|  |
| --- |
|  |

|  |
| --- |
|  |
| Bachelor of Science in  Information Technology Capstone |
| [insert your code]  Analysis of security appliances log files based on public lists of malicious internet protocol addresses  by  Panagiotis Krommydakis |
| New Hampshire, 2014 |

|  |
| --- |
|  |
| Bachelor of Science in  Information Technology Capstone |
| [insert your code]  Analysis of security appliances log files based on public lists of malicious internet protocol addresses  by  Panagiotis Krommydakis  Approved by:   |  |  |  | | --- | --- | --- | | Panagiotis Karampelas, PhD  Assistant Professor of Information Technology  Hellenic American University |  | Christos Pavlatos, PhD  Assistant Professor of Information Technology  Hellenic American University |   submitted: [insert date]  approved: [insert date]  A capstone submitted in partial fulfillment of the requirements for the degree of BSIT |
| New Hampshire, 2014 |

Abstract

This capstone project aims to investigate the effectiveness and efficiency of a detective safeguard, in the form of a software tool, which will help security administrators and/or security officers identifying –in a timely fashion- external attacks occurring at the perimeter of an enterprise network.

**Keywords**:

FIREWALL

SECURITY APPLIANCE

ATTACK

THREAT AGENT

Acknowledgements

I would like to thank the following people for the role they played throughout the entire process of completing this capstone:

* My Significant Other, Georgia, for her words of encouragement, and her love. Thanks for your support, patience and trust in me.
* To the faculty members and employees of Hellenic American University, who have guided, supported and assisted me along my undergraduate studies—thank you all for this precious journey.
* To Dr. Zoi Ekaterinidi and my colleagues at “Software Competitiveness International” who have always “been there” for me.
* To the kind, decent and hardworking people who have touched my life. You will always encourage me to be the best I can be.

TABLE OF CONTENTS

[Chapter 1. Introduction 1](#_Toc391034533)

[1.1. Write subtitle 1](#_Toc391034534)

[Chapter 2. Insert title 2](#_Toc391034535)

[Chapter 3. Insert a title 3](#_Toc391034536)

[Chapter 4. Insert a title 4](#_Toc391034537)

[Chapter 5. Introduction 5](#_Toc391034538)

[Chapter 6. Information Security Management 6](#_Toc391034539)

[Chapter 7. Research and Literature Review 7](#_Toc391034540)

[Chapter 8. The Business Scenario 8](#_Toc391034541)

[Chapter 9. Project Management 9](#_Toc391034542)

[Chapter 10. Requirements Engineering 10](#_Toc391034543)

[Chapter 11. Functional and Non-Functional Specification 12](#_Toc391034544)

[Chapter 12. System Architecture 13](#_Toc391034545)

[Chapter 13. Design Specification 14](#_Toc391034546)

[Chapter 14. Prototype Implementation 15](#_Toc391034547)

[Chapter 15. Test Specification 16](#_Toc391034548)

[Chapter 16. Prototype Evaluation 17](#_Toc391034549)

[Chapter 17. Future Work 18](#_Toc391034550)

[Chapter 18. Conclusion 19](#_Toc391034551)

List of figures

[Figure 1: Labelling page 3](#_Toc278814685)

List of Tables

[Table 1: Database Tables and Attributes 4](#_Toc278814691)

# Introduction

## Write subtitle

# Insert title

Here is an example of a footnote[[1]](#footnote-1) and an example of a reference (Datta & Joshi & Li & Wang, 2008).

Here is an example of bulleted list

* The same image can mean different things to different people (Chen & Rasmussen, 1999; Enser 2000, as cited in Matusiak 2006).
* Images can have several layers of meaning from specific to more abstract (Enser 2000; Jorgensen et al. 2001; Layne 1994, as cited in Matusiak 2006).

