IOT-BASED MEDIA FRAMEWORK FOR PUBLIC SERVICE MEDIA: EXPANSION OF CURRENT DIGITAL BROADCASTING SYSTEMS

Hiroki Endo, Tatsuya Sato, Hiromu Ogawa, Shinya Abe, Satoshi Fujitsu, Kinji Matsumura, and Hiroshi Fujisawa

NHK (Japan Broadcasting Corporation), Japan

ABSTRACT

This paper proposes “IoT-based Media Framework,” a media presentation framework that autonomously presents content using appropriate devices via suitable presentation methods for users. The proposed framework replaces the existing vertically integrated broadcast system for dedicated receivers, with a new mechanism using semantic descriptions that represent three important aspects of media presentation—content, environment (including devices), and user context—to provide a user-centric experience that is essential for public service media. This concept is ideal for extending content expression and for increasing accessibility to people with disabilities thus furthering the provision of universal services. In this paper, we describe the novel concept and mechanism of the proposed framework. Advanced implementation of the prototype using the Web of Things (WoT) recommendation of World Wide Web Consortium (W3C) is successfully achieved. Furthermore, we report on the feasibility testing of our framework including the testing for W3C WoT Plugfest, an interoperability testing event with IoT device manufacturers.

INTRODUCTION

Broadcasters have been providing universal services that entertain people and promote safety and security as public service media, and they will continue this role in the future. Conventionally, broadcasters have been providing stable services only to dedicated receivers. Nowadays, a wide variety of Internet of Things (IoT) devices, such as smartphones, connected TVs, smart speakers, lights, wearable devices, and smart appliances, are becoming more common. Broadcasters and other content providers are now offering universal services to these devices with various forms of media content separately for each service such as broadcast programs, radio, video on demand (VOD), podcasts, news applications, and social networking services, based on their target users and the information they want to convey. However, considering the production cost and operational management, it is extremely difficult to produce and distribute content that is suitable for all these infinitely expanding options using a conventional vertically integrated system consisting of dedicated receivers and distribution equipment.

From the user's perspective, there is another important issue. Two typical patterns for a user to consume content are as follows: a) The user decides what content to consume, remembers the corresponding device for the content, and then activates the device. b) The user activates a well-used device and consumes the content on it. Thus, the user’s prior knowledge is used to link the device to the content. If the user is not familiar with the functions of the device, does not have sufficient knowledge about the content delivery media, or has a habit of using only a particular device, the opportunity for content consumption will be lost (1). In the smart society of the future, this issue will become more serious as a greater variety of devices is expected to present media content.

To solve these issues for both users and content providers, we propose IoT-based Media Framework, a user-centric media presentation framework for future public service media that treats three elements involved in the media experience: 1) the content provided by broadcasting services, 2) the environment that includes the devices, and 3) the user context within the same framework. The proposed framework replaces the present broadcasting system using a novel system architecture wherein the three elements are loosely coupled. This framework has the following features.

- This framework abstracts user requirements for content consumption, which are now specified for specific locations on specific devices, and autonomously presents content according to the user's environment. This content can be presented such that the user's original intent is maximally met, or the expectations are exceeded.
- The framework can be developed; thus, incorporating advances in the areas of the three elements. This will ensure that broadcasting services are sustainable even with the unexpected appearance of new types of devices and changes in the lifestyle of users.

In this paper, we introduce the concept of the IoT-based Media Framework. Then, we describe two use-case categories for public service media and their concrete examples. Subsequently, we extract the requirements. Thereafter, a system architecture that fulfills the requirements is presented. In addition, we report on the feasibility of the framework through two tests: one at the World Wide Web Consortium (W3C) Web of Things (WoT) Plugfest, an interoperability test of the WoT standard (2), and the other at an in-house testing using actual devices.

IOT-BASED MEDIA FRAMEWORK—KEY CONCEPT FOR FUTURE PUBLIC SERVICE MEDIA

We classify the elements involved in the media experience to compose future universal services wherein the provided content reaches a wider range of users than at present, is easily interpreted, and exhibits the value of the content provided. We define the three elements involved in media presentation: the content as the object that fulfills the user's intentions, the environment, which includes the device to be used as the interface that fulfills the content, and the user context that enables the presentation of media in accordance with the user's profile and behavior

We propose IoT-based Media Framework that loosely connects these independent entities. Herein, information is exchanged according to each semantic description, and devices can autonomously deliver content and information to users. We believe that this functional design of a user-centric media service will fulfill the requirements of the future public service media (Figure 1).

