Design of Wireless Audio Real Time Transmission Model Based on Body Area Network Technology

Qing-li Niu[1] (✉) and Hong Xu[2]

[1] College of Information Engineering, Zhengzhou University of Science and Technology, Zhengzhou 450064, China
jsj32989317@163.com

[2] School of Software, Zhengzhou University, Zhengzhou 314500, China

Abstract. In order to achieve real-time compression of audio signals and establish a more stable wireless network transmission mode, this study designed a wireless audio real-time transmission model based on body area network technology. First, with the support of wireless body area network technology, a complete information aggregation processing system is built, and then the body area network environment is built by configuring the downloaded audio resources. Based on this, the digital signal processor is selected, and a complete data structure of multiple audio frames is established, so as to realize the real-time transmission of wireless audio. Experimental results show that compared with the traditional multi-channel audio transmission technology, the proposed model has higher effectiveness and stronger signal matching ability, which proves that it can compress audio signals in real time and establish a more stable wireless network transmission environment at the same time.

Keywords: Body area network technology · Wireless audio · Real-time transmission · Convergence processing layer · Resource download · Resource configuration · Signal processor · Protocol message

1 Introduction

Audio is a technical term used to describe sound-related devices in the audio range and their functions. All sounds that human beings can hear are called audio, which may include noise, etc. After the sound is recorded, whether it is voice, singing, musical instruments can be processed by digital music software, or it can be made into a CD. At this time, all the sounds have not changed, because CD is a type of audio file. Audio is just the sound stored in the computer. If there is a computer and the corresponding audio card, which is often called the sound card, we can record all the sounds, and the acoustic characteristics of the sound, such as the sound level, can be stored in the form of computer hard disk files. Audio signals are the carrier of regular frequency and amplitude changes of sound waves with voice, music and sound effects. According to

S. Wang et al. (Eds.): IoTCare 2021, LNICST 415, pp. 103–114, 2022.
https://doi.org/10.1007/978-3-030-94182-6_8

the characteristics of sound waves, audio information can be classified into regular audio and irregular sound. Regular audio can be divided into voice, music and sound effects. Regular audio is a continuous changing analog signal, which can be represented by a continuous curve, called sound wave. The three elements of sound are tone, intensity and timbre.

Body area network is a basic technology that can monitor and record human health signals for a long time. Its early application is mainly used to continuously monitor and record the health parameters of patients with chronic diseases and provide some way of automatic therapy control. In the future, body area network can also be widely used in consumer electronics, entertainment, sports, environmental intelligence, animal husbandry, ubiquitous computing, military or security fields. Body area network is a special wireless communication system, which takes the devices around the human body, such as watches, sensors and mobile phones, as well as the internal human body (implantable devices) [1].

Therefore, in order to realize the real-time compression of audio signals and establish a more stable wireless network transmission mode, this study designed a wireless audio real-time transmission model based on volume domain network technology. The design idea of this model is as follows:

(1) Build a complete information aggregation processing system by using the wireless body area network system architecture, and then complete the construction of the body area network environment by configuring the downloaded audio resources.
(2) The digital signal processor is selected and processed. According to the data structure of multiple audio frames, the smooth application of wireless audio real-time transmission model based on volume domain network technology is realized.

2 Building Body Area Network Environment

The body area network application environment is composed of system architecture, information aggregation processing layer, audio resource download and configuration system. The specific construction method is as follows.

2.1 Wireless Body Area Network System Architecture

Wireless body area network system architecture is divided into three levels, namely information collection layer, information aggregation processing layer and cloud information layer.

The audio information acquisition layer contains many sensor nodes, which have the function of collecting audio information, mainly for the audio wavelength, amplitude, frequency, etc. The collected information is then transmitted to the upper control node. In this study, a simple star topology is used as the system structure of the information acquisition layer in the volume domain network.

