BENCHMARKING FRAMEWORK FOR IDS CLASSIFIERS IN TERM OF SECURITY AND PERFORMANCE BASED ON MULTICRITERIA ANALYSIS

AMNEH HUSSEIN MOHD ALAMLEH

SULTAN IDRIS EDUCATION UNIVERSITY

2022

BENCHMARKING FRAMEWORK FOR IDS CLASSIFIERS IN TERM OF SECURITY AND PERFORMANCE BASED ON MULTICRITERIA ANALYSIS

AMNEH HUSSEIN MOHD ALAMLEH

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FORTHE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF ART, COMPUTING & CREATIVE INDUSTRY
SULTAN IDRIS EDUCATION UNIVERSITY

2022

UNIVERSITI PENDIDIKAN SULTAN IDRIS
اونيۏرسيتي قندديقن سلطان ادريس
SULTAN IDRIS EDUCATION UNIVERSITY

UPSI/IPS-3/BO 32

Please tick (✓)
Project Paper
Masters by Research
Masters by Mix Mode
Ph.D. ✓

INSTITUTE OF GRADUATE STUDIES
DECLARATION OF ORIGINAL WORK

This declaration is made on the 17/5/2022 .

i.Student's Declaration:
I'm Amneh Hussein Mohd Alamleh - P20171001172 - Faculty of Art, Computing, and Creative Industry
Hereby declares that the dissertation / thesis for titled (BENCHMARKING FRAMEWORK FOR IDS CLASSIFIERS IN TERM OF SECURITY AND PERFORMANCE BASED ON MULTICRITERIA ANALYSIS) is my original work. I have not plagiarized from any other scholar's work and any sources that contain copyright had been cited properly for the permitted meanings. Any quotations, excerpt, reference or re-publication from or any works that have copyright had been clearly and well cited.

Amneh Alamleh

Signature of the student

ii. Supervisor's Declaration:
I'm Dr. Aws Alaa hereby certify that the work entitled (benchmarking methodology for multiclass classifiers based on multi-criteria decision analysis) was prepared by the above-named student, and was submitted to the Institute of Graduate Studies as a partial / full fulfillment for the conferment of the requirements for Doctor of Philosophy (By Research), and the aforementioned work, to the best of my knowledge, is the said student's work.

18/5/2022

Certified True Copy
DR. AWS ALAA ZAIDAN
Department of Computing
Faculty of Art, Computing and Creative Industry
Sultan Idris Education University

Date

Signature of the Supervisor

INSTITUT PENGAJIAN SISWAZAH /
INSTITUTE OF GRADUATE STUDIES

BORANG PENGESAHAN PENYERAHAN TESIS/DISERTASI/LAPORAN KERTAS PROJEK
DECLARATION OF THESIS/DISSERTATION/*PROJECT PAPER FORM*

Tajuk / *Title*: BENCHMARKING FRAMEWORK FOR IDS CLASSIFIERS IN TERM OF SECURITY AND PERFORMANCE BASED ON MULTICRITERIA ANALYSIS

No. Matrik /*Matric No.*: P20171001172

Saya / *I* : Amneh Husseing Mohd Alamleh

mengaku membenarkan Tesis/Disertasi/Laporan Kertas Projek (Kedoktoran/Sarjana)* ini disimpan di Universiti Pendidikan Sultan Idris (Perpustakaan Tuanku Bainun) dengan syarat-syarat kegunaan seperti berikut:-
acknowledged that Universiti Pendidikan Sultan Idris (Tuanku Bainun Library) reserves the right as follows:-

1. Tesis/Disertasi/Laporan Kertas Projek ini adalah hak milik UPSI.
The thesis is the property of Universiti Pendidikan Sultan Idris

2. Perpustakaan Tuanku Bainun dibenarkan membuat salinan untuk tujuan rujukan dan penyelidikan.
Tuanku Bainun Library has the right to make copies for the purpose of reference and research.

3. Perpustakaan dibenarkan membuat salinan Tesis/Disertasi ini sebagai bahan pertukaran antara Institusi Pengajian Tinggi.
The Library has the right to make copies of the thesis for academic exchange.