RELATED WORK

The integrated broadcast-broadband systems, especially Hybridcast (3) and HbbTV (4), use web applications called broadcast-related applications, which are displayed on a TV

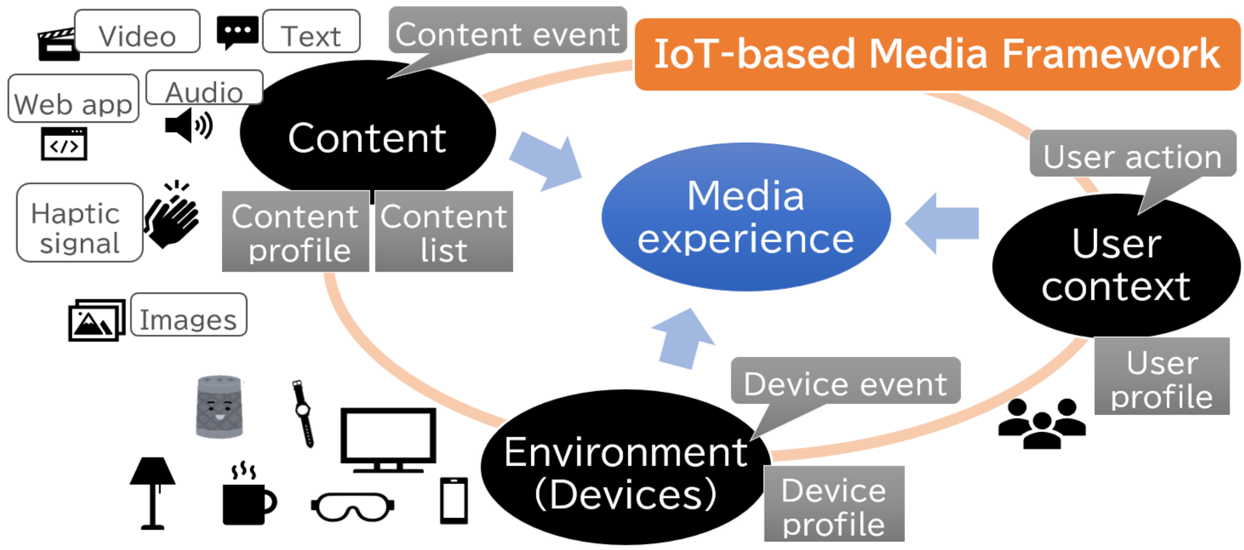

Figure 1 – Concept of IoT-based Media Framework

screen or a companion device that works with the TV. These services are based on the experience of watching TV programs on a TV screen. As such, they are not intended for the use cases envisioned herein, wherein a variety of devices are used to present content.

RadioDNS (5) and DVB-I (6) exemplify previous efforts to combine broadcasting and the Internet. These systems are designed to use content via the Internet with traditional broadcast systems. Conventionally, they were intended to be location-based services in terms of broadcast channels. The aim is to move away from location-based content management to manage content in a network of mixed media in a unified manner.

Object-based media (7) aims to provide a personalized experience by decomposing and rearranging content into its constituent object units.

Information-centric networking (ICN) (8) is a new network technology that provides seamless access to content by shifting from host-oriented to content-oriented design. The current architecture of the Internet would have to be changed to realize ICN.

USE CASE FOR IOT-BASED MEDIA FRAMEWORK

Broadcasting is a public service media required to deliver accurate information to many people. To realize user-centric services that present appropriate content according to the content consumption conditions of various users, two groups of use cases can be categorized as “alternative media” and “enhanced media experience” (Figure 2).

The use case categorized as “alternative media” aims to deliver as much information as possible that could not be delivered based on the user's situation and profile. This is a use case that improves accessibility by reducing the chances of the user not being able to consume the content or doing so depending on the situation or user context. Ensuring accessibility of information is a core task that public service media must universally realize. For example, when a soccer game is being broadcast, the content is presented to the user through the living room TV. For devices that cannot receive the broadcast, the video and audio will be delivered via the Internet. For speaker devices that do not have a screen, for people with visual impairments, or those who are working and cannot see the display, audio descriptions will provide audio commentary that includes the video’s content. This

