Designing an Elderly Virtual Caregiver Using Dialogue Agents and WebRTC

Chen, Sinan

Nakamura, Masahide

Designing an Elderly Virtual Caregiver Using Dialogue Agents and WebRTC

Sinan Chen [1], Masahide Nakamura [1,2]
[1]Graduate School of System Informatics, Kobe University, 1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
[2]RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo, 103-0027, Japan
Email: chensinan@ws.cs.kobe-u.ac.jp, masa-n@cs.kobe-u.ac.jp

***Abstract*—An aging society has gradually become a worldwide topic. To assist elderly people in daily dialogue and aid memory, a virtual caregiver with dialogue agents has been actively developed. In most virtual caregivers, limitations of system environments or devices exist. Share-ability for elderly people who are so far apart also requires improvement urgently. The purpose of this paper is to design a novel architecture of an elderly virtual caregiver. Our key idea is to regard the online chat room as a virtual house for elderly people, integrating dialogue agents and Web real-time communication (WebRTC). In the proposed architecture, we first extend its useable places and devices by sharing its interface to an online chat room. Then, we regard the action of calling a partner to chat as one of the dialogue scenarios by integrating dialogue agents and WebRTC.**

***Index Terms*—aging society, virtual caregiver, dialogue agents, WebRTC, smart healthcare**

I. Introduction

As social aging intensifies, *healthcare* for elderly people has currently become one of the key issues. It is always challenging to reduce the burden of family members and understand the mental status of elderly people. Especially for those super-aged people with long-term care or living alone, various assistive technologies are flourishing. Within them, using engineering technology to build a `Virtual Caregiver` has been rapidly conducting [1] [2]. It generally refers to using the form of voice interaction between human and agent, to monitor and improve the mental of elderly people. The related work includes *mind sensing* with LINE chatbot [3].

Our interest is in studying and developing a virtual caregiver with dialogue agents. More specifically, we focus on the *spoken* dialogue agent. For this, we are introducing a `Listening Service` (see Section II-B and Figure 1) for elderly people. However, there are two challenges that require improvement: (1) *Limitations* of system environments or devices. (2) *Share-ability* for elderly people who are so far apart. It still cannot be implemented that the elderly chat with a virtual caregiver when him/her are undergoing an outing. Moreover, pushing *human-to-human* dialogue also still cannot be implemented with dialogue agents yet. The goal of this paper is to design a novel architecture of an elderly virtual caregiver. To achieve this goal, we are currently integrating dialogue agents and *Web Real-Time Communication (WebRTC)* [1]. We first propose two modules to extend useable places and devices of virtual caregivers and to push human-to-human dialogues. Then, we clarify that `Address Book Service`, `Chat Room API`, and `Chat Room Management Service` are required in Module 2. The advantage and limitation are contained in Section IV.

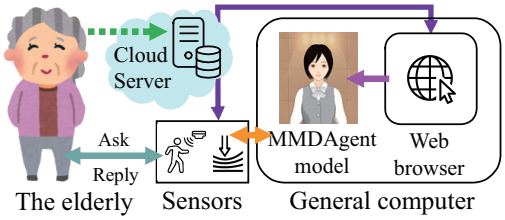

Fig. 1. Overview of a `Listening Service`. Copyright 2009-2018 Nagoya Institute of Technology (MMDAgent Model "Mei").

II. Preliminaries

A. Aging Society and Elderly Care in Japan

As of October 1, 2019, Japan's total population was 126.17 million, with 35.89 million people aged 65 and over, accounting for 28.4% of the total population (the aging rate) [2]. According to the ratio of the population aged 65 and above to the population aged 15 to 64, in 1950 there were 12.1 working-age people (those aged 15 to 64) for every one person aged 65 and above, while in 2015 there were 2.3 working-age people for every one person aged 65 and above. In the future, the aging rate will increase and the ratio of working-age people will decrease, so that by 2065, the ratio will be 1.3 working-age people for every 1 person aged 65 and over.

