CYBER THREAT DETECTION, SECURING AND STORING CONFIDENTIAL FILES IN BYOD

Madushanth R.K.U
IT19053042
Bachelor of Science (Hons.) in Information Technology Specializing in Cyber Security
Department of Information System Engineering
Sri Lanka Institute of Information Technology
Sri Lanka
IT19053042@my.sliit.lk

Anjana D.M.I.S
IT19027166
Bachelor of Science (Hons.) in Information Technology Specializing in Cyber Security
Department of Information System Engineeting
Sri Lanka Institute of Information Technology
Sri Lanka
IT19027166@my.sliit.lk

B.M.U.Sandaken
IT19034546
Bachelor of Science (Hons.) in Information Technology Specializing in Cyber Security
Department of Information System Engineeting
Sri Lanka Institute of Information Technology
Sri Lanka
IT19034546@my.sliit.lk

R.P.A.Tharuka
IT19034614
Bachelor of Science (Hons.) in Information Technology Specializing in Cyber Security
Department of Information System Engineering
Sri Lanka Institute of Information Technology
Sri Lanka
IT19034614@my.sliit.lk

***Abstract*—** As the demand for Bring Your Own Device (BYOD) service enablement grows, it becomes an important element of the business. Because of the increased business agility, essential infrastructure is becoming more important. Employee productivity and happiness as a metric different way of working surroundings but harmful access to vital infrastructure from an unauthorized source. The use of BYOD has increased the number of cyber-attacks. As a result, it has become a major cybersecurity issue for most businesses.

As a result, the corporate ecosystem fragments. There are numerous options. In recent years, new technologies and techniques for reducing pollution have been created. This section has been subjected to cyber attacks, which has resulted in the resolution of a number of concerns. However, due to new assault strategies and tools, work has resumed.is a condition that must be met on a regular basis. The majority of the company issuing a certificate-based authentication system that is secure to enable a bring-your-own-device (BYOD) environment However, the certificate-based system There are a slew of issues with the authentication system that need to be addressed. A security flaw that allows unauthorized access.

Keywords—BYOD, Cyber Attack, Security

I. INTRODUCTION

One of the most widely used technologies in the information technology industry today is the development of mobile computing, which later gave rise to Bring Your Own Device (BYOD), which permits employees to use their privately owned information technology resources, such as computer hardware devices or software, in corporate-related projects. Employees link their smart equipment to the company network in order to perform daily business operations and access corporate data. Using any endpoint devices and smartphones in a way that is independent of time and location enables employees to connect more during work-related tasks. Businesses have benefited significantly from and found convenience in the BYOD process.

Apart from using IT-issued devices, employees are substantially happier when using the mobile devices, they already like. Employing BYOD procedures helps businesses lower their operating and device costs. Businesses providing services can save a lot of money because employees pay the majority of the cost of mobile devices. Another fantastic benefit offered by the BYOD process is an increase in participation both at work and outside of normal business hours.

Employee response to business functions increases as they have access to anything from anywhere in a collaborative setting. Additionally, it significantly boosts employment productivity. Apart from these benefits, using BYOD in an enterprise raises serious security issues that need to be taken into account. According to Drury, BYOD is utilized by around 80% of businesses in nations like Spain, Brazil, Malaysia, and Singapore. In reality, all of the advantages of BYOD may be lost without a proper security solution. Due to their vulnerability to attacks, these mobile devices need to be equipped with suitable security measures.

Information that is processed, recorded, or stored by the organization and is in its control is a very important asset. The two key security concerns when deploying BYOD in an organization are confidentiality and integrity. Contrary to IT devices provided by the company, BYOD devices can have a variety of programs loaded, which is a genuine risk to organizational data. When BYOD is used in an organization, this research will help address security concerns that IT personnel may have and new tactics.

