2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress

HiMessage: An Interactive Voice Mail System With the Humanoid Robot Pepper

Paola Barra
Dept. of Computer Science
University of Salerno
Salerno, Italy
pbarra@unisa.it

Carmen Bisogni
Dept. of Computer Science
University of Salerno
Salerno, Italy
cbisogni@unisa.it

Antonio Rapuano
Dept. of Computer Science
University of Salerno
Salerno, Italy
arapuano@unisa.it

Andrea Francesco Abate
Dept. of Computer Science
University of Salerno
Salerno, Italy
abate@unisa.it

Gerardo Iovane
Dept. of Computer Science
University of Salerno
Salerno, Italy
giovane@unisa.it

***Abstract*—This paper presents a method aimed at providing a voice mail system encrypted by a biometric key, using the humanoid robot Pepper. Thanks to a facial recognition algorithm, Pepper is able to recognize the people who approach him. A user recognized by Pepper can record a message to be sent to another user, or can listen to messages in memory sent to him. Sensitive data, such as messages or names stored in memory, are protected by a cryptographic key on the server. In order to evaluate the presented application we selected five users between the ages of 23 and 30. Each user interacted with Pepper ten times and Pepper has been able to end the interaction as user desired in 100% of cases. It shows that the robot can work with Face Recognition and Words Recognition properly.**

***Index Terms*—Humanoid Robot, Pepper, Face Recognition, Voice Mail, Cryptography, Cloud.**

I. Introduction

The growing development of humanoid robotics and human interaction has given the possibility to develop applications at the service of society adapting to every need. Non-invasive sensors allow to collect and use the extracted information, but it creates the need to protect human data.

This paper reports a cloud services application developed for Pepper [2]–[4], a new humanoid robot designed and developed by SoftBank Robotics Corp. and Aldebaran Robotics SAS. in Fig. 1.

A. The personal robot "Pepper"

Pepper [2]–[4] is a personal humanoid robot created by Aldebaran Robotics SAS. and launched by SoftBank Robotics Corp. in 2015, and is the successor of NAO, humanoid robot of the same company [5]. Both humanoid robots have the NAOqi operating system [17]. The robot is particularly suitable for daily interaction with humans, leading people to smarter, safer, healthier, and happier existence, as shown in [1]. Peppers main specifications are shown in the Table I.

Pepper was designed for use with information acquisition and cloud databases, which enables users to expand Pepper's functions by installing new software and various applications (robo-appli).

Below is an overview of the applications that use Pepper's hardware and the sensor component. The three main categories are:

1) Applications for interaction with the environment: focuses on applications to improve the sensors of the robot and its engines. In [6] the problem of the dysfunction of GPS sensors in indoor envoironments is addressed. This problem is solved by increasing the information with the use of the visual sensors of the robot, in particular a 3D mapping of the environment is made to allow the robot to move autonomously avoiding obstacles. The work in [7], in a way similar to the previous one, has the aim of exploiting a mapping of the known indoor environment to allow the robot to travel along a preset trajectory. The purpose of these works is to create an intelligent workspace through the use of Pepper's visual sensors. More technical is the work of [8] which deals with stressing the maximum inclinations of the robot in terms of engines and stability of the latter. With regard to sensors, the work [9] deals with managing physical contacts with the robot, using proprioceptive sensors.

2) Applications for use in medical care: Several works show how it is possible to use humanoid robots in the medical field. The work [10] presents an exploratory survey of the usefulness of the humanoid robot in healthcare especially by anxious patients. The robot can also be seen as an assistant to medical figures. For example, at [11] the robot was used for daily patient data collection, an operation normally performed by human operators. The staff saved time and the patients found the robot able of replacing their work. In the work [12] instead, the robots helped the patients to take drugs correctly, these robots were used as support in a pharmacy in a teaching program to the patients. In addition to hospital patients, the use of humanoid robots also extends to the elderly. In the work [13] it is highlighted how the robot can be made able

978-1-7281-3024-8/19/$31.00 ©2019 IEEE
DOI 10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00123

to perform various tasks such as helping, serving or having a conversation. In this case, the robot learning process is reinforced by user feedback. The rate of human satisfaction in this context exceeds 90%.

