MOS), Perceptual Speech Quality Measure (PSQM) and Perceptual Evaluation of Video Quality (PEVQ).

Here a subset of QoS is presented, namely the pair of: **Quality of Fidelity (QoF)** and **Probability of Blocking**, which comprises aspects of a connection related to the capacity and coverage of a network, blocking probability and fidelity of the received digitized voice or video signal. Here is assumed a standard information source with a priori known probability distribution P . The information could be divided in half (or other fraction) and sent over two routes. If either route is blocked or disconnected, a reduced fidelity reproduction is still available to the receiver. As mentioned above, the probability of a failure (Blocking Event) of one channel is denoted by p .

This concept can be applied to the Quality of Service requirements. The values of the distortion vector D_0 , D_1 and D_2 , that guarantee a level of reliable communication system should be weighted by the appropriate probabilities of channel blocking events. The result of the pair

$$P_{total_block} = (p^2, QoF) \quad (29)$$

can be considered as a new approach to QoS (Quality of Service) taking into account all of blocking and loss of fidelity factors even for the expected peak traffic load. Here the probability of total failure (Blocking Event) of one channel is denoted by p , the probability of total failure (Blocking event) of two channels is denoted by p^2 and the probability of blocking the whole system is denoted by P_{total_block} .

Let P_0 is the probability that both channels function well, P_1 is the probability that only the first channel functions well and P_2 is the probability that only the second channel functions well. Obviously the probability of total failure (Blocking Event), that is both channels are blocked is the probability $P_{total_block} = 1 - P_0 - P_1 - P_2$. The value of Quality of Fidelity (QoF) is defined as the average fidelity in the system

$$QoF = \frac{P_0 D_0 + P_1 D_1 + P_2 D_2}{P_0 + P_1 + P_2} \quad (30)$$

The pair (P_{total_block}, QoF) can be considered as a new subset of QoS (Quality of service) taking into account all possibilities of blocking probabilities and loss of fidelity even for the peak traffic load in multi terminal network.

As an example let's consider a network with specific distortion vectors D_0, D_1 and D_2 . The distortion measure functions d_0, d_1 , and d_2 are known. The designer should specify the parameters of the channels: probability p of channel failure (due to blocking or other reason) and capacity C of the channels, according to requirements defined by the new definition of Quality of Fidelity QoF. Here is assumed a standard source with a priori known probability distribution P of all letters in alphabet U . Let's assume that all channels are identical, have the same unknown capacity C and the probability of total failure (like Blocking Event) of a channel is denoted by p . Let's assume that Failure Events (Blocking) of each channel are independent.

Thus, if there are have two available channels, the probability that both channels are blocked simultaneously is, assuming independence of the events of blocking a channel p^2 , and it is much lower than the single channel case. Obviously, $P_0 = (1-p)(1-p) = (1-p)^2 = 1 - 2p + p^2$ is the probability that both channels function well, and $P_1 = P_2 = p(1-p) = p - p^2$ are the probabilities that only the first, or the second channel are functioning well.

The rates (R_1, R_2) in the optimal system will be close to the channel capacity C . Here it is necessary to calculate the optimal rates which are actually equal to the channel capacity $R_1 = R_2 = C$ for a known standard source with probability distribution P of all letters in alphabet U (for example: typical human voice speaking English), known values of the distortion vector D_0, D_1 and D_2 and known distortion measure functions d_0, d_1 , and d_2 .

The optimal solution for C is located on the boundary of the Achievable Rate Region (ARR). To find the boundary region means to determine the Infimum over all the deterministic mapping from the input process u to the output processes y, v and w represented by all possible conditional per letter probability distributions

$$Q = Q_{vwy|u} \quad (30)$$

$$R_1 + R_2 = 2C = I(u; y) \quad (31)$$

$$R_1 = C = I(u; v) \quad (32)$$

$$R_2 = C = I(u; w) \quad (33)$$

such that the average per letter distortions are less or equal to D_0, D_1, D_2 , which are given with the equations (15), (16) and (17). The value of QoF, i.e. average fidelity in the system is described by the equation (30).

Finally, the value of probability channel failure (due to Blocking Event or something else) denoted by $P_{block} = p$ is easily calculated for a given value of Quality of Fidelity QoF. The conclusion is that the pair

$$P_{total\ block} = (p^2, QoF). \quad (34)$$

is a good criterion for the Quality of Service of communication system characterized by typical standard source u with probability distribution P of letters in alphabet U .

