

Transmisión inalámbrica multimedia coordinada con DASH-SAND

Román Belda⁽¹⁾, Ismael de Fez⁽¹⁾, Juan Carlos Guerri⁽¹⁾

robeltor@iteam.upv.es, isdefez@iteam.upv.es, jcguerri@dcom.upv.es

⁽¹⁾ Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTEAM), Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia

Abstract- The adoption of DASH (Dynamic Adaptive Streaming over HTTP) as a video transmission standard has introduced new challenges in environments with high density of Wi-Fi clients. If video players do not take into account the playback status of other devices in the network unfair bandwidth distribution may cause playback problems (stalls, low playback quality, etc.). This paper proposes a collaborative playback technique based on DASH-SAND (MPEG's Server and Network Assisted DASH) that aims to reduce, in a collaborative way, the number of interruptions in the playback of multimedia content in Wi-Fi environments with high density of users. Results prove that the proposed solution reduces both the number and duration of stalls at the expense of decreasing the average representation of the playback.

I. INTRODUCCIÓN

DASH [1] se ha convertido en uno de los principales mecanismos para la transmisión de vídeo sobre redes IP. La adopción de DASH supone un cambio de modelo que traslada toda la lógica de la reproducción al player de vídeo. Adicionalmente, la utilización de HTTP para la transmisión de los segmentos ha eliminado la necesidad de un servicio específico de streaming que, a su vez, facilita la utilización de CDNs (*Content Delivery Networks*) para escalar los servicios. Sin embargo, DASH también introduce nuevos retos, sobre todo en el lado del receptor cuando el último salto es Wi-Fi y además está compartido por un gran número de dispositivos.

En este tipo de escenarios como, por ejemplo, salas de eventos, estadios y medios de transporte (autobuses, aviones, barcos...) es frecuente ofrecer servicios de streaming de vídeo relacionados con la propia actividad o el entretenimiento. En estos entornos, los reproductores de vídeo compiten entre ellos por el acceso al medio, lo cual tiene efectos adversos como la inestabilidad y la desigualdad en la distribución del ancho de banda disponible [2]. A este problema se suma que la utilización de redes Wi-Fi introduce otro factor de variabilidad y de desigualdad en el acceso a los contenidos.

Con el objetivo de mitigar los problemas introducidos por DASH en algunos escenarios surge la tecnología *Server and Network Assisted DASH* (SAND), definida en la ISO/IEC FDIS 23009-5:2016 – Part 5 [3].

En este trabajo se presenta una implementación de DASH-SAND que tiene como objetivo la coordinación del recurso del ancho de banda entre los reproductores de vídeo que se encuentran en una misma red Wi-Fi con alta concentración de usuarios.

El resto del artículo tiene la siguiente estructura: la Sección II presenta el estado del arte relacionado; la Sección III detalla

la propuesta presentada en este artículo; la Sección IV describe la metodología empleada para obtener los resultados mostrados en la Sección V; finalmente las conclusiones se presentan en la Sección VI.

II. ESTADO DEL ARTE

MPEG SAND tiene como objetivo permitir una mejor cooperación del cliente, servidor y red en la transmisión de contenido DASH. La tecnología SAND ofrece mensajes de estandarización y protocolos de intercambio para que proveedores de servicio y operadores de red puedan ofrecer mejoras en el funcionamiento y la utilización de recursos que, de otro modo, no sería posible utilizando únicamente DASH.

Existen en la literatura diversos trabajos relacionados con DASH-SAND aunque, debido a que es un estándar relativamente novedoso, no existen estudios que traten el escenario que se contempla en este trabajo. Cabe destacar, por ejemplo, el trabajo de Khorov et al. [4] que propone añadir un servicio DANE (*DASH-Aware Network Element*) en el propio punto de acceso. Este servicio DANE sería el encargado de modificar la forma en la que el punto de acceso pone el medio de transmisión a disposición de los distintos clientes. La forma en la que este artículo propone distribuir el ancho de banda disponible entre los clientes de vídeo es la utilización del *Transmit Opportunity* (TXOP) del punto del acceso. Si bien este tipo de modificaciones es interesante para ser estudiada, en la práctica existen multitud de dispositivos Wi-Fi que pueden no formar parte de la red en cuestión y es difícil modificar la forma en la que los dispositivos comerciales acceden al medio.