The audio information aggregation processing layer is similar to a personal server, and its requirements on energy consumption and size are different from the acquisition sensor of the audio information acquisition layer. In the star topology, the monitored audio information is directly received from the front-end information acquisition nodes. In general, personal PDA devices with display and processing functions are selected as the aggregation nodes in the information aggregation processing layer. On the one hand, the wireless audio information collected by the front end can be stored and analyzed. On the other hand, the information can also be sent to the cloud data information layer through the wireless network [2].

The schematic diagram of the wireless body area network system architecture is shown in Fig. 1. The audio information cloud mainly includes a server that supports remote transmission, which is mainly used to receive the monitoring audio feature values forwarded by the sink node, process and analyze the received monitoring data, and store it in the background database or upload it to the network.

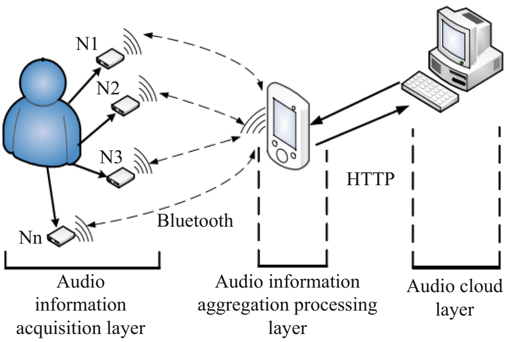

Fig. 1. Schematic diagram of wireless body area network system architecture

2.2 Information Aggregation Processing Layer

The audio information aggregation processing layer, as its name implies, refers to the aggregation node. In addition to receiving the monitoring information transmitted by the information collection layer, it also has the functions of UI display and audio data processing. For this reason, the body area network host is selected as the information aggregation processing layer. The convergence center is mainly considering that the positioning and remote communication functions of the body area network host can meet the characteristics of wireless audio instantaneity in emergency situations, and the body area network host has occupied with its complete openness and unlimited user equality. Most of the Internet application network [3]. For the communication connection with the Bluetooth module of the information collection layer, the body area network host has already supported Bluetooth development in the 2.0 version of the software development kit. In the communication with Bluetooth, eight categories are mainly used, and the location is under the Bluetooth package. The eight categories are shown in Table 1.

Table 1. Application capability description of information aggregation processing layer

Class name	Explain
BluetoothAdapter	Bluetooth adapter for local audio
BluetoothClass	Bluetooth
BluetoothClass.Device	Bluetooth transmission devices
BluetoothClass. Device.Major	Bluetooth transmission device management
BluetoothClass. Service	On the class of Bluetooth transmission service
Bluetooth Device	Bluetooth device for wireless remote audio
BluetoothServerSocket	Wireless audio monitoring Bluetooth connection class
BluetoothSocket	Bluetooth connection class, mainly the client

When communicating with Bluetooth module, we mainly use Bluetooth adapter class, Bluetooth device class and Bluetooth socket class. Bluetooth adapter class is used to obtain the local wireless audio adapter structure; Bluetooth device class is used to represent a remote Bluetooth device, that is, the Bluetooth module of the Bluetooth information collection layer; Bluetooth socket class is used to represent the output stream of the Bluetooth connection.

2.3 Download and Configuration of Audio Resources

In the wireless body area network system application development process, in addition to mastering the development technology of audio resource download, it is also necessary to understand the structure of audio configuration relationships.

The audio resource application program is generally constructed by four functional modules: active object, intention, service and content provider. But not all applications must use the above four functional modules, and can also be composed of one or more of the above. The only thing in common is that no matter which function module is used, it must be registered in advance in the configuration file of the program.

Activity

Activity is the most basic application component of the body area network system. Although each application is not necessarily composed of the above four components, as long as there is an interface display, the application must contain the activity application component. In a popular sense, each activity corresponds to an independent screen in the body area network system. It is an interface used to realize human-computer interaction, and is used to store display controls to realize various functions. Each activity can be regarded as a class. You can see that most applications are composed of many activities, in order to distinguish which one Activity is the current interface in the application program. The body area network system especially sets seven operation states for activity: create, start, pause, pause resume, stop, stop resume and destroy. These seven different operation states constitute the life cycle of each activity [4].

