# CERTIFICATE

# COMPLIANCE CERTIFICATE

# PAPER PUBLICATION CERTIFICATE

# THESIS APPROVAL CERTIFICATE

# STATEMENT OF ORIGINALITY

# Acknowledgements

Table of Contents

[CERTIFICATE ii](#_Toc74976446)

[COMPLIANCE CERTIFICATE iii](#_Toc74976447)

[PAPER PUBLICATION CERTIFICATE iv](#_Toc74976448)

[THESIS APPROVAL CERTIFICATE v](#_Toc74976449)

[STATEMENT OF ORIGINALITY vi](#_Toc74976450)

[Acknowledgements vii](#_Toc74976451)

[CHAPTER 1: PROPOSAL 1](#_Toc74976452)

[1.1. Introduction 1](#_Toc74976453)

[1.2. Background 1](#_Toc74976454)

[1.2.1. Overview 1](#_Toc74976455)

[1.2.2. FIM Global Market Share 2](#_Toc74976456)

[Global FIM Market: Segmentation Analysis 2](#_Toc74976457)

[1.3. Problem Statement 3](#_Toc74976458)

[1.4. Aims 4](#_Toc74976459)

[1.5. Objectives 4](#_Toc74976460)

[1.6. Signiﬁcance of the Project 4](#_Toc74976461)

[1.7. Methodology 4](#_Toc74976462)

[i. Authentication of Users 5](#_Toc74976463)

[ii. Identification of files to be monitored 5](#_Toc74976464)

[iii. Creation of Baseline values for files to be monitored 5](#_Toc74976465)

[iv. Monitoring changes 5](#_Toc74976466)

[v. Sending an alert 5](#_Toc74976467)

[vi. Reporting results 6](#_Toc74976468)

[1.8. Scope of the Project 6](#_Toc74976469)

[i. Host Machine Environments 6](#_Toc74976470)

[ii. Files and file types 6](#_Toc74976471)

[1.9. Deﬁnition of Key Variables 6](#_Toc74976472)

[i. Files 6](#_Toc74976473)

[ii. Environment 6](#_Toc74976474)

[iii. FIM 7](#_Toc74976475)

[iv. Baseline Value 7](#_Toc74976476)

[v. Compromised File 7](#_Toc74976477)

[1.10. Conclusion 8](#_Toc74976478)

[CHAPTER 2: LITERATURE REVIEW 9](#_Toc74976479)

[2.1. Introduction 9](#_Toc74976480)

[REFERENCES 10](#_Toc74976481)

List of Figures

List of Tables

Abbreviations

Abstract

# CHAPTER 1: PROPOSAL

## Introduction

File integrity monitoring (FIM) is simply the technique of keeping watch of the data, who accesses and makes changes to it as well as how it has been changed. Integrity is one of the pillars of the CIA triad of information security; hence the scope of this project is cyber security. To better understand this concept, the phrase can be broken down into individual words – file, integrity and monitor. The term *file* refers to data or information that is actually stored on the computer or network. The *integrity* of the file means that it has not been maliciously altered or modified. *Monitoring* refers to constantly checking the file to see whether there have been any modifications done to it. This check will verify whether the integrity of the data is maintained or not [1].

## Background

### Overview

In the business world, the most important assets to any firm are the records of its clients. Hackers are tirelessly working towards illegally gaining access to and modifying these critical assets. Cyber security personnel are therefore devising counter-measures to detect these intrusions and contain the impact of any malicious modification. FIM is an internal control or process that performs the act of validating the integrity of operating system and application software files using a verification method between the current file state and a known, good baseline. This comparison method often involves calculating a known cryptographic checksum of the file's original baseline and comparing with the calculated checksum of the current state of the file. Other file attributes such as timestamps can also be used to monitor integrity. [2] says the general act of performing file integrity monitoring is automated using internal controls such as an application or process and that such monitoring can be performed randomly, at a defined polling interval, or in real-time.

The FIM solutions also involve examination of the files and displaying information about when and how it was changed and who changed it, as well as any further action that can be done to restore those files if modifications are unauthorized.

