PAPER NAME

Arduino_Raspberry_Pi_and_Smartphone_Usage_Comparison_for_Voice-based_Virtual_Assistant.pdf

WORD COUNT	CHARACTER COUNT
3355 Words	**18409 Characters**
PAGE COUNT	FILE SIZE
5 Pages	**349.1KB**
SUBMISSION DATE	REPORT DATE
Nov 27, 2023 2:33 PM GMT+7	**Nov 27, 2023 2:34 PM GMT+7**

7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 3% Internet database
- 3% Publications database
- Crossref database
- Crossref Posted Content database
- 3% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 8 words)
- Manually excluded sources
- Manually excluded text blocks

Arduino, Raspberry Pi, and Smartphone Usage Comparison for Voice-based Virtual Assistant

Ridwan Sanjaya
Information System Department
Soegijapranata Catholic University
Semarang, Indonesia
ridwan@unika.ac.id

Christin Wibhowo
Psychology Department
Soegijapranata Catholic University
Semarang, Indonesia
christine@unika.ac.id

2022 6th International Conference on Information Technology (InCIT) | 978-1-6654-8912-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/InCIT56086.2022.10067741

Abstract—**A virtual assistant or an artificial intelligence chatbot is one of the technological interventions to help individuals with borderline personality disorder (BPD). The virtual assistant must have a bank of knowledge to answer the questions from individuals with BPD to accompany and chat with them like a digital friend. As part of artificial intelligence software, Program-O can handle the interaction between users and the bank of knowledge in the server. The voice recognition and text-to-speech features can be used to enhance the virtual assistant's function to become even closer as a friend. There are three possibilities for implementing voice recognition and text-to-speech features on virtual assistants, which are using Arduino, Raspberry Pi, and smartphone devices. All these three devices can be connected to the speech recognition module or Application Protocol Interface (API) for voice recognition on the internet. This paper will explore and compare several aspects of the implementation of voice recognition and text-to-speech features in the virtual assistant. The result will be a recommendation for the application of voice recognition and text-to-speech features on a virtual assistant for individuals with BPD.**

Keywords- borderline personality disorder; chatbot; digital friend; text to speech; virtual assistant, voice recognition

I. INTRODUCTION

Individuals with borderline personality disorder (BPD) have a high risk of suicide. In previous studies [1], 9-33% of suicides were committed by individuals with BPD. There is no single most effective and sustainable approach to prevent them from committing suicide [2]. However, accompanying them will create opportunities to make them more stable, less hopeless, and less alone.

Creating a digital friend such as a virtual assistant with artificial intelligence and the ability to respond to them when they need someone to talk to is one approach that has been tried in previous research [3]. The advantage of the virtual assistant is that they can accompany them within 24 hours when BPD symptoms appear. In addition, their confidentiality is also better maintained because the virtual assistant does not inform other people about their conversation.

Technically, a voice-based virtual assistant shown consists of a program with a bank of knowledge that can answer based on question patterns, a voice recognizer, and text-to-voice. In previous studies, mobile applications whose gadgets are equipped with voice recognition and text-to-speech features that are connected to a computer server equipped with a knowledge database have been able to answer these needs. The bank of knowledge can be improved according to the needs and development of the case.

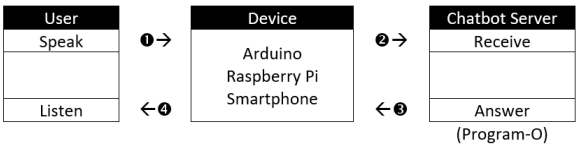

Figure 1. Communication flow in voice-based chatbot

Portable devices like Google Nest and Amazon Alexa should be the solution to answer their needs more interactively. Google Nest and Amazon Alexa are examples of smart speakers that come with Google Assistant. Some activities, such as playing music, turning on the television, asking about the weather, or planning activities, can be done using the user's voice.

However, the device is limited to some functions and cannot connect to the chatbot. To realize a virtual assistant in the form of a portable device, Arduino and Raspberry Pi were explored to be used and connected to a server with a knowledge database. These devices can connect to the server and have the ability to recognize voice using an additional module, a minimum requirement for the voice function on the virtual assistant.

