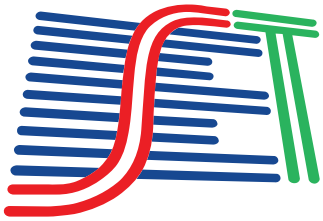
# SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING

SET IJBE V. 6, 2020

ISSN print: 2446-9246
ISSN online: 2446-9432



# SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING

SET IJBE V. 6, 2020

ISSN print: 2446-9246
ISSN online: 2446-9432

# SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING

ISSN PRINT: 2446-9246 | ISSN ONLINE: 2446-9432

INDEXED IN:

Google's system that offers tools for search of academic literature.

International Cataloging Data in the Publication - CIP - Librarian Zoraide Gasparini CRB/9 1529

S517 SET International Journal of Broadcast Engineering —— vol. 6, (Dec. 2020). – São Paulo: Brazilian Society of Television Engineering —— SET, 2020

Annual frequency
ISSN Print 2446- 9246
ISSN online 2446- 9432
Available at: http://www.set.org.br/ijbe

Broadcasting – Periodic. 2. Transmission Engineering. 3. Digital TV. I. SET. II. Title.

CDD: 384.54

THE CONCEPTS SUBMITTED IN THE MANUSCRIPTS ARE THE SOLE RESPONSIBILITY OF THE AUTHOR (S), NOT NECESSARILY REFLECTING THE MAGAZINE'S OPINION

THE SPELLING AND GRAMMAR REVISION OF THE WORKS IS THE RESPONSIBILITY OF THE AUTHOR(S).

ANNUAL | CIRCULATION : 500 EXEMPLARY | GRAPHIC DESIGN AND DIAGRAM: SOLANGE LORENZO

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

# **EDITORIAL** BOARD


**EDITOR IN CHIEF**

**Yuzo Iano**
State University of Campinas – Brazil

**STEERING COMMITTEE**

**Carla Liberal Pagliari**
Military Engineering Institute – Brazil

**Cristiano Akamine**
Mackenzie Presbyterian University – Brazil

**José Frederico Rehme**
Positivo University – Brazil

**Marcelo Ferreira Moreno**
Federal University of Juiz de Fora – Brazil

**ASSOCIATE EDITORS**

**Alexandre de Almeida Prado Pohl**
Federal Technological University of Paraná – Brazil

**Edgard Luciano Oliveira da Silva**
State University of Amazonas – Brazil

**Gustavo de Melo Valeira**
Mackenzie Presbyterian University – Brazil

**Henry Glock**
Federal University of Semiarido – Brazil

**Luís Geraldo Pedroso Meloni**
State University of Campinas – Brazil

**Rangel Arthur**
State University of Campinas – Brazil

**Thiago Genez**
University of Bern – Switzerland

**Valdecir Becker**
Federal University of Paraíba – Brazil



**CORPORATE AUTHOR AND EDITOR**

SET – Brazilian Society of Television Engineering, or, in Portuguese, SET – Sociedade Brasileira de Engenharia de Televisão.
Address: Av. Auro Soares de Moura Andrade, 252, suites 31 and 32 - Barra Funda District - São Paulo - SP - Brazil - Postal Code: 01156-001

# **SET BOARD OF** DIRECTORS

## Deliberative Council

2019 - 2020

**President:** Carlos Fini

**Vice-President**: Claudio Eduardo Younis

| **OFFICE HOLDER** | **SUBSTITUTE** |
|---|---|
| Carlos Fini | Ivan Miranda |
| Luiz Bellarmino Polak Padilha | David Estevam de Britto |
| Claudio Eduardo Younis | José Carlos Aronchi de Souza |
| Claudio Alberto Borgo | Luis Otavio Marchezetti |
| Mauro Alves Garcia | Almir Antonio Rosa |
| Daniela Helena Machado e Souza | José Salustiano Fagundes de Souza |
| Luis Alberto Menoni Popienia | Vinicius Augusto da Silva Vasconcellos |
| Raymundo Costa Pinto Barros | Marcelo Santos Wance de Souza |
| Roberto Dias Lima Franco | Valderez de Almeida Donzelli |
| José Antonio de Souza Garcia | Paulo Henrique Corona Viveiros de Castro |
| Sergio Eduardo di Santoro Bruzetti | José Chaves Felippe de Oliveira |
| José Eduardo Marti Cappia | Marco Tulio Nascimento |
| José Raimundo Lima da Cunha | Esdras Miranda de Araujo |
| Marcio Rogério Herman | Israel de Moraes Guratti |
| Cristiano Akamine | Carla Liberal Pagliari |

## Fiscal Council

| | |
|---|---|
| Cintia Leite Nascimento | Rafael Silviera Leal |
| João Braz Borges | Sandra Regina Rogenfisch |
| Alexandre Aparecido Barrocal | Eduardo de Oliveira S. Bicudo |

## Council of Former Presidents

| | |
|---|---|
| Adilson Pontes Malta | Liliana Nakonechnyj |
| Carlos Eduardo Oliveira Capelão | Olímpio José Franco |
| Fernando Mattoso Bittencourt Filho | Roberto Dias Lima Franco |
| José Munhoz | |

# ABOUT THE JOURNAL

## SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING

The SET IJBE, (SET International Journal of Broadcast Engineering) is an open access, peer-reviewed article--at-a-time international scientific journal whose objective is to cover knowledge about communications engineering in the field of broadcasting. The SET IJBE seeks the latest and most compelling research articles and state-of-the-art technologies.

### Publishing schedule and schema

**On-line version** – Once an article is accepted and its final version approved by the Editorial Board, it will be published immediately on-line on a one article-at-a-time basis

**Printed version** – Once a year, all articles accepted and published on-line over the previous twelve months will be compiled for publication in a printed version.

### Types of papers

Regular (Full) Papers: Traditional and original research [from 6 to 20 pages]

Tutorial Papers: Brand new OTT (Over-The-Top) detailed implementation and fully set up state-of-the art systems [4 – 6 pages]

Letters: Short notes and consideration about current and relevant techniques, technologies and implementations involving engineering solutions [1-3 pages]

### Open Access Policy

If an article is accepted for publication, it will be made available to be read and re-used under a Creative Commons Attribution (CC-BY) license.

### Editorial Office:

If you require any additional information, please contact the SET IBJE (SET International Journal of Broadcast Engineering) administration staff:

Address: Av. Auro Soares de Moura Andrade, 252, suites 31/32, Postal Code:01156-001, São Paulo – SP – Brazil; E-mail: IJBE@set.org.br

## Aims and Scope include, but are not limited to:

Advanced audio technology and processing
Advanced display technologies
Advanced RF Modulation Technologies
Advanced technologies and systems for emerging broadcasting applications
Applying IT Networks in Broadcast Facilities
Broadcast spectrum issues – re-packing, sharing
Cable & Satellite interconnection with terrestrial broadcasters
Cellular broadcast technologies
Communication, Networking & Broadcasting
Content Delivery Networks – CDN
Digital radio and television systems: Terrestrial, Cable, Satellite, Internet, Wireless.
Electromagnetic compatibility issues between collocated services (e.g. broadcast and LTE)
General Topics for Engineers (Math, Science & Engineering)
Hybrid receiver technology
Interactive Technology for broadcast
IP Networks management and configuration
Metadata systems and management
Mobile DTV systems (all aspects, both transmission and reception)
Mobile/dashboard technology
Next-gen broadcast platforms and standards development
Non-real time (NRT) broadcast services
Ratings technology, second screen technology and services
Secondary service system design; mitigation of interference in primary services
Securing Broadcast IT Networks
Signal Processing & Analysis
Software Defined Radio – SDR Technologies
Streaming delivery of broadcast content
Transmission, propagation, reception, re-distribution of broadcast signals AM, FM, and TV transmitter and antenna systems
Transport stream issues – ancillary services
Unlicensed device operation in TV white spaces

We wish to inform you that the activities, events and publications of the Brazilian Society of Television Engineering – SET, including this one, enjoy international support, under formal agreements, from the following international organizations. We also take this opportunity to thank them and reiterate how proud we are that they support our work.

## SUMMARY

SET IJBE v. 6, 2020, 74 pages, 6 articles.

# EDITORIAL

Dear reader,

In the IJBE 2020 edition, we present several papers dealing with the thematic of the next digital terrestrial television broadcast and convergence to broadband, including its technological innovations in the physical, transport, and application coding layer.

The pursuit of quality in the development of technologies that make digital communication systems more efficient is the focus of researchers to whom the IJBE provides an opportunity to publish their studies, experiments, and research in the scientific and technological areas of production and distribution of information content.

We hope you enjoy these papers and feel motivated to submit a paper to us.

Best wishes,
**SET IJBE Editors**

# Fixed Reception Performance of FDM-based Transmission System for Advanced ISDB-T

Takuya Shitomi, Shingo Asakura, Shogo Kawashima, Akihiko Sato, Hiroaki Miyasaka, Noriyuki Shirai, Yoshikazu Narikiyo, Tomoaki Takeuchi, Kohei Kambara, Madoka Nakamura, Tsuyoshi Nakatogawa, Kenichi Murayama, Masahiro Okano and Kenichi Tsuchida

**CITE THIS ARTICLE**

Shitomi, Takuya; Asakura, Shingo; Kawashima, Shogo; Sato, Akihiko; Miyasaka, Hiroaki; Shirai, Noriyuki ; Narikiyo, Yoshikazu; Takeuchi,Tomoaki; Kambara, Kohei; Nakamura,Madoka; Nakatogawa, Tsuyoshi; Murayama,Kenichi; Okano, Masahiro and Tsuchida, Kenichi ; 2020. Fixed Reception Performance of FDM-based Transmission System for Advanced ISDB-T. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2020.1. Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

**COPYRIGHT** This work is made available under the Creative Commons - 4.0 International License. Reproduction in whole or in part is permitted provided the source is acknowledged.

# Fixed Reception Performance of FDM-based Transmission System for Advanced ISDB-T

Takuya Shitomi, Shingo Asakura, Shogo Kawashima, Akihiko Sato, Hiroaki Miyasaka, Noriyuki Shirai, Yoshikazu Narikiyo, Tomoaki Takeuchi, Kohei Kambara, Madoka Nakamura, Tsuyoshi Nakatogawa, Kenichi Murayama, Masahiro Okano, and Kenichi Tsuchida

***Abstract*—With the aim of improving the quality and expanding the functions of digital terrestrial television broadcasting services, we have been developing an advanced transmission system that inherits key features of the current Integrated Services Digital Broadcasting-Terrestrial (ISDB-T) system, which employs hierarchical transmission based on frequency division multiplexing (FDM) and a segment structure. The advanced ISDB-T system has a new signal frame structure that enables bandwidth to be flexibly allocated to multiple services for different reception scenarios, such as fixed reception and mobile reception, compared with ISDB-T. By introducing transmission technologies such as the latest forward error correction and modulation scheme, this specification has high spectral efficiency and transmission robustness, i.e., the transmission capacity increases by about 10 Mbps for the same required carrier to noise ratio (CNR) in comparison with the current ISDB-T system, or the required CNR can be reduced by about 7 dB for the same transmission capacity. We describe the channel coding scheme and evaluated the performance of bit-interleaved coded modulation (BICM) in simulations. This paper provides a BICM selection guideline based on the simulation results for fixed reception scenarios toward the practical application of advanced ISDB-T.**

***Index Terms*—digital terrestrial television broadcasting, advanced ISDB-T, bit-interleaved coded modulation**

## I. Introduction

THE current transmission scheme for digital terrestrial television broadcasting (DTTB) in Japan was first reported by the Telecommunications Technology Council in May 1999, and it was established as an Association of Radio Industries and Businesses (ARIB) standard in May 2001. In the spirit of terrestrial integrated digital broadcasting, this scheme is called "Integrated Services Digital Broadcasting-Terrestrial" (ISDB-T) [1].

A feature of the ISDB-T scheme, which handles different receiving modes (such as fixed reception and mobile reception), is that it can provide multiple services with different transmission capacities and levels of transmission robustness within one channel (6 MHz) used for terrestrial television broadcasting. Currently, many broadcasting stations simultaneously provide high-definition broadcasts for fixed reception and one-segment broadcasts for mobile reception in Japan. By utilizing orthogonal frequency division multiplexing (OFDM) [2] as the modulation scheme with time/frequency interleaving, we can make ISDB-T robust against (i) multipath echoes reflected off buildings, mountains, etc. in the propagation path and (ii) fluctuation in the electric field strength in the case of mobile reception. Furthermore, ISDB-T is compatible with the Moving Picture Experts Group 2 (MPEG-2) systems adopted for multiplexing video, audio, data, etc., constituting broadcasting satellite (BS) and communication satellite (CS) digital broadcast programs, so it has high interoperability.

At NHK, aiming to improve the functionality of digital terrestrial television broadcasting and attain higher-quality video and audio, we have been developing a frequency division multiplexing (FDM)-based transmission system inheriting the above-described features of the ISDB-T scheme while considering the specification of the next-generation terrestrial-broadcasting transmission scheme (hereafter referred to as "advanced ISDB-T") [3][4]. When we formulated the specification, we stipulated a service requirement that Super Hi-Vision broadcasting for fixed reception and high-definition broadcasting for mobile reception be simultaneously provided on one channel because the radio frequency spectrum is tight in Japan. To satisfy this requirement mentioned above, we introduced a new signal structure and the latest technologies for improving transmission characteristics. Additionally, we kept in mind that the advanced ISDB-T scheme must achieve the increased spectral efficiency, i.e., the transmission capacity per unit frequency, and excellent transmission robustness. Supposing that the MPEG Media Transport (MMT) and type-length-value (TLV) adopted for the multiplexing schemes for 4K/8K satellite broadcasting [5], which started in December 2018, are used, we investigated the physical layer scheme, such as the configuration of the forward error correction (FEC) block and the OFDM frame structure.

Advanced ISDB-T adopts bit-interleaved coded modulation (BICM) [6] and has a wide-ranging transmission capacity and robustness using six types of carrier modulation and thirteen types of low-density parity check (LDPC) code

Manuscript received October 12, 2020 and revised December 16, 2020. This research was partly funded by the Ministry of Internal Affairs and Communications, Japan as part of its program titled "Research and Development for Advanced Digital Terrestrial Television Broadcasting System."

T. Shitomi (e-mail: shitomi.t-gy@nhk.or.jp) is with the Science and Technologies Research Laboratories, NHK (Japan Broadcasting Corporation), Tokyo, Japan and with the iTEAM Research Institute, Universitat Politecnica de Valencia, 46022 Valencia, Spain.

S. Asakura, S. Kawashima, A. Sato, H. Miyasaka, N. Shirai, Y. Narikiyo, T. Takeuchi, K. Kambara, M. Nakamura, T. Nakatogawa, K. Murayama, M. Okano, and K. Tsuchida are with NHK (Japan Broadcasting Corporation), Tokyo, Japan.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

rates. To formulate a link budget for advanced ISDB-T, it is necessary to develop BICM that achieves the desired transmission capacity and robustness. In this paper, we report the results of a computer simulation on the transmission performance of all combinations (78 options) of BICM in the additive white Gaussian noise (AWGN) and multipath environments assumed for the fixed reception of DTTB. The evaluation and selection of BICM for practical use have not been researched thus far for advanced ISDB-T, and a method of selecting BICM toward the practical application of advanced ISDB-T is provided.

This paper is structured as follows. Section II reviews the basic configuration of FDM-based advanced ISDB-T. Section III describes the channel coding scheme in the physical layer specification. The methodologies and the simulation setup are presented in Section IV. Section V presents the simulation results and a discussion on the fixed reception performance. The paper is concluded in Section VI.

## II. Overview of FDM-based advanced ISDB-T

### A. Inherited Features of ISDB-T

The current ISDB-T utilizes OFDM, which has excellent transmission performance against multipath echoes, as a modulation scheme and adopts a segment structure in which a signal block is formed by dividing the band of an OFDM signal into signal blocks, and the block, called an "OFDM segment," is used as a unit for data transmission. This segment structure enables OFDM signals to be partially received, i.e., only part of the spectrum can be demodulated and decodable with a narrowband receiver. By concatenating multiple OFDM segments, a transmission signal bandwidth suitable for the target service can be flexibly configured on a segment basis, and television services for fixed reception and mobile reception can be transmitted simultaneously on one channel.

In the case of ISDB-T, one OFDM segment occupies 1/14 of the channel bandwidth, i.e., 6 MHz, and the transmission signal of ISDB-T is composed of 13 OFDM segments. In addition, it is possible to perform hierarchical transmission that simultaneously transmits up to three layers (layers A, B, and C) with different transmission parameters, such as the carrier modulation, code rate of the FEC, and time-interleaving length. Among the 13 segments, the segment (with a bandwidth of 0.43 MHz, i.e., 1/14×6 MHz) in the center is set as a partial-reception band.

In regard to advanced ISDB-T, functions such as hierarchical transmission and partial reception are inherited. That is, the segment structure is adopted so that hierarchical transmission up to three layers with different transmission capacities and levels of transmission robustness is possible allowing partial reception. Hierarchical transmission and partial reception in advanced ISDB-T are illustrated in Fig. 1. For advanced ISDB-T, the number of divisions of the channel bandwidth is increased from 14 to 36, and the transmission signal is composed of up to 35 OFDM segments. A narrowband receiver that carries out partial reception is supposed to receive 9 OFDM segments (with a bandwidth of 1.50 MHz) in the center from among 35 OFDM segments, and the number of segments in layer A can be set in the range

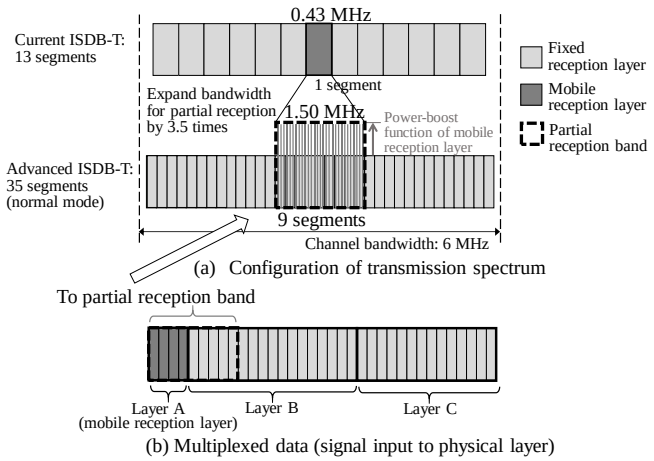



Fig. 1. Hierarchical transmission and partial reception in advanced ISDB-T (in case of 4 segments in layer A).

from 1 to 9 to increase the flexibility of the transmission capacity for partial reception. Moreover, by increasing the partial-reception bandwidth to 3.5 times that of ISDB-T, the effect of frequency interleaving for partial reception is enhanced. The example configuration in Fig. 1 shows hierarchical transmission in which four segments are allocated to a service for mobile reception and transmitted by layer A, and the other segments are allocated to a service for fixed reception and transmitted by layers B and C with different transmission parameters. Figure 1 (b) illustrates multiplexed data that is input to the physical layer and configured to the transmission spectrum as shown in Figure 1 (a). Here, layer A accounts for the transmission capacity of four out of the nine segments of the partial-reception band, and the carriers of layer A are dispersed throughout the partial-reception band in the physical layer processing. The carriers of layer A have a power boost function that sets the transmission power higher than that of the other carriers, thereby improving the robustness of layer A.

In addition to this function, a low-delay transmission path called "LLch" (low-latency channel) is provided in such a way that critical information, such as emergency earthquake warnings, can be transmitted with short delay.

### B. Enhanced Signal Structure

Comparing advanced ISDB-T (which uses 35 of 36 segments of the 6-MHz channel) with ISDB-T (which uses 13 of 14 segments) reveals that, together with improving the flexibility in allocating bandwidth for fixed or mobile reception, the increase in the number of divisions of the channel bandwidth reduces the guard band, resulting in an increase in capacity by about 5%.

In the case of advanced ISDB-T, spectral efficiency is improved by reducing the ratio of the guard interval (GI), which does not contribute to information transmission. Although the current ISDB-T can be operated with a fast Fourier transform (FFT) size of 8,192 ($2^{13}$) points, advanced ISDB-T adopts an FFT size of up to 32,768 ($2^{15}$) points. Increasing the FFT size reduces the carrier frequency spacing. Also, as the effective symbol length increases, even when the GI duration is set to be the same, the ratio of the GI to the effective symbol length (hereafter referred to as "GI ratio") becomes small, and the transmission capacity can be

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

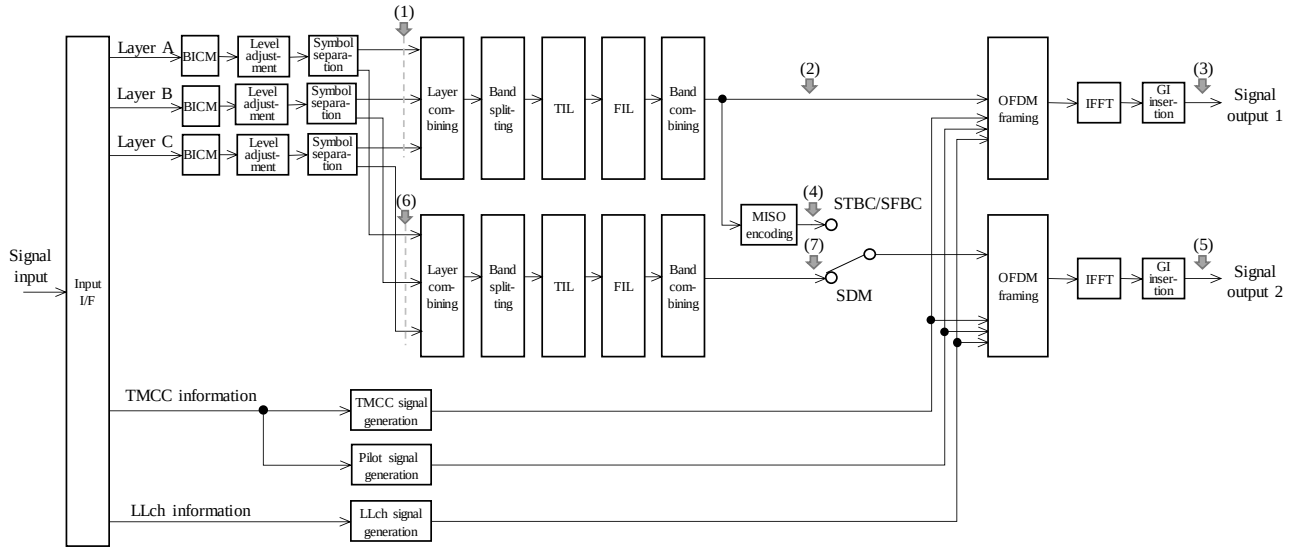



Fig. 2. Block diagram of channel coding in advanced ISDB-T.

increased. However, it should be noted that increasing the FFT size requires higher computational complexity. Compared with ISDB-T, the maximum FFT size is quadrupled from 8,192 to 32,768 in advanced ISDB-T, so the complexity is increased. In addition, considering the compatibility of the advanced ISDB-T with the current ISDB-T, parameters whose bandwidth and GI duration match those of ISDB-T are included in the advanced ISDB-T.

Moreover, in the case of ISDB-T, the pilot signal pattern used for channel estimation is the same in all layers; in contrast, in the case of advanced ISDB-T, optimal patterns can be separately selected for fixed reception and for mobile reception. Additionally, on the premise that the broadcasting coverages are equivalent in all layers, the arrangement interval in the frequency direction (Dx) is set to the same value for all layers, and that in the time direction (Dy) can be set to a different value for each layer.

### *C. Improved Transmission Performance and Capacity*

In the case of ISDB-T, a concatenated code composed of a convolutional code and a Reed-Solomon (RS) code is adopted as an FEC. By correcting errors that cannot be corrected by the convolutional code with the RS code, robustness to noise, interference, etc., is improved. In the case of advanced ISDB-T, to further reduce the required carrier to noise ratio (CNR), an LDPC code that has a higher error-correction capability than a convolutional code is adopted, and transmission performance is improved by using it concatenated with a Bose-Chaudhuri-Hocquenghem (BCH) code.

Regarding the carrier modulation scheme in the case of ISDB-T, three options are available: QPSK (quadrature phase shift keying), 16QAM (quadrature amplitude modulation), and 64QAM, and the signal points are arranged in a uniform constellation (UC). In the case of advanced ISDB-T, by increasing the modulation order ranging from QPSK to 4096QAM, a higher transmission capacity is achieved. In addition to UC, robustness against noise is improved by introducing a non-uniform constellation (NUC), especially in high order modulation, which makes the arrangement of signal points non-uniform [7].

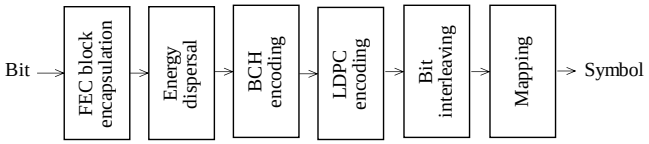



Fig. 3. Configuration of BICM block.

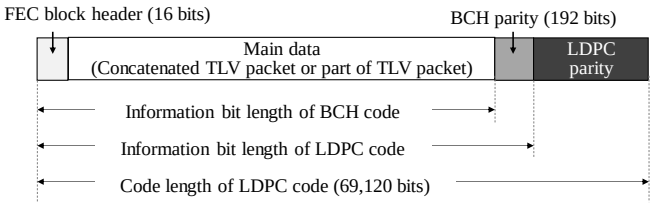



Fig. 4. Configuration of FEC block.

ISDB-T is a single-input single-output (SISO) transmission scheme in which one transmitting antenna and one receiving antenna are configured. It is assumed that a signal stream is transmitted by either a horizontally polarized wave or a vertically polarized wave. In contrast, in the case of the advanced ISDB-T, it is possible to improve transmission robustness by using the 2×1 multiple-input single-output (MISO) scheme—using two transmitting antennas and one receiving antenna—in addition to the SISO scheme and to increase the transmission capacity and/or transmission robustness by using the 2×2 multiple-input multiple-output (MIMO) scheme—using two transmitting antennas and two receiving antennas.

## III. CHANNEL CODING

### *A. Basic Configuration of Channel Coding*

The basic configuration of the channel coding of advanced ISDB-T is shown in Fig. 2. Three hierarchical-frame signals (on layers A, B, and C), transmission and multiplexing configuration control (TMCC) information, and LLch information are output from the input interface (Input I/F). The former signals are input into BICM blocks and are

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

TABLE I
BIT LENGTH OF FEC BLOCK FOR EACH CODE RATE

| LDPC code | | | | BCH code | | FEC block header | Main data |
|---|---|---|---|---|---|---|---|
| Code rate | Code length | Parity bit length | Information bit length | Parity bit length | Information bit length | | |
| 2/16 | 69,120 | 60,480 | 8,640 | 192 | 8,448 | 16 | 8,432 |
| 3/16 | 69,120 | 56,160 | 12,960 | 192 | 12,768 | 16 | 12,752 |
| 4/16 | 69,120 | 51,840 | 17,280 | 192 | 17,088 | 16 | 17,072 |
| 5/16 | 69,120 | 47,520 | 21,600 | 192 | 21,408 | 16 | 21,392 |
| 6/16 | 69,120 | 43,200 | 25,920 | 192 | 25,728 | 16 | 25,712 |
| 7/16 | 69,120 | 38,880 | 30,240 | 192 | 30,048 | 16 | 30,032 |
| 8/16 | 69,120 | 34,560 | 34,560 | 192 | 34,368 | 16 | 34,352 |
| 9/16 | 69,120 | 30,240 | 38,880 | 192 | 38,688 | 16 | 38,672 |
| 10/16 | 69,120 | 25,920 | 43,200 | 192 | 43,008 | 16 | 42,992 |
| 11/16 | 69,120 | 21,600 | 47,520 | 192 | 47,328 | 16 | 47,312 |
| 12/16 | 69,120 | 17,280 | 51,840 | 192 | 51,648 | 16 | 51,632 |
| 13/16 | 69,120 | 12,960 | 56,160 | 192 | 55,968 | 16 | 55,952 |
| 14/16 | 69,120 | 8,640 | 60,480 | 192 | 60,288 | 16 | 60,272 |

subjected to error-correction coding and mapping to carrier symbols. Furthermore, the power for each layer is adjusted, e.g., the power of layer A can be boosted, at the level-adjustment blocks. After that, the layer combining blocks combine the carrier symbols for hierarchical layers A, B, and C. Then, band splitting, time interleaving (TIL), frequency interleaving (FIL), and band combining are performed to form data segments.

To achieve low-latency transmission, processing that differs from the hierarchical frame of the three layers is applied to LLch information, and differential binary phase-shift keying (DBPSK) modulation is performed after the differential reference bit is added to the beginning of the LLch. Some predetermined carriers are assigned to LLch for each OFDM segment, and these carriers can be used for various purposes according to service requirements. At the OFDM framing blocks, a pilot signal, LLch signal, and TMCC signal are added to the data segments to construct an OFDM frame. OFDM modulation is then applied to the frame by using an inverse fast Fourier transform (IFFT), a GI is added, and the output OFDM signal is generated. The signal for SISO transmission is generated via path (1)-(2)-(3) in Fig. 2.

In the case of MISO transmission, the MISO encoding unit applies space time block code (STBC) or space time frequency code (SFBC) to the carrier symbols for data after band combining [8]. The OFDM frame is then formed and subjected to OFDM modulation. Signal output 1 for MISO transmission is generated by the same path (1)-(2)-(3) as for SISO transmission, and signal output 2 is generated by path (1)-(4)-(5). MISO encoding aims to improve the transmission robustness by exploiting the transmission-diversity in a poor reception environment, and the transmission capacity is equivalent to that in the case of SISO.

Moreover, in the case of advanced ISDB-T, it is possible to configure space-division multiplexing (SDM) MIMO [9], which transmits different information from signal outputs 1 and 2, to expand the transmission capacity. For MIMO transmission, it is advantageous that the transmission capacity can be doubled compared with that possible with SISO; however, to achieve satisfactory transmission characteristics by reducing the correlation between propagation paths 1 and 2 in a line-of-sight reception environment, it is necessary to install transmitting and receiving antennas that support orthogonal polarization (horizontal and vertical polarization), which requires an initial cost for broadcasters and viewers. In MIMO transmission, the level-adjusted carrier symbols are separated by symbol-separation blocks into two streams, OFDM frames are formed and subjected to OFDM modulation, and the two signals are output. Signal outputs 1 and 2 of the MIMO transmission are generated via paths (1)-(2)-(3) and (6)-(7)-(5), respectively. For the FIL in MIMO transmission, different interleaving processes are applied for signal outputs 1 and 2 to improve the interleaving effect.

Although SISO, 2×1 MISO, and 2×2 MIMO are introduced in advanced ISDB-T, in the case of using one transmitting antenna, it is supposed that all layers are transmitted via SISO. In the case where two transmitting antennas are available, MISO or MIMO can be selected for each layer. For example, it is possible to improve the reception robustness with layer A (for mobile reception) using MISO and to expand the transmission capacity with layer B (for fixed reception) using MIMO.

### *B. Bit-Interleaved Coded Modulation*

A detailed configuration of the BICM block is shown in Fig. 3. For FEC, a combination of concatenated codes (with the LDPC code as the inner code and BCH code as the outer code) is adopted. Transmission capacity close to the Shannon limit [10] can be obtained by optimizing LDPC codes in combination with bit interleaving and NUCs. For high code rates of 8/16 or higher, the irregular repeat accumulate (IRA) type parity check matrix is used [11]. For low coding rates of 7/16 or less, the multi edge type (MET) parity check matrix characterized by a structure in which two parity check matrices are combined is used [12].

In the block for FEC block encapsulation shown in Fig. 3, the variable-length TLV packet [13], which is the input data, is encapsulated in a fixed-length FEC block. The configuration of the FEC block is shown in Fig. 4. In the case of advanced ISDB-T, to increase the error-correction capability, codes were designed with an FEC block length of 69,120 bits. This block length is longer than that of ISDB-S3 [4], DVB-T2 [14], and ATSC3.0 [15]. The parity length of the LDPC code is determined according to the code rate of the LDPC code. The information bits of the LDPC code include the parity of the BCH code, and the length of the BCH parity is 192 bits regardless of the code rate of the LDPC code.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

TABLE II
TRANSMISSION PARAMETERS

| | | Advanced ISDB-T | | ISDB-T |
|---|---|---|---|---|
| Channel bandwidth | | 6 MHz | | |
| Number of segment divisions | | 36 | | 14 |
| Bandwidth of segment | | 166.7 kHz | | 428.6 kHz |
| Number of segments | | 35 | 33 + adjusting band | 13 |
| Signal bandwidth | | 5.83 MHz | 5.57 MHz | 5.57 MHz |
| Number of layers | | Up to 3 layers | | |
| Number of segments in partial reception band | | 9 | | 1 |
| FFT size ($N_{FFT}$) | | 8,192, 16,384, 32,768 | | 2,048, 4,096, 8,192 |
| FFT sampling frequency | | 6.321 MHz (=512/81) | | 8.127 MHz (=512/63) |
| Guard interval ratio (GI ratio) | | 1/4, 1/8, 1/16, 1/32, 1/256, $800/N_{FFT}$ | | 1/4, 1/8, 1/16, 1/32 |
| Carrier modulation | | QPSK, 16QAM, 64QAM, 256QAM, 1024QAM, 4096QAM (Uniform, Non-uniform) | | DQPSK, QPSK, 16QAM, 64QAM (Uniform) |
| FEC | Inner code | LDPC code | | Convolutional code |
| | Outer code | BCH code | | RS code |
| System configuration | | SISO, 2×1 MISO, 2×2 MIMO | | SISO |

The variable-length TLV packets are stored in the main data area shown in Fig. 4. When a TLV packet cannot fit in this area, it is divided and partly stored there, while the rest is stored in the main data area of the next FEC block. At that time, the start position of the TLV packets in each FEC block is designated by the TLV packet pointer in the FEC block header shown in Fig. 4. The bit length of each part of the FEC block of advanced ISDB-T is listed in Table I.

In general, as the number of decoding iterations and the block length of the LDPC code increase, a larger amount of calculation is required at the receiver. Although it depends on the decoding algorithm, the computational complexity of LDPC code decoding in advanced ISDB-T is considered larger than that of Viterbi decoding in ISDB-T.

### C. *Transmission Parameters*

The transmission parameters of advanced ISDB-T are listed in Table II. Although the new scheme attempts to expand the transmission capacity compared with that of ISDB-T by enhancing the frame structure (by increasing the FFT size and reducing the GI ratio while inheriting the functions of hierarchical transmission), considering the migration from the current terrestrial broadcasting to the next-generation system, a GI ratio of $800/N_{FFT}$ was included for setting the GI length to 126 μs (which is the same as the operational parameter used in Japan for the current ISDB-T). Here, $N_{FFT}$ indicates the FFT size, which is 8,192 (8k), 16,384 (16k), or 32,768 (32k).

Regarding the signal bandwidth, it is possible to choose either the normal mode (bandwidth: 5.83 MHz, where the guard band is reduced compared with that for ISDB-T) using 35/36 of the channel bandwidth or a compatible mode (5.57 MHz: the same bandwidth as for ISDB-T).

TABLE III
BICM SPECTRAL EFFICIENCY (BIT/SYMBOL)

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | 0.25 | 0.38 | 0.50 | 0.63 | 0.75 | 0.88 | 1.00 | 1.13 | 1.25 | 1.38 | 1.50 | 1.63 | 1.75 |
| 16QAM (4) | 0.50 | 0.75 | 1.00 | 1.25 | 1.50 | 1.75 | 2.00 | 2.25 | 2.50 | 2.75 | 3.00 | 3.25 | 3.50 |
| 64QAM (6) | 0.75 | 1.13 | 1.50 | 1.88 | 2.25 | 2.63 | 3.00 | 3.38 | 3.75 | 4.13 | 4.50 | 4.88 | 5.25 |
| 256QAM (8) | 1.00 | 1.50 | 2.00 | 2.50 | 3.00 | 3.50 | 4.00 | 4.50 | 5.00 | 5.50 | 6.00 | 6.50 | 7.00 |
| 1024QAM (10) | 1.25 | 1.88 | 2.50 | 3.13 | 3.75 | 4.38 | 5.00 | 5.63 | 6.25 | 6.88 | 7.50 | 8.13 | 8.75 |
| 4096QAM (12) | 1.50 | 2.25 | 3.00 | 3.75 | 4.50 | 5.25 | 6.00 | 6.75 | 7.50 | 8.25 | 9.00 | 9.75 | 10.50 |

TABLE IV
TRANSMISSION PARAMETERS FOR PERFORMANCE EVALUATION

| | |
|---|---|
| Number of segments (allocated for layer A: layer B) | 35 (4:31) |
| Signal bandwidth | 5.83 MHz |
| Number of layers | 2 layers |
| FFT size ($N_{FFT}$) | 16,384 |
| Sub-carrier spacing | 385.8 Hz |
| Guard interval ratio | 800/16,384 |
| Effective symbol length | 2592 μs |
| Guard interval length | 126.56 μs |
| Scattered pilot ratio (Dx,Dy) | 1/12 (6,2) |
| Scattered pilot boost (amplitude) | 1.29 |
| System configuration | SISO |

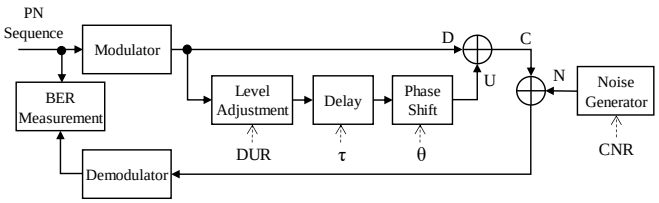



Fig. 5. Block diagram of simulation.

## IV. METHODOLOGY AND SIMULATION SETUP

### A. *Methodology and Transmission Parameters*

As shown in Table II, there are six options for carrier modulation {QPSK, 16QAM, 64QAM, 256QAM, 1024QAM, 4096QAM} corresponding to the number of transmission bits per symbol $V$= {2, 4, 6, 8, 10, 12}. There are thirteen types of LDPC code rates $R$= {2/16, 3/16, …, 14/16}. In this paper, the spectral efficiency of BICM is defined as $V \times R$ [bit/symbol]. Table III shows the spectral efficiency, and combinations having the same spectral efficiency are shown in the same color. As shown, there are multiple combinations with the same spectral efficiency. Fixed reception characteristics were evaluated for the AWGN channel and echo channels for all combinations of BICM with actual channel estimation. The transmission parameters used in the simulations are listed in Table IV. For LDPC code decoding, the layered belief propagation [16] was applied with 25 iterations. For comparison with the current DTTB system, the transmission performance was extensively evaluated with the SISO system. Since MISO is introduced to improve the robustness for mobile reception, this paper focuses on the performance evaluation on SISO in fixed reception. In this paper, we evaluated the fixed reception characteristics of BICM using NUCs and UCs.

### B. *Simulation Setup for Fixed Reception Evaluation*

Figure 5 shows a block diagram of the simulation for evaluating the fixed reception performance. The bit error rate (BER) of layer B (for fixed reception) and corresponding

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

TABLE V
CRITERIA OF BICM SELECTION METHODS FOR FIXED RECEPTION

| | |
|---|---|
| Method 1-1 | Required CNR for AWGN |
| Method 1-2 | Average required CNR for AWGN and 10-dB echo channels |
| Method 1-3 | Average required CNR for AWGN and 3-dB echo channels |
| Method 1-4 | Average required CNR for AWGN, 10-dB echo, and 3-dB echo channels |
| Method 2 | Step 1: required CNR for AWGN<br>Step 2: spectral efficiency and required CNR for 3-dB echo |

CNR were evaluated. We defined the signal power "C" as the total power of main wave "D" and the echo wave "U". The CNR was calculated with the total signal power of layers A and B, and the power boost function of layer A was deactivated. The minimum CNR that achieves quasi error free (QEF) data transmission was defined as the required CNR, and the QEF condition was defined as the BER after LDPC code decoding being lower than $10^{-7}$. The power ratio between the main wave and the echo in the multipath environment was defined as the desired-to-undesired ratio (DUR), and the phase θ of the echo signal was set to π/2. The delay time between the main wave and the echo was set to a delay time τ = 63 μsec, which is half the GI duration and the value used in the verification of current ISDB-T as a typical value for a practical reception [1]. To confirm the effect of the delay time, the degradation of the required CNR was verified with a short delay time τ = 1 μsec.

### C. BICM Selection

Methods for selecting BICM were verified on the basis of the evaluated transmission performance of all BICM combinations (78 options). For all BICMs, the required CNRs for the AWGN and echo channels were compared, and recommended BICM combinations that had the lowest required CNRs were selected. The following two methods were compared.

#### 1) Method 1: One-step selection

Method 1 is a straightforward selection procedure based on the required CNRs for a specific transmission channel. If there are BICM combinations that have the same spectral efficiency, the one that achieves the lowest required CNR should be selected. We considered four typical criteria that are considered reasonable for the fixed reception environment. They are summarized in Table V. Method 1-1 considers only the required CNR for the AWGN channel. Method 1-2 is based on the average required CNR for the AWGN and 10-dB echo channels. Method 1-3 uses the average required CNR for the AWGN and 3-dB echo channels. Method 1-4 is based on the three required CNRs for the AWGN, 3-dB echo, and 10-dB echo channels. The channels of 3-dB echo and 10-dB echo were used in the verification of current ISDB-T as typical values for a practical reception [1].

#### 2) Method 2: Two-step selection

Method 2 is a two-step selection procedure. First, BICM is selected for the AWGN channel in the same manner as Method 1-1. For the surviving BICM combinations in step 1, which have unique spectral efficiency, the final selection is conducted in consideration of the spectral efficiency and the required CNR for the 3-dB echo channel. Specifically, we compared the required CNRs for the 3-dB echo channel of the surviving BICM combinations and when there is a BICM that outperforms others in terms of both the required CNR and spectral efficiency, the inferior one is excluded.