Kleinrouweler et al. [5] evalúan la utilización de una gestión de colas de tráfico para reducir el número de parones en la reproducción de vídeo en redes cableadas con un elevado número de clientes de vídeo (600). A pesar de que tanto el entorno de pruebas como la aproximación al problema es diferente, en dicho estudio se pueden encontrar puntos comunes con el trabajo realizado en la propuesta aquí presentada ya que, en ambos casos, existe una necesidad de que los clientes de vídeo informen a un servicio central de cierta información y de que exista un ancho de banda disponible en todo momento para la transmisión de las comunicaciones de control entre el servicio de gestión y los clientes de vídeo. Este último punto, en el caso que del escenario considerado en este artículo es, si cabe, más crítico debido a la utilización de redes Wi-Fi.

De los mismos autores que el trabajo anterior, en [6] se propone como mecanismo de obtención de la información

relativa a las características del tráfico de los clientes de vídeo la utilización de un dispositivo SDN (*Software Defined Network*). Otra solución basada en SDN es la presentada por Lee et al. en [7], donde se propone la utilización de un punto de acceso con capacidades de SDN que intercepta las peticiones de los clientes de vídeo. Gracias a esta interceptación, el sistema es capaz de estimar el tamaño del búfer de los clientes, así como sus anchos de banda. Esta propuesta tiene varios puntos que dificultan su implementación en un despliegue real. La más importante es que el sistema no puede funcionar utilizando la versión segura de HTTP, es decir, HTTPS. Otros puntos difícilmente implementables serían: la modificación de la MPD (*Media Presentation Description*) de DASH para incluir información relativa al PSNR de los segmentos, y la interceptación y modificación de la representación solicitada por los clientes.

Finalmente, en la propuesta de Cofano et al. [8] se evalúan tres aproximaciones diferentes a la adaptación de streaming asistida por la red: 1) por reserva de ancho de banda; 2) estimación del ancho de banda disponible; y 3) utilización conjunta de los tipos anteriores. La utilización del modo 2 es, según este artículo, el mecanismo de colaboración cliente-red que mejores resultados proporciona teniendo en cuenta el parámetro VQF (*Video Quality Fairness*).

III. PROPUESTA

Con el objetivo de mejorar la experiencia de los usuarios en el acceso al contenido multimedia en entornos Wi-Fi, se propone una solución basada en DASH-SAND que permite la coordinación de los reproductores de vídeo de forma que se produzcan un menor número de paradas en la reproducción. La propuesta consta de dos partes, tal como refleja la Fig. 1:

- Servidor de contenido. Aparte de servir contenido, realiza funciones de DANE, tal como se comenta a continuación.
- Clientes de vídeo. Son reproductores DASH a los que se les ha añadido capacidades de utilizar SAND para comunicar su estado y recibir indicaciones sobre el estado de la red.

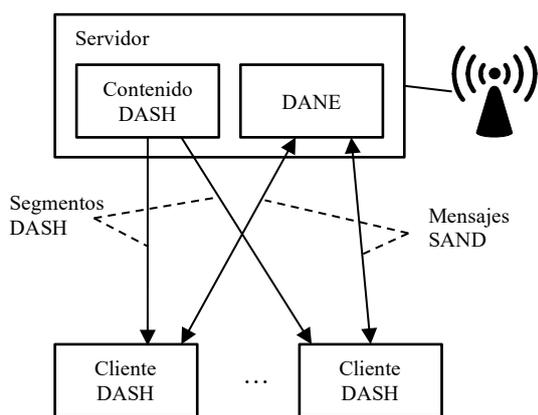


Fig. 1. Solución SAND para la mejora del servicio DASH

En la Fig. 1 se puede observar como, adicionalmente a la descarga de los segmentos de DASH, los clientes de vídeo mantienen una comunicación bidireccional (a través de conexiones *WebSocket*) con el elemento DANE desarrollado. En particular, los clientes de vídeo realizan dos tipos de notificaciones: 1) periódicas, utilizadas a modo de *heartbeat* y para indicar el tamaño del búfer si está en estado de