Arduino is an open-source electronic circuit board whose main component is a microcontroller chip with additional components that can be customized according to specific needs. Arduino can be added with a Wi-Fi module to have functionality on an internet connection. The circuit board can be added with a speech recognition module to understand the user's voice. Each addition of modules and functions will increase the power needed by Arduino to work.

The Raspberry Pi is a single-board pocket-sized minicomputer that is equipped with an operating system and can be used to run computer applications. It has Bluetooth, Wi-Fi, and ethernet built-in devices. Like the Arduino board, the Raspberry Pi board can be connected to additional modules to gain access to input/output devices used by desktop computers.

Several aspects, such as the total cost of providing device components, ease of assembling the device, and device functionality of these two devices, will be compared with a smartphone to determine which

978-1-6654-8912-6/22/$31.00 ©2022 IEEE

devices are recommended to serve the function as virtual assistants for individuals with BPD. The results can also determine the possibility of the device being developed in other forms, such as smart speakers or other smart devices.

II. CHATBOT KNOWLEDGE MANAGEMENT

Program-O is an open-source web-based application used for knowledge management needed by chatbots in responding to individuals with BPD [3]. The source of the application can be downloaded via the URL https://github.com/Program-O/Program-O. It is used to manage knowledge in Artificial Intelligence Markup Language (AIML) format in terms of adding, changing, deleting, displaying, searching, training, and compiling knowledge formats in AIML.

AIML as a source of knowledge for chatbots must be created using unique question patterns from each other, previous conversation references, or conversation topic groupings. When the patterns of questions are arranged to overlap each other, the resulting responses can be confusing and different from the expectations. For the response to meet expectations, it is necessary to take full advantage of the various tags used in AIML, especially in terms of separating question topics and associating the responses to the context of the previous conversation.

To give different responses to each topic's conversation, AIML has the <topic> tag. Even though the questions are the same, the answers will differ based on the topic set in the previous conversation using the <set> tag. The example of using the <topic> tag in AIML can be seen in Figure 2. When the users talk about movies, the following responses about the marvel and comedy keywords will be different when the users talk about music. The conversation about the pop keyword can be responded to only when the users talk about music.

```
<aiml>
   <category>
      <pattern>* MOVIES</pattern>
      <template>Let's talk about <set name="topic">movies</set></template>
   </category>
   <category>
      <pattern>* MUSIC</pattern>
      <template>Talk about <set name="topic">music</set></template>
   </category>
   <topic name="music">
      <category>
         <pattern>* POP</pattern>
         <template>I do not like pop music</template>
      </category>
      <category>
         <pattern>* MARVEL</pattern>
         <template>I do not have any marvel music soundtrack.</template>
      </category>
      <category>
         <pattern>* COMEDY</pattern>
         <template>Comedy music? I prefer podcast.</template>
      </category>
   </topic>
   <topic name="movies">
      <category>
         <pattern>* MARVEL</pattern>
         <template>Watching Marvel movie stimulates our minds.</template>
      </category>
      <category>
         <pattern>* COMEDY</pattern>
         <template>I like comedy movies too.</template>
      </category>
   </topic>
</aiml>
```

Figure 2. Using tag <topic> to separate topics.

To continue the previous conversation, AIML provides <that> tag to associate the answers with the previous chatbot's responses. Even though the questions are the same, the answers will differ based on the previous chatbot's responses mentioned in <that> tag, as seen in Figure 3.

```
<category>
   <pattern>* HOBBIES</pattern>
   <template>What hobbies do you like?</template>
</category>
<category>
   <pattern>* BASKET</pattern>
   <that>What hobbies do you like</that>
   <template>Do you know Michael Jordan?</template>
</category>
<category>
   <pattern>YES</pattern>
   <that>Do you know Michael Jordan</that>
   <template>Wow I know your age</template>
</category>
<category>
   <pattern>NO</pattern>
   <that>Do you know Michael Jordan</that>
   <template>I understand</template>
</category>
```

Figure 3. Using tag <that> to associate with the previous conversation.

