

# Priority-Aware Packet Pre-marking for DiffServ Architecture Based on H.264/SVC Video Stream Structure

Slawomir Przylucki<sup>1</sup> · Dariusz Czerwinski<sup>1</sup>

Published online: 1 October 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** The H264/SVC codec allows for generation of hierarchical video streams. In the stream of this type video data belonging to different layers have different priority depending on their importance to the quality of the video and the decoding process. This creates new demands on the mechanisms of packet marking, and thus new challenges for the policy guaranteeing QoS parameters, such as those defined in the DiffServ architecture. Therefore, mechanisms of the traffic engineering used in the DiffServ network should, as far as possible, take into account internal distribution of priorities inside video streams. This may be achieved by implementing an appropriate method for packet pre-marking. The paper describes the Weighted Priority Pre-marking (WPP) algorithm for priority-aware SVC video streaming over a DiffServ network. Our solution takes into account the relative importance of the Network Abstraction Layer Units. It also does not require any changes in the implementation of the DiffServ marker algorithm. The results presented confirm that video transmission in the DiffServ domain, based on the WPP packet pre-marking, can provide better perceived video quality than the standard (best effort) streaming of multi-layered SVC video. In addition, a comparison with the transmission of the same video content encoded with the H264/AVC codec also points to the superiority of our proposed method.

**Keywords** QoS · DiffServ · Packet pre-marking · SVC video

---

✉ Slawomir Przylucki  
s.przylucki@pollub.pl

Dariusz Czerwinski  
d.czerwinski@pollub.pl

<sup>1</sup> Institute of Computer Science, Technical University of Lublin, ul. Nadbystrzycka 36b, 20-618 Lublin, Poland

## 1 Introduction

IP networks were developed as a transmission environment providing no built-in Quality of Service (QoS). With the development of Internet services, the problem of defining mechanisms to guarantee the transmission parameters became an important issue. The Internet Engineering Task Force (IETF) recommended the Differentiated Service (DiffServ) architecture [1, 2] as one of the methods to guarantee QoS parameters. The idea of differentiated services is based on a simple model, which classifies IP packets to the appropriate groups (aggregates) and defines the allocation of a certain amount of forwarding resources for each aggregate. In recent years DiffServ confirmed its usefulness in IP networks, offering scalability and manageability. One of the key issues associated with the efficiency of a DiffServ domain is the process of classification and marking of IP packets. This process can take place at the data source (pre-marking) or at the edge of the DiffServ domain. For many years the second solution has been dominant. Edge routers used the packets marking algorithms in colour-blind mode, i.e. they did not take into account any outside information about the priority of the analysed IP packets.

The growing popularity of Internet multimedia services was one of the reasons to change this situation [3]. In this context the crucial role has been played by the development of video streaming services and the significant progress in the area of video encoding algorithms. Currently, the internal structure of video streams is complex and depends not only on the type of video codec but also on the specifics of a streaming service. The priority of an individual packet inside the stream is known at the video source but its reassignment to the packet on the edge of a DiffServ domain becomes extremely difficult if at all possible. This has resulted in renewed interest in the use of packet pre-marking and implementation of algorithms for marking in colour-aware mode at the edge of a DiffServ domain [4, 5]. Pre-marking is especially promising because it allows the source of a video stream to indicate priority. This opens the possibility to take into account the specific characteristics of video coding, and protection of these packets, which are of great importance to maintain an acceptable level of the perceived video quality [6].

Recently, the question of guaranteeing the QoS parameters in video streaming services has become even more complex. The last few years have brought tremendous development of multimedia devices and applications. Therefore, it becomes necessary to deliver a real-time content not only in accordance with static rules defined for a given class of the network traffic. Attention should also be paid to the different technical parameters of receivers and thus different expectations of clients regarding to the quality of the received video. A good example of the solution that meets these requirements is the Dynamic Adaptive Streaming over HTTP (DASH) standard published by the IETF in 2012 [7]. The mechanisms of pre-marking and the rules of packet forwarding inside DiffServ domains should meet the requirements of these relatively new solutions.

Taking into account the diversity of end-user expectations and problems posed by the characteristics of modern video transmission systems, the Scalable Video Coding (SVC) is a highly attractive solution for video streaming. It enables the scalability in spatial, temporal, and quality (SNR) domains, while keeping compression at high efficiency [8]. However, at the same time scalability leads to a complex, hierarchical stream structure of the SVC video. In this structure, the priority of the data depends on the position within the individual layers as well as on the inter-layer relationships. For this reason, pre-marking methods developed for H.264/AVC may not be simply and directly applied to the case of a system for the streaming of the layered video. Therefore, any proposal of the priority-

aware pre-marking method must be based on the analysis of the new structure of a SVC stream. Moreover, the verification of a solution should take into account SVC video streams generated by the standard streaming services as well as by more advanced services such as DASH. Based on these assumptions, we develop a new Weighted Priority Pre-marking (WPP) algorithm, which takes into account the relative importance of data within the SVC video stream and does not require any changes in the DiffServ marker algorithm. It allows to obtain a better perceived video quality than for SVC video transmission without pre-marking. The streaming system employing the WPP algorithm also showed superiority (in terms of the video quality) over the system for the H.264/AVC video transmission with the TypeMapping pre-marking [9, 10].

The remainder of this paper is organised as follows. The principles of the DiffServ architecture and the process of the packet marking for video transmission systems is presented in Sect. 2. Additionally, Sect. 2.2 presents an overview of the promising proposals that relate to the streaming video in a DiffServ domain. Section 3 describes the coding rules for the Scalable Video Coding. It also contains discussion of the SVC bit stream ordering and the gradation of the video quality Sect. 3.1. There is the principle of defining the Operating Point (OP) and the structure of the SVC stream used in video streaming services based on the DASH standard. The developed WPP pre-marking algorithm is presented in Sect. 4. The description of the testbed, assumptions used during the simulation (Sect. 5.1) and the experimental results (Sect. 5.2) are in Sect. 5. Finally, a brief conclusion is made in Sect. 6.