II. Objectives

A. *Network Cyber Threat Detection in BYOD*

The word Bring Your Own Device has started being used since early 1990's but only vastly being used since 2011. Many companies use Bring Your Own Device (BYOD) namely laptops, smartphones, tablets and Personal digital assistant to expand their computing resources, particularly in terms of hardware. It's also thought that allowing employees to bring their own devices boosts productivity. Whether BYOD not implemented in a company, mobile devices are extensively used in the workplace [1]. In 2022, there are approximately 15 billion devices available and set to increase 18 billion by 2025 [2], implying that every individual on the planet will have multiple devices, and 69 percent of employees, even in non-BYOD organizations, use personal mobile devices to access to workplace networks. As the company allows to employee to use the BYOD into the enterprise network, BYOD security guidelines and endpoint security solutions becomes more crucial. Many companies do not sufficiently protect their network from the cyber threats that employee's owned device invites [3]. It is important implementing threat detecting system at the endpoint of the networks.

According to previous research works we found many limitations in securing BYOD techniques in organization. Most of the outcomes of the BYOD device did not provide proper solution to secure the BYOD techniques. Still many security flaws are to be addressed to protect the BYOD and organization data. Many security methods are secure the important data using existing techniques which are not suits with the increase of BYOD usage. In the literature, there are various descriptions of network attack detection systems involving various intelligent-based techniques including machine learning (ML) and deep learning (DL) models, can be found in the literature. However, while some strategies have shown to be effective in specific areas, no technique has yet proven to be effective in preventing all types of network attacks. This is due to the fact that some intelligent-based techniques lack critical features that make them trustworthy systems capable of dealing with various forms of network attacks. This was the driving force behind this study, which looked at current intelligent-based research directions in order to close a gap in the area. With the effect of Covid Pandemic many Critical infrastructure sectors do allow employee to use their personal device without implement proper security measures. Although they have implemented some existing BYOD protecting mechanisms which are introduced years ago and now they are fail to overcome against some advanced cyber-attacks. Also now employees are using some devices with advance which do not support for the old BYOD security protocols. Mobile Device Management is used to monitor, control and secure corporate or personally used devices that used multiple operating system. MDM has preferred BYOD solution that employees needs to allow installation of security software's on their personal devices. When the devices aren't actually owned by the corporation, things get a little more complicated. MDM software fails to detect the presence of malware on laptops and tablets. Blocking and prevention techniques are highly reliant on advanced attacks, and their effectiveness is dwindling.

Existing blocking and prevention capabilities are insufficient to protect against motivated, advanced attackers, but most businesses continue to place an excessive amount of emphasis on protection-only measures. When computers are company-owned, information security lacks the continuous visibility it requires to detect advanced assaults, and BYOD exacerbates the problem. What's needed is a system that can detect threats that elude and circumvent the limited protection perimeter-based defensive technologies can provide against highly motivated, sophisticated attackers.

Many available tools are not real time threat detectors. Real-time threat detection capabilities that can sit deep inside the network to detect when an attack has evaded perimeter defenses while it is happening. Our research project gives solution for implementing new real time threat detection tool. It also consists of threat monitoring, alert system and threat prevention system. Can be used in both cloud based and on premise.

Cyber threats are the major issues faced by BYOD techniques in corporate environment. In Future usage of BYOD will increase rapidly with the current trend. All the organization must be in a position to allow BYOD techniques to their employees. So they have to be prepared with proper BYDO security measures. Main purpose of the BYOD security is to prevent and detect cyber-attacks. This research will be give the solution for the above requirement. Many techniques have been offered to cope with such attack threads in cybersecurity. One of the most effective options is attack detection, which provides a comprehensive and dynamic security framework for monitoring, preventing, and resisting attacks. Attack detection, in particular, would collect data by monitoring the network, system state, behavior, and usage of the system, allowing it to detect unauthorized system usage and external attacker attacks on the system automatically.

Previously, the majority of network attack detection systems relied on a set of pre-defined signature-based assaults. This was a significant setback because the database of assaults needed to be updated on a regular basis as attackers discovered new ways to exploit network security. The predicted accuracy of recognizing and categorizing network assaults has substantially increased with the growth of intelligent-based approaches such as ML and DL. As a result, applying intelligent-based strategies to network security is a burgeoning subject of study that must be pursued.