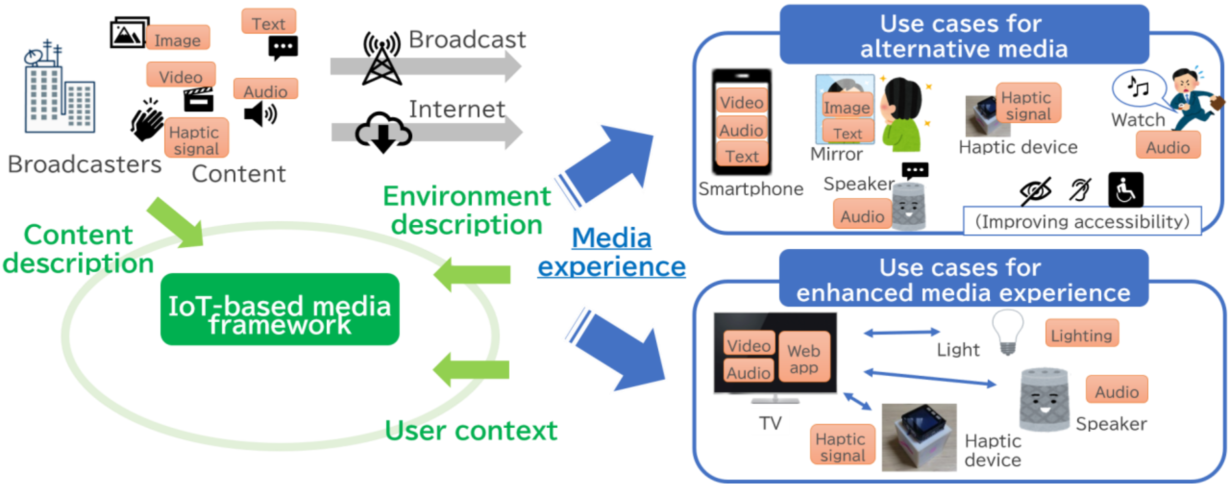

Figure 2 – Overview of use cases for IoT-based Media Framework

means that the content related to the delivery media, device functions, and user context will be realized by alternative means to deliver the program content correctly.

Use cases categorized as "enhanced media experience" are those that further improve the media experience in the current state by incorporating multiple devices or referring to the user context. For example, in a conventional live soccer program, supplementary information such as the progress of the game, player names, and results are overlaid on the screen, but can be displayed on a tablet in the user's hand at the user's desired timing. The audio commentary cannot be output from the TV speakers but from another speaker device. Another example of this use case is the enhancement of production effects, such as dimming the lights in the room, or the vibration of the sofa as the game progress.

From these two perspectives, we list typical use cases that represent user-centric content consumption. As a system configuration derived from these use cases, the three elements involved in the media experience are organized in Table 1 as three entities: content, environment, and user.

Alternative Media Use Case

- Alternative presentation based on the distribution media function
  - A) While broadcast content is played on TVs, live-streamed or VOD content is presented on tablets and smartphones without tuners via the Internet.
  - B) The radio broadcast program is played back on a radio, while the live-streamed audio content is presented on a smart speaker without radio tuners.
- Alternative presentation based on the presentation function of the device
  - C) Video content produced with a layout for large screens is played on TVs, content produced with a layout for small screens is provided on smartphones, and audio content is provided on smart speakers.
  - D) When the early warning of an earthquake is provided as a video on TV, the information is presented as audio information on radios and smart speakers, as changes in brightness and color on smart lights, and as vibrations on haptic presentation devices.
- Alternative presentation based on the environment of devices

E) In an environment where broadcast signals can be received, the broadcast will be used; in an environment where broadcast signals cannot be received, the Internet simulcast of the broadcast will be received and presented.

F) Depending on the speed of the network it is connected to, such as high-speed optical network, 5G, and 4G/LTE, the content can be reduced in terms of video resolution, changed to audio only, or, changed to still images and text.

G) A smartphone that normally plays a video will turn off the display and play only audio when the battery power is low.

H) When a device leaves the environment, such as when a user takes a mobile device away from home or the device's battery runs out, the way the content is presented changes to fit the new device environment.

- Alternative presentation based on user context

I) Video information will be supplemented with visual information with audio description for the visually impaired. For the hearing-impaired, the information will be supplemented by video and vibration. The supplementary information may be generated by content providers or by user devices.

J) In situations where no display device is available, video information is supplemented by audio description. In places such as libraries where users need to maintain silence or in noisy environments where it is difficult to hear, audio is replaced by visual, textual, or haptic information.

K) When users do not have time to pay attention to the content, such as when they are cooking or when they are on the move, the program is served instead by a digest video or a synopsis site consisting of still images and text.

L) When a user is watching a program alone and one of the family members enters the same space, the content will automatically change to a common preference content (the user's goal is not to watch a certain content but to relax, watching any content).