Against this background, the number of people requiring care has increased 1.6% from 6.58 million to 6.69 million over March 2019 to March 2021. While the Japanese government's financial outlays for medical and welfare facilities are increasing every year [3], there is still a shortage of nursing manpower to deal with the number of elderly patients with chronic diseases (mild cognitive impairment, disuse syndrome, etc.). Specifically, according to the long-term care situation in August 2020, there are 3.9 million people who are conducting care nursing at home [4]. Moreover, 59.7% of these people are the elderly take care of each other and 33.1% are over 75

[1]https://en.wikipedia.org/wiki/WebRTC

[2]https://www8.cao.go.jp/kourei/english/annualreport/2020/pdf/2020.pdf
[3]https://www.mhlw.go.jp/english/wp/wp-hw13/dl/10e.pdf
[4]https://www.jili.or.jp/lifeplan/lifesecurity/1115.html (in Japanese)

TABLE I
DIFFERENCE BETWEEN WEBRTC AND WEBSOCKET TECHNOLOGIES.

WebRTC	WebSocket
Reliable & UnReliable (TCP/UDP)	Reliable (TCP)
Peer ⟷ Peer Model	Client ⟷ Server Model
Low Flexibility	High Flexibility
High-Level APIs	Low-Level APIs

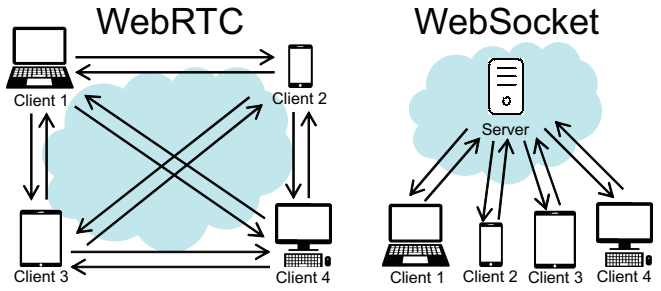

Fig. 2. Example of architecture in WebRTC and WebSocket.

years old. In order to assist the home care of the elderly, many studies around developing virtual caregivers using engineering technology have been rapidly progressing [4] [5] [6] [7] [8].

B. Previous Study: Dialogue Agent as a Virtual Caregiver

In our previous research, to assist the elderly in daily dialogue and aid memory, we have been studying the concept of a dialogue agent as a virtual caregiver for the elderly in smart homes [9] [10], and have been designing architecture [11] [12], agent appearance [13] [14], dialogue scenarios [15] [16], and extended microservices [17]. A dialogue agent is a system that interacts with humans in the form of a *dialogue* using a machine as an intermediary.

Especially in recent research [17] [16], we focus mainly on *spoken* dialogue systems, which are being developed and used as a `Listening Service` for the elderly. Figure 1 shows the overview of a Listening Service. By detecting the presence status of the elderly from a motion sensor and connecting it to a general computer [18], a dialogue agent interface can be driven. This interface consists of MMDAgent (agent appearance, speech synthesis) [19] and a web browser (dialogue window, speech recognition [20]). In this way, the dialogue agent greets the elderly periodically and asks questions (e.g., diet contents, mental or physical state, etc.) on a daily basis. The reply contents from the dialogue agent currently include two modes as follows: (1) Echo a word at random (e.g., Yeah, etc.). (2) Related log summarization and sentiment analysis.

C. Technical Challenges

In the virtual caregiver with dialogue agents, we consider that there are the following two big challenges.

C1: Limitations of system environments or devices
Most existing virtual caregiver services commonly rely on a specified single form, such as mobile applications (i.e., IOS, Android) [5], nursing robots [21], or smart speakers [6]. To realize virtual caregiver services that are widely used for general households at any place freely, a significant issue on how to smoothly execute it with various mobile devices, tablets, and computers has been considered.