## 2 Service Differentiation

The primary purpose of the DiffServ architecture was the implementation of scalable service differentiation in the Internet. The principles of DiffServ have been described in the RFC2475 [1]. The scalability is achieved by aggregating traffic classification state and IP packets marking using the Differentiated Services (DS) field. That field is located in the IP packet header. The first three bits of the DS field are used to determine the traffic class, while the next three define the packets rejection probability. The last two bits are left unused [11]. All these 6 bits create so called Differentiated Services Code Point (DSCP). The traffic management, according to the DiffServ model, is carried out only at the boundary of the DiffServ domain. This means that all operations, such as classification, marking, policing, and shaping need only be implemented on the border nodes (routers). Also, classification and marking can be a part of functionality of source hosts (on the nodes that are sources of the network traffic associated with a given service). Traffic streams of marked IP packets receive a particular per-hop behaviour (PHB). A particular PHB defines the allocation of a certain amount of forwarding resources (buffer space and bandwidth) to these traffic streams along their path. The marked packets may belong to one of four basic PHB groups depending on DSCP values [11, 12]. These PHB groups are respectively:

- Best Effort—BE
- Assured Forwarding—AF
- Expedited Forwarding—EF
- Class Selector—CS

According to the IETF, a PHB group designed for video streaming is the Assured Forwarding (AF) [13, 14]. This group offers different levels of forwarding assurances for

IP packets, while accomplishing a target throughput for each network aggregate. Within the AF PHB group, IP packets are marked and then forwarded with a specified value of drop precedence. The DiffServ defines four independent AF classes. Within each AF class, an IP packet is assigned one of three different levels of drop precedence. In other words, a single AF PHB group consists of three PHBs, and uses three DSCPs as is shown in Fig. 1.

### 2.1 Packet Marking

Generally, packet marking is often called packet colouring and refers to the setting of bits in the DS field which represent dropping precedences. Packets are coloured in green, yellow and red. In the case of the AF PHB group green means AFx1, yellow AFx2 and red AFx3. The IETF in RFC4594 [11] recommends AF3x class for services that require near-real-time packet forwarding of network traffic and are not delay sensitive. These characteristics are consistent with the requirements of video streaming applications. This class has been featured in Fig. 1 by a thin black border. IETF also recommends that the applications or IP end points should pre-mark their packets with DSCP values (so called colour-aware marking) or the router topologically closest to video source should perform the classification and mark all packets as AF3x (so called blind marking). The most popular, standard algorithms of packet colouring are as follows:

- Single Rate Three Colour Marker (srTCM) [15],
- Two Rate Three Colour Marker (trTCM) [16],
- The Time Sliding Window Three Colour Marker (TSW3CM) [17].

In order to carry out the task of marking packets belonging to the AF class, IETF recommended a Two Rate Three Colour Marker (trTCM). The trTCM is a combined metering and marking algorithm. It consists of two token buckets with the token accumulation rate dedicated to each one of them. The first token bucket (indicated in Fig. 2 as B1) has the Committed Information Rate (CIR) and the second token bucket (indicated as B2) has the Peak Information Rate (PIR). The Peak Burst Size (PBS) is used as the size of the B2 bucket and Committed Burst Size (CBS) is used as the size of the B1 bucket. The B1 bucket is incremented with the rate of CIR while the B2 bucket is incremented with a PIR rate. The trTCM can operate in two modes, colour-blind and colour aware respectively. In the case of the colour aware mode, the trTCM assumes that the packets have already been colored by any previous entity (pre-marking). The algorithm of this operating

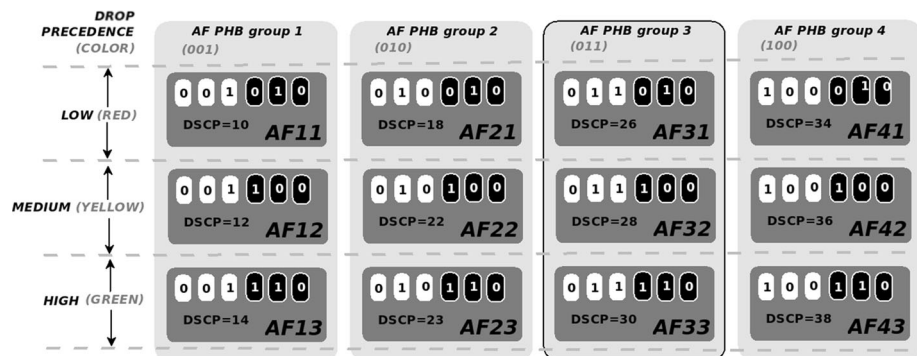
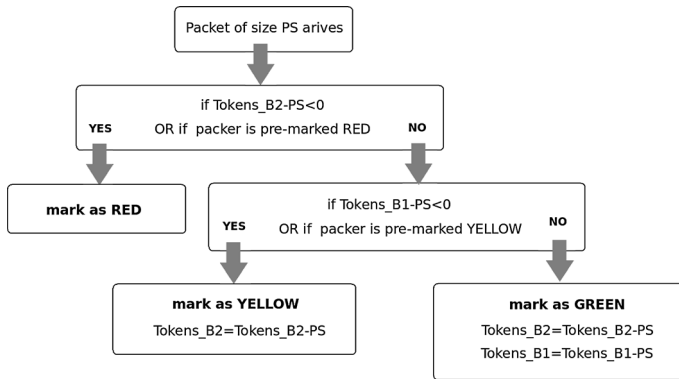


Fig. 1 Assured Forwarding (AF) PHB groups as defined in the RFC 2597



**Fig. 2** Color-aware mode algorithm for TRTCM

mode is presented in Fig. 2. In the colour-blind mode all incoming packets are treated equally.

## 2.2 Related Work

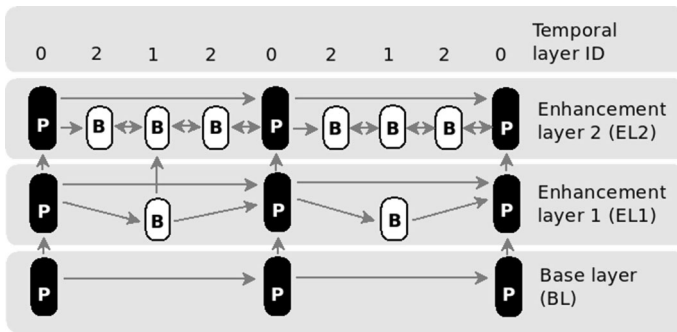
The pre-marking scenario can potentially apply the hierarchical structure of the modern video stream. It offers a possibility to protect that part of the video stream which is the most important in terms of quality and effectively react to congestion on transmission links. Unfortunately, despite the above-mentioned mechanisms and recommendations, defining the method of pre-marking packets by video source is not a trivial task. Especially in the case of a SVC video, it requires the analysis of the specific schema of coding, taking into account the structure of Group of Pictures (GOP) and relevance of GOP components to the perceived quality of video. Over the last few years, many classifications and marking strategies have been proposed for different types of video codecs. Regarding today's most popular video coding algorithms it is worth pointing out a few promising proposals that relate to the streaming video data through the DiffServ domain. In the case of the H.264/AVC codec which is the widely used MPEG standard for video encoding [18, 19] one of the earliest proposed solutions is to include information about the type of frames in the packet classification process [20]. Usually, this method involves the simple frame type mapping (I,P,B) to DCSP values (packet colours) [9, 10]. Other popular proposals are based on the analysis of the loss impact of frames or slices on video quality. In [21] that loss is estimated by means of counting how many times the frame is taken as reference by previous or future elements. Further studies have used the idea of the classification and marking of packets based on the identification of perceptually important video regions [22]. Their extension was the proposal of two dimensional analysis of packet loss [23]. In the first, temporal dimension the significance of a lost packet is computed based on the estimated error propagation and in the second, spatial dimension, the algorithm computes the packet significance based on the content complexity. Among the recent and most commonly cited proposals it is worth paying attention to the solution published in [24]. It introduces the concept of marking probabilities and methods for their estimations in conjunction with the relative importance of the IP packet in terms of perceived video quality and the traffic conditions along the forwarding path.