The main objective of the proposed solution is to implement a network threat detection tool. It can be integrated to all Bring your own devices. Having the ability to be customized while also preserving the confidentiality, integrity and availability of the system. The tool will identify the unauthorized network traffic which enter to the network of the organization. Solution will help to detect unauthorized usage of the system users and from the external attacks.

The concept of examining the entirety of a security ecosystem to discover any malicious behavior that could compromise the network is known as threat detection. If a danger is discovered, mitigating measures must be taken to effectively neutralize the threat before it can exploit any existing threats. When it comes to identifying and resolving risks, time is precious. In order for attackers to have enough time to search around in sensitive data, security solutions must be possible to perceive threats quickly and efficiently. An organization defensive programs should be able to stop the most of threats. The tool designed in a manner to detect threats across your devices and services using automated procedures, reducing the requirement for manual detection. Once the device is connected to the network of the organization by VPN then the all the networks must pass the detection tool to analyze and detect the network flow. All the report will be saving in the log.

B. Automated Confidential files detection, classification and analysis

File detection is one of the main parts of an Operating system. Due to the higher expansions of files and data usage over the years keeping track of selected data has been critical to the organizations while handling big data in their systems. When it comes to handling data, critical information such as restricted, internal only data is considered as critical to the organizations because of their sensitivity of the data and the threat these data might present to the organization in case of a leakage. This is where the problem of how to identify critical data exposed. Most organizations deal with BYOD devices and these devices came to be attackers' favorite targets because of the value of the data it carries and the most exploitable and weakest security link in any system it exposes; the Human Error. To avoid this issue an organization must be protecting and keeping tracked all the critical files in BYOD devices to guarantee its safety. To make sure the identification of critical/confidential files is a dire requirement and collecting these files into one single environment using a detection method, classification and analysis is required.

Previous research has shown that the gap of file classification exists when it comes to critical file classification where a paramount gap exists just because of classification of file haven't covered the areas except discovering information such as Personally Identifiable Information (PII [1]). Besides finding critical files in a system manually might take resources including time, energy, etc. which can be used to other important tasks. Since automation and file classification exists in the information technology field already, this will provide the opportunities to attach these systems and technologies to provide viable solution to critical files detection issue. There are other methods available for identifying data such as pattern recognition [2] which was used for various purposes but identifying critical data is unlikely used by this technology.

Before identifying the critical files, the system needs to presort previously classified documents and new documents. To achieve this task the system needs to be implemented with a tokenization method and database to store tokenized data. While presorting, system tokenized the file data so it cannot be used to identify important information, then compares the tokenized values to previously processed values exist in the tokenized database. By using this this method time, energy management can be increased since there is no need to process previously processed files which identified as critical/not critical by the machine learning critical file detection system which is a relatively higher resource consuming task than tokenization. Python natural language toolkit (NLTK) is ideal for tokenization and after tokenization values will be stored into the connected database for future use of identification. In tokenization algorithm file operations, data insertion, database updating is happened.

In this research paper researchers express a method to identify critical files exist as a text files method including .txt, docx and pdf as per popular examples. Supervised and unsupervised Machine learning methods will be used in text categorizing in the text exist in selected documents automatically based on its content in this case, criticality. Since accuracy, precision and sensitivity of these documents are essential for the matter categorizing documents will use a trust/confident score to classify document further into three categories including High, Medium, Low according to threat score it might present to the organization in a security incident in case detected by the machine learning based critical files detection system.

While execution and file analysis process system identify files without having to be sorted previously or having the requirement to exist all relevant data in one page. After analysis the system sends the documents identified as critical or confidential to relevant category based on the trust score the document content has generated. Google collabs has been a used for machine learning model training and data set for the proposed solution has been identified as information security field sensitive data which will be required to gather cooperate information including confidential documents, Security Operations Center reports, templates, rules, alerts and other information security related cooperate data and data structures which can be vulnerable for attacks in user devices. Dataset will be used as a dictionary to create inter-sentences relationship between words which will be later assigned with a threat score of the criticality.