Some of the possible applications can be in the field of mobile multimedia, cellular networks and cloud computing in general. A digitized voice or video signal could be divided in half and sent over two routes. If either route is blocked or disconnected, a reduced fidelity reproduction is still available to the receiver. Multiple Description Coding improves Quality of Service and provides new service of rate adaptive streaming.

6. CONCLUSION

This paper demonstrated that Multiple Description Coding (MDC) is the right approach to pave the way for Mobile multimedia and Cloud Computing. Firstly it was provided the advantages of MDC “Send and Forget” method over technologies, such as TCP/IP, UDP and RTP. Then the Achievable Rate Region of MDC System was provided. Then a new approach for improving the quality of multimedia and other services in the cloud was presented. Here, a new subset of Quality of Service that takes into account issues of blocking in multi-terminal multimedia network and fidelity losses was introduced. Finally it was concluded that Multiple Description Coding system is an efficient and reliable communication system aiming to improve QoS (Quality of Service) by breaking a message into more distorted replicas, which together contain enough information to reconstruct the original message. A low fidelity reproduction can be obtained if only one channel is intact. Some of the possible applications can be in the field of mobile multimedia, cellular networks and cloud computing in general. A digitized voice or video signal could be divided in half and sent over two routes. If either route is blocked or disconnected, a reduced fidelity reproduction is still available to the receiver.

In order to provide high-quality video services to the users in wireless mobile terminals it is necessary to design video coding that has low coding complexity and strong error resilient ability. One possible solution is to make a full separation between multimedia and textual communication: TCP/IP, or similar protocol for textual communication, and Multiple Description Coding (MDC) Send and Forget method for Multimedia communication.

Despite the aforementioned advantages of MDC, SD codecs for multimedia communications still dominate in industry. However the cloud computing and next generation of mobile and wireless networks are technology that should adopt this breakthrough, since it is clear that the old protocols such as TCP/IP do not fit, and there will be increased demand for multimedia streaming applications. Mobile multimedia and cloud computing provide a very good area for application of Real-Time Multiple description coding.

In future, we plan to test the quality of multimedia content by using the Quality of Experience (QoE) metrics to determine if the new improved QoS parameters are correct.

REFERENCES

- [1] Antonopoulos, N.; Antonopoulos; Gillam, L.; , “Cloud Computing: Principles, Systems and Applications,” Springer Verlag, 2010.
- [2] Hayes, B.; , “Cloud Computing,” *Communications of the ACM*, vol. 51, pp. 9-11, 2008.
- [3] Dihn, T., H.; Lee, C; Niyato, D.; Wang, P.; , “A Survey of Mobile Cloud Computing: Architecture, Applications, and Approaches,” *Wireless Communications and Mobile Computing – Wiley*, DOI:10.1002 WCM/1203, pp. 1 – 38, 11 October 2011.
- [4] Kumar, J. P.; Badam, R.; , “Optimized Solutions for Mobile Cloud Computing,” *International Journal Of Computational Engineering Research (IJCER)*, Vol. 2 Issue. 4, ISSN 2250-3005, August 2012.
- [5] Ortiz, C., E.; , “On (Mobile) Cloud Computing – Multiple Perspectives to its Benefits, Drivers, and Economics,” 25 December 2009. <http://weblog.cenriqueortiz.com/mobility/2009/12/25/on-mobile-cloud-computing-multiple-angles-to-benefits-drivers-and-economics/> [retrieved: 01, 2014]
- [6] Peuhkuri, T. M.; , “IP Quality of Service,” *Helsinki University of Technology, Laboratory of Telecommunications*, 1999.
- [7] Xia, P.; Chan, S. H. G.; , “Optimal Bandwidth Assisgnment for Multiple Description Coded Video,” *IEEE CCNC 2009*, January 2009.
- [8] Kumar, N.; Tripathi, J.; Dubey, R.; Srivastava, A.; , “Optimal Bandwidth Assisgnment for Multiple Description Coded Video,” *International Journal of Engineering and Computer Science (IJECS)*, ISSN: 2319-7242, Vol. 2 Issue 5, pp. 1434-1438, May 2013.
- [9] Tang, Y.; Luo, J.-G.; Zhang, Q.; Zhang; Yang, S.-Q.; , “Deploying P2P networks for large-scale live video-streaming service,” *IEEE Communications Magazine*, vol. 45, no. 6, pp. 100–106, June 2007.