Enhanced Media Experience Use Case

- Presentation of supplemental information for the main content

O) Supplemental text information will be displayed on a nearby table display as the broadcasted content progresses on a TV.

P) As the user becomes more interested in the content, more detailed information will be presented on the TV and peripheral devices along with the broadcast content.

Q) When a device is added to the environment, such as when a user brings a mobile device home or when a device's battery is charged, supplemental content information is added.

- Provision of emphasis on main content

P) The color and brightness of the lighting will be changed to match the content, and haptic information will be presented. Ventilation fans and robot vacuum cleaners will be paused at important scenes.

Q) During the halftime period of the soccer game, the user's favorite coffee machine will automatically make coffee.

SYSTEM ARCHITECTURE

Based on the analysis of the use cases presented in Table 1, we defined the functions of each entity. We then presented the IoT-based Media Framework, a framework for linking these three elements together.

System Requirements

We determined that the “alternative media” and “expansion media” use cases were equivalent in terms of requirements and could be realized within the same framework. Based on our analysis of the use cases, we defined three media elements: content, environment, and user context. Content is the object that fulfills the user's intentions, and the environment includes the device that is the interface between the content and user. User context includes the user profile and behavior.

System Configuration

Figure 3 shows the IoT-based Media Framework consisting of three elements—content, environment, and user context. These three elements interact with a media presentation application. These elements are described semantically, in static or dynamic form (Table 2).

The application receives or reads these descriptive events and presents the content to the user as appropriate to the situation. The application can run on an independent entity such as a home server, be implemented on each device, be embedded in Content, or a combination of these implementations.

When an event occurs in one of the three elements, it queries the other two elements to reconstruct the media experience. Figure 4 shows the processing sequence.

Use case categories	***Use case***	***Triggered event***	***Elements mainly used for the decision of each presentation method***		
			Content	***Environment***	***User***
Alternative Media	D, I, J	Content		√	√
	E, F	Environment	√		
	G	Environment		√	√
	H	Environment	√	√	√
	A, B, C	User	√	√	
	K, L	User	√	√	√
Expansion Media	M, P	Content	√	√	√
	O	Environment	√	√	√
	N, Q	User	√	√	√