AIML also has the <condition> tag to set the chatbot's response based on the first conversation. Consistent responses will keep users talking in the context of the previous conversation and not having to speak in complete sentences. The example of using the <condition> tag can be seen in Figure 4. When users talk about Bangkok, the following answer is all about Bangkok. The same condition can also be seen when the user talks about the city of Manila and the city of Semarang. The responses will all be related to that city.

```
<category>
    <pattern>* KNOW *</pattern>
    <template>
        <think><set name="quest"><star index="2"/></set></think>
        <condition name="quest">
            <li value="Bangkok">Bangkok is a City of Angels</li>
            <li value="Manila">Manila is a Pearl of Orient</li>
            <li value="Semarang">Semarang is the Venice of Java</li>
            <li>Huum I do not know</li>
        </condition>
    </template>
</category>
<category>
    <pattern>WHERE *</pattern>
    <template>
        <condition name="quest">
            <li value="Bangkok">Bangkok is in Thailand</li>
            <li value="Manila">Manila is in the Philippines</li>
            <li value="Semarang">Semarang is in Indonesia</li>
            <li>I am so sorry</li>
        </condition>
    </template>
</category>
```

Figure 4. Using tag <condition> to follow the conversation.

The chatbot knowledge management should be done to make the individuals with BPD feel comfortable and not feel annoyed by having to repeat questions or because of responses that are not contextually appropriate. The work result of chatbot knowledge management is expected to increase the capability of responding to users. The chatbot responses can be processed in other forms according to the expected target.

Other applications can read the responses from Program-O by setting the output as JSON or XML format. The output will be used to relate the question from the user with the answer stored in the O-Program using any programming language. An example of an application that uses JSON output from Program-O is a LINE-based Virtual Friend named Sovi Lau, which stands for "Sobat Virtual Anti Galau" or "Anti-Stress Virtual Friend" [4]. LINE Messaging API will communicate with the chatbot server by sending the user's message, receiving the JSON-based response,

and displaying responses in sentences humans can understand.

Chatbot responses can be engineered into voice to give users new experiences like talking to friends and might be needed by individuals with BPD.

III. VOICE-BASED CHATBOT DEVELOPMENT

In implementing a voice-based chatbot on all those devices as virtual friends for individuals with BPD, as shown in Figure 5, several additional components are needed to perform the following functions: listening to the user's voice, converting voice to text, sending a text to the chatbot server, receiving text responses from the server chatbots, and converting text to voice.

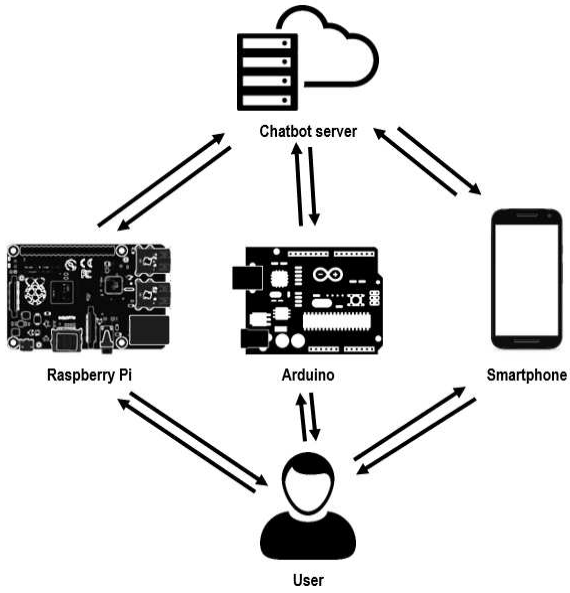

Figure 5. Voice-based chatbot Scenario

A. Arduino

The use of Arduino devices requires additional modules for voice recognition and the internet of things (IoT) [5] as shown in Figure 6. The voice recognition module is used to mediate the user's voice recognition and process it in the microcontroller. While the IoT module is used to send commands from the microcontroller to the chatbot server and receive responses from the chatbot server. The ESPduino-32 board can be used to facilitate the provision of an Arduino that has been equipped with an ESP-32 module for Wi-Fi connection [6]. Voice Recognition Module v3.1 can detect and change voices in certain activities [7]. Approximately all those devices will cost around 30-40 USD.