- [10] Li, B.; Liu, J.; , "Multirate video multicast over the Internet: An overview," *IEEE Network*, vol. 17, no. 1, pp. 24–29, January/February, 2003.
- [11] Goyal, V., K.; , "Multiple Description Coding: Compression Meets the Network," *IEEE Signal Processing Magazine*, September 2001.
- [12] Bai, H.; Wang, A.; Zhao, Y.; Pan, J., Abraham, A.; , "Distributed Multiple Description Coding Principles, Algorithms and Systems," *Springer-Verlag London Limited* ISBN 978-1-4471-2247-0, DOI 10.1007/978-1-4471-2248-7, 2011.
- [13] Fitzek, F., H., P.; Can, B.; Prasad, R.; Park, D.; Cho, Y.; , "Application of Multiple Description Coding in 4G Wireless Communication Systems," *World Wireless Research Forum (WWRF)* 8bis February. 2004.
- [14] Wang, Y.; Orchard, M., T.; Vaishampayan, V.; Reibman, A., R.; , "Multiple Description Coding Using Pairwise Correlating Transforms," *IEEE Transactions Image Processing*, vol. 10, pp. 351–366, March 2001.
- [15] "JPEG Standard, JPEG ISO/IEC 10918–1 ITU-T Recommendation T.81," 1992.
- [16] "JPEG 2000 Image Coding System, ISO/IEC International Standard 15444-1, ITU Recommendation T.800," 2000.
- [17] "ISO/IEC. JCT1/SC29 CD11172–2 MPEG1. International Standard for Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to 1.5 Mbps," 1991.
- [18] "ISO/IEC. JCT1/SC29 CD13818–2 MPEG2. Coding of Moving Pictures and Associated Audio for Digital Storage," 1993.
- [19] "ISO/IEC. JCT1/SC29 WG11/N3536 MPEG4. Overview V.15," 2000.
- [20] "ITU-T Draft ITU-T Recommendation H.261: Video Codec for Audio/Visual Communications at p x 64 kbps," 1993.
- [21] "ITU-T Draft ITU-T Recommendation H.263 (V1, V2, V3): Video Coding for Low Bit Rate Communications," 1996 – 2000kl/.
- [22] "JVT & ITU-T VCEG, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification H.264 (MPEG-4 Part 10)[S]," 7 – 14 March 2003.
- [23] Shannon, C. E.; , "A Mathematical theory of Communication," *Bell System Technical Journal*, 27. pages 379-423, 623-656.1948.
- [24] Shannon, C.E.; , "Coding Theorems for a Discrete Source with a Fidelity Criterion," *IRE Nat. Conv. Rec.* pages 142-163. 1959.
- [25] McMillan, B.; , "The basic theorems of Information Theory," *Annals Math. Stat* 24. pages 196-219, 1953.
- [26] Breiman, L.; , "The individual Ergodic Theorem of Information Theory," *Annals Math. Stat Vol.* 28 pages 809-811. 1957, and correction in Vol 31 pages 809-810, 1960.
- [27] Ornstein, D., S.; Shields, P., C.; , "Universal almost sure data compression," *Annals of Probability* 18: 441-452, 1990.
- [28] Sadeh, I.; , "PhD Thesis," *Tel-Aviv University*, 1993.
- [29] Sadeh, I.; , "The Rate Distortion Region for Coding in Stationary Systems," *Journal of Applied Mathematics and Computer Science*, Vol. 6, No.1, pp. 123 – 136, 1996.
- [30] Sadeh, I.; , "Universal Data Compression Algorithm based on Approximate String Matching," *Probability in the Engineering and Informational Sciences*, 10, pp. 465 – 486, 1996.
- [31] Sadeh, I.; , "Bound on Data Compression Ratio with a Given Tolerable Error Probability," *Probability in the Engineering and Informational Sciences*, 12, pp. 189 – 210, 1998.
- [32] Blahut, R., E.; , "Principles and Practice of Information Theory," *Addison-Wesley*, 1987.
- [33] Cover, T., M.; Gamal, E., A.; , "Achievable Rates for Multiple Descriptions," *IEEE Transactions Information Theory*, 1982.
- [34] Gray, R., M.; , "Sliding block source coding," *IEEE Transactions Information Theory*, 1975.
- [35] Ahlswede R.; , "The rate distortion region of a binary source for multiple descriptions without excess rate," *IEEE Transactions Information Theory*, 1985.
- [36] Zhang, Z.; Berger, T.; , "New Results in Binary Multiple Descriptions," *IEEE Transactions on Information Theory*, 33: 502-521, 1987.
- [37] Effros, M., F., H.; , "On the Rate Loss of Multiple Description Source Codes and Additive Successive Refinement Codes," *Proceedings ISIT*, 2002.
- [38] Comer, D., E.; , "Interworking with TCP/IP: Principles, Protocols and Architecture Vol. 1," *Prentice Hall*, 5th Edition, 2006.
- [39] Gray, R., M.; Neuhoff, D., L.; Omura J., K.; , "Process Definitions of Distortion Rate Functions and Source Coding Theorems," *IEEE Transactions Information Theory*, 1975.
- [40] "Teletraffic Engineering Handbook," *ITU-T Study Group 2*, Geneva, December 2003.
- [41] Franken, L.; , "Quality of Service Management: A Model-Based Approach," PhD thesis, *Centre for Telematics and Information Technology*, 1996.