Table 1 – Analysis of use cases

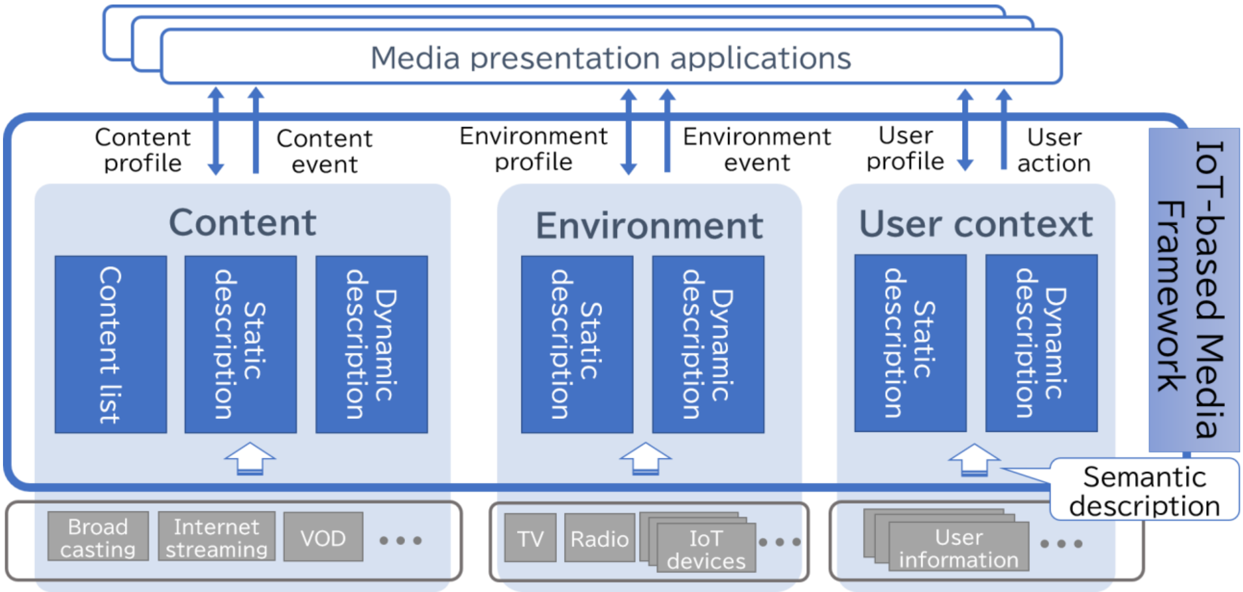

Figure 3 – System architecture of IoT-based Media Platform

Elements	Static information	Dynamic information
Content	Content list: description of the relationship between contents for discovery and grouping. Content profile: description of the title, genre, distribution location, media format, etc.	Content event: event information that occurs within the content, such as goals and half-time in soccer broadcasts
Environment	Device profile: supported functions related to the delivery media, display and speaker functions, screen size, etc.	Device event: battery level notification, broadcast and communication reception environment, and so on.
User context	User profile: information on accessibility, preferences for content and media presentation	User action: who is where and what they are doing.

Table 2 – Three elements involved in media presentations

PROTOTYPE

Based on the system schematic illustrated in the previous section, we implemented a prototype system to realize the service scenario based on simple parameters (Figure 5). Regarding content, for several programs, we listed the forms of content, such as video, audio, text, and web applications, as well as meta-information about the location and content. In addition, there are trigger events related to the content, such as event messages multiplexed in the broadcast signal, and media timed events (9) for synchronization of events with Internet-streamed content.