User		Device		Chatbot Server
Speak	❶→	ESPduino-32 Voice Recognition Module v3.1	❷→	Receive
Listen	←❹	Text-to-Speech API	←❸	Answer

(Program-O)

Figure 6. Arduino in voice-based chatbot

B. Raspberry Pi

Raspberry Pi, as a minicomputer [8], has complete functions to process every task, including the Wi-Fi function for transferring the text or any requests to the chatbot server. The board only needs a GPIO expansion board, a ReSpeaker 4-Mic Array module, and a USB microphone as additional devices as shown in Figure 7. The ReSpeaker 4-Mic Array module functions to recognize the user's voice [9], the USB microphone functions as the user's voice input, and the GPIO expansion board functions to mediate the connection between the Raspberry Pi board and ReSpeaker 4-Mic Array module. Approximately all those devices will cost around 200-210 USD.

User		Device		Chatbot Server
Speak	❶→	Raspberry Pi 4 Expansion Board GPIO ReSpeaker 4-Mic Array Mic USB Microphone	❷→	Receive
Listen	←❹	Text-to-Speech API	←❸	Answer

(Program-O)

Figure 7. Raspberry Pi in voice-based chatbot

C. Smartphone

Current smartphones can recognize the voice and facilitate text-to-speech. It only requires the mobile application to receive the user's voice, convert the voice to text, send the text to the chatbot server, wait for the text response from the chatbot server, and convert the text to voice, as seen in Figure 8. The application developers can only use MIT App Inventor to provide these functions within the application [10]. This tool has limited skill requirements to operate website-based applications to create Android-based mobile applications [11]. The minimum specifications of a smartphone with the required capabilities will cost around 30-35 USD, depending on the smartphone's brand.

User		Device		Chatbot Server
Speak	❶→	Smartphone's Voice Recognition Text-to-Speech	❷→	Receive
Listen	←❹		←❸	Answer

(Program-O)

Figure 8. Smartphone in voice-based chatbot

IV. EVALUATION OF DEVICES

Several factors must be considered when deciding which voice-based chatbot device to develop. These factors are cost, difficulty, and functionality. The cost factor is the money spent to develop the device's intended function. The difficulty factor is the effort of providing hardware, technical aspects, and developing the supporting application. The functional factor is the function of understanding the voice, connecting with the chatbot server, and displaying the chatbot response.

A. Device Cost

In assembling the supporting components of a voice-based chatbot using Arduino as the main device, the cost is around 30-40 USD. Assembling a

Raspberry Pi with a voice-based chatbot support devices costs around 200-210 USD. The cheapest smartphone purchases cost around 30-35 USD or more depending on the smartphone brand. The most expensive expense is in assembling the Raspberry Pi. The other two options cost in the same range. The cost comparison of the three devices can be seen in table I.

TABLE I. COST RANK

Device	Cost	Rank
Arduino	30-40 USD	2
Raspberry Pi	200-210 USD	1
Smartphone	30-35 USD	2

B. Device Assembly Difficulty

Raspberry Pi as a minicomputer has the minimum specifications required for office work, but for multimedia purposes, additional components are needed. In terms of providing additional components, the provision of a ReSpeaker 4-mic Array module connected to the Raspberry Pi via the GPIO Expansion Board and voice input using a USB Microphone Mic connected to the Raspberry Pi has the same difficulty as adding components on the Arduino board circuit with ESP32 components to Wi-Fi and Voice Module v3.1 components for voice recognition. However, the use of ESPduino-32 which is a combination of Arduino board and ESP-32 module makes it a little easier and reduces the number of components that need to be provided. The current smartphones have all the functions needed so additional devices to work as planned is not needed. The order of providing the most difficult hardware is Raspberry Pi, Arduino, and smartphone.

After providing the hardware, it is needed to assemble all the components with the main board. Even though Raspberry Pi and Arduino almost have similar effort to assemble, Raspberry Pi has more components to manage than Arduino. smartphones are ready to be used. Users can use all functions needed without any technical intervention. The order of the most difficult technical aspects is Raspberry Pi, Arduino, and smartphone.