Total percentage of the threat alerts will be used for categorizing the documents into its respective categories. Using Python as the programming language is important because it is flexible and it highly support Machine learning. After the critical files detection, report generates using the metadata files stores and system detects. This information will be collected from file and presented as a human readable document using pandas library and matplotlib library in Python. This report is significantly important for the users who try to determine the importance of the file and proceed with further proceedings such as uploading it into the secure storage.

Encryption is required to preserve the confidentiality of information in the detected documents and until upload to secure storage or user decides otherwise files need to be kept unreadable in case of a threat actor exploits the device or a human error to grab the files required and try to read it. Encryption function has been implemented to address this specific issue of sensitive files exist in an unprotected directory waiting to be exploited.

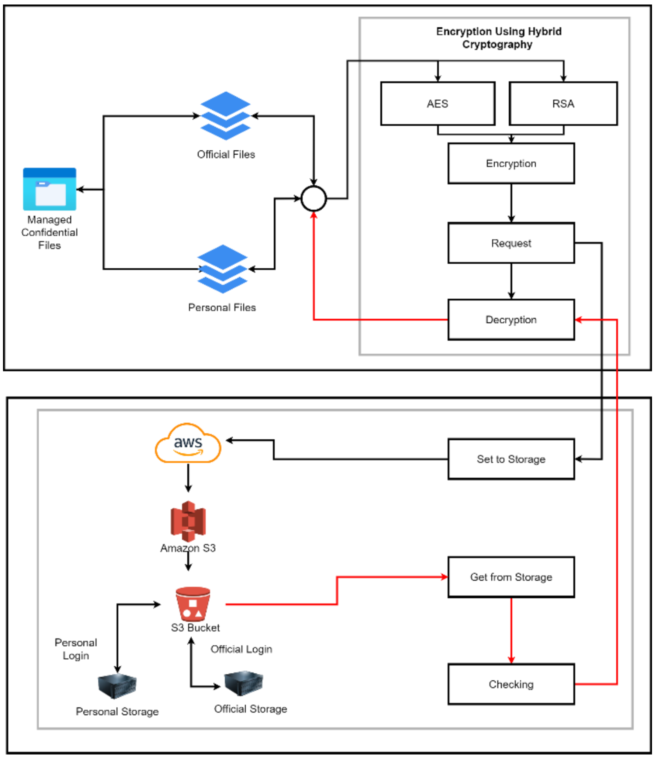
Dependencies of the product can be listed as follows. Pandas library for creating data structures, matplotlib for create pdf files, NLTK for language processing, sqlite3 for database structure development and Pypdf for pdf encryption and metadata extraction.

As per system requirements of the system fully functional BYOD device to test and deploy as the environment, internet accessibility, storage to deploy the system and database purposes. Functional requirements include platform independence because different users tend to use different types of devices to work and system needs to compliant with the existing laws and regulations of the required organization, state, country, etc. system should have an authentication method so unauthorized users cannot

C. Create Safe Storage using c cloud computing approach for confidential files used in BYOD

A secure storage has become an essential necessity for organizations which deal with sensitive information. Creating a safe storage using cloud computing approach is because not everyone has access to the cloud storage and the files stored on the cloud servers are encrypted which enables high security for the sensitive information that is being stored in the storages. However, the main objective is to secure the confidential files in a cost-effective manner, using only a limited number of resources and providing the users maximum security for the files and applications they are storing in their systems. Currently due to the changes in the working enviroment due various reasons like the adaptation to the advancements in technology companies have assigned people to work from home which have increased the use of cloud computing drastically. Thus, using the system we are allowing employees to have the capability to access their files, confidential or personal securely from their own devices from anywhere they are working from. This has brought great benefits to the companies as well as their employees, because working has become more convenient and effective. In order, to ensure this security we have developed a cloud computing storage for the use of companies.