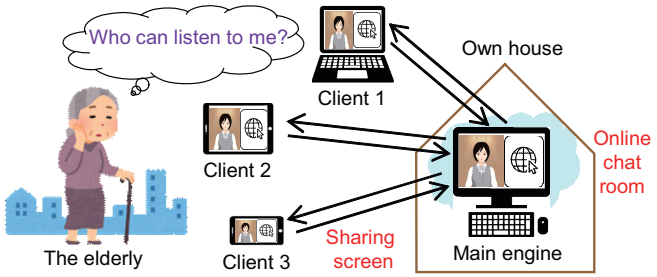

Fig. 3. Exampling *human-to-agent* dialogues extending places and devices.

C2: Share-ability for elderly people who are so far apart
Although a variety of audio telephone [22] and video chat applications [23] with their mature technology are currently pervasive and commonplace, they are still challenging apply to super-aged people and the elderly with dementia. Especially for the elderly living alone, improving remote chat by sharing the same virtual caregiver is urgently needed. It should be simpler to show to the elderly.

D. Web Real-Time Communication

Web real-time communication (RTC) refers to a free and open-source project providing web browsers and mobile applications with RTC via simple *application programming interfaces (APIs)*. It supports video, audio, and generic data to be sent between peers that allows developers to build stronger communication [7]. As depicted in Table I and Figure 2, unlike the conventional approach with *WebSocket* [24], the latter can provide faster communication because its communication directly connects client terminals together without through the server [8]. Hence, develop an elderly virtual caregiver service using dialogue agents and WebRTC is promising.

III. DESIGN OF AN ELDERLY VIRTUAL CAREGIVER

A. Goal and Key Idea

In this section, we focus on improving the two challenges in Section II-C, aiming at designing a more convenient virtual caregiver service. Our key idea is to regard the online chat room as a virtual house for elderly people, integrating dialogue agents and WebRTC. The proposed concept is different from traditional *human-to-agent* dialogue mainly in two aspects: (1) Driving human-to-agent dialogue using the web browser of various system environments and devices (see Figure 3). (2)

[5]https://vitalk.health/
[6]https://assistant.google.com/intl/en/
[7]https://webrtc.org/
[8]http://www.qcontokyo.com/data_2013/ToruYoshikawa_QConTokyo2013.pdf (in Japanese)

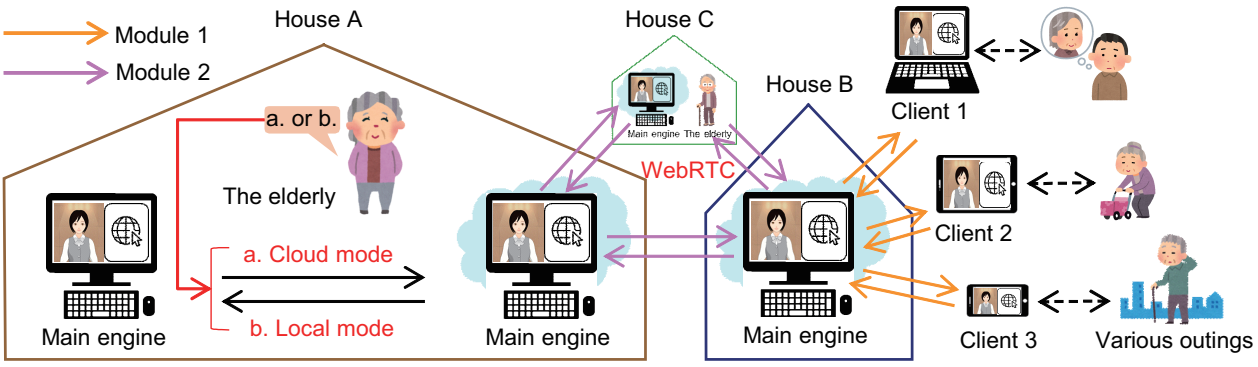

Fig. 4. The proposed architecture includes Module 1 and Module 2.