4. Sila tandakan (√) bagi pilihan kategori di bawah / Please tick (√) from the categories below:-

☐	**SULIT/*CONFIDENTIAL***	Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub dalam Akta Rahsia Rasmi 1972. / *Contains confidential information under the Official Secret Act 1972*
☐	**TERHAD/*RESTRICTED***	Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan ini dijalankan. / *Contains restricted information as specified by the organization where research was done.*
☑	**TIDAK TERHAD / *OPEN ACCESS***	

Amneh Alamleh

(Tandatangan Pelajar/ Signature)

Tarikh: 18/ 5 /2022

Certified True Copy
DR. AWS ALAA ZAIDAN
Department of Computing
Faculty of Art, Computing and Creative Industry
Sultan Idris Education University

(Tandatangan Penyelia / *Signature of Supervisor*)

& (Nama & Cop Rasmi / *Name & Official Stamp*)

Catatan: Jika Tesis/Disertasi ini **SULIT** @ **TERHAD**, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan ini perlu dikelaskan sebagai **SULIT** dan **TERHAD**.

Notes: If the thesis is CONFIDENTAL or RESTRICTED, please attach with the letter from the related authority/organization mentioning the period of confidentiality and reasons for the said confidentiality or restriction.

ACKNOWLEDGMENT

"In the name of Allah, the Most Gracious and the Most Merciful"
Alhamdulillah, first and foremost, praise be Allah, the Cherisher and Sustainer of the World and to the Prophet Muhammad (Peace and Blessings of Allah Be Upon Him) who was sent by Allah to be a great teacher to the mankind.

I would like to extend my appreciation to those who involved and give a helpful hand in ensuring the success of this research. This research would not have come to fruition without all your help and supports.

I am very grateful to my main supervisor Dr. Aos Alaa Zaidan for guiding me during my work on this research. I would like to express my sincere thanks and gratitude to him for his continuous guidance, support and patience.

I would like to express my sincere thanks to my co supervisors; Dr. Bilal Bahaa Zaidan, Dr. Osamah Abahrey and Dr. Mashitoh Hashim for their help and support as well, my appreciation and gratitude for them.

I am also very grateful to my brothers and sisters for their support blessing, patience, love, and encouragement.

Finally, I would like to thank all friends who have helped me and encouraged me. Thank you. Allah blesses you.

DEDICATION

To the souls of my parents. To my beloved brothers; Muhammad, Taha and Maryam. I dedicate this work. Without your support, inspiration and love, this work would not have been possible.

ABSTRACT

This research aims to assist the developers of intrusion detection systems (IDS) to make the right selection decision of a suitable classification model. Many classification algorithms have been developed to be used in an IDS detection engine. Developers of IDS have been facing challenges in how to evaluate and benchmark classifiers. Different perspectives and multiple, conflicting importance evaluation criteria represent the challenges in evaluation, benchmarking and selecting suitable IDS classifiers. The current evaluation studies depend on evaluating the IDS classifier from a single incomplete perspective. In each study, the evaluations have been achieved with reference to some security-related evaluation criteria and ignore performance criteria. Furthermore, the weighting process that reflects the importance of each criterion depended on a personal subjective perspective. The goal of this thesis is to set a new standardisation and benchmarking framework based on a set of standardised criteria and set of unified multi-criteria decision-making (MCDM) methods that overcome the shortage. This study attempts to establish and standardise IDS classifier evaluation criteria and construct a decision matrix (DM) based on crossover of the standardised criteria and 12 classifiers. This DM was evaluated using datasets consist of 125,973 records; each record consists of 41 features. Subsequently, the classifiers are evaluated and ranked using unified MCDM techniques. The proposed framework consists of three main parts: the first for standardising evaluation criteria, the second for constructing the DM and the third for developing weighting and ranking unified MCDM methods and IDS classifiers evaluation and benchmarking. The fuzzy Delphi method (FDM) has been used for criteria standardisation. Integrated weighting methods using direct rating and the entropy objective method are developed to calculate the weights of the criteria. The Vlse Kriterijumska Optimizacija Kompromisno Resenje (VIKOR) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) ranking methods were integrated into a unified method for ranking the selected classifiers. The Borda voting method was used to unify the different ranks and perform a group ranking context. An objective validation process has been used to validate the ranking results. The mean ± standard deviation was computed to ensure that the classifier ranking underwent systematic ranking. The following results were confirmed. (1) FDM is a suitable way to reach a standard set of evaluation criteria. (2) Using an integrated (subjective, objective) weighting method can find the suitable criteria weights. (3) A unified ranking method that integrates VIKOR and TOPSIS effectively solves the classifier selection problem and (4) the objective validation shows significant differences between the groups' scores, indicating indicates that the ranking results of the proposed framework were valid. (5) The evaluation of the proposed framework shows an advantage over the benchmarked works with a percentage of 100%. The implications of this study benefit IDS developers in making the right decisions in selecting the best classification model. Researchers can use the proposed framework for evaluation and selection in similar evaluation problems.