The introduction and acceptance of the new algorithm H.264/SVC [8] made the selection of a method for packet marking an even more complex issue. One of the visible changes is the focus of attention not only for pre-marking mechanisms but also for modification of marking algorithms inside the DiffServ domain. A good example of this approach is an Enhanced Token Bucket Three Colour Marker (ETBTCM) presented by Ke et al. [25]. A similar approach can be found in the another proposal, published in [26]. The Improved Two-Rate Three-Colour Marker (ITRTCM) marks video packets according to the current vacancy degrees of the token buckets and the relative significance of the packets. Both markers use a number of thresholds for dividing the relative significance into a few grades. Unfortunately, the more significant the divisions, the higher the possibility of a mistake and degradation of perceived video quality. That problem has been partially solved in [27]. The authors there describe a new marker called Priority-Aware Two-Rate Three-Colour Marker (PATRTCM) which allows to minimise inaccuracy when the token count is close to the thresholds. Further development of the use of the ITRTCM marker has been proposed in [28]. The solution consists of a source marking scheme based on NALU priorities and ITRTCM as an edge router marker. Another group of solutions is focused on the QoE-based traffic management techniques for SVC video streaming. Study in [29] demonstrates the QoE-aware traffic management for scalable mobile video delivery within the MEDIEVAL architecture. The neuro-fuzzy scheme, described in [30], regulates output rate using a buffer and ensures that video streams from host to client conform to desired traffic conditions. The latest developments in this area are also using artificial intelligence algorithms and based on Software Defined Network principles [31]. Finally, we should also mention the very promising results of the analysis of traffic patterns [32] that could potentially allow bonding marking policies and IP traffic behaviour.

### 3 SVC Coding

Scalable video coding (SVC) introduces the concept of layered video stream. The fundamental idea of a SVC is to enable the removal of parts of the coded video data by rejecting certain layers. The quality and throughput of the video stream depend on the number of its layers. On the other hand, the process of the layers selection ensures that the resulting stream will continue to be correctly decoded by the receiver. In order to achieve this in practice, multi-layer stream of SVC video consists of one base layer (BL) and several enhancement layers (EL). The terminal devices with different technical parameters can choose to receive partial stream, e.g., base layer in case of mobile devices, and all layers for HD screens [8, 33]. The H.264/SVC standard was created as an extension of the H.264/AVC codec. For this reason, individual layers rely on the principles of the H264/AVC coding while simultaneously intra-layer relationships provide three types of scalability: temporal, spatial and quality (SNR), respectively. The hierarchical structure of the SVC stream is presented in Fig. 3.

According to the H.264/SVC, each spatial dependency layer requires a dedicated prediction module to perform both motion-compensated prediction and intra prediction within the layer. Additionally, the SVC coding algorithm introduces new modules closely related to the video quality. The first one, the SNR refinement module, provides the mechanisms for quality scalability within each layer, and the second one, the inter-layer prediction module, is responsible for the dependency management between subsequent spatial layers.



**Fig. 3** Hierarchical SVC stream

As the end result, different temporal, spatial and SNR levels are simultaneously integrated into a single scalable video stream.

Functionally, the H.264/SVC is divided into two parts, the Video Coding Layer (VCL) and the Network Abstraction Layer (NAL) [19]. The VCL produces the coded representation of the source video and the NAL formats these video data by means of the header information [8]. A NAL unit consists of a header and a payload part. The three fields inside the header are relevant to the issues discussed: DID (dependency id) which indicates the inter-layer coding dependency level of a layer representation, QID (quality id) which indicates the quality level of an SNR layer representation and TID (temporal id) which indicates the temporal level of a layer representation.

### 3.1 SVC Bit Stream Ordering and Gradation of Video Quality

For obvious reasons, marking packets belonging to the SVC video stream should lead to the protection of these layers, whose contribution to the perceived quality of the received video content is highest. In order to assess the level of this contribution, the relationship between video quality and the structure of the SVC stream must be analysed. As it was mentioned earlier, SVC coding provides three types of scalability: temporal, spatial and quality, respectively. Spatial scalability allows for selection of the video frame size. On the basis of the tests we carried out, it can be stated that changes of resolution definitely lead to the negative quality assessment by recipients of a given video content. Even the existing upscaling algorithms (eg. inside mobile devices and TV sets) only partially reduce this adverse effect. An impact of changes in resolution, particularly their frequent occurrences during video playback on the perceived video quality was also investigated in [34, 35]. The published results indicate a similar observations. Another type of scalability, temporal scalability, allows for a change of the video stream bit rate by changing the number of frames per second. Practical implementation of this scalability is doubling the number of frames per second in each subsequent enhancement layer relative to the base layer. Unfortunately, the short-term lowering of the video frame rate is not accepted by the end-users [35, 36]. The SNR scalability is the third type of scalability offered by the SVC encoding. The H264/SVC standard provides two ways of its implementation. The first one is the Coarse Grain Scalability (CGS). It employs inter-layer prediction mechanisms (residual, motion parameters and macroblock mode predictions). The second implementation is the Medium Grain Scalability (MGS) and it splits each SNR enhancement layer

into several sublayers (MGS layers). Thanks to that schema, the finer quality granularity can be obtained [37].

The hierarchical structure of the SVC stream stresses the need to decide on the method of determining the priority of the single NALU. This issue is directly related to the methods of the Fix Priority Ordering (FPO) [39]. Taking into account the specific characteristics of the SVC encoding, obvious solutions are: temporal-based (or frame-based), spatial-based and SNR-based algorithms, where the video bit stream is arranged first by the temporal, the spatial and the SNR layer, respectively. It should be noted that in the JSVM reference software [38] the default bit stream ordering is spatial-based (layer order: spatial-SNR-temporal). Xiao et al. in [40] presented test results for various FPO configurations. Their research indicates that for a wide range of bit rates, the best video quality is achieved for the temporal-based FPO (layer order: temporal-spatial-SNR) or SNR-based (layer order: SNR-temporal-spatial). They also proposed the adaptive FPO (APO) configuration. The APO method arranges the H.264/SVC bit stream according to contribution of different layers to the whole performance within a given Group of Pictures (GoP). Unfortunately, since each GoP may have different characteristics, the optimal bit stream may vary from one GoP to the other in the same video sequence.