reproducción; 2) al producirse un cambio en el ancho de banda estimado o en la representación que se está visualizando.

El objetivo del DANE es mantener actualizada una estimación de ancho de banda, la cual se envía de forma proactiva a todos los clientes de vídeo cuando ésta se modifica. Por su parte, los clientes de vídeo, tras la descarga de un segmento, comunican su estimación de ancho de banda al DANE, desencadenando la actualización de la estimación de ancho de banda del propio DANE.

De esta forma, cada cliente de vídeo i , a la hora de tomar la decisión de qué representación descargar en el instante t , actualiza su estimación efectiva (\widehat{bw}'_i) al valor menor entre su propia estimación (\widehat{bw}_i) y la indicada por el DANE (\widehat{bw}_{DANE}), siempre y cuando exista esta última:

$$\widehat{bw}'_i(t) = \begin{cases} \min(\widehat{bw}_i(t), \widehat{bw}_{DANE}(t)) & \text{si } \exists \widehat{bw}_{DANE}(t) \\ \widehat{bw}_i(t) & \text{si } \nexists \widehat{bw}_{DANE}(t) \end{cases} \quad (1)$$

Para llegar a su estimación, el DANE calcula la media aritmética de los anchos de banda estimados por cada uno de los N clientes de vídeo con una reproducción en curso (Ecuación 2) y les aplica una minoración en función de cada uno de los estados en los que se encuentren sus búferes de reproducción según Ecuación 3.

$$bw'_{DANE}(t) = \frac{1}{N} \sum_{i=1}^N \widehat{bw}_i(t), \quad (2)$$

$$bw_{DANE}(t) = \max\left[0, bw'_{DANE}(t) * \left(1 - 3 * \left(\frac{n_r}{N} + \frac{n_y}{2N}\right)\right)\right]. \quad (3)$$

Donde:

- n_y : número de clientes que su nivel de búfer en el instante t está entre 10 y 20 segundos (zona amarilla).
- n_r : número de clientes que su nivel de búfer en el instante t es menor de 10 segundos (zona roja).

Así, los clientes que estén en zona roja provocarán una reducción del ancho de banda el doble que aquellos que están en zona amarilla. Según (3), si un tercio de los clientes está en zona roja, o dos tercios se encuentra en zona amarilla, el valor de bw_{DANE} será 0, por lo que \widehat{bw}'_i será 0 –según (1)–, y el algoritmo de adaptación de DASH actuará en consecuencia (eligiendo la representación de menor *bitrate*).

IV. METODOLOGÍA

Para evaluar la implicación de la introducción de SAND en el escenario propuesto, se ha efectuado una comparativa entre el funcionamiento del servicio de streaming con y sin el mecanismo SAND activado. Para ello, se emplean vídeos codificados a diferentes calidades utilizando el estándar DASH, concretamente tres vídeos de acceso libre creados por la Blender Foundation: “Elephants Dream”, “Tears of Steel” y “Sintel”.

Los vídeos se reproducen de forma pseudo-aleatoria, de manera que cada tableta reproduce el mismo número de veces cada uno de los vídeos. En las pruebas realizadas, cada tableta realiza un total de 9 reproducciones por lo que, teniendo en cuenta que la duración aproximada de cada vídeo es de 10 minutos, la duración de las pruebas ha sido 90 minutos aproximadamente. En total, se han realizado tres pruebas de cada caso, haciendo un promedio para obtener los datos presentados en la sección de evaluación.