The below figure shows the system architecture of the storing process of the cloud computing storage, which comprises four main processors that are jointing working to secure the confidential files being stored. The first process is the process where the storage system receives the personal and confidential files after they have been managed by the managing confidential file system. After the files are received, they will be sent through the Hybrid Cryptography encryption method. As for the two types of encryption methods, the AES algorithm and RSA algorithm are used double encrypt the files, which makes them much more secure than other files. The files will be encrypted using the keys generated. After the encryption process is finished the files will be sent to the storage. The cloud storage used in this project is the Amazon S3 Bucket, which is run using a server called the AWS. There are two types of authentication channels created for the purpose of storing personal and official files and different authentication methods are used for both personal and official storage. In order to decrypt the stored files, there are two separate approaches being used for the personal and official files. For the personal files a personal code will be sent to the mobile device according to the user credentials given at the beginning and for the official files the decryption code will be received an email. This decryption happens using two keys; the private key and the AES Symmetric Key. After this process only can the files be downloaded by using those authentications and give the private and public keys for decryption only.



To achieve this objective, there are a few steps that are needed to be taken to achieve the ultimate objective. They are explained below:

1) USING A HYBRID CRYPTOGRAPHY METHOD

Hybrid cryptography allows the use of two or more encryption methods. In comparison to other methods, a hybrid approach enables encryption to take the shortest possible time and has the highest throughput for both encryption and decryption. Furthermore, adopting the hybrid cryptography strategy makes a remote server more secure after it is deployed to the cloud environment. Through the use of cryptography, readable material can be encrypted into a form that cannot be read without first decrypting it. AES and RSA are the two types of encryption algorithms that will be utilized in this project since we need to secure the secret data in an economical manner.

2) ENCRYPTION METHOD TO SECURE FILES

Since the contents are already encrypted using hybrid cryptography before being delivered to the cloud storage, the confidential files are already well-protected. However, it can be quite helpful to make it easier to communicate with these devices and to further improve the security of these files saved in the cloud storage. For instance, keeping data in cloud storage has extra advantages for businesses since it offers security from hackers and data loss. Furthermore, cloud storages make it simple for users to exchange and retrieve their files stored there more securely. They are also incredibly affordable and safe. For instance, whereas mass file uploads on other storage devices take a long time, they can be done rapidly with cloud storage by using a link. As a result, it is more useful for the users.

3) CREATING TWO AUTHENTICATION CHANNELS

Users may simply access, store, edit, and remove data that are kept on the cloud. Therefore, during a time like this, they may unintentionally destroy a crucial sensitive file, which could cause the companies to suffer severe losses. Therefore, by managing them independently and more effectively, the staff may provide stronger protection for the secret information by having two distinct storages for personal files and confidential files. Employees may manage files more efficiently and organize them appropriately by having separate storage. They can more easily access and save files thanks to this.

4) USER AUTHENTICATION

Users can access their saved data more quickly by segregating their personal and confidential files in storage. However, these staff members have the potential to unintentionally delete the files, and even if users keep the file open, another staff member has the power to view it, which could lead to a data breach. As a result, each of these folders is separately locked, offering great security. Due to the need that employees check in using their work email, confidential files are given top priority during this process as well. Because each of these folders is separately locked, excellent security is provided. Since employees are required to check in using their work email, confidential data are also given top priority during this process.

5) PERSONAL FILES

The user has to access the personal login in order to store personal files in the cloud storage, After the successful completion of the login process the user can store files, after the personal files are separately stored to the personal file folder they will get authenticated and when the file or files are getting stored to the cloud storage a message will be sent to the user’s mobile device with the decryption code. The user’s mobile number will be obtained that the process of user authentication where user details are being collected. This code is the code that must be used by the user in order to decrypt the stored file. This ensures that on one else will have the privileged to decrypt our stored personal files.

a) OFFICIAL FILES

The user has to login to their official login using the official user credentials through the official login the user has the privilege to use the official email of the user to upload a file or files to the cloud storage. In the official file storing process the decryption codes of the files are sent to the user’s official email addresses. Unlike for personal files for official files there are two decrypt keys are being sent. They are the private key and the AES Symmetric Key. Then the user can use these keys to decrypt and obtain the files securely.