Here is an example of numbering

1. Primitive features (content-based image retrieval):
2. Objects
3. Inductive interpretation (subjective belief):

# Insert a title

Here is an example of a figure



Figure 1: Labeling page

# Insert a title

Here is an example of a table

Table 1: Database Tables and Attributes

|  |  |  |
| --- | --- | --- |
| ***Table name*** | ***Attribute*** | ***Description*** |
| TAGS | ID | The unique identifier of the inputted label |
|  | USER\_ID | The ID of the user |
| TAGS\_TYPE | ID | The unique identifier of the label type |
|  | TYPE | The description of the label type |
| PHOTOGRAPHS | ID | The unique identifier of the photograph |
|  | TITLE | The name of the photograph |
| USERS | USER\_ID | The unique identifier of the user |
|  | USERNAME | The user name |
|  | PASSWORD | The password |

# Introduction

Security appliances (such as firewalls), which are usually deployed at the network perimeter of an enterprise, filter inbound traffic based on statically configured rules and policies. Consequently, their preventive capabilities rely heavily upon the experience and competence of the security administrator. To that end, it is always a best practice to collect the log files, of such devices, in a syslog server for further (offline) analysis, by means of detective safeguard. Nevertheless, the information stored in the security log files is not usually correlated with and evaluated against an external database of threats. Manual inspection of the log files is also a non-reliable process for offline incident detection. On the other hand, major IT Security Organizations and Institutions maintain lists of malicious IPs, which are collected from various sources, evaluated/validated and finally published on the Internet. A tool that will automatically parse the security appliances log files against these lists could be a trustworthy companion of the security administrators and/or security officers in their everyday task to protect the company perimeter.

One can find off-the-self software packages, implemented by major international security companies, which are capable of real-time security incident detection. However, such solutions are very expensive and require highly trained personnel to operate them. Therefore, the introduction of such system, by a company, is normally accompanied by the respective Managed Security Services, and a 24/7 SLA, which is also very expensive. Small–Medium Businesses (SMBs) cannot afford the overall cost of such solution. This capstone project has the aforementioned target group of companies in focus, with a view to propose a cost-effective but reliable solution for offline incident detection.

The overall objective of the capstone is to achieve a significant level of integration between the security appliances, which are used in the network perimeter of enterprises as preventive security controls, and the up-to-date information published by trusted IT Security Organizations/Institutions.

As far as the expected outcome of this project is concerned, the target is to create a proof of concept application that will analyze (by means of text parsing) syslog files retrieved from security appliances (i.e. ASA etc.), based on info collected from various lists of malicious IPs (i.e. SANS list), and identify the potential attacks against the enterprise perimeter.

# Information Security Management

Nowadays, enterprises have to deal with numerous compliance requirements (regulations) while they should operate in a hostile environment in terms of external attacks and malicious insiders (Majdalawieh, 2014). It is essential to identify the information risk profile of the organization. According to J. Pironti (2013), development of an information risk profile is an enterprise-wide process. Many stakeholders shall be involved, such as upper management, auditors, legal and compliance departments as well as various data and process owners. The information risk profile depicts the nature, the priority, the probability and the potential impact of all the risks the organization is exposed to (Pironti, 2013). In a second step senior management should decide which of these risks are acceptable and initiate remedy actions, through the risk mitigation process, for the rest of them (Pironti, 2013). The information risk profile should be developed and modeled in such a way it is easy to comprehend, covers the entire organization and adds value as it is a very useful tool for decision support on the senior management and stakeholders’ levels (Conrad, Misenar, & Feldman, 2012). The information risk profile provides guidance that supports the strategic and tactical objectives of the enterprise through alignment of the company’s goals with the means to achieve them (Harris, 2008). The guiding principles to establish an information risk management program include the following (Conrad, Misenar, & Feldman, 2012):

* Identify the regulations the organization shall comply with;
* Identify all the critical information assets and the corresponding business processes;
* Identify the data, application and system owners as well as the custodians, which are assigned with the operational responsibility;
* Categorize, classify and prioritize the assets according to their value for the organization;
* Ensure the availability of critical business processes and of related data;
* Identify and evaluate accurately the applicable threats, possible vulnerabilities and their associated risks. Thus, the senior management will be able to make appropriate decisions for risk management.
* Ensure that the right controls and safeguards are implemented. Moreover, monitor in a permanent basis these controls in terms of functionality so as to be sure they remain up to date and valid.
* Ensure that budget and resource allocation is done efficiently to maximize the Return-On-Investment (ROI).