RANGKA KERJA PENANDA ARAS UNTUK PENGELAS IDS DARI SEGI KESELAMATAN DAN PRESTASI BERDASARKAN ANALISIS PELBAGAI KRITERIA

ABSTRAK

Penyelidikan ini bertujuan untuk membantu pembangun sistem pengesanan pencerobohan (IDS) membuat keputusan yang tepat untuk model pengelasan yang sesuai. Pelbagai algoritma pengelasan telah dibangunkan untuk digunakan dalam enjin pengesanan IDS. Pembangun IDS menghadapi beberapa cabaran dalam proses menilai dan menanda aras pengelas. Perspektif yang berbeza dan pelbagai kriteria penilaian yang bercanggah mewakili cabaran-cabaran dalam menilai, menanda aras dan memilih pengelas IDS yang sesuai. Kajian penilaian terkini bergantung kepada penilaian pengelas IDS daripada satu perspektif yang tidak lengkap. Dalam setiap kajian, penilaian dicapai dengan merujuk kepada beberapa kriteria penilaian berkaitan keselamatan dan mengabaikan kriteria prestasi. Tambahan pula, proses pemberatan yang mencerminkan kepentingan setiap kriteria bergantung kepada perspektif subjektif peribadi. Matlamat kajian ini adalah untuk menetapkan satu rangka kerja penyeragaman dan penandaarasan baharu berdasarkan satu set kriteria piawai dan satu set kaedah membuat keputusan berbilang kriteria bersatu (MCDM) untuk mengatasi kekurangan tersebut. Kajian ini telah mewujudkan dan menyeragamkan kriteria penilaian pengelas IDS dan membina matriks keputusan (DM) berdasarkan persilangan bagi kriteria piawai dan 12 pengelas. DM ini dinilai menggunakan set-set data yang terdiri daripada 125,973 rekod dan setiap rekod mengandungi 41 ciri. Seterusnya, pengelas dinilai dan ditarafkan menggunakan teknik MCDM bersatu. Rangka kerja yang dicadangkan terdiri daripada tiga bahagian utama: bahagian pertama adalah untuk menyeragamkan kriteria penilaian; bahagian kedua adalah untuk membina DM; dan bahagian ketiga adalah untuk membangunkan pemberat dan menarafkan kaedah-kaedah MCDM bersatu, selain daripada menilai dan menanda aras pengelas IDS. Kaedah Delphi kabur (FDM) telah digunakan untuk penyeragaman kriteria. Kaedah pemberat bersepadu yang menggunakan penarafan langsung dan kaedah objektif entropi telah dibangunkan untuk mengira wajaran kriteria. Kaedah penarafan Vlse Kriterijumska Optimizacija Kompromisno Resenje (VIKOR) dan Teknik untuk Keutamaan Pesanan mengikut Keserupaan kepada Penyelesaian Ideal (TOPSIS) telah disepadukan menjadi kaedah bersatu untuk menaraf pengelas terpilih. Kaedah pengundian Borda telah digunakan untuk menyatukan taraf-taraf yang berbeza dan melaksanakan konteks penarafan kumpulan. Proses pengesahan objektif telah digunakan untuk mengesahkan keputusan penarafan. Purata sisihan piawai dikira untuk memastikan penarafan pengelas menjalani penarafan yang sistematik. Keputusan berikut telah disahkan: (1) Dengan menggunakan FDM, 17 daripada 20 kriteria penilaian (14 kriteria untuk keselamatan dan 3 kriteria untuk prestasi) mencapai konsensus pakar; (2) Wajaran keseluruhan menunjukkan bahawa kawasan di bawah lengkung mempunyai berat terendah (0.036, 0.025, 0.020), manakala masa CPU mempunyai berat tertinggi (0.118, 0.196, 0.235); (3) Penarafan kumpulan kaedah VIKOR-TOPSIS bersepadu menunjukkan bahawa BayesNet ialah pilihan terbaik, manakala SVM ialah pilihan yang paling teruk; dan (4) Proses penilaian menunjukkan bahawa rangka kerja yang dicadangkan mengatasi keputusan kajian-kajian berkaitan dengan 83% mata perbandingan. Implikasi kajian ini boleh memberi manfaat kepada pembangun IDS dalam membuat keputusan yang tepat ketika memilih model pengelasan yang terbaik.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Acc	Accuracy
AHP	Analytical Hierarchy Process
ANN	Artificial Neural Network
AUC	Area Under the ROC Curve
CPUT	CPU Time
CPUU	CPU Usage
DT	Decision Tree
FDM	Fuzzy Delphi Method
FNR	False Negative Rate
FPR	False Positive Rate
GDM	Group Decision-Making
IBK	Instance-Based Knowledge
ICC	Incorrectly classified instances
IDES	Intrusion Detection Expert System
IDS	Intrusion Detection System
IRS	Intrusion Response System
MAE	Mean Absolute Error
MCC	Matthews Correlation Coefficient
MCDM	Multi-criteria decision-making
ML	Machine Learning
MLP	Multilayer Perceptron