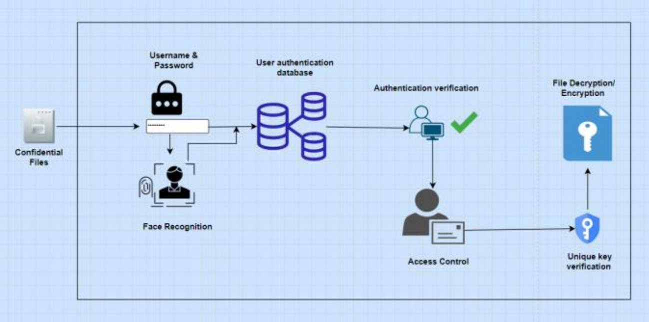
D. **Managing confidential files in BYOD**

The majority of existing software for managing BYOD devices, such as Microsoft Intune, a cloud-based service that focuses on mobile device management (MDM) and mobile application management (MAM), employs encryption to secure personal information, which is a questionable strategy. If the secret data owner is a malevolent insider, encrypted information can be sent to unauthorized outsiders or unauthorized insiders such as low-level staff. This is a reality that none of the software developers have taken into account.

Biometrics, the use of fingerprints, face recognition, voice recognition, or retina scans to access information are far more trustworthy than safe password methods. These are significantly more difficult to steal from a database, and they also save time by eliminating the need to memorize passwords or waste time trying to retrieve them. Continued biometric authentication and/or multi-factor authentication (MFA), can be used to offer an extra layer of security. It should be noted that biometrics can be used to prevent unauthorized viewing of a secret document.

In this proposed system the 'Biometric check' is used to identify whether the accessing user is not a malicious person because if the device got stolen any malicious person can access the files. If the entered biometric data is not in the database, the system consider it as a malicious and blocks the activity. Categorizing access control for employee titles, enabling real time protection, even though the device gets stolen or any malicious user accesses there is no chance of leaking confidential documents because of the biometric check. This biometric check helps to identify the authorized devices and confidential files will only be readable in identified devices. In my system I decided to apply face recognition. Biometric data is becoming increasingly common in the workplace, and it appears to be a "quick fix" modern security or fraud prevention solution. The greater usage of biometric data for personal purposes appears to have enhanced the perceived acceptance of biometric data use.

An authorized device can access to the confidential files and do necessary activities according to their status. Using this system, it will be categorized permissions of the authorized employees in to four tiers. If the device is unauthorized (i.e., not assess the developed biometric check), user can't access to the relevant files, as the system itself will automatically block the activity. When the system detects unusual activity, it switches to mitigation mode, which disables the questionable activity until its behavior is properly studied. Users' suspicious behavior also prompts action since it might suggest an incursion or an insider threat. While the system investigates, those concerning user accounts are locked. Even if the gadget is lost or stolen, this method is simple.



This system helps users save files in encrypted format at their local storage. The authentication is 2-step with an access control.

i. Authentication by username and password from the SQLite database.
ii. Authentication using Face Recognition
iii. Access Control using four tier architecture

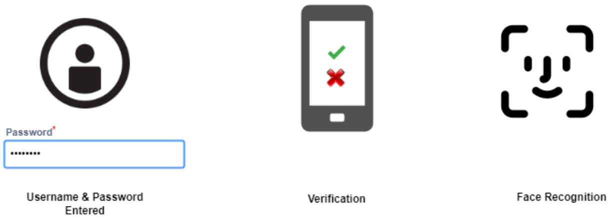


Figure 1

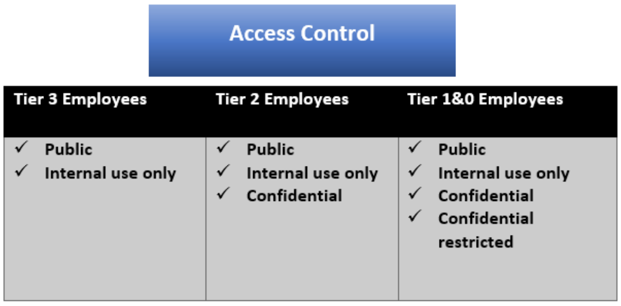
a) 2- step authentication method

First is authentication by username and password from the SQLite database. The password stored in database is hashed using MD5 Hash. Second is Face Recognition. If both the authentication is successful, then the user is taken to operations page where encryption and decryption can be performed. The encryption-decryption is performed using the user's unique key which makes it impossible to decrypt the encrypted files by any other means as the user's unique key is not revealed to even the user. Once the user logs out or closes the application all the files decrypted during that session are deleted for added security.