### 3.2 Gradation of Video Quality

Another issue to be considered in the context of the application of SVC coding in video streaming applications is that the GOP structure does not necessarily include all possible layers. A good example of such behaviour is the adaptation mechanism defined in the DASH standard [7]. It can be described as a process of elimination of certain enhancement layers or as a change of Operation Point (OP) of a given SVC stream. In the first Operation Point OP1 each GOP includes a full set of video data. Removal of the individual enhancement layers leads to define the next OPs. Of course, the process of removal of certain layers must take into account the existing dependencies between them. For example, one cannot remove the layers which are used as references for any others. Switching between the OPs allows to adjust the the bit rate of SVC video stream to conditions on the transmission link. Figure 4 illustrates an exemplary internal structure of the GOP for OP1 and OP5.

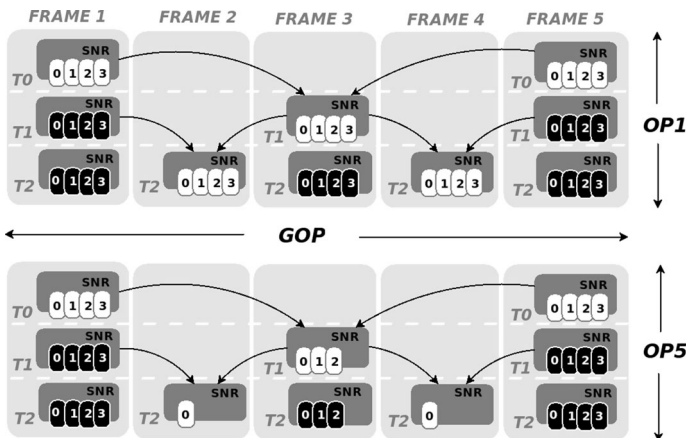


Fig. 4 The internal structure of the GOP for OP1 and OP5



In the case illustrated in Fig. 4, the GOP consists of five frames and contains three temporal layers (T0–T2) and four SNR layers (SNR0 - SNR3). The number 0 indicates the base layer. The Fig. 4 does not show the spatial layer because it was assumed that the stream represents video with a resolution of the near (but not exceeding) resolution supported by the receiving device. This choice should not be changed by the adaptation process. According to the information mentioned above and our own research results, temporal layers should also be protected. This leads to the solution in which the decoder should have, for each of temporary layer, at least a minimum set of video data (base layer SNR0). Because each temporal layer is associated with a double number of frames per second, this assumption maintains a high level of video smoothness. In summary, the set of OPs are defined by removing successive SNR layers. This scheme to define the structure of the SVC stream allows to provide the best quality of received video content [40, 41]. The internal structure of the GOP for OP5, defined according to the above assumption is also shown in Fig. 4. As was stated at the beginning, the illustration presented applies if the spatial and temporal layers are to be protected. Of course, in the same way different sets of OPs can be created, preferring other layers of the SVC stream. Detailed analysis of quality gradation by OP section can be found in [41]. All possible OPs for a particular SVC video stream can be specified as it is done in Fig. 5. The illustrated exemplary structure of the video stream is analogous to the previous example (includes three temporal layers and four SNR layers for each of them).

The combination of the need to determine the priority of the single NALU (to protect the selected layers) and the possibilities for creating OPs leads to defining a set of scalable SVC video streams which can potentially be transmitted over an IP network using DiffServ rules. Therefore, three FPO configurations (JSVM default and best two from [40]) and a full set of OPs for each of FPO were selected for further testing.

### 4 Weighted Priority Pre-marking

The video packets need to be classified into different priorities according to their relative importance before any pre-marking algorithm can be applied. Let us assume that the relative importance of the NALU is represented by  $(S_i, SNR_j, T_k)$  where  $S_i$  is the  $i$ -th spatial layer,  $SNR_j$  is the  $j$ -th quality layer and  $T_k$  is the  $k$ -th temporal layer. Let  $P(S_i, SNR_j, T_k)$  be the priority of the NALU. We assign the weights to the individual layers by giving values 1, 2 and

OP10	X				X				X				
OP9	X	X			X				X				
OP8	X	X	X		X				X				
OP7	X	X	X	X	X				X				
OP6	X	X	X	X	X	X			X				
OP5	X	X	X	X	X	X	X		X				
OP4	X	X	X	X	X	X	X	X	X				
OP3	X	X	X	X	X	X	X	X	X	X			
OP2	X	X	X	X	X	X	X	X	X	X	X		
OP1	X	X	X	X	X	X	X	X	X	X	X	X	X
	SNR0 SNR1 SNR2 SNR3 SNR0 SNR1 SNR2 SNR3 SNR0 SNR1 SNR2 SNR3												
	← T0 →				← T1 →				← T2 →				

Fig. 5 Definition of OPs for SVC video stream

3 to the most important scalability layer, less important scalability layer and the least important scalability layer, respectively. In such a case, the priority for a given NALU can be expressed by Formula (1)

$$P(S_i, SNR_j, T_k) = W_S S_i + W_{SNR} SNR_j + W_T T_k \tag{1}$$

where  $W_S, W_{SNR}, W_T$  are weights for spatial, SNR and temporal layers, respectively.

Assigning weight values for each type of scalability layers of the video stream depends on the particular provider of video streaming service. This choice is important because it affects the priority of the subsequent extended layers. The higher the weight value, the quicker decreases the priority assigned to successive, extended layers, which represent a particular scalability. In other words, this choice has a direct influence on the quality of the video. Therefore, Sect. 5 presents the experimental verification of our theoretical assumptions (see Sect. 3) for the most popular streaming schemes based on the H264/SVC.

In order to use the DiffServ marking algorithm trTCM in colour-aware mode, the mapping of priorities to the three colours (DSCP codes) is necessary. Constructing the principle of such mapping, we took into account two main reasons:

- if an operator of a video service decides to increase (or decrease) the number of extended layers inside the SVC streams, the policy of the priority mapping should also be modified. A number of protected layers that are most important to preserve the best possible perceived quality of the received video, should grow with the increasing number of OPs,
- in a case of congestion on the transmission path, the final decision to change the priority belongs to the mechanism of packet colouring which is implemented on the edge of a DiffServ domain. The even distribution of the packets between different groups (colors) should be a neutral solution from the point of view of the potential competition for the available bandwidth between the multiple data and video streams.

