 https://doi.org/10.24191/jeesr.v21i1.008

Feasible Analysis on Hardware Security Module With Digital Voice Assistants Device

Rizzo Mungka Rechie, Yusnani Mohd Yussoff, Lucyantie Mazalan and Suhairi Mohd Jawi @ Said

***Abstract*—The advancement of the internet of things technology in the recent years is moving forward drastically due to the recent global pandemic started on the year 2020. Consumer product such as Digital Voice Assistant Device (DVAD) has become popular and is highly in demand. A security vulnerability is one of the major problems related to the IoT devices technology. One of the problem is the user data integrity issue, where DVAD users have no awareness of how their data is being used in the DVAD and if their DVAD are being tampered. With this in mind, the objective of this study is to investigate the feasibility and effect of hardware security module (HSM) onto a DVAD. A layer of security hardware is proposed, and the analysis focuses on the encryption ability, energy, and resource consumption are investigated. The findings revealed the HSM stack upon the DVAD is highly feasible and secure. The contribution of this study will add value to the existing literature on security for IoT technology. Besides, it is important for hardware-based security to consider the effect of feasibility which helps to improve the performance of such devices.**

***Index Terms*— cybersecurity, IoT, voice assistant**

I. INTRODUCTION

DIGITAL Voice Assistant Device is a digital assistant software that runs or performs task based on the user's command. The function is to provide information or to run tasks such as detecting the music name, reporting the weather forecast for a given day, setting an alarm, checking work schedule, and many more. Nevertheless, several important information such as the user's location, contact number, and the audio or video service of the DVAD device are always turned on to enable the DVAD service running smoothly.

This manuscript is submitted on 31st January 2022 and accepted on 12th April 2022. This research study is supported by Universiti Teknologi MARA under the research grant Geran Penyelidikan Lestari no: 600-RMC/MyRA 5/3/LESTARI (098/2020) in collaboration with MIMOS Berhad, Malaysia. R. Mungka, Y. Yussoff and L. Mazalan is with the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor. S. Jawi is with Cybersecurity Malaysia, 63000 Cyberjaya, Selangor.

*Corresponding author
Email address: mungkarizzo@gmail.com

The voice assistant runs on smartphones, smart TVs, and DVAD devices such as Amazon Echo Dot or Google Home. This device is connected to our home wireless router for it to work seamlessly. Apart from that, the DVAD is also able to perform home automation such as controlling the home lights, a smart TV, and smart air conditioner. All the home devices that can to be controlled by the DVAD must be configured with the DVAD applications itself and must be compatible with the system software for it to work flawlessly.

Apart from its usefulness, the major concern is the security issue, which involves the integrity of the user data when it is being kept inside the device. The need of hardware security module (HSM) inside a DVAD is crucial due to its high usage in the year 2020. This HSM, Zymkey 4i to be specific, can be implemented and become an added security layer towards the home or office device. This security device will secure the user file and perform an integrity check for every time the device boots up.

This paper will discuss the objectives and contribution of the study. A review of the voice assistant known vulnerability and different areas of studies regarding HSM. In regards to that, a known vulnerability to voice assistant is tabulated and highlight the focus of a vulnerability this study. Afterwards a function comparison between different HSM are being made to see how different HSM protects its devices. At the methodology section, it explains on how is the DVAD environment set up and its security scenario. Then at the result section, it shows the RAM usage when the HSM are being implemented onto the DVAD and the security test when the device is being tampered.

This study is focused on the security and the integrity of the device data. As a result, the study had shown that the implementation of Zymkey 4i as HSM into the DVAD is feasible and secure.

The main contribution of this paper are summarized as follows:

1) Integrate the Zymkey 4i as HSM onto a DVAD.
2) Mitigate DVAD tampering issue on data integrity.
3) Investigate the feasibility of the HSM onto the DVAD to run daily task with the enhancement of security features.

This study is aimed at enhancing the Digital Voice Assistant Device at home or at the office for personal or group use. This paper is highly motivated and focused on the device data

encryption security features. The objectives are as follows:

1) To investigate the encryption impact performance of the HSM on the DVAD.
2) To study the feasibility of the HSM with the DVAD based on energy and resource consumption.

II. Related Works

A. Vulnerability of The Digital Voice Assistant Device

The vulnerability of the DVAD varies in all aspects, such as its protocol and hardware security. The most common vulnerabilities found in the device derive from hardware components of the device. This is due to the microphone and speaker of the device are designed to accept input or produce output automatically. Without any security measure for microphone input, the device can accept and run any task from anonymous users without verifying the integrity or authentication of the users' identity.

A study had been made on the concerning security issue [1], which is the Voice Squatting Attack towards DVAD. This attack will take advantage of the user's voice commands and mislead its services into the attackers preference. The example of the voice squatting attack is as capturing the command as "what is the weder today" instead of "what is the weather today" as intended. As the attack is being made, the user command will mislead the user voice command and will run other task rather than telling the exact weather of the day.

This study is based on the vulnerability of a voice-driven interface is very much related to this DVAD security issue. One of the concerning issues is the ecosystem of the DVAD. The DVAD ecosystem poses security vulnerabilities as well as threats, as the device enables user data to circulate [2] around the smart home or office. In regards to the data circulation around the home premise due to the smart home device ecosystem. Another threat to users is the smart assistant device data security, whereby a study had shown that the Amazon Alexa [3] can constantly hear users' voices in standby mode and the data stored are anonymously in their online database. This produces a concern for users as these smart devices are built to make their lives easier, but in return the home privacy is being compromised by a single microphone contained on the smart assistant device. This has further strengthened the need to enhance the hardware security to further secure the device from any voice attacks.