Protection of information is based on ensuring the effectiveness of crucial security features. More particularly, Confidentiality, Integrity, and Availability (called the CIA triad) are considered as the cornerstones of information security (Conrad, Misenar, & Feldman, 2012). Confidentiality aims to prevent the unauthorized disclosure of information by means of unauthorized read access to data (Harris, 2008).

Integrity aims to prevent unauthorized modification of information by means of unauthorized write access to data (Harris, 2008). There are two distinguished variants of integrity: (a) data integrity and (b) system integrity (Harris, 2008). Data integrity seeks to protect information against unauthorized modification during data processing while system integrity seeks to protect a system that processes critical data (Harris, 2008). Availability ensures that both systems and information are available when needed and moreover in the form requested by the user for the latter case (Harris, 2008). Protecting the CIA triad has significant impact on the usability factor. Therefore, it is crucial to always balance the need to protect data and systems with the need to keep them accessible for the end users (proportionality). The opposite of the CIA triad is Disclosure, Alteration, and destruction (DAD) (Conrad, Misenar, & Feldman, 2012). Disclosure is the unauthorized disclosure of information, alteration is the unauthorized modification of data, and destruction refers to intentional or unintentional unavailability of data and/or systems (Conrad, Misenar, & Feldman, 2012).

Supplementary features applicable to information security are Authentication, Authorization, and Accountability (AAA). AAA features are applied in a strict sequential order (Conrad, Misenar, & Feldman, 2012). Authentication consists of a verified (that means authenticated) identity claim (Harris, 2008). Identity without proof is a very week – almost meaningless control. At a glance, there are three types of identity verification: (a) something you know (i.e. password), (b) something you have (i.e. smartcard or security token) and (c) something you are (biometrical characteristics) (Harris, 2008). Strong authentication consists of the combination of two or even more verification methods such as a smartcard that can be used after a valid password has been inserted (Harris, 2008). Authorization determines the actions one can perform on a system, and its data, once she has been identified and authenticated (Conrad, Misenar, & Feldman, 2012). Actions may include read, write, or execute rights on data files and system resources (Conrad, Misenar, & Feldman, 2012). Accountability keeps authenticated users accountable for their actions (either authorized or unauthorized and in some cases malicious) (Conrad, Misenar, & Feldman, 2012). This is typically done using system logs/traces and analyzing audit data (transaction logs) (Conrad, Misenar, & Feldman, 2012).  Nonrepudiation ensures that no user could possibly deny having performed an action (Conrad, Misenar, & Feldman, 2012).

Security concepts should by always applied in the context of security valid methods. For instance, least privilege ensures that users are granted the absolutely minimum access rights, which are necessary to perform efficiently their tasks (Harris, 2008). Need to know ensures that users have access only to the specific pieces of information they need to execute their tasks (according to their authorization level always) (Harris, 2008). Defense-in-Depth (also called layered security or bull’s-eye) applies multiple safeguards (also called controls: measures taken to reduce risk) to protect an organization (Harris, 2008). While any security control may fail under certain circumstances deploying multiple controls on various systems improve the confidentiality, integrity, and availability of the enterprise information (Conrad, Misenar, & Feldman, 2012).