3) Applications for social interaction: is one of the most widespread use of Pepper. Social interaction aims at offering services or entertainment. When the robot interacts with an human, he needs to consider all the emotional feedback of the interacting human being. A wide variety of applications have been created to explore the interactions between the robot and the human. The cultural implication of this type of application was discussed at [14], which shows that robots can learn to behave according to the cultural background of the subjects. Also in this case, human feedback represents an essential role for the robot learning and can be emotional, at the same time the robot can respond to stimuli in a way to simulate its emotions [15]. With these tools it becomes possible to use the robots also with children, allowing them to become an interactive educator [16].

Thanks to this overview we have a broad picture of many of the potential of this device, and it is clear that the quantity and quality of data that Pepper can acquire gives us the need to protect them.

The proposed application allows users to use the robot as an answering machine: after the facial identification by the robot, the user logs a message for another registered user; when the robot subsequently meets the user to whom the message was addressed, it will allow the user to listen to the messages left to him or to record new ones. The security of this system is guaranteed by biometric access to messages, the user is allowed to listen or record messages only after the robot has successfully carried out the facial recognition.

The paper is organized as follows: in the next section we address the state of the art, related work in the cloud and in non-humanoid robots that offer answering machine services. Then, in Section III will be presented the method and the interaction experiments with the robot. Finally, Section IV presents our conclusions and the possible future developments of the application will be presented.

II. Related works

It is not trivial to compare our system with the state of the art, currently the secretarial work is easily delegated to voice assistants like Alexa, Siri, Cortana, and more. In the article [18] we find a broad overview of the features offered by these systems, which are exclusively vocal unlike Pepper which also uses motion sensors and cameras. We cannot therefore make a comparison with existing systems but we can investigate the features of the current voice assistants. These offer services such as sending and reading text messages, making phone calls, and sending and reading email. As said before, these systems mainly consists in voice assistants not featuring biometric recognition. Therefore, even ensuring to send the message to a recognized device, they cannot exclusively send the message to a recognized recipient.

Fig. 1. Pepper, the personal robot.

TABLE I
PEPPER TECHNICAL SPECIFICATIONS

Size (H x D x W)	1210 x 425 x 485 [mm]
Weight	28kg
Battery	Li-ion 30.0Ah / 795Wh
Sensors (head)	Mic 4, RGB camera 2 3D sensor 1, touch sensor 3
Sensors (trunk)	Gyroscope sensor 1
Sensors (hand)	Touch sensor 2
Sensors (leg)	Ultrasonic sensor 2, laser sensor 6, bumper sensor 3, gyroscope sensor 1
DOF	20
Display	10.1 inches touchable display
OS	NAOqi OS
Network	Wireless / wired interfaces
Velocity	Max. 3km / h

III. Method and Experiments

The proposed method aims at providing a voicemail system encrypted by a biometric key. This application allows users to interact with the robot in order to record messages for other users or to listen to messages intended for them.

The sensitive information exchanges by users and the system are protected by the biometric access. In fact, to record and listen to messages, the user must first be identified by his/her face.

The information sent by Pepper to the server, and vice versa, is also protected by cryptographic methods.

The method is examined below in its two main aspects, human interaction and the technical details of data processing.

A. Human Interaction

The first step of human interaction with the robot concerns the learning of the user's face, in order to store it within Pepper's memory. This step must be carried out previously with the help of a collaborator and separately from the algorithm flow. It consists in the pre-registration of the face

and the name of the user in order to be recognized later. This procedure must be authorized by the owner of the robot for two main reasons: firstly to prevent this procedure from being carried out by unauthorized users, secondly a human being is needed to ensure the success of the procedure, to avoid non-consensual learning of a face by the Robot. After this enrollment step, the robot is able to recognize the qualified user approaching the field of view of the robot. When a user approaches, Pepper performs the scan and perform a matching between the acquired face and dataset of known faces, if so he turns to the user asking "Hello *Name*, how can I help you?".