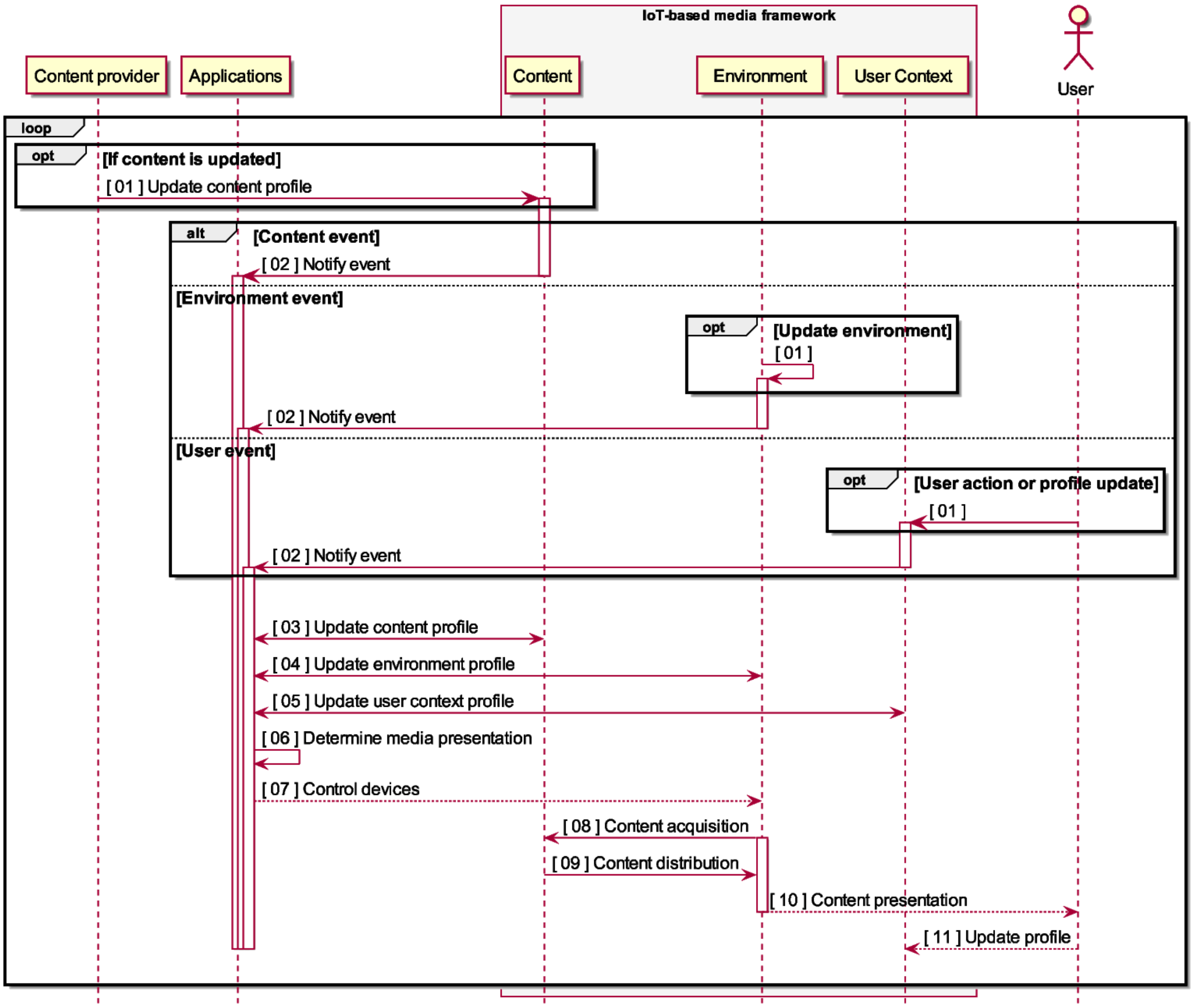

Figure 4 – Procedure sequence of IoT-based Media Platform

We obtained event messages multiplexed into the broadcasting content of the Japanese Integrated Broadcast Broadband system using the Hybridcast Connect API. The environment is described as available device information based on the W3C WoT recommendations, and it implements device properties, possible actions, and generated events. Regarding the implementation devices, we adopted a Hybridcast TV, a device with a web browser, smart speakers, and a cube-shaped haptic device that improves accessibility for people with disabilities (10) and a smart coffee machine.

The description of the user context was simple here, showing the user's location in their home (living room, kitchen, entrance) and their preferred genre of content. The application is implemented using a visual programming tool, i.e., Node-RED (11). The registered devices are controlled based on the results of a retrieval of the events of each element and referencing of the profile information. The vocabulary of the description of the three elements was referenced from schema.org (12).

VALIDATION TEST

The prototype's implementation confirmed the feasibility of our system, which dynamically changes the way media is presented to the user depending on the events based on content (broadcast programs), events based on user behavior, device functions, and other related events.

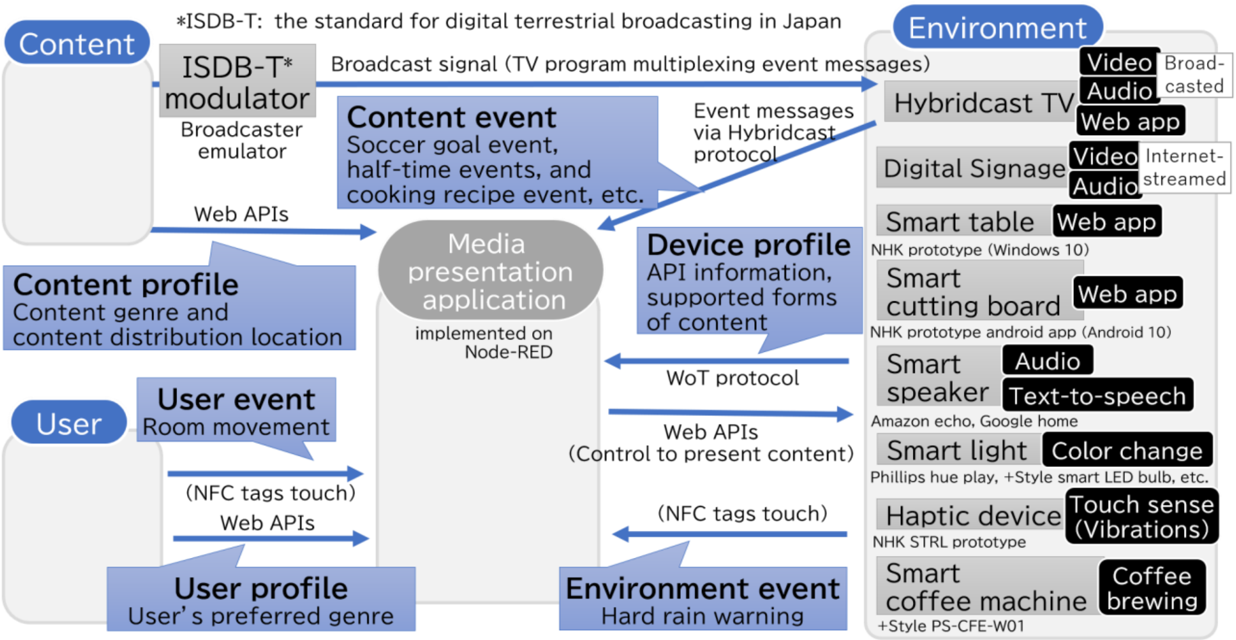

Figure 5 – System configuration of prototype

W3C WoT Plugfest

We participated in WoT Plugfest in September 2020. We used the Hybridcast emulator (13) to create an event that controlled other participants' devices by broadcasting events (Figure 6). At Plugfest, we were able to verify basic scenarios, such as the scenario in which a coffee machine in a viewer's home is controlled by a broadcasting station event and the scenario in which a TV channel is controlled by a sensor of an IoT device within the home (Figure 7).