Another voice driven vulnerability [4] on a voice assistant device is the inaudible attack, which is known as the DolphinAttack. The studies show that the attack is more focused on the hardware instead of software. The attack is focused on taking advantage of the voice assistant's microphone and the way to mitigate it is by doing the microphone enhancement and baseband cancellation. The microphone is to be enhanced and designed to suppress any acoustic signals whose frequencies are in the ultrasound range to secure the voice assistant. In the case to make a Inaudible Voice Command Cancellation, a module can be enhanced prior to lowpass filters (LPF) to detect the DolphinAttack and cancel the demodulated voice commands. The cancelled signals are signals that are within the ultrasound frequency range that exhibit AM modulation characteristics. The Inaudible Voice Command Cancellation method is designed to capture sounds. Whereas, the voice capture subsystem is designed to suppress signals out of the frequency range of audible sounds that operates from 20 Hz to 20 kHz. Prior to the DVAD study, the need for device or hardware enhancement to further secure the device is highlighted, where some attacks can be mitigated efficiently by enhancing it.

Figure 1 shows the DVAD Ecosystem from the smart device such as laptop, smart light connected to the DVAD and subsequently, to the Internet cloud from the gateway. The highlight of this ecosystem is the user's privacy regarding browsing or command histories, which are contained inside the DVAD prohibits users from observing and editing their history files. The analysis has been done by using Forensic Toolkit (FTK Imager) towards the DVAD and it was found that the user's browsing history, contacts, history, and location that data are kept can be revealed.

From preliminary publications relating to this research, it is apparent that the confidentiality of a system storage is very crucial as it involves a human personal data in a commercial item. A further security enhancement is needed at the device to secure users' data.

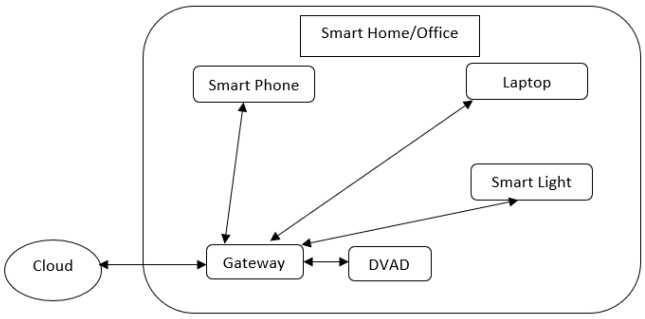

Fig. 1 DVAD Ecosystem

B. Voice Assistants Security

There are multiple number of voice assistants to be used by global consumers. In regards to that, a number of studies on voice assistants security issue has been made.

Table 1 shows the HSM security function comparison based on the device security issue. The study on replay attack mitigation had emphasise on the device microphone input where they use HOSA features [5] to identify the replay attack that is being carried out towards the device.

Another study on the DolphinAttack is focusing on the same voice asstant component which is the microphone. The DolphinAttack is an acoustic signals whose frequencies are in the ultrasound range, which is inaudible, but the voice asstant accept it as a voice input. The only way to mitigate this is by making an Inaudible Voice Command Cancellation, a module can be enhanced prior to lowpass filters (LPF) to detect the DolphinAttack and cancel the demodulated voice commands.

Another study that focuses on the microphone input is the

voice squatting attack. It is an attack where the voice assistant misheard the command and the misheard command is processed to be a malicious attack. To counter this, this study provide a detector that takes a skill's response and/or the user's input is to be determine whether an impersonation risk is present, and alerts the user once detected.

TABLE I
HSM SECURITY FUNCTION COMPARISON

Study/Desc.	Security Issue	Security features	
This Study	Device tampering	Device authentication on bootup process by using device digital signature check.	HSM's hierarchical key management setup to unlock the device.
Replay Attack Study	Replay attack on voice assistant's microphone.	The use of higher-order spectral analysis (HOSA) features to capture traces of replay attack distortion and detection.	
DolphinAttack Study	Inaudible voice asstant's microphone attack.	By making a microphone enhancement and baseband cancellation.	
Voice Squatting Attack Study	Voice Squatting Attack on voice assistant	Skill Response Checker captures suspicious responses from a malicious skill such as a fake skill recommendation that mimics the service.	Alerts the user once the skill impersonation risk is detected.

From these study, it is crucial for the DVAD to have its security features enhanced at a different point. Which is its data integrity and authenticity when the device is being tampered by an attacker. The study provided in this paper is focused on the DVAD device data integrity and a way to mitigate it by implementing a HSM onto it to further secure the device. The HSM will perform a secure boot and device integrity check to secure the device prior to boot up. It is crucial to have this study where the DVAD vulnerability have a number of securities issues, and the device tampering is one of it.