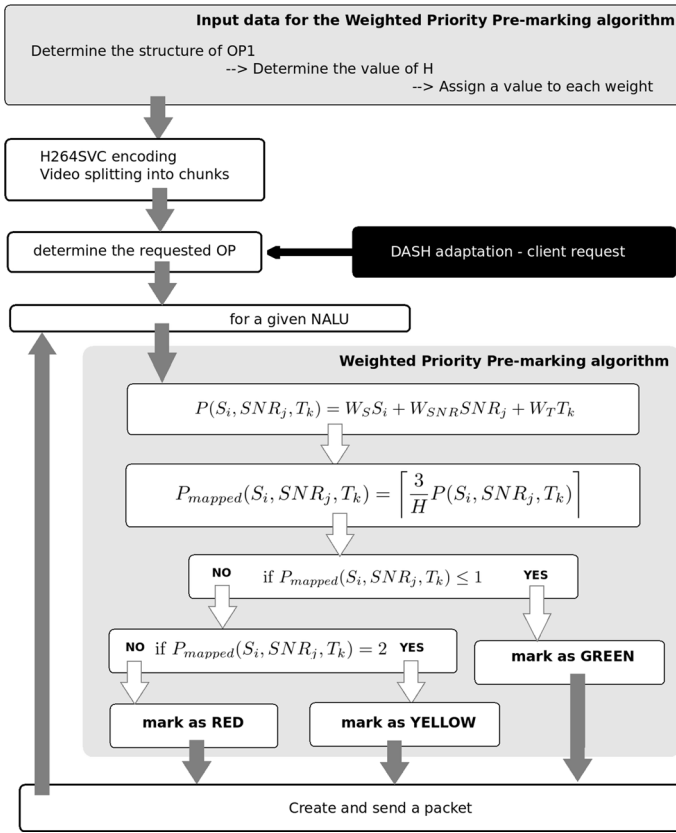
Taking into account the above issues, we proposed the following solution. Let  $H$  be the highest value of priority calculated according to Formula (1) for a given structure of the OP. The priority after mapping can be obtained using the formula as follows.

$$P_{mapped}(S_i, SNR_j, T_k) = \left\lceil \frac{3}{H} P(S_i, SNR_j, T_k) \right\rceil \tag{2}$$

The last step is to apply the following rules of marking:

- if  $P_{mapped}(S_i, SNR_j, T_k) \leq 1$  then use pre-marking as green,
- if  $P_{mapped}(S_i, SNR_j, T_k) = 2$  then use pre-marking as yellow,
- if  $P_{mapped}(S_i, SNR_j, T_k) = 3$  then use pre-marking as red.

That algorithm consists of Formula (1), Formula (2) and the scheme of packet colouring, we named Weighted Priority Pre-marking (WPP). Figure 6 shows the WPP algorithm and its use in a video streaming system based on the H264/SVC and the principles of the standard DASH. The provider of the DASH streaming service must determine the structure of the SVC stream (OP1) and the priorities attached to particular types of scalability (weight values). Based on these assumptions, the highest value of priority is calculated (value of H). Weight values and the parameter H are necessary for operation of the WPP algorithm. Next, the video content is encoded using encoder H264/SVC and divided into chunks in accordance with the standard DASH. Each request generated by a client of the DASH streaming service leads to the necessity of sending the video fragment with specific parameters. According to the



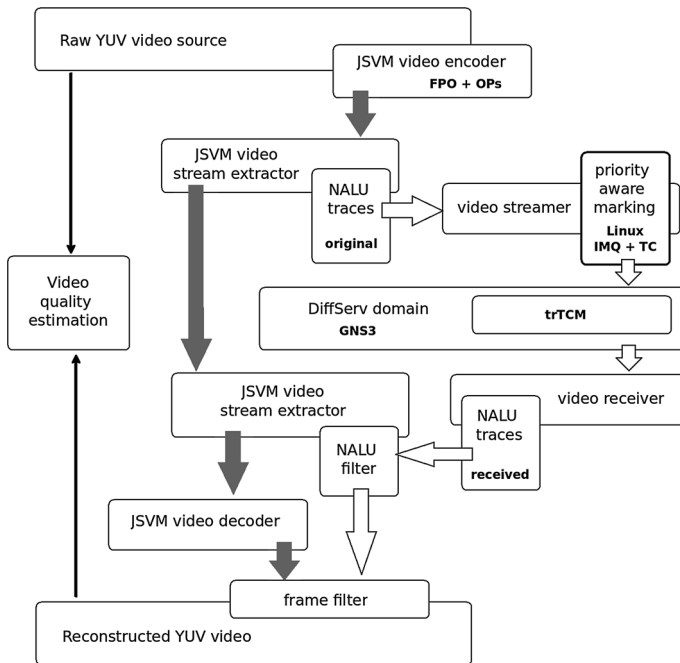
**Fig. 6** The use of the WPP algorithm in a video streaming system based on the H264/SVC coder and the principles of the DASH

principles of adaptation DASH these parameters can vary between successive requests. This means that it is necessary to determine, in each such case, the new OP of the SVC stream. At this stage, the final structure of the video stream is known. The system is ready for the process of packet pre-marking according to the algorithm WPP.

### 5 Test Scenarios

The main aim of the tests is to determine the properties of the proposed WPP algorithm in combination with different methods of packet pre-marking. For this purpose, the video streams have been generated for the three SVC FPO bit streams ordering with a full set of OPs. The packets belonging to these streams were then pre-marked (colored). Finally, they were sent through the DiffServ domain, which uses the trTCM marker in colour-aware mode. Therefore, the video quality was analysed for following schemas:

- S-SNR-T: WPP pre-marking order S-green, SNR-yellow, T-red (JSVM default)
- T-S-SNR: WPP pre-marking order T-green, S-yellow, SNR-red
- SNR-T-S: WPP pre-marking order SNR-green, T-yellow, S-red



**Fig. 7** The testbed structure

The conclusion of these tests we used as a starting point for the comparative analysis of the proposed WPP algorithm. The best WPP configuration is compared with two other scenarios of the video transmission in the DiffServ domain. The first one is the transmission of the pre-marked stream of single layer H.264/AVC video [9, 10] and the second one is the transmission of the SVC stream without pre-marking. All the compared solutions were tested at the comparable network configuration and used the same video source.

## 5.1 Testbed Configuration

The coding part of the testbed system is built up on the basis of the JVM [38, 42]. The network part consists of two elements: the software framework SVEF [43] and the network simulator GNS3. The first one, the SVEF, allows to obtain the desired order in the SVC bit streams and was responsible for the proper generation and processing of the SVC traces. This packet has also been used to estimate the quality of the transmitted video. The DiffServ domain was implemented in GNS3. The trTCM marker (in color-aware mode) was configured on the ingress router. The structure of the developed testbed is presented in Fig. 7.