BIOGRAPHY OF AUTHORS



Ilan Sadeh received BSc Degree in Electrical Engineering and Electronics at Technion Aifa, Israel. He obtained Master of Science Degree in Computer and Electrical Engineering at Tel Aviv University, Israel and Master of Science Degree in Mathematics and Computer Science at Tel Aviv University, Israel. He received his PhD Degree in Mathematics and Computer Science at Tel Aviv University, Israel. From 1981 to 1985 he was working as an engineer at Research and Development Department at Israeli High Tech Industry. From 1985 to 1987 he was a manager at the Research and Development Department of Control, Communications, and Navigation Systems at Melitav Engineering Ltd. From 1990 to 1993 he was research and lecturer at Computer Science Department at Tel Aviv University, Israel. From 1996 to 2010 he was CEO at Research and Development Department for Video Transmission, Video Compression and Digital Signal Processing at Visnet Ltd in Israel. From 2000 to 2004 he was CEO at the research and development department in Control Systems, Smart CODECs, Video Compression at Vipeg Ltd in Israel. From 2011 to 2013 he was professor at the University for Information Science and Technology “St. Paul the Apostle” in Ohrid, Macedonia. He received many awards, among them the INTEL award for research in telecommunications and computer science in 1994 University for Information Science and Technology “St. Paul the Apostle” in Ohrid, Macedonia,. His research interests are video compression, and coding, video and image processing, automata theory, digital signal processing, etc.



Stojan Kitanov received Bachelor's Degree in Electrical Engineering in Electronics and Telecommunications at Ss Cyril and Methodius University, Faculty of Electrical Engineering in Skopje, Macedonia. He obtained Master of Science Degree in the area of 4G Mobile Networks at the University of Essex, in United Kingdom. Currently he is enrolled as a PhD student at the Ss Cyril and Methodius University, Faculty of Electrical Engineering and Information Technologies in Skopje, Macedonia. From October 2005 to August 2010 he was Systems Support Engineer in Motorola. He provided technical support to ONE, the 2nd mobile operator in Macedonia. He also participated in the deployment of many GSM Base Stations, GSM EDGE deployment, WiMAX deployment, as well as software upgrade of GSM and WiMAX live networks. During his employment he completed many training courses at the Motorola Training Center in Swindon, UK, and obtained many qualifications for GSM and WiMAX Networks, and Quality Control. Currently he works as a teaching and research assistant at the University for Information Science and Technology “St. Paul the Apostle” in Ohrid, Macedonia, since December 2010. His research interests include next generation mobile and wireless networks, future optical networks, distributed systems, GRID Networks, cloud computing, multimedia networks and services, and future internet architecture.



Danco Davcev obtained the degree in Engineer in Electronics and Computer Science from the Faculty of Electrical Engineering at the University of Belgrade (Yugoslavia) in 1972. He received his “Doctor of Ingenieur” degree in Computer Science from the Paris XI University (Paris, France) in 1975, and PhD degree in Informatics from the University of Belgrade (Yugoslavia) in 1981. From 1984 – 1985 he was a Fulbright postdoctoral research scientist at the University of California, San Diego (USA). In 2001/2002 he was a visiting Fulbright Research Professor at several universities in USA, where he was working on distance learning projects, cross-cultural educational programs, augmented reality systems, and multi agent systems. He participated on several international conferences and gave several invited lectures and seminars. At the moment he is professor in Computer Science and Engineering, at the Ss Cyril and Methodius University in Skopje, Macedonia. He has more than 300 papers presented on international conferences, or published in international journals in computer science and informatics. His research interests include multimedia databases and educational systems, spatio-temporal data and 3D objects, intelligent information systems and data fusion, distributed systems, cloud computing, wireless sensor systems, mobile computing, etc. He is a senior member of ACM and senior member of IEEE.