C. Secure Boot

Secure boot is a security measure that had been developed in the PC industry to ensure that the device only boots using software that is trusted by the Original Equipment Manufacturer (OEM). An authenticated device prior to a boot process has a very crucial role in a secure premise, as a single processor may handle a highly confidential data that should not be edited or erased. A research which highlighted the same importance of a secure boot as the DVAD research had shown that one of the ways to secure a user data is by performing a secure boot method. This research [6] had shown a four-stage boot process to securely boot up a processor state at its initial state and during the shut down process at the processor. The first stage involves the initialization of Trusted Memory-Interface Unit (TMIU) to authenticate the Programmable System-on-Chip (PSoC) by a unique non-volatile device identifier. In any case of the identifier not matching a cryptographic hash compiled into the TMIU bitstream itself, the system will go into a secure lockdown mode. The system will then authenticate the Non-Volatile Memory (NVM). An SD card works as a non-volatile mass storage device. This card identifies itself via its unique 128-bit Card Identification number, which to be identified before the system authenticate the NVM content. As the system authenticate the NVM, the system then authenticate the NVM content by using a secure generation of the secret AES key denoted as to decrypt the boot image and other data stored inside the NVM. After the boot partition has been successfully loaded, the TMIU will hand over full control to the system on the PSoC to set up the remaining operating system (OS) process and starts the user application. From this study, it had shown that the importance of securing a Non-Volatile Memory such as SD Card and flash memory by checking its integrity and authenticity before the system boots up have a crucial role to secure an organization whenever any storage or devices on a processor are being compromised before a boot up.

Another study had shown the high requirement for a secure chip on an embedded system. The security process that is being studied is the bootrom security [7], which is to prevent the loading of fake system images and prevent malicious attacks towards the device. This research design adds AES, SHA, ECC cryptographic algorithm in the chip bootrom and checks the loaded image for security and integrity, to provide personalized key management for chips to prevent potential attacks. This study is highly related to the study of DVAD to further secure the device data by using AES-256 and SHA-256. This had proven that the need of embedded device to be secured is in high demand and various studies had been made to further ensure its security.

D. Hardware Security Module

A HSM is a module that is to be ported or to be implemented onto a control unit or a processing unit to enhance its security features. Its functionality varies across devices such as a personal computer or a vehicle electronic control unit. The application for the HSM onto a device has one common goal, which is to have better security features such as authentication or added encryption onto its data. A study on HSM [8] implementation onto a personal computer has added the device security function to have a clean boot in the case of device tampering or malicious software has infect the device.

Another study on HSM has added a security functionality onto a vehicle to secure its data transmission between the vehicle to the server or to another passing by vehicle by using Physically Unclonable Functions (PUF) [9], which is a type of hash function in which a given input will result in a specific output. This unique method is to secure the device data

communication and data storage. The PUF module uses ECC-256 for asymmetric cryptography, AES-128 for symmetric encryption and decryption, and WHIRLPOOL as its hash function.

In comparison to the DVAD studies, the DVAD uses the HSM to secure the device from being tampered, where its data can be altered and user log data or live activity can be seen if infected. The HSM will provide a pre boot check for any modification made towards the device.

E. Data Security

Data security is the scope of this study, where the DVAD users' data need to be encrypted in the case of the device is being stolen or compromised. The use of Advanced Encryption Standard (key size of 128, 192, or 256 bits) has enabled Internet of Things (IoT) users to apply encryption standard widely used and recognized by most countries as an approved cryptographic algorithm to protect their top secret information in a database. The need to secure data contained inside a memory block is in demand for any institution or organization that is setting up their infrastructure. AES has shown its feasible application for non-volatile memory, especially when it is implemented on mobile devices. This study had shown that the AES in memory is feasible [10] and the encryption process time is quicker and cost minimal energy usage. This relates back to the DVAD research that the hardware security module can use AES to encrypt or decrypt the NVM file contain inside an SD card in between boot and shut down process.

Additionally to this, a hardware based security is proven crucial in the drone industry. This industry had grown drastically for the past few years due to its fast geographical monitoring service. The need for data confidentiality [11] onto a drone is crucial as the device may take to the air anywhere with the risk of being intercepted or lost – the data contained inside it may then be breached. This method had shown that the data encryption in the drone security module is suitable for small load telemetry vehicle in consideration of the small computing power of the security components inside micro controller unit. This study had shown the relation between the drone security module and the DVAD security module as its objective is to secure the users' data when it is lost. The outcome of users' data security enhancement will keep their data safe in the case of the device is being stolen or sent to be repaired, in the event of a service personnel replacing the memory unit without the user authorization.

F. Digital Voice Assistant Device

A digital voice assistant device function [12] makes the life of its users easier. Some of the examples for this device are determining the weather based on the live location, playing the music specifically asked by the user, switch on the lights at the living room, switch off the air conditioning at the bedroom, and set up an important meeting at the office for the user. Apart from aiding the average users in performing their daily tasks, voice assistants [13] also help users with disabilities with their daily tasks as well. All of these data are being collected live at the device itself without the user's awareness, and these small data collection can represent the lifestyle of its users. With this in mind, this poses a security issue for this device as its data is transferred around its ecosystem waiting to be intercepted. Although this device brings an easy lifestyle for the users because it automates some of the users' tasks, but if the device breaks down and is sent to the service center to be fixed, this poses as a potential threat as anyone can tamper the device as a spy gadget or open the system to copy and/ or steal the user data.

G. Security Comparision

A HSM comparison are to be made to observe how different HSM enhance the device security that it is implemented to. There are 2 HSM devices to be compared with the current study HSM implementation towards the DVAD. The comparison are as follows:

Table 2 shows the HSM functions comparison. The PnP HSM implementation onto a personal computer has added the device security function to have a clean boot in the case of device tampering or malicious software has infect the device.