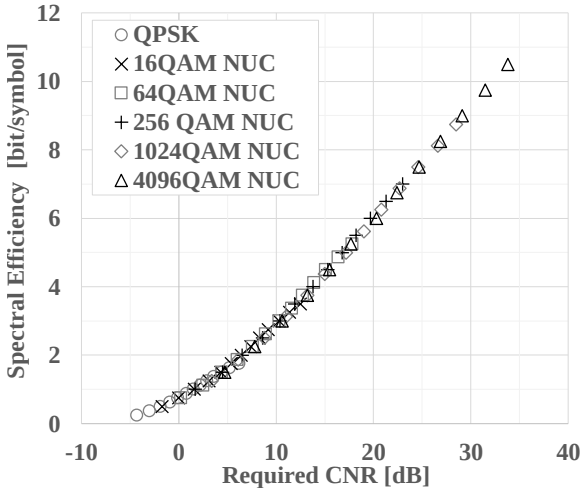


Fig. 6. Required CNR of NUCs for AWGN channel.

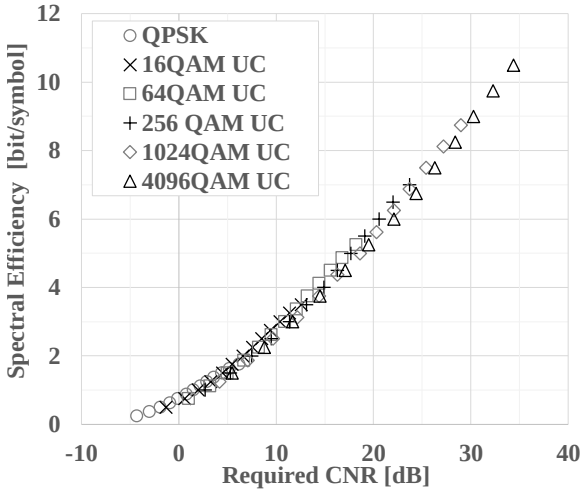


Fig. 7. Required CNR of UCs for AWGN channel.

## V. SIMULATION RESULTS

### A. Transmission Performance for AWGN Channel

Figure 6 shows the results for the required CNR calculated in simulation for the AWGN channel for all BICM combinations with NUCs. From Fig. 6, each BICM was distributed linearly, and the required CNR corresponding to a spectral efficiency of 0.25 to 10.5 [bps/symbol] was from -5 dB to 35 dB. Figure 7 shows the required CNR for the UCs. It can be seen that the required CNR slightly increased (degraded) for the combination of high-order modulation and low code rate of LDPC codes.

The relationship between the transmission capacity per channel of advanced ISDB-T and the required CNR is plotted in Fig. 8. The results were calculated with the parameters in Table IV with the exception that all 35 segments were allocated for one layer. The capacity of the MIMO system is also shown in Fig. 8. When the advanced ISDB-T SISO (circles) and ISDB-T scheme (diamond) are compared, it can be seen that, for the former scheme, (i) the transmission capacity increased by about 10 Mbps with the same required CNR, and (ii) the required CNR was reduced by about 7 dB

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

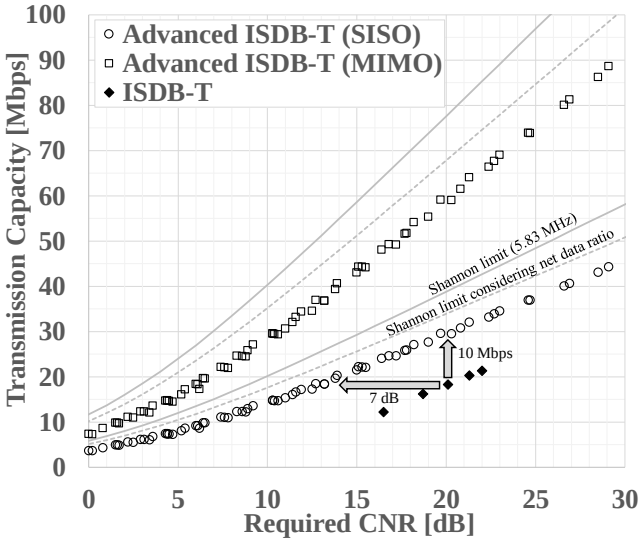



Fig. 8. Transmission capacity for SISO and MIMO systems with NUCs for AWGN channel. Advanced ISDB-T with 35 segments (5.83 MHz) and ISDB-T with 13 segments (5.57 MHz).

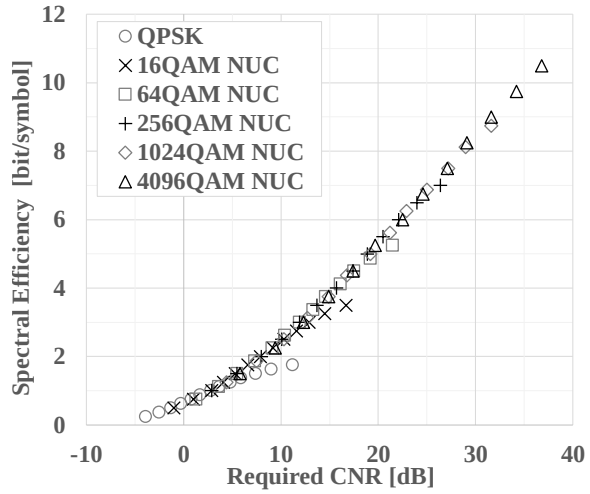



Fig. 9. Required CNR of NUCs for 3-dB echo channel.

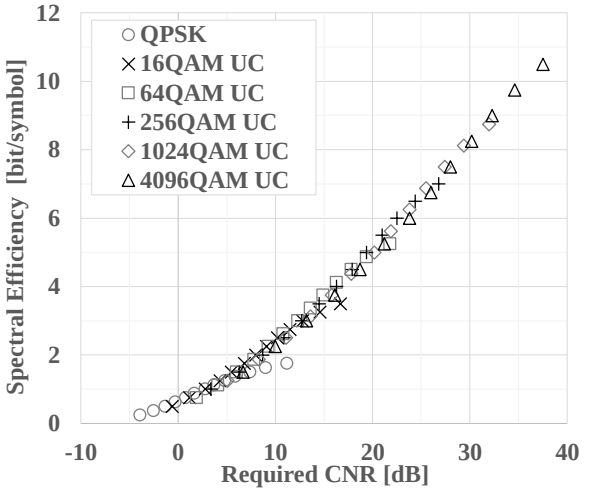



Fig. 10. Required CNR of UCs for 3-dB echo channel.

while keeping the same transmission capacity. As to MIMO configuration, the transmission capacity was doubled compared with SISO configuration, when the overhead signal ratio is the same as SISO.

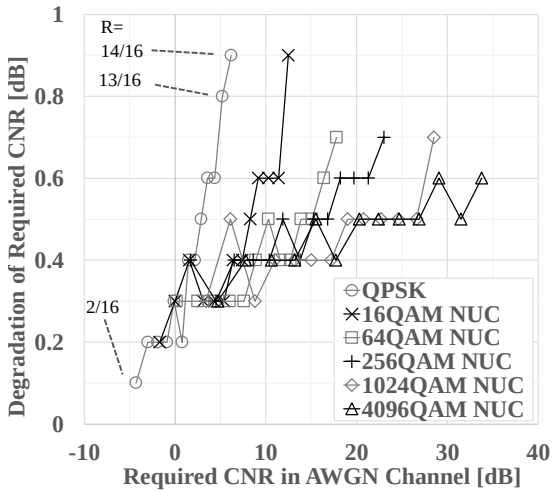



Fig. 11. Degradation of required CNR with NUCs for 10-dB echo channel.

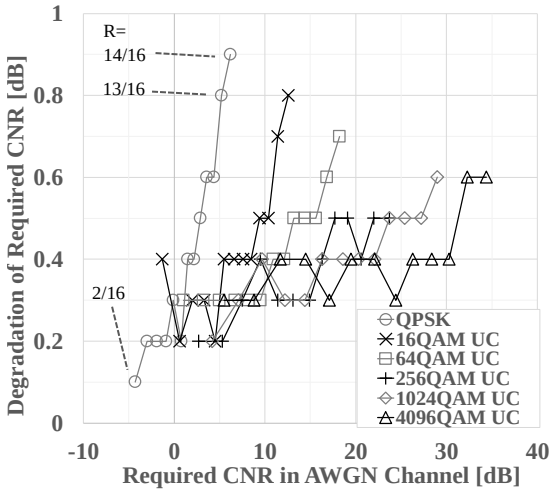



Fig. 12. Degradation of required CNR with UCs for 10-dB echo channel.

### *B. Transmission Performance in Echo Channel*

Figures 9 and 10 show the results for the required CNR for the 3-dB echo channel for NUCs and UCs, respectively. Almost all the BICMs were distributed on a straight line; however, some with high code rates diverged slightly to the right. Table VI shows a summary of the required CNR for the AWGN and echo channels when NUCs were configured. In addition to the 3-dB echo channel, the required CNRs for the 10-dB echo channel were also evaluated. The deterioration in the required CNR from that for the AWGN channel was calculated, and the degradation values are plotted in Figs. 11 and 12, which show the degradation of NUCs and UCs, respectively. Regardless of the carrier modulation whether NUC or UC was used, the degradation of the required CNR for the 10-dB echo channel became within 1 dB. These results indicate that the multipath margin integrated with the required electric field strength used for the link budget of advanced ISDB-T can be 1 dB or less assuming the same 10-dB echo channel used in the current ISDB-T planning used in Japan [1].

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

TABLE VI
SUMMARY OF REQUIRED CNR FOR NUCS IN AWGN CHANNEL AND ECHO CHANNELS

| Channel | Carrier modulation | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| AWGN | QPSK | -4.3 | -3 | -1.9 | -0.9 | -0.1 | 0.8 | 1.5 | 2.2 | 2.9 | 3.6 | 4.4 | 5.2 | 6.2 |
| | 16QAM | -1.7 | 0 | 1.6 | 3.1 | 4.3 | 5.4 | 6.4 | 7.4 | 8.3 | 9.2 | 10.3 | 11.4 | 12.5 |
| | 64QAM | 0.2 | 2.5 | 4.4 | 6 | 7.6 | 8.9 | 10.3 | 11.6 | 12.7 | 13.9 | 15.1 | 16.4 | 17.8 |
| | 256QAM | 1.7 | 4.5 | 6.5 | 8.6 | 10.4 | 11.9 | 13.8 | 15.3 | 16.8 | 18.2 | 19.7 | 21.3 | 23 |
| | 1024QAM | 3.4 | 6.1 | 8.8 | 11 | 13.2 | 15 | 17.2 | 19 | 20.8 | 22.7 | 24.6 | 26.6 | 28.5 |
| | 4096QAM | 4.7 | 7.8 | 10.6 | 13.2 | 15.5 | 17.7 | 20.3 | 22.4 | 24.7 | 26.9 | 29.1 | 31.5 | 33.8 |
| 3-dB echo | QPSK | -3.9 | -2.5 | -1.3 | -0.3 | 0.8 | 1.7 | 2.8 | 3.7 | 4.8 | 5.9 | 7.4 | 9 | 11.2 |
| | 16QAM | -1 | 1 | 2.9 | 4.1 | 5.3 | 6.6 | 7.9 | 9.2 | 10.3 | 11.6 | 12.9 | 14.5 | 16.7 |
| | 64QAM | 1.3 | 3.6 | 5.5 | 7.3 | 9.1 | 10.4 | 11.9 | 13.3 | 14.6 | 16.1 | 17.5 | 19.2 | 21.5 |
| | 256QAM | 2.9 | 5.5 | 8 | 10.1 | 11.9 | 13.7 | 15.7 | 17.4 | 18.9 | 20.5 | 22.1 | 24 | 26.4 |
| | 1024QAM | 4.4 | 7.5 | 10.3 | 12.7 | 14.9 | 16.8 | 19.2 | 21.2 | 22.9 | 25 | 27.2 | 29 | 31.6 |
| | 4096QAM | 5.8 | 9.4 | 12.3 | 14.9 | 17.4 | 19.7 | 22.5 | 24.6 | 27.1 | 29.1 | 31.6 | 34.2 | 36.8 |
| 10-dB echo | QPSK | -4.2 | -2.8 | -1.7 | -0.7 | 0.2 | 1 | 1.9 | 2.6 | 3.4 | 4.2 | 5 | 6 | 7.1 |
| | 16QAM | -1.5 | 0.3 | 2 | 3.4 | 4.6 | 5.7 | 6.8 | 7.8 | 8.8 | 9.8 | 10.9 | 12 | 13.4 |
| | 64QAM | 0.5 | 2.8 | 4.7 | 6.3 | 7.9 | 9.3 | 10.8 | 12 | 13.1 | 14.4 | 15.6 | 17 | 18.5 |
| | 256QAM | 2.1 | 4.8 | 6.9 | 9 | 10.8 | 12.4 | 14.2 | 15.8 | 17.3 | 18.8 | 20.3 | 21.9 | 23.7 |
| | 1024QAM | 3.7 | 6.6 | 9.1 | 11.4 | 13.6 | 15.4 | 17.6 | 19.5 | 21.3 | 23.2 | 25.1 | 27.1 | 29.2 |
| | 4096QAM | 5 | 8.2 | 11 | 13.6 | 16 | 18.1 | 20.8 | 22.9 | 25.2 | 27.4 | 29.7 | 32 | 34.4 |

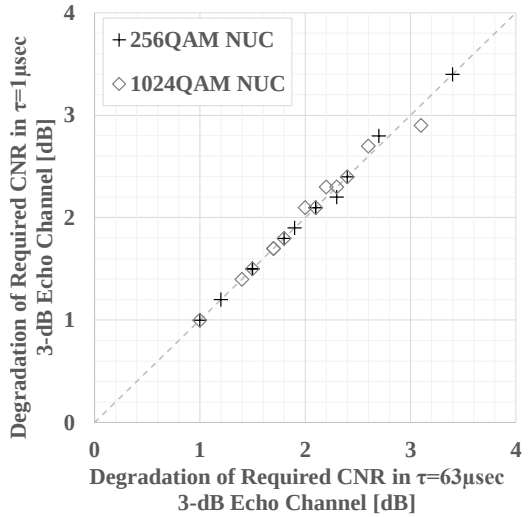



Fig. 13. Comparison of degradation of required CNR with NUCs in short and long delay 3-dB echo channels.

Figure 13 shows the degradation of the required CNR for the 3-dB echo channel with NUCs for delay time τ = 1 μsec and 63 μsec. It was observed that even when the delay time was very short, the degradation amount was almost the same as that observed for the τ = 63 μsec echo channel.

## C. *BICM Selection*

### 1) *Method 1: One-step selection*

For NUCs, the BICM selection results with Methods 1-1, 1-2, 1-3, and 1-4 are shown in Tables VII, VIII, IX, and X. The excluded combinations are shown in gray. As an example of a spectral efficiency of 1.00 bps/Hz, the BICMs are shown in brown in Table III, i.e., (QPSK, 8/16), (16QAM, 4/16), (256QAM, 2/16). From Table V, the required CNRs for AWGN were 1.5, 1.6, and 1.7 dB, respectively. In this case, (QPSK, 8/16) was selected with Method 1-1, and the remaining two combinations were excluded. It was confirmed that the difference between the BICMs selected by Methods 1-1 and 1-2 was (16QAM, 12/16). With Method 1-3, two combinations, (QPSK, 10/16) and (16QAM, 12/16), were excluded, and two combinations, (16QAM, 5/16) and (4096QAM, 10/16), survived when compared with the results of Method 1-1. In the results for Method 1-4, (QPSK, 10/16)

TABLE VII
METHOD 1-1: NUC BICM SELECTION RESULTS AND REQUIRED CNR FOR AWGN CHANNEL

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -4.3 | -3.0 | -1.9 | -0.9 | -0.1 | 0.8 | 1.5 | 2.2 | 2.9 | 3.6 | 4.4 | 5.2 | 6.2 |
| 16QAM (4) | -1.7 | 0.0 | 1.6 | 3.1 | 4.3 | 5.4 | 6.4 | 7.4 | 8.3 | 9.2 | 10.3 | 11.4 | 12.5 |
| 64QAM (6) | 0.2 | 2.5 | 4.4 | 6.0 | 7.6 | 8.9 | 10.3 | 11.6 | 12.7 | 13.9 | 15.1 | 16.4 | 17.8 |
| 256QAM (8) | 1.7 | 4.5 | 6.5 | 8.6 | 10.4 | 11.9 | 13.8 | 15.3 | 16.8 | 18.2 | 19.7 | 21.3 | 23.0 |
| 1024QAM (10) | 3.4 | 6.1 | 8.8 | 11.0 | 13.2 | 15.0 | 17.2 | 19.0 | 20.8 | 22.7 | 24.6 | 26.6 | 28.5 |
| 4096QAM (12) | 4.7 | 7.8 | 10.6 | 13.2 | 15.5 | 17.7 | 20.3 | 22.4 | 24.7 | 26.9 | 29.1 | 31.5 | 33.8 |

TABLE VIII
METHOD 1-2: NUCS BICM SELECTION RESULTS AND AVERAGE REQUIRED CNR FOR AWGN AND 10-DB ECHO CHANNELS

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -4.25 | -2.90 | -1.80 | -0.80 | 0.05 | 0.90 | 1.70 | 2.40 | 3.15 | 3.90 | 4.70 | 5.60 | 6.65 |
| 16QAM (4) | -1.60 | 0.15 | 1.80 | 3.25 | 4.45 | 5.55 | 6.60 | 7.60 | 8.55 | 9.50 | 10.60 | 11.70 | 12.95 |
| 64QAM (6) | 0.35 | 2.65 | 4.55 | 6.15 | 7.75 | 9.10 | 10.55 | 11.80 | 12.90 | 14.15 | 15.35 | 16.70 | 18.15 |
| 256QAM (8) | 1.90 | 4.65 | 6.70 | 8.80 | 10.60 | 12.15 | 14.00 | 15.55 | 17.05 | 18.50 | 20.00 | 21.60 | 23.35 |
| 1024QAM (10) | 3.55 | 6.35 | 8.95 | 11.20 | 13.40 | 15.20 | 17.40 | 19.25 | 21.05 | 22.95 | 24.85 | 26.85 | 28.85 |
| 4096QAM (12) | 4.85 | 8.00 | 10.80 | 13.40 | 15.75 | 17.90 | 20.55 | 22.65 | 24.95 | 27.15 | 29.40 | 31.75 | 34.10 |

TABLE IX
METHOD 1-3: NUCS BICM SELECTION RESULTS AND AVERAGE REQUIRED CNR FOR AWGN AND 3-DB ECHO CHANNELS

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -4.10 | -2.75 | -1.60 | -0.60 | 0.35 | 1.25 | 2.15 | 2.95 | 3.85 | 4.75 | 5.90 | 7.10 | 8.70 |
| 16QAM (4) | -1.35 | 0.50 | 2.25 | 3.60 | 4.80 | 6.00 | 7.15 | 8.30 | 9.30 | 10.40 | 11.60 | 12.95 | 14.60 |
| 64QAM (6) | 0.75 | 3.05 | 4.95 | 6.65 | 8.35 | 9.65 | 11.10 | 12.45 | 13.65 | 15.00 | 16.30 | 17.80 | 19.65 |
| 256QAM (8) | 2.30 | 5.00 | 7.25 | 9.35 | 11.15 | 12.80 | 14.75 | 16.35 | 17.85 | 19.35 | 20.90 | 22.65 | 24.70 |
| 1024QAM (10) | 3.90 | 6.80 | 9.55 | 11.85 | 14.05 | 15.90 | 18.20 | 20.10 | 21.85 | 23.85 | 25.90 | 27.80 | 30.05 |
| 4096QAM (12) | 5.25 | 8.60 | 11.45 | 14.05 | 16.45 | 18.70 | 21.40 | 23.50 | 25.90 | 28.00 | 30.35 | 32.85 | 35.30 |

TABLE X
METHOD 1-4: NUC BICM SELECTION RESULTS AND AVERAGE REQUIRED CNR FOR AWGN, 3-DB, 10-DB ECHO CHANNELS

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -4.13 | -2.77 | -1.63 | -0.63 | 0.30 | 1.17 | 2.07 | 2.83 | 3.70 | 4.57 | 5.60 | 6.73 | 8.17 |
| 16QAM (4) | -1.40 | 0.43 | 2.17 | 3.53 | 4.73 | 5.90 | 7.03 | 8.13 | 9.13 | 10.20 | 11.37 | 12.63 | 14.20 |
| 64QAM (6) | 0.67 | 2.97 | 4.87 | 6.53 | 8.20 | 9.53 | 11.00 | 12.30 | 13.47 | 14.80 | 16.07 | 17.53 | 19.27 |
| 256QAM (8) | 2.23 | 4.93 | 7.13 | 9.23 | 11.03 | 12.67 | 14.57 | 16.17 | 17.67 | 19.17 | 20.70 | 22.40 | 24.37 |
| 1024QAM (10) | 3.83 | 6.73 | 9.40 | 11.70 | 13.90 | 15.73 | 18.00 | 19.90 | 21.67 | 23.63 | 25.63 | 27.57 | 29.77 |
| 4096QAM (12) | 5.17 | 8.47 | 11.30 | 13.90 | 16.30 | 18.50 | 21.20 | 23.30 | 25.67 | 27.80 | 30.13 | 32.57 | 35.00 |

TABLE XI
METHOD 2: NUC BICM SELECTION RESULTS AND REQUIRED CNR FOR 3-DB ECHO CHANNEL

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -3.9 | -2.5 | -1.3 | -0.3 | 0.8 | 1.7 | 2.8 | 3.7 | 4.8 | 5.9 | 7.4 | 9.0 | 11.2 |
| 16QAM (4) | -1.0 | 1.0 | 2.9 | 4.1 | 5.3 | 6.6 | 7.9 | 9.2 | 10.3 | 11.6 | 12.9 | 14.5 | 16.7 |
| 64QAM (6) | 1.3 | 3.6 | 5.5 | 7.3 | 9.1 | 10.4 | 11.9 | 13.3 | 14.6 | 16.1 | 17.5 | 19.2 | 21.5 |
| 256QAM (8) | 2.9 | 5.5 | 8.0 | 10.1 | 11.9 | 13.7 | 15.7 | 17.4 | 18.9 | 20.5 | 22.1 | 24.0 | 26.4 |
| 1024QAM (10) | 4.4 | 7.5 | 10.3 | 12.7 | 14.9 | 16.8 | 19.2 | 21.2 | 22.9 | 25.0 | 27.2 | 29.0 | 31.6 |
| 4096QAM (12) | 5.8 | 9.4 | 12.3 | 14.9 | 17.4 | 19.7 | 22.5 | 24.6 | 27.1 | 29.1 | 31.6 | 34.2 | 36.8 |

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

TABLE XII
METHOD 1-1: UC BICM SELECTION RESULTS AND REQUIRED CNR FOR AWGN CHANNEL

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -4.3 | -3.0 | -1.9 | -0.9 | -0.1 | 0.8 | 1.5 | 2.2 | 2.9 | 3.6 | 4.4 | 5.2 | 6.2 |
| 16QAM (4) | -1.3 | 0.6 | 2.0 | 3.3 | 4.5 | 5.5 | 6.6 | 7.6 | 8.5 | 9.4 | 10.4 | 11.4 | 12.6 |
| 64QAM (6) | 1.0 | 3.2 | 5.0 | 6.7 | 8.2 | 9.5 | 10.9 | 12.1 | 13.2 | 14.4 | 15.6 | 16.8 | 18.2 |
| 256QAM (8) | 2.7 | 5.3 | 7.5 | 9.5 | 11.4 | 13.1 | 14.9 | 16.3 | 17.7 | 19.1 | 20.6 | 22.0 | 23.7 |
| 1024QAM (10) | 4.2 | 7.1 | 9.7 | 12.2 | 14.4 | 16.3 | 18.6 | 20.3 | 22.1 | 23.7 | 25.4 | 27.2 | 29.0 |
| 4096QAM (12) | 5.5 | 8.8 | 11.7 | 14.5 | 17.1 | 19.5 | 22.1 | 24.4 | 26.3 | 28.4 | 30.3 | 32.3 | 34.4 |

TABLE XIII
METHOD 2: UC BICM SELECTION RESULTS AND REQUIRED CNR FOR 3-DB ECHO CHANNEL

| Carrier modulation (bit/symbol) | Code rate of LDPC code (x/16) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| QPSK(2) | -3.9 | -2.5 | -1.3 | -0.3 | 0.8 | 1.7 | 2.8 | 3.7 | 4.8 | 5.9 | 7.4 | 9.0 | 11.2 |
| 16QAM (4) | -0.6 | 1.2 | 2.8 | 4.3 | 5.5 | 6.8 | 8.0 | 9.1 | 10.2 | 11.5 | 12.8 | 14.6 | 16.7 |
| 64QAM (6) | 1.9 | 4.1 | 6.0 | 7.8 | 9.3 | 10.8 | 12.3 | 13.6 | 14.9 | 16.3 | 17.8 | 19.4 | 21.8 |
| 256QAM (8) | 3.4 | 6.3 | 8.7 | 10.9 | 12.7 | 14.5 | 16.3 | 17.9 | 19.4 | 21.0 | 22.5 | 24.4 | 26.8 |
| 1024QAM (10) | 5.1 | 8.3 | 11.1 | 13.6 | 15.8 | 17.8 | 20.2 | 21.9 | 23.8 | 25.5 | 27.4 | 29.4 | 32.0 |
| 4096QAM (12) | 6.7 | 10.0 | 13.2 | 16.1 | 18.7 | 21.2 | 23.8 | 26.0 | 28.0 | 30.2 | 32.3 | 34.6 | 37.5 |

and (16QAM, 12/16) were excluded, and (16QAM, 5/16) survived. Method 1 is a simple procedure, but the BICM combination excluded was limited to BICMs that had overlapping spectral efficiency.

*2) Method 2: Two-step selection*

For step 1, BICM combinations having the same spectral efficiency with a higher required CNR for the AWGN channel were excluded as with Method 1-1 in step 1.

Table XI shows the BICM selection results for NUCs after step 2 for Method 2. The combinations excluded in step 1 (Method 1-1) are shown in black, and those cut out in step 2 are shown in gray. Explaining the excluded combination example, (QPSK, 11/16) had a spectral efficiency of 1.38 bit/symbol, and the required CNR for the 3-dB echo channel was 5.9 dB. In comparison, (16QAM, 6/16) showed a spectral efficiency of 1.50 bit/symbol with the required CNR of 5.3 dB. In this example, (QPSK, 11/16) that had a smaller spectral efficiency and a higher required CNR than (16QAM, 6/16) was excluded. As in the case just described, if other BICM combinations with high spectral efficiency and low required CNR were available, the undesirable BICMs would be excluded. From Table XI, a BICM combinations with a high code rate and low order modulation tended to be excluded in the BICM selection in step 2 due to their worse transmission performance in echo channels.

For UCs, the BICM selection results for step 1 (Method 1-1) and step 2 are shown in Tables XII and XIII. The combinations excluded in step 1 are shown in black, and those cut out in step 2 are shown in gray in Table XIII. The BICM selection results of UCs were not completely the same as those for NUCs, but BICM combinations with high code rate and low order modulation or low code rate and high order modulation are not recommended for practical use.

## VI. CONCLUSION

Aiming to enhance the functionality and quality of digital terrestrial television broadcasting, NHK has been studying a new transmission scheme for simultaneously providing Super Hi-Vision broadcasting for fixed reception and high-definition broadcasting for mobile reception on a single channel. While inheriting the features of the ISDB-T scheme such as the hierarchical transmission function based on a segment structure, the new advanced ISDB-T has new features: an enhanced signal structure and the latest technology for improving transmission performance and capacity. In particular, the flexibility of allocating bandwidth to multiple services for different receiving modes such as fixed reception and mobile reception is improved. Moreover, by introducing the latest LDPC code, NUC, MISO/MIMO technology for improving performance and capacity, the new transmission scheme attains a higher spectral efficiency, larger capacity, and better transmission robustness than ISDB-T.

The fixed reception performance of all BICM combinations of the advanced ISDB-T was evaluated for AWGN and echo channels. As a result of an evaluation by computer simulation, several BICM selection methods were demonstrated to select the recommended combination for fixed reception. It was confirmed that BICM combinations with a high code rate and low modulation order or low code rate and high modulation order were not recommended. In an echo channel with DUR = 10 dB, the degradation in the required CNR was below 1 dB for all combinations of BICM regardless of whether NUCs or UC were used.

In the future, we will evaluate other transmission parameters, verify them in indoor experiments and field experiments, and proceed with demonstrations that will contribute to the formulation of a link budget and planning parameters for the practical application of advanced ISDB-T broadcasting networks.

## ACKNOWLEDGEMENT

This research was partly performed under the auspices of the Ministry of Internal Affairs and Communications, Japan, as a part of its program “Research and Development for Advanced Digital Terrestrial TV Broadcasting System.”

The authors are grateful to Sony Co. for collaborating on the design of the forward error correction code, bit interleave, and non-uniform constellation.

## REFERENCES

[1] ARIB STD-B31 ver. 2.2, Transmission System for Digital Terrestrial Television Broadcasting, 2014.
[2] R.W. Chang and R.A. Gabby, “A Theoretical Study of Performance of an Orthogonal Multiplexing Data Transmission Scheme,” IEEE Trans on Commun., COM-16, pp. 529-540, 1968.
[3] M. Nakamura et al., “A Study on the Transmission System for Advanced ISDB-T,” Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, 2019.
[4] N. Shirai et al., “Transmission System Design of UHD-1/4K and UHD-2/8K Terrestrial Television Broadcasting and its Performance Proof by Large-scaled Field Experiments,” SMPTE Motion Imaging Journal, vol. 126, issue 6, pp. 35-42, 2020.
[5] ARIB STD-B44 ver. 2.1, Transmission System for Advanced Wide Band Digital Satellite Broadcasting (ISDB-S3), 2016.
[6] A. Guillén i Fàbregas, A. Martinez, and G. Caire, “Bit-interleaved Coded Modulation,” Foundations and Trends in Communications and Information Theory, vol. 5, no. 1-2, pp. 1-144, 2008.
[7] G.D. Forney Jr. and L-F Wei, “Multidimensional Constellations - Part I: Introduction, Figures of Merit, and Generalized Cross Constellations,” IEEE Journal on Selected Areas in Communications, vol. 7, no. 6, pp. 877-892, 1989.
[8] S.M. Alamouti, “A Simple Transmit Diversity Technique for Wireless Communications,” IEEE J. Select. Areas Commun., vol. 16, no. 8, pp. 1451-1458, 1998.
[9] R.V.L. Hartley, “Transmission of Information,” Bell System Technical Journal, 1928.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

[10] G.J. Foschini, "Layered Space-time Architecture of Wireless Communication in a Fading Environment when using Multiple Antennas," Bell Labs. Tech. J., vol. 1, no. 2, pp. 41-59, 1996.
[11] H. Jin, A. Khandekar, and R. McEliece, "Irregular Repeat-accumulate Codes," Proc. 2nd Int. Symp. Turbo Codes, Related Topics, pp. 1-8, 2000.
[12] T. Richardson and R. Urbanke, "Multi-edge Type LDPC Codes," Proc. Workshop Honoring Prof. Bob McEliece 60th Birthday, Pasadena, CA, USA, pp. 24–25, 2002.
[13] Recommendation. ITU-R BT.1869-0, "Multiplexing Scheme for Variable-length Packets in Digital Multimedia Broadcasting Systems," 2010.
[14] ETSI EN 302 755 V1.4.1, Digital Video Broadcasting (DVB); Frame Structure, Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2), 2015.
[15] A/322:2017, ATSC Standard: Physical Layer Protocol, 2017.
[16] D. E. Hocevar, "A Reduced Complexity Decoder Architecture via Layered Decoding of LDPC Codes," Proc. IEEE Workshop on Signal Processing Systems, 2004.

**Takuya Shitomi** received ME degree in system science from the Tokyo Institute of Technology, Japan in 2005. He joined NHK in 2005. Since 2009, he has been working on next-generation terrestrial broadcasting system with NHK STRL. He is a research engineer in the Advanced Transmission Systems Research Division. He held visiting research appointment with the iTEAM, Universitat Politecnica de Valencia, Spain, in 2016. He has participated in the standardization of broadcast technologies in ATSC 3.0 and International Telecommunication Union – Radiocommunication (ITU-R) Study Group 6.

**Shingo Asakura** received BS, ME and PhD degrees from the Tokyo Institute of Technology, Tokyo, Japan. He joined NHK in 2006. Since 2010, he has been working for NHK STRL. He is a research engineer for the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems for UHDTV. He is a member of the ITE.

**Shogo Kawashima** received BE and ME degrees in electrical engineering from Kyoto University, Kyoto, Japan. He joined NHK in 2018 and has been working for NHK STRL. He is a research engineer for the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems for UHDTV. He is a member of the ITE.

**Akihiko Sato** received BE and ME degrees in electrical engineering from Waseda University, Tokyo, Japan. He joined Japan Broadcasting Corporation (NHK), Tokyo in 2011. Since 2014, he has been working for NHK STRL. He is a research engineer for the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems for UHDTV. He is a member of the ITE.

**Hiroaki Miyasaka** received BE and ME degrees in electrical engineering from the Tokyo University of Science, Chiba, Japan. He joined NHK in 2009. Since 2013, he has been working for NHK STRL. He is a research engineer for the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems for UHDTV.

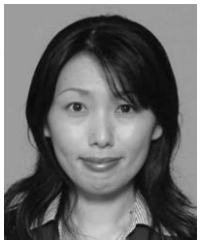
**Noriyuki Shirai** received BE and ME degrees in electrical engineering from the Tokyo University of Science, Chiba, Japan. He joined NHK in 2004. From 2013 to 2020, he was a research engineer with the NHK STRL, Tokyo. His research interest included the next-generation DTTB systems. Currently, he is a senior manager with the Nagoya Station. He is a member of the ITE.

**Yoshikazu Narikiyo** received BE and ME degrees in electrical and computer engineering form Yokohama National University, Kanagawa, Japan, in 2000 and 2002, respectively. He joined NHK in 2002 and researched mobile reception for ISDB-T and next-generation terrestrial broadcasting system in NHK STRL until 2017.Currently, He is a senior manager with the Engineering Administration Department. He is a member of the ITE.

**Tomoaki Takeuchi** received BE, ME, and PhD degrees from Keio University, Japan in 1997, 1999, and 2013, respectively. He joined NHK in 1999. Since 2002, he has been with the NHK STRL. During 2013–2014, he was with NHK Engineering System, Inc. His research interests include digital signal processing and terrestrial broadcasting systems.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

**Kohei Kambara** received BE and ME degrees in electrical and computer engineering from Yokohama National University, Kanagawa, Japan, in 1999 and 2001. He joined NHK in 2001. From 2001-2009 and since 2019, he has been working for NHK STRL. He is a senior research engineer of the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems for UHDTV.



**Madoka Nakamura** received BE and ME degrees in electric engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1998 and 2000, respectively. She joined NHK in 2000. She worked for NHK STRL from 2006 to 2019. She researched optical transmission systems for UHDTV signals, mobile and handheld reception for ISDB-TSB, and the next-generation digital terrestrial television broadcasting (DTTB) systems. Currently, she is a senior manager with the Engineering Administration Department, NHK. She is a member of the ITE and the Institute of Electronics, Information and Communication Engineers (IEICE).

**Tsuyoshi Nakatogawa** received BE and ME degrees in electrical and computer engineering from Yokohama National University, Kanagawa, Japan, in 1998 and 2000. He joined NHK in 2000.
Since 2003, he has been working for NHK STRL. He is a senior research engineer for the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems for UHDTV.

**Kenichi Murayama** received his ME degree in mechanical engineering from Niigata University, Japan in 1996. He joined Japan Broadcasting Corporation (NHK) in 2002. From 2008 to 2018, he worked for NHK STRL and engaged in research and development related to the next generation of digital terrestrial broadcasting. He participated in the standardization process of ATSC 3.0. Currently, he is the senior manager of NHK Engineering Administration Department.

**Masahiro Okano** received BE degree from the University of Electro-Communications, Tokyo, Japan, in 1993. He joined NHK, in 1993, has been with NHK STRL, since 1995. He has been engaged in research on digital terrestrial television broadcasting systems. His current research interests include terrestrial wireless transmission and next-generation terrestrial broadcasting systems. He is a member of the ITE.

**Kenichi Tsuchida** received BE and ME degrees from Tohoku University, Sendai, Japan, in 1988 and 1990, respectively. He joined NHK in 1990.Since 2011, he has been working for NHK STRL. He is the head of the Advanced Transmission Systems Research Division and is engaged in the development of next-generation terrestrial broadcasting systems. He is a member of the ITE.

Received in 2020-10-12 | Approved in 2020-11-16

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.1 Web Link: http://dx.doi.org/10.18580/setijbe.2020.1

# ATSC 3.0 for Future Broadcasting: Features and Extensibility

Sungjun Ahn,
Sunhyoung Kwon,
Seok-Ki Ahn,
Hoiyoon Jung and
Sung-Ik Park

**CITE THIS ARTICLE**

Ahn, Sungjun; Kwon, Sunhyoung; Ahn, Seok-Ki; Jung, Hoiyoon; Park, Sung-Ik; 2020. ATSC 3.0 for Future Broadcasting: Features and Extensibility. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2020.2. Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

**COPYRIGHT** This work is made available under the Creative Commons - 4.0 International License. Reproduction in whole or in part is permitted provided the source is acknowledged.

# ATSC 3.0 for Future Broadcasting: Features and Extensibility

Sungjun Ahn, Sunhyoung Kwon, Seok-Ki Ahn, Hoiyoon Jung, and Sung-Ik Park

***Abstract*— A recent development of Advanced Television Systems Committee (ATSC) 3.0 has made over-the-air services of a 4K ultra-high-definition and a simultaneous multiple high-definition soft-landed to reality. However, ATSC 3.0 is in essence designed to be forward compatible, and is hence able to provide better extensible features to enrich the next media era beyond this initial deployment. This paper introduces several selected features of ATSC 3.0 that could play a prominent role in the near future. This investigation encompasses the optional technologies defined in ATSC 3.0 physical-layer which can improve single frequency network integrity, enable additional features in other domains, or further widen the throughput capability. Majorly focused on the extensibility of ATSC 3.0, we also elaborate on possible inter-network cooperation with broadband and cellular systems.**

***Index Terms*—ATSC 3.0, physical-layer, digital terrestrial broadcasting, broadcast-broadband convergence**

## I. Introduction

ONE of the most notable evolution in modern terrestrial broadcasting is a development of new digital terrestrial television (DTT) standard, so-called Advanced Television Systems Committee (ATSC) 3.0 [1]. To satisfy a wide variety of requirements for the upcoming media era, this new standard has been designed without considering backward compatibility to the legacy ATSC networks [2]. From the first Call for Proposals (CfP) in March 2013, the overall establishment had been progressed across the entire system layers until the first finalized release in January 2018. The first commercial launch of ATSC 3.0 has been held in the Republic of Korea since the mid of 2017 [3]. Driven by a group of major broadcasters, the United States has as well commenced ATSC 3.0 broadcasting since 2020. In the wake of the 600 MHz spectrum repack recently concluded in the United States [4], ATSC 3.0 is planned to be distributed nationwide by the end of 2020. Currently, the on-air services in Korea are in the majority oriented to streaming a 4K ultra-high-definition (UHD) video, while the United States stakeholders focus on providing dozens of service channels over a single 6 MHz-width radio frequency (RF) band, which may boost a cord-cutting faster.

Thanks to a manifold of cutting-edge physical-layer technologies integrated therein, ATSC 3.0 ensures 30% enhanced capacity over ATSC 1.0 for the same coverage [5]. More than this distinct throughput improvement, the new standard also provides an intensely flexible transmission unlike the previous DTT standards existed ever [6]-[8]. For example, even the combination of forward error correction (FEC) and constellation mapping alone has 360 different options. Up to 60 Mbps[1] over up to 64 different physical-layer pipes (PLPs) are capable within a single RF channel thereon [9]. A combination of various interleaving and waveform parameters are also given selectable by purpose, thereby allowing accurate customization for each use case and single frequency network (SFN) topology [10], [11]. Such physical-layer flexibility enables an SFN infrastructure to offer mobile and stationary services or the mixture thereof without undesirable leakage of resources, hence acquiring substantial network fidelity and resource/cost efficiency. Impressively, this physical-layer flexibility can potentially be extended wider by leveraging a future compatible versioning of bootstrap signals [12].

Another noticeable feature of ATSC 3.0 is an Internet Protocol (IP)-based design across the entire system [13]. The overall IP compatibility grants ATSC 3.0 an extensibility to cooperate with other radio or wired access networks [14], [15]. This possibility, so-called broadcast-broadband convergence, has brought enormous attention from the industry since it can create new service opportunities for which improves and enriches the experience [16]. Beyond the physical-layer flexibility within an ATSC 3.0 standalone network, such converged networks can open the door toward another dimension of flexibility upon system configuration [17]. As pointed out in various publications, the converged network can become better enhanced by being combined with functional gadgets in ATSC 3.0, such as a layered division multiplexing (LDM) [18], [19].

Focusing on the service extensibility of ATSC 3.0, this paper introduces several selected features of ATSC 3.0 that could play a prominent role beyond the initial deployment. This investigation encompasses the optional technologies defined in ATSC 3.0 physical-layer which can improve SFN integrity, enable additional features in other domains, or further widen the throughput capability. We also elaborate on possible inter-network cooperation with broadband and

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2020-0-00846, Development of Convergence Transmission Technology for 5G and ATSC 3.0 Networks). *(Corresponding author: Sung-Ik Park.)*

Sungjun Ahn, Sunhyoung Kwon, Seok-Ki Ahn, Hoiyoon Jung, and Sung-Ik Park are with the Media Research Division, Electronics and Telecommunications Research Institute (ETRI), 218 Gajeong-ro, Yuseong-gu, Daejeon, 305-700 S KOREA (e-mail: {sjahn, shkwon, seokki.ahn, junghy, psi76}@etri.re.kr)

[1] This maximum throughput is subjected to a single physical channel usage. Multiple-input multiple-output (MIMO) [20] or channel bonding (CB) [21] operations could double the capacity, as will be described below.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

cellular systems. We therein present a particular service example that realizes an interplay between the boundaries of broadcast and broadband planes. The discussions in this paper are expected to foster a progressive DTT environment that can readily evolve to support emerging media demands.

This paper is organized as follows. Section II introduces several forward-looking features of ATSC 3.0 physical-layer, and Section III discusses the extensibility of ATSC 3.0 mainly in a broadband convergence perspective. The verification test results for ATSC 3.0 system performance are summarized in Section IV, and Section V finally concludes the paper with some remarks.