Pushing remote *human-to-human* dialogue by commanding the keyword into spoken dialogue agents.

B. Proposed Architecture

In the proposed architecture, the novel elderly virtual caregiver with dialogue agents includes two modes: local mode and cloud mode. `Listening Service` (see Section II-B) developed so far belong to the local mode. Its feature is that the elderly can have voice interaction with the main engine only when they are at home. In this paper, we concentrated on cloud mode, including the following two modules: (1) Adding every virtual caregiver to a specified online chat room and sharing its interface, which is called an *external extension*. (2) Developing a `Chat Room API` for dialogue agents, which is called an *internal extension*. The proposed architecture including Module 1 and Module 2 is shown in Figure 4.

C. Module 1: External extension

The purpose of this module is to enhance that the virtual caregiver can be chatted with even when the elderly are undergoing various outings. Its core part is to join the virtual caregiver into the specified chat room and to share the dialogue agent interface. For this, we first set up the virtual caregiver ID as an online chat room ID to easily manage. Building a chat room service with audio, video, and text can refer to the description [9]. Then, a novel feature that shares the screen is required. Unlike the conventional approach to calling the video from the camera device, the key step here is to use a `getDisplayMedia` function. Its source code in detail is shown in Listing 1. In this way, elderly people can chat with a virtual caregiver using the web browser of various devices (e.g., smart phone, tablet, computer) even if they are not at home.

Listing 1. Code sample of getting video stream from the display screen.

```
<script type="text/javascript">
   navigator.mediaDevices
       .getDisplayMedia({ cursor: true })
       .then(stream => {
           const myScreen =
              document.createElement("video");
           addVideoStream(myScreen,
               stream, stream.id);
   });
</script>
```

[9]https://levelup.gitconnected.com/building-a-video-chat-app-with-node-js-socket-io-webrtc-26f46b213017

D. Module 2: Internal extension

The purpose of this module is to push *human-to-human* dialogue using dialogue agents and WebRTC. The core part is to regard the action of calling a partner to chat as one of the dialogue scenarios. Figure 5 shows an example of the dialogue scenario of calling a partner to chat in the web browser. More specifically, developing the following components is required:

- Developing an `Address Book Service` that allows elderly people to add, edit or delete the address ID (i.e., virtual caregiver ID) of the contact with voice interaction.
- Extending a `Chat Room API` that transfers data stream from the partner for embedding dialogue scenarios.
- Creating a `Chat Room Management Service` where the address ID active status of every chat room is monitored.

IV. Conclusion

In this paper, a novel architecture of a virtual caregiver is designed with dialogue agents and WebRTC. The main advantage includes two aspects: (1) By sharing it in an online chat room, elderly people can chat with it using various systems in any place. (2) By integrating chat rooms and dialogue scenarios, the elderly can conveniently call any partner using a virtual caregiver. However, some limitations and lack also exist. For example, the main engine must be kept in communication at home when elderly people are out and would like to chat with them. This will cause a lot of unnecessary consumption of electricity. Moreover, for building the server, privacy and security issues also must be taken seriously. As future work, we will not only start from the technical point but from the current elderly people to develop a smarter virtual caregiver.

Acknowledgment

This research was partially supported by JSPS KAKENHI Grant Numbers JP19H01138, JP18H03242, JP18H03342, JP19H04154, JP19K02973, JP20K11059, JP20H04014, JP20H05706 and Tateishi Science and Technology Foundation (C) (No.2207004).