Another study on HSM has added a security functionality onto a vehicle to secure its data transmission between the vehicle to the server or to another passing by vehicle through Physically Unclonable Functions (PUF), which is a type of hash function in which a given input will result in a specific output.

In comparison to the DVAD studies, the DVAD uses the HSM to secure the device from being tampered, where its data can be altered and user log data or live activity can be seen if infected. The HSM will provide a pre boot check for any modification made towards the device.

TABLE II
HSM FUNCTION COMPARISON

	DVAD	**PnP HSM**	**PUF**
Main function	Secure boot device to prevent infection	Clean boot device to boot device normally after being infected	Secure device communication
Security features	Secure device data by SD Card encryption	Secure device data by encryption	Secure device data by encryption
	Device binding process Use of hierarchical key management setup to protect MasterKey for decryption	End-to-end encryption Nothing further discussed	ECC256, AES128 and WHRILPOOL

III. METHODOLOGY

The proposed architecture for the method consists of three key components: Raspberry Pi, Zymkey 4i as Hardware Security Module (HSM), and an SD Card. These components are inter-connected as a device for Digital Voice Assistant. This device is being secured with an extra security layer, with the

HSM.

Figure 2 shows the DVAD set up environment where the key components are Raspberry Pi 4, SD Card and Zymkey 4i with a speaker as an essential component. The reason why the SD Card, Raspberry Pi, and Zymkey 4i is interconnected to each other is because the binding process that is being made towards each other. The Zymkey unique features which is to generate a unique identity (ID) for the host system is based upon a fingerprint that measures specific system components. Once all the devices have gone through the binding process, the alteration of any of these devices will cause the Raspberry Pi as a whole will not boot up properly during boot process.

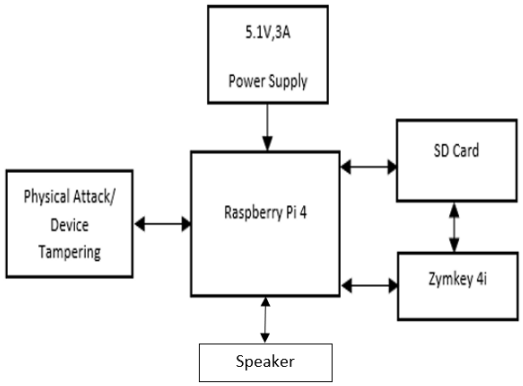

Fig. 2 DVAD set up environment

This study found that when the DVAD is being secured by the HSM, the binded memory unit inside the device is being secured with a digital signature to verify its identity. With this in mind, when the device boots up, the HSM will perform the digital signature check before it decrypts the SD Card data. If the device is not being tampered or altered, the device will boot up normally, otherwise the device will stop after the signature authentication process, and does not proceed with the decryption of files then boot up.

Figure 3 shows the DVAD block diagram. This diagram shows the interaction between the HSM and the Raspberry Pi as a DVAD. The HSM enables multifactor device ID and authentication, data encryption and signing, key storage and generation, and physical tamper detection towards the HDVA. The features such as secure element root of trust, real-time clock, and true random number generator are also available.

In order to properly set up this module for the research, the Raspberry Pi 4 needs an SD card and the HSM to be attached to it. With this setup, the Raspberry Pi 4 will act as a DVAD device that comes along with its security module attached, which is the HSM. The HSM, Raspberry Pi 4, and the SD card will be bound together to successfully generate a unique Digital Identity (ID) for the host system.

The Linux Unified Key Setup (LUKS) is the popular key management setup for dm-crypt, where it only has a single master key. The initial method of encrypting or decrypting the root file system is through the single master key that the dm-crypt has. The main disadvantage of this method is, for every new iteration of the master key during changing authorized users or servers, the master key will need to be re-encrypted everytime the changes occur. This result is inconvenient for the DVAD users, as authentication for its users tends to change for every usage. To make it more user-friendly, a unique security functionality can be provided by the DVAD by using the HSM's hierarchical key management setup, whereby each user or service is given a User Key that can be used to decrypt the MasterKey used for the SD card encryption and decryption. This will provide convenience without the need to re-encrypt the file when there are multiple users using the device.

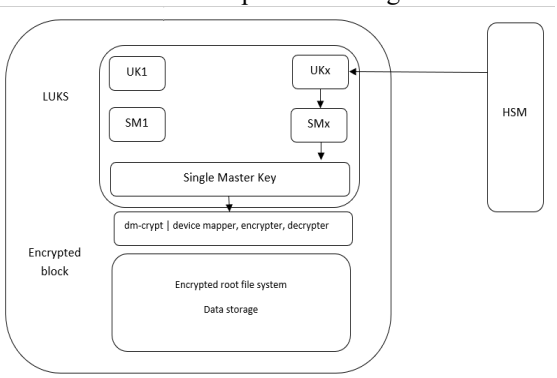

Fig. 3 DVAD System Block Diagram

Figure 4 shows the system boot up process. The system boot-up process is unique in its own way due to the implementation of Zymkey 4i into the DVAD. The process flow shows how the security features work when the Zymkey 4i has been added into the device. This process protects the data or the root file system in the SD card only to the authorised user when boot up. With this in mind, the process involves digital signature verification, which involves the three bound components altogether.