Intent
An application is often composed of multiple activities. In order to realize the jump between different activities, the body area network system uses the Intent component to complete the communication of different activities, such as starting a new activity, and can pass data to this new activity. The Activity o Intent can also be used to call applications provided by the body area network system itself, such as calls to dial-up programs and sending text messages. The data transferred in the Intent includes 7 types, namely operation, data, data type, operation category, additional information, components and signs.

Service
In the application, there are some programs running in the background, such as sending SMS or playing a music player in the background. These operation components without interface display are actually a service, and the main functions provide some necessary support for activity Service also has its own life cycle. Compared with activity, the basic life cycle of service only needs to use two operation methods: staxt sexvioe() and stop sexvioe().

Content Provider
The body area network system provides five storage methods for data, the most commonly used is to save in files and SQLite data, but when some applications developed need to call the data of existing applications, such as a program to obtain the phone book Since different applications are working in different processes, it is generally difficult to transfer data from one virtual machine to another virtual machine [5]. Therefore, in order to meet the data sharing between different programs, the body area network system specifically defines the content provider, a special data storage class, and provides a set of standard interfaces for obtaining and operating data. Content Provider also supports data sharing. Adding, querying, updating, and deleting operations facilitate the mutual access of data between different programs.

3 Wireless Audio Real Time Transmission Model

Based on body area network environment, according to the selection of digital signal processor, RTP/RTCP protocol message format definition, audio frame data structure connection processing flow, the design of wireless audio real-time transmission model based on body area network technology is completed.

3.1 Selection of Digital Signal Processor

In the design of wireless audio real-time transmission model based on body area network technology, a popular multimedia audio and video processing chip TMS320DM642 is selected for the digital signal microprocessor. It is a 32-bit fixed-point DSP chip of TMS320C6000 series which is suitable for processing digital audio and video stream launched by Texas TI company in 2003, instead of tms320c5000 series which is commonly used in embedded device system The design needs to process multi-channel audio data, and the tms320c5000 series digital signal processor chip has at most three McBSP ports to receive digital audio data. Moreover, the McBSP port has other uses, which cannot be used to collect digital audio data completely [6]. TMS320DM642 chip adds a multi-channel audio mcasp port, which can accommodate up to eight audio data streams, just to meet the requirements of multi-channel audio processing, four digital audio input streams and four audio data output streams.

TMS320DM642 also has a 10/1OOM Ethernet interface that meets the IEEE802.3 standard, eliminating the need for an Ethernet controller, thereby simplifying circuit design and reducing circuit costs. Therefore, the core processor chip designed in this research selects TMS320DM642 as the audio processor. Facing the rapid development of multimedia networks, the requirements of multi-channel data transmission and the expansion of on-chip memory capacity, in 2003, Texas Instruments (TI) launched a new generation of digital multimedia processor TMS320DM642 to meet the requirements of the market., It is a 32-bit fixed-point DSP chip belonging to TI's C6000 series. TMS320DM642 retains the original core structure of C64x, using advanced Harvard structure memory [7]. Its main features are as follows:

1. The working frequency of the DSP chip is input by the external crystal oscillator, and the internal PLL frequency doubling setting can achieve the clock frequency of 720 MHz, the number of executable instructions per second is 5760mips, and 8 32-bit instructions are executed per cycle, which has high-speed digital processing speed.
2. There are six ALUs in DSP, and each Alu supports 32-bit arithmetic operation per clock cycle.
3. A multi-channel audio serial interface (mcasp), 8 serial data pins, can accept up to 8 channels of audio data input or output, complete support I2S, S/PDIF and aes-3 digital audio stream audio transmitter.
4. An Ethernet controller interface (EMAC) conforming to IEEE 802.3 standard is integrated, including MII and MDIO.
5. A 64 bit external memory interface (EMIF), which can be seamlessly connected with external nonvolatile memory such as flash, SDRAM and EPROM. The maximum addressable space of external memory is 1024MB.