## II. New Features of ATSC 3.0 Physical-Layer

ATSC 3.0 physical-layer in essence provides a broad range of improvement over throughput, reliability, flexibility, and future extensibility by using state-of-the-art bit-interleaved coded modulation (BICM) and orthogonal frequency division multiplexing (OFDM) framing technologies. Meanwhile, ATSC 3.0 offers more advanced options for service opportunities by supporting further distinguished features that have hardly been considered in previous DTT systems. This section describes several selected features which have been newly adopted in ATSC 3.0 system and made the new standard further promising.

### *A. Layered Division Multiplexing (LDM)*

For decades, the idea of non-orthogonal multiplexing (NOM) has been probed to be a superseder of traditional time- and frequency-domain orthogonal multiplexing schemes [22]-[24]. The efforts have been mainly focused on the power-domain NOM, i.e., superposition coding, and have revealed its notable gain on spectral efficiency[2]. The first standardized form of NOM was the multi-user simultaneous transmission (MUST) [3] in long-term evolution-advanced (LTE-A) [25]. However, due to some practical problems including a difficulty at uplink scheduling, MUST unfortunately has not yet been commercialized in the industry.

On the other hand, the 2nd effort for standardizing NOM, LDM of ATSC 3.0, has attracted a significant attention to fit in the broadcasters' service plans well [11],[26]-[39]. Particularly the fundamental excellence of power-domain NOM in unequal error protection (UEP) scenarios has been highlighted [40]. Celebrated use cases multiplexing mobile, indoor HD, and fixed UHD services are generally referred to promote LDM, where different coverage or quality of service (QoS) is intended for each service therein [4]. Since broadcasters may attempt to serve as many programs as possible to various types of users within limited spectral resources (that they are subjected to), such spectral efficiency gain from LDM would be significantly desirable in a commercial sense. In addition, unlike MUST in the cellular access system, LDM for DTT is already viable in practice because broadcasting does not have any difficulties on user pairing and scheduling. To be mentioned, the 2018 Winter Olympic Games were broadcasted over ATSC 3.0 LDM in

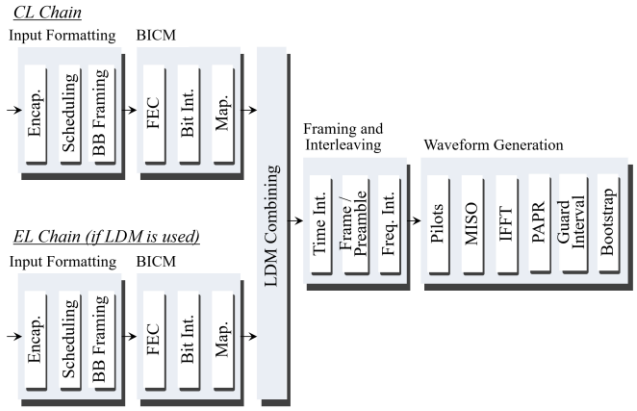



Fig. 1. Block diagram for ATSC 3.0 LDM transmission.

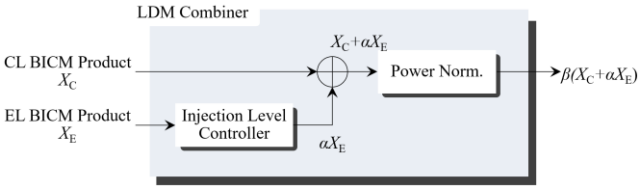



Fig. 2. Description of LDM combining block.

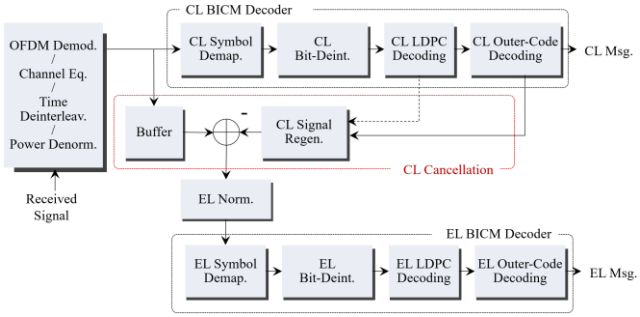



Fig. 3. Block diagram for ATSC 3.0 LDM decoding procedure.

Raleigh, United States. In this on-air demonstration, HD and 4K UHD programs with different target coverages were transmitted on the same radio channel [42].

ATSC 3.0 LDM is specifically a two-layer NOM that combines each, differently powered, layer signal cell-by-cell. The unit *cell* herein refers to a single constellation-mapped signal component that occupies a single subcarrier part within an OFDM symbol. The signal layers of ATSC 3.0 LDM are specified into a *core layer (CL)* and an *enhanced layer (EL)*, which denote the stronger and the weaker signal layers, respectively. The baseband power ratio between those layers is determined by a variable so-called an *injection level*, which is explicitly informed through L1-Detail signaling.

As shown in Fig. 1 and Fig. 2, the post-BICM products of two independent baseband streams are arithmetically combined in a weighted-sum manner. The overall constellation figure of the LDM-combined signal then becomes a form of which the EL constellation is duplicated for each point of the CL constellation. Note herein that every participating PLP undergoes the BICM process independently. The BICM configurations for CL and EL

[2] A superposition coding has been decently proved to be optimal for broadcast channels, in an information-theoretic sense [41].

[3] In specific words, MUST is a non-orthogonal multiple access (NOMA) scheme rather than a simple NOM.

[4] One convincing use case is a scalable video service over LDM, which views UHD to stationary user and moves into HD when the user goes through somewhat harsh environment, such as mobile reception.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

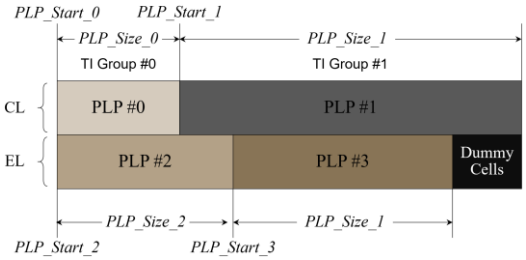



Fig.4. Example of cell positioning and TI under LDM configuration.

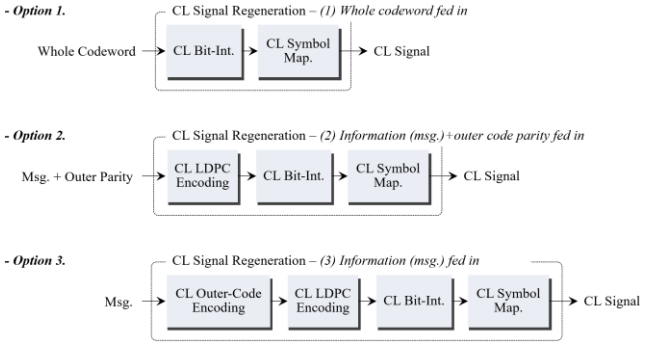



Fig. 5. Implementation options for CL regeneration within SIC block.

PLPs and the contents therein are therefore not subjected to each other. This independency empowers ATSC 3.0 LDM to be versatile to various service attempts, by flexibly satisfying each intended coverage and throughput per PLP.

To offer reasonable signal processing complexity, memory requirement, and end-to-end latency, the interleaving, framing, and waveform generation are applied after the layer combining [27]. That is, for instance, the CL and EL cells paired to each other go through exactly the same time-interleaving (TI). This arrangement reduces the entailed signaling overhead also. The shared TI is defined within the signaling fields subjected to CL PLPs, and each EL cell is just led to follow the same sequence with the CL cell superposed to it. To be clarified, it is not necessary to align the boundaries of CL and EL PLPs to each other. The TI for each EL cell is determined by the cell position within a subframe, and particularly depends on which CL PLP is superposed on it. An auxiliary unit *TI group* is accordingly introduced to clarify TI coordination in complex scenarios. For the sake of clarity, convolutional TI over an LDM frame is carried out on a TI group basis as depicted in Fig. 4 [11].

Since the signal division of LDM does not reside in the time and frequency domain[5], it is also possible to use LDM in combination with time- or frequency-division multiplexing (TDM/FDM). One possible example is a *time-divided LDM (TLDM)*, whose TDM slices are LDM-combined PLPs. In a similar sense, *layered-TDM (LTDM)*, *frequency-divided LDM (FLDM)*, and *layered-FDM (LFDM)* are available as well. Such combinatorial use further fleshes out the transmission flexibility, bringing more opportunity to enhance spectral efficiency. For a specific example transmitting immersive audio, mobile HD, and fixed UHD within a frame, [11] reported that LTDM could achieve 120% capacity gain over TDM for the mobile HD PLP.

Although there is no stipulated restriction on BICM configuration for ATSC 3.0 LDM, it is typical to use a better robust modulation-and-coding (ModCod) combination for CL than that of EL. [8], the recommended guideline document for ATSC 3.0 physical-layer, further suggests to use up to 64 non-uniform constellation (NUC) for CL PLP, and to use up to 7/15-rate low-density parity check (LDPC) code if the CL PLP is modulated with 64 NUC.

Such considerations for CL robustness (relative to EL PLPs) attributes to inter-layer interference (ILI) due to the EL injection. In general, a typical ATSC 3.0 receiver attempts to retrieve the CL signal directly from the received signal but does not extract any constructive information from EL components during then. When a CL PLP is intended, EL signals are therefore regarded as inevitable ILI that degrades the signal quality. Sufficiently robust BICM configurations are hence considered for CL PLPs in order to make up such degradation. However, not to be confused, this does not imply that using LDM penalizes the spectral efficiency compared to TDM (and FDM). Though there is an implicit throughput penalty from ILI, LDM CL can instead exploit every data cell resource within a frame whereas only a part is available to the TDM counterpart.

On the contrary, EL decoding proceeds in an ILI-free condition, thanks to successive interference cancellation (SIC)[6]. According to a typical ATSC 3.0 receiver architecture described in Fig. 5, the terminal restores CL signals first and then uses them to remove the corresponding component from the received signal prior to the EL decoding. The *effective* EL signal fed into the CL BICM decoder is thereby *refined* to have a substantially higher signal-to-interference plus noise ratio (SINR) than the ILI-included signal previously fed into CL BICM decoder. In this end, desirable reliability is ensured for ultra-high throughput transmissions over EL.

As can be noticed, LDM decoding involves some additional computations and buffering due to SIC. However, it has been verified that the entailed extra latency, memory and complexity lie in an acceptable, may be even negligible, range. Thanks to a careful design of ATSC 3.0 that allows CL and EL to share a common signal structure (e.g., TI and other OFDM/reference signal organizations), the additional burdens were practically demonstrated to be insignificant via official plug-in tests supervised by ATSC.

Depending on implementations, the receiver complexity with respect to SIC can be further reduced more. To this end, [27] and [8] have introduced two alternative SIC implementations that do not regenerate the CL codeword down from the raw data level. As described in [27], the low-complexity receiver can be implemented by directly extracting partially decoded codeword from the CL BICM decoder for canceling out the CL signal. Compared to the base strategy that uses the end product of CL BICM decoder, those alternative schemes omit decoding and re-encoding processes for FEC, and then significantly reduce the computation and latency. Such procedure-drops may bring the possibility of cross-layer interference (CLI) from CL to EL, i.e., the remainders after SIC. However, EL decoding is

[5] Note that LDM does not reside in the same place with TDM and FDM even in a transmit chain perspective. TDM and FDM are carried out as a part of framing block, whereas LDM occurs earlier.

[6] A superposition coding generally operates in a pair with SIC at the receiver-side.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

typically intended in a stationary environment with high signal-to-noise ratio (SNR), since the ModCods would be configured to have a sufficient SNR headroom between the CL and EL decoding. Therefore, quasi-error-free demapping of CL PLPs would probably be achieved, so that would get rid of the CLI problem in practice.

In order to assure a complete success of CL decoding as a premise to EL decoding instances, [8] recommends setting the injection level to be at least 3 dB greater than the non-LDM case threshold-of-visibility (ToV) of the least robust CL PLP. This additional headroom of 3 dB is heuristically determined in a rule-of-thumb sense, particularly to cover possible implementation losses.

### *B. Transmitter Identification (TxID)*

Under a critical spectrum scarcity these days, DTT over an SFN has turned out significantly attractive due to its minimal spectrum usage. A coordinated transmission of SFN makes it possible to serve even a nation-wide territory with a single RF channel, but requires cautious deployment and orchestration among the transmitters.

For better accuracy of the SFN coordination, ATSC has introduced an idea to radiograph the impulse response from each transmitter, so that makes each transmitter's channel component distinguishable from the others. This idea, referred to as TxID, is achieved by inserting additive radio watermark signals (so-called TxID signals) into the transmit signals [43]-[52]. A code orthogonality principle of code-division multiplexing (CDM) underlies this feature. When the TxID signals are uniquely assigned to transmitters, it becomes able to figure out the contribution (and interference as well) of each transmitter's emission to the received signal at once. It is an impressive benefit that this does not require any hassle of isolating the signal from each transmitter, i.e., turning the others off. Worthwhile to mention in addition, TxID signals are transparent to the legacy ATSC 3.0 receiver. Truly being a watermark, the TxID signals are seen as background noise for decoding the message-bearing frames.

ATSC 3.0 uses a set of binary phase-shift keying (BPSK)-modulated Gold code sequences for TxID [48]. The length of a single sequence block therein is 8191 (= $2^{13} - 1$) samples, while the TxID signal in a single frame can consist of multiple repetition of such a unit sequence. Let us henceforth denote those unit sequences as *TxID sequences*. Based on the CDM capability of TxID sequences, up to 8192 (= $2^{13}$) different transmitters can be uniquely indicated at the same time.

After being BPSK-modulated, the TxID signal is arithmetically combined with the baseband waveform of the 1st preamble symbol, sample-by-sample. To avoid possible data contamination that TxID signal injection can bring out, ATSC 3.0 restricted the TxID signals to be injected only in the preamble period, whose FEC protection is typically given stronger than payloads. Even so, a broadcaster should pay particularly careful attention to determining the TxID injection level (i.e., the relative power level of TxID signal compared to the preamble transmission) to balance well between the performances of service acquisition (i.e., preamble decoding) and TxID signal detection.

As depicted in Fig. 6, the beginning of every TxID signal is precisely aligned with the 1st sample part of the preamble

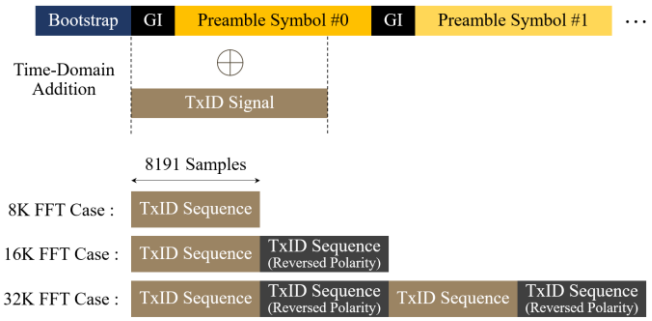



Fig. 6. TxID signal injection for various FFT sizes

interval (guard interval included). Once mentioned above, an 8191 samples-length TxID sequence can be repeated multiple times, thereby allowing the TxID signal to be better detectable. While then, the TxID signal should coexist only with the 1st preamble symbol. The number of repetitions is therefore determined by the size of fast Fourier transform (FFT) applied to preamble signals.

In case an 8K FFT is applied to preambles, for example, a TxID sequence is injected once per frame. Note herein that the duration of the 1st 8K FFT-modulated preamble symbol can contain up to one single TxID sequence. On the other hand, the 1st preamble symbol's duration is doubled if 16K FFT is used. In this 16K FFT case, the TxID sequence is duplicated into two sequence blocks and then consecutively transmitted. The 2nd sequence block therein is arranged to have the opposite polarity to the 1st TxID sequence block, specifically to remove out the direct current component. As for 32K FFT configurations, a TxID sequence block is repeated four times. Likewise, the 2nd and 4th TxID sequence blocks are modulated to have the opposite polarity to that of the 1st and 3rd blocks, while the 3rd block is just a repetition of the 1st block. Such topologies are briefly visualized in Fig. 6.

The TxID signals can be detected by means of auto- and cross-correlations, and tend to be detected better when they are injected with higher power level [49]-[52]. However, excessively high TxID injection levels could rather disturb decoding preamble signals, and therefore should be avoided. Even more, a regulation issue restricts the TxID injection level as well. Since transmitters do not normalize the power of baseband signals after combining TxID signals, high-powered TxID signals could increase the net transmit power much. In effect, the transmit power is in practice strictly regulated by governments. Generally speaking, it is not allowed to go beyond the granted transmit power by more than its 5%. The TxID injection level should hence be lower than -15 dB. In this regard, TxID injection levels from -45 dB to -15 dB are recommended to comply with the regulation.

In fact, TxID signals may not be detected that clearly for some instances in the real field. For such tough circumstances, the receiver can consider averaging the received TxID signals over a plurality of physical-layer frames. This method, so-called ensemble averaging, can reduce the effective noise seen in the received TxID signals, especially when non-TxID signal components are uncorrelated across frames. Back in the days with ATSC 1.0 systems, data frames (excluding pilot tones) were barely correlated with each other, and ensemble averaging could thereby achieve its maximum ability [43]-[47]. However, on the contrary, the preamble blocks of ATSC

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

3.0 are pretty much correlated among frames, specifically for pilot elements scattered therein. Precisely, according to [51], the 1st preamble symbol has turned out to have 52.58% ~ 76.65% randomness. For this reason, ensemble averaging for ATSC 3.0 cannot improve the visibility of TxID signals by more than 4 dB, whereas it could do significantly better in ATSC 1.0 systems.

To overcome such a limited ability for detecting the ATSC 3.0 TxID signal, [49]-[52] brought an idea of interference canceling into TxID detection. The idea is to remove preamble signal components from the received signal before feeding it into the correlation procedure. In particular, [51] proposed three types of detection methods: Pilot cancellation; whole preamble cancellation (WPC) with full LDPC decoding; and hard decision-based WPC. Those methodological branches arose from the tradeoff between detection performance and computational complexity. For example, the pilot cancellation method requires the least computation but is feasible only when ensemble averaging is executed over many frames. On the contrary, the WPC based on thorough decoding enhances TxID detection the best. Still, this exhaustive method involves much more computation which LDPC decoding brings with. The last option came out from this context, as a simplified version of WPC. This simplified technique regenerates preamble cancellation signals back from hard decision demapping over the preamble cells, which are modulated by quadrature phase-shift keying (QPSK), instead of performing complete decoding for FEC codewords. This simplified WPC lies between the previous two methods in terms of the performance-complexity tradeoff, providing acceptable detection performance and reasonable complexity.

Did not remain only as a concept, the aid of preamble cancellation has also been implemented in real hardware TxID detectors [51]. According to the laboratory and field test results in [51], WPC could assist the TxID detector to detect the TxID signals injected even 15 dB lower than what the conventional, correlation-only, method can detect. This enhanced detection capability allows the ATSC 3.0 broadcasters to reduce the injection power of TxID signals, so that relieves interference for decoding preambles. The network-scale gain of preamble cancellation was also theoretically proved in [50].

To be emphasized, the field experiments were done in the actual commercial network as well. Note again that the public ATSC 3.0 services in South Korea are currently on-air. The SFNs for those services include abundant transmitters in them, which are deployed over wide geographical areas. To manage such a complex network topology, TxID is often employed in the major broadcaster's on-air transmission in reality. The study [53], for example, presented the field measurement results that were drawn from a public ATSC 3.0 SFN in the Seoul metropolitan area. There, in [53], some snapshots of TxID-analyzed channel profiles were provided to show the real environment of dense SFN in the Seoul metropolitan.

### *C. MIMO and Channel Bonding (CB) for Higher Data Rate Transmission*

In another vein, there also has been an effort to extend the pipe-width of ATSC 3.0. Although multiple videos with 4K

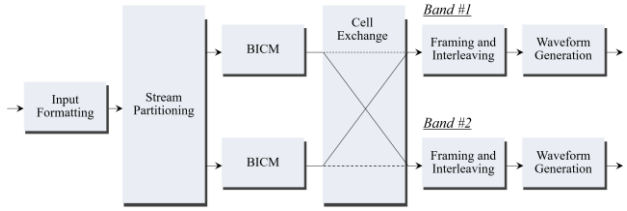



Fig. 7. ATSC 3.0 CB transmitter structure.

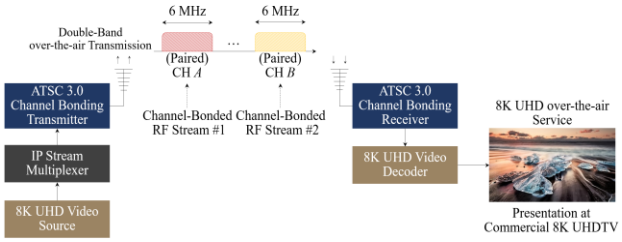



Fig. 8. Block diagram of ATSC 3.0 CB system.

Fig. 9. 8K UHD system based on ATSC 3.0 CB.

resolution were already viable in baseline system configurations, ATSC 3.0 additionally defined two optional technologies for increased capacity: MIMO and CB [20], [21]. The intent behind those optional systems was to support future plans beyond 4K UHD, such as 8K videos and virtual reality, by doubling the pipelines. In specific, ATSC 3.0 MIMO transmissions exploit two polarization channels by employing two antennas each for the transmitter and receiver. On the other hand, CB technology utilizes two RF channels for increasing system capacity.

As for the CB in ATSC 3.0, two types of CB transmission are defined: A plain CB and a CB with SNR averaging. The plain CB transmits the data streams separately for each RF channel by configuring them independently, instead of applying the same system parameter set. Otherwise, if SNR averaging is applied to CB, data cells are exchanged between two paired radio channels to obtain frequency diversity. However, transmissions over both radio channels shall share the same system parameters in this case [7].

Fig. 7 describes the procedure of ATSC 3.0 CB. The stream partitioning block first splits the input stream into two streams, and each stream is delivered to a different BICM block. The cell exchange block is enabled if SNR averaging is used, but remains disabled for the plain CB. If enabled, the cell exchange block selectively exchanges the even-index data cells between the data stream paths. That is, the cell exchange is executed once for every two data cells, and hence the odd-index cells remain in the same path. This cell exchange

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

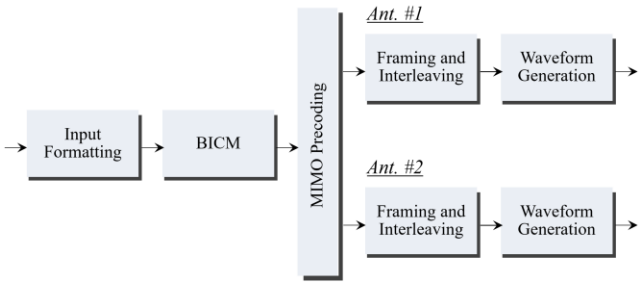



Fig. 10. ATSC 3.0 MIMO transmitter structure

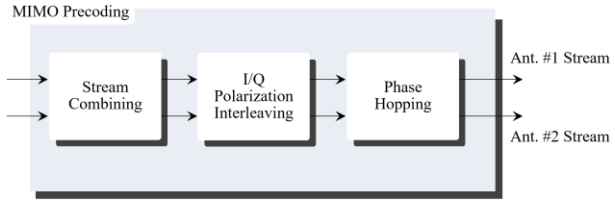



Fig. 11. ATSC 3.0 MIMO precoding structure

between separated channels disperses the frequency-selective fading effect, and as a result, acquires frequency diversity in an overall sense [54], [55]. To be noted, every CB transmission in ATSC 3.0 does not require the intended RF channels to be adjacent to each other.

An end-to-end system for ATSC 3.0 CB was implemented in reality as well, particularly to demonstrate a delivery of 8K UHD or multiple 4K UHD services over CB [56]. Fig. 8 briefly describes the 8K-over-CB system presented in [56]. An 8K UHD video, encoded by High Efficiency Video Codec (HEVC), was first IP-encapsulated and then fed into the ATSC 3.0 CB transmitter. The CB transmit signals, which went through the dedicated modulation chains shown in Fig. 7, were transmitted over two different wireless bands. Again, those RF bands could be either contiguous or non-contiguous. On the counterpart, the ATSC 3.0 CB receiver demodulated the received dual-channel signals to produce a combined output. The 8K video decoder then retrieved the video source from this output, and finally presented 8K imageries at the commercial 8K UHD television (UHDTV) display.

The implementation in [56] was in practice found to serve 8K UHD program reliably over two 6 MHz wireless channels. Fig. 9 shows an instance of laboratory trials. Particularly, transmissions of an 80 Mbps bit-rate video was demonstrated in those trials. At the time of this writing, the implemented CB system was examined to be capable of transceiving a 110 Mbps video seamlessly in laboratory environments.

In another vein, ATSC 3.0 also supports a cross-polarized MIMO technology for increased capacity [57]. This technique utilizes only one single RF channel band but requires a pair of horizontally and vertically polarized antennas to be installed at each transmitter and receiver [58]. Fig. 10 depicts the transmitter architecture of ATSC 3.0 MIMO. It reuses most of the system blocks from the single-antenna operation, but MIMO mapping (MAP)[7] and MIMO precoding blocks are newly introduced.

ATSC 3.0 MIMO carries out three types of precoding for transmission: Stream combining, in-phase/quadrature-phase

[7] The MIMO MAP block consists of MIMO demultiplexing and constellation mapping sub-blocks, and is included in the BICM block.

Fig. 12. ATSC 3.0 MIMO field experiments – Facilities and reception status.

(I/Q) polarization interleaving, and phase hopping (see Fig.11.). Stream combining reprojects a pair of the demultiplexed constellation symbols by applying a rotation matrix. The rotation angle is determined according to ModCod configuration. I/Q polarization interleaving then switches the imaginary part of the symbol between two data paths. In subsequence, the phase-hopping rotates the phase of the second data path's symbol. By means of those precoding procedures, MIMO transmissions can be better protected from imbalanced fading in an interleaving manner.

An end-to-end MIMO system fully compliant with ATSC 3.0 physical-layer standard has been implemented as well. The real environment performance of the implemented system has been examined by field testing, as shown in Fig. 12. Seamless delivery of 110 Mbps data has been confirmed in the laboratory, and 105 Mbps delivery has been exhibited through field trials.

## III. Extensibility of ATSC 3.0

One prominent feature that allows ATSC 3.0 to be distinct from previous standards is extensibility. This extensibility mainly owes to the overall IP-based design of ATSC 3.0, so that enables efficient cooperation with other non-ATSC 3.0 networks, such as a broadband cellular. This section summarizes recent evolutions on ATSC 3.0 extensibility, particularly focusing on the convergence with up-to-date broadband networks.

### *A. Convergence with Broadband*

A part of the underlying philosophy on the design of ATSC 3.0 is to make an extensible system finding itself compatible with its future versions, or to the interplays with other networks' interfaces. A definition of bootstrap versioning and all-IP-based system design is the result of such principle. This consideration of extensibility lies in the same vein to the global trend in network designs: To exemplify thereof, the 5$^{th}$ generation wireless standard (5G) can first be referred, which pursues establishing a macro topology to interact with various heterogeneous (including non-5G) networks. Since the user environments and service purposes are coming highly diverged, such an inter-network collaboration is foreseen to have an influential edge to optimize deliveries and broaden the system capability.

The recent interest of the media and entertainment (M&E) industry has been particularly on the conjunction between ATSC 3.0 and broadband. Such an idea, popularly called as a *broadcast-broadband convergence (BcBbC)*, has been seen as an opportunity for DTT to grab user interactivity at hand [16]. As one may come up with the first, this interactivity can

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

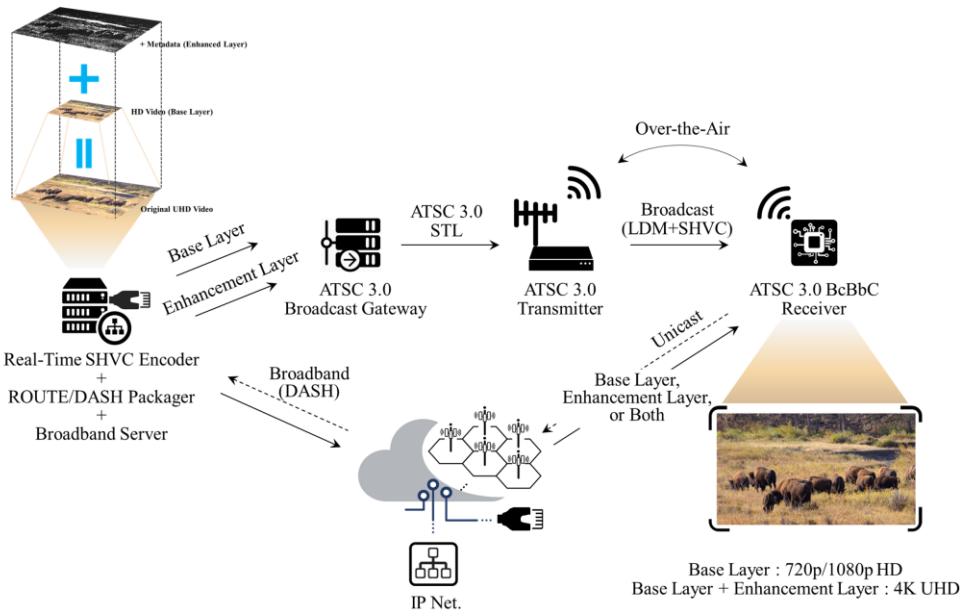



Fig. 13. Equipment configuration for the ATSC 3.0 BcBbC use case proposed in [18].

be embodied in presentation- and application-layers to deliver personalized contents, e.g., target advertisement, on-demand supplementary data service [59]. In fact, the figures of BcBbC would include but not limited to the upper layer data bonding. Diverse forms of the convergence are currently being discussed in every system layer, owing to the extensibility of ATSC 3.0. As itemized in [16], BcBbC embodiments are specifically envisioned to offer enhanced media quality, data traffic stability, mobile media continuity, personalized service, cost efficiency and coverage extension based on tower overlaying. Those features are expected to play versatile roles in a wide range of verticals, more than remaining in the M&E industry. In this end, the future ecosystem could be assisted by the BcBbC evolution in many aspects, such as intelligent traffic system (ITS), mobile positioning, and so on.

One rising example of BcBbC is seamless mobile streaming based on dual connectivity. As aforementioned in Section I and II, ATSC 3.0 is strongly anticipated to bring mobile DTT service of rich, over-HD quality videos into the picture. This mobile-supporting feature is seen further promising when we remind the drastic advance of automotive technology. However, there undesirable shadowing and coverage gaps, e.g., a tunnel, would still exist and hence may lay some temporal stream loss. If more than 4K UHD quality is intended, such connection losses will be more significant because more than 10 dB SNR is typically required to receive over-15 Mbps ATSC 3.0 stream successfully. To resolve this problem, [18] and [19] proposed an idea to use a broadband link as a backup channel.

In the proposal [18], the convergence has been accomplished in transport- and application-layers. In parallel to the broadcast-side transmission, the studio was designed to upload timely segments of the same (live-encoded) video program streams to the on-line server also[8]. On the other hand, the video decoder application installed to an ATSC 3.0

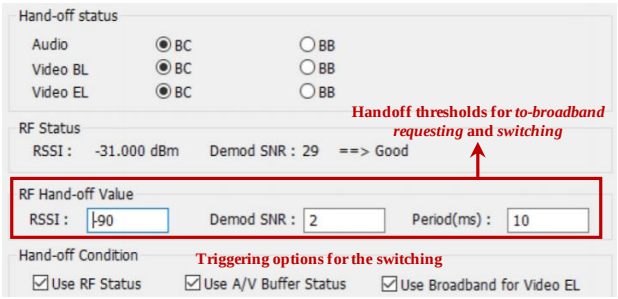



Fig. 14. A controller software to monitor the RF signal status and determine the switching between broadcast and broadband connections [18].

receiver was implemented to request an apt video segment whenever the receiver is about to lose ATSC 3.0 signal.

To this end, the receiver keeps monitoring the statuses of the physical channel (i.e., SNR and received signal strength) and video buffer. Should be noted, the requesting physically lags to the reception of the broadcast signal and there also exists a negligible round-trip delay in the broadband delivery. The broadband requests were therefore executed in somewhat a conservative manner, *prior* to the expected broadcast signal loss. That is, the threshold SNR which triggers the request was set to be reasonably higher than the ToV of the broadcast signal.

The video segments received via broadband linkage were registered into the buffer, and then the receiver became able to keep the playback stable even while the ATSC 3.0 signal is blocked out. Conversely, the receiver was led to switch back into the broadcast link seamlessly (i.e., to idle the requesting) whenever ATSC 3.0 physical-layer connectivity is recovered. Note here that the handoffs to broadband were executed opportunistically and temporally, which therefore eased the users' financial burden for using paid networks. Keeping in mind the difference of end-to-end latency between broadcast and broadband deliveries, it was requested the packets whose presentation times precede to that of the

[8] At the IP-multiplexing component, encoded video stream was duplicated and then in parallel encapsulated into two different formats at the same time. UDP/IP over Dynamic Adaptive Streaming over HTTP (DASH) was used for the broadband path, and TCP/IP over MPEG Media Transport (MMT) or Real-time Object delivery over Unidirectional Transport (ROUTE) was applied to the copy fed into the ATSC 3.0 physical-layer chain (See Fig. 13).

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2


(a) (b) (c)

Fig. 15. Field trial instants in [18]: (a) Seamless switching between HD and UHD in broadcast-only mode, (b) hybrid service combining base layer from broadcast and enhancement layer from broadband, and (c) seamless switching between broadcast and broadband networks for HD only.

via-broadcast packet received at the *triggered* instant. To this end, the presentation time offset was heuristically determined by experiments.

Notably, the idea for BcBbC-assisted seamless mobile has been further extended decently, by being combined with scalable video coding (SVC). Precisely speaking, the implementation in [18] aimed at providing a scalable service based on H.265 Scalable High-Efficiency Video Codec (SHVC)[9]. A 720p HD service was encoded to be the SHVC *base layer*, and the SHVC *enhancement layer* consisted of metadata that supplement the base layer stream to elevate the video quality into 4K UHD[10].

This figure of plugging SHVC into the proposed BcBbC system allowed even a 4K UHD content to be seamlessly streamed in mobile environments. Where the base and enhancement layers were loaded into different PLPs, the receiver was able to combine the packets from both broadcast and broadband channels depending on the broadcast signal availability. For example, as shown in Fig. 15(b), base layer-over-broadcast and enhancement layer-over-broadband packets could be combined to rebuild 4K UHD when the broadcast channel quality was good enough to decode the robust PLP containing base layer but not sufficient to decode the enhancement layer's PLP. Worth to mention, the user can choose whether to pursue maintaining the enhanced-quality video or to remain at base-quality service, depending on its affordability.

To expand the coverage more, this system was built on top of a cross-layer combination of SHVC+LDM [61]. The base and enhancement layers were there conveyed by CL and EL PLPs, respectively. The layered, hierarchical architectures of those two gears in essence fit into each other, and their engagement thereby lubricates a switching between base- and enhanced-quality services. In addition, a bit-reduction gain from SHVC and a spectral efficiency gain from LDM synergistically improve the system coverage of broadcast itself. This enhancement fortunately reduces the possibility to handover both base and enhancement layers entirely to broadband, hence bringing better economic efficiency to users.

A real-field practice of this convergence was demonstrated in Jeju Island, South Korea, by using a public LTE-A Pro network for the broadband part. Through the field-test carried out in 2019, it was well-verified that BcBbC guarantees a stable experience to on-vehicle users. In addition, the reliability gain of this BcBbC concept has been as well theoretically proved in [19].

### *B. Traffic Stability and NOMA in Convergence Perspective*

Although 5G enhanced mobile broadband (eMBB) has raised the net throughput capability impressively, rapidly increasing demands on massive video data are still a nuisance to cellular networks. DTT networks have again been explored in this context, to offload the traffic upon its super-efficient point-to-multi-point (PTM) capability. Several groups of manufacturers are already in progress to develop the broadcast-offloading solutions, particularly concentrating on medium access control (MAC) layer functionalities.

Beyond the idea of traffic offloading, the convergence further down to the physical-layer level has been introduced as well [16], [17], [42], [62], [63]. The idea of physical-layer level cooperation emerged from an expectation that it would create new market opportunities. Specifically, broadcasters are envisioning an infra-leasing business based on tower overlay. At the broadband carrier's perspective, routing high-power high-towers (HPHTs) of the DTT network is an efficient option to reduce operating and capital expenditures required for widening its service coverage. If this physical-layer convergence is achieved above the advance of cloud-based network orchestration technologies, the BcBbC could

[9] Note that this development was fully compliant to ATSC 3.0 standard. The video specifications of ATSC 3.0 have adopted SHVC as well as HEVC.

[10] Following the terminologies in SHVC, a base layer is defined to be the smallest subset of a video content representing the lowest quality product. An enhancement layer (multiple enhancement layers can exist in principle, but up to one enhancement layer is allowed in ATSC 3.0) on the other hand indicates the larger, metadata subset which references the base layer. SHVC is designed in a hierarchical manner so that a better-quality video can be retrieved by combining base and enhancement layers during decoding.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

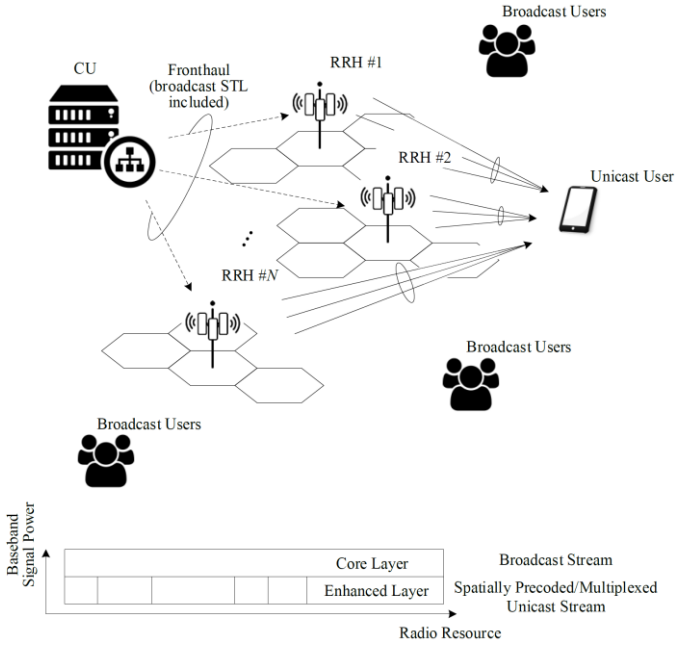



Fig. 16. Conceptual description in [42] for the C-RAN with BMUST.

come with an organically integrated network that can coordinate HPHTs and low-power low-towers (LPLTs) at every moment.

When we turn back to the fundamentals, BcBbC can be viewed as cooperation between broadcast and unicast transmissions. From this end, some pioneer works have analyzed the theoretical gain which can be obtained by a conjunction between broadcasting and unicasting. Those studies have illustrated abstract figures which can pave the way toward the convergence below Layer-3.

In particular, [42], [62]-[66] have focused on the usage of NOM for co-transmitting broadcast and unicast messages jointly. Let us henceforth refer to this class of NOM/NOMA as *broadcast/multicast and unicast superposed transmission (BMUST)*, as in [63]. Those works have pointed out the possible gain from BMUST amplifying the benefit of BcBbC, by underlining the throughput advantage from NOM. In addition, BMUST has a desirable edge at scheduling simplicity, because BMUST does not require a user pairing unlike MUST solely for unicasting.

The approach of [62] would especially be interesting to service and network providers. To take the offloading gain into account, [62] has brought the popularity of each content into the picture. Accordingly, the selection of the content set that is transmitted over broadcast channel, namely a *broadcast content profile (BCP)*, has been optimized to maximize the stochastic throughput (i.e., successful delivery rate) of the network. Alongside this, the BMUST frame structure has been proposed to convey broadcast and unicast signals by CL and EL pipes, respectively. Two different MAC policies have then been considered: (1) A *generous MAC* that allows transmitting the broadcast contents via unicast; and (2) an *offloading-oriented MAC* that restricts the system not to use unicast transmissions for delivering the broadcast contents.

In turn, closed-form formulas for the optimal BCP selection strategy have been analytically derived in a probabilistic form. The optimal strategy has been found to be dependent on network density, network access load, bit-rate and popularity of each content, and so on. This optimal BCP selection has been shown to be definitely beneficial than the *maximum popularity selection* and *uniform probability selection* strategies regardless of the popularity distribution. Numerical results in [62] have verified a significant offloading gain of BcBbC, and BMUST has also been shown to achieve over 170% of network throughput gain over orthogonal multiplexing-based BcBbC.

On the other hand, [42] has delved into the radio transmission chain more. Unlike the stochastic geometry framework of [62], [42] has considered a deterministic system topology. Instead, the coordinated multi-point (CoMP) signal processing has been addressed to facilitate BMUST within cloud-radio access networks (C-RANs). As also been pointed out in [16], the centralized natures of broadcasting and C-RAN have been highlighted to be fit well with each other. The effect of broadcast signal coexistence has been carefully investigated for both central unit-to-frontend (i.e., fronthaul) and frontend-to-user connections, and reflected to the compressions over fronthauls and the spatial multiplexing over physical channels. Precisely, in [42], precoding and fronthaul compression for each remote frontend have been jointly optimized to maximize the sum-rate of unicast transmissions.

Note that every BMUST studies above have considered the CLI penalty that turns out when the CL decoding is failed. Impressively, BMUST has been shown still beneficial than the traditional radio access technologies despite the destructive CLI effect. Owing to this delightful throughput advantage, BMUST will turn out to be a super-appealing feature in the future BcBbC era, in which intimate convergences are realized in various system layers.

In addition to applications at the frontend transmissions, using LDM for wireless fronthauls/backhauls could also help to cope with resource scarcity. Spectrally efficient LDM transportations would possibly reduce the static resource occupation of wireless backhauls, and therefore would allow the users to experience better QoS, in average [67].

### C. BICM Comparison with FeMBMS and 5G NR

Through the BICM chain, input data is protected by applying FEC and then mapped to constellation points. BICM chain, which is closely related to spectral-efficient and robust transmission, should be designed to meet the requirements of the systems.