MU	Memory Usage
NN	Neural Network
NN RBF	Neural network radial basis function
NPV	Negative Predicted Value
PPV	Positive Predicted Value
PROMETHEEII	Ranking Organization Method of Enrichment Evaluation
RAE	Relative Absolute Error
RF	Random Forest
RMSE	Root Mean Squared Error
ROC	Receiver Operating Characteristic
RRSE	Root Relative Squared Error
SD	ScienceDirect
SOM	Self-Organising Map
SVM	Support Vector Machine
TesT	Testing Time
TNR	True Negative Rate
TOPSIS	Techniques for Order of Preference by Similarity to Ideal Solution
TPR	True Positive Rate
TraT	Training Time
VIKOR	Vlsekriterijumska Optimizcija I Kaompromisno Resenje
WoS	Web of Science
WSM	Weighted Sum Model

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter provides an overview of the fundamentals of this research. In addition to the topic background, this chapter presents the problem that will be addressed, as well as the scope, goals, objectives and outline of the thesis.

This chapter will be organised as follows. A brief background of the research components is presented in Section 1.2. The problem statement, on which the direction of the research is based, is identified and introduced in Section 1.3. The research questions are presented in Section 1.4. The research objectives are determined in Section 1.5. The research objectives, research questions and the specific and general problem are connected in Section 1.6. The scope of the study is discussed in Section 1.7. The significance of the study is discussed in Section 1.8. The organisation of the thesis is outlined in Section 1.9. Finally, this chapter is summarised in Section 1.10.

1.2 Research Background

An intrusion detection system (IDS) is a crucial technique used to protect a network against attacks (Arnaboldi & Morisset, 2021). The first pattern of a real-time IDS was established by Dorothy Denning and Peter Neumann between 1984 and 1986. Their model was originally called an intrusion detection expert system (IDES) The IDES was initially an expert system with rule-based training to expose a recognised malicious activity (Bruneau, 2001; Yost, 