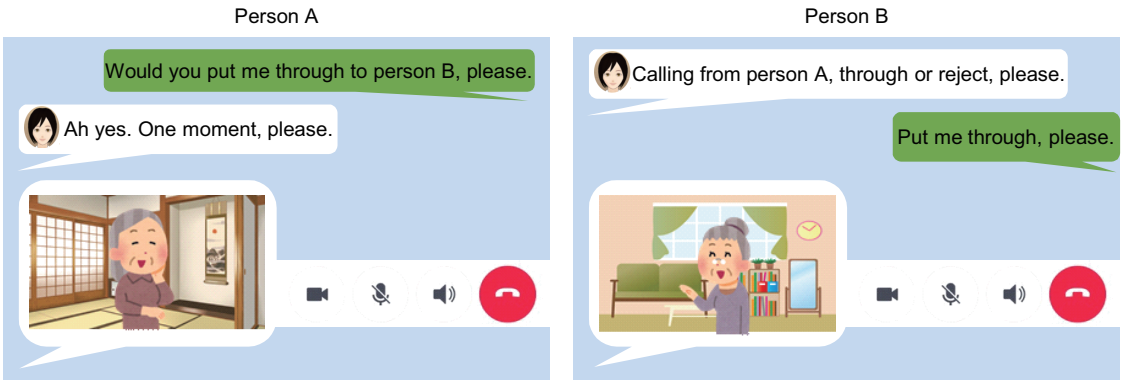

Fig. 5. An example of the dialogue scenario of calling a partner to chat in the web browser.

References

[1] F. Lanza, V. Seidita, and A. Chella, "Agents and robots for collaborating and supporting physicians in healthcare scenarios," *Journal of biomedical informatics*, vol. 108, p. 103483, 2020.

[2] C. Z. Li, E. M. Borycki, and A. W. Kushniruk, "Connecting the world of healthcare virtually: A scoping review on virtual care delivery," in *Healthcare*, vol. 9, no. 10. Multidisciplinary Digital Publishing Institute, 2021, p. 1325.

[3] C. Miura, S. Saiki, M. Nakamura, and K. Yasuda, "Implementing and evaluating feedback feature of mind monitoring service for elderly people at home," in *The 22nd International Conference on Information Integration and Web-based Applications & Services (iiWAS2020)*, November 2020, pp. 390–395.

[4] S. Soda, M. Nakamura, S. Matsumoto, S. Izumi, H. Kawaguchi, and M. Yoshimoto, "Implementing virtual agent as an interface for smart home voice control," in *Asia-Pacific Software Engineering Conference (APSEC2012)*, December 2012, pp. 342–345.

[5] D. Takatsuki, S. Saiki, and M. Nakamura, "Using virtual agent for facilitating online questionnaire surveys," in *International Conference on Signal Processing and Information Security (ICSPIS2018)*, November 2018, dubai, UAE.

[6] S. Nakatani, S. Saiki, M. Nakamura, and K. Yasuda, "Implementation and evaluation of personal ontology building system with virtual agent," in *10th International Conference, DHM 2019, Held as Part of the 21st HCI International Conference, HCII 2019*, vol. LNCS 11582, no. 2. Springer, July 2019, pp. 391–403, orlando, FL, USA.

[7] H. Maeda, S. Saiki, M. Nakamura, and K. Yasuda, "Memory aid service using mind sensing and daily retrospective by virtual agent," in *10th International Conference, DHM 2019, Held as Part of the 21st HCI International Conference, HCII 2019*, vol. LNCS 11582, no. 2. Springer, July 2019, pp. 353–364, orlando, FL, USA.

[8] M. Nakamura, K. Hatano, J. Miyazaki, K. Yasuda, N. Kuwahara, H. KazuiI, S. Saiki, S. Tokunaga, M. Otake, N. Kodama, and N. Kosugi, "Developing a smart system encouraging self-aid and mutual voluntary aid for elderly people at home," in *ISG's 12th World Conference of Gerontechnology*, vol. 19(suppl), October 2020.

[9] H. Horiuchi, S. Saiki, S. Matsumoto, and M. Nakamura, "Virtual agent as a user interface for home network system," *International Journal of Software Innovation*, vol. 3, no. 2, pp. 24–34, April 2015.