A similar situation also occurs when installing programs to make the three devices function as planned. Installing some programs and training the hardware are needed to understand the users' voice. Almost the same as the other two aspects, Raspberry Pi and Arduino almost have the same difficulty, but the Raspberry Pi has more programs to be installed than in the Arduino because Raspberry Pi has an operating system, but Arduino does not. However, each hardware has to be connected to an API to have the needed voice function. The smartphone has operating system features and can work as targeted without any additional APIs. Users only use the existing features through easy procedures. The order of the most difficult application development is Raspberry Pi, Arduino, and smartphone. The comparison of difficulty levels in hardware provision, technical aspects, and application development discussed above can be seen in table II.

TABLE II. LEVEL OF DIFFICULTY

Device	Level of Difficulty		
	Hardware Provision	*Technical Aspects*	*Application Development*
Arduino	2	2	2
Raspberry Pi	1	1	1
Smartphone	3	3	3

C. Device Functionality

Those three tools can achieve the most targeted functions but at different levels of implementation. For example, Arduino can recognize voice only in the form of a word but not in a sentence. Each recognized word will trigger an action on the Arduino. However, Raspberry Pi and Smartphone can recognize the voice in the form of a word or a sentence.

Arduino can perform its functions on the chatbot server connection through the ESP32 module that functions as Wi-Fi. While on the Raspberry Pi, this feature is embedded in the minimum Raspberry Pi package. The same thing is found on smartphones with Wi-Fi by default features. It can be said all three can perform the function of connecting to the chatbot server.

However, compared to the two device's functional experiments above, different results are obtained in getting responses from the chatbot server. Arduino is failed to provide chatbot responses because the voice recognition function through Voice Recognition Module v3.1 is limited to a keyword. The complete words cannot be recognized and an appropriate response cannot be given. Arduino cannot give the expected response because it can only detect one word for each action. The expected answer from responding to complete sentences cannot be achieved. Both Raspberry and smartphones can recognize long sentences and can provide responses from appropriate chatbots. The ReSpeaker 4-Mic Array module connected to the Raspberry Pi enables speech recognition and functions within Amazon Alexa and Google Home. While the voice features that have been embedded in the smartphone make it a facility from the start. The comparison of functionality in voice recognition, connection to chatbot server, and chatbot responses for the three devices discussed above can be seen in table III.

TABLE III. DEVICE'S FUNCTIONALITY

Device	Functionality		
	Voice recognition	*Connection to chatbot server*	*Chatbot responses*
Arduino	V	V	X
Raspberry Pi	V	V	V
Smartphone	V	V	V

V. CONCLUSIONS

The devices used to integrate voice recognition features with AIML-based chatbots that function as needed are smartphones and Raspberry Pi. But in terms of cost, a smartphone with low specifications for 30-35 USD can be used for these needs. While the cost of Raspberry Pi reaches 200-210 USD. In terms of difficulty level, Raspberry Pi has a greater level of difficulty than a smartphone.

The integration of voice recognition features with an AIML-based chatbot can be done by utilizing the speech recognition and text-to-speech features in smartphones and available features in making applications using MIT App Inventor. Speech recognition translates the user's voice into text and sends it to the chatbot server. Then, the chatbot server will receive the sent sentence and answer based on its word pattern. By a smartphone, the answer will be converted into voice through the text-to-speech feature and played for the user.

Overall, based on the benchmarking of the three devices, the use of smartphones is recommended in presenting voice recognition features in chatbots as it is faster, cheaper, and easier. Making a smartphone application using MIT App Inventor can already activate the voice recognition and text-to-speech functions needed for a voice-based chatbot.

ACKNOWLEDGMENT

This research was funded by the Indonesian Ministry of Research and Higher Education under the scheme of Higher Education Fundamental Research Excellence Grant in 2022 titled Borderline Personality Disorder Assistance Model using Virtual Assistant Technology.

REFERENCES

[1] M. Pompili, P. Girardi, A. Ruberto, and R. Tatarelli, "Suicide in borderline personality disorder: a meta-analysis," *Nord. J. Psychiatry*, vol. 59, no. 5, pp. 319–324, 2005.

[2] J. Paris, "Suicidality in Borderline Personality Disorder," *Medicina (B. Aires).*, vol. 55, no. 6, p. 223, 2019.

[3] C. Wibhowo and S. Ridwan, "Virtual Assistant to Suicide Prevention in Individuals with Borderline Personality Disorder," in *International Conference on Computer and Information Science*, 2020.