“The average person does a poor job of accurately analyzing risk: if you fear the risk of dying while traveling, and drive from New York to Florida instead of flying to mitigate that risk, you have done a poor job of analyzing risk” (Conrad, Misenar, & Feldman, 2012). Travelling by car is much riskier than by airplane when taking into consideration the related accident statistics. Accurate Risk Analysis is a critical skill for every professional assigned with such responsibilities. Risk analysis decisions will dictate which safeguards should be deployed to protect the critical assets as well as the amount of money and resources to be spent (Conrad, Misenar, & Feldman, 2012). Quantitative and Qualitative Risk Analysis are two methods for analyzing risk (Scherling, 2011). Quantitative Risk Analysis uses hard metrics, such as money figures. Qualitative Risk Analysis uses simple describing values. While quantitative is more objective, qualitative is 100% subjective and relies on the experience and the background of the managers performing it (Scherling, 2011). Quantitative Risk Analysis is more difficult (Harris, 2008). Assets are valuable tangible and intangible enterprise property that shall be protected. Assets can be data, systems, people, buildings etc (Scherling, 2011). The value of an asset for the organization will finally determine which and how sophisticated safeguards should be deployed. People are a priori the most valuable asset (Harris 2008). A threat is defined as a potential negative (or even catastrophic) incident, like earthquakes, power or network outages (Conrad, Misenar, & Feldman, 2012). Vulnerability is a weakness that may lead to a threat realization (Conrad, Misenar, & Feldman, 2012). “Examples of vulnerabilities are data center without proper backup power, or a Microsoft Windows XP system that has not been patched in a few years” (Conrad, Misenar, & Feldman, 2012). Risk value occurs as the multiplication of a specific threat with the related vulnerability: “Risk = Threat × Vulnerability” (Harris, 2008). Usually, a number in the range 1-5, is assigned to both threats and vulnerabilities (Harris, 2008). While the range is a matter of definition it is essential to keep it consistent when comparing different risks (Conrad, Misenar, & Feldman, 2012). An additional parameter, called impact, can be also taken under consideration when calculating the risk as follows: “Risk = Threat × Vulnerability × Impact” (Harris, 2008). Impact depicts the severity of the damage when a specific risk is realized (Harris, 2008). “The Risk Analysis Matrix uses a quadrant to map the likelihood of a risk occurring against the consequences (or impact) that risk would have” (Conrad, Misenar, & Feldman, 2012). The Risk Analysis Matrix allows to perform Qualitative Risk Analysis based on likelihood (from "rare" to "almost certain") and consequences (or impact), from "insignificant" to "catastrophic" (Conrad, Misenar, & Feldman, 2012). The resulting indicators are Low (L), Medium (M), High (H), and Extreme Risk (E) (Conrad, Misenar, & Feldman, 2012). Low risks can be managed via the already available processes without escalation (Scherling, 2011). Moderate risk requires medium-level management involvement, while high risks require escalation to senior management (c-level) (Scherling, 2011). Extreme high risks require immediate action by senior management including a strict mitigation and follow up plan (Scherling, 2011). The goal of the matrix is to reveal high likelihood/high impact risks, and mitigate them to acceptable levels (Scherling, 2011). The Annualized Loss Expectancy (ALE) calculation helps to calculate the annual cost of a loss due to a risk realization (Harris, 2008). Therefore, ALE is primary parameter to be considered so as to make decisions to mitigate the risk (Scherling, 2011). ALE depends on the Asset Value (AV) and the so-called exposure factor (Harris, 2008). The Exposure Factor (EF) is defined as the percentage of the asset value that is lost due to an event (Harris, 2008). Accordingly, the Single Loss Expectancy (SLE) is the cost for a single occurrence as follows: SLE is the Asset Value (AV) times the Exposure Factor (EF) (Harris, 2008).  The Annual Rate of Occurrence (ARO) is the number of incidents occurring in yearly basis (Harris, 2008). The Annualized Loss Expectancy (ALE) is the yearly cost due to a certain risk (Conrad, Misenar, & Feldman, 2012). It is calculated by multiplying the Single Loss Expectancy (SLE) times the Annual Rate of Occurrence (ARO) (Conrad, Misenar, & Feldman, 2012). The Total Cost of Ownership (TCO) is the budget required to deploy a safeguard so as to mitigate a risk to acceptable levels (Conrad, Misenar, & Feldman, 2012). The Return on Investment (ROI) is the amount of money saved by deploying a safeguard. Consequently, if the annual Total Cost of Ownership (TCO), for maintaining a safeguard in operational state, is less than the Annualized Loss Expectancy (ALE), then ROI is positive (Conrad, Misenar, & Feldman, 2012). Otherwise, alternatives for the specific risk handling shall be considered.