3.2 Message Format of RTP/RTCP Protocol

RTCP sends control packets periodically. RTP needs RTCP to provide guarantee for its quality of service. RTCP and RTP are all transmitted by UDP. RTCP control packets contain short control information and are encapsulated in a UDP packet. RTCP includes five different types of data packets (Table 2).

RTP/RTCP protocol message format definition includes the following processing steps:

1. Version: indicates which RTP version is used. If version 2 is used, the value is set to 2, occupying the first two bits of the first byte.
2. Fill: the end of RTP packet can contain one or more fill bits, which are not included in the load and occupy the third bit.

Table 2. Five packet types of RTCP

Types of	Data pack	Purpose
200	SR	Audio sender report
201	RR	Audio receiver report
202	SDES	Audio origin description
203	BYE	End audio transmission
204	APP	Specific audio transmission applications

3. Extension: if the fixed header of RTP package is followed by an extended header, this bit will be assigned to occupy the fourth bit.
4. CSR counter: the number of CSR after RTP packet header, occupying the last 4 bits of the first byte.
5. Mark: its content is specified by a specific protocol and occupies the first bit of the second byte.
6. Load type: it is an important data bit in RTP packet. It identifies the data type of RTP packet and defines the load format, occupying the last seven bits of the second byte.
7. Sequence number: It is an important data bit in the RTP packet. The sequence number increases by 1 after sending an RTP packet. The receiver can use this to detect the number of lost packets and restore the data packets out of order, occupying two bytes.
8. Timestamp: It is also an important data in the RTP packet, indicating the sampling time of the first byte of the RTP packet data. The value of the timestamp must be continuously accumulated. The timestamp matches the serial number to correctly play the audio to achieve jitter removal and synchronization, occupying four bytes.
9. Synchronization source identifier: used to distinguish RTP data packets from different sources. A single RTP session can only have a unique SSRC value, which is also four bytes.
10. Contribution source list: used to indicate the RTP sources contributed in the data packet. The source SSRC identifier in the data packet is identified, and the receiving end can correctly identify the participant, including 0–15 items, each with four bytes [8].

3.3 Audio Frame Data Structure

IP audio means that digital audio signal is transmitted by IP data structure in Ethernet network, which provides a short delay, low cost and high precision solution for digital audio real-time transmission system. The goal of IP audio transmission is to deliver high-precision clock signals and professional uncompressed digital audio signals over the Ethernet network, and to carry out complex routing at the network layer. IP based network audio transmission process uses transport layer UDP protocol and RTP protocol, of course, also need to use RTCP protocol.

Ethernet (EtherNet) is a baseband local area network specification jointly developed by Xerox, Intel and DEC. It is the most widely used communication protocol standard for local area networks. In recent years, most of the network audio real-time transmission solutions that have emerged are based on Ethernet technology standards, which are currently the most popular embedded device technology used, and have replaced token ring, FDDI and ARCNET LAN standards to a large extent. The Ethernet-based TCP/IP protocol occupies a dominant position in the field of audio transmission.

Ethernet is the basis for determining the efficiency and performance of the audio network. The analog audio signal transmitted on the standard Ethernet cannot be easily converted into a data format. In the audio network transmission, the audio signal requires a relatively strong timeliness, and the data sent The delay of the packet will cause incoherence and loss of the audio signal. Ethernet is an asynchronous technology and does not have the concept of real-time. Transmission management is uncertain, which means that Ethernet cannot guarantee the timely delivery of data packets. Therefore, in order to transmit audio data in real time and stably, the network must have a certain deterministic time-sensitive transmission technology [9]. For example, EtherSound, CobraNet and Dante technologies can provide such technologies.

The audio frame data structure definition form is shown in Fig. 2.