The test video was a Foreman sequence, which has 2000 frames with a GoP size of 8. The structure of the stream consists of the following layers:

- spatial (SL0—QCIF, SL1—CIF, SL2—CIF)
- temporal for SL0 and SL1 (TL0—3.75 Hz, TL1—7.5 Hz, TL2—15.0 Hz)
- temporal for SL2 (TL0—3.75 Hz, TL1—7.5Hz, TL2—15.0 Hz, TL3—30Hz)
- SNR (3 SNRL layers: BL and two EL for each spatial-temporal layer)

The above video bit stream has competed with one ON-OFF background traffic flow, which had an exponential distribution with the mean packet size of 1000 bytes, burst time 200 ms, idle time of 50 ms, and the rate of 500 kbps. The test network also transmitted one FTP traffic flow of 640 kbps. The DiffServ routers implemented the Weighted Random Early Detection (WRED) mechanism for active queue management. The WRED parameters include a minimum threshold, a maximum threshold, and a maximum drop probability. In our simulations, these parameters were specified respectively as 2, 4, 0.1 for red packets, 4, 6, 0.05 for yellow packets, and 6, 8, 0.025 for green packets. The final assessment of video quality was based on PSNR metrics.

## 5.2 Simulation Results

*The selection of WPP configuration.* At this phase, all three selected pre-marking strategies have been simulated for the AF PHB overload ranging from 1.0 to 1.2. Each test was repeated 20 times and the average values of the received PSNR are presented in Fig. 8.

Analysing the results shown in the Fig. 8, it is difficult to identify a clear winner. From the perspective of the typical IP network behaviour, it seems to be reasonable to concentrate on the area of small and medium-size congestions. With respect to the simulation performed, it is the range of congestion from 1.05 to 1.15. In this area the best quality guarantee methods of protecting data are associated mainly with the SNR and the spatial layers. This is also true for OPs. The protection of SNR layers allows for maintaining a relatively small decline in the video quality for a few first OPs. Similarly, in the case of small and medium-sized congestion, adaptation mechanisms that operate on the OPs should mostly use the first few of them. That was confirmed by our separate research [41], when the standard DASH adaptation mechanism was configured for the same video configurations and network conditions, operating on the first three OPs. We can then state that the OPs from OP1 to OP3 are the most probable structure of the SVC GOP for video streaming applications. For this reason, in the next phase, the WPP configuration SNR-T-S was selected for the comparative analysis.

*Comparative analysis of proposed WPP algorithm.* Based on results from previous tests, the weights in Eq. 1 are assigned values, 1 for the SNR layer, 2 for the spatial layer and 3 for the temporal layer. Next, the priority for any given NALU were calculated according to the Eq. 1. The  $H$  coefficient has value 15 for the video sequences used (the least important triple in the Foreman sequence is (2,2,3) therefore  $H = 1 * 2 + 2 * 2 + 3 * 3$ ). The last step was to apply the rules of marking presented in Sect. 4.

To justify our algorithm, the simulation results for the video transmission of the SVC stream with the WPP pre-marking have been compared to the transmission of the SVC sequence without pre-marking (trTCM was configured in blind mode) and the video coded by the H.264/AVC coder with pre-marking based on simple frame type mapping (I,P,B) [9, 10]. In all the cases the Foreman video sequence was used. The simulation results are shown in Table 1.

In the case of using WPP pre-marking, the video quality improvement is observed for relatively small values of congestion (especially in the range from 1.0 to 1.1). This is due to better protection of spatial and SNR layers. Without pre-marking, the mechanism inside the transmission system cannot protect the low and the lowest layers very well and at the same time losses of higher spatial and SNR layers are relatively high so end-user has little or no benefit from the SVC coding. The transmission of the H.264/AVC video with random losses of P and B macroblocks (pre-marking algorithm preferred I frames) causes numerous errors in the mechanisms of motion vectors reconstruction and inter-frame

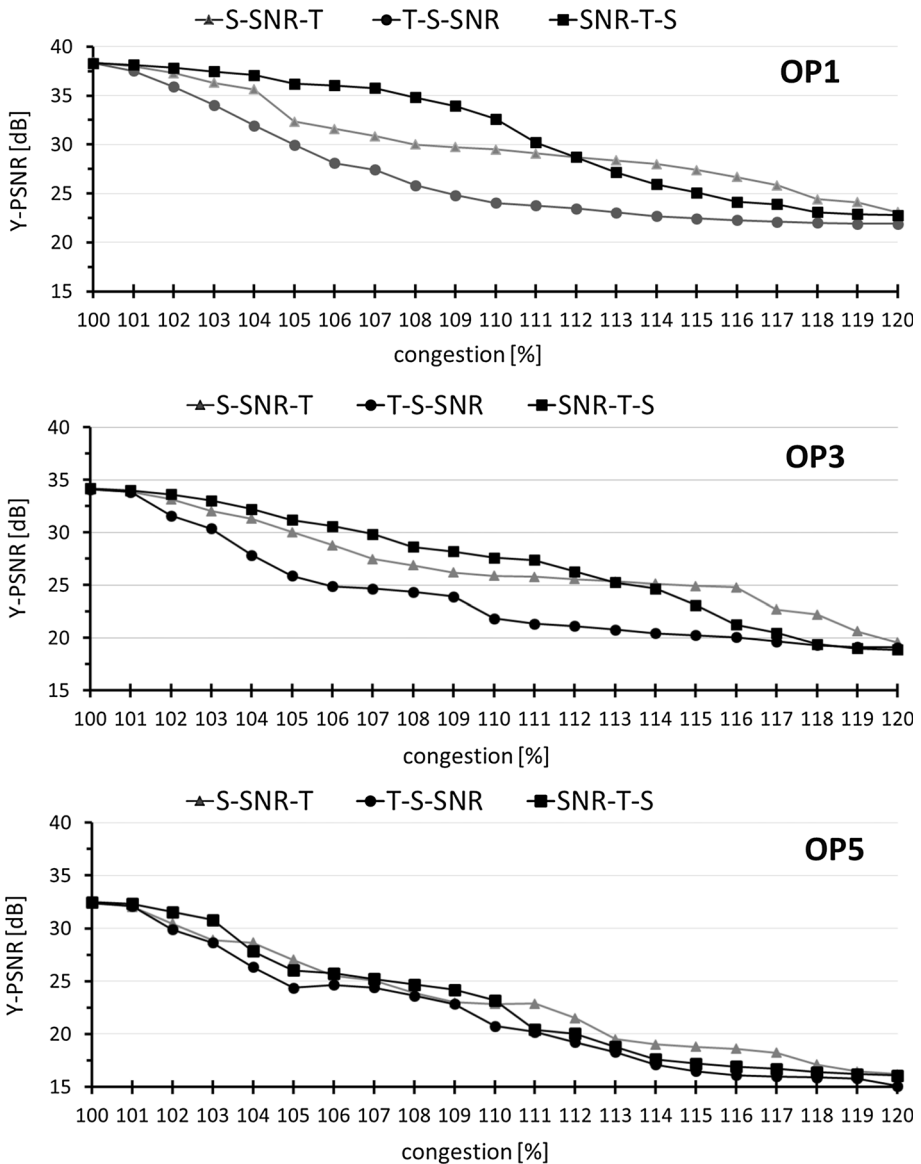


Fig. 8 Y-PSNR for different test scenarios and for different values of AF PHB overload obtained for selected OPs