As soon as the organization has assessed risk the next steps in the risk management process have to be decided. Valid options include accepting the risk, mitigating the risk, transferring the risk, and avoiding the risk (Harris, 2008). First of all, some risks may be accepted because it is cheaper to leave an asset unprotected rather than spending a ton of money to protect it. Upper management notification and approval is always needed in such cases. Low likelihood and low impact risks are good cases for risk acceptance (Conrad, Misenar, & Feldman, 2012). However, high and extreme risks can never be accepted as this is a recipe for disaster. Furthermore, risks related to regulations or safety should be confronted otherwise legal penalties are applied. Thus, accepting the risk is not really an available option in such cases. Mitigating the risk means lowering the risk to acceptable levels. Transferring the risk means to purchase insurance so as to deal with the results of the risk occurrence (Conrad, Misenar, & Feldman, 2012). What is necessary to bear in mind is that while risks may be transferred, responsibility remains always with the organization.

Effective information risk management is a critical success factor for todays’ enterprises (Samuel and Sanchez, 2013). Establishment of the risk management program should take place using the top-bottom approach in order to never question the senior management commitment to the necessity of always improving the security posture of the company. But to be successful it has to be embedded into the company’s culture and every employee has to understand she should identify the potential risk areas on the functions she works. Furthermore, senior management should ensure the appropriate controls and safeguards are in place to quickly identify risk areas and mitigate the risk findings.

# Research and Literature Review

# The Business Scenario

# Project Management

# Requirements Engineering

A far as requirements definition is concerned the industry standard process of Requirements Engineering has been followed. At a glance, requirements engineering can be divided into the following discrete but sequential steps:

* Requirements elicitation,
* Requirements analysis and negotiation,
* Requirements specification,
* System modeling,
* Requirements validation,
* Requirements management.

As no separated requirements specification will be created, the entire process results are incorporated in the upcoming chapters.

# Functional and Non-Functional Specification

# System Architecture

# Design Specification

# Prototype Implementation

# Test Specification

# Prototype Evaluation

# Future Work

# Conclusion

#### Appendix I – Public Lists of Malicious IPs

#### Appendix II – Sample Firewall Log Files

#### Appendix II – Use Case Template

The **IBM® Rational® Modeler** template for use cases has been used:

**Purpose**

*[This section aims at presenting the purpose of the use case and what this functionality does briefly.]*

**Subsystem**

*[The Subsystem section defines the subsystem where the use case belongs to.]*

**Primary Actor(s)**

*[The actor(s) executing the actions specified in the use case specification.]*

**Precondition(s)**

*[A precondition (of a use case) is the state of the excise operation that must be present prior to a use case being performed.]*

**Post-condition(s)**

*[A post-condition (of a use case) is a list of possible states the excise operation can be in immediately after a use case has finished.]*

**Associated Use Case(s)**

*[Use cases to which the current use case is associated. In this paragraph also the use cases that need to be executed prior to the current use case are listed]*

**Basic Flow**

*[The detailed steps followed during the execution of the use case.]*

Step 1 Format: Step [The number of the Basic Flow]

*[This use case starts when the actor performs an action. The use case describes what the actor does and what the system does in response.]*

[Alternative Flow – name] optional

[Exception Flow – name] optional

[Extension Point – name] optional

Step 2

**Alternative Flow – name**

Step 1 Format Step n.[The number refers to the nth step in this alternative flow]

*[This part of the use case describes an alternative system behavior referred by a specific step of the Basic Flow]*

**Exception Flow - name**

Step 1 Format Step n. [The number refers to the nth step in this exception flow]

*[This part of the use case describes an exception system behavior referred by a specific step of the Basic Flow and/or alternative flows]*

**Extension Point – name**

*[This part of the use case describes an extension system behavior referred by a specific step of the Basic Flow and/or alternative flows]*

**Special Requirements**

*[Supplementary requirements the current use-case is associated with.]*

**Activity Diagram (optional)**

*[Activity diagram are used to support the documented Primary and Alterative flows. The textual description takes precedence over the diagram.]*

**Additional Info**

*[Additional information is used to support the use case, e.g. screenshots]*

References

Datta, R., Joshi, D., Li, J., & Wang, J. (2008, April). Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys*, *40*(2), 1-60. Retrieved August 2, 2009, from Academic Search Complete database.

1. http://www.airliners.net/ [↑](#footnote-ref-1)