Quasi error-free performance, e.g., block error rate (BLER) $\leq 10^{-6}$, is required for broadcasting systems, and therefore

TABLE I. BICM COMPONENTS OF ATSC 3.0, FEMBMS, AND 5G NR

| | FEC for data channel | FEC for control channel | Constellation |
|---|---|---|---|
| ATSC 3.0 | LDPC codes | S/P LDPC codes | NUC |
| FeMBMS | Turbo codes | Convolutional codes | QAM |
| 5G NR | LDPC codes | Polar codes | QAM |

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

TABLE II. PERFORMANCE OF ATSC 3.0, FEMBMS, AND 5G NR [75]

| Code Rate | ATSC ($n$=64800) | ATSC ($n$=16200) | 5G NR ($n$=16200) | FeMBMS ($k$=6144) |
|---|---|---|---|---|
| 3/15 | -4.31 | -3.80 | -3.98 | -3.48 |
| 4/15 | -2.87 | -2.37 | -2.52 | -2.16 |
| 5/15 | -1.66 | -1.33 | -1.38 | -1.36 |
| 6/15 | -0.48 | -0.30 | -0.35 | -0.28 |
| 7/15 | 0.32 | 0.56 | 0.52 | 0.63 |
| 8/15 | 1.20 | 1.36 | 1.50 | 1.48 |
| 9/15 | 2.00 | 2.19 | 2.23 | 2.35 |
| 10/15 | 2.81 | 2.96 | 3.05 | 3.23 |
| 11/15 | 3.63 | 3.81 | 3.87 | 4.05 |
| 12/15 | 4.51 | 4.70 | 4.73 | 4.93 |
| 13/15 | 5.54 | 5.75 | 5.77 | 5.98 |

LDPC codes of length 16,200 and 64,800 bits are employed in both DVB-T2/S2 and ATSC 3.0 systems. During their standardization, it is focused on optimizing the performance of LDPC codes for the combinations of fixed code lengths and rates [68]-[75]. Contrary to DTT broadcasting systems, 3GPP, which is developing a standard for wireless communications, should support the high flexibility required for adapting to the channel quality and the size of the payload [76], [77]. For this reason, the flexibility in terms of code rate and length is the critical requirement in the design of the BICM chain for 3GPP.

In 3GPP, further evolved multimedia broadcast multicast service (FeMBMS) is referred to as the most advanced technique for terrestrial broadcasting, which is based on eMBMS first standardized in Rel-9. FeMBMS is further evolved in Rel-16 in order to satisfy the 5G requirements [76], which requests large inter-site distance (ISD) up to 100 km and high mobility up to 250 km/h for SFN. As a result, a new numerology, which uses 0.37 kHz subcarrier spacing (SCS) and 300 us cyclic prefix (CP), is endorsed so that LTE-based 5G broadcast (i.e., evolved FeMBMS) can meet the requirements for 5G multimedia broadcast/multicast service (MBMS). However, LTE-based 5G broadcast is based on FeMBMS which employs turbo codes and convolutional codes as its FEC schemes, and therefore it is considered as limitations of FeMBMS from the viewpoint of physical-layer performance.

One innovative technology adopted in 5G new radio (NR) physical-layer is to employ LDPC codes and polar codes as FEC schemes for protecting data and control channels. As a result, 5G link-level performances are better than those of LTE which uses turbo codes and convolutional codes as its physical-layer FEC schemes. Future terrestrial broadcasting in 3GPP based on NR physical-layer can take advantage of not only recent enhancements for terrestrial broadcasting in Rel-16 but also the advanced NR BICM chain. Table I represents components of the BICM chain of ATSC 3.0, FeMBMS, and 5G NR.

Link-level performances of FEC schemes employed for data channels of ATSC 3.0, FeMBMS, and 5G NR are compared in Table II, which represents the required SNRs to achieve BLER =$10^{-2}$ over the additive white Gaussian noise (AWGN) channels. Since the number of information bits $k$ supported by the 5G NR LDPC codes is limited to 8448, LDPC codes are generated by lifting base graph1 (BG1) with a larger size than that defined in the spec. [77]. Table II shows that ATSC 3.0 with 64,800 bits-length LDPC codes has the best performance thanks to dedicatedly designed long LDPC codes. In addition, the performance advantage of ATSC 3.0 LDPC codes gets better compared to FeMBMS when target BLER is less than $10^{-6}$ because turbo codes suffer from error floor phenomenon. Moreover, when transmitting data with high-order modulations (HOMs), ATSC 3.0 has a strictly better performance than FeMBMS and 5G NR employing QAM constellation due to the shaping gain of the NUC [7]. The performance gain from NUC reaches up to about 1 dB when the modulation order is 256.

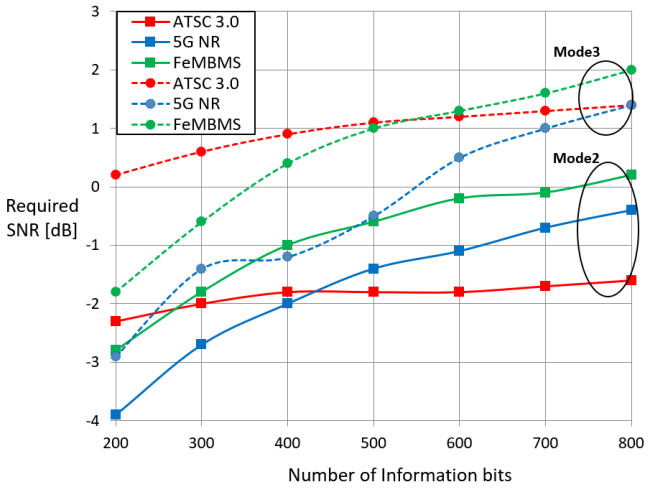



Fig. 17. Performance of BICM chains for control information: Comparisons with L1-protection mode 2 and mode 3 of ATSC 3.0.

In order to properly decode the data at the receiver, the control information should be decoded properly because it includes essential information for receiving data channels. In the case of FeMBMS, it is transmitted in a cell acquisition subframe (CAS) [78]. In the case of ATSC 3.0, it is called layer-1 (L1) signaling which is divided into L1-basic and L1-detail. These signaling information are strongly protected by FEC schemes, unlike the ones used for data channels. FeMBMS and 5G NR have adopted convolutional codes and polar codes for protecting control information, respectively. On the other hand, ATSC 3.0 makes use of the same LDPC codes adopted for data channels with information shortening and parity puncturing, which is necessary to support the required number of bits and code rate for control channels [7].

Link-level performance of FEC schemes for control channels of ATSC 3.0, FeMBMS, and 5G NR are compared in Fig. 17, which represents the required SNRs to achieve BLER =$10^{-2}$ over AWGN channels. To ease the comparison, the number of information bits is chosen from 200 to 800 and the number of codeword bits is determined to support mode 2 and mode 3 of L1-detail protection [73], [7]. Note that the code rates and net codeword lengths of FeMBMS and 5G NR frames were controlled to be the same with ATSC 3.0 frames for each comparison.

As shown in Fig. 17, the performance of ATSC 3.0 is rather uniform according to the number of information bits. When

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

the number of information bits is small, the performance of ATSC 3.0 is the worst because massive shortening and puncturing are inevitable to support the small number of signaling bits. Even though the patterns for shortening and puncturing is delicately designed for the given length-16,200 LDPC codes, it cannot reach up to the performance of polar codes and convolutional codes. As the number of information bits increase, however, the BLER performance of 5G NR and FeMBMS is much more degraded than that of ATSC 3.0 because 5G NR polar codes is designed up to length-1024 and convolutional codes cannot have benefit of long code-length. It means that the physical-layer performance of the FEC scheme for L1 signaling is comparable with that of polar codes and convolutional codes in 5G NR and FeMBMS and it can be improved by designing new dedicated LDPC codes for protecting L1 signaling.

## IV. Field Verification

A number of active research groups have examined the end-to-end system performance of ATSC 3.0 to provide a practical guideline for implementation and service [8]. Those verifications were not limited to computer simulation and laboratory test, but also encompassed the field trials in real environments. In particular, [9] and [3] have drawn exhaustive evaluations for the reliability of every ModCod combination in ATSC 3.0 physical-layer[11]. Whereas the baseline performance based on non-LDM transmissions have been addressed in [9] and [3], the throughput/reliability gain of LDM over TDM has additionally been measured and reported in [79]. Those results have been officially published in [8] to be the reference measures. This section gives a brief review of those celebrated results, particularly on field test results. The revisited results may be especially informative to potential network organizers, as those offer an intuition upon the practice of ATSC 3.0 physical-layer.

### A. *Reliability of ATSC 3.0 Physical-Layer*

The reports [9] and [8] provide an exhaustive list of the ToVs for ATSC 3.0 transmissions measured in several representative stationary channel models. Motivated by [80], the laboratory experiments were carried out over three reference channel models: An AWGN channel; a Rician channel with a single line-of-sight component and 19 multipaths (RC20); and a Rayleigh channel with 20 multipath components (RL20). To follow-up with coherency, the experiments continued in the field measured the results in carefully chosen reception points whose channel profiles were seen *AWGN-like*, *RC20-like*, or *RL20-like*.

The field experiments of [9] took place in Jeju Island, South Korea. To this end, the full-chain studio and transmission facilities were constructed for ATSC 3.0 broadcasting (See Fig. 18). For the receiver-side, a professional testing vehicle was built as shown in Fig. 19. A specialized measurement system was implemented therein, so as to integrate professional ATSC 3.0 physical-layer receiver, spectrum analyzer, noise generator, low noise amplifier, antenna elevator and rotor, media decoder, Global Positioning System (GPS) receiver, and signal attenuator

Fig. 18. Facility configurations: Studio/transmitter-side infrastructure [9].

Fig. 19. Facility configurations: Customized vehicle for measurements [9].

under a unified software controller so-called *the Integrated Measurement Analysis System (IMAS)*. Using this measurement system, decoding errors were recoded bit-by-bit and frame-by-frame at every instant, together with the information of the overall reception environment profile.

The signal robustness of each ModCod combination was quantified by its ToV, the SINR at least required to assure QEF. Complying with ATSC's recommendations for fixed service, [9] measured the ToVs at the SNR points where the frame error rate (FER) equals to $10^{-4}$ in approximate. Through the presented results, [9] demonstrated that a variety of transmission capacities in ATSC 3.0 are feasible in the real world. Every BICM configuration of ATSC 3.0 stably delivered the guaranteed throughput if sufficient signal strength was acquired. For example, an HD video with about 1.1 Mbps could be successfully received even at -3 dB SNR environment. While the coverage was correspondingly reduced, over 50 Mbps transmissions were successfully verified in the field also. As for the promising use case, 4K UHD, around 15 Mbps delivery was available at 12.8 dB ~ 14.5 dB SNR, depending on the multipath profile.

By and large, the ToV was shown in [9] to be increased by LDPC code rate and constellation order. However, an interesting property was found at the same time: Combining the high-rate LDPC codes with the low-order constellation, e.g., QPSK, sometimes could not protect the signal so well as the other ModCod options which achieve the analogous throughput. This aspect was once discussed during the standardization, and in this regard, the physical-layer specification [7] of ATSC 3.0 has not included the combinations of high-rate LDPC codes and low-order constellations in the mandatory ModCod set.

Such performance degradation turned out significant especially in fading channel environments, as shown in Fig. 20 [9]. Can be noticed from the result in Fig. 20, the

[11] The evaluations have been carried out under a 16K-FFT configuration.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

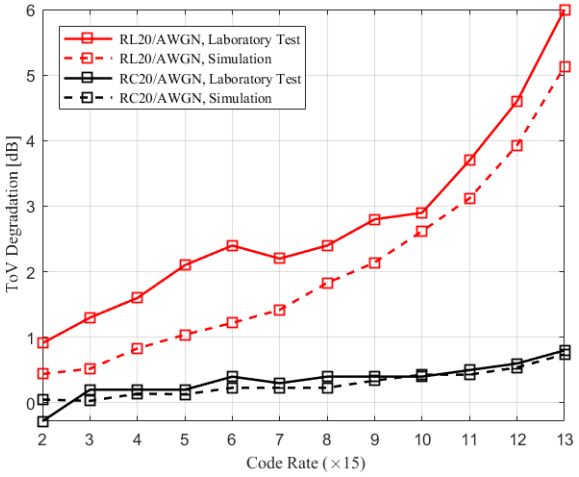



Fig. 20. Fading channel impact on ToV with respect to LDPC code rates (QPSK-modulated) [9]

degradation was found more significant in RL20 channel case than that of RC20 case. Interpreting this result into a relative vulnerability against harsh channel environments, it may be incurred that the combinations of high-rate LDPC+low-order constellation would not be adequate for mobile channels although they could show a good performance in AWGN or mild multipath channels.

### *B. Field Comparison between LDM and TDM in ATSC 3.0*

#### *1) Fixed Reception Results*

Based on the decent facility construction in Jeju Island, the field feasibility of ATSC 3.0 LDM was also validated. The report [79], [81] demonstrated that LDM truly achieves substantial coverage and capacity gains over TDM in practice, as was touted widely. A reasonable UEP scenario was considered to serve two video contents with different bit-rates. A 720p HD video and a 4K UHD video, both HEVC-coded, were broadcasted over-the-air via the CH 50 (centered at 689 MHz) radio channel. Both video services were contained in every physical-layer frame, where each video service was loaded into its dedicated PLP. The detailed physical-layer configurations are described in Table III.

Table IV recalls the field comparison results presented in [81][12]. To be noted, ATSC 3.0 LDM assured better reliability than TDM for both HD and UHD services simultaneously, where the throughput of each PLP was set analogous between the LDM and TDM cases. LDM precisely attained about 3.3 dB ToV gain over TDM for the HD service and reduced nearly 2.2 dB of ToV for the UHD service.

Such superiority on error protection capability was reflected to a coverage extension in quality. As for the rooftop receptions conducted across the 40 selected test points in Jeju City, LDM CL reception was succeeded at every test point while 5% of the test points failed at retrieving the HD service in TDM frames. In addition, the UHD video in LDM EL was available at 72.5% test points while the TDM counterpart showed a 70% success rate.

Indoor environments were shown more challenging, whereby the indoor RSS (averaged over 20 test sites) was

[12] At each measurement point, the ToV was measured by injecting artificial AWGN until the reception cannot maintain an errorless video playback over 60 seconds.

TABLE III. ATSC 3.0 PHYSICAL-LAYER PARAMETERS USED IN LDM/TDM FIELD COMPARISON TESTS OF [81] AND [82]

| | | LDM | TDM |
|---|---|---|---|
| Frame Length | | 251.33 ms | |
| Occupied Bandwidth | | 5.832844 MHz | |
| Time Interleaver | | CTI, 1024-sample depth | |
| Frequency Interleaver | | ON | |
| HD Service | | CL | 1st Subframe |
| | FFT Size | 16k | 8k |
| | Guard Interval | 148.15 us (GI5_1024) | |
| | Pilot Pattern | SP Dx=6, Dy=2 | |
| | # of Payload Symbols | 98 | 75 (38% of each frame) |
| | Outer Code | BCH | |
| | Inner Code | 4/15 LDPC (64K Length) | 11/15 LDPC (64K Length) |
| | Constellation | QPSK | |
| | Data Rate | 2.59 Mbps | 2.75 Mbps |
| UHD Service | | EL | 2nd Subframe |
| | FFT Size | 16k | 32k |
| | Guard Interval | 148.15 us (GI5_1024) | |
| | # of Payload Symbols | 98 | 30 (62% of each frame) |
| | Inner Code | 10/15 LDPC (64K Length) | 12/15 LDPC (64K Length) |
| | Outer Code | BCH | |
| | Constellation | 64 NUC | 256 NUC |
| | Data Rate | 19.56 Mbps | 19.90 Mbps |
| Injection Level | | -4 dB | - |

TABLE IV. LDM/TDM FIELD COMPARISON RESULTS PRESENTED IN [81] (FIXED RECEPTION, AVG. SIGNAL STRENGTH = -63.53 dBm)

| | | Throughput [Mbps] | Measured ToV [dB] |
|---|---|---|---|
| HD Service (720p) | LDM CL | 2.59 | 2.29 |
| | TDM 1st Subframe | 2.75 | 5.77 |
| UHD Service (4K) | LDM EL | 19.56 | 20.61 |
| | TDM 2nd Subframe | 19.90 | 22.91 |

nearly 19 dB lower than the average RSS measured at rooftop reception sites. The average RSS of the indoor case was -82.33 dBm and that of the rooftop case was -63.53 dBm. As for the indoor experiments, LDM brought 15%p more places into the HD coverage, compared to the TDM counterpart. Due to a significant signal strength loss from penetration, only 10% of the test sites were available of the UHD PLP, for both LDM and TDM cases.

#### *2) Mobile Reception Results*

As once mentioned above, one of the most featured use cases over ATSC 3.0 LDM is the coexistence of mobile and fixed services within the same physical-layer frame. The HD service configurations in Table III were designed to be a versatile carrier that could serve mobile recipients also, in addition to indoor and coverage-edge users or the *global service* recipients residing between the coverages of *locally inserted* services [30], [39].

Aside from the comparison for fixed receptions, the mobile reliability was compared in the same project [81] (and the related works [82], [79]). Results exhibited pragmatic evidence whereby touted an excellent mobile coverage extension opportunity of using LDM.

To summarize the experiment of [82], an ATSC 3.0 measurement system mounted in the test vehicle recorded every instant decoding errors that arose during each drive across the downtown of Jeju City. The configurations in

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

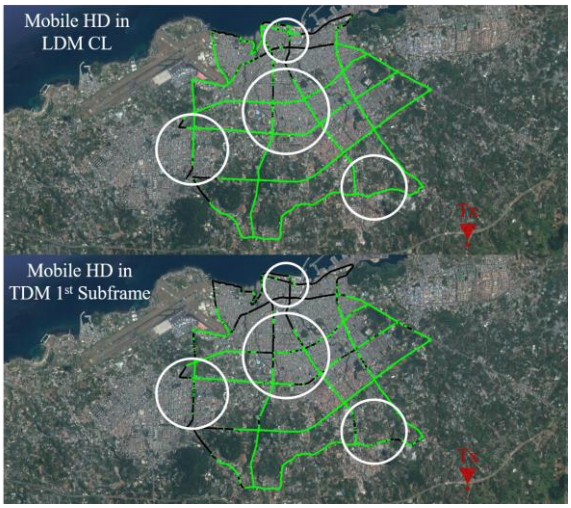



Fig. 21. Reception performance comparison between the mobile services over LDM and TDM: Reception success (green), reception failure (black). LDM case (upper), TDM case (lower) [82].

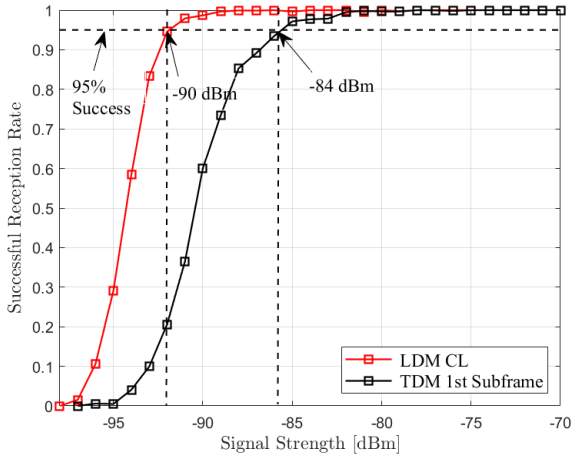



Fig. 22. Field comparison test result: ESR curves for ATSC 3.0 mobile LDM and TDM services [82].

Table III were reused for those experiments. At every second, physical-layer frame errors were recorded in pairs with the location (GPS) and RSS information at that instant[13]. In turn, every reception success or failure could be marked on a map as presented in Fig. 21.

As shown in Fig. 21 clearly, applying LDM evidently reduced the reception failures, compared to TDM. Such an improvement was especially notable in the areas pointed by white circles. It was shown that most of the points of interest were brought into the coverage by applying LDM, despite the challenging penalty from mobility. This advantage was also well evaluated in terms of the successful reception rate. To this end, the erroneous second ratio (ESR)-5 criterion was used, which is recommended by the International Telecommunication Union-Radiocommunication Sector (ITU-R) guideline. Accordingly, Fig. 22 plots the ESR curve that rearranged the results of [82]. According to Fig. 22, the usage of LDM could reduce the RSS requirement by 6 dB, subject to 95% success of HD service playback during the driving time.

As well, there were many other efforts to verify the mobile feasibility of ATSC 3.0 and the gain of LDM. For instance, [3] thoroughly presented the mobile performance of every ATSC 3.0 ModCod recommended in [8], which was measured by computer simulations (laboratory tests were also included partly) under a 6-tap typical urban (TU-6) channel. In the same paper, comprehensive discussions on the effect of various system components, e.g., pilot boosting, LDM transmission, time-domain windowing, and SFN operation, were carried out for ATSC 3.0 mobile service. [83] figured out the effect of transmit diversity code filter set (TDCFS) filtering into an SFN-based mobile service over LDM, through field tests in Jeju Island. As for a theoretical approach, [84] proved that LDM could cope better with receiver mobility than TDM when multiple services are intended in a single RF channel.

[13] Every trial was carefully but heuristically controlled to keep the vehicle speed and environmental conditions at each location to be the same between trials as possible.

## V. Conclusion

This paper introduced several selected features of ATSC 3.0 that could play a prominent role in the near future. This investigation encompassed the optional technologies defined in ATSC 3.0 physical-layer which can improve single frequency network integrity, enable additional features in other domains, or further widen the throughput capability. Majorly focused on the extensibility of ATSC 3.0, we also elaborated on possible inter-network cooperation with broadband and cellular systems.

## References

[1] L. Fay *et al.*, “An overview of the ATSC 3.0 physical layer specification,” *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 159-171, Mar. 2016.
[2] R. Chernock *et al.*, “ATSC 3.0 next generation digital TV standard - An overview and preview of the issue,” *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 154-158, Mar. 2016.
[3] S. Ahn *et al.*, “Mobile performance evaluation for ATSC 3.0 physical layer modulation and code combinations under TU-6 channel,” accepted for publication in *IEEE Trans. Broadcast.*
[4] D. Gómez-Barquero and M. W. Caldwell, “Television spectrum incentive auctions in the U.S.: Trends, challenges, and opportunities,” *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 50-56, Jul. 2015.
[5] L. Michael and D. Gómez-Barquero, “Bit-interleaved coded modulation (BICM) for ATSC 3.0,” *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 181-188, Mar. 2016.
[6] Advanced Television Systems Committee, *ATSC Standard: A/321, System Discovery and Signaling*, Doc. A/321, Mar. 2016.
[7] Advanced Television Systems Committee, *ATSC Standard: A/322, Physical Layer Protocol*, Doc. A/322, Jan. 2020.
[8] Advanced Television Systems Committee, *ATSC Standard: A/327, ATSC Recommended Practice: Guidelines for the Physical Layer Protocol*, Doc. A/327, Jan. 2020.
[9] S.–I. Park *et al.*, ``Performance analysis of all modulation and code combinations in ATSC 3.0 physical layer protocol,'' *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 197-210, Jun. 2019.
[10] M. Earnshaw *et al.*, “Physical layer framing for ATSC 3.0,” *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 263-270, Mar. 2016.
[11] J. Lee *et al.*, “Layered division multiplexing for ATSC 3.0: Implementation and memory use aspects,” *IEEE Trans. Broadcast.*, vol. 65, no. 3, pp. 496-503, Sep. 2019.
[12] D. He *et al.*, “System discovery and signaling transmission using bootstrap in ATSC 3.0,” *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 172-180, Mar. 2016.
[13] Advanced Television Systems Committee, *ATSC Standard: A/331, Signaling, Delivery, Synchronization, and Error Protection*, Doc. A/331, Jan. 2020.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

[14] W. Kwon *et al.*, "The ATSC link-layer protocol (ALP): Design and efficiency evaluation," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 316-327, Mar. 2016.
[15] G. K. Walker *et al.*, "ROUTE/DASH IP streaming-based system for delivery of broadcast, broadband, and hybrid services," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 328-337, Mar. 2016.
[16] D. Gómez-Barquero *et al.*, "IEEE transactions on broadcasting special issue on: Convergence of broadcast and broadband in the 5G era," *IEEE Trans. Broadcast.*, vol. 66, no. 2, part 2, Jun. 2020.
[17] M. Simon *et al*., "ATSC 3.0 broadcast 5G unicast heterogeneous network converged services starting Release 16," *IEEE Trans. Broadcast*., vol. 66, no. 2, pp. 449-458, Jun. 2020.
[18] J. Lee *et al.*, "IP-based cooperative services using ATSC 3.0 broadcast and broadband," *IEEE Trans. Broadcast.*, vol. 66, no. 2, part 2, Jun. 2020.
[19] S. Ahn *et al*., "Cooperation between LDM-based terrestrial broadcast and broadband unicast: On scalable video streaming applications," accepted for publication in *IEEE Trans. Broadcast.*
[20] D. Gómez-Barquero *et al*., "MIMO for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 298-305, Mar. 2016.
[21] L. Stadelmeier *et al*., "ATSC 3.0 channel bonding," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 289-297, Mar. 2016.
[22] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2-14, Jan. 1972.
[23] R. Zhang and L. Hanzo, "A unified treatment of superposition coding aided communications: Theory and Practice," *IEEE Trans. Wireless Commun*., vol. 11, no. 7, pp.2628-2639, Jul. 2012.
[24] S. Vanka *et al*., "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun*., vol. 11, no. 7, pp. 2628-2639, Jul. 2012.
[25] *Study on Downlink Multiuser Superposition Transmission (MUST) for LTE*, document TR 36.859, 3GPP, Dec. 2015.
[26] L. Zhang *et al*., "Layered-division-multiplexing: Theory and practice," *IEEE Trans. Broadcast*., vol. 62, no. 1, pp. 216-232, Mar. 2016.
[27] S.–I. Park *et al*., "Low complexity layered division multiplexing for ATSC 3.0," *IEEE Trans. Broadcast*., vol. 62, no. 1, pp. 233-243, Mar. 2016.
[28] C. Regueiro *et al*., "LDM core services performance in ATSC 3.0," *IEEE Trans. Broadcast*., vol. 62, no. 1, pp. 244-252, Mar. 2016.
[29] E. Garro *et al*., "Layered division multiplexing with multi-radio-frequency channel technologies," *IEEE Trans. Broadcast*., vol. 62, no. 2, pp. 365-374, Jun. 2016.
[30] W. Li *et al*., "Using LDM to achieve seamless local service insertion and local program coverage in SFN environment," *IEEE Trans. Broadcast*., vol. 63, no. 1, pp. 250-259, Mar. 2017.
[31] J. Lee *et al*., "Multiple service configurations based on layered division multiplexing," *IEEE Trans. Broadcast*., vol. 63, no. 1, pp. 267-274, March 2017.
[32] L. Zhang *et al*., "Layered-division-multiplexing: An enabling technology for multicast/broadcast service delivery in 5G," *IEEE Commun. Mag*., vol. 56, no. 3, pp. 82-90, Mar. 2018.
[33] E. Garro *et al*., "Layered division multiplexing with distributed multiple-input single-output schemes," *IEEE Trans. Broadcast*., vol. 65, no. 1, pp. 30-39, Mar. 2019.
[34] L. Zhang *et al*., "Using layered-division-multiplexing to deliver multi-layer mobile services in ATSC 3.0," *IEEE Trans. Broadcast*., vol. 65, no. 1, pp. 40-52, Mar. 2019.
[35] L. Zhang *et al*., "Layered-division-multiplexing for high spectrum efficiency and service flexibility in next generation ATSC 3.0 broadcast system," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 116-123, Apr. 2019.
[36] E. Garro *et al.*, "Layered division multiplexing with co-located multiple-input multiple-output schemes," *IEEE Trans. Broadcast.*, vol. 66, no. 1, pp. 9-20, Mar. 2020.
[37] H. Yamamoto *et al.*, "A Study on LDM-BST-OFDM transmission for the next-generation terrestrial broadcasting," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 205-215, Jun. 2020.
[38] L. Zhang *et al.*, "Using layered division multiplexing for wireless in-band distribution links in next generation broadcast systems," accepted for publication in *IEEE Trans. Broadcast*.
[39] J. Montalban *et al.*, "Improved semi-blind channel estimation with time domain cancellation for LDM-LSI," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 613-619, Sep. 2020.
[40] D. Gómez-Barquero and O. Simeone, "LDM versus FDM/TDM for unequal error protection in terrestrial broadcasting systems: An information theoretic view," *IEEE Trans. Broadcast*., vol. 61, no. 4, pp. 571-579, Dec. 2015.
[41] T. M. Cover and J. A. Thomas, *Element of Information Theory*, New York, NY, USA: Wiley, 2006.
[42] S. Ahn *et al*., "Fronthaul compression and precoding optimization for NOMA-based joint transmission of broadcast and unicast services in CRAN," accepted for publication in *IEEE Trans. Broadcast.*
[43] S.-I. Park *et al*., "Transmitter identification signal analyzer for single frequency network," *IEEE Trans. Broadcast.*, vol. 54, no. 3, pp. 383-393, Sep. 2008.
[44] S.-I. Park *et al*., "An efficient receiver structure for robust data transmission using TxID signal in the ATSC DTV system," *IEEE Trans. Consum. Electron.*, vol. 56, no. 2, pp. 408-414, May 2010.
[45] Y.-W. Suh *et al.*, "A novel TxID insertion system for ATSC DTV auxiliary data transmission," *IEEE Consum. Electron*., vol. 57, no. 1, pp. 35-39, Feb. 2011.
[46] S.-I. Park *et al*., "RF watermark backward compatibility tests for the ATSC terrestrial DTV receivers," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 246-252, Jun. 2011.
[47] B. Rong *et al*., "Signal cancellation techniques for RF watermark detection in ATSC mobile DTV system," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 4070-4076, Oct. 2011.
[48] S.-I. Park *et al*., "ATSC 3.0 transmitter identification signals and applications," *IEEE Trans. Broadcast.*, vol. 63, no. 1, pp. 240-249, Mar. 2017.
[49] J. Lee *et al*., "Efficient ATSC 3.0 TxID signal detection for single frequency networks," *IEEE Trans. Broadcast.*, vol. 65, no. 2, pp. 326-332, Jun. 2019.
[50] S. Ahn *et al*., "Probabilistic analysis on RF-watermark TxID detection in SFN with randomly distributed co-channel interferers and preamble cancellation," *IEEE Access*, vol. 8, pp. 56300-56311, Mar. 2020.
[51] S. Kwon *et al*., "Detection schemes for ATSC 3.0 transmitter identification in single frequency network," *IEEE Trans. Broadcast*, vol. 66, no. 2, pp. 229-240, Jun. 2020.
[52] J. Lee *et al.*, "Transmitter identification signal detection algorithm for ATSC 3.0 single frequency networks," *IEEE Trans. Broadcast*, vol. 66, no. 3, pp. 737-743, Sep. 2020.
[53] S. Jeon *et al*., "Field trial results for ATSC 3.0 TxID transmission and detection in single frequency network of Seoul, in *Proc. IEEE BMSB*, Valencia, Spain, Jun. 2018.
[54] H. Jung *et al.*, "ATSC 3.0 channel bonding performance with equal PLP rate in fixed channel environment," in *Proc. IEEE BMSB*, Valencia, Spain, Jun. 2018.
[55] H. Jung *et al*., "ATSC 3.0 Channel bonding performance in mobile channel environments," in *Proc. IEEE BMSB*, Jeju Island, South Korea, Jun. 2019.
[56] H. Jung *et al.*, "Implementation of terrestrial 8K broadcast system using ATSC 3.0 channel bonding," in *Proc. IEEE BMSB*, Paris, France, Oct. 2020.
[57] D. Gómez-Barquero *et al*., "MIMO for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 298-305, Mar. 2016.
[58] S. Saito *et al*., "8K Terrestrial Transmission field tests using dual-polarized MIMO and higher-order modulation OFDM," *IEEE Trans.Broadcast.*, vol. 62, no. 1, pp. 306-315, Mar. 2016.
[59] H.-J. Yim *et al*., "Implementation of targeted advertisement services on ATSC 3.0 runtime environment," *IEEE BMSB*, Jeju Island, South Korea, Jun. 2019.
[60] H.-J. Yim *et al.*, "Application-based targeted advertisement system for ATSC 3.0 UHD service," accepted for publication in *IEEE Trans. Broadcast*.
[61] J. Lee *et al.*, "Efficient transmission of multiple broadcasting services using LDM and SHVC," *IEEE Trans. Broadcast.*, vol. 64, no. 2, pp. 177-178, Jun. 2018.
[62] S. Ahn *et al.*, "Large-scale network analysis on NOMA-aided broadcast/unicast joint transmission scenarios considering content probability," accepted for publication in *IEEE Trans. Broadcast.*
[63] D. Vargas and Y. J. D. Kim, "Two-layered superposition of broadcast/multicast and unicast signals in multiuser OFDMA systems," *IEEE Trans. Wireless Commun*., vol. 19, no. 2, pp. 979-994, Feb. 2020.
[64] J. Zhao *et al.*, "Non-orthogonal unicast and broadcast transmission via joint beamforming and LDM in cellular networks," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 216-228, Jun. 2020.
[65] L. Zhang *et al.*, "Using non-orthogonal multiplexing in 5G-MBMS to achieve broadband-broadcast convergence with high spectral efficiency," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 490-502, Jun. 2020.
[66] E. Iradier *et al.*, "Using NOMA for enabling broadcast/unicast convergence in 5G networks," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 503-514, Jun. 2020.
[67] L. Zhang *et al*., "Using layered division multiplexing for wireless in-band distribution links in next generation broadcast systems," accepted for publication in *IEEE Trans. Broadcast*.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.2 Web Link: http://dx.doi.org/10.18580/setijbe.2020.2

[68] S.-I. Park *et al.*, "Augmented data transmission based on low density parity check code for the ATSC terrestrial DTV system," *IEEE Trans. Broadcast.*, vol. 58, no. 4, pp. 637-641, Dec. 2012.
[69] S.-I. Park *et al.*, "A newly designed quarter-rate QC-LDPC code for the cloud transmission system," *IEEE Trans. Broadcast.*, vol. 59, no.1, pp. 155-159, Mar. 2013.
[70] S.-I. Park *et al.*, "Raptor-like rate compatible LDPC codes and their puncturing performance for the cloud transmission system," *IEEE Trans. Broadcast.*, vol. 60, no.2, pp. 239-245, Jun. 2014.
[71] K.-J. Kim *et al.*, "Low-density parity-check codes for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no.1, pp. 189-196, Mar. 2016.
[72] N.-S. Loghin *et al.*, "Non-uniform constellations for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no.1, pp. 197-203, Mar. 2016.
[73] H. Jeong *et al.*, "Flexible and robust transmission for physical layer signaling of ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no.1, pp. 204-215, Mar. 2016.
[74] J. Barrueco *et al.*, "Constellation design for bit-interleaved coded modulation (BICM) systems in advanced broadcast standards," *IEEE Trans. Broadcast.*, vol. 63, no.4, pp. 603-614, Dec. 2017.
[75] S.-K. Ahn *et al.*, "Comparison of low-density parity-check codes in ATSC 3.0 and 5G Standards," *IEEE Trans. Broadcast.*, vol. 65, no.3, pp. 489-495, Sep. 2019.
[76] 3GPP TR 38.913 V15.0.0, "Study on scenarios and requirements for next generation access technologies (Release 15)", Jun. 2018.
[77] 3GPP TS 38.212 V15.7.0, "Multiplexing and channel coding (release 15)," Sep. 2019.
[78] 3GPP TR 36.976 V1.0.0, "Overall description of LTE-based 5G broadcast (Release 16)", Dec. 2019.
[79] S. –I. Park *et al.*, "Field test results of layered division multiplexing for the next generation DTV system," *IEEE Trans. Broadcast.*, vol.63, no. 1, pp. 260-266, Mar. 2017.
[80] ETSI TS 102 831: "Digital Video Broadcasting (DVB); Implementation guidelines for a second generation digital terrestrial television broadcasting system (DVB-T2)," V1.2.1, Aug. 2012.
[81] S. –I. Park *et al.*, "Field comparison tests of LDM and TDM in ATSC 3.0," *IEEE Trans. Broadcast.*, vol.64, no. 3, pp. 637-647, Sep. 2018.
[82] B.-m. Lim *et al.*, "Mobile field comparison test of LDM and TDM based on ATSC 3.0," in *Proc. IEEE BMSB*, Cagliari, Italy, Jun. 2017.
[83] B.-m. Lim *et al.*, "Mobile Testing of ATSC 3.0 MISO in SFN," in *Proc. IEEE BMSB*, Jeju Island, South Korea, Jun. 2019.
[84] S. Ahn *et al.*, "Multi-antenna diversity gain in terrestrial broadcasting receivers on vehicles: A coverage probability perspective," *submitted to ETRI J.*

**Sungjun Ahn** received B.S. and M.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2015 and 2017, respectively. He has been with Media Research Division, Electronics and Telecommunications Research Institute (ETRI) since 2017, where he is currently a Research Engineer. His research interests include stochastic geometry analysis, signal processing, and optimization for wireless communications and digital broadcasting.

**Sunhyoung Kwon** received the B.S. degree in electrical communications engineering from Information and Communications University, Daejeon, Korea, in 2008, and the M.S. and Ph.D. degree in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, in 2010 and 2020, respectively. Since 2010, he has been with the Media Research Division, Electronics and Telecommunication Research Institute, where he is currently a senior member of research staff. His research interests are in the area of digital broadcasting and communications.

**Seok-Ki Ahn** received the B.S., M.S., and Ph.D. degrees in electronics and electrical engineering from Pohang University of Science and Technology (POSTECH) in 2006, 2008, and 2013, respectively. From 2010 to 2013, he was a student on scholarship at the Digital Media and Communications R&D Center, Samsung Electronics. From 2013 to 2018, he was a Senior Engineer with Samsung Electronics, Suwon, South Korea. He is currently with Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea. His research interests include channel coding, MIMO transceiver design, and broadband communications.

**Hoiyoon Jung** received the B.S. and M.S. degree in information and communication engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2006 and 2008, respectively. He has been with Media Research Division, Electronics and Telecommunications Research Institute (ETRI) since 2008, where he is currently a Senior Researcher. His research interests include 5G cellular system and next generation terrestrial broadcasting system.

**Sung-Ik Park** received the B.S.E.E. degree from Hanyang University, Seoul, South Korea, in 2000 and the M.S.E.E. degree from POSTECH, Pohang, South Korea, in 2002, and the Ph.D. degree from Chungnam National University, Daejeon, South Korea, in 2011. Since 2002, he has been with the Media Research Division, Electronics and Telecommunication Research Institute (ETRI), where he is Project Leader and Principal Member of Research Staff. His research interests are in the area of error correction codes and digital communications, in particular, signal processing for digital television. He has over 200 peer-reviewed journal and conference publications, and multiple best paper and contribution awards for his work on broadcasting technologies. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON BROADCASTING and ETRI Journal, and a Distinguished Lecturer of IEEE Broadcasting Technology Society.

Received in 2020-10-12 | Approved in 2020-11-16

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

# TDS-OFDM based Digital Television Terrestrial Multimedia Broadcasting Standards

Jian Song, Chao Zhang,
Jintao Wang, Yonglin Xue,
Changyong Pan, Kewu Peng,
Fang Yang, Jun Wang,
Hui Yang, Yu Zhang
and Zhixing Yang

**CITE THIS ARTICLE**

Song, Jian; Zhang, Chao; Wang, Jintao; Xue, Yonglin; Pan, Changyong; Peng, Kewu; Yang, Fang; Wang, Jun; Yang, Hui; Zhang, Yu and Yang, Zhixing; 2020. TDS-OFDM based Digital Television Terrestrial Multimedia Broadcasting Standards. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2020.3. Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

**COPYRIGHT** This work is made available under the Creative Commons - 4.0 International License. Reproduction in whole or in part is permitted provided the source is acknowledged.

# TDS-OFDM based Digital Television Terrestrial Multimedia Broadcasting Standards

Jian Song, Chao Zhang, Jintao Wang, Yonglin Xue, Changyong Pan, Kewu Peng, Fang Yang, Jun Wang, Hui Yang, Yu Zhang, and Zhixing Yang

***Abstract*—As the most popularly utilized broadcasting network, digital terrestrial television broadcasting (DTTB) can provide multimedia information coverage for the broad audience in a very efficient way because of its characteristic of wide-range coverage and mobile reception ability. After promulgating the first generation DTTB standard, digital terrestrial/television multimedia broadcasting (DTMB), in 2006, China began to research and develop the next generation DTTB standard, namely DTMB-advanced (DTMB-A), aiming to support higher spectrum efficiency and further improve transmission reliability. In 2019, DTMB-A was accepted by ITU as the second generation international DTTB standard (as System C). Similar to DTMB, time-domain synchronous - orthogonal frequency division multiplexing (TDS-OFDM) based multi-carrier modulation scheme is adopted by DTMB-A. Thanks to the more flexible frame structure, advanced error correction coding and improved constellation mapping, DTMB-A offers 30% higher transmission capacity than DTMB under the same transmission conditions. Thus, DTMB-A can support both fixed and mobile reception more efficiently, and provide users with higher quality services such as ultra-high definition television (UHDTV). This paper first gives details of key technologies at the transmitter of DTMB/DTMB-A and introduce core algorithms at the receiver. Both laboratory test and field trial results will then be provided and analyzed, especially for the application of 4K UHDTV and single frequency network (SFN).**

***Index Terms*—digital terrestrial television broadcasting (DTTB), digital terrestrial/television broadcasting (DTMB), DTMB-advanced (DTMB-A), time-domain synchronous orthogonal frequency division multiplexing (TDS-OFDM), ultra-high definition TV (UHDTV)**

## I. Introduction

Digital terrestrial television broadcasting (DTTB) can meet the basic demands of "information to person" in modern society with the advantages of high power and spectrum efficiency, which offers great value. Therefore, DTTB has been extensively adopted in many countries. The first generation DTTB standards, including American advanced television systems committee (ATSC) standard [1], European digital video broadcasting - terrestrial (DVB-T) standard [2], Japanese integrated service digital broadcasting - terrestrial (ISDB-T) standard [3] and Chinese digital terrestrial/ television multimedia broadcasting (DTMB) standard [4]-[6], have facilitated the popularization of DTTB around the world. With the continuous development of multimedia information technology, a variety of new media services and high-quality audio and video services such as augmented reality (AR), virtual reality (VR), 3-dimensional TV (3DTV), ultra-high definition TV (UHDTV), etc. are emerging. Such challenges from the new broadcasting services become great driving forces to promote the evolution of DTTB standards.