The LUKS is a popular key management setup for dm-crypt, which is responsible for block device encryption by using Linux. While dm-crypt is a transparent disk encryption subsystem in Linux kernel and a part of device mapper infrastructure. Taking this into consideration, the security efficacy of the DVDA with LUKS and encrypted root file system is highly dependent upon how the user keys are generated and where their storage are located.

As observed in Figure 4, the study shows that Zymkey 4i provides a general locking service, whereby a block of plaintext data is encrypted and digitally signed. When LUKS is used, the User key is sent to the Zymkey 4i to be encrypted and signed when the filesystem is created. Later on, when the device boots up or restart, the system needs to decrypt the root file system. To do this, the locked LUKS key signature will be verified first. The content is then decrypted, and afterwards, it is being presented to dm-crypt. As a result, if the key was unlocked successfully, the boot process continues normally. This is the boot process that involves LUKS/dm-crypt filesystem where the key is protected by the Zymkey 4i.

In the event of an attack towards the DVAD, namely data tampering from the SD Card, the data contain within the SD card of the device is securely encrypted and prevent an attacker from stealing or compromising the user's data. As a result, any modification towards the data inside the SD card or an SD card replacement will result in failure to boot up for the DVAD as the device initial digital signature verification fails.

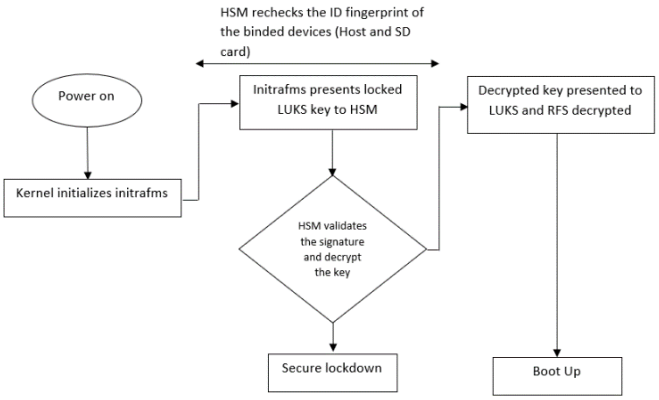

Fig. 4 DVAD Secure Boot Flow

IV. Result

The study had shown that the feasibility and security of the DVAD with HSM are highly stable and secured. The findings for the memory, CPU, and energy consumption on the device are shown below.

Table 3 shows the comparison of energy consumption between each task when the DVAD is running on the Raspberry Pi OS without HSM with Amazon Alexa is activated in listening mode and doing task such as reporting the weather. Then, the energy consumption is being measured again when the HSM is being implemented. In this DVAD feasibility study, the energy consumption when it is generating, encrypting, and decrypting the block is being analyzed too, with the intention of measuring how well can the DVAD cope when the HSM performs its encryption tasks. It is also observed when the HSM is being pulled out of the DVAD, the energy consumption usage is low, as the boot process is not complete.

The DVAD energy consumption is also being compared with the common Google Home Mini, as it is a widely used Smart Home Device by global consumer besides Amazon Alexa. Note that the energy usage by the DVAD is slightly higher, as its OS runs a desktop environment compared to Google Home Mini, which does not have a graphical installation program. The possible reason behind this is because the DVAD is required to run a desktop environment for analysis purpose as the study requires configuration of the device and a direct monitoring of its OS.

TABLE III
ENERGY CONSUMPTION BY DVAD

Task	V	I	W
OS	5V	0.61A	3.05W
DVAD.Idle	5V	0.74	3.7W
DVAD.Listening	5V	0.98A	4.9W
DVAD.Task	5V	1.04A	5.2W
OS.HSM	5V	0.82A	4.1W
HDVA.HSM	5V	0.90A	4.5W
DVAD.HSM.L	5V	1.15A	5.75W
DVAD.HSM.T	5V	1.20A	6.0W
TESTING UTILS	5V	0.90A	4.5W
Generate Block	5V	0.93A	4.65W
Encrypt Block	5V	0.87A	4.35W
Decrypt Block	5V	0.90A	4.5W
HSM pulled out	5V	0.67A	3.35W
GOOGLE.Idle	5V	0.34A	1.7W
GOOGLE.Listening	5V	0.46A	2.3W
GOOGLE.Task	5V	0.48A	2.4W

Table 4 shows the CPU consumption comparison between each task when the DVAD is only running the Raspberry Pi OS, having its Amazon Alexa activated, when it is in listening mode and doing task such as reporting the weather. The CPU consumption is then being measured again when the HSM is being implemented. In this DVAD feasibility study, the CPU consumption when it is generating, encrypting, and decrypting the memory block is being analyzed too, to measure how well can the DVAD cope when the HSM performs its encryption tasks. Note that the CPU measurement becomes null when HSM is pulled out from DVAD. The command line interface can only be accessed if the boot is operates normally in order to measure the CPU consumption.