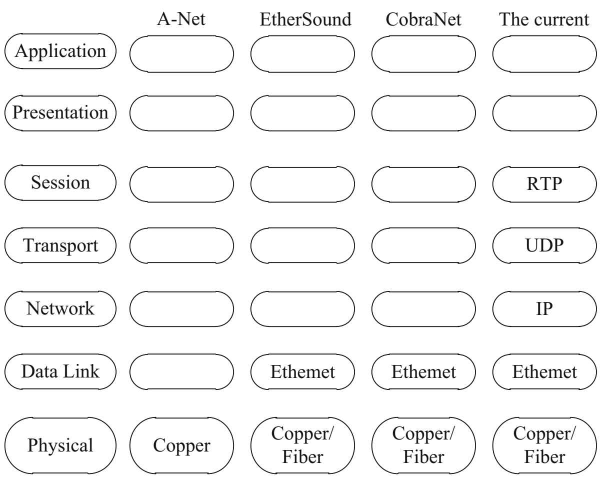

Fig. 2. Definition form of audio frame data structure

A complete Ethernet frame usually consists of synchronization bit PR, split bit SD, destination MAC address Da, source MAC address SA, type field type, MAC data field and frame check sequence. Among them, synchronization bit and split bit are also called preamble and frame start character. The content of the first seven bytes is 0xaaaaaa, and the content of the eighth byte is oxab. They play the role of synchronization at the receiving end, marking the formal start of transmission of an Ethernet frame. Followed by the six byte destination address DA and six byte transmission source address Sa of Ethernet frame transmission, the two byte type field indicates the protocol type used for network layer transmission [10, 11]. The next 461500 bytes are the data packets transmitted, which are the key parts of frame transmission. The last four bytes are the frame validation fields of Ethernet frame, which verify whether the data bits transmitted before are correct or not.

CobraNet transmits PCM data of uncompressed digital audio stream without data encoding and decoding, so it is used for real-time audio transmission in radio and television industry. CobraNet data packet is located in data link layer, namely MAC layer, including clock data packet, audio data packet and reservation data packet. The structured of audio data packet is shown in Fig. 3.

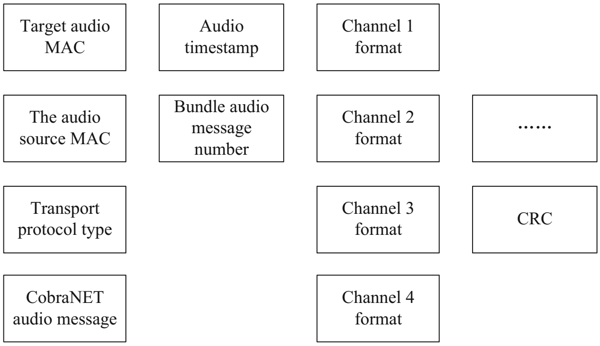

Fig. 3. Audio frame packet definition format

4 Model Application Ability Detection

In order to verify the practical application value of the wireless audio real-time transmission model based on volume area network technology designed above, the following comparative experiments are designed.

The experimental group and the control group were connected to the audio transmitter respectively. Among them, the experimental group was equipped with wireless audio real-time transmission model based on body area network technology, while the control group was equipped with multi-channel audio transmission technology. In the same experimental environment, the specific changes of the experimental indicators were analyzed.

It is known that the audio data transmission time can reflect the effectiveness of wireless audio information transmission. Generally, the shorter the consumption time, the higher the effectiveness of wireless audio information transmission, and vice versa. The specific experimental conditions are shown in Table 3.

Analysis of Table 3 shows that as the total amount of audio data increases, the transmission time of the experimental group maintains a trend of first increasing and then stable, and the global maximum can only reach 2.40 s. The transmission time of the control group has always maintained a rising trend, and the global maximum value reached 4.52 s, which is an increase of 2.12 s compared with the maximum value of the experimental group.