New types of multimedia services always require higher transmission data rates, which means the DTTB system needs to provide higher spectrum efficiency within the same bandwidth and the requirement for the transmission reliability also increases significantly, which is beyond the capability of the first generation DTTB standards. Therefore, the study and development of the second generation DTTB standard become very important and are now in progress in many countries.

Europe began the research of the second generation DTTB standard first and promulgated DVB-T2 in 2008 [7]. DVB-T2 adopts the latest breakthrough for modulation and coding technologies at that time, providing efficient and reliable audio/video and data transmission for fixed, portable and mobile devices. The maximum data rate of DVB-T2 is about 50.1Mbps in 8MHz bandwidth using extended bandwidth mode. Under the same planning restrictions and working conditions, DVB-T2 increases the transmission capacity by over 30% compared with DVB-T.

ATSC started the preparation of the new generation DTTB standard ATSC 3.0 in 2011. ATSC opened the solicitation of technical proposals in 2013, and technical status was basically frozen in 2015. In June 2017, the physical layer transmission standard of ATSC 3.0 (A/322) [8] has been finally released. Without the backward compatibility requirement, orthogonal frequency division multiplexing (OFDM) technology combining with advanced constellation mapping and channel coding are utilized by ATSC 3.0, supporting higher spectrum efficiency and throughput. Therefore, ATSC 3.0 gives broadcasters more choices of technical solutions.

China launched the research and development of a new generation DTTB standard, DTMB - advanced (DTMB-A) [9][10], based on the key technologies of DTMB since 2008. DTMB-A retains time-domain synchronous - OFDM (TDS-OFDM) as its basic transmission technology similar to DTMB.

This work was supported by the China National Science Foundation under Grant 61931015.

The authors are with the Beijing National Research Center for Information Science and Technology (BNRist), Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: z_c@tsinghua.edu.cn).

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

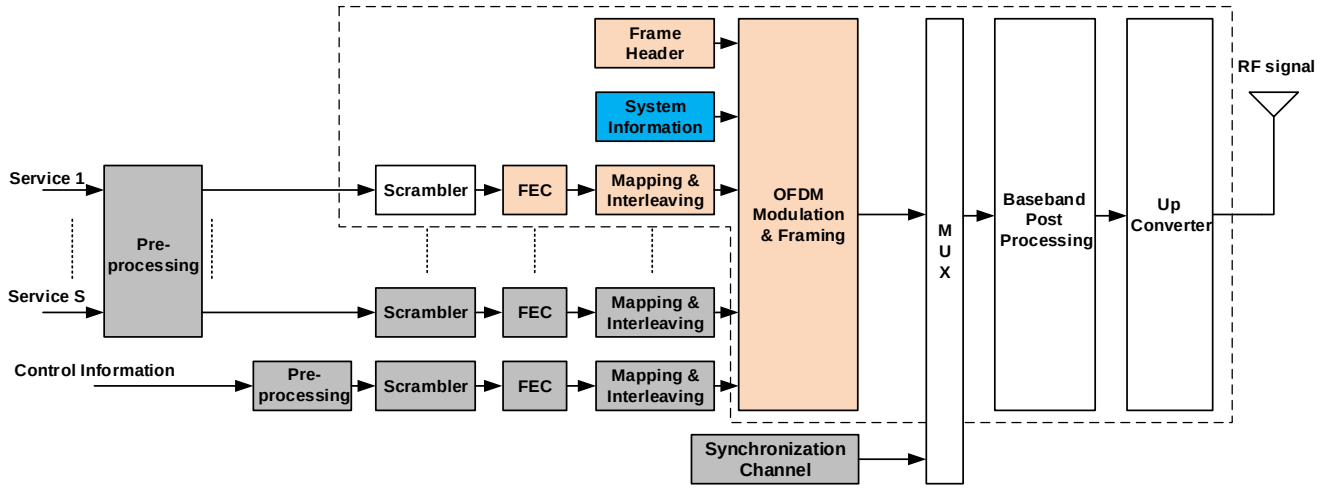



Fig. 1. Functional block diagram of the transmitter for DTMB and DTMB-A.

Besides, the more flexible frame structure incorporating amplitude phase shift keying with Gray mapping (Gray-APSK) and bit-interleaved coded modulation (BICM) based on low density parity check (LDPC) code provides DTMB-A more powerful performance. Benefit from the above techniques, DTMB-A has over 30% higher spectrum efficiency than DTMB under the same transmission conditions, or in another word, possesses better immunity to interference and lower receiving threshold under the same spectrum efficiency. DTMB-A can support both fixed and mobile reception, and can support UHDTV broadcasting and high-performance single frequency network (SFN). In December 2019, DTMB-A was adopted by ITU as the System C of the second generation international DTTB standards with DVB-T2 as System A and ATSC 3.0 as System B [11].

Extensive laboratory tests and field trials have been conducted during the development of the DTMB-A system [12]-[14]. In 2018, the first 4K UHDTV broadcasting experimental network based on DTMB-A was established in Jiaxing City, Zhejiang Province. Moreover, in August 2019, invited by Hong Kong Television Broadcasts Limited (TVB), Tsinghua University and Beijing Digital Television National Engineering Laboratory (DTNEL) organized the field trials to evaluate the performance of SFN and UHDTV broadcasting in Hong Kong. The test results prove that DTMB-A can support reliable 4K UHDTV program transmission under the fairly complex receiving environment in Hong Kong. In August 2019, the DTMB-A based 4K transmission system with 6MHz bandwidth was exhibited at the SET EXPO in Sao Paulo, Brazil.

Both DTMB and DTMB-A adopt TDS-OFDM as their key technology. Hence, this paper will start from TDS-OFDM, and systematically introduce the system architectures and the key techniques used by DTMB and DTMB-A. In addition, details of algorithms utilized at the receiver are also discussed as an important part of the whole DTTB system. Laboratory test and field trial results of the DTMB-A system will be then provided, especially for the application of the 4K UHDTV demonstration network.

The rest of this paper is organized as follows. Section II introduces the system architecture of the transmitter. Key technologies of transmitter and receiver are provided in Section III and Section IV respectively. Section V shows the laboratory test and field trial results. A conclusion is drawn in Section VI.

## II. System Architecture

### A. *Transmitter Structures*

In order to show the structures clearly and give a comparison between DTMB and DTMB-A system, a block diagram combining DTMB and DTMB-A transmitter modules is shown in Fig. 1. Since only the multi-carrier transmission mode of TDS-OFDM is used in DTMB-A, only the structure of the OFDM transmission is described. In Fig. 1, the parts in the dotted box constitute the DTMB transmitter, while the gray modules are extra parts for the DTMB-A transmitter. The blue module is only available in the DTMB system and the orange ones in the dotted box are common functional blocks for both systems but improved distinctly for the DTMB-A system with newly developed technologies. Thus, differences between the two systems are clear from Fig. 1.

It is illustrated that the DTMB system only supports single service broadcasting. After multiplexing the service data into standard MPEG-TS packet, scrambling, forward error correction (FEC) coding, constellation mapping and interleaving are executed successively and data constellation symbols are obtained. These symbols are multiplexed with the signaling of transmission parameters (system information) and then allocated to $C$ sub-carriers. After that, inverse fast Fourier transform (IFFT) is performed to get the time-domain signal of the frame body. Then, the frame header sequence and frame body are combined to form the signal frame. After the pulse shaping filtering and orthogonal up-conversion, the radio frequency (RF) signal is obtained and ready to be transmitted over the wireless channel.

DTMB-A has distinct improvements in comparison with DTMB. DTMB-A uses super-frame as its important

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

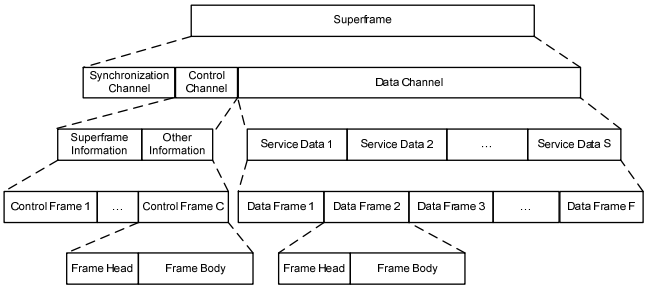



Fig. 2. Super-frame structure of DTMB-A.

transmission unit in the physical layer, and it comprises of logical channels for super-frame synchronization, control if needed, and service data. Furthermore, DTMB-A is able to support simultaneous transmission of multiple services with different quality requirements and provide independent sub-channels for each of them, which fundamentally distinguishes from single service transmission mode by program multiplexing.

DTMB-A has a flexible design of frame structure based on TDS-OFDM similar to DTMB, but it pads the frame header with multi-carrier pseudo-random noise (PN-MC) sequence instead of time-domain PN sequence. It can help the receiver perform channel estimation and equalization with lower complexity and higher accuracy. Moreover, in order to adapt to different application scenarios, DTMB-A provides various optional frame specifications. The frame header supports three options including 512, 1024 and 2048 symbols, while the frame body length can be 4096 (4K), 8192 (8K) and 32768 (32K) symbols.

Amplitude phase shift keying with Gray mapping (Gray-APSK) is utilized in DTMB-A. Significant shaping gain is attained compared to traditional quadrature amplitude modulation (QAM), which enables reliable information recovery at a lower demodulation threshold. With that, DTMB-A can support high-order constellation up to 256, which provides the maximum payload data rate of 49.57Mbps in regular bandwidth mode, or 51.00Mbps in extended mode. It means that under 8MHz RF channel and adopting the newest H.265 video compression coding, the DTMB-A system can transmit 10 HDTV programs simultaneously or one 4K UHDTV program.

DTMB-A uses improved LDPC coding. The compatibility of multiple code rates and multiple code lengths is fully considered in the design. Three different code rates, including 1/2, 2/3 and 5/6, and two different code lengths, including 15360 and 61440, have been provided. At the receiver, a high-throughput parallel iterative decoder can be implemented to improve the decoding performance at the same time as keeping low hardware complexity. Additionally, LDPC decoding and constellation demapping can be combined and processed iteratively, which further lowers the threshold compared with conventional independent decoding and demapping.

DTMB-A also supports transmit diversity with two antennas as an option, which improves the robustness under deep fading channel. The diversity mode can also be applied in SFN. By applying diversity codes from different transmitters, the influence of the artificial multipath in traditional SFN can be effectively eliminated, which improves the coverage of the SFN.

### *B. Frame Structure*

DTMB standard has a hierarchical super-frame structure, including signal frame, super-frame, minute frame and day frame from bottom to top. The signal structure has periodicity and is kept synchronized with nature time. As the basic transmission unit of DTMB, the signal frame consists of frame header and frame body, where the frame header is constituted by time-domain PN sequence and the frame body is dependent on the number of sub-carriers.

In DTMB-A, the more flexible physical layer super-frame structure is utilized which is depicted in Fig. 2. As the fundamental unit transmitted in the physical layer, the super-frame is made up of synchronization channel, control channel and data channel. The synchronization channel is used for primary synchronization of the super-frame and transmitting basic parameters. *S* services are allocated on the data channel, each of them occupies an integral number of data frames. Every super-frame has *F* data frames, where *F* is determined by transmission parameters. Control channel has C control frames, which carry the service configuration information, parameters for demapping and decoding, and fast real-time information (e.g., short message). Data frame and control frame have the same frame structure consisted of frame header and frame body, and both adopt TDS-OFDM.

The preamble of each super-frame is the synchronization channel, which is used for fast signal capture, coarse timing, estimation of carrier frequency offset, and basic transmission parameters acquisition.

Control frame is generated by using low-rate of FEC coding, quadrature phase shift keying (QPSK) mapping, symbol interleaving, IFFT, and frame header insertion. Control channel transmits the information of frame structure, service composition and other necessary information for multiple service broadcasting. Each super-frame can contain C control frames, which are placed right after the synchronization channel one by one. For single service mode, the control frame isn't needed and can be eliminated.

Each data frame within the data channel may have independent mode of coding and modulation, which can be configured flexibly according to the application. Similar processes in control channel are carried out to obtain data frames. Moreover, both the control frame and the data frame can have different lengths of either frame header or frame body.

Synchronization channel, data frames of all the services and the control frames are combined to compose a super-frame. Baseband post-processing is implemented to get the baseband transmission signal, which is followed by orthogonal up-conversion and the RF signal is obtained.

### *C. System Parameters*

Table I lists some basic transmission parameters of DTMB and DTMB-A systems. It is shown that DTMB-A has made

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

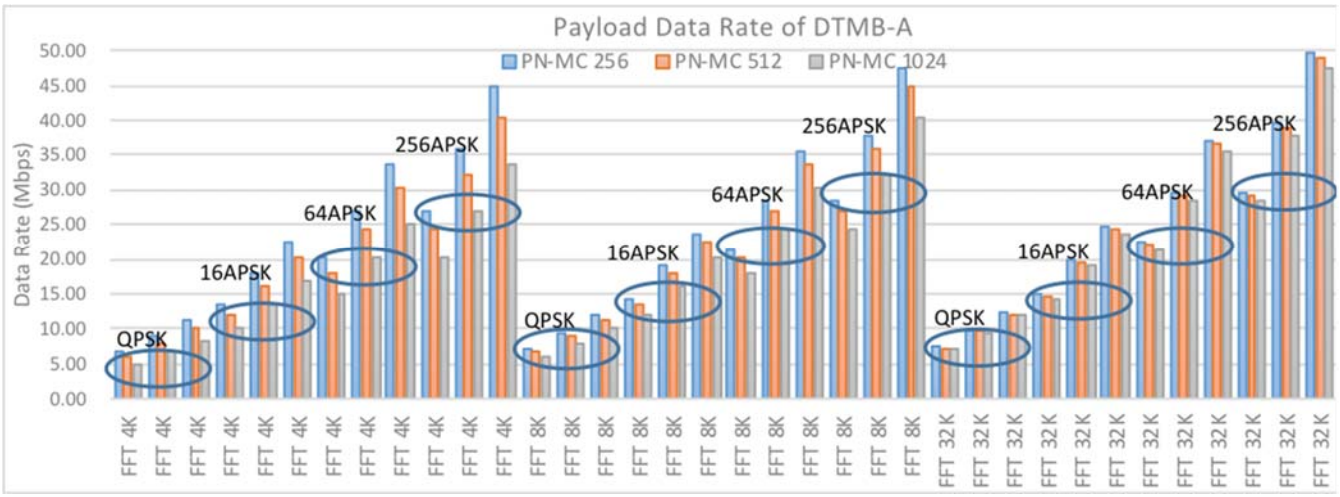



Fig. 3. Payload data rate of DTMB-A in 7.56MHz bandwidth mode.

TABLE I
TRANSMISSION PARAMETERS OF DTMB AND DTMB-A

| Parameter | DTMB | DTMB-A |
|---|---|---|
| Modulation | TDS-OFDM, Single-carrier | TDS-OFDM |
| FEC | BCH, LDPC | BCH (option), LDPC |
| LDPC code length | 7493 | 15360, 61440 |
| LDPC code rate | 0.4, 0.6, 0.8 | 1/2, 2/3, 5/6 |
| Constellation | QPSK, QPSK-NR, 16QAM, 32QAM, 64QAM | QPSK, 16APSK, 64APSK, 256APSK |
| Guard interval | PN420, PN595, PN945 | 2PN-MC 256, 2PN-MC 512, 2PN-MC 1024 |
| FFT size | 3780 | 4096, 8192,32768 |
| Roll-off factor | 0.05 | 0.05, 0.025 |
| Bandwidth | 7.56 MHz | 7.56 MHz, 70/9 MHz |
| Payload data rate | 4.813~32.486 Mbps | 5.00~51.00 Mbps |
| Multi-service | No | Yes |
| Transmit diversity | No | Yes |
| Enhanced SFN | No | Yes |
| PAPR reduction | No | Yes |

certain enhancement on channel coding and modulation, frame structure, therefore, has higher spectrum efficiency, and can support more functions like multiple service broadcasting, enhanced SFN and so on, which makes DTMB-A satisfy the requirements of new types of broadcasting services well.

### *D. Payload Data Rates*

Both DTMB and DTMB-A provide a number of options of parameters to deal with different applications. The payload data rate is determined by the choice of the parameter set, which includes frame header and frame body lengths, constellation order, channel coding rate, system bandwidth etc. Taking DTMB-A as an example, the payload data rates under different working modes in 7.56MHz bandwidth channel are shown in Fig. 3.

Data in Fig. 3 are divided into three groups with the respect to the FFT length of 4K, 8K and 32K for the horizontal axis. Every group is further divided into four sub-groups by the constellation order. Data in the sub-groups with the same color

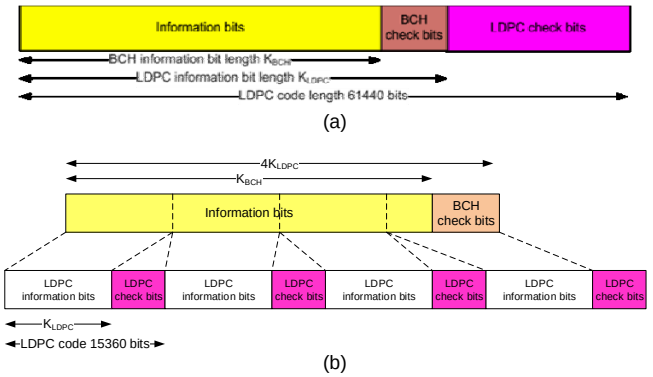



Fig. 4. The codeword structure of the LDPC code. (a) LDPC code of 61440 bits. (b) LDPC code of 15360 bits.

are arranged from left to right according to the coding rate of 1/2, 2/3 and 5/6 respectively, while different colors represent different frame header lengths. Fig. 3 indicates that DTMB-A can reach the maximum data rate with PN-MC 256, 32K FFT, 256APSK and 5/6 code rate, and has the minimum data rate with PN-MC 1024, 4K FFT, QPSK and 1/2 code rate. Under the fixed frame header length and FEC code rate, the working mode with longer FFT length and higher constellation order is more suitable for the fixed reception, while mobile applications prefer shorter FFT length and lower constellation order.

## III. BIT INTERLEAVED CODED MODULATION

### *A. Forward Error Correction Code*

After randomization, the data stream will be encoded by FEC coding to enhance the transmission robustness.

DTMB-A uses concatenated FEC coding as DTMB, where the inner code is LDPC code and the optional outer code is BCH code. Using BCH code as the outer code can provide stronger error correction ability at a slightly lower data rate.

Comparing with Turbo code, LDPC code has lower decoding complexity [15]. By further optimizing the existing LDPC code in DTMB standard, DTMB-A uses quasi-cyclic LDPC (QC-

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

TABLE II
PARAMETERS OF LDPC CODE

| Code Rate | Information Bit Length | |
|---|---|---|
| | Code Length 61440 | Code Length 15360 |
| 1/2 | 30720 | 7680 |
| 2/3 | 40960 | 10240 |
| 5/6 | 51200 | 12800 |

TABLE III
PARAMETERS OF QC-LDPC CODE

| Code Rate | Code Length 15360, b=128 | | Code Length 61440, b=512 | |
|---|---|---|---|---|
| | *n* | *k* | *n* | *k* |
| 1/2 | 120 | 60 | 120 | 60 |
| 2/3 | 120 | 80 | 120 | 80 |
| 5/6 | 120 | 100 | 120 | 100 |

LDPC) code, which further lowers both the carrier-to-noise ratio (C/N) threshold and implementation complexity, and is also suitable for the multiple code rates and code lengths, aiming at supporting fixed and mobile reception [16]. Considering the structure of the QC-LDPC encoder, receiver performance, storage occupation, and power consumption, the optimal time-frequency interleaving scheme is applied to DTMB-A for better performance.

DTMB-A has two code length options of 15360 and 61440 and three code rate options of 1/2, 2/3 and 5/6, which are shown in Table II. The code word structures of two code lengths are shown in Fig. 4 (a) and (b) respectively.

QC-LDPC code is an important subset of LDPC code and its parity check matrix has a quasi-cyclic structure. This kind of LDPC code is simple in structure, easy to design and has excellent performance. Because of the regular structure of the parity check matrix, the QC-LDPC encoder can be implemented by a relatively simple circuit while parallelization can be adopted in decoding to lower the hardware complexity. The generation matrix of the QC-LDPC code with systematic structure used by DTMB-A can be expressed as

$$\mathbf{G}_{\mathrm{qc}} = \begin{bmatrix} \mathbf{I} & \mathbf{O} & \cdots & \mathbf{O} & \mathbf{G}_{0,0} & \cdots & \mathbf{G}_{0,n-k-1} \\ \mathbf{O} & \mathbf{I} & \cdots & \mathbf{O} & \mathbf{G}_{1,0} & \cdots & \mathbf{G}_{1,n-k-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \mathbf{G}_{i,j} & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{I} & \mathbf{G}_{k-1,0} & \cdots & \mathbf{G}_{k-1,n-k-1} \end{bmatrix} \quad (1)$$

where **I** is *b*×*b* identity matrix, **O** is *b*×*b* zero matrix, $G_{i,j}$ is *b*×*b* circulant matrix with $0 \le i \le k-1$, $0 \le j \le c-1$. The values of *n*, *k*, and *b* under different coding parameters are shown in Table III.

### B. Constellation Mapping

DTMB standard adopts traditional QAM constellation mapping with maximum order of 64.

Information theory suggests that only Gaussian-distributed input can achieve the channel capacity over power-limited additive white Gaussian noise (AWGN) or fading channel. Restricted by constellation, traditional regular QAM mapping won't generate the signal with Gaussian distribution, so there is a gap between theoretical channel capacity and reliable data rate under the constellation constraint. Accordingly, the technique that makes mapping output closer to the Gaussian distribution is called shaping, and the resulting gain is called shaping gain.

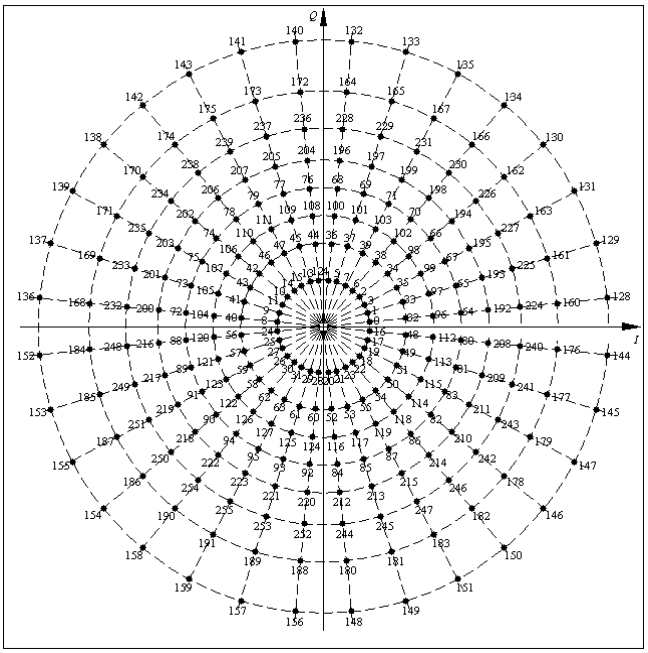


Fig. 5. Gray-APSK constellation of 256 points.

APSK constellation has a typical non-uniform distribution characteristic, which gives it notable shaping gain compared with regular QAM mapping. Gray-APSK design based on mutual information optimization theory can significantly improve channel capacity under constellation constraint and obtain shaping gain with both independent and iterative demapping [17].

Four different constellation orders are defined in DTMB-A, including QPSK, 16APSK, 64APSK, and 256APSK, in which 256APSK can support higher transmission rate and be suitable for fixed reception with high spectrum efficiency, while QPSK can provide higher robustness and be suitable for severe transmission conditions such as high-speed mobile reception. The typical constellation of 256 Gray-APSK used in DTMB-A is shown in Fig. 5.

### C. BICM-ID

It is proved that interleaving in bit-level can bring higher diversity order to improve the performance. Thus, combining the channel coding and constellation mapping by bit-interleaving can provide better performance in fading channel. Such technique is called bit-interleaved coded modulation (BICM). If at the receiving side, the demapping process of BICM doesn't use the constraint introduced by channel coding among bits, performance loss, especially with high order constellation mapping will occur. Therefore, the output of decoding can be fed back to the demapping module to assist demapping, which forms the process of BICM with iterative decoding/demapping (BICM-ID) [18][19]. The block diagrams of both BICM and BICM-ID systems are shown in Fig. 6 (a) and (b) respectively.

Gray-APSK adopted in DTMB-A means the amplitude and phase are independent at each constellation point and both of

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

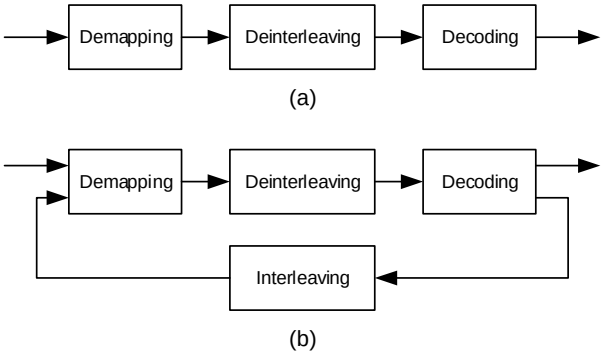



Fig. 6. Block diagram of the demapping and decoding. (a) BICM. (b) BICM-ID.

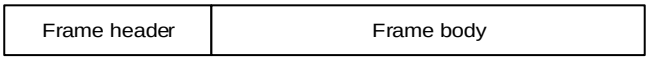



Fig. 7. The signal frame structure of TDS-OFDM.

them are Gray mapped respectively. Thus, DTMB-A is more suitable for both independent demapping and LDPC or Turbo code based BICM-ID, which further enhances the iterative gain.

## IV. TDS-OFDM BASED FRAME STRUCTURE

### A. *TDS-OFDM Technology*

One of the biggest challenges of terrestrial broadcast wireless transmission is the frequency selective fading caused by multipath. OFDM has natural advantages in combating it. However, a strict requirement of synchronization is needed to ensure the orthogonality among subcarriers when applying OFDM.

TDS-OFDM is adopted both in DTMB and DTMB-A as the fundamental transmission technique. TDS-OFDM is a multi-carrier transmission scheme that can flexibly support time and frequency domain processing when needed. Through the combined processing of time and frequency domains, it can easily realize fast signal acquisition and robust synchronization, and also supports time-frequency two-dimensional segment for the physical layer resource allocation. The signal frame structure of TDS-OFDM illustrated in Fig. 7 consists of frame header and frame body. The frame header is a known training sequence, usually a PN sequence. It is used as the guard interval between two adjacent OFDM data blocks instead of traditional cyclic-prefix or zero-padding. Since the sequence is known by the receiver, the frame header is also used for parameter synchronization and channel estimation. Thus, TDS-OFDM does not need to insert any frequency-domain pilot which improves the spectrum efficiency obviously.

### B. *Signal Frame Structure Based on PN-MC*

DTTB system usually works under complex multipath conditions. For example, the receiver will receive the many reflected signal from buildings and complex terrain. When SFN is applied, artificial multipath also exists and the receiver will receive the transmitted signals from several nearby transmitters

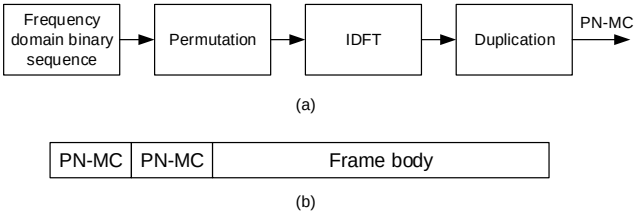



Fig. 8. PN-MC structure. (a) Generation method of PN-MC. (b) TDS-OFDM signal frame structure using dual PN-MC as the frame header.

through different paths and processes them as echoes. The multipath signal will significantly influence the performance of the receiver, leading to a higher bit error rate and even failure of reception.

In TDS-OFDM system, the frame header can be used to facilitate fast synchronization, efficient channel estimation and equalization. Moreover, since the frame header adopts known training sequences, the accuracy of the parameter estimation can be improved remarkably under low SNR conditions. The longer the frame header is, the better it is to resist long delay echo signal, at the cost of lower payload data rate. In addition, the long frame header mode can be a preferable alternative for wide range SFN.

DTMB system has three different frame header lengths, including PN420, PN595 and PN945. These three kinds of frame headers are based on the time-domain PN sequence, and generated by cyclic expansion (PN420, PN945) or truncation (PN595). However, in multipath channel, there is intrinsic interference between the frame header and the frame body when TDS-OFDM is used. Iterative interference cancellation algorithm is necessary to obtain high precision channel estimation, which increases the hardware complexity of the receiver.

In order to further improve the precision and also reduce the implementation complexity, DTMB-A adopts an improved frame header design method, using two repeated PN-MC sequences as the frame header. The details of the generation procedure and the frame structure are depicted in Fig. 8 (a) and (b) respectively. As illustrated in Fig. 8 (a), the known frequency-domain binary sequence is firstly interleaved in the frequency-domain by permutation operation. Then, IFFT is performed and the time-domain PN-MC sequence is obtained. Finally, the PN-MC sequence is repeated and put in front of the frame body as the frame header, creating the frame structure shown in Fig. 8 (b). DTMB-A supports three PN-MC sequence lengths of 256, 512 and 1024, as well as three frame body lengths of 4096, 8192 and 32768.

Since the frame header of DTMB-A contains two identical PN-MC sequences, the first one can be regarded as the cyclic prefix of the second PN-MC, which effectively avoids the interference from the frame body of the former signal under multipath channels. Therefore, precise channel estimation can be obtained at slightly lower spectrum efficiency by using the second PN-MC sequence without any interference, and the iterative interference cancellation can be eliminated for the receiver compared with traditional TDS-OFDM system

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

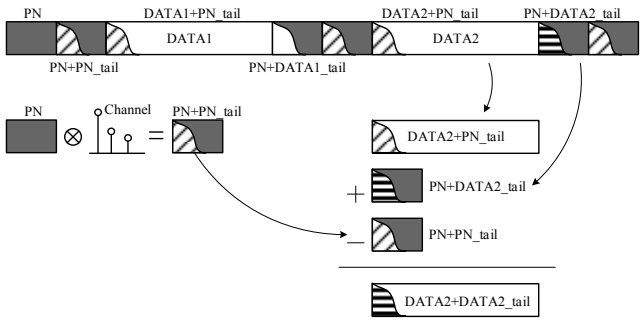



Fig. 9. The algorithm of cyclic reconstruction used in DTMB-A.

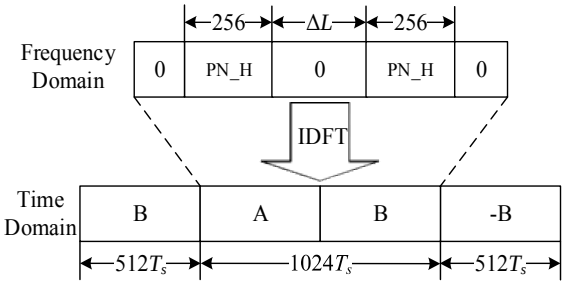



Fig. 10. The time and frequency domain structure of the synchronization channel.

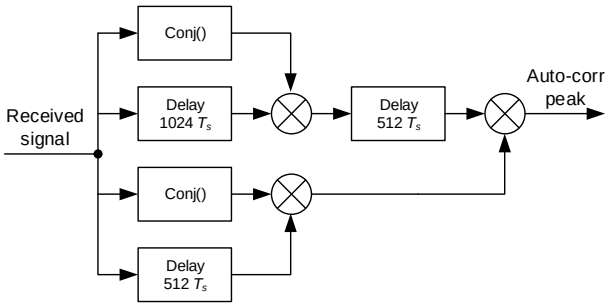



Fig. 11. The algorithm of signal detection using the synchronization channel.

[20][21]. After the accurate channel estimating, the receiver can utilize a simple overlap-and-add algorithm to complete the cyclic reconstruction of the OFDM data block. The schematic diagram of the algorithm is shown in Fig. 9.

### *C. Synchronization Channel*

The super-frame synchronization channel is a novel design in the DTMB-A standard [22]. It is placed at the beginning of each super-frame and has a special structure in both time and frequency domains. It can provide fast and robust signal detection, coarse timing and carrier estimation in low SNR conditions. At the same time, it can support fundamental signaling transmission, which helps the receiver get basic parameters of the physical layer signal and guide the subsequent demodulation and decoding.

The structure of the super-frame synchronization channel is displayed in Fig. 10. The basic unit is an OFDM symbol with the length of 1024 samples. In the frequency domain, the OFDM symbol has two identical PN sequences, namely PN_H, whose length is 256, while other subcarriers in the symbol are padded with 0. There are $\Delta L$ subcarriers between these two PN_H sequences where $\Delta L \in [64,\ 319]$. Therefore, it means that $\Delta L$ has 256 options which can be used to transmit 8-bit signaling information. Let $\{Z_m\}_{m=0}^{1023}$ denotes the 1024-point frequency-domain signal, and it can be expressed as

$$Z_m = \begin{cases} 0 & 0 \le m < 256 - \lceil \Delta L/2 \rceil \\ \text{PN\_H}_{m-256+\lceil \Delta L/2 \rceil} & 256 - \lceil \Delta L/2 \rceil \le m < 512 - \lceil \Delta L/2 \rceil \\ 0 & 512 - \lceil \Delta L/2 \rceil \le m < 512 + \lfloor \Delta L/2 \rfloor \\ \text{PN\_H}_{m-512-\lfloor \Delta/2 \rfloor} & 512 + \lfloor \Delta L/2 \rfloor \le m < 768 + \lfloor \Delta L/2 \rfloor \\ 0 & 768 + \lfloor \Delta L/2 \rfloor \le m < 1024 \end{cases} \tag{2}$$

where $\lceil \cdot \rceil$ denotes round towards plus infinity, and $\lfloor \cdot \rfloor$ presents denotes round towards minus infinity.

This kind of signaling transmission method based on distance variation is one of the important features of DTMB-A.

The frequency-domain OFDM symbol is converted to the time domain signal of length 1024 by IFFT. After that, the time domain signal is divided into two parts of A and B with the same length. Then, part B and the minus version of B are copied to the front and the end of the time domain OFDM symbol respectively as the pre and postambles, which yields the super-frame synchronization channel with the special structure of [B,A,B,-B] in the time domain.

One of the important functions of the super-frame synchronization channel is to realize the detection of DTMB-A signal from the received signal without any prior information, and extract the basic transmission parameter signaling for demodulation and decoding. Since there are three repeated segments of B (including an inverse one) in the time domain, the low complexity time-domain auto-correlation can be used for the signal detection. Fig. 11 gives an illustration of the algorithm of the signal detection process using the super-frame synchronization channel. The correlation between the first and the third parts of B, and between the third and the fourth parts of B, are calculated respectively. These two correlation results are combined by multiplication or addition to ensure a more sharp auto-correlation peak [23].

The super-frame synchronous channel can realize signal detection and recognition under the noisy environment, and ensures the fast detection of the DTMB-A signal.

## V. Field Trial Results

The prototype verification system and commercial set-top box of DTMB-A have been fully tested in the laboratory, and field trials have been carried out in many cities and districts in China, including Changsha, Kunming, Urumchi, Jiaxing and Hong Kong. The reception performance of the DTMB-A system in various typical urban environments has been evaluated sufficiently. In this section, we will focus on the test results of 4K in Jiaxing and SFN plus 4K in Hong Kong, respectively.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

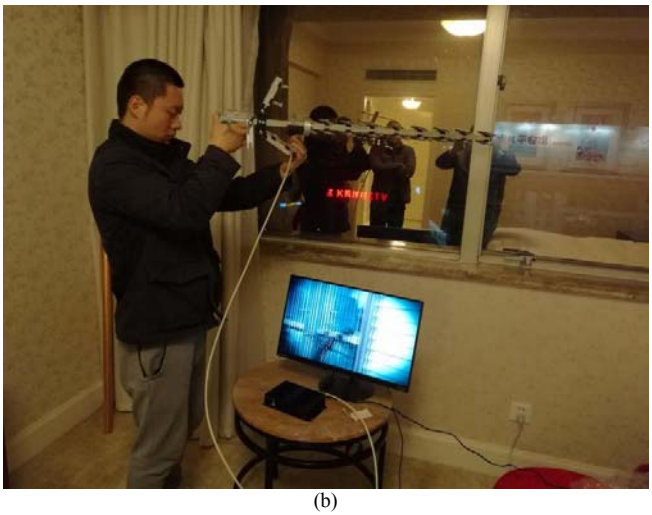
TABLE IV
PARAMETERS OF THE FIELD TRIALS IN JIAXING CITY

| Parameter | Value |
|---|---|
| Transmission system | DTMB-A |
| Bandwidth | 7.56MHz |
| Constellation | 256APSK |
| FFT size | 32K |
| Guard interval | Dual PN-MC 256 |
| FEC | LDPC 61440 |
| FEC rate | 2/3 |
| Payload data rate | 39.7Mbps |
| Center frequency | 562MHz |
| Transmission power | 1000W |
| Transmission antenna | Horizontal polarized |
| Video coding | H.265 |
| Audio coding | MPEG-4 AAC |
| Transmission antenna height | 145m above ground level |
| Receiving antenna height | 10m above ground level |

TABLE V
TEST RESULTS IN JIAXING CITY

| Outdoor reception | | | Indoor reception | | |
|---|---|---|---|---|---|
| No. | Signal strength (dBm) | Margin (dB) | No. | Signal strength (dBm) | Margin (dB) |
| 1 | -52.9 | 28 | 1 | -50.6 | 32 |
| 2 | -65.5 | 13 | 2 | -57.9 | 18 |
| 3 | -65.5 | 10 | 3 | -36 | 46 |
| 4 | -72.9 | 2 | 4 | -55.6 | 27 |
| 5 | -68.4 | 10 | 5 | -40.8 | 37 |
| 6 | -65.2 | 11 | 6 | -39 | 40 |
| 7 | -66.6 | 14 | 7 | -35.2 | 48 |
| 8 | -50.2 | 32 | | | |
| 9 | -61.8 | 14 | | | |
| 10 | -70.2 | 4 | | | |

(a)



(b)

Fig. 12. The field trial in Jiaxing City. (a) Outdoor reception. (b) Indoor reception.

### A. *4K UHDTV Broadcasting in Jiaxing City*

In order to validate the feasibility and reliability of DTMB-A system for 4K UHDTV program broadcasting, the experimental DTMB-A network was set up in Jiaxing City, Zhejiang Province in August 2018 [14]. Reception performance is tested at several typical sites in Jiaxing City, including both indoor and outdoor receptions. The transmission mode and the transmitter parameters are summarized in Table IV and the situations of the test sites are shown in Fig. 12.

During the measurements, 10 typical fixed outdoor and 7 representative fixed indoor receiving points were chosen, with the locations of the outdoor and indoor testing points and the corresponding transmitting stations shown in Fig. 13 (a) and (b) respectively. The received signal strength and the margin to the reception threshold are measured at each testing point and the results are given in Table V. Among the chosen points, the maximum distance from the receiver to the transmitter was about 39km, which proves the good coverage effect of DTMB-A in the application of 4K UHDTV thanks to its excellent performance.

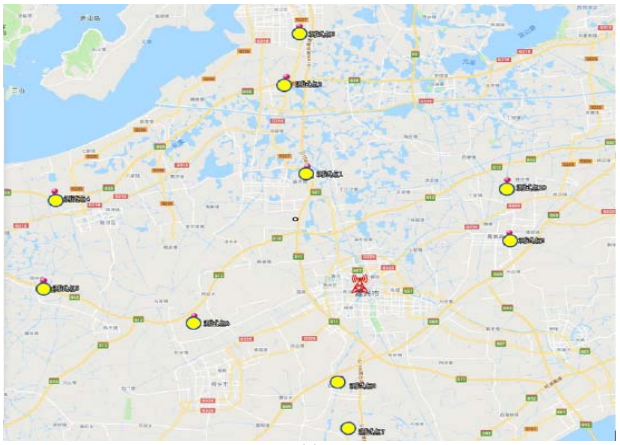
### B. *Field Trial Results of SFN in Hong Kong*

Hong Kong is characterized by concentrated high-rise buildings and dense population, complex hilly urban terrain, and also unpredictable tidal effect, which gives a serious challenge to the coverage of the DTTB signal. Since 2007, Hong Kong has successfully established the SFN covering the whole area with good performance using DTMB.

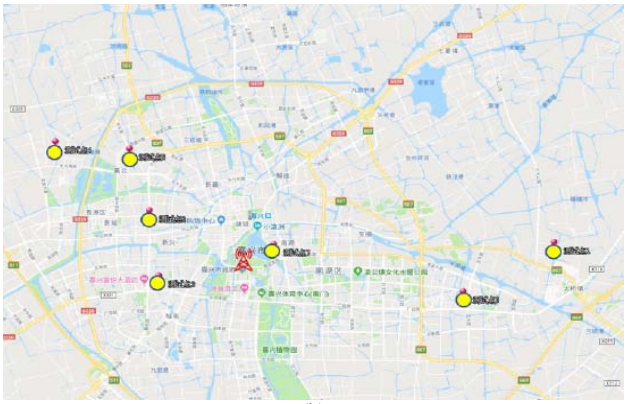
In order to further prove the coverage performance of the DTMB-A system in Hong Kong, the SFN field trials were conducted in August 2019 using three DTMB-A transmitting stations to form the SFN [13]. Outdoor fixed reception, mobile reception, and the tidal fading tests were included for the measurements. One 4K UHDTV program (data rate of 25.5Mbps) and one 2K HDTV program (data rate of 6Mbps) using H.265 were selected for the fixed and mobile receptions, respectively.

Three transmitting stations of the tested SFN were located at Temple Hill, Golden Hill and Kowloon Peak, and their basic technical parameters are listed in Table VI. Three typical working modes were chosen, one for fixed reception and two for mobile reception. The detailed transmission parameters are

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.



(a)



(b)

Fig. 13. The locations of the test points. (a) Outdoor reception. (b) Indoor reception.


Fig. 14. The locations of the transmission stations in Hong Kong.


(a)

(b)

Fig. 15. The field trial conditions in Hong Kong. (a) Transmission stations. (b) Survey Car.