TABLE IV
CPU CONSUMPTION BY DVAD

Task	CPU 1 (%)	CPU 2 (%)	CPU 3 (%)	CPU 4 (%)
OS	0.0	0.0	0.6	2.6
DVAD	8.3	2.0	0.0	0.6
DVAD.Listening	58.7	40.3	22.4	36.9
DVAD.Task	31.3	25.8	10.6	11.7
OS.HSM	0.6	0.0	0.0	2.6
HDVA.HSM	7.7	3.2	0.6	0.0
DVAD.HSM.L	39.5	34.2	25.7	48.7
DVAD.HSM.T	44.1	52.6	17.9	24.7
TESTING UTILS	1.9	1.3	2.6	2.6
Generate Block	2.5	2.6	0.0	0.0
Encrypt Block	5.7	3.3	0.0	1.3
Decrypt Block	1.9	3.2	0.6	1.9
HSM pulled out	-	-	-	-
GOOGLE.Idle	-	-	-	-
GOOGLE.Listening	-	-	-	-
GOOGLE.Task	-	-	-	-

Table 5 shows a similar study as the previous one, which consists of memory consumption comparison between each task when the DVAD is only running the Raspberry Pi OS, having its Amazon Alexa activated, when it is in listening mode and doing task such as reporting the current weather.

Instead of CPU consumption in Table 2, Table 3 shows the memory consumption when it is generating, encrypting, and decrypting the content inside memory block. The analysis measures how well the DVAD can cope when the HSM performs its encryption tasks. Similar to before, the memory measurement becomes null if HSM is pulled out from DVAD and can be only read in normal operation to measure the memory consumption.

Figure 5 shows the RAM consumption before HSM implementation where the DVAD is in an idle state. As observed the RAM usage only consumes 146.7 MB for the device to run in an idle DVAD state where most of the process comes from the root.

Figure 6 shows the RAM consumption after HSM implementation where the DVAD is in an idle state with the HSM active state. As observed the RAM usage only consumes 208 MB for the device to run in an idle DVAD state where Zymkey process are added, which contributes to the additional 61.3 MB resources being consumed. This minor increase had proven the implementation of HSM into a DVAD is feasible where the RAM usage are not increasing drastically and does

not cause RAM memory shortage when user wants to perform other task with the DVAD. The free remaining memory of 3.53 GB is more than sufficient for the user to perform other heavy task using the DVAD.

TABLE V
MEMORY CONSUMPTION BY DVAD

Task	RAM (3.81G)	RAM (%)	Swap (100.0M)
OS	230M	6%	0M
DVAD	253M	7%	0M
DVAD.Listening	270M	7%	0M
DVAD.Task	275M	7%	0M
OS.HSM	208M	5%	0M
HDVA.HSM	211M	6%	0M
DVAD.HSM.L	222M	6%	0M
DVAD.HSM.T	223M	6%	0M
TESTING UTILS	175M	5%	0M
Generate Block	177M	5%	0M
Encrypt Block	176M	5%	0M
Decrypt Block	177M	5%	0M
HSM pulled out	-	-	-
GOOGLE.Idle	-	-	-
GOOGLE.Listening	-	-	-
GOOGLE.Task	-	-	-

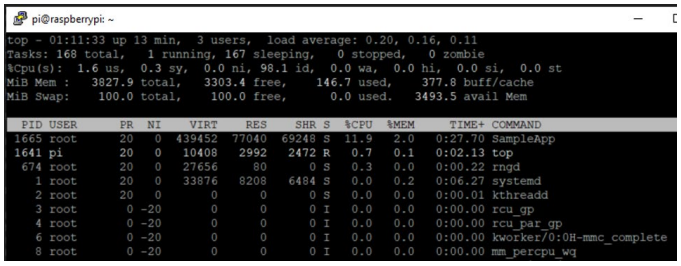

Fig. 5 RAM consumption before HSM implementation

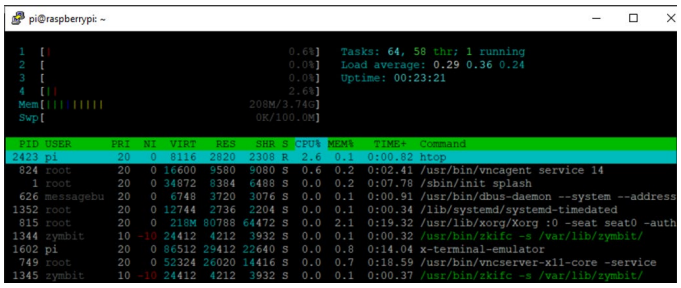

Fig. 6 RAM consumption after HSM implementation

Figure 7 shows the RAM consumption on running task before HSM impelementation onto the DVAD. The task being carried out is the execution of the "joke" command. Where the DVAD quote a random joke from the internet, in reply for the user "Tell me a joke" command. The RAM only consumes 148.8 MB of its resource which is a slight 2.1 MB increase from the previous DVAD idle state.

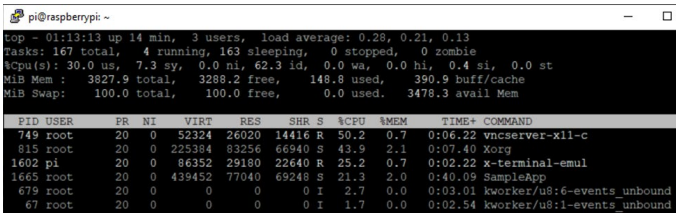

Fig. 7 RAM consumption on running task before HSM implementation

Figure 8 shows the RAM consumption on the same running task after HSM impelementation onto the DVAD. The task being carried out is the execution of the same "joke" command as the previous test. Where the DVAD quote a random joke from the internet, in reply for the user "Tell me a joke" command.