After recognition, the user can choose via voice commands if: (1) recording a message to be sent to an user within the database, (2) or listening to a message from another user. The described flow is shown in Fig. 2.

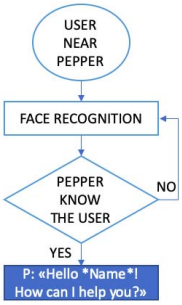

Fig. 2. First interaction with user

1) Leave a message: If the user chooses to leave a message with the words "I want to leave a message", the flow described below will be activated, as in Fig. 3. The information that the robot asks the user for this service are the recipient of the message and the audio recording of the message. The robot asks directly "To whom is the message?" and the user responds with the name of the recipient of the message. Then the robot checks if the recipient is an enrolled user.

- If the matching fails, the robot warns the user and ends the flow "Leave a message". Then return to the initial phase of Face Recognition.
- If the recipient is registered, Pepper notifies the user that it is ready to record the messages.

Recording lasts 10 seconds, this duration is established during the robot installation phase. After registration, the robot warns the user and the "Leave a Message" stream ends. The robot returns to the "Face Recognition" phase in order to serve a new user. The technical details of the registration phase are in paragraph III-B.

2) Listen to a message: If the user chooses to listen to a message with the words "I would to listen to a message", the flow described below will be activated, as in Fig. 4. In this case, the robot needs to know who is the sender of the message and he asks "Who is it from?". When the user tells the sender's name an algorithm check if there is a sender-receiver correspondence in the list of recorded messages.

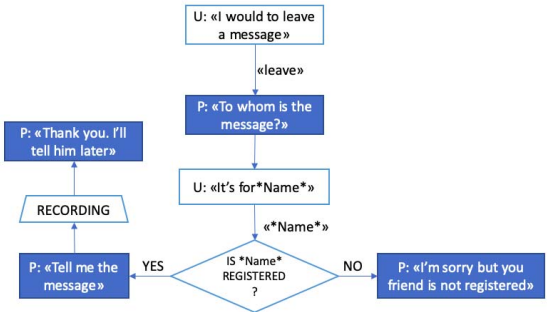

Fig. 3. Interaction to leave a message

- If there is not such a correspondence, the robot warns the user and ends the flow "Listen to a message". Then it returns to the initial phase of Face Recognition.
- If there is a message from sender to user, Pepper performs the face recognition:
  - If the user is the same as the initial interaction, the robot will reproduce the message and at the end it will asks the user if he/she wants to listen to the message again.
  - If the user is not the same as the initial interaction, the robot will perform face recognition two more times. If there is no correspondence, it will return to the initial phase of Face Recognition.

Facial recognition is repeated twice for security reasons. If the user of the initial interaction get away from the robot during the interaction, the audio reproduction will not start. This would ensure the privacy both the sender and the receiver. Furthermore, this dual face recognition, without notice the user, can prevent possible spoofing attacks.

The technical details of the audio reproduction phase are in paragraph III-B.

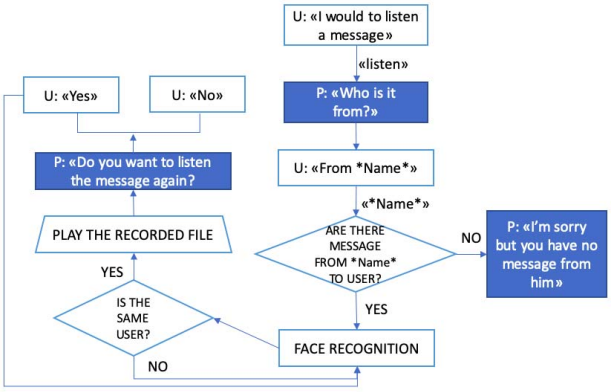

Fig. 4. Interaction to listen a message

B. Storage and listening technical Flow

The proposed method is divided in two systems (see Fig. 5). The first system is stored on Pepper's Android OS and will be

executed from the server through SSH protocol. It has the aim of executing the following steps through Pepper's modules:

- perform the face recognition;
- as user has been recognized, Pepper asks him if he wants to "listen" or "leave" a message;
- as the "leave" keyword is recognized and the recipient of the message is registered, Pepper will listen to the message;
- now, the script will perform an HTTPS request to the server with the parameters: *sender's ID*, *recipient's ID* and the *audio file* in encrypted way;
- the script will delete the file from Pepper's memory.

The second system is stored on a server, it will listen to Pepper's HTTP requests and deals with Google Drive APIs with the aim of uploading the file on Google Drive cloud system. It will store a crypted array containing *sender's ID*, *recipient's ID* and the *audio file*. The array has been crypted with AES symmetric method. The server also executes the "listen" module through NAOqi APIs, this module has the following steps:

- check if the listener and the sender are in the list;
- perform the face recognition;
- if the listener has been recognized, execute the Pepper's tablet module and open the Google Drive audio link.

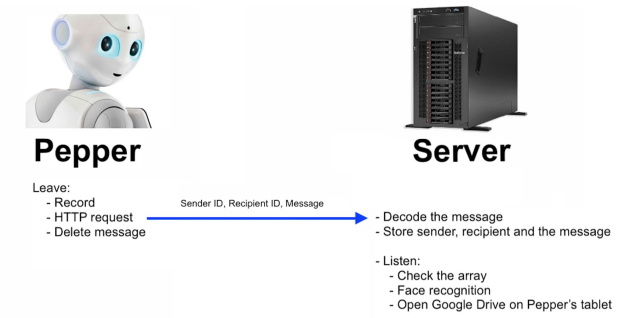

Fig. 5. Pepper and Server interaction

C. Experiments and user experience

In Human-Robot interaction the qualitative indexes are the feelings of the user during the interaction and the ability of the robot to conclude the interaction as user's desired. In order to evaluate these aspects we selected five users between the ages of 23 and 30. After the registration step, we explained them the abilities of the robot and the correct way to interact with it, as key word to leave and listen to a message and the fixed length of the recording. Each user interact with pepper at least ten times and the results are the following:

- Pepper has been able to end the interaction as user desired in 100% of cases. It shows that the robot can work with Face Recognition and Words Recognition properly.
- Users enjoyed themselves during the interaction with Pepper because of the design of the robot, colloquial phrases and easy-to-understand interaction flow.

IV. Conclusions and future works

In conclusion, the idea of using a humanoid robot as a voice mail instead of non humanoid systems has proved to be very useful and funny for the users. In future improvements, we want to make the humanoid robot able to move along path in order to reach the receiver room/office. Another useful improvement will be the automated check of the message of users that the robot meet during a path, in order to notify them and ask if they want to listen to the message. Finally, a totally new functionalities may be added to the robot, like asking to people if they know the user location in order to deliver the message directly. This functionality may be built in the path follow upgrade.

References

[1] Amit Kumar Pandey and Rodolphe Gelin, Pepper: "A Mass-Produced Sociable Humanoid Robot: Pepper: The First Machine of Its Kind", IEEE Robotics & Automation Magazine, Volume: 25 , Issue: 3 , Sept. 2018, pp. 40 - 48.
DOI: 10.1109/MRA.2018.2833157

[2] CNN. Meet Pepper, the emotional robot. Retrieved February 24, 2015, from https://edition.cnn.com/2014/06/06/tech/innovation/pepper-robot-emotions/index.html, 2014.