TABLE VI
TRANSMITTER PARAMETERS IN HONG KONG

| | Temple Hill | Golden Hill | Kowloon Peak |
|---|---|---|---|
| ERP | 1000W | 320W | 320W |
| Frequency | 602MHz | | |
| Antenna type | 4 dipoles on the reflector | | |
| Polarization | Horizontal polarization | | |
| Antenna height | 527m | 404m | 651m |
| Location | 22.35595°N 114.20665°E | 22.36258°N 114.14534°E | 22.34089°N 114.22333°E |

TABLE VII
PARAMETERS OF THE FIELD TRIALS IN HONG KONG

| Parameter | Fixed | Mobile1 | Mobile2 |
|---|---|---|---|
| Bandwidth | 7.56MHz | | |
| Constellation | 256APSK | 16APSK | 64APSK |
| FFT size | 32K | 4K | 4K |
| Guard interval | Dual PN-MC 1024 | | |
| LDPC length | 61440 | | |
| LDPC rate | 2/3 | 1/2 | 1/2 |
| Payload data rate | 37.89Mbps | 10.07 | 15.10 |

shown in Table VII.

The failure criterion of subjective evaluation was adopted in the fixed reception test. A failure according to the threshold of visibility (TOV) criterion was counted if three times of mosaic was observed within one minute. For the mobile reception, the DTMB-A receiver would give an error package indication signal from the LDPC decoder to a dedicated mobile test recorder. The recorder would record the time, the coordinates of the location, and the corresponding field strength of the received signal when the error package appeared.

In order to establish DTMB-A SFN, the SFN adapter and the TS player were equipped in Temple Hill station besides DTMB-A transmitter. The generated TS from the TS player is processed by the SFN adapter and a special packet carrying SFN synchronization information is inserted into the TS signal. The output TS is transmitted to Golden Hill and Kowloon Peak stations through the program distribution link, and radiated by the DTMB-A transmitters. In order to guarantee the clock

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

© 2020 SET - Brazilian Society of Television Engineering / ISSN (Print): 2446-9246/ISSN (Online): 2446-9432

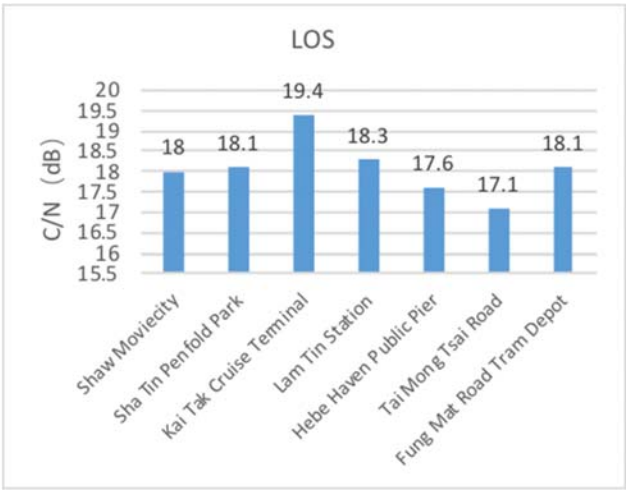



Fig. 16. Measurement results of fixed reception under LOS.

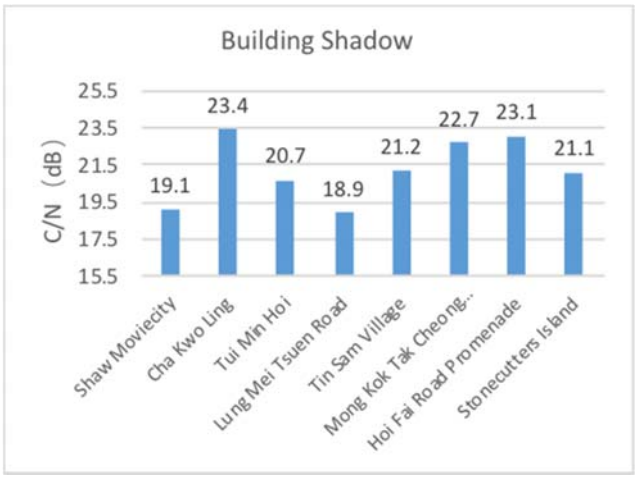



Fig. 17. Measurement results of fixed reception under building shadow.

synchronization of all stations, a global position system (GPS) timing receiver is set up at each station to provide stable clock reference for all equipment. The geographical location of the three stations is shown in Fig. 14. The photos of the equipment construction at the stations and the situations of the field trials are presented in Fig. 15.

*1) Fixed Reception*

The performance of the fixed reception is meaningful to check the coverage effect of the DTMB-A signal in Hong Kong. Thus, a large number of fixed reception test points were selected to evaluate the receiving quality. The C/N threshold and the minimum reception level under line-of-sight (LOS) and building shadow conditions were mainly measured in the test.

Figs. 16 and 17 give the C/N of various testing points under LOS and building shadow conditions, respectively. The figures indicate that DTMB-A with 8MHz bandwidth can provide reliable signal reception and support 4K UHDTV under the above mentioned two conditions.

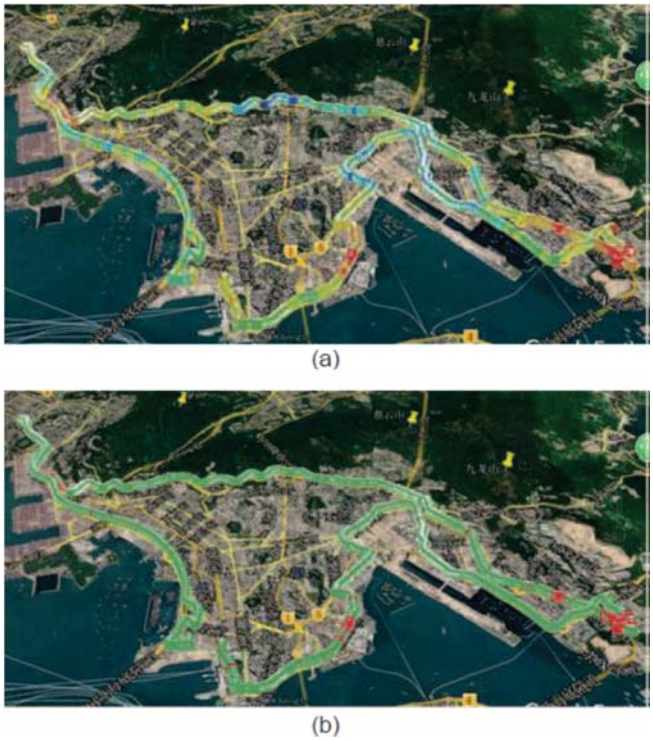



Fig. 18. Measurement results for mobile reception. (a) Signal Strength. (b) Block error.

*2) Mobile Reception*

Mobile reception is important to mobile terminals such as vehicle TVs and smartphones. Thus, the performance of mobile reception was fully tested on typical routes including Hong Kong Island Expressway, Kowloon Island loop line and Nathan Road.

Fig. 18 shows the results of mobile reception tests on the Kowloon Island loop line. Fig. 18 (a) gives the received signal strengths recorded at one-second intervals over the range from -90 dBm to -55 dBm. Different colors indicate different signal strengths, for example, the colors from red to yellow, green and blue mean the signal power changing from weak to strong. Fig. 18 (b) gives the block error results of the same points as (a), where green denotes error-free and red means the occurrence of decoding error. It is clear that except for the positions where the received signal strength was about or lower than -90dBm, the DTMB-A receiver could recover the video programs without error on most parts of the test route.

The selected routes of the whole test referred to the field trials of DTMB. The test results fully prove that DTMB-A can achieve the same coverage effect as DTMB with over 30% higher spectrum efficiency. In other words, DTMB-A can provide better performance or larger coverage area under the same spectrum efficiency as DTMB.

## VI. Conclusion

DTMB and DTMB-A both adopt TDS-OFDM as their basic physical layer transmission scheme. Compared to DTMB, DTMB-A can provide higher spectrum efficiency, better

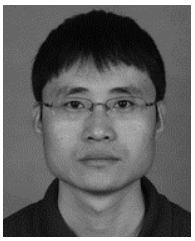
This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

receiving reliability, and stronger robustness to the multipath. Thus, DTMB-A can support more applications with better service quality. This paper focuses on the details of the physical layer transmission techniques of DTMB and DTMB-A, showing the similarities and differences, and presents field trial results of the DTMB-A system. The results confirm that DTMB-A can provide reliable signal reception under various typical complicated receiving conditions, reliable coverage for both fixed and mobile reception. DTMB-A based SFN works well and 4K UHDTV programs can be transmitted through a DTMB-A network successfully.

## REFERENCES

[1] *ATSC Digital Television Standard*, ATSC Standard, A/53 parts 1–6, Jan. 2007.
[2] *Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*, ETSI Standard EN 300 744 v1.6.2, Oct. 2015.
[3] *Transmission System for Digital Terrestrial Television Broadcasting*, ARIB Standard STD-B31 Version 2.2, Mar. 2014.
[4] *Framing Structure, Channel Coding and Modulation for Digital Television Terrestrial Broadcasting System*, Chinese National Standard GB20600-2006, Aug. 2006.
[5] J. Song *et al.*, "Technical review on Chinese digital terrestrial television broadcasting standard and measurements on some working modes," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 1–7, Mar. 2007.
[6] J. Song, Z. Yang, and J. Wang, *Digital Terrestrial Television Broadcasting: Technology and System*. Piscataway, NJ, USA: Wiley, 2015.
[7] *Digital Video Broadcasting (DVB); Frame Structure Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2)*, ETSI EN 302 755 v1.4.1, Jul. 2015.
[8] *ATSC Standard: Physical Layer Protocol (A/322)*, Adv. Television Syst. Committee, Washington, DC, USA, Jun. 2017.
[9] J. Song and C. Zhang, "Technical review on DTMB-advanced (DTMB-A) standard," in *Proc. Int. Conf. Eng. Telecommun. (EnT)*, Nov. 2016, pp. 128–133.
[10] J. Song, C. Zhang, K. Peng, J. Wang, C. Pan, F. Yang, J. Wang, H. Yang, Y. Xue, Y. Zhang and Z. Yang, "Key technologies and measurements for DTMB-A system," *IEEE Trans. Broadcast.*, vol. 65, no. 1, pp. 53-64, Mar. 2019.
[11] *Error-correction, data framing, modulation and emission methods and selection guidance for second generation digital terrestrial television broadcasting systems*, Recommendation ITU-R BT.1877-2, Dec. 2019.
[12] C. Pan, J. Wang, H. Fang and J. Song, " Field trial of advanced DTMB system DTMB-A in Hong Kong," in *Proc. Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, pp. 1–4, Jun. 2013.
[13] C. Pan, C. Zhang, H. Yang, J. Wang and X. Li, "Results of the DTMB-A field trials in Hong Kong," in *Proc.Int. Conf. Engineering and Telecommunication (EnT)*, Nov. 2019.
[14] Y. Xue, H. Yang, C. Pan and J. Song, "Field trial of UHDTV over digital television terrestrial broadcasting network," in *Proc. Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Jun. 2019.
[15] K.-J. Kim *et al.*, "Low-density parity-check codes for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 189–196, Mar. 2016.
[16] A. I. V. Casado, W.-Y. Weng, S. Valle, and R. D. Wesel, "Multiple-rate low-density parity-check codes with constant blocklength," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 75–83, Jan. 2009.
[17] Z. Liu, Q. Xie, K. Peng, and Z. Yang, "APSK constellation with gray mapping," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1271–1273, Dec. 2011.
[18] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 1, no. 6, pp. 169–171, Nov. 1997.
[19] T. Cheng, K. Peng, Z. Liu, and Z. Yang, "Efficient receiver architecture for LDPC coded BICM-ID system," *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1089–1092, Jul. 2015.
[20] Z. Yang, X. Wang, Z. Wang, J. Wang, and J. Wang, "Improved channel estimation for TDS-OFDM based on flexible frequency-binary padding," *IEEE Trans. Broadcast.*, vol. 56, no. 3, pp. 418–424, Sep. 2010.
[21] J. Fu, J. Wang, J. Song, C.-Y. Pan, and Z.-X. Yang, "A simplified equal-ization method for dual PN-sequence padding TDS-OFDM systems," *IEEE Trans. Broadcast.*, vol. 54, no. 4, pp. 825–830, Dec. 2008.
[22] Z. Gao, C. Zhang, and Z. Wang, "Robust preamble design for synchronization, signaling transmission, and channel estimation," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 98–104, Mar. 2015.
[23] J. Liu, C. Zhang and C. Pan, "Reliable transmission parameter signaling detection for DTMB-A standard," *IEICE Trans. Commun*., vol. E100-B, no. 12, pp. 2156-2163, Dec. 2017.

**Jian Song** received all his degrees in electrical engineering from Tsinghua University, Beijing China. He was with the Chinese University of Hong Kong, Hong Kong and University of Waterloo, Canada, in 1996 and 1997, respectively. He worked for industry in the US for seven years before joining the faculty team at Tsinghua University as a Professor in 2005. He is currently the Director of the Tsinghua's DTV Technology Research and Development Center. His research interests include digital television terrestrial broadcasting, wireless communication, power line communication, and visible light communication, and he is the recipient of the IEEE Scott Helt Memorial Award in 2015. He has published over 240 peer-reviewed journal and conference papers in the aforementioned areas and one book in DTV area by Wiley, holds two U.S. and over 40 Chinese patents. He is a Fellow of IEEE, IET and CIE.



**Chao Zhang** received the B.S. and Ph.D. degrees from Beihang University in 2001 and 2008, respectively. From 2008 to 2010, he was a Post-Doctoral Fellow with the Department of Electronic Engineering, Tsinghua University, Beijing, China. From 2011, he was an Assistant Professor with the Department of Electronic Engineering, Tsinghua University. He is currently an Associate Professor with the Department of Electronic Engineering, Tsinghua University. He has authored over 50 journal and conference papers. He holds over 20 Chinese patents. His research interests are in wireless and visible light communications with an emphasis on OFDM, MIMO, synchronization, channel estimation and channel equalization. He received the IEEE Scott Helt Memorial Award (Best Paper Award in IEEE TRANSACTIONS ON BROADCASTING) in 2016.

**Jintao Wang** received the B.Eng. and Ph.D. degrees in electrical engineering both from Tsinghua University, Beijing, China, in 2001 and 2006, respectively. From 2006 to 2009, he was an Assistant Professor in the Department of Electronic Engineering at Tsinghua University. Since 2009, he has been an Associate Professor and Ph.D. Supervisor. He is a [24] Standard Committee Member for the Chinese national digital terrestrial television broadcasting standard. His current research interests include space-time coding, MIMO, OFDM, and SFN systems. He has published more than 100 journal and conference papers and holds more than 40 national invention patents.

**Yonglin Xue** got his Bachelor and Master degrees from the Department of Electronic Engineering, Tsinghua University and Chinese Academy of Sciences in 1989 and 1992, respectively. He is now an Associate Professor with the Research Institute of Information Technology of Tsinghua University with major research interest in the digital TV broadcasting, video coding and communication. He has published more than 70 academic papers.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

**Changyong Pan**, born in Anhui province, China, in 1975, is a full professor in the Research Institute of Information Technology and deputy director of DTV R&D Center of Tsinghua University. Prof. Pan has authored or co-authored more than 180 technical papers and published 8 technical books. He holds 34 patents, won national technical awards three times and is also the winner of numerous other awards.

**Kewu Peng** was born in Hefei, China. He received the B. Eng degree in Electronics Engineering from Hefei University of Technology in 1993, the M.E. degree in Electronics Engineering from Tsinghua University in 1996, and the Ph.D. degree in Electrical Engineering from the University of Minnesota, Minneapolis, in 2003. He was a Researcher and a Lecturer with the Department of Electronic Engineering, Tsinghua University, from Aug 1996 to Aug 1999. Since Jan 2005, he joins the Digital Television Research Center at Tsinghua University as a Research Staff, Assistant Researcher('06), and Associate Professor('09). His research interests include mobile/wireless communications, digital terrestrial/television broadcasting, and image/video transmission. Currently, he is working on LDPC-coded non-orthogonal multiple access. Dr. Peng has published more than 90 journal and conference papers and holds more than 60 China patents.

**Yang Fang** received the B.S.E. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2005 and 2009, respectively. He is currently an Associate Professor with the Research Institute of Information Technology, Tsinghua University. He has authored over 120 peer-reviewed journal and conference papers. He holds over 40 Chinese patents and two PCT patents. His research interests lie in the fields of channel coding, channel estimation, interference cancelation, and signal processing techniques for communication systems, especially in power line communication, visible light communication, and digital television terrestrial broadcasting. He received the IEEE Scott Helt Memorial Award (Best Paper Award in IEEE TRANSACTIONS ON BROADCASTING) in 2015. He is the Secretary General of Sub-Committee 25 of the China National Information Technology Standardization (SAC/TC28/SC25).

**Jun Wang** was born in Henan, China, on October 5, 1975. He received the B.Eng. and Ph.D. degrees from the Department of Electronic Engineering at Tsinghua University, Beijing, China, in 1999 and 2003, respectively. He has been Assistant Professor and member of the DTV Technology R&D Center of Tsinghua University since 2000. His main research interests focus on broadband wireless transmission techniques, especially synchronization and channel estimation. He is actively involved in the Chinese national standard on the Digital Terrestrial Television Broadcasting technical activities, and has been selected by the Standardization Administration of China as the Standard committee member for drafting.

**Hui Yang** was born in Guangxi province, China, in 1967. He received his Master degree from the Department of Electronic Engineering, Tsinghua University. Now, he is the Senior Engineering with the Research Institute of Information Technology of Tsinghua University. His main research interests include digital television transmission, visible light communications and power-line communications, etc. He has published more than 100 technical papers and holds more than 20 patents.

**Yu Zhang** received the B.E. and M.S. degrees in electronics engineering both from Tsinghua University, Beijing, China, in 1999 and 2002, respectively, and the Ph.D. degree in electrical and computer engineering from the Oregon State University, Corvallis, OR, USA, in 2006. From 2007, he was an Assistant Professor with the Research Institute of Information Technology, Tsinghua University, for eight months. He is currently an Associate Professor with the Department of Electronic Engineering, Tsinghua University. His current research interests include the performance analysis and detection schemes for MIMO-OFDM systems over doubly-selective fading channels, transmitter and receiver diversity techniques, and channel estimation and equalization algorithm.

**Zhixing Yang**, born in Hunan China in 1946, is a full professor and PhD advisor with the Department of Electronic Engineering, Tsinghua University, Beijing, China. He is the deputy director of the State Key Lab on Microwave and Digital Communications, Director of the Digital TV (DTV) transmission technology R&D center at Tsinghua University. He is also a member of the special committee of National DTV standardization, group leader of the coordinating working group for DTV standardization of MII, the invited committee member of the science and technology committee of SARFT. He is now the chairman of the National Engineering Lab. for DTV (Beijing).
He has been focusing on the various transmission technologies for both telecommunications and broadcasting areas. Prof. Yang has won national technical invention awards three times and is also the winner of numerous other awards. Prof. Yang is the first draftsman of the Chinese Digital Television Terrestrial Broadcasting Standard, DTMB (GB20600-2006).

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.3 Web Link: http://dx.doi.org/10.18580/setijbe.2020.3

# Future Vision of Interactive and Intelligent TV Systems using Edge AI

Álan L. V. Guedes,
Antonio J. G. Busson,
João Paulo Navarro
and Sérgio Colcher

**CITE THIS ARTICLE**

Guedes, Álan L.V.; Busson, Antonio J.G.; Navarro, João Paulo; Colcher, Sérgio; 2020. Future Vision of Interactive and Intelligent TV Systems using Edge AI. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2020.4. Web Link: http://dx.doi.org/10.18580/setijbe.2020.4

**COPYRIGHT** This work is made available under the Creative Commons - 4.0 International License. Reproduction in whole or in part is permitted provided the source is acknowledged.

# Future Vision of Interactive and Intelligent TV Systems using Edge AI

Álan L. V. Guedes, Antonio J. G. Busson, João Paulo Navarro, and Sérgio Colcher

alan@telemidia.puc-rio.br, busson@telemidia.puc-rio.br,
jpnavarro@nvidia.com, colcher@inf.puc-rio.br,

***Abstract*—Recently the Brazilian DTV system standards have been upgraded, called TV 2.5, in order to provide a better integration between broadcast and broadband services. The next Brazilian DTV system evolution, called TV 3.0, will address more deeply this convergence of TV systems not only at low-level network layers but also at the application layer. One of the new features to be addressed by this future application layer is the use of Artificial Intelligence technologies. Recently, there have been practical applications using Artificial Intelligence (AI) deployed to improve TV production efficiency and correlated cost reduction. The success in operationalize and evaluate these applications is a strong indication of the interest and relevance of AI in TV. This paper presents TeleMídia Lab's future vision on interactive and intelligent TV Systems, with particular focus on edge AI. Edge AI means use in-device capabilities to run AI applications instead of running them in cloud.**

***Index Terms*—Deep Learning; Video Analysis; TV; Edge AI.**

## I. Introduction

Recently the Brazilian DTV system standards have been upgraded, called TV 2.5, in order to provide an integration between broadcast and broadband services. It is due to the widespread adoption of high-speed Internet services, TV devices also use both linear and nonlinear (on-demand) content. ITU calls this combination scenario as IBB (Integrated Broadcast-Broadband) service, which takes advantage of strong points from each one. The former has advantages regarding live events and high audiences, while the last benefits from better navigability and recommendations. The next Brazilian DTV system evolution, called TV, 3.0[1], will address more deeply this convergence of TV systems not only at low-level network layers but also at the application layer. One of the new features to be addressed by this future application layer is the use of Artificial Intelligence technologies.

According to ITU [1], there have been recently numerous practical applications of Machine Learning (ML) and Artificial Intelligence (AI) in TV They supported the broadcast program and production to improve production efficiency and correlated cost reduction. These applications include: automated programming; streaming Optimization; social media analysis; sign language synthesis; content creation from legacy archives; target advertisement; Personalized content; and others. The advances by industry to successfully operationalize and evaluate these applications is a strong indication of the interest and relevance.

This paper presents TeleMídia Lab's future vision on interactive and intelligent TV Systems with particular focus on edge AI. Edge AI means use in-device capabilities to run AI applications instead of running them in the cloud. In order to present this vision, we firstly present the state of art of ML/AI for TV (Section II). Then we discuss our future vision and present our final remarks (Section III and Section IV).

## II. State of the Art

### A. Media Analysis

Methods based on Deep Learning (DL) became the state-of-the-art in various segments related to automatic media analysis. More specifically, Convolutional Neural Networks (CNN) architectures, or ConvNets, have become the primary method used for audio-visual pattern recognition. The classification task consists of mapping media content into one or more distinct categories. Deep Learning architectures based on CNN's (Convolutional neural network or ConvNets) have become the main method used for recognizing audiovisual patterns. Next, we present some method DL methods for image and video analysis.

*Image Analysis*. Since the victory of the AlexNet [2] in the ImageNet 2012 challenge, new and more accurate CNN-based architectures have emerged. The winner of ImageNet 2014, for example, was the InceptionNet [3], which proposed the use of the Inception block, a block that uses several filters of different sizes at the same level to solve the problem of finding information in images. One year later, the ResNet [4] network was the winner of ImageNet 2015, introducing the concept of residual connections, which increased performance and reduced the training time for CNNs. Later, the Inception-Resnet [5] architecture was proposed as a combination of the Inception blocks with the residual connections, producing one of the most popular models that form the basis for many other CNN architectures for extracting features. The SE-Net architecture (Squeeze-and-Excitation Network) [6] is the state-of-the-art in the image classification task,2 obtaining 2.25% error top-5 at ImageNet 2017. SE-Net proposes a new type of block called SE, which improves the network's power of representation by highlighting the inter-dependencies between the image channels and their features maps. For this, SE-Net uses a mechanism that allows the network to re-calibrate features, through which it uses global information to emphasize the most informative features and suppress the features less useful.

[1] https://forumsbtvd.org.br/tv3_0/

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.4 Web Link: http://dx.doi.org/10.18580/setijbe.2020.4

*Video Analysis*. Unlike images, videos are not only visual but also have audible content. Current methods for video classification are generally divided into two stages: (1) the CNNs stage, called backbone, used to extract audio-visual features from the video content; (2) the after-extraction stage, comprised of sophisticated methods for aggregating features, such as NetVLAD [7] and NetFV [8], that can be applied to undermine audio-visual features and perform classification. These methods achieve state-of-theart performance on the YouTube-8M video classification task [9]. To extract the visual features from video, CNNs (e.g., Inception [3], ResNet) pre-trained in the dataset ImageNet are often used. For the extraction of features from the audio, CNNs adapted for the audio domain, such as AudioVGG [10] or WaveNet [11], pre-trained in dataset AudioSet are the most used models. Currently, most of the video analysis community is devoted to the task of classifying video at the segment level (i.e., temporally locating and classifying the segments in the videos). The Video Action Transformer Network [12] is a proposal to localize and classify actions in space and time.

### B. AI for TV

Practical AI applications in broadcast program and production are used to improve production efficiency and correlated cost reduction [1]. It was made possible given the methods such as the image and video analysis discussed in the previous subsection. We discuss some of such applications in what follows.

*Automated Video Previews*. Previews of programs and digest videos give viewers brief introductions of their content. NHK has developed image analysis to recognize characters and performers in order to create previews and digests. Wimbledon Championships also have efforts, in partnership with IBM, to generate highlight clips. Automated Programming. BBC has projects that aim using ML/AI algorithms to mine through thousands of hours of legacy archived content dating back to 1953 to generate programming. In particular, they use information from past scheduling, past audience, content metadata, and other program attributes to learn and mine the archived content for optimal targeted demographic programming.

*Automated Camera*. Broadcasters traditionally use multiple ultra-high-definition cameras for capturing during live events. The capture from these cameras has been used as a feed to a highly reduced or even single-operator human-driven system. The BBC has implemented ML efforts to allow fewer cameras and camera operators to capture a richer scene and environment. Moreover, Endemol Shine Group used ML methods in the Spanish version of the reality show “Big Brother” to capture patterns of interactions happening in the relationships and dynamics of the house members. This way, they anticipate relationship dynamics of the group interactions and direct camera and recordings costs. Compliance tracking and content creation. Companies have also targeted news workflows to help facilitate and improve FCC mandated compliance in production and delivery. TVU Networks, for instance, provides a transcriber service that assures that all video content is US FCC compliant prior to on-air broadcast Additionally, they make use of ML/AI algorithms to mute audio during any profanity or excluded speech.

*Social Media Analysis*. Social media can influence and become a critical part of broadcasting, particularly in breaking news. ML systems can automatically search for trends and notable information for news production from a massive amount of Twitter data and subsequently judge the authenticity of such posts to determine the presence of any target terms typically associated with news-worthy breaking events. These terms (e.g., "fire" and "accident") are categorized information that is representative of types of information that can be featured in news programs Additionally, research is underway to improve image identification of relevant events and objects, from user image Twitter data.

*Streaming Optimization*. The industry is also using AI to improve video encoder efficiency. BitMovin[2] is successfully using AI to enable learning the video content complexity and other features from prior video encode to improve quality and efficiency during later stage encoding. This may result in optimized quality at maximum bandwidth efficiency


(a) Digital Announcer (b) Transcription

Figure 1: NHK Speech Applications [1]

*Digital Announcer*. NHK has successfully used an AI-driven announcer in one of the news programs. It is a CG-generated announcer (Fig. 1a) with a high-quality synthesized voice that reads out news manuscripts about topics being discussed on social networks. As input training data for ML voice synthesis, it uses text data collected from a large number of past news manuscripts and audio data as read by professional announcers This training facilitates the automated reading of news manuscripts in a natural cadence and prosody, reflecting the characteristics of professional newsreaders. NHK has also used the same system in an automatic voice generation scenario for weather information programs. The information is spoken by the system use a style similar to a professional announcer.

*Transcription support*. Live content such as conversations and interviews also require a way to effectively and3 accurately transcribe speech. NHK transcription system (Fig. 1b) greatly improve their transcription workflow. It uses ML speech recognition algorithms to not only generate text but also categorize each individual in the program cast.

*Sign language CG synthesis*. Sign Language accessibility is required for hearing-impaired individuals. There are some ML method efforts to translate the text to Sign language using a CG animation agent. This translation is enabled through the acquisition of training data from content with Sign Language. However, the corpus vocabulary available is small and still cause the translation to be inaccurate. Before achieving a larger corpus, NHK currently uses a speech recognition-

[2] https://bitmovin.com/

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

based tool to support experts fixing Sign Language translation.

*Content Creation from Legacy Archives*. Colourisation of monochrome images in historical or documentary programs can enable more realistic, engaging, and immersive content production. NHK has developed a colourisation technology for monochrome video images that uses AI to learn the colors of various objects in advance, which learns the colors prevalent in the sky, mountains, or buildings.

*Metadata Creation*. Program producers often search for archived video footage and audio files for possible reuse during program production. To support this search, ML methods have become quite effective at successfully automating metadata generation in legacy and new content. Image and Video Analysis (Section II-A) technologies make it possible to automatically describe relevant information and features of the scenes. For instance, multiple TV affiliates use Prime Focus system to improve media asset management by recognizing elements within audio and video content and automatically generate associated metadata. Moreover, NHK has used text detection to add metadata about scenes depicted in TV programs, such as public signs, that can be used to identify the location or address of a building. In a similar way, Nippon TV uses ML methods to detect the player profiles and uniforms and then learns to recognize "individual players" in live event sports.

*Content Personalization*. Content personalization efforts can be deeply supported by ML/AI algorithms. The relevant efforts include demographically targeted and optimized content for different audiences. For instance, BBC used AI algorithms over archived content to create optimized programming to the user demographic. Its methods of learning include scene identification (e.g. landscapes, objects, people); text metadata, including subtitles; and motion activity of videos.

## III. Future Vision

Today, most ML/AI algorithms run and use data stored in cloud services. This way, they take benefit from the scalability, redundancy, and better costs of those services. The cloud-based ecosystem has demonstrated itself as a practical platform to serve some AI applications. However, it has limitations that might prevent the adoption of other AI fields [13]. For instance, we cite autonomous driving that requires robustness and latency to enable vehicle safety and prevent collisions. In this direction, we recently see the rise of Artificial Intelligence at Edge (Edge AI). This approach proposes that the ML/AI algorithms be processed locally on a hardware device instead of the cloud. Such a device uses data, e.g. sensor data or signals, that are created/captured on the device. In particular, the device does not require to be always connected to the internet. This way, Edge AI may reduce costs for data communication, because fewer data will be transmitted because it is processing locally. This is also important regarding streaming and personal data application to prevent making the user vulnerable from a privacy perspective. The expectations from devices at the Edge are ever-growing especially the new in-device hardware specifically designed for AI tasks.

In our future vision for TV, we envisioned Edge-AI services running at the TV receiver. So, TV will also embed hardware specifically designed for AI tasks. To enable this vision, it is required to extend the current TV application layer to take advantage of such features. As previously mentioned, the current TV application layer focuses on support media presentation for both linear and nonlinear (on-demand) content. The extension for Edge-AI is required supporting programmers to describe learning data and recognize content semantics of the media/data at the receiver.

Some efforts have already tried to extend the NCL language, standardized in SBTVD application layer, with AI features. Moreno *et al.* [14] propose an NCL extension that supports the specification of relationships between knowledge concepts. Their approach is inspired by knowledge engineering standards, such as RDF and OWL. To recognize concepts during the presentation of media, they propose a new type of trigger event (called *InferenceRole*), which is an expected concept to be recognized in the media. Guedes *et al.* [15] extend the NCL to support the development of NCL applications with multimodal interactions. It defines expected user interactions through multimodal descriptions (e.g., SRGS for voice recognition). These descriptions are used in a new <*input*> element to represent input devices. Such elements may have a virtual anchor, called *RecognitionAnchor*, that triggers a recognition event when an expected interaction is recognized from the input device. Abreu and Santos [16] propose the *AbstractAnchor*, which is an anchor type that represents parts of a content node where concepts are detected. Then, during the document parsing, the processor analyses all the media and create the timestamps relative to the time interval where expected concepts are recognized in each media. Busson *et al.* [17] propose the *SemanticAnchor* to allow recognition events in run-time. This way, we also can perform recognition events even in applications of live streams. Additionally, they also offer access to properties of recognition events, such as: identifier of the recognized event, part of media where that concept appears, time of recognition event trigger, etc.

*UC-01 In-device TV Program Analyses*. ML methods for recognition, scene classification, and speech recognition, make it possible for TV to understand the semantics of the content of the transmitted video. By using a concept-aware application layer, TV applications can use the semantics of video content as an anchor to trigger events. For example, imagine an e-commerce TV application that notifies viewers of the availability of purchasing a product whenever they appear on the TV program.

*UC-02 In-device Sensitive video filtering*. TV applications may also present user-generated live streaming (e.g., music, lifestyle and gaming), which are available on social media. There are, however, many examples of uncomfortable situations during video conferences when a participant forgets to disable his/her microphone or camera and goes to the bathroom or change clothes. Moreover, some private video services leave room for the spread of sensitive and inappropriate content for certain ages, such as pornography, violence, or other potentially offensive content. This kind of exposure is especially concerning when considering the vulnerability of children spectators. This way, current AI methods for sensitive video analysis [18] can be used to

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.4 Web Link: http://dx.doi.org/10.18580/setijbe.2020.4

automatically block this during their presentation on TV device.

*UC-03 In-device Personalized Content Creation*. With the TV empowered by AI, personalized content based on audio-visual content and subtitles can be generated automatically. Some useful applications involve content translation and auxiliary content for people with disabilities, such as sign language video [19] and audio-description [20]. Imagine, for example, instead of the broadcaster sending the video of sign languages superimposed on the main video, sign languages could be generated in a personalized way only by TV of persons with hearing disabilities.

*UC-04 Spectator Understanding*. TV application may gather user's identification (face recognition) or context information (sentimental analyses) to trigger pre-configured actions, such as: turn-off TV when the user gets slept; suggest new content when the user is not happy with the current one; suggest animation movies when the children are present. Moreover, the user may require to not store personal data in broadcasters' servers. This way, TV application must process ML task in user data (profile, channels choices) locally to provide recommendations.

*UC-05 Virtual TV Assistant*. A Bot is a conversational software that interacts with users. It uses natural language processing in users' textual input, or intents, to perform actions or responses with information. Today, they are used in different industry fields and especially in services regarding consumer support. A Bot may act as a Virtual TV Assistant running in-device and reacting to user actions or behavior. It may support meeting actions such as stop/start recording or even configure background music. In particular, the questions may be regarding the currently selected video, such as: "what is the name of this character"; and "what happens with him in the last episode".

## IV. Final Remarks

In this paper, we presented an overview of the state of-the-art in Deep Learning for media content analysis (image, audio, and video) and described recent works that propose the integration of ML/AI and TV systems. Then we presented our future vision regarding the Edge-AI in the TV receiver. More precisely, we envisioned the ML/AI use on TV to enable in-device media and data analysis. To support such a vision, we describe a set of envisaged use cases to define new requirements and API design. As future work, we cite performing participatory design with TV programmers.

## References

[1] ITU. (2019) Artificial intelligence systems for programme production and exchange. Available: https://www.itu.int/pub/R-REP-BT.2447-2019

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[3] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

[5] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning." in AAAI, vol. 4, 2017.

[6] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," 2019.

[7] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "Netvlad: Cnn architecture for weakly supervised place recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 5297–5307.

[8] A. Miech, I. Laptev, and J. Sivic, "Learnable pooling with context gating for video classification," arXiv preprint arXiv:1706.06905, 2017.

[9] M. Skalic and D. Austin, "Building a size constrained predictive model for video classification," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 0–0.

[10] S. Hershey, S. Chaudhuri, D. P. W. Ellis, J. F. Gemmeke, A. Jansen, C. Moore, M. Plakal, D. Platt, R. A. Saurous, B. Seybold, M. Slaney, R. Weiss, and K. Wilson, "Cnn architectures for large-scale audio classification," in International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017. Available: https://arxiv.org/abs/1609.09430

[11] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," arXiv preprint arXiv:1609.03499, 2016.

[12] R. Girdhar, J. Carreira, C. Doersch, and A. Zisserman, "Video action transformer network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 244–253.

[13] Y.-L. Lee, P.-K. Tsung, and M. Wu, "Techology trend of edge ai," in 2018 International Symposium on VLSI Design, Automation and Test (VLSI-DAT). IEEE, 2018, pp. 1–2.

[14] M. F. Moreno, R. Brandao, and R. Cerqueira, "Extending hypermedia conceptual models to support hyperknowledge specifications," International Journal of Semantic Computing, vol. 11, no. 01, pp. 43–64, 2017.5

[15] Á. L. V. Guedes, R. G. de Albuquerque Azevedo, and S. D. J. Barbosa, "Extending multimedia languages to support multimodal user interactions," Multimedia Tools and Applications, vol. 76, no. 4, pp. 5691–5720, 2017.

[16] R. Abreu and J. A. dos Santos, "Using abstract anchors to aid the development of multimedia applications with sensory effects," in Proceedings of the 2017 ACM Symposium on Document Engineering, 2017, pp. 211–218.

[17] A. J. G. Busson, Á. L. V. Guedes, S. Colcher, R. L. Milidiú, and E. H. Haeusler, "Embedding deep learning models into hypermedia applications," in Special Topics in Multimedia, IoT and Web Technologies. Springer, 2020, pp. 91–111.

[18] P. Almeida, A. Busson, L. V. Guedes, and S. Colcher, "A deep learning approach to detect pornography videos in educational repositories," in Brazilian Symposium on Informatics in Education (Simpósio Brasileiro de Informática na Educação - SBIE), vol. 30, 2020, p. 912.

[19] V. Veríssimo, C. Silva, V. Hanael, C. Moraes, R. Costa, T. Maritan, M. Aschoff, and T. Gaudêncio, "A study on

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

the use of sequence-to-sequence neural networks for automatic translation of brazilian portuguese to libras," in Proceedings of the 25th Brazillian M. Hasegawa-Johnson, A. Black, L. Ondel, O. Scharenborg, and F. Ciannella, "Image2speech: Automatically generating audio descriptions of images," Proceedings of ICNLSSP, Casablanca, Morocco, 2017.

**Álan L. V. Guedes** is researcher at PUC-Rio TeleMídia laboratory and contributed Ginga and NCL standards to the SBTVD and ITU Forum. His research interests include Interactive Multimedia, Immersive Media, and Multimedia Analysis.

**Antonio Busson** is a PhD student in Informatics at PUC-Rio and a researcher at TeleMídia laboratory. Graduated (2012) and Master (2015) in Computing at UFMA. His research interests include multimedia/hypermedia systems and pattern recognition

**João Paulo** is a Solution Architect at NVIDIA with a focus on high-performance computing and Deep Learning. He has extensive experience in the development of algorithms and visualization techniques aimed at geophysical processing.

**Sérgio Colcher** is professor at the PUC-Rio Department of Informatics (DI) and coordinator of the TeleMídia laboratory. He was also professor of MBA courses in Telecommunications Management and MBA in e-Business at Fundação Getúlio Vargas. His areas of interest include computer networks, performance analysis of computer systems, multimedia systems and pattern recognition

Received in 2020-10-12 | Approved in 2020-11-17

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.4 Web Link: http://dx.doi.org/10.18580/setijbe.2020.4

# Practical tests with MMT and ROUTE/DASH on the transport layer of ATSC 3.0

Allan Seiti Sassaqui Chaubet,
George Henrique Maranhão Garcia de Oliveira,
Gustavo de Melo Valeira
and Cristiano Akamine

**CITE THIS ARTICLE**

Chaubet, Allan Seiti Sassaqui; de Oliveira, George Henrique Maranhão Garcia; Valeira, Gustavo de Melo; Akamine, Cristiano; 2020. Practical tests with MMT and ROUTE/DASH on the transport layer of ATSC 3.0. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2020.5. Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

**COPYRIGHT** This work is made available under the Creative Commons - 4.0 International License. Reproduction in whole or in part is permitted provided the source is acknowledged.

# Practical tests with MMT and ROUTE/DASH on the transport layer of ATSC 3.0

Allan Seiti Sassaqui Chaubet, George Henrique Maranhão Garcia de Oliveira, Gustavo de Melo Valeira and Cristiano Akamine

Electrical Engineering and Computing Program
Mackenzie Presbyterian University
Sao Paulo, Brazil

allanseiti.chaubet@mackenzista.com.br, george.oliveira@mackenzie.br, gustavo.valeira@mackenzie.br, cristiano.akamine@mackenzie.br

***Abstract*— The development of new technologies allowed television systems to evolve over time. For the transport layer, the Movie Pictures Expert Group (MPEG) developed several standards to deliver multimedia content, including the MPEG-2 Transport Stream (TS), which has been widely explored for years. However, it was developed before the spread of the internet, which led to the interest in a new standard that could fulfill the needs from a connected world. One of these standards is the MPEG Multimedia Transport (MMT), which has inherited some features from the MPEG-2 TS, adapting them to be compatible with the Internet Protocol (IP). The broadband systems also needed a new standard compatible with Hypertext Transfer Protocol (HTTP), resulting in the development of the MPEG-Dynamic Adaptive Streaming over HTTP (DASH). To deliver DASH on broadcast channels, it was combined with the Real-time Object delivery over Unidirectional Transport (ROUTE) protocol. The Advanced Television Systems Committee 3.0 (ATSC 3.0) adopted many technologies to attend the requirements for the next generation of television systems, including both MMT and ROUTE/DASH. This paper presents a historic background of these delivery methods, as well as a brief technical review, focusing on a practical setup to test the methods mentioned, analyzing the differences and similarities of their properties.**

***Index Terms*—MMT, ROUTE/DASH, ATSC 3.0, Transport protocols.**

## I. Introduction

NEW television systems have been developed to be compatible with new technologies, either for better user experience or better broadcasting conditions. From black and white to color systems, from analog to digital, allowing better video and audio quality, error detection and correction methods, power and spectrum efficiency and different antenna configurations.

For the next television generation, some of the expected improvements are the transmission in Ultra High Definition (UHD), immersive audio, internet compatibility, hybrid services, personalized content and interactivity. For that, improvements on the transport layer were needed.

The Moving Pictures Expert Group (MPEG), has developed several standards for multimedia content delivery, such as the MPEG-2 Transport Stream (TS), which was adopted in many broadcast systems, including the Advanced Television Systems Committee 1.0 (ATSC 1.0), the Digital Video Broadcasting (DVB), and the Integrated Services Digital Broadcasting - Terrestrial (ISDB-T/Tb), some of which are still in operation [1]. Even though MPEG-2 TS has several characteristics that made it appropriate for broadcasting, the popularization of the Internet presented complications for this standard, with the interest in new technologies such as streaming services, Internet Protocol Television (IPTV) and hybrid systems. Thus, it was observed the need to carry out new studies that could result in the proposal of new transport protocols [2].