prediction. These phenomena very quickly (for relatively small values of overload) manifest themselves as important video quality degradation. For the higher values of congestion, the advantage of SVC over AVC coding slowly disappears. The same can be said about the relationship between transmission with and without pre-marking. Even the WPP algorithm cannot prevent loss of a substantial part of video data.

*Practical aspects of the implementation of the algorithm WPP* During all of the tests, the elements of a typical video streaming system were used. This applies to both

**Table 1** Comparison of Y-PSNR in different video streaming scenarios

Scenario	Link overload		
	105 %	110 %	115 %
AVC with frame-type based pre-marking	32.4 dB	27.2 dB	25.5 dB
SVC without pre-marking (OP1)	32.5 dB	29.6 dB	27.1 dB
SVC with WPP pre-marking (OP1)	36.1 dB	31.2 dB	27.8 dB

components responsible for the distribution of video and data transmission. The DiffServ domain was based on the operating system fully compatible with commercial systems used by the routers Cisco 2900 series. The algorithms of the packet marking, queuing and the rules of the dynamic routing are configured as recommended for typical network operators. Also, the video distribution system was based on the reference software JSVM [42]. For this reason, focusing on practical aspects of the implementation of the algorithm WPP, we can conclude that:

- The proposed solution can be implemented in existing video streaming services based on the standard H264/SVC and DASH,
- WPP algorithm works fine with a typical network infrastructure that supports DiffServ mechanisms and the packet marking in accordance with the algorithm trTCM (colour-aware mode).

## 6 Conclusion

In this paper the relationship between the relative importance of NALUs and the packet pre-marking for the H.264/SVC video has been studied. We proposed the Weighted Priority Pre-marking algorithm for colour-aware SVC video streaming over the DiffServ network. This algorithm has been tested for different bit stream ordering and operation point scenarios. Selected scenarios reflect the typical use of SVC coding in today's video streaming applications. In contrast to other proposed solutions, our approach is consistent with the DiffServ model and does not require changing the marking schema at the edge of the DiffServ domain. Thus, the proposed algorithm can be applied to any IP network using the principles of the service differentiation.

By comparing the simulation results with the standard streaming solution based on single layer H.264/AVC and best-effort H.264/SVC transmission, a simple conclusion can be drawn that the proposed pre-marking algorithm can well reflect the relative importance inside the SVC video stream and allows users to take advantage of the scalability extension of H.264/SVC.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., & Weiss, W. (1998). An architecture for differentiated services IETF RFC2475.
2. Zhang, F., Macnicol, J., Pickering, M. R., Frater, M. R., & Arnold, J. F. (2006). Efficient streaming packet video over differentiated services networks. *IEEE Transactions on Multimedia*, 8(5), 1005.
3. Wu, D., Hou, Y., & Zhang, Y. Q. (2000). Transporting real-time video over the Internet: Challenges and approaches. *Proceedings of the IEEE*, 88(2), 1855.
4. Fan Zhai, Luna, C.E., Eisenberg, Y., Pappas, T.N., Berry, R., & Katsaggelos, A.K. (2003). A novel cost-distortion optimization framework for video streaming over differentiated services networks. In *Proceedings of IEEE International Conference on Image Processing* (Vol. 3, pp.III–293).
5. Ziviani, A., Rezende, J. F., Duarte, O. C., & Fdida, S. (2002). Improving the delivery quality of MPEG video streams by using differentiated services. In *2nd European Conference on Universal Multiservice Networks (ECUMN02)* (p. 107).
6. P.10/G.100. (2007). Amendment 1, New Appendix I - Definition of Quality of Experience (QoE), ITU-T.
7. ISO/IEC 23009-1:2014. (2014). Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats, [http://standards.iso.org/ittf/PubliclyAvailableStandards/c065274\\_ISO\\_IEC\\_23009-1\\_2014.zip](http://standards.iso.org/ittf/PubliclyAvailableStandards/c065274_ISO_IEC_23009-1_2014.zip) (2014).
8. Schwarz, H., Marpe, D., & Wiegand, T. (2007). Overview of the scalable video coding extension of the H.264/AVC standard. *IEEE Transactions on Circuits and System Video Technology*, 17(9), 1103.
9. Ziviani, A., Wolfinger, B. E., De Rezende, J. F., Duarte, O., & Fdida, S. (2005). Joint adoption of QoS schemes for MPEG streams. *Multimedia Tools and Applications*, 26(1), 59.
10. Przylucki, S. (2012). *Efficiency of IP Packets pre-marking for H.264 video quality guarantees in streaming applications, computer networks* (Vol. 291). Berlin: Springer.
11. Babiarz, J., Chan, K., & Baker, F. (2006). Configuration guidelines for DiffServ service classes IETF RFC4594.
12. Baker, F., Polk, J., & Dolly, M. (2010). A Differentiated Services Code Point (DSCP) capacity-admitted traffic IETF RFC5865.
13. Heinanen, J., Baker, F., Weiss, W., & Wroclawski, J. (1999). Assured forwarding PHB Group, IETF RFC2597.
14. Grossman, D. (2002). New terminology and clarifications for DiffServ IETF RFC3260.
15. Heinanen, J., & Guerin, R.A. (1999). A single Rate Three Color Marker RFC 2697.
16. Heinanen, J., & Guerin, R.A. (1999). A two rate three color marker RFC 2698.
17. Fang, W. & Seddigh, N. (2000). Time sliding window three color marker RFC 2859.
18. ITU-T ISO/IEC JTC-1. (2005). *Advanced Video Coding for generic audiovisual services*. ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 part 10).
19. Wiegand, T., Sullivan, G. J., Bjontegaard, G., & Luthra, A. (2003). Overview of the H.264/AVC video coding standard. *IEEE Transactions Circuits and Systems for Video Technology*, 13(7), 560.
20. Zoi, S., Loukatos, D., Sarakis, L., Stathopoulos, P., & Mitrou, N. (2003). *Extending an open MPEG-4 video streaming platform to exploit a differentiated services network, high-speed networks and multimedia communications* (Vol. 2720). Berlin: Springer.
21. Tan, W., & Zakhor, A. (2001). Packet classification schemes for streaming MPEG video over delay and loss differentiated networks. In *Proceedings of Packet Video Workshop*. <http://citeseerx.ist.psu.edu/viewdoc/download?sessionid=3E5A22A76186633E09327F159113E8E0?doi=10.1.1.23.7506&rep=rep1&type=pdf>.
22. DeVito, F., Quaglia, D., & DeMartin, J.C. (2004). Model-based distortion estimation for perceptual classification of video packets. In *Proceedings of IEEE International Workshop on Multimedia Signal Processing (MMSP)* (Vol. 1, p. 79).
23. Lee, C C Yu., Ya, Ju, & Chang, P. C. (2010). Adaptive significance classification mechanism for streaming video over differentiated service networks. *Journal of Computers*, 21(2), 3.
24. Chen, L., & Liu, G. (2010). A delivery system for streaming video over DiffServ networks. *IEEE Transactins Circuits and Systems for Video Technology*, 20(9), 1255.
25. Ke, C. H., Shieh, C. K., Hwang, W. S., & Ziviani, A. (2005). A two markers system for improved mpeg video delivery in a DiffServ network. *IEEE Communications Letters*, 9(4), 381.
26. Chen, L., Liu, G., & Zhao, F. (2007). *An improved marking mechanism for real-time video over DiffServ networks, advances in multimedia information processing PCM 2007* (Vol. 4810). Berlin: Springer.