[10] S. Tokunaga, K. Tamamizu, S. Saiki, M. Nakamura, and K. Yasuda, "VirtualCareGiver: Personalized smart elderly care," *International Journal of Software Innovation (IJSI)*, vol. 5, no. 1, pp. 30–43, October 2016, dOI: 10.4018/IJSI.2017010103, http://www.igi-global.com/journals/ abstract-announcement/158780.

[11] H. Horiuchi, S. Saiki, S. Matsumoto, and M. Nakamura, "Designing and implementing service framework for virtual agents in home network system," in *2014 15th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD2014)*, June 2014, pp. 343–348.

[12] S. Tokunaga, K. Tamamizu, S. Saiki, M. Nakamura, and K. Yasuda, "Cloud-based personalized home elderly care using smart agent," in *10th World conference of Gerontechnology (ISG2016)*, vol. 15, September 2016, p. 98s.

[13] S. Nakatani, S. Saiki, and M. Nakamura, "Integrating 3d facial model with person-centered care support system for people with dementia," in *the 1st International Conference on Intelligent Human Systems Integration: Integrating People and Intelligent Systems (iHSI 2018)*, January 2018, pp. 216–222.

[14] S. Nakatani, S. Saiki, M. Nakamura, and K. Yasuda, "Generating personalized virtual agent in speech dialogue system for people with dementia," in *Digital Human Modeling 2018 (DHM 2018), Held as Part of HCI International 2018*, vol. LNCS 10917. Springer, July 2018, pp. 326–337.

[15] S. Chen, M. Nakamura, and S. Saiki, "Developing a platform of personalized conversation scenarios for in -home care assistance," in *The 2021 IEEE International Conference on Industry 4.0, Artificial, Intelligence, and Communication Technology (IAICT'21)*, July 2021, pp. 148–153.

[16] S. Chen and M. Nakamura, "Generating personalized dialogues based on conversation log summarization and sentiment analysis," in *The 23nd International Conference on Information Integration and Web-based Applications & Services (iiWAS2021)*, (to appear).

[17] H. Ozono, S. Chen, and M. Nakamura, "Study of microservice execution framework using spoken dialogue agents," in *In 22th IEEE-ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel Distributed Computing (SNPD2021)*, (to appear).

[18] Y. Kashio, S. Matsumoto, S. Tokunaga, S. Saiki, and M. Nakamura, "Design and implementation of service framework for presence sensing in home network system," in *The Third International Conference on Digital Information, Networking, and Wireless Communications (DINWC2015)*, February 2015, pp. 109–114.

[19] A. Lee, K. Oura, and K. Tokuda, "MMDAgent-A fully open-source toolkit for voice interaction systems," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 8382–8385.

[20] J. Adorf, "Web speech api," *KTH Royal Institute of Technology*, 2013.

[21] K. Obayashi and S. Masuyama, "Pilot and feasibility study on elderly support services using communicative robots and monitoring sensors integrated with cloud robotics," *Clinical therapeutics*, vol. 42, no. 2, pp. 364–371, 2020.

[22] C. A. Befort, J. J. VanWormer, C. Desouza, E. F. Ellerbeck, B. Gajewski, K. S. Kimminau, K. A. Greiner, M. G. Perri, A. R. Brown, R. D. Pathak *et al.*, "Effect of behavioral therapy with in-clinic or telephone group visits vs in-clinic individual visits on weight loss among patients with obesity in rural clinical practice: a randomized clinical trial," *JAMA*, vol. 325, no. 4, pp. 363–372, 2021.

[23] M. K. Miller, M. Johannes Dechant, and R. L. Mandryk, "Meeting you, seeing me: The role of social anxiety, visual feedback, and interface layout in a get-to-know-you task via video chat." in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–14.

[24] P. Murley, Z. Ma, J. Mason, M. Bailey, and A. Kharraz, "Websocket adoption and the landscape of the real-time web," in *Proceedings of the Web Conference 2021*, 2021, pp. 1192–1203.