So, in 2009, MPEG hosted workshops to identify the requirements introduced by the new technologies and change of service environments. One of the major observations was that some features from MPEG-2 TS are still convenient for the next generation of television systems and therefore should be maintained or upgraded. However, other features should be simplified and adapted to consider the use of Internet Protocol (IP) as its delivery protocol, also requiring compatibility with Hypertext Markup Language 5 (HTML5). This led to the development of MPEG Multimedia Transport (MMT), which was published in 2014 as part 1 of ISO/IEC 23008 [3], [4]. MMT was adopted by the ATSC 3.0 [5] and the Integrated Services Digital Services for Satellite, 3rd generation (ISDB-S3) [6].

In the same year, MPEG noticed a market demand for a new standard for streaming services, that is, for broadband use. The reason for that was that the Real-time Transport Protocol (RTP) started to face difficulties, because it was not supported by many Content Delivery Networks (CDN), which replaced managed IP networks. Other issues were being blocked by firewalls and not being friendly for large-scale usage, since it required an individual session for each client. Hypertext Transfer Protocol (HTTP) solves these situations [7].

This led to a Call for Proposals from MPEG for a streaming

This work was supported in part by the National Counsel of Technological and Scientific Development (CNPq) and Grupo Globo.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

standard based on HTTP, which resulted as the MPEG-Dynamic Adaptive Streaming over HTTP (DASH) standard. It was published as ISO/IEC 23009 in 2012 and became very popular, being used on YouTube and Netflix, for example [8].

However, HTTP is not appropriate for broadcast delivery, so a new transport method was elaborated to deliver DASH on broadcast channels, which is called Real-time Object delivery over Unidirectional Transport (ROUTE), a work in progress by the Internet Engineering Task Force (IETF). ROUTE has improved certain features from File Delivery over Unidirectional Transport (FLUTE), such as more flexibility to deliver multiple streams at the same time, making ROUTE/DASH a logical extension from FLUTE/DASH [9].

The ATSC 3.0 meets the requirements for the next generation of television systems by combining modern technologies for all its layers. For the transport layer, it adopted both MMT and ROUTE/DASH, making it possible to compare these standards using the same setup, considering the usage rules for this standard.

For this paper, a complete setup for the ATSC 3.0 allowed the analysis of the standards mentioned above. With this practical approach, it was possible to observe the characteristics of each standard on a transport analyzer, considering the usage rules for the ATSC 3.0.

This paper is organized in six sections. Sections II, III and IV give a brief overview on the ATSC 3.0, MMT and ROUTE/DASH, respectively. Section V explains the setup used and contains the analysis made. Section VI presents the conclusion of this work.

## II. ATSC 3.0

The ATSC 3.0 has been developed to meet the requirements for the next generation of broadcast systems. It has been designed in a layered architecture, making flexibility and extensibility easier. The three layers that define the system are the Physical layer, the Management and Protocols layer, and the Application and Presentation layer [10].

The physical layer has been built with some the latest technologies available at the moment, such as Low-Density Parity Check (LDPC) codes, modulation with non-uniform constellations (NUCs), Layered Division Multiplexing (LDM), Multiple-Input Single-Output (MISO) and Multiple-Input Multiple-Output (MIMO) frame types and channel bonding. The broadcaster can set different configurations that can suit for several scenarios, aiming for more robustness or higher bitrates [8].

The Management and Protocols layer groups two system layers, becoming the transport layer of ATSC 3.0, which is based on IP and is responsible for the delivery of services and content over broadcast, broadband and hybrid networks. Signalling, media synchronization and application-layer forward error correction (AL-FEC) are also included in this layer. The methods specified for service delivery are the MMT and the MPEG-DASH. MMT delivers Media Processing Units (MPUs) using the MMT protocol (MMTP), while MPEG-DASH delivers DASH Segments using the ROUTE protocol. ROUTE is also used to deliver content that does not require real-time rendering. Both MMTP and/or ROUTE can be used for signalling delivery, and the Service List Table (SLT) provides the Bootstrap Signalling information. DASH is also used on the broadband side for the delivery of hybrid services. Figure 1 shows a conceptual model for the system.

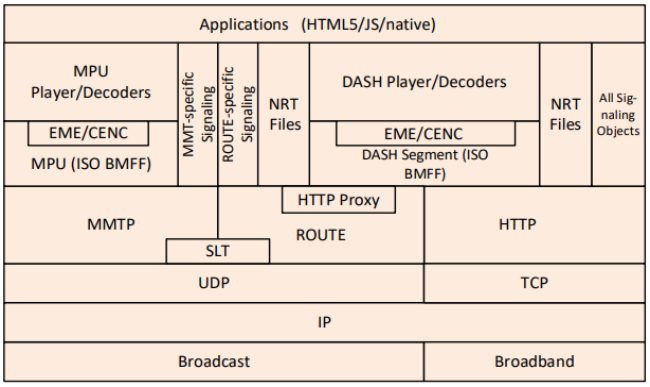



Fig. 1. Conceptual protocol stack for ATSC 3.0 (ATSC A/331:2020).

ROUTE and MMPT usage shall follow the rules:

- ROUTE and MMTP sessions shall not be used simultaneously for a broadcast delivery of a Linear Service without an app-based feature.
- At least one ROUTE session shall be used for a broadcast delivery of a Linear Service with an app-based feature (additional ROUTE and MMTP sessions are allowed, except for streaming media components in the same service).
- Only ROUTE sessions shall be used for a broadcast delivery of an app-based service, an Electronic Service Guide (ESG) Service, an Emergency Alert (EA) Service, or a Digital Rights Management (DRM) Data Service [9].

These rules consider the following terms: Linear Service, App-Based Feature and App-Based Service.

A Linear Service is a service that consists of at least one continuous video component associated with at least one continuous audio component which is then associated with at least one closed caption component and may contain app-based features.

An App-Based Feature is an application that can be directed to take individual actions on specific times using optional files and notifications.

An App-Based Service is a service that consists completely of app-based features, providing interface with the service for the user.

The application and presentation layer consists of the technologies for Audio/Video (A/V) coding, interactivity, accessibility, captions, subtitles, and other services [12].

## III. MMT

As stated before, the development of MMT started because MPEG-2 TS faced complications for personalized content and hybrid systems. However, it has inherited some features from MPEG-2 TS. One of its main characteristics is the ability to multiplex multiple streams in a single stream, interleaving data packets to allow a synchronized playback.

The MMT standard [13] defines the tools to encapsulate, organize, signal and deliver media and data. Its architecture is composed of three functional areas: the MPU format, delivery and signalling, as shown by Figure 2. Depending on the needs of a particular multimedia service, only part of these tools can be used.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

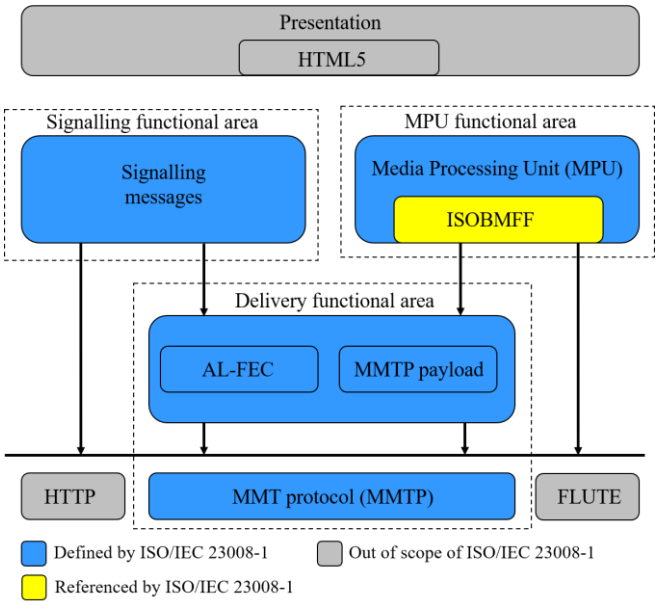



Fig. 2. MMT architecture (ISO/IEC 23008-1).

The MPU functional area consists of the logical arrangement of media, Package and the format of data units, using the ISO-based media file format (ISOBMFF). The Package provides the information about the compression on the media content and the format of data units that encapsulates the encoded media. In more detail, a Package is composed of at least one Asset and one Presentation Information (PI).

An Asset contains one or more MPUs with the same Asset identification (ID), that carries timed or non-timed media content, which can be an audio, a video or a web page, for example. Each Asset is associated with an asset delivery characteristics (ADC). A single ADC can be associated with several Assets, but each Asset can only be associated with one ADC. Also, MPUs of timed content contained in the same Asset shall have distinct presentation times.

The PI is a document that specifies how Assets are related for consumption, with freedom for spatial and temporal arrangements, thus providing an improvement from MPEG-2 TS, which assumed a single video and audio synchronized from the beginning in a fixed arrangement. An example of a PI document is the combination of HTML5 and composition information (CI) documents. Another possibility is to use a media presentation description (MPD), which is a manifest that describes the content and its characteristics [4], [7].

Figure 3 presents an overview for the MMT Package.

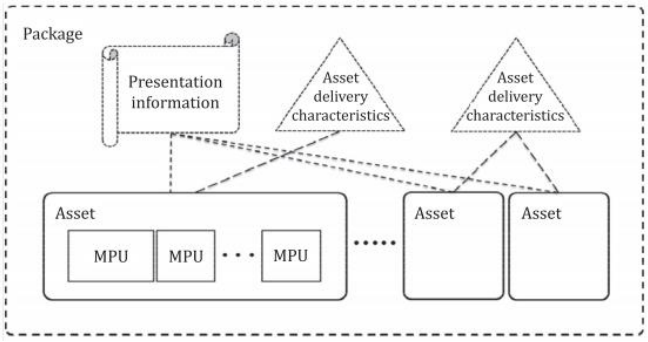



Fig. 3. MMT Package overview (ISO/IEC 23008-1).

The delivery functional area contains the MMTP as the application layer transport protocol and the MMTP payload, which is the payload format and is agnostic to media codecs, therefore supporting new codecs by requiring signalling messages that identify them.

The signalling functional area manages the consumption of media data from the Package and the delivery for the MMTP and the MMTP payload.

## IV. ROUTE/DASH

MPEG-DASH is based on a client-server model to allow adaptive streaming. A HTTP server contains a MPD and the multimedia content that consists of multiple A/V segments. The MPD is then delivered to a DASH client, which becomes aware of the content available and its characteristics. After analyzing such information, the DASH client can start streaming any chosen content, being able to change to a different configuration on the next available switching point. The system constantly monitors the available bandwidth. Whenever the bandwidth is lower than the current bitrate being streamed, it switches to a lower bitrate at the next available switching point. It can also switch to another bitrate or a different audio content if the user decides to do so.

Figure 4 presents the architecture for MPEG-DASH, which only defines the red blocks. The delivery of the MPD, the encoding format, the heuristics control and the media player are out of the standard's scope [7].

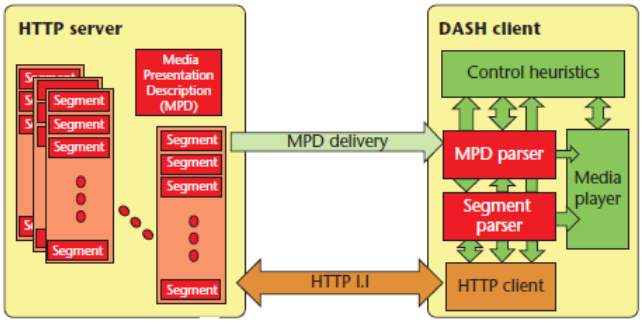



Fig. 4. MPEG-DASH architecture [7].

ROUTE is a delivery method that supports AL-FEC and can be combined with MPEG-DASH to deliver it over broadcast. ROUTE consists of two major components, a source protocol and a repair protocol. The source protocol is used to deliver objects and is aligned with FLUTE but has been optimized for the delivery of real-time media. The repair protocol protects one or multiple objects that are delivered through the source protocol. Source and repair protocols are independent, meaning that the source protocol can be used alone. Using only the repair protocol is also allowed and could be used on specific scenarios where it is expected that all receivers are AL-FEC capable, such as the delivery exclusively for mobile reception and the delivery for targeted regions. Since all packets in ROUTE are Layered Coding Transport (LCT) packets, which are defined in RFC 5651 [14], source and repair packets can be differentiated by the use of distinct ROUTE sessions, distinct LCT channels that use different Transport Session Identifiers (TSIs) values in the LCT header, and the Protocol-Specific Indication (PSI) bits in the LCT when both source and repair protocols are carried in the same LCT channel [9].

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

## V. Setup and Analysis

To test MMT and ROUTE/DASH on the transport layer of ATSC 3.0, the setup presented on Figure 5 was used, where two services are being transmitted at the same time, one with MMTP (identified with Service ID 6001) and the other with ROUTE (identified with Service ID 6002). The description of the equipment is provided below.

The A/V source delivers a content to the encoder, which accepts the following input types: IP, Serial Digital Interface (SDI), SDI 4K, Real Time Messaging Protocol (RTMP), Receiver Demodulator, Zixi, Society of Motion Picture and Television Engineers (SMPTE) 2022-6 and SMPTE-2110. For this setup, a source with the SDI format was used.

The encoder output forwards the A/V encoded stream to the multiplexer (MUX) in the Web Distributed Authoring and Versioning (WebDAV) format, which is an extension to HTTP for the execution of content authoring operations remotely [16]. In this stage, signalling information is added and the signal is forwarded to the scheduler (in our setup, the MUX and the scheduler are integrated). The scheduler adds transmitting data in the stream for the modulator, with the intention that the receiver finds the desired services, which is a function of the Service Layer Signalling (SLS) table. The Low Level Signalling (LLS) table is also added, in order to improve the channel scan speed [17].

The scheduler feeds the modulator with Studio-to-Transmitter Link (STL) packets, which is an interface between the transport and physical layers [18]. So, the modulator can convert the STL stream in a Radio Frequency (RF) channel, considering the modulation parameters configured.

The transport analyzer can monitor the STL stream from the scheduler or the MMTP and ROUTE output from the MUX, where it is possible to check several information such as the received tables, codec information, bit rate, resolution, frame rate and the audio and video synchronization.

Figure 6 presents part of the signalling configuration of the MUX for the MMT service, showing information about the Assets contained. The *Asset type* specifies the type of the Asset using a four-character code (4CC), according to the MPEG-4 Registration Authority (MP4RA) [19]. In this setup Asset #1 was set as High Efficiency Video Coding (HEVC), Asset #2 was set as Advanced Audio Coding (AAC), and Asset #3 was set as MPEG-4 Audio (MP4A). The *Asset Schema* specifies the identification mechanism and either Universally Unique Identifier (UUID) or Uniform Resource Identifier (URI) can be used for MMT. For this setup, all Assets were set as UUIDs. The *Default Asset Flag* indicates if an Asset should be a default Asset [13].

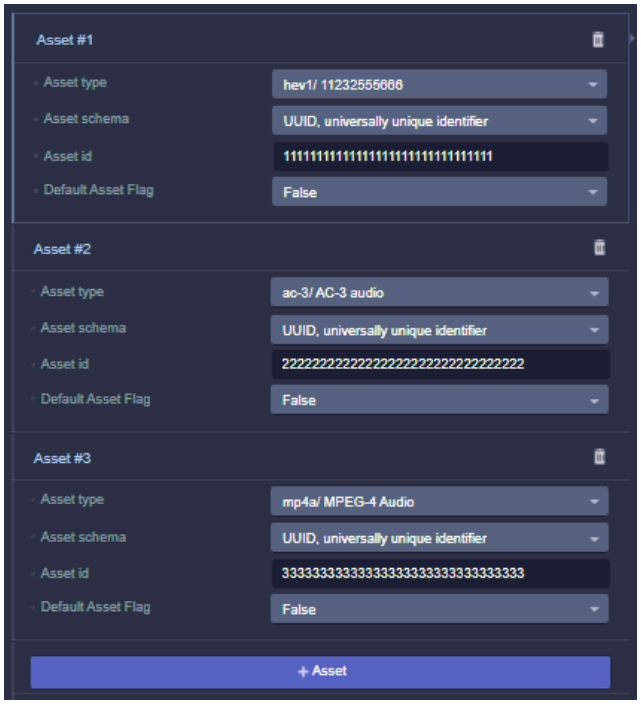



Fig. 6. MUX configuration for the Assets of the MMTP service.

Figures 7 and 8 show the respective configuration for the ROUTE/DASH service, with information of the Route Sessions (RSs) and LCT Sessions (LSs). For the configuration of the Route Sessions, the *sIpAddr* sets the IP Address of the source, the *dIpAddr* sets the IP Address of the destination, and the *dPort* sets the destination port. For the LCT Sessions, the TSI must have unique values for each Session. The *BW* value can set the maximum bandwidth required by the LCT channel [5].

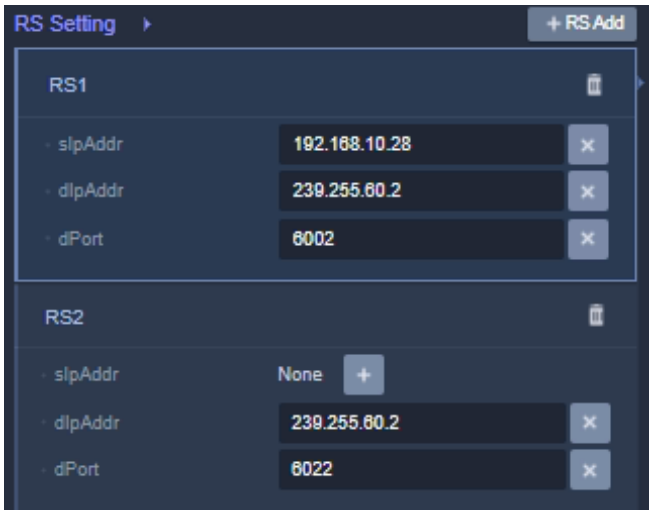



Fig. 7. MUX configuration for the RSs of the ROUTE service.

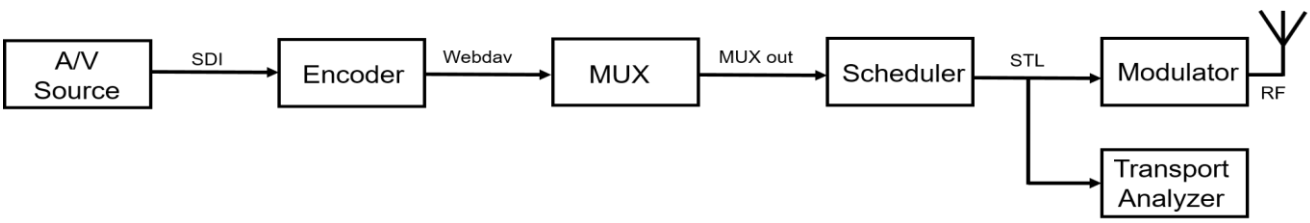



Fig. 5. Setup used to test both MMT and ROUTE/DASH delivery methods.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

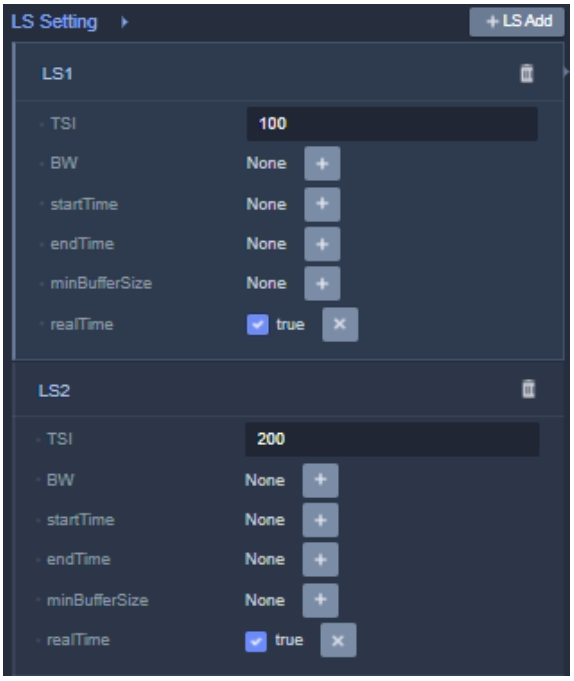



Fig. 8. MUX configuration for the LSs of the ROUTE service.

Figure 9 shows the STL information obtained from the transport analyzer, where the parameters for the Physical layer can be observed. For this setup, the Fast Fourier Transform (FFT) size is 16K, the Guard Interval (GI) is 7/2048, the modulation order is 64QAM-NUC, FEC is 10/15 and LDM was not used.

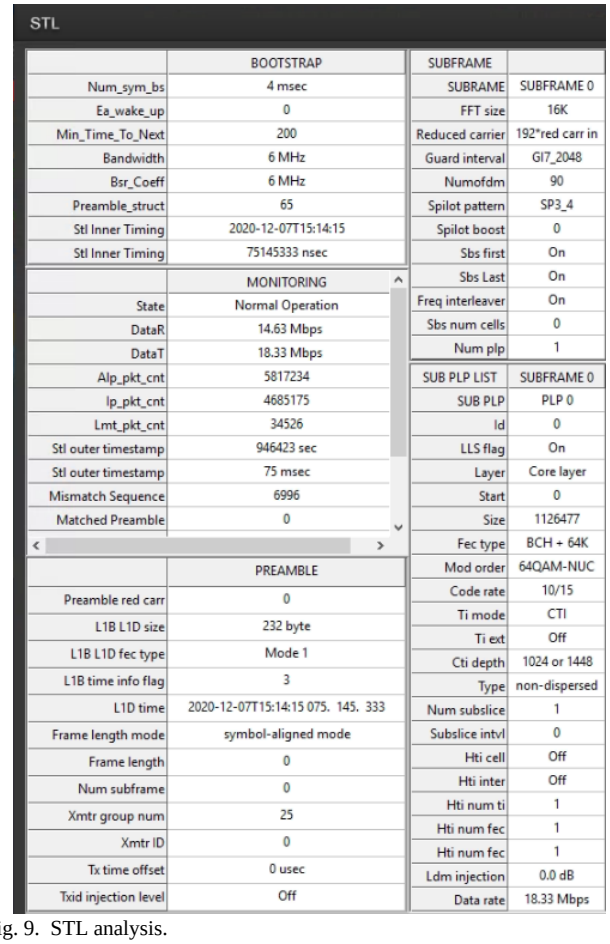



Fig. 9. STL analysis.

Figure 10 shows the input stream info obtained from the monitor of the transport analyzer, where it can be observed that for this setup, the audio was coded using Advanced AAC at 128Kbps, the video was coded using HEVC at 2.35Mbps. Since the transport analyzer is constantly monitoring the streams, the Total Rate of the streams is different because the images were obtained at different moments, however, the average rate for both streams was around 3.00Mbps. The Signal Rate for each stream also changed during time but not significantly, meaning that the Signal Rate for the ROUTE service was higher.

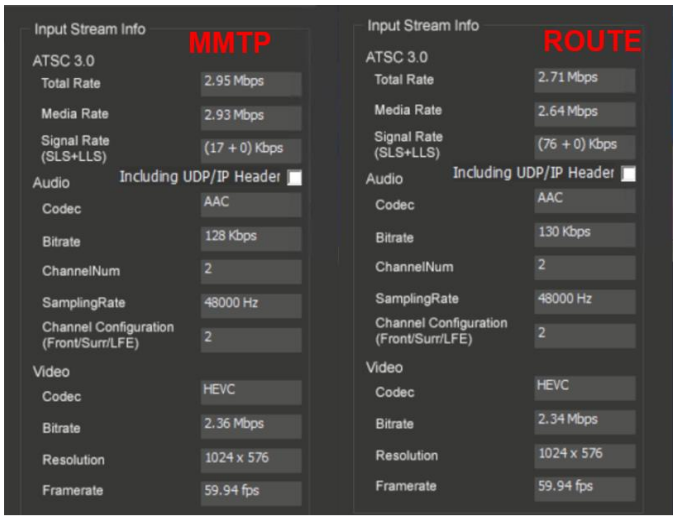



Fig. 10. Input Stream Info from the transport analyzer for both services.

In the setup used in this paper, only the Service List Table (SLT) and *SystemTime* signalling are transmitted inside of the LLS, as shown in the Figure 11. The *Version* is incremented by 1 when a change occurs on the table identified.

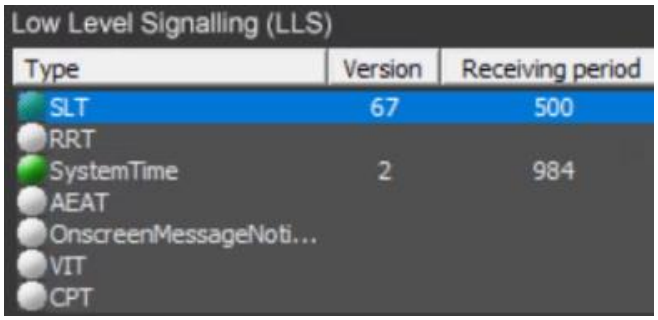



Fig. 11. LLS Signalling information.

The SLT provides the description of the service, identifying the used protocol, major channel (physical channel), minor channel (virtual channel) and SLS destiny information. It shall be repeated in the LLS at least every 5 seconds, however, repeating it more frequently (preferably every second, but not more than that) reduces the channel scan time. Figure 12 presents the content of SLT. It is possible to note that the *slsProtocol* value is 1 for ROUTE and 2 for MMTP, as determined by the ATSC 3.0 standard [5].

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

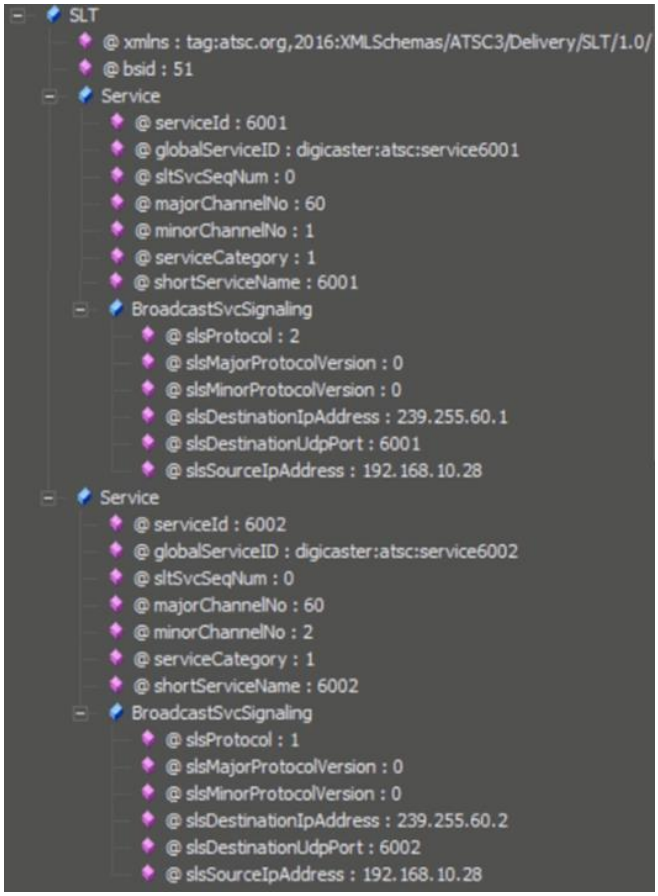



Fig. 12. SLT XML (Extensible Markup Language).

In SystemTime signalling, *currentUtcOffset* contains the number of leap seconds, *utcLocalOffset* indicates the offset between Coordinated Universal Time (UTC) and the originating broadcast station's time zone, and *dsStatus* indicates whether the transmitter location is currently on daylight saving time or not [5].

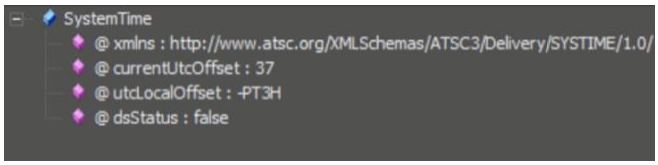



Fig. 13. SystemTime XML.

The SLS from Service 6001 (MMT) is shown in the Figure 14. Is it composed by the User Service Bundle Description (USBD) and the Complete MMT Package Table (MPT).


Fig. 14. SLS Signalling information – Service 6001 (MMTP).

The USBD links the MPU components to allow audio and video playback. The Figure 15 presents the USBD information.

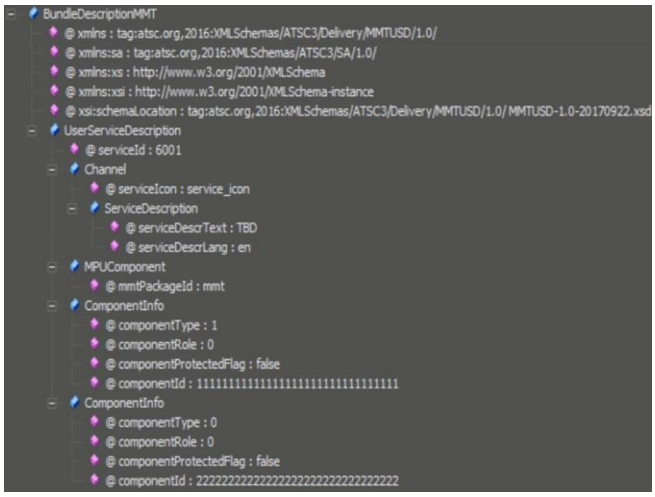



Fig. 15. USBD – Service 6001 (MMTP).

The Complete MPT contains the information in order to permit the consumption of the received packages and is presented is Figure 16.

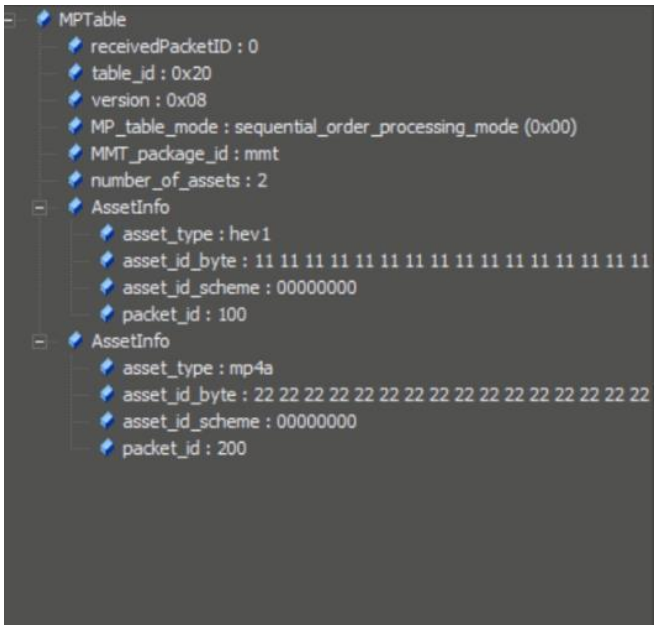



Fig. 16. Complete MPT – Service 6001 (MMTP).

The SLS from Service 6002 (ROUTE/DASH) is shown in Figure 17. This signalling is composed by the USBD, the Service-based Transport Session Instance Description (S-TSID) and the Media Presentation Description (MPD).

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

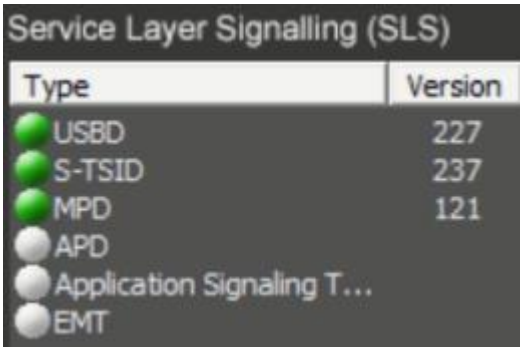



Fig. 17. SLS Signalling information – Service 6002 (ROUTE).

Figure 18 presents the USBD for service 6002 (ROUTE/DASH), in which *BasePattern* contains a character pattern to be used by the ATSC 3.0 receiver that matches with any portion of the Segment URL (Uniform Resource Locator) used by the DASH client.

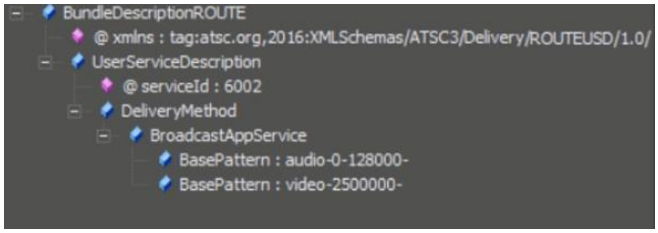



Fig. 18. USBD – Service 6002 (ROUTE).

Figure 19 shows the S-TSID information. This signalling lists the RSs and LSs, where the *SrcFlow* contain elements that provide information about the A/V flow from service 6002.

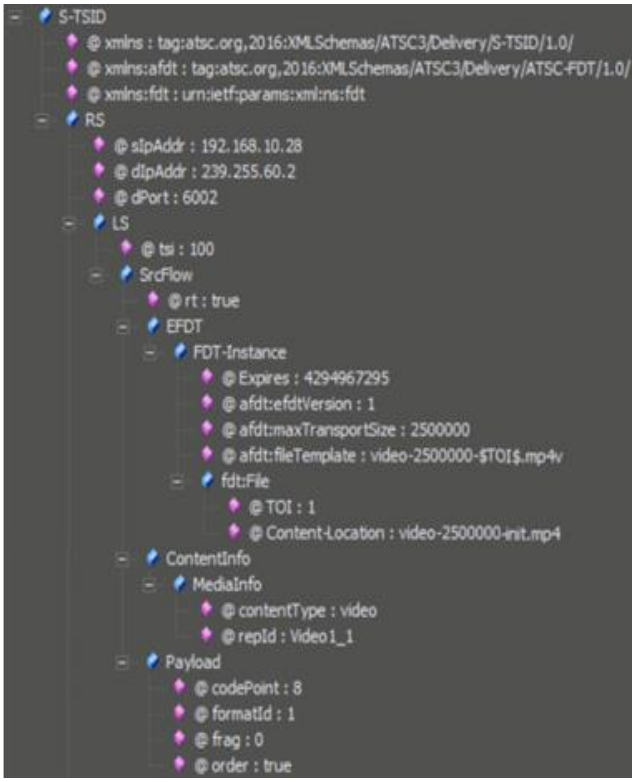



Fig. 19. S-TSID – Service 6002 (ROUTE).

Figure 20 shows the MPD table for service 6002. This signalling presents general media information.

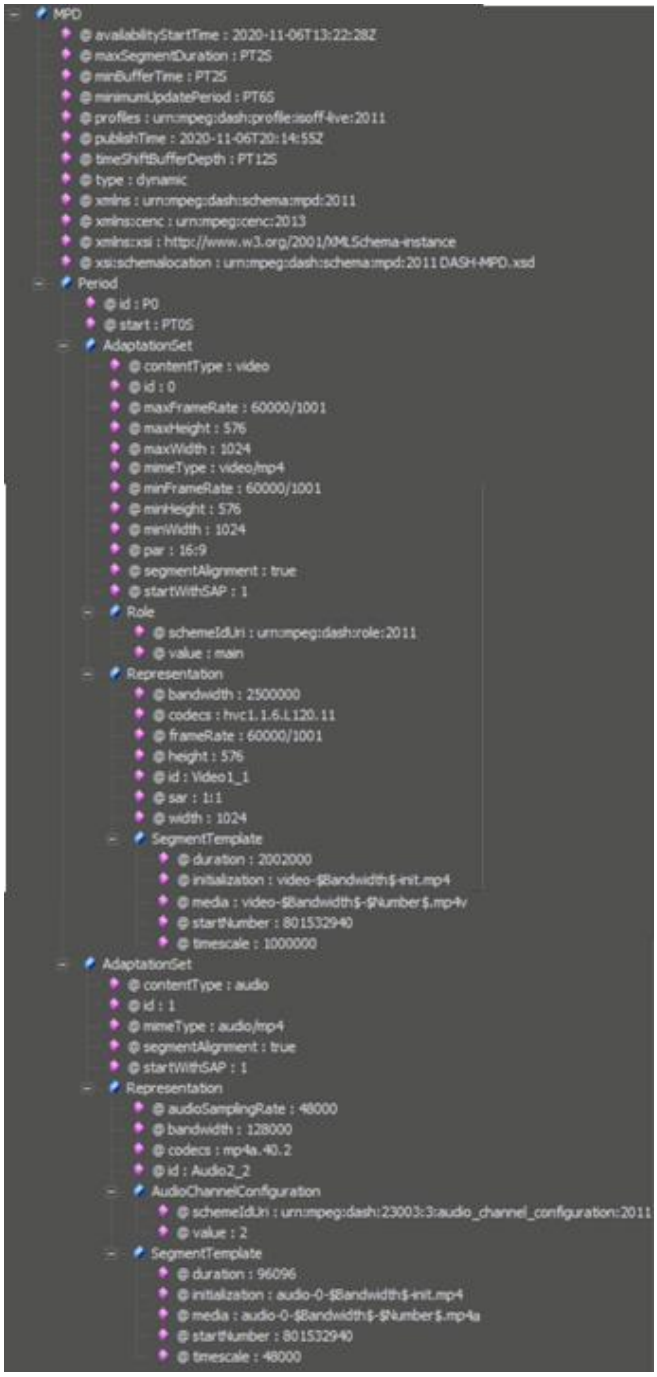



Fig. 20. MPD – Service 6002 (ROUTE).

## VI. Conclusion

From a conception point of view, it is possible to observe that these standards were initially developed for different purposes. MMT was developed due to new broadcast systems that needed IP connection to work with hybrid services, while DASH itself was designed for modern broadband applications. Then, ROUTE was developed to enable the delivery of DASH over broadcast, and therefore making ROUTE/DASH also compatible with hybrid systems. This resulted in distinct architectures.

Both methods are compatible with ISOBMFF and AL-FEC, and the delivery protocol used in each service was signalled in *slsProtocol*. Also, they are both agnostic to media codecs, hence allowing future extensions.

MPEG-DASH uses MPD for content description, while MMT can also use a combination of CI and HTML5.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

For the setup configuration used for this paper, ROUTE presented a larger overhead (since the Signal Rate shown in Figure 9 was larger for ROUTE).

For this configuration, the stream from the ROUTE service was delayed in relation to the MMTP service.

MPEG-DASH has been vastly deployed, due to its popularity for broadband applications, so there are many more tools and equipment compatible with it.

For future works, some possibilities are to test with other A/V sources, and for different configurations, such as targeted advertising, multi-view, multiple audio streams, and with additional services like the emergency alert. This would provide a more complete comparison of these methods, and possibly allow to understand if the delay observed in this paper is constant. A more complete analysis of each protocol could also contribute for a better comprehension of their characteristics. And, as a final opportunity, different setups could be elaborated to observe other aspects, such as measuring the latency of these systems for the scenarios mentioned above.

## ACKNOWLEDGMENT

The authors would like to thank PPGEEC and their colleagues at Mackenzie's Digital TV Research Laboratory.

## REFERENCES

[1] K. Park, Y. Lim and D. Y; Suh, "Delivery of ATSC 3.0 Services With MPEG Media Transport Standard Considering Redistribution in MPEG-2 TS Format," IEEE Transactions on Broadcasting, vol. 62, no. 1, March 2016.

[2] C. Diniz, C. Akamine and G. d. M. Valeira, "Development of MMT Analyzer for ATSC 3.0", 2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1-5, 2019.

[3] Y. Lim, "MMT, new alternative to MPEG-2 TS and RTP," 2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1-5, 2013.

[4] Y. Lim, S. Aoki, I. Bouazizi and J. Song, "New MPEG Transport Standard for Next Generation Hybrid Broadcasting," IEEE Transactions on Broadcasting, vol. 60, no. 2, June 2014.

[5] ATSC, "ATSC Standard: Signaling, Delivery, Synchronization, and Error Protection with Amendment No. 1", Doc. A/331:2020, 27 October 2020.

[6] ITU-R, "Satellite transmissions for UHDTV satellite broadcasting", Report ITU-R BO.2397-0, October 2016.

[7] I. Sodagar, "The MPEG-DASH Standard", IEEE Multimedia, vol. 18, no. 4, 2011, pp. 62-67.

[8] Bitmovin, "Why YouTube & Netflix use MPEG-DASH in HTML5" [Online]. Available at <https://bitmovin.com/status-mpeg-dash-today-youtube-netflix-use-html5-beyond/>. Accessed on November 2020.

[9] G. K. Walker, T. Stockhammer, G. Mandyam, Y. K. Wang and C. Lo, "ROUTE/DASH IP Streaming-Based System for Delivery of Broadcast, Broadband and Hybrid Services", IEEE Transactions on Broadcasting, vol. 62, no. 1, pp. 328-337, March 2016.

[10] ATSC, "ATSC Standard: A/300:2020, ATSC 3.0 System", 15 May 2020.

[11] L. Fay, L. Michael, D. Gomez-Barquero, N. Ammar and M. W. Caldwell, "An Overview of the ATSC 3.0 Physical Layer Specification", IEEE Transactions on Broadcasting, vol. 62, no. 1, March 2016.

[12] R. Chernock, D. Gomez-Barquero, J. Whitaker, S. Park and Y. Wu, "ATSC 3.0 Next Generation Digital TV Standard - An Overview and Preview of the Issue", IEEE Transactions on Broadcasting, vol. 62, no.1, March 2016.

[13] Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 1: MPEG media transport (MMT), ISO/IEC 23008-1:2017, August 2017.

[14] IETF, "RFC 5651: Layered Coding Transport (LCT) Building Block", October 2009. [Online]. Available at <https://tools.ietf.org/html/rfc5651>. Accessed on November 2020.

[15] ATEME, "TITAN Live User Manual", Release 4.1.22, June 2020.

[16] IETF, "RFC 4918: HTTP Extensions for Web Distributed Authoring and Versioning (WebDAV)", June 2007. [Online]. Available at <https://tools.ietf.org/html/rfc4918 >. Accessed on November 2020.

[17] ATSC, "ATSC Recommended Practice: Techniques for Signaling, Delivery, and Synchronization, with Amendment No. 1", Doc. A/351:2019 20 August 2019, Amendment No. 1 approved 20 January 2020.