We verified the effectiveness of the basic use cases involving media and IoT devices, and implemented the basic integration of the proposed system with other devices that belong to other companies at the public event.

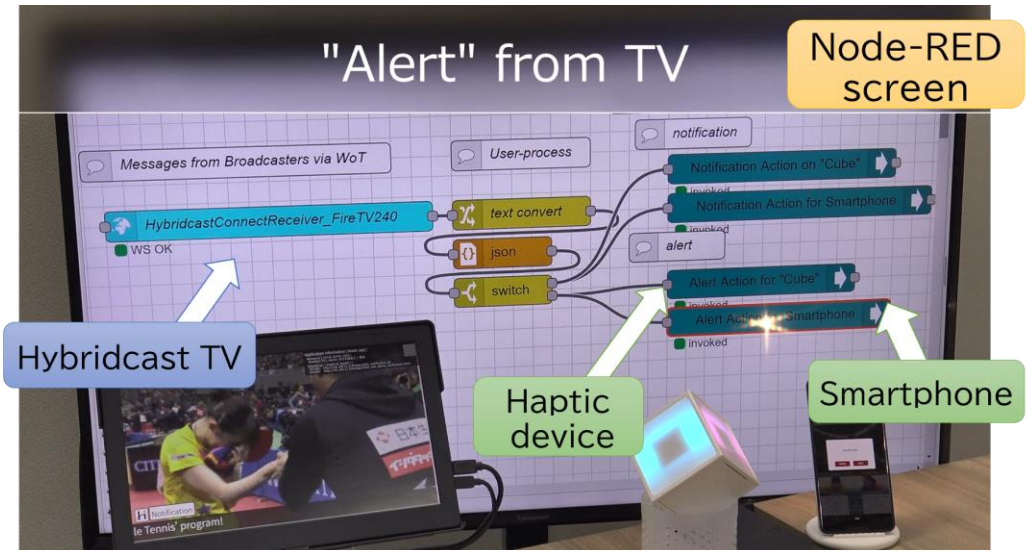

Figure 6 – Alternative presentation using IoT devices

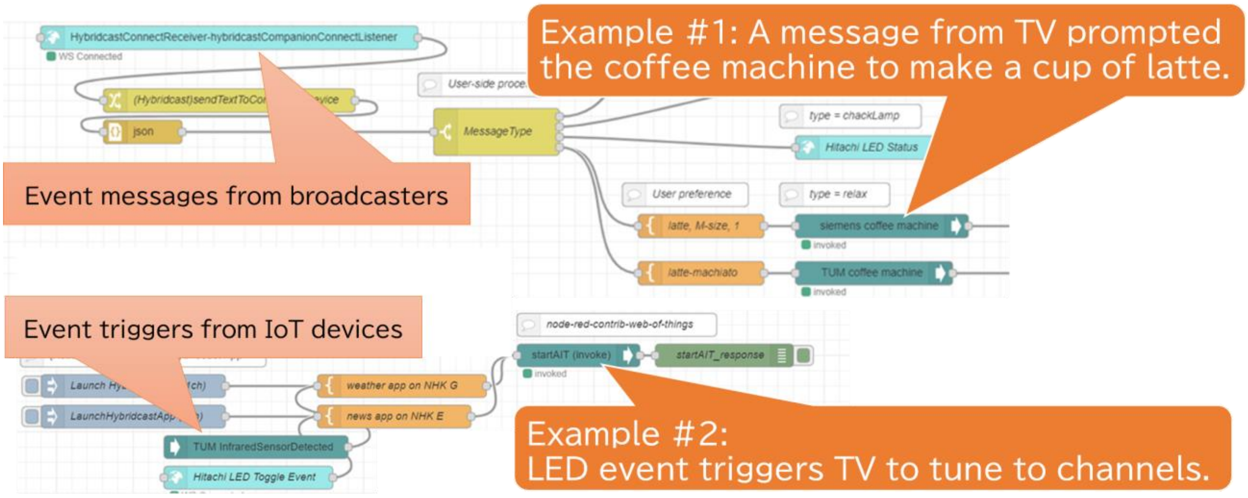

Figure 7 – Prototype of media-device interaction (Node-RED screen capture)

In-house Testing

In an experimental environment that emulated a home, we implemented a scenario in which multiple forms of content were automatically presented to various device environments according to events in the three elements of content, environment, and user. Figures 8, 9, and 10 show one aspect of the operational verification.