As observed the overall RAM usage only consumes 222 MB for the device to run a task in DVAD state, where Zymkey process are being added, which contributes to the additional 14 MB resources being consumed. This minor increase had proven the implementation of HSM into a DVAD is feasible where the RAM usage are not increasing drastically and does not cause RAM memory shortage when user wants to perform a task with the DVAD. The free remaining memory of 3.52 GB is more than sufficient for the user to perform other heavy task using the DVAD.

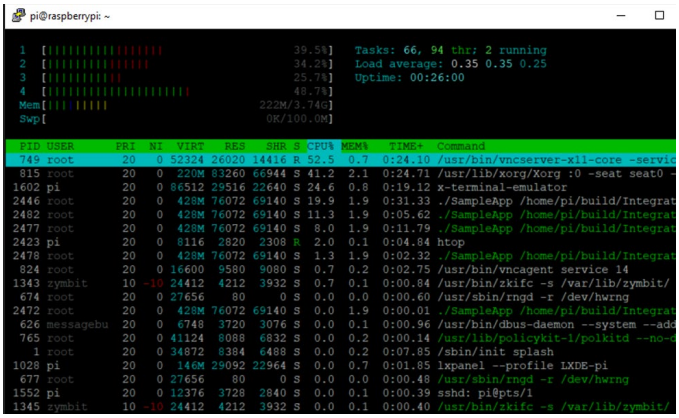

Fig. 8 RAM consumption on running task after HSM implementation

Figure 9 shows the idle state of Alexa at the DVAD at background. Amazon Alexa is built inside the DVAD and associated by using a token to verify and register the device from Amazon Web Services. At idle state, the DVAD microphone will be at standby mode to receive any wake command of "Hey Alexa" for it be at active state and to listen to any commands afterwards from the user.

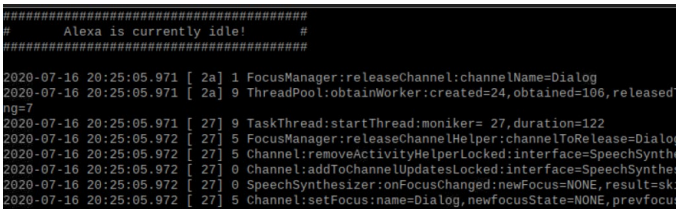

Fig. 9 Alexa idle state at DVAD

Figure 10 shows the background status when Alexa is executing a "joke" task. This is the situation when the user ask the DVAD "Tell me a joke" command. The DVAD then

proceed to search for random joke from its source then proceed to notify the user about the joke. The Alexa active state and its responsiveness to a task had been proven a success while the HSM is attached.

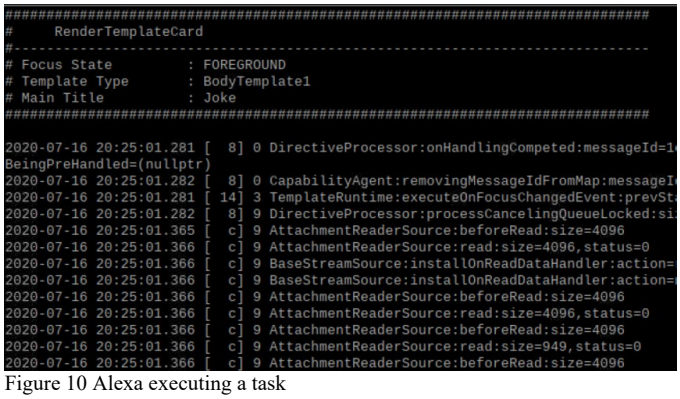

Figure 10 Alexa executing a task

Figure 11 shows the DVAD executing a crypto test after the success of the bind process. This test had shown that the encryption and decryption of a random block process is a success where the HSM is being installed and integrated properly with the DVAD.

Fig. 11 DVAD executing Crypto test

Figure 12 shows the DVAD boot scenario when the device is being tempered. This is the scenario when the SD Card or Zymkey 4i are being altered by an attacker. The DVAD boot process will stop at the boot page where it does not proceed to decrypt the SD card, as the locked LUKS key are not abled to be verified.

When LUKS is used, the User key is sent to the Zymkey 4i to be encrypted and signed when the filesystem is created. Later on, when the device boots up, the system needs to decrypt the root file system. To do this, the locked LUKS key signature will be verified first. Till then the content is then decrypted, and afterwards, it is being presented to dm-crypt. As a result, if the key was unlocked successfully, the boot process continues normally.

Figure 12 DVAD when device is being tempered

V. Conclusion

In conclusion, the study had shown that the DVAD with the implementation of Zymkey 4i as HSM is revealed highly feasible and secure, in terms of its memory and CPU consumption, and its secure boot function. The contribution of this study will add value to the existing literature on security for IoT technology. It is important for hardware-based security to consider the effect of feasibility which helps to improve the performance of such devices. In addition, the DVAD has also shown its feasibility in terms of energy usage and memory consumption when the HSM has been implemented. The security features of SD card encryption and secure boot up by the Zymkey 4i have further secured the DVAD data. This additional feature will further secure this device when it is not being monitored, leaving the users at ease from device tampering.

Acknowledgement

This research study is funded by Universiti Teknologi MARA under the research grant Geran Penyelidikan Lestari no: 600-RMC/MyRA 5/3/LESTARI (098/2020) in collaboration with MIMOS Berhad, Malaysia.