[3] TIME. Meet Pepper, the Robot Who Can Read Your Emotions. Retrieved February 24, 2015, from https://time.com/2845040/robotemotions-pepper-softbank/, 2014.

[4] IEEE SPECTRUM. How Aldebaran Robotics Built Its Friendly Humanoid Robot, Pepper. Retrieved February 24, 2015, from https://spectrum.ieee.org/robotics/home-robots/how-aldebaran-robotics-built-its-friendly-humanoid-robot-pepper, 2014

[5] Arkadiusz Gardecki and Michal Podpora, "Experience from the operation of the Pepper humanoid robots", 2017 Progress in Applied Electrical Engineering (PAEE), 15 August 2017.
DOI: 10.1109/PAEE.2017.8008994

[6] Eiji Kaneko and Nobuyuki Umezu, "Rapid Construction of Coarse Indoor Map for Mobile Robots", 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE 2017), 21 December 2017.
DOI: 10.1109/GCCE.2017.8229367

[7] Dorota Belanová, Marián Mach, Peter Sinčák and Kaori Yoshida, "Path Planning on Robot Based on D* Lite Algorithm", 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA), August 2018.
DOI: 10.1109/DISA.2018.8490605

[8] Jory Lafaye, Cyrille Collette and Pierre-Brice Wieber, "Model predictive control for tilt recovery of an omnidirectional wheeled humanoid robot", 2015 IEEE International Conference on Robotics and Automation (ICRA), 02 July 2015.
DOI: 10.1109/ICRA.2015.7139914

[9] Anastasia Bolotnikova, Sbastien Courtois and Abderrahmane Kheddar, "Contact Observer for Humanoid Robot Pepper based on Tracking Joint Position Discrepancies", 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 08 November 2018.
DOI: 10.1109/ROMAN.2018.8525774

[10] Sachie Yamada, Tatsuya Nomura and Takayuki Kanda, "Healthcare Support by a Humanoid Robot", 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 25 March 2019.
DOI: 10.1109/HRI.2019.8673072

[11] Daisy van der Putte, Roel Boumans, Mark Neerincx, Marcel Olde Rikkert and Marleen de Mul, "A Social Robot for Autonomous Health Data Acquisition Among Hospitalized Patients: An Exploratory Field Study", 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 25 March 2019.
DOI: 10.1109/HRI.2019.8673280

[12] Keitaro Ishiguro, Saki Minamino, Jun Kawahara and Yukie Majima, "Development of a Robot Intervention Program in Medication Instruction at a Pharmacy", 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI), 18 April 2019.
DOI: 10.1109/IIAI-AAI.2018.00198

[13] Chiao-Yu Yang, Ming-Jen Lu, Shih-Huan Tseng and Li-Chen Fu, "A companion robot for daily care of elders based on homeostasis", 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 13 November 2017.
DOI: 10.23919/SICE.2017.8105748

[14] Thi Le Quyen Dang, Nguyen Tan Viet Tuyen, Sungmoon Jeong and Nak Young Chong, "Encoding cultures in robot emotion representation", 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 14 December 2017.
DOI: 10.1109/ROMAN.2017.8172356

[15] Wen-Feng Shih, Keitaro Naruse and Shih-Hung Wu, Implement human-robot interaction via robot's emotion model, 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST), 15 January 2018.
DOI: 10.1109/ICAwST.2017.8256522

[16] Fumihide Tanaka, Kyosuke Isshiki, Fumiki Takahashi, Manabu Uekusa, Rumiko Sei and Kaname Hayashi, "Pepper learns together with children: Development of an educational application", 2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids), 28 December 2015.
DOI: 10.1109/HUMANOIDS.2015.7363546

[17] Softbank Robotics Documentation http://doc.aldebaran.com/2-5/index.html

[18] Matthew B. Hoy, "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants", Medical Reference Services Quarterly, Volume 37, 2018 - Issue 1, pp 81-88.
DOI:10.1080/02763869.2018.1404391