b) Access Control

The biometric verification will be conducted for confidential data when they are transferred to BYOD. An approved employee's biometric data will be loaded into the database tool installed on their device, granting them access to such private information based on their status. When the biometric data identifies the user, an alert message confirming the user's validity is presented automatically. When a user attempts to open a file without the right biometric data, the secret file instantly prevents the action without even allowing the user to authenticate into the device.

The biometric check and an access control is designed to protect confidential files from harmful activities that may be injected into them, and it is installed on the employee's BYOD devices, which are then used to access the private files, thereby eliminating the threat. When the unique biometric authentication is entered during the process, the system determines the validity of the user and grants access to the necessary categories of files based on the employee's status, which has been classified into four tiers. Finally, employing this instrument, the following services may be achieved, which will be quite beneficial in protecting private files in the firm.



Access Control

Tier 3 Employees	Tier 2 Employees	Tier 1&0 Employees
✓ Public ✓ Internal use only	✓ Public ✓ Internal use only ✓ Confidential	✓ Public ✓ Internal use only ✓ Confidential ✓ Confidential restricted

Employees are divided into four tiers in this system, with the 0th tier being the highest and the 3rd tier being the lowest, based on their standing. Because each employee has a unique identification number that is stored in the database, his status may be accessed from the database via this method. When an authorized employee opens the merged confidential file, the system asks for his employee id. As a result, when the identification number is entered, the system will determine his status automatically. Employees will be categorized into four tiers,

1. The following Tier 3 staff have access to the files: -
   a) Open to the public.
   b) restricted to internal usage only

2. The following Tier 2 staff have access to the files: -
   a) The general public
   b) Only for internal usage
   c) Private and confidential

3. Tier 1 and Tier 0 staff have access to the files: -
   a) The general public
   b) Only for internal usage
   c) Private and confidential

d) Restricted confidentiality (Secret)

An authorized device can access to the confidential files and do necessary activities according to their status. Using this system, it will be categorized permissions of the authorized employees in to four tiers. If the device is unauthorized (i.e., not assess the developed biometric check), user can't access to the relevant files, as the system itself will automatically block the activity.

ACKNOWLEDGMENT

This paper is done as a requirement for bachelor's degree in Information Technology Specialization in Cyber Security provided by Sri Lanka Institute of Information Technology. We would like to express our gratitude to Ms. Dulani Lunugalage and Ms.Pipuni Wijesiri, lecturer at Sri Lanka Institute of Information Technology for mentoring this research project.

REFERENCES

[1] F. Jamal, M. Taufik, A. A. Abdullah, and Z. M. Hanapi, "A systematic review of Bring Your Own Device (BYOD) authentication technique," J. Phys. Conf. Ser., vol. 1529, no. 4, p. 042071, 2020.

[2] J. Cruz, "74 BYOD statistics 2021: Forecasts, benefits and market share," THE iNFLUENCER FORUM, 21-Dec-2021. [Online]. Available: https://theinfluencerforum.com/74-byod-statistics-2021-forecasts-benefits-and-market-share/. [Accessed: 10-Feb-2022]

[3] "BYOD endpoint security," Mcafee.com. [Online]. Available: https://www.mcafee.com/enterprise/en-us/security-awareness/endpoint/byod-endpoint-security.html. [Accessed: 10-Feb-2022]

[4] Cornell University, "Confidential Data Scanning," 19 January 2022. [Online]. Available: https://it.cornell.edu/cit-intranet/confidential-data-scanning.

[5] D. Maiorca, G. Giacinto and I. Corona, "A Pattern Recognition System for Malicious PDF Files Detection," MLDM, pp. 510-524, 2012.