Table 3. Comparison of audio data transmission duration

Total audio data/(Gb)	Audio data transmission time/(s)	
	Experience group	Control group
1.0	2.31	4.18
2.0	2.33	4.21
3.0	2.35	4.24
4.0	2.36	4.27
5.0	2.38	4.29
6.0	2.39	4.33
7.0	2.40	4.35
8.0	2.40	4.38
9.0	2.40	4.52

The PSR index can describe the real-time matching ability of wireless audio signals. Generally, the larger the PSR index value, the stronger the real-time matching ability of wireless audio signals, and vice versa. Table 4 records the actual changes in the PSR indicators of the experimental group and the control group.

Table 4. Comparison of PSR indexes

Experiment time/(min)	PSR index/(%)	
	Experience group	Control group
5	57.38	41.32
10	76.19	42.08
15	58.24	42.79
20	78.87	43.16
25	56.73	44.87
30	79.62	45.20
35	56.25	46.13
40	80.36	46.65
45	56.41	47.41
50	79.50	48.54

According to the analysis of Table 4, with the extension of the experimental time, the PSR index of the experimental group kept the trend of rising and falling alternately, and the global maximum reached 80.36%. The global maximum of PSR in the control group was only 48.54%, which decreased by 31.82% compared with that in the experimental group.

By comprehensive analysis of the above results, it can be seen that the transmission consumption time of audio data is effectively controlled after the application of the model in this paper, and the value of PSR index also increases significantly, which proves that the model in this paper can realize real-time compression of audio signals while establishing a stable wireless network transmission environment. The reason for the above results is that the model in this paper has an independent information aggregation processing layer system, which can establish a complete audio frame data structure, thus improving the PSR index and reducing the time consumption of transmission process.

5 Conclusion

Compared with the multiplexed audio transmission technology, the wireless audio real-time transmission model designed in this study has an independent information aggregation processing layer system supported by the volume domain network environment, which can realize the selection of digital signal processor and establish a complete audio frame data structure at the same time. From the practical point of view, after the application of the model in this paper, the audio data transmission time is shortened and the PSR index value increases, which proves that the model in this paper can meet the real-time compression requirements of audio signals more effectively and has a strong application and promotion significance.

In the following research, we will consider enhancing the anti-noise ability of the audio signal while using the model to efficiently transmit the audio data, so as to complete the further optimization of the model.

Fund Projects. Research and Practice on learning performance evaluation model of online education under the background of smart campus The Key research projects of Henan Provincial Department of Education NO:21B880041.

References

1. El-Zoghdy, S.F., El-Sayed, H.S., Faragallah, O.S.: Transmission of chaotic-based encrypted audio through OFDM. Wireless Pers. Commun. **113**(10), 1–21 (2020)
2. Kaur, A., Dutta, M.K.: High embedding capacity and robust audio watermarking for secure transmission using tamper detection. ETRI J. **40**(1), 133–145 (2018)
3. Bakeva, V., et al.: Gaussian channel transmission of images and audio files using cryptcoding. IET Commun. **13**(11), 1625–1632 (2019)
4. Telagam, N., et al.: USRP based digital audio broadcasting using OFDM in virtual and remote laboratory. Int. J. Onl. Eng. (iJOE) **15**(13), 77–85 (2019)
5. Anzinger, S., et al.: Acoustic transmission line based modelling of microscaled channels and enclosures. J. Acoust. Soc. Ame. **145**(2), 968–976 (2019)
6. Jain, A., et al.: A route selection approach for variable data transmission in wireless sensor networks. Clust. Comput. **23**(3), 1697–1709 (2020)
7. Xu, Y., et al.: An information integration and transmission model of multi-source data for product quality and safety. Inf. Syst. Front. **21**(1), 191–212 (2019)

8. Liu, S., et al.: A fast fractal based compression for MRI images. IEEE Access **7**, 62412–62420 (2019)
9. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. Entropy **21**(9), 902 (2019)
10. Liu, S., et al.: Prediction of gene expression patterns with generalized linear regression model. Front. Genet. **10**, 120 (2019)
11. Zheng, Y.: Formal process virtual machine for smart contracts verification. Int. J. Perform. Eng. **14**(8), 1–9 (2018)