J. v. Ogden summarises the components of FIM into three distinct features:

1. A Database: This database stores information on the original state of your files and configurations as cryptographic hashes.
2. Agents: These technical components measure your hardware and applications and send data back to your database for comparison.
3. User Interface: This is the visual component of the FIM for administrative users, which serves as the centralized portal for reporting, evaluation, change monitoring, and change control [3].

FIM technology is considered as a major part of cyber security processes and technology, owing to its ability to scan, analyse, and report unexpected changes to important files in an IT environment such as operating system (OS), database, and application software files. [3] elaborates on some of the benefits of FIM such as faster incident response times and real-time, continuous monitoring.

### FIM Global Market Share

According to [4], the File Integrity Monitoring Market was valued at USD 703.98 Million in 2019 and is projected to reach USD 1869.87 Million by 2027, growing at a CAGR of 14.01% from 2020 to 2027. [4] explains how the involvement of bodies such as the PCI-DSS and FISMA has led to FIM becoming one of the major requirements and act as a major factor driving the global file integrity monitoring market growth. They go on to describe how the rapid complexity of cyber-attacks and increasing threats to IT infrastructure are also among some of the major factors driving the global file integrity monitoring market. However, market growth is being hampered by the high costs associated with setting up advanced file integrity monitoring solutions and financial constraints among SMEs [4].

### Global FIM Market: Segmentation Analysis

The global segmentation is done based on Organisation Outlook, End-Use Outlook and Geography [4].

1. By Organisation Outlook
2. SMEs
3. Large Enterprises

Cutting-edge FIM solution is costly in terms of development and this has become a major restraint for the growth of the market. The high cost of Research and Development (R&D) expenses needed to develop advanced FIM solutions has led to high pricing of the security solutions. As the frequency of security breaches has increased over the past 5 years, organizations have increased their IT security investments to protect against advanced threats [4].

1. By End-Use Outlook
2. BFSI
3. Government
4. Healthcare
5. Education
6. IT & Telecom
7. Others

FIM solutions and services are provided to end-users to cater for their specific business requirements, compliance and security needs.

1. By Geography
2. North America
3. Europe
4. Asia Pacific
5. Rest of the world

According to [4], North America is estimated to hold the largest market size. Increasing penetration of the internet and incidents of attacks on enterprise IT infrastructure has driven the need for FIM solutions. Furthermore, rapid economic growth in the developing countries, along with improving regulatory reforms and economic stability is driving the file integrity monitoring market growth in APAC [4].

## Problem Statement

Enterprises firms and institutions store critical information about their customers on their organisational servers. Each server performs some form of response to client or user requests. If a response is given to a malicious user, it could lead to loss of critical data. The need for scalable security monitoring in such dynamic environments has become a critical need.

## Aims

This project aims to provide system and device administrators with a centralised tool where they can be able to monitor the files kept on their computers and network. Each file’s existence or lifetime on the computer can be monitored for any malicious alterations or modifications. Once a file is flagged to have been altered, the system will alert the responsible administrative agent to make changes or take precautionary action against the incident.

## Objectives

1. To generate and store a baseline for all the files to be monitored
2. To verify the baseline of stored files by comparing the stored hash against the current hash
3. To display the baseline status of files
4. To generate an alert if a compromised file is detected

## Signiﬁcance of the Project

With the rapid increase in workflow in the workspace, there is a huge demand for automation in order to cope with and ease the pressure on responsible personnel. This solution does that by automating the whole process of monitoring a host machine’s files and triggering the necessary action when a compromised file is detected. This means IT and Cyber Security personnel need only perform analytical tasks and not carry out the whole process of keeping an eye on each individual file.

## Methodology

This section describes and clarifies the entire FIM process from the initial setup stages to the point where a file is flagged as compromised and recommended action taken to restore the previous state of the file. As mentioned before, FIM is a process and is broken down into small sequential stages as follows:

### Authentication of Users

As a basic part of every secure system, we would need to authorize and keep track of users and their activities. The first step is to have responsible staff members register to use the system. Having users sign up onto the system allows for analytics at a later stage in the project.

### Identification of files to be monitored

Once a user has been granted access into the system, the next step is for them to identify the files they would like the system to monitor. As different departments have different personnel handling different workloads and files, this step would need to be divided into departmental level. Responsible personnel locate the directory that they would like for the tool to monitor. However, the overall FIM process would need to be overseen by the IT engineer responsible for managing the whole process.