References

[1] N. Zhang, X. Mi, X. Feng, X. Wang, Y. Tian, and F. Qian, "Dangerous skills: Understanding and mitigating security risks of voice-controlled third-party functions on virtual personal assistant systems," *Proceedings - IEEE Symposium on Security and Privacy*, vol. 2019-May, pp. 1381–1396, 2019, doi: 10.1109/SP.2019.00016.

[2] G. Germanos, D. Kavallieros, N. Kolokotronis, and N. Georgiou, "Privacy Issues in Voice Assistant Ecosystems," *Proceedings - 2020 IEEE World Congress on Services, SERVICES 2020*, pp. 205–212, 2020, doi: 10.1109/SERVICES48979.2020.00050.

[3] C. Jackson and A. Orebaugh, "A study of security and privacy issues associated with the Amazon Echo," *International Journal of Internet of Things and Cyber-Assurance*, vol. 1, no. 1, p. 91, 2018, doi: 10.1504/ijitca.2018.090172.

[4] C. Yan, G. Zhang, X. Ji, T. Zhang, T. Zhang, and W. Xu, "The Feasibility of Injecting Inaudible Voice Commands to Voice Assistants," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 3, pp. 1108–1124, May 2021, doi: 10.1109/TDSC.2019.2906165.

[5] K. M. Malik, H. Malik, and R. Baumann, "Towards Vulnerability Analysis of Voice-Driven Interfaces and Countermeasures for Replay Attacks," *Proceedings - 2nd International Conference on Multimedia Information Processing and Retrieval, MIPR 2019*, pp. 523–528, 2019, doi: 10.1109/MIPR.2019.00106.

[6] F. J. Streit *et al.*, "Secure Boot from Non-Volatile Memory for Programmable SoC Architectures," *Proceedings of the IEEE International Symposium on Hardware Oriented Security and Trust, HOST 2020*, pp. 102–110, 2020, doi: 10.1109/HOST45689.2020.9300126.

[7] N. Bin *et al.*, "Research and design of Bootrom supporting secure boot mode," *Proceedings - 2020 International Symposium on Computer Engineering and Intelligent Communications, ISCEIC 2020*, pp. 5–8, 2020, doi: 10.1109/ISCEIC51027.2020.00009.

[8] A. Asaduzzaman, M. F. Mridh, and M. N. Uddin, "An inexpensive plug-and-play hardware security module to restore systems from malware attacks," 2013. doi: 10.1109/ICIEV.2013.6572565.

[9] C. Labrado and H. Thapliyal, "Hardware Security Primitives for Vehicles," *IEEE Consumer Electronics Magazine*, vol. 8, no. 6, pp. 99–103, Nov. 2019, doi: 10.1109/MCE.2019.2941392.

[10] M. Xie, Y. Wu, Z. Jia, and J. Hu, "In-memory AES Implementation for Emerging Non-Volatile Main Memory," in *Proceedings of IEEE Computer Society Annual Symposium on VLSI, ISVLSI*, Jul. 2019, vol. 2019-July, p. 103. doi: 10.1109/ISVLSI.2019.00027.

[11] K. Kim and Y. Kang, "Drone security module for UAV data encryption," in *International Conference on ICT Convergence*, Oct. 2020, vol. 2020-October, pp. 1672–1674. doi: 10.1109/ICTC49870.2020.9289387.

[12] H. Mauny, D. Panchal, M. Bhavsar, and N. Shah, "A prototype of smart virtual assistant integrated with automation," *Proceedings of the 3rd International Conference on Inventive Research in Computing Applications, ICIRCA 2021*, pp. 952–957, 2021, doi: 10.1109/ICIRCA51532.2021.9544101.

[13] RVS Technical Campus, IEEE Aerospace and Electronic Systems Society, and Institute of Electrical and Electronics Engineers, *Raspberry Pi based voice-operated personal assistant (Neobot)*. Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology [ICECA 2019], 2019.

Rizzo Mungka Anak Rechie is a technology enthusiast that wanted to make a change in the world by using technology. He enjoys exploring and learning innovative ideas from all around the world and implement it to the local government and private sector of his country for the better of it. He represents the Universiti Teknologi MARA Shah Alam, Malaysia for this paper. He is also a member of World Merit Malaysia which runs innovative projects with Sustainable Development Goals.

PM. Ir. Dr. Yusnani Mohd Yussoff is a senior lecturer in the Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam Malaysia. She has 20 years working experience as a lecturer and researcher. Her research area focusses on Wireless Sensor Network, Trusted Authentication, Embedded Security and Internet of Things. She has graduated few PhD and Master students and currently involve few researchs related to the area. She is currently the head of Information, Security and Trusted Infrastructure Laboratory or InSTIL reseach group. She has authored and co-authored 41 indexed publications with SCOPUS h-index of 7. She is currently a member of IEEE, IAENG and Board of Engineer Malaysia.

Dr. Lucyantie Mazalan is a senior lecturer at the Faculty of Electrical Engineering, Universiti Teknologi MARA since year 2011. She received her PhD from The University of Sheffield. Her research interest expands to image processing, artificial intelligence, trusted computing and chemoinformatics.

Suhairi Mohd Jawi is a Specialist at Cryptography Department, Cybersecurity Malaysia. Currently, his specific responsibilities are in implementation of cryptography-related technology and cryptographic module evaluation according to well-known standards. Suhairi has hands-on experiences on programming and databases under UNIX and Windows environments. Formerly, he was in charge in handling computer incidents from Malaysian Internet users and security auditing on web application vulnerabilities.