Los vídeos han sido codificados a 720p con H.264, utilizando CRFs (*Constant Rate Factor*) entre 5 (mejor calidad) y 55 (menor calidad), en intervalos de 5, esto es, un total de 11 calidades de vídeo.

Para comprobar el funcionamiento de la propuesta en diferentes entornos, se han utilizado dos algoritmos de adaptación: el algoritmo adaptativo que incluye la librería ExoPlayer y el algoritmo Look Ahead [9].

Los parámetros evaluados son los siguientes: 1) número de interrupciones/paradas durante la reproducción de un vídeo; 2) duración de las interrupciones; 3) representación (calidad) media de la reproducción; y 4) número de cambios de calidad, definido como número de veces que se muestra un segmento de una calidad/representación diferente a la del segmento anterior.

El entorno de pruebas está formado por: servidor, punto de acceso Wi-Fi (802.11n 5GHz) y un panel de 36 tabletas (18 tabletas de 7 pulgadas y 18 tabletas de 9 pulgadas, como se observa en la Fig. 2).



Fig. 2. Panel de tabletas de pruebas

El player de vídeo ha sido desarrollado a partir del proyecto ExoPlayer v2. Al iniciar, el player de vídeo busca la presencia del elemento DANE. Si lo encuentra, establece una comunicación con él a través de un WebSocket. Esta conexión será utilizada durante toda la sesión de pruebas para la comunicación de las tabletas con el DANE.

Por otro lado, para facilitar la realización de los tests, se ha desarrollado una aplicación que permite la ejecución de las pruebas en las tabletas de forma desatendida. La aplicación muestra, a través de una interfaz web, información del estado de cada una de las tabletas, por ejemplo, el estado del búfer. En la Fig. 3 se ve una captura de pantalla con la web informativa de la aplicación de pruebas. La información que se muestra es, de izquierda a derecha:

- IP de cada una de las tabletas.
- Tiempo (en segundos) que ha pasado desde que se recibió la última comunicación de una tableta en particular (*heartbeat*).
- Indicación de si la tableta está en estado de reproducción de un contenido.
- Ancho de banda estimado calculado por la propia tableta.
- Tiempo (en segundos) que ha pasado desde que se recibió la última comunicación de ancho de banda estimado de una tableta en particular.
- Tasa de bitrate de la representación en reproducción.
- Longitud del búfer (en segundos) de cada cliente.

Una vez todas las tabletas han iniciado la aplicación de pruebas, se puede enviar un test al conjunto de ellas. Al recibir dicho mensaje de test, las tabletas inician la reproducción, de

manera paulatina, de los contenidos que se encuentran en el mensaje.