[4] L. V. Oey, R. Sanjaya, and C. Wibhowo, "LINE-based Virtual Friend Development for Borderline Personality Disorder," in *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 2021, pp. 80–85.

[5] M. Schwartz, *Internet of Things with Arduino Cookbook*. Birmingham, UK: Packt Publishing Ltd, 2016.

[6] A. Hidayat *et al.*, "Designing IoT-Based Independent Pulse Oximetry Kit as an Early Detection Tool for Covid-19 Symptoms," in *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)*, 2020, pp. 443–448.

[7] K. Khotimah *et al.*, "Validation of Voice Recognition in Various Google Voice Languages using Voice Recognition Module V3 Based on Microcontroller," in *2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE)*, 2020, pp. 1–6.

[8] J. C. Shovic, *Raspberry Pi IoT Projects*, Second. Apress, 2021.

[9] D. Yang, L. Ma, and F. Liao, "An Intelligent Voice Interaction System Based on Raspberry Pi," in *2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2019, pp. 237–240.

[10] D. D. Prasad, G. J. Mallika, S. U. Farooq, A. Tanmaie, D. R. Krishna, and S. Pramod, "Voice Controlled Home Automation," in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2021, pp. 673–678.

[11] D. Wolber, H. Abelson, E. Spertus, and L. Looney, *App Inventor 2: Create Your Own Android Apps*. Sebastopol, CA: O'Reilly Media, 2015.

7% Overall Similarity

Top sources found in the following databases:

- 3% Internet database
- Crossref database
- 3% Submitted Works database
- 3% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Christin Wibhowo, Ridwan Sanjaya. "E-Learning Design for Psychologis... Crossref	2%
2	University of Queensland on 2016-05-06 Submitted works	<1%
3	Florida International University on 2023-04-22 Submitted works	<1%
4	The University of Wolverhampton on 2023-05-04 Submitted works	<1%
5	University of Queensland on 2019-10-10 Submitted works	<1%
6	seaninstitute.org Internet	<1%
7	Christin Wibhowo, Ridwan Sanjaya. "Virtual Assistant to Suicide Preven... Crossref	<1%
8	Ping-Yi Yen, Yong-Yi Fanjiang, Chi-Huang Hung, Tzu-Min Zhang. "Desig... Crossref	<1%

9	University of Wales Swansea on 2023-04-28 Submitted works	<1%
10	Rochester Institute of Technology on 2002-11-16 Submitted works	<1%
11	University of Glasgow on 2012-11-25 Submitted works	<1%
12	thesai.org Internet	<1%
13	psych.org Internet	<1%
14	researchgate.net Internet	<1%
15	Basith Adli, Pranoto Hidaya Rusmin. "Vibration Measuring Tools For Ro... Crossref	<1%
16	National College of Ireland on 2019-12-14 Submitted works	<1%
17	ejurnal.stmik-budidarma.ac.id Internet	<1%

Excluded from Similarity Report

- Bibliographic material
- Cited material
- Manually excluded sources
- Quoted material
- Small Matches (Less then 8 words)
- Manually excluded text blocks

EXCLUDED SOURCES

Ridwan Sanjaya, Christin Wibhowo. "Arduino, Raspberry Pi, and Smartphone U... 85%
Crossref

repository.unika.ac.id 85%
Internet

EXCLUDED TEXT BLOCKS

Arduino, Raspberry Pi, and Smartphone UsageComparison for Voice-based Virtual ...
sinta.lldikti6.id

Soegijapranata Catholic UniversitySemarang, Indonesiaridwan@unika.ac.idAbstract
technodocbox.com

Christin WibhowoPsychology DepartmentSoegijapranata Catholic UniversitySemar...
Christin Wibhowo, Ridwan Sanjaya. "E-Learning Design for Psychologists to Implement Chatbots for Clients ...

2022 6th International Conference on Information Technology (InCIT) | 978-1-6654...
Theresia Dwi Hastuti, Ridwan Sanjaya, Benediktus Danang Setianto. "Blowing the Truth: Strengths and Weak...

Authorized licensed use limited to: National Yang Ming Chiao Tung University. Do...
people.cs.nctu.edu.tw

Authorized licensed use limited to: National Yang Ming Chiao Tung University. Do...
people.cs.nctu.edu.tw