27. Wang, H., Liu, G., Chen, L., & Wang, Q. (2011). A novel marking mechanism for packet video delivery over DiffServ networks. In *IEEE International Conference on Multimedia and Expo (ICME), 2011* (p. 1).
28. Chen, L., & Liu, G. (2008). A delivery system for scalable video streaming using the scalability extension of H.264/AVC over DiffServ networks. In *IHMSP '08 International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (p. 665).
29. Amram, N., Fu, B., Kunzmann, G., Melia, T., Munaretto, D., Randriamasy, S., Sayadi, B., Widmer, J., & Zorzi, M. (2011). QoE-based transport optimization for video delivery over next generation cellular networks. In *IEEE Symposium on Computers and Communications (ISCC)* (p.19).
30. Mushtaq, M. S., Augustin, B., & Mellouk, A. (2012). Empirical study based on machine learning approach to assess the QoS/QoE correlation. In *Networks and Optical Communications (NOC), 17th European Conference* (p. 1).
31. Alreshoodi, M., Woods, J., & Musa, I.K. (2014). QoE-enabled transport optimisation scheme for real-time SVC video delivery. In *Communication Systems, Networks and Digital Signal Processing (CSNDSP), 9th International Symposium on* (p. 865).
32. Sierszen, A. (2014). Reduction of reference set for network data analyzing using the bubble algorithm. *Image Processing and Communications*, 313, 319.
33. Unanue, I., Urteaga, I., Husemann, R., DelSer, J., Roesler, V., Rodriguez, A., & Sanchez, P. (2011). A tutorial on H.264/SVC scalable video coding and its tradeoff between quality, coding efficiency and performance. InTech. <http://cdn.intechopen.com/pdfs-wm/15138>.
34. Daronco, L.C., Roesler, V., & Lima, J.V. (2010). Subjective video quality assessment applied to scalable video coding and transmission instability. In *Proceedings of ACM Symposium on Applied Computing SAC'10* (p. 1898).
35. Ni, P., Eg, R., Eichhorn, A., Griwodz, C., & Halvorsen, P. (2011). Flicker effects in adaptive video streaming to handheld devices. In *Proceedings of 19th ACM international conference on Multimedia MM'11* (p. 463).
36. Mok, R., Luo, X., Chan, E., & Chang, R. (2012). QDASH: a QoE-aware DASH system. In *Proceedings of 3rd annual ACM Conference on Multimedia Systems MMSys'12* (p. 11).
37. Gupta, R., Pulipaka, A., Seeling, P., Karam, L. J., & Reisslein, M. (2012). H.264 coarse grain scalable (CGS) and medium grain scalable (MGS) encoded video: A trace based traffic and quality evaluation. *IEEE Transactions on Broadcasting*, 58(3), 428.
38. Reichel, J., Schwarz, H., Wien, M., & Vieron, J. (2007). Joint scalable video model 9 of ISO/IEC 14496-10:2005/AMD3 scalable video coding. Joint Video Team (JVT), Docs. JVT-X202.
39. Xiao, S., Wang, H., & Jay-Kuo, C. C. (2007). *Priority ordering and packetization for scalable video multicast with network coding, advances in multimedia information processing PCM 2007* (520th ed., Vol. 4810). Berlin: Springer.
40. Xiao, S., Wu, Ch., Li, Y., & Du, J. (2008). Priority ordering algorithm for scalable video coding transmission over heterogeneous network. In *Proceedings of 22nd International Conference on Advanced Information Networking and Applications, AINA 2008* (p. 896).
41. Przylucki, S., & Sierszen, A. (2015). Gradation of video quality in DASH services using the H264/SVC coding. *Image Processing and Communications*, 20(3), 15–24.
42. Joint Scalable Video Model - reference software, <http://www.hhi.fraunhofer.de/en/fields-of-competence/image-processing/research-groups/image-video-coding/svc-extension-of-h264avc/jsvm-reference-software.html?NL=0>.
43. SVEF: Scalable Video-streaming Evaluation Framework, <http://svf.netgroup.uniroma2.it/#download>.



**Sławomir Przylucki** received his M.Sc. and Ph.D. degrees in Electrotechnics from Lublin University of Technology (LUT), Poland, in 1991 and 1999, respectively. He has been employed at LUT since 1991, currently as an Assistant Professor at the Institute of Computer Science. His scientific interests cover IP traffic engineering, wireless networks, sensor networks and video streaming systems. S. Przylucki is an author of more than 50 papers published in communication journals and presented at national and international conferences. He participated in several R&D projects funded by national authority as well as by European Union. Since 2013, he has been served as an expert in the field of IT projects for the government of the Republic of Polish and the Czech Republic.



**Dariusz Czerwinski** currently working as a Professor and Head of Institute of Computer Science at Lublin University of Technology (LUT). Graduated from the Faculty of Electrical Engineering LUT and student of Kanazawa University in Japan. He received his Ph.D. in 2001 and D.Sc. in 2014. Scientific activities include: numerical modelling of electromagnetic devices, the use of cloud computing, data security in network systems, data mining systems using MapReduce paradigm, the use of Artificial Neural Networks, Wireless Sensory Networks. . He is an author of over 100 papers, and author/co-author/editor of 2 books. He is also a member of IEEE Communications Society.