To test this, we employed events in the content, such as goal events in soccer broadcasts, as well as events emitted by devices, such as user requests to view content, movement between rooms, and the issuing of heavy rain warnings. The prototype realized two use case groups: alternative media and enhanced media experiences. The feasibility of the proposed system architecture was confirmed through verification tests compared with actual devices.

	
A TV shows broadcasted video and audio. A table displays the goal event animation.	An internet-streamed video is displayed on in-home digital signage.
	
A vibration is presented to notify goal events using a cube-shaped haptic device.	A coffee machine is triggered by a halftime event.

Figure 8 - Verification 1: Improving accessibility and entertainment of soccer games

	
A TV presents broadcasted content	A cutting board presents a recipe

Figure 9 - Verification 2: Presenting cooking programs through various means

	
A TV presents broadcasted warnings Lights emit warning colors	A table displays an alarm notification

Figure 10 – Verification 3: Effective weather warning by TV and various IoT devices

DISCUSSION

At the WoT Plugfest, we tested the effectiveness of a basic use case that involved media and IoT devices and conducted a basic integration of the proposed system with the devices from other companies at a public event. In addition, in-house testing verified the feasibility of the proposed system architecture by means of some scenario with the use of actual devices. Two validation tests confirmed the feasibility of the proposed system. The framework was found to maximize the satisfaction of the user's original intentions or exceed their expectations by presenting content according to the user's environment. It was also found that this framework can be developed to incorporate advances in each of the elemental areas, such as WoT recommendation.

For the description of the three elements, information based on the scenarios was prepared in advance. We referred to WoT recommendation and schema.org for the descriptions, but some definitions related to media presentation had to be defined independently. It is expected that the definition of the devices will be enriched by various contributors. By having broadcasters provide semantic metadata for media presentation, this framework can be more generic and contribute to universal services.

CONCLUSIONS

This study proposes IoT-based Media Framework, which realizes user-centric media experiences in diverse device environments. We present a system architecture that enables the realization of two use cases in a common framework: the presentation of alternative media that satisfies the user's intentions, and the provision of enhanced media experience that exceeds the user's expectations. We designed a system architecture that satisfies the requirements extracted from the assumed use cases and implemented a prototype.

We verified the prototype on two testing. Accordingly, the feasibility of the proposed system was verified based on the confirmation that the system can automatically present contents on various devices while referring to information related to each other, triggered by content, environment, and user events. We found that our framework satisfies the requirements of universal services for public service media.

It was also observed that the three elements that comprise the framework are loosely coupled and will continue to develop, thus incorporating advances in each area. This will ensure that broadcasting services are sustainable. This proposal has the potential to make a technological contribution to solve the digital divide caused by the diversification of our living environment.

In the future, we will continue to study more specific, practical descriptions of content, environment, and user context information, as well as conduct user evaluations, to realize universal services based on the proposed framework.

REFERENCES

1. Endo, H., et al., 2020. Distribution media-independent multidevice cooperation system for content viewing. 2020 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). C6-1.
2. W3C, 2020. Web of Things (WoT) architecture. https://www.w3.org/TR/wot-architecture/.
3. IPTV FORUM JAPAN, 2020. IPTVFJ STD-0010 Integrated broadcast-broadband system specification version 2.3.
4. ETSI, 2018. ETSI TS 102 796 v1.5.1 Hybrid broadcast broadband TV.
5. ETSI, 2015. ETSI TS 103 270 V1.1.1 RadioDNS hybrid radio: hybrid lookup for radio services.
6. DVB, 2019. DVB document A177 digital video broadcasting (DVB); service discovery and programme metadata for DVB-I services.
7. Armstrong, M., et al., 2020. Taking object-based media from the research environment into mainstream production. SMPTE Motion Imaging J. vol. 129, no. 5, pp. 30 to 38.
8. Xylomenos, G., et al., 2013. Survey of information-centric networking research. IEEE Trans. on Consum. Electron. Vol. 52, no.1, pp. 1 to 26.
9. W3C, 2020. Media timed events. https://www.w3.org/TR/media-timed-events/.
10. Azuma, M., et al., 2017. Development of vibration cube to convey information by haptic stimuli. Proceedings of the 24th Display Workshops (IDW'17). pp. 128 to 130.
11. Node-RED. https://nodered.org/.
12. Schema.org. https://schema.org/.
13. Antwapp4hc. https://github.com/nhkrd/antwapp4hc.