IP	Last heartbeat	Playing	Estimated Bandwidth	Estimated Bandwidth Age	Representation's Bitrate	Buffer Length
10.10.0.101	0.4s true		4.7Mbps	8s	0.7Mbps	24s
10.10.0.102	1.0s true		6.0Mbps	7s	2.1Mbps	25s
10.10.0.103	1.2s true		0.0Mbps	22s	0.1Mbps	24s
10.10.0.104	0.7s true		6.2Mbps	7s	2.1Mbps	26s
10.10.0.105	1.1s true		5.0Mbps	7s	1.1Mbps	25s
10.10.0.106	0.7s true		5.9Mbps	7s	1.1Mbps	24s
10.10.0.107	0.5s true		5.2Mbps	7s	1.1Mbps	25s
10.10.0.108	0.3s true		7.9Mbps	6s	2.1Mbps	26s
10.10.0.109	2.0s true		2.1Mbps	0s	20.4Mbps	12s
10.10.0.110	1.3s true		13.1Mbps	14s	1.1Mbps	26s
10.10.0.111	0.5s true		6.4Mbps	7s	1.1Mbps	26s
10.10.0.112	1.0s true		5.9Mbps	7s	2.1Mbps	25s
10.10.0.113	2.7s true		5.6Mbps	8s	0.7Mbps	24s
10.10.0.114	0.8s true		14.4Mbps	4s	2.1Mbps	34s
10.10.0.115	0.9s true		16.1Mbps	2s	2.1Mbps	29s
10.10.0.116	1.3s true		3.2Mbps	5s	2.1Mbps	27s
10.10.0.117	1.5s true		0.0Mbps	9s	0.0Mbps	0s
10.10.0.118	0.3s true		5.2Mbps	2s	1.1Mbps	35s
10.10.0.201	1.4s true		6.2Mbps	7s	2.1Mbps	26s
10.10.0.202	2.7s true		5.2Mbps	8s	0.7Mbps	25s
10.10.0.203	1.8s true		5.2Mbps	57s	2.1Mbps	0s
10.10.0.204	1.6s true		6.1Mbps	7s	2.1Mbps	24s
10.10.0.205	1.1s true		10.6Mbps	4s	2.1Mbps	34s
10.10.0.206	1.4s true		6.6Mbps	6s	1.1Mbps	25s
10.10.0.207	1.9s true		6.1Mbps	7s	2.1Mbps	25s
10.10.0.208	1.9s true		1.7Mbps	6s	0.4Mbps	24s
10.10.0.209	1.6s true		5.9Mbps	7s	1.1Mbps	25s
10.10.0.210	0.5s true		20.7Mbps	3s	1.1Mbps	30s
10.10.0.211	1.5s true		5.8Mbps	10s	1.1Mbps	24s
10.10.0.212	1.0s true		1.5Mbps	19s	0.6Mbps	23s
10.10.0.213	1.7s true		5.9Mbps	7s	1.4Mbps	24s
10.10.0.214	1.4s true		3.6Mbps	12s	2.1Mbps	25s
10.10.0.215	0.8s true		15.1Mbps	4s	2.1Mbps	34s
10.10.0.216	1.5s true		6.9Mbps	8s	2.1Mbps	27s
10.10.0.217	1.1s true		5.9Mbps	14s	30.4Mbps	19s
10.10.0.218	0.1s true		5.9Mbps	6s	1.1Mbps	25s

Fig. 3. Web de información de estado del panel de tabletas

V. RESULTADOS

En este apartado se analiza el funcionamiento del módulo SAND desarrollado, evaluando diferentes parámetros como el número y duración de interrupciones, la calidad media de reproducción y el número de cambios de representación.

En primer lugar, la Fig. 4 muestra el número medio de paradas por contenido reproducido. Se puede observar que, mientras con ExoPlayer sin usar SAND el número medio de interrupciones por cada tableta es superior a 2.5, cuando se usan mecanismos SAND el número de interrupciones es 0.6, lo que representa una reducción del 75%, mientras que en el caso de Look Ahead la reducción es, aunque no tan elevada, bastante significativa (un 35%).

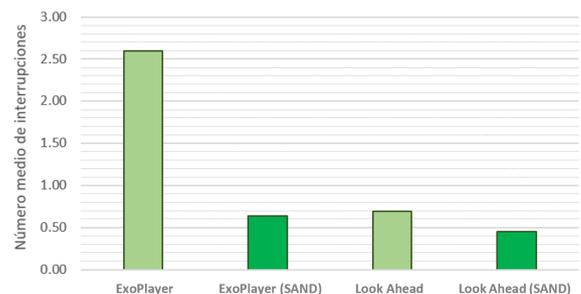


Fig. 4. Evaluación del número de paradas

En relación con la duración media de las paradas, la Fig. 5 muestra la misma tendencia que con el número de paradas. De esta forma, el uso de SAND hace reducir la duración media de las paradas por reproducción de 35.4 a 9.8 segundos (-72%) cuando se usa el algoritmo ExoPlayer, y de 12.3 a 7.7 segundos (-37%) cuando se emplea Look Ahead.