### Creation of Baseline values for files to be monitored

Before they can be actively monitored for changes, the files must have an initial baseline based on their current state. The system therefore generates hashes for the files within the specified location. This is achieved by performing an MD5 Hash calculation for each individual file. In addition, the creation date and modification date of the file will also be taken into account.

### Monitoring changes

With a detailed baseline, we can proceed to monitor all designated files for changes. A regular interval after which a system scan will be carried out will be set, for example a minute.

### Sending an alert

This step is dependent on the results from the scan in (*iv)*. If a file’s scanned MD5 checksum calculated in *(iv)* is different from that generated in *(iii)*, it means that the file has been modified, in which case the system will send an alert to the user that a change has been detected.

### Reporting results

For analytics purposes, the system will log all activities and allow for a report of user activity to be generated.

## Scope of the Project

It should be taken into consideration that the project is built after the following factors:

### Host Machine Environments

This project is limited to monitoring files based on the host machine on which the tool will be installed. It does not focus on cloud-based files or server files that are hosted elsewhere. Administration and monitoring is limited only to the host machine or environment in which the software tool is running.

### Files and file types

Since, when a file is changed or updated, its hash value changes, the project focuses on files that are considered as archive files, whose information need not be changed. If archived file’s hash values are changed, then it means the files have been tampered with thus meaning the files have been compromised. For dynamic files, the system would only generate false positives as each alteration or modification of the original document would result in a different baseline value for the file, hence triggering an alert.

## Deﬁnition of Key Variables

### Files

A ‘file’ in this project refers to any document or database record contained and stored in the host machine. These files are organisational documents that are of utmost importance to them and can be of, but not limited to, the following file extensions: ***.pdf, .docx, .xlsx, .sqlite3, .py****.*

### Environment

The environment refers to the platform on which the system will run. We will create a virtual environment using Python in which Django, a Python Web Framework, and other dependencies and libraries, will be installed.

### FIM

FIM is an abbreviation that stands for File Integrity Monitoring. This refers to the whole process of tracking the integrity of files contained in a specified directory in the host machine by constantly comparing the current hash value against a baseline value already saved in the database.

### Baseline Value

The baseline value is the unique MD5 checksum that is assigned to each file within the monitored directory. The stored baseline will be compared with the current checksum at any given time. If the file’s contents are changed, the hash value changes and will therefore be different from the saved baseline value meaning the file will have been compromised.

### Compromised File

A compromised file is one that has been maliciously accessed and had its contents modified. For the purposes of this project and in order to prevent false positives, we will be focusing on archived records whose baseline value is set to be constant because there would be no need to alter their contents.

## Conclusion

This chapter sought to summarise the research problem and describe the proposed solution, its functionality and how it is intended to solve the existing problem. Having analysed the current technologies in the FIM industry and the gap that has been identified, it can be concluded that a more viable, robust and convenient solution is needed, hence this FIM tool.

# CHAPTER 2: LITERATURE REVIEW

## Introduction

This chapter analyses existing systems and critiques them in order to point out gaps or room for improvement that the proposed system will cover. The purpose then is to find a resolution for the identified flaws in the current and existing systems.

## Synthesis of Literature

## Conclusion

# REFERENCES

|  |  |
| --- | --- |
| [1] | C. Peter, "What Is File Integrity Monitoring and How Does It Work? - DZone Big Data," 25 November 2020. [Online]. Available: https://dzone.com/articles/what-is-file-integrity-monitoring-and-how-it-works. [Accessed 18 June 2021]. |
| [2] | WHOA.com, "File Integrity Monitoring PCI - DSS - FIM - PCI - DSS Compliance | WHOA.com," WHOA.com, 23 October 2013. [Online]. Available: https://www.whoa.com/file-integrity-monitoring. [Accessed 18 June 2021]. |
| [3] | J. v. Ogden, "How Does File Integrity Monitoring Work?," CIMCOR, 14 6 2016. [Online]. Available: https://www.cimcor.com/blog/how-does-file-integrity-monitoring-work. [Accessed 18 6 2021]. |
| [4] | V. M. Research, "Global File Integrity Monitoring Market Size By Organization Outlook, By End-Use Outlook, By Geographic Scope And Forecast," Verified Market Research, Amsterdam, 2021. |