[18] ATSC, "ATSC Standard: Scheduler/Studio to Transmitter Link", Doc. A/324:2018 5 January 2018.

[19] MP4RA, Official Registration Authority for the ISOBMFF family of standards. Available at < http://mp4ra.org/>. Accessed on November 2020.

**Allan Seiti Sassaqui Chaubet** was born in São Paulo, in 1994. He received the B.Sc. degree in electrical engineering in 2019 from Mackenzie Presbyterian University, São Paulo, Brazil, with a sandwich period at Rose-Hulman Institute of Technology, Terre Haute, IN. He is currently pursuing the Ph.D. degree in electrical engineering at Mackenzie Presbyterian University.

He was an intern at the Digital TV Research Laboratory of Mackenzie Presbyterian University, where he later participated in two development projects. His research interests include broadcast television with emphasis on the transport layer.

**George Henrique Maranhão Garcia de Oliveira** received the B.Sc. and M.Sc. degrees in electrical engineering from Mackenzie Presbyterian University, São Paulo, Brazil, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at Mackenzie Presbyterian University.

He has experience in electrical engineering with emphasis on characterization of ISDB-T receivers, testing and configuration of SFN networks and LTE interference tests on digital TV. He also has experience in Electrical, Magnetic and Electronic Measurements. His fields of study are broadcasting, Error correction codes and Software Defined Radio.

**Gustavo de Melo Valeira** received his B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering from Mackenzie Presbyterian University, São Paulo, Brazil, in 2007, 2010 and 2015, respectively.

He has been a researcher in the Digital TV Research Laboratory at Mackenzie Presbyterian University since 2008 and his primary research area is on FPGA and embedded processors programming for digital television. He also had the opportunity to test, work and study the ISDB-T system, including modulation, multiplexing/re-multiplexing and BTS compressor/decompressor parts, since the first transmission of ISDB-T in Brazil. His research interests include broadcast television and FPGA programming.

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

**Cristiano Akamine** received a Ph.D. degree in electrical engineering from the State University of Campinas, Brazil, in 2011.

He is a Professor at Mackenzie Presbyterian University, where he is a Coordinator of the Digital TV Research Laboratory. He is a member of the Board of the Brazilian Digital Terrestrial Television Forum and Society of Brazilian Broadcast Engineers (SET). He works in the ISDB-TB broadcasting standardization and holds several patents, licensing of intellectual property, numerous articles published. He has also served as a reviewer for several periodicals and conferences. He has participated as a Guest Editor in the Special Issue Point-to-Multipoint Communications and Broadcasting in 5G of IEEE Communications Magazine and the Special Issue on 5G for Broadband Multimedia Systems and Broadcasting of IEEE Transactions on Broadcasting. His research currently focuses on digital terrestrial broadcasting, software-defined radio, channel codes, embedded systems, and 5G.

Received in 2020-11-06 | Approved in 2020-11-16

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.5 Web Link: http://dx.doi.org/10.18580/setijbe.2020.5

# BemTV: Hybrid CDN/Peer-to-Peer Architecture for Live Video Distribution over the Internet

Flávio Ribeiro Nogueira Barbosa
and Guido Lemos de Souza Filho

**CITE THIS ARTICLE**

Barbosa, Flávio Ribeiro Nogueira; de Souza Filho, Guido Lemos; 2020. BemTV: Hybrid CDN/Peer-to-Peer Architecture for Live Video Distribution over the Internet. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2020.6. Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

**COPYRIGHT** This work is made available under the Creative Commons - 4.0 International License. Reproduction in whole or in part is permitted provided the source is acknowledged.

# BemTV: Hybrid CDN/Peer-to-Peer Architecture for Live Video Distribution over the Internet

Flávio Ribeiro Nogueira Barbosa, Guido Lemos de Souza Filho

***Abstract* — Assuming that video streaming is now responsible for the absolute majority of the Internet traffic and considering that the audience uses WebRTC-enabled web browsers and mobile devices to access and retrieve content, this work proposes the development of a peer-to-peer overlay network to assist the delivery of video streaming events that use HTTP-based protocols without the need to install additional software. Using the peer-to-peer network, the client/server model becomes hybrid, where network nodes that are watching the same event can retrieve portions of the video content directly from the server or neighboring nodes. This approach has two main objectives; decrease the client/server traffic and consequently the economic cost of delivery while improving the quality of the users' experience, given that communication between neighboring nodes can support the flow of better quality videos between the points**

***Index Terms* — peer-to-peer, webrtc, http live streaming, HLS.**

## I. Introduction

The growth of Internet use, followed by the increase of the preference for multimedia over text-based content consumption by the users in major portals, social networks, and video streaming services, supports the fact that the distribution of videos is now responsible for the majority of the data traffic on the Internet. In some countries, such as the United States, video consumption already exceeds 60% of all traffic on the network, and it is estimated that in 2022, 82% of all IP traffic in the world will be dedicated to the transmission of video content, 15 times higher than the video traffic in 2017.

The latest live events broadcast on the Internet have beat audience records, and it is expected that this phenomenon will continue to happen. For example, CBS' NFL Super Bowl 53 internet streaming for home users in the United States in 2019 reached 7.5 million unique devices. The users consumed more than 560 million minutes of video, increasing more than 20% when compared to the previous year, as shown in Fig 1. As another example, the FIFA Football World Cup in 2018 also set a new audience record for online streaming in Brazil. It peaked at more than 1.2 million simultaneous users, presented in Figure 2.

The high demand for bandwidth, quality of service (QoS), and high traffic inherent in video distribution applications have brought significant challenges to today's Internet. The implementation of large-scale infrastructures, the cost of transmission, and the QoS became a central issue for the industry. It is also observed that, as video consumption increases, rebuffering problems increase and users' session time decreases.

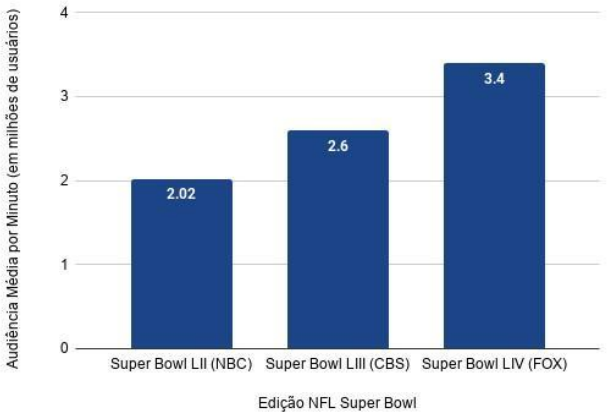



Fig 1: Average audience growth per minute between editions of the NFL Super Bowl.

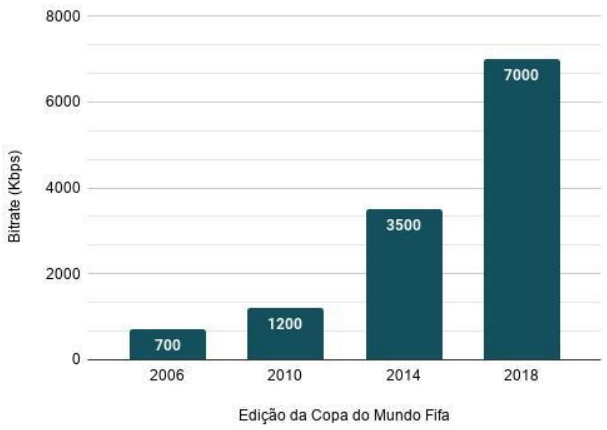



Fig. 2: Quality growth (bitrate) between editions of the World Cup, Grupo Globo.

Typically, large companies that stream video to hundreds, thousands, or even millions of people, use one or multiple *Content Delivery Networks* (*CDN*) to meet public demand. The main objective of *CDN* 's is to distribute content through sets of agglomerated servers called points of presence or *PoP*'s that are usually distributed geographically to guarantee a reliable, scalable, and efficient delivery to end-users [1]. This approach has some disadvantages:

- Scalability. CDN's use PoP's for delivery and load

F. R. N. Barbosa - Video Technology Group, CBS Interactive , New York, NY - United States (e-mail: flavio.ribeiro@cbsinteractive.com).

G. L. Souza Filho - LAVID, Universidade Federal da Paraíba (João Pessoa, PB - Brazil (e-mail: guido@lalavid.ufpb.br).

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

balancing. When a user is geographically far from the PoP, she depends on the Internet's communication channels and exchange points (PTT) to retrieve the requested content. In countries with poor telecommunications networks and slow exchange points, there is a need for an increase in the deployment of PoP's. In Brazil, the most popular CDN's concentrate their presence in the South and Southeast, damaging the experience of video consumers in the Midwest, North, and Northeast.

- Cost. The cost of building and managing a CDN is very high, both economically and in terms of complexity [12]. In addition to the need for specific hardware infrastructure, the point of presence needs to operate without interruption. According to [19], Google Youtube is believed to spend an average of $ 1 million a day on distribution infrastructure costs.
- Quality of Experience (QoE). The last significant event transmissions based on the traditional client/server model, using only CDN's, presented major scalability problems, making the audience's consumption experience unfeasible or degrading [27] [14]. In the context of Brazil, as shown in Fig. 3, in May 2013, users connected from the telecommunications provider Embratel in Caruaru - a city located in the Northeast of the country obtained an average of 9 2 buffer underflow events during the transmission of the semifinal of the European Champions League [5].

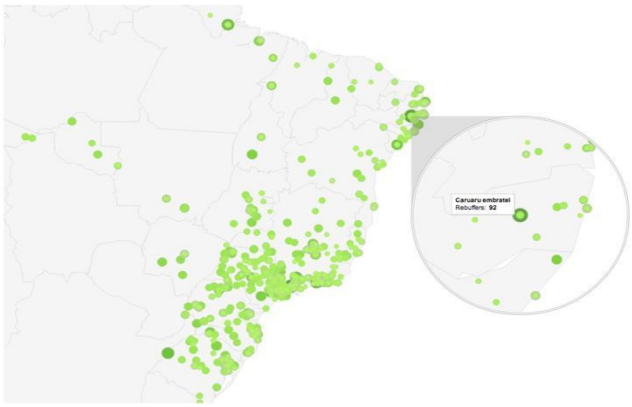



Fig. 3: User with 92 *rebuffers* events when watching video on Globo.com using Embratel provider in Caruaru, Pernambuco.

Given the considerably intensive requirements and resources imposed by streaming services in addition to the points already discussed, it can be said that traditional client/server systems do not scale sufficiently [25]. It is known that peer-to-peer architectures such as CoolStreaming, GridMedia, PPTV, PPLive, PPStream, UUSee, and TVAnts present themselves as alternatives to the problem of large-scale transmissions [8]. However, with the development of protocols for streaming video on the Internet, and migration in recent years of large companies (including Microsoft, Adobe and Apple) protocols *stateless* based on HTTP (English *Hypertext Transfer* Protocol), many state-of-the-art have become obsolete [16]. Since then, *peer-to-peer* models supporting HTTP delivery have been developed.

Assuming that the absolute majority of the audience consumes the videos in web browsers or mobile devices capable of operating and supporting traffic via WebRTC (*Web Real-Time Communications*), this work proposes the development of a *peer-to-peer* overlay network to assist delivery of video content in live streams that use HTTP-based protocols without the need to install additional software. Using the *peer-to-peer network*, the client/server model becomes a hybrid model, where network nodes that are watching the same transmission can retrieve portions of the video content directly from the server or neighboring nodes. This approach has two main objectives; decrease the client/server traffic and consequently the economic cost of delivery while improving the quality of the users' experience, given that communication between neighboring nodes can favor the flow of better quality videos between the points.

Thus, it is expected that in addition to the traffic between users and the CDN decrease, the system will become more consistent and fault-tolerant, increasing the throughput of the network, improving the user experience, and decreasing the cost of transmission as a whole.

## II. THEORETICAL FRAMEWORK

*Peer-to-peer* meshes for solving scalability problems are widely studied and used in several segments. In the context of streaming services, the transmission of video streams where there is a high need for resource availability, the traditional client / server-based system proves to be insufficient for the growing demand [27] [14]. An alternative explored by the research and industry fields in recent years focuses on using users' upload bandwidth to exchange video segments through a peer-to-peer strategy in conjunction with the traditional client/server approach. According to [3], sharing data with the help of a peer-to-peer mesh can decrease traffic between users and servers by up to 96%, promoting a drastic decrease in transmission costs and reducing bottlenecks, resulting in improving the experience of broadcast users.

It is important to note that, for the exchange of video segments to occur, a fundamental requirement for users participating in the peer-to-peer system is the capacity to store portions of the watched content, to disseminate them to the other participants.

CoolStreaming [26] is one of the first streaming systems of large-scale video using P2P technology. The system was developed using the Python programming language and applies a gossip protocol [24] for the transmission of buffer maps between nodes. A buffer map reflects the portions of the video that are stored on a given node [6]. When starting the player, it connects to an origin node to receive a set of nodes available for that transmission. Upon receiving the set, the player randomly draws a subset of nodes and promotes them to partners. Once the partners are defined, the receiver receives their buffer maps and requests the first video clips from this subset. Since the choice of partners is random, the probability that the chosen nodes are not optimal for the recipient in question is high. Strategies such as the calculation

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

of the RTT ( Round- trip time ), elapsed time for sending a message and arrival of the acknowledgment that the destination received it correctly in the handshake, the establishment of the initial connection between nodes or information about the location of the system components in the origin node could be implemented to optimize the choice of this subset.

PPlive [22] is one of P2P video streaming systems unstructured most popular today, especially in China, which serves about 3 million daily users in 300 different transmissions at a rate ranging from 250Kbps and 400kbps [6]. The system is non-structured (mesh), and the content exchange mechanism between nodes of the overlay network (overlay network) is done by breaking the streaming video into small pieces (chunks) [23]. Although widely used, PPLive has some shortcomings. The startup delay (the time between the press play until the first video frame) is in the order of 20 seconds for popular broadcasts and reaches up to a few minutes for less popular broadcasts. Another known limitation is the delay ratio between the nodes of the PPLive mesh, which reaches 150 seconds [7], degrading the experience in cases of live transmissions, such as a football game, for example. The protocol used in PPLive is proprietary, making it difficult to evolve. Still, [Li, Chen 2008] proposed a segment recovery strategy called Threshold Bipolar (TB). To conduct the study, it was necessary to reverse engineer the protocol without concrete conclusions about its improvement.

It is known that since the appearance of CoolStreaming and PPLive, several other architectures, models and peer-to-peer systems have presented themselves as alternatives to the problem of large-scale transmissions. Among the published models, we can highlight:

- P2PSP [13] uses a structured model where the nodes are classified into three layers: Broadcasting, Integrity, and MultiChannel. Broadcasting layer nodes are responsible for segmenting and adapting the video stream before being propagated over the network. The users that make up the Integrity layer manage the input and behavior of the nodes, detecting poisoned and hostile peers. Users who belong to the MultiChannel usually t is m grid conditions and resources to consume more than one stream video, serving as a relay for more than one transmission.
- mTreeBone and CorePeer [4] [Want 2007] are hybrid systems that utilize an array of unstructured networks (mesh) and structured (trees). Both classify the component nodes of the network as stable or not based on different criteria. mTreeBone considers the time assisted in distinguishing and promoting as the CorePeer analyzes the width of upload bandwidth thereof to promote them. A node, when promoted to stable, connects directly to CDN functioning as an extension of it, serving the other nodes of the mesh.
- BiToS [21] uses the sharing service network architecture files peer-to-peer best known globally, BitTorrent, to assist the flow of video segments between users. BiToS changes the order in which *chunks* are to be trafficked. Common BitTorrent uses a rarest-*first* strategy (where less popular chunks are prioritized, seeking a balance in their availability). In contrast, BiToS prioritizes chunks in an order similar to the progressive download, where the most critical segments to be sent and received are the next to be played in the video player.
- OpenSatRelaying [20] is a hybrid model where geographically distributed users can receive and repeat a broadcast signal via satellite and serve as entry points for regional IP transmission. Users who receive the IP stream also act as a relay for others. The main challenge of this approach is the synchronism between users since the transmission via satellite promotes a low delay of sending and receiving, unlike IP transmission.

With the evolution of protocols for streaming video on the Internet and the migration in the last years of large companies (including Microsoft, Adobe, and Apple) to stateless protocols based on HTTP, lots of the progress on the state of the art became obsolete [17].

### *A. Hybrid Peer-to-Peer Systems for HTTP-based Transmissions*

SmoothCache [18] system is considered the first peer-to-peer built-in support for content transmissions on-demand and live via the HTTP protocol and model your solution using the concepts caching distributed.

In the first approach of the system, classified as Baseline Caching, each node maintains a table with a subset of the network's nodes and which video clips each has. Thus, when requesting a segment that does not exist in the distributed cache, the player directs its request directly to the server at the CDN and, when retrieving the requested segment, reports to a large number of other nodes so that they update their table using the method of gossip. The construction of the connections between the nodes and their maintenance is done at random, in an unstructured scheme (mesh). It is important to note that, as previously mentioned, the scheme takes advantage of the lack of synchronism between users inherent in the implementation of the players in this protocol - making the exchange of segments between them possible.

The Improved caching the SmoothCache adds an array of features. First, in addition to the natural desynchronization between users, the model positions users with greater direct bandwidth with CDN ahead of the transmission - limiting this positioning to up to 4 seconds ahead of other users. Each node that can connect to another node is defined as a neighbor (neighbor), and from the set of neighbors are randomly chosen node partners ( partner), a subset of neighbors. Each chosen partner receives an initial score composed of the amount of data transferred successfully so far, and the current quality of the video streaming consumed in the case gave an adaptive streaming scenario. It is important to note that favoring nodes that are in the same quality as the user is dangerous: Once the node changes quality, it will be isolated. Therefore, the version Improved Cached the SmoothCache keeps its set of partners, we are in all the qualities available in a Gaussian distribution centered on the current quality of the given node.

In a multi-bitrate stream, the user's player maintains a periodic calculation of the band with the CDN - by measuring the time required to download a segment compared to its size

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

in seconds and its quality in bits per pixel (bpp) to define which correct quality of the stream to be consumed at that time. Since SmoothCache shares segments between nodes, if the desired segment is already in the local cache, the implementation adds a delay in delivery to prevent the calculation from resulting in a false positive for the quality of the band at the moment, maintaining the download time of the segment approximately equal to the time it would take if it were recovered from CDN.

The implementaçãoo the SmoothCache uses the standard Smooth -Streaming Microsoft and *player headless* that runs natively in the operating system.

The pDASH [10] is a P2P overlay proposal to assist in the delivery of video clips on-demand using the DASH (Dynamic Adaptive Streaming over HTTP) standard. Similar to other HTTP standards and drafts, a DASH video stream is segmented on the server and distributed among users. To coordinate the order of the segments, the server builds a playlist - known as the Media Presentation Description (MPD) in the DASH standard - that will be accessed by the player before requesting the segments. In pDASH, this server that builds and serves the MPD also monitors accesses and records the addresses of users who have successfully downloaded a segment. When successfully serving a stretch, the server adds the user's address to the MPD, allowing users who request the MPD after this modification to choose which source they want to retrieve that stretch from. It is important to note that this addition is allowed by the standard. For pDASH to work, clients need to run an HTTP server locally in addition to having a public IP or NAT route for external access from this server. The implementation of simulation of pDASH used parts of HTTPTools Framework [9] to the HTTP server, and parts of the DASH VLC Plugin [2] to the native video playback.

[15] presents a P2P network using APIs from WebRTC. The publication does not clarify whether the transmission is made using any HTTP protocol, but describes how the network is built. The system uses a middleware called PeerJS that abstracts the direct calls to the native API's WebRTC and maintains compatibility if the method signatures are changed. The system also uses a server to coordinate and connect incoming users. Thus, when rendering the video player, the peer generates a unique user ID and registers with the WebRTC coordinator. The coordinator, in turn, selects two other users from the same transmission and promotes the connection between them. When closing the connection, the nodes use the RTCDataChannel API to exchange transmission data. Since the coordinator randomly selects the peers, an unstructured mesh is created and the users who are in the transmission the longest are selected more often and, therefore, are more stable and maintain more connections.

## III. PROPOSED ARCHITECTURE

Based on the academic contributions and technological developments on the Internet already mentioned, and given the market scenario, where the adoption and migration to HTTP-based protocols are being made and based on the successful migration of the RTMP protocol to HTTP Live Streaming (HLS ) by Globo.com, which resulted in 33% more time watched, we believe that the future of live broadcasts on the Internet will be using stateless protocols.

Analyzing the state of the art hybrid models CDN/Peer-to-Peer, we believe that non-structured models are more effective in dealing with highly dynamic networks and an alt turnover rate (node churn), an inherent feature of online live broadcasts large-scale. With the support for peer-to-peer communication added by WebRTC, we believe it is possible to implement a model entirely within the browser, without the need to install clients or plug-ins on devices, using only JavaScript technologies, already supported in most browsers used on the Internet. Therefore, we propose a hybrid architecture model CDN/Peer-to-Peer unstructured (mesh) for live video streams using only the technologies available in modern browsers, including WebRTC where we consumers help to CDN delivery of video segments that use the HTTP Live Streaming (HLS) protocol.

To validate our research work, we decided to create two architectures. One, initial, for proof of concept, and a second where we improve the protocol with the election of partners.

The goal of the development of this proof of concept was to validate the technologies chosen for the evolution of a peer-to-peer overlay network where users who consume the video cannot have their consumption experience impaired. In this way, the work focused on a hybrid model adding what already exists in current infrastructures, enabling the evolution of a distribution platform with the adoption of the described system without offering impact or downtime on the infrastructure.

### *A. BemTV Baseline*

#### *1) Creation of the Peer-to-Peer Mesh*

For nodes to be able to establish a connection, it is necessary to a stage where the nodes become aware of each other. For BemTV Baseline, it was defined that only nodes from the same geographic region and Internet provider would meet each other. This set of nodes was given the name of a location-aware swarm (geo-aware swarm).

The node then makes an HTTP GET request to a server that will identify you based on your city and provider and will return a swarm name as shown in Fig. 4.


```
{
    asn: "AS28604 Globo Comunicação e Participaçoes SA",
    city: "Rio De Janeiro",
    ip: "186.192.87.75",
    swarm_name: "UmlvIERlIEphbmVpcm9BUzI4NjA0"
}
```


Fig. 4: HTTP Response by swarm name.

To identify the ISP and city, the request's IP is used as the input for libraries GeoIPASNum and MaxMind's GeoLiteCity. The server consists of a Python application using the Flask framework and is served by an Nginx server.

In possession of the swarm name, the node uses the RTCPeerConnection API to publish. An rtc-switchboard server receives the publication and initiates the attempt to switch via STUN ( Session Traversal Utilities for NAT ) to find a route through NAT's ( Network Address Translation Port Mapping Protocol ) where the two nodes can exchange

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

data. It is important to note that, with each new node in the swarm, a connection to each node of the swarm will be stabilized, increasing the number of commutations exponentially, that is, considering n as the number of nodes the number of connections is equal to n x ( n - 1).

*2) Requesting Video Segments*

Once the nodes already have connections with the others in the same cluster, the video player starts playback by requesting the playlist for CDN. The direct request from the player to the CDN, in this case, is not intercepted, since we want to keep the nodes updated with the live stretches, reducing the playback delay for the live event.

After receiving and interpreting the data ( *parsing* ) of the playlist, the player initializes the request to the chunks. Instead of making HTTP requests directly to the CDN, the node sends a DESIRE message to all nodes in the cluster, and each receiving node searches its local cache for the requested video segment. The implementation stores the last 10 stretches of video watched, and if the chunk is contained in the cache, the node sends back a DESACK message.

The node that wants the segment receives all DESACK's and analyzes which is the most appropriate to receive that segment. In the current implementation, the node that sends the first DESACK is chosen. Considering the node that wants the chunk to be Peer A, and the node that sent the first DESACK being Peer B, the negotiation process follows as shown in Fig. 5. A REQ message is sent from Peer A to Peer B, which in turn sends an OFFER message with the content of the chunk.

When Peer A sends the DESIRE message to the cluster, it waits 0.7 seconds until some other node sends the DESACK. If it does not, Peer A requests the chunk of the video directly n the CDN using the common HTTP scheme. The same is true for sending the REQ. If the node does not send the content of the chunk within 1 second, the node will resort to CDN.

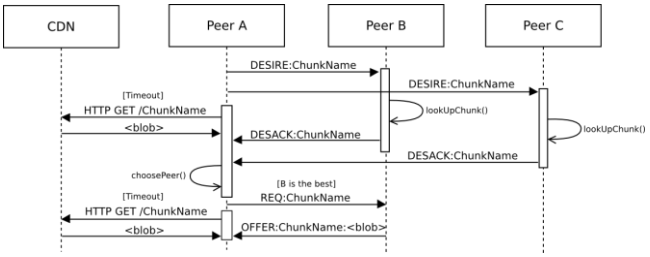



Fig. 5: BemTV Baseline Video Snapshot Trading Scheme

*3) BemTV Baseline Results*

To validate the proposed architecture, an experiment was carried out with the transmission of a live video stream lasting 60 min. 10 Apple MacBook White with 2 CPU cores and 2GB of SDRAM were used. The browser chosen was Mozilla Firefox 27.1, compatible with the latest draft of the WebRTC specification. The comparison test was also done with the same computers and with the Apple Safari 6.0.5 browser, capable of reproducing HLS streams natively.

All computers were on the same 10/100 Mbps Wireless network, which means that they were in the same geographic location and internet provider and, consequently, in the same cluster. The video stream was broken into 5-second chunks at a rate of 600 Kbps bitrate using the HTTP Live Streaming protocol. The CDN was represented by a server with 1 core and 512MB of SDRAM. Fig. 6 shows computer screens during the experiment.

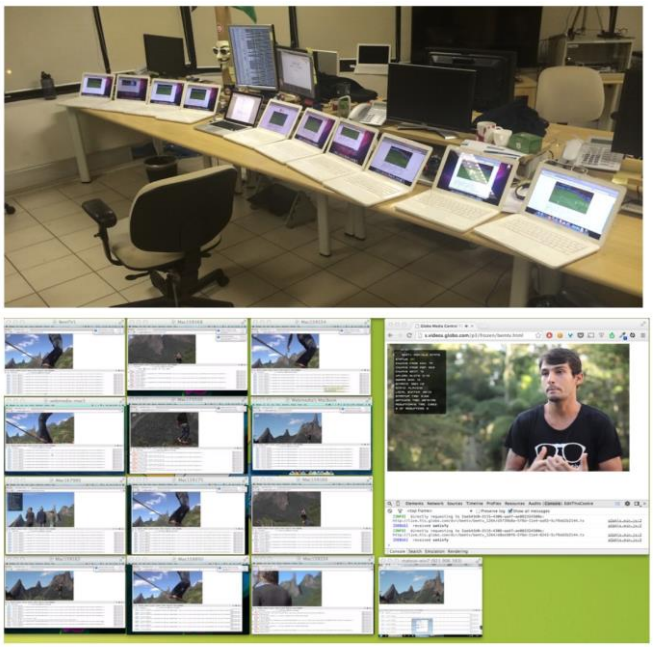


Fig. 6: BemTV Baseline Controlled Experiment

Computers were subjected to an hour of streaming using the native Apple Safari browser and then compared to a streaming consumption of time using the implementation of BemTV Baseline and were analyzed by the number of requests made directly in the CDN. The smaller the number of requests in BemTV Baseline, the more effective the peer-to-peer mesh of changing video chunks will be.

Fig. 7 shows 745 requests made by customers in an hour of experiment, of which 448 were directed to CDN and 297 were directed to BemTV peers. This result indicates promising results of the proposed strategy because, in this experiment, savings of 39.86 % of the egress traffic from CDN were found pointing to an important reduction in the cost of operation of the evaluated video distribution service.

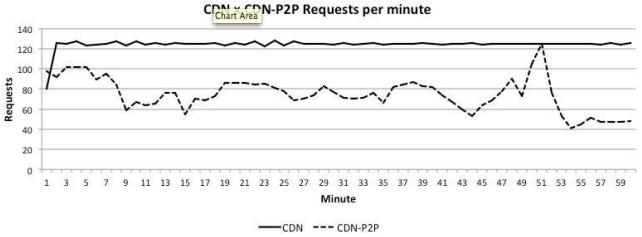



Fig. 7: Comparative Result of Direct Requests to CDN in the conventional model and BemTV Baseline.

*B. BemTV Enhanced*

With the proof of concept validated, we evaluated several protocols for the dynamic promotion of partners and created a balancing heuristic. The final idea is that each node in the

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

peer-to-peer cluster has an optimized view for its current state of connectivity with both the CDN and neighboring nodes in an unstructured (mesh) architecture.

*1) Creation of the Peer-to-Peer mesh*

A reputation-based mechanism was created, where all users of the P2P loop have different scores associated with each point-to-point connection. When a user starts their session to watch the live video, they request information from a central server that returns up to 50 other peers that are part of that cluster, based on what was described in BemTV Baseline. Upon receiving the connection information with the 50 nodes, the peer will try to establish a WebRTC DataChannel with all of them and assign an initial score of 1000.

During this process, various connectivity problems happen (traverse *NAT, firewalls, proxies),* and the final quantity of successfully established connections is significantly lower than the number of attempts, as illustrated in Fig. 8.

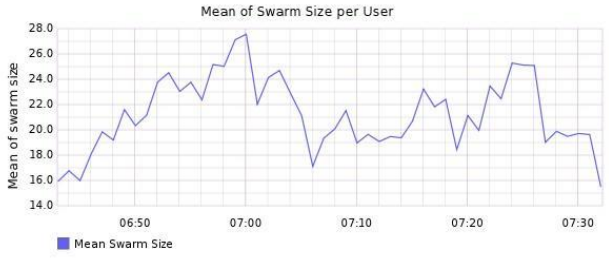



Fig. 8: Swarm size by users during the transmission of Bom Dia São Paulo, Globo.com

With the connection established with other nodes, the user sends a PING and is expected to return a PONG During this stage, the user counts the Round- Trip Delay Time (RTT) or the time it took between this initial handshake and the score to be shot The higher the RTT, the greater the impact on the user's score, which directly translates to the distance in the network topology between the two users. After exchanging PING and PONG messages the user has a picture of the current topology and promotes the top 10 scores to partners

*2) Video Segment Requisition and Scores Management*

The request to *manifest* file with the playlist of chunks video available in the partner (partners) remains similar to that used in the basic architecture ( *baseline* ) and does not add delay to the end-user, but the request of the video segments was broken into two stages

*3) Requesting Video Segments*

To request segments (chunks) of video, the user sends an INTERESTED message to the swarm, and waits for answers from the peers. If it receives CONTAIN back, *interested* and awaiting response s *shock* of partners. When you receive a message *containing the node* decrements of 1 the score of the peer and promotes the candidate partner. When it does not receive a response from the *peer*, the node removes it from the swarm.

Once the candidate with the highest score is identified, the node sends him a REQUEST message. If he receives a choke as a reply, he decreases that partner's score by 1. If, on the other hand, you receive a SATISFY message in response, it is already accompanied by a binary file with the content of the video chunk. This content is finally sent to the playback buffer and the score of the partner is incremented by 1. On the other side, and the partner does not respond it is removed from the cluster.

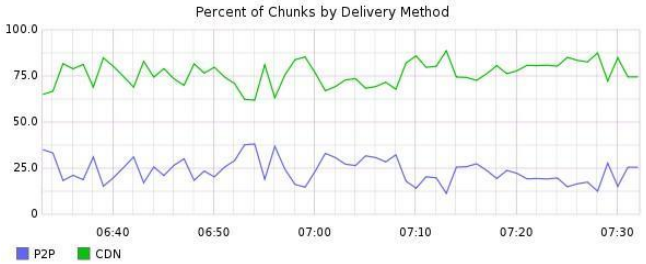
Then the partner with the highest score is promoted to sender avoiding sending INTERESTED messages to him.

Finally, with each new segment request, the array of peers is reordered promoting peers to partners.


Figure 9: BemTV Enhanced in a production environment for Bom Dia São Paulo on Globo.com.

## IV. FINAL RESULTS

The implementation of the advanced architecture was tested in a production environment at Globo.com during the exhibition of a Bom Dia SP show. In this experiment, the number of chunks trafficked in the overlay of pairs was measured, which reached the value of 19759 video chunks. The number of chunks trafficked at the CDN was also measured, resulting in a total value of 22522 video segments. These figures point to an economy of 46.7 3 % of chunks video transmitted during the experiment. Fig. 9 illustrates the screen of a browser running BemTV Enhanced code during the experiment and Fig. 10 shows the number of chunks received by users of CDN and other nodes in the BemTV overlay network.




This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6

Figure 10: Percentage of Video Segments by Delivery Method in BemTV Enhance.

## V. CONCLUSIONS AND FUTURE WORK

Based on the experiment and considering that browsers are constantly evolving, it is possible to conclude that the use of WebRTC as a basis for implementing a hybrid CDNP2P architecture to offer a live streaming video service is not only feasible but a very promising strategy. Optimizations in the data exchange protocol can increase the percentage of chunks negotiated between peers, reducing the cost of transmissions, increasing the scalability of the system, and offering a better user experience for the service consumers. However, several challenges remain open and will be addressed in future work:

- Convergence between pairs and over swarming. The approach described in this article proved to be effective for a relatively small number of users. However, building swarms that aggregate all users in a given location or connected to the same Internet access service provider can generate demands for grouping hundreds or thousands of users leading to an excessive exchange of DESIRE and DESACK messages. The use of reputation [Xiong and Liu 2004], partnership [11], or election of leader [Kutten 2013] are strategies that can be applied to address these problems.
- Video Chunks exchange protocol. With the introduction of improvements in the formation of swarms, we believe that it will be necessary to improve the protocol's stability, and the number of chunks over the P2P network tends to increase. However, research to find the ideal chunks size, video transmission rate, and node cache size will be necessary. The protocol must additionally ensure that the negotiation and transfer of chunks do not interfere with the experience of s users.
- Content security. The protocol presented in this article does not guarantee that the content exchanged between peers is identical to that distributed by CDN. Algorithms for detecting chunks adulterated by DoS starvation as presented [13] need to be analyzed and used.

## REFERENCES

[1] Bronzino, F. Gaeta, R. Grangetto, M. Pau, G. (2012) "An Adaptive Hybrid CDN/P2P Solution for Content Delivery Networks". VCIP, páginas 1-6, IEEE.

[2] C.' Müller,' C.' Timmerer,' "A' VLC' Media' Player' Plugin' enabling' Dynamic' Adaptive' Streaming' over' HTTP,"' In' Proceedings' of' the' ACM Multimedia 2011, Scottsdale, Arizona, November 28, 2011.

[3] Cho, S., Cho, J., Shin, S. (2010) "Playback Latency Reduction for Internet Live Video Services in CDN-P2P Hybrid Architecture". 2013 IEEE International Conference on Communications.

[4] Deng, H., Xu, J. (2013). CorePeer: A P2P Mechanism for Hybrid CDN-P2P Architecture. In WebAge Information Management (pp. 278-286). Springer Berlin Heidelberg.

[5] Globo.com (2013) "Rebuffers por Cidade e Operadora – Bayern x Barcelona". HYPERLINK http://goo.gl/VwqXFK, acessado em Outubro, 2014.

[6] Hei, X., Liang, C., Liang, J., Liu, Y., Ross, K. W. (2007). A measurement study of a large-scale P2P IPTV system. Multimedia, IEEE Transactions on,9(8), 1672-1687.

[7] Hei, X., Liu, Y., and Ross, K. (2008). IPTV over P2P streaming networks: the mesh-pull approach. Communications Magazine, IEEE, 46(2):86–92.

[8] Jurca, D., Chakareski, J., Wagner, J. P., Frossard, P. (2007). Enabling adaptive video streaming in P2P systems. IEEE Communications Magazine,45(EPFL-ARTICLE-98484), 108-114.

[9] K. Jonsson, "HttpTools: A Toolkit for Simulation of Web Hosts in OMNeT++", In proceedings of the 2nd OMNeT++ workshop, Rome, Italy, 2009.

[10] Lederer, S., Muller, C., Timmerer, C. (2012, May). Towards peer-assisted dynamic adaptive streaming over HTTP. In Packet Video Workshop (PV), 2012 19th International (pp. 161-166). IEEE.

[11] Li, B., Keung, G. Y., Xie, S., Liu, F., Sun, Y., Yin, H. (2008, November). An empirical study of flash crowd dynamics in a p2p-based live video streaming system. In Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE (pp. 1-5). IEEE.

[12] Lu, Z., Wang, Y., Yang, Y. R. (2012) "An analysis and comparison of CDN-P2P-hybrid content delivery system and model". journal of communications,7(3), 232-245.

[13] Medina-López, C., Naranjo, J.A.M., García-Ortiz, J. P., Casado, L. G., González-Ruiz, V. (2013) "Execution of the P2PSP protocol in parallel environments". XXIV Jornadas de Paralelismo. Madrid, Spain.

[14] Nordyke, K, (2014) "HBO Go Crashes During 'True Detective' Finale". http://goo.gl/zixUr7, acessado em março, 2014.

[15] Rhinow, F., Veloso, P. P., Puyelo, C., Barrett, S., Nuallain, E. O. (2014, January). P2P live video streaming in WebRTC. In Computer Applications and Information Systems (WCCAIS), 2014 World Congress on (pp. 1-6). IEEE.

[16] Roverso R, El-Ansary S., Hogqvist M. (2013-a) "On HTTP live streaming in large enterprises". SIGCOMM '13 Proceedings of the ACM SIGCOMM 2013.

[17] Roverso R. (2013-b) "A System, Tools and Algorithms for Adaptive HTTP-live Streaming on Peer-to-peer Overlays."Doctoral Thesis in Information and Communication Technology, KTH - Royal Institute of Technology Stockholm.

[18] Roverso, R., El-Ansary, S., Haridi, S. (2012) "SmoothCache: HTTP-Live streaming goes peer-to-peer."In Proceedings of the 11th international IFIP TC 6 conference on Networking - Volume Part II (IFIP'12), Robert Bestak, Lukas Kencl, Li Erran Li, Joerg Widmer, and Hao Yin (Eds.), Vol. Part II. Springer-Verlag, Berlin, Heidelberg, 29-43.

[19] Spangler, T. (2009) "YouTube May Lose $470 Million in 2009: Analysts". http://goo.gl/oNgAzY, acessado em Março, 2014.

[20] Tommasi, F., Melle, C., De Luca, V. (2014). OpenSatRelaying: a Hybrid Approach to Real-Time Audio-Video Distribution over the Internet. Journal of Communications, 9(3), 248-261.

[21] Vlavianos, A; Iliofotou, M.; Faloutsos, Michalis, "BiToS: Enhancing BitTorrent for Supporting Streaming Applications," INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings, vol., no., pp.1,6, 23-29 April 2006. doi: 10.1109/INFOCOM.2006.43

[22] Vu, L., Gupta, I., Liang, J., Nahrstedt, K. (2006). Mapping the PPLive network: Studying the impacts of media streaming on P2P overlays.

[23] Vu, L., Gupta, I., Nahrstedt, K., Liang, J. (2010). Understanding overlay characteristics of a largescale peer-to-peer IPTV system. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP),6(4), 31.

[24] Wuhib, Fetahi, Rolf Stadler, and Mike Spreitzer. "A gossip protocol for dynamic resource management in large cloud environments." Network and Service Management, IEEE Transactions on 9.2 (2012): 213-225.

[25] Yu-Kwong Kwok, (2011) "Peer-to-Peer Computing: Applications, Architecture, Protocols, and Challenges", Chapman and Hall/CRC Computational Science Series, CRC Press, Taylor and Francis Group, 200 pp., ISBN: 978-1-4398-0934-1.

[26] Zhang, Xinyan, et al. "CoolStreaming/DONet: a data-driven overlay network for peer-to-peer live media streaming." INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE. Vol. 3. IEEE, 2005.

[27] Zimmerman, A., (2014) "ABC Promised to Livestream the Oscars and Totally Failed". http://goo.gl/sTNb7d, accessed in March, 2014.

**Flavio Ribeiro** is an MSc. Student in Computer Science at the Federal University of Paraiba and Senior

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.

Director of Engineering at ViacomCBS Inc. working on research, development and deployment of large-scale online video systems. Flavio participated on the design and delivery of large-scale video projects, such as the live streaming stack responsible for the 2014 FIFA World Cup for Grupo Globo, the U.S. Presidential Election Debates for The New York Times and the NFL SuperBowl 53 for CBS.

**Guido Lemos de Souza Filho**, is a Full Professor in the Department of Computer Systems at the Center for Informatics at the Federal University of Paraíba (DI-CI-UFPB) holds a PhD in Informatics from the Pontifical Catholic University of Rio de Janeiro (PUC-RIO). Coordinates LAVID (Center for Research and Extension in Digital Video Applications) where he works in research on the following topics: digital television, digital cinema, distributed multimedia applications, video distribution networks, distributed artistic performances, accessibility, information security, fakenews and blockchain applications. Worked on the development of the Ginga middleware, published as ITU-T and ITU-R recommendations, and adopted as a standard in the Brazilian Digital Television System and in several countries in Latin America, whose implementation is software currently installed on about 100 million devices TV. The results of their research also include the development of a system for the storage, transmission and display of 4K 3D videos called Fogo Player, the development of a platform to support the realization of distributed dance, theater and music shows called Arthron, the development of video servers for live transmission and on demand, called DLive and DVod, which were used in RNP 's Digital Video Network and in the IPTV service of USP - SP, the VLibras accessibility software (used on www. brasil.gov.br, senado.leg.br and Câmara.leg.br) output 1.5 billion times a year; the development of technologies for registration, validation and preservation of Digital Diploma based on blockchain that will be used in 170 Brazilian public universities; finally, the development of the V4H health video system that uses digital signature technologies, blockchain registration and preservation to add security in the use of videos generated in appointments. He was coordinator of REUNI at UFPB and participated in its creation and implementation foreseen in the increase from 20 to 45 thousand in the number of undergraduate and graduate students. He also serves as a member of the Deliberative Council of the Brazilian Digital Television System Forum and guest of Ancine's Technical Accessibility Chamber.

Received in 2020-11-17 | Approved in 2020-11-21

This open access article is distributed under a Creative Commons Attribution (CC-BY) license.
http://www.set.org.br/ijbe/ doi: 10.18580/setijbe.2020.6 Web Link: http://dx.doi.org/10.18580/setijbe.2020.6