Para reducir el número medio de interrupciones, el mecanismo SAND implementado se basa en que el servidor insta a las tabletas a que reduzcan su calidad (representación) media de reproducción, para así reducir el tráfico en la red y dar más opciones de acceso al medio a las tabletas que tienen problemas en la recepción de contenido. Por tanto, la calidad media de reproducción de los vídeos se reduce, tal como se refleja en la Fig. 6. La figura muestra en el eje Y la

representación media, que en este caso va de 0 a 10, ya que para las pruebas se han utilizado un total de 11 representaciones. Como se puede observar, aunque la calidad media se reduce, la reducción no es muy elevada (un 20% en ExoPlayer y un 21% en Look Ahead), sobre todo teniendo en cuenta que la reducción en términos de duración de las interrupciones ha sido mucho mayor (un 72% en ExoPlayer y un 37% en Look Ahead).

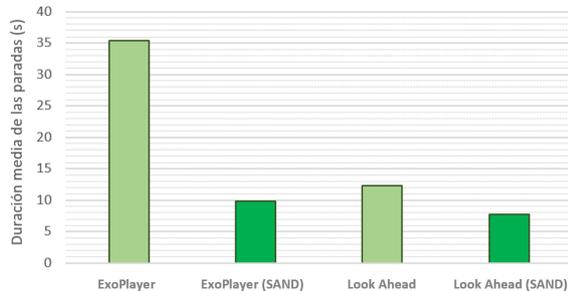


Fig. 5. Evaluación de la duración de las paradas

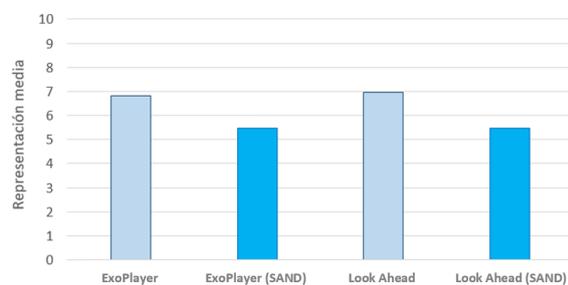


Fig. 6. Evaluación de la representación media de la reproducción

Finalmente, la Fig. 7 muestra el número medio de cambios de representación por contenido reproducido. En este caso, la utilización de la propuesta basada en SAND produce un mayor número de cambios de calidad, lo cual puede empeorar la Calidad de Experiencia de usuario si dichos cambios son muy significativos (por ejemplo, cambiar de una representación alta a una muy baja). Como vemos, el número de cambios de representación afecta más a ExoPlayer (25 cambios de calidad más utilizando SAND) que a Look Ahead (18 cambios).

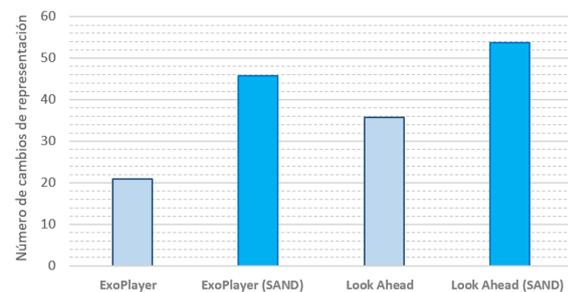


Fig. 7. Evaluación del número medio de cambios de representación

VI. CONCLUSIONES

La utilización de DASH supone un nuevo reto en escenarios con redes Wi-Fi y una alta densidad de usuarios de streaming de vídeo. Es por esto por lo que se está trabajando en la estandarización de mecanismos que permitan coordinar la reproducción del contenido, como la tecnología SAND.

En este artículo se presenta un mecanismo de coordinación de clientes de vídeo, basado en SAND, que utiliza la información de estado del conjunto de los dispositivos para ofrecer una alternativa de estimación de ancho a la calculada

por los propios reproductores de vídeo. Con esta estimación, mientras unos clientes reducen su tasa de reproducción, el resto tiene más opciones de acceder al medio, mejorando así la Calidad de Experiencia (QoE) media de los usuarios.

La solución propuesta, a diferencia de las encontradas en la bibliografía, no requiere de modificación alguna en los algoritmos de acceso al medio ni de un control, ni por parte del servidor ni de los clientes, del ancho de banda percibido por los clientes de vídeo.

La evaluación llevada a cabo muestra como, en presencia del elemento DANE propuesto, se reduce de forma significativa el número de paradas en cada reproducción, así como la duración de estas. Al mismo tiempo, la representación media de las reproducciones desciende en torno a un 20%. En cuanto al número de cambios, éstos también sufren un incremento debido a las indicaciones del DANE.

A pesar de la menor representación media y de un mayor número de cambios de calidad, a falta de un análisis más exhaustivo de la QoE, se puede considerar que la solución propuesta supone una mejora en la calidad percibida debido a la gran importancia que las interrupciones en la visualización de un contenido tienen sobre la QoE [10].

AGRADECIMIENTOS

Este trabajo ha sido subvencionado por el Programa PAID-10-18 así como por el Proyecto Línea de I+D+i “Tecnologías de distribución y procesado de información multimedia y QoE” de la Universitat Politècnica de València.

REFERENCIAS

- [1] ISO/IEC 23009-1:2014, Dynamic adaptive streaming over HTTP (DASH) - Part 1: media presentation description and segment formats, 2014.
- [2] S. Akhshabi, L. Anantakrishnan, A. C. Begen, and C. Dovrolis, “What happens when HTTP adaptive streaming players compete for bandwidth?”, in Proc. of the Int. Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV), Toronto, Canada, Jun. 2012, pp. 9-14.
- [3] International Organization for Standardization, “Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 5: Server and network assisted DASH (SAND)”, ISO/IEC 203009-5:2017, disponible en línea: <https://www.iso.org/standard/69079.html>, Feb. 2017.
- [4] E. Khorov, A. Krasilov, M. Liubogoshchev, and S. Tang, “SEBRA: SAND-enabled bitrate and resource allocation algorithm for network-assisted video streaming,” in International Conference on Wireless and Mobile Computing, Networking and Communications, 2017.
- [5] J. W. Kleinrouweler, B. Meixner, and P. Cesar, “Improving Video Quality in Crowded Networks Using a DANE,” in Proc. of the 27th Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV), Taipei, Taiwan, 2017, pp. 73-78.
- [6] J. W. Kleinrouweler, S. Cabrero, and P. Cesar, “Delivering stable high-quality video: an SDN architecture with DASH assisting network elements,” in Proc. of the 7th International Conference on Multimedia Systems (MMSys), Klagenfurt, Austria, May 2016.
- [7] H. Lee, Y. Go, and H. Song, “SDN-assisted HTTP adaptive streaming over Wi-Fi network,” in Proc. of the International Conference on Software Defined Systems (SDS), Barcelona, Spain, Apr. 2018.
- [8] G. Cofano, L. De Cicco, T. Zinner, A. Nguyen-Ngoc, P. Tran-Gia, and S. Mascolo, “Design and experimental evaluation of network-assisted strategies for HTTP adaptive streaming,” in Proc. of the 7th International Conference on Multimedia Systems, Klagenfurt, Austria, May 2016, article no. 3.
- [9] R. Belda, I. de Fez, P. Arce, and J. C. Guerri, “Look Ahead: a DASH adaptation algorithm,” in Proc. of the IEEE Int. Symp. On Broadband Multimedia Systems and Broadcasting (BMSB), Valencia (Spain), Jun. 2018, article no. 158.
- [10] I. de Fez, R. Belda, and J. C. Guerri, “New objective QoE models for evaluating ABR algorithms in DASH,” Computer Communications, vol. 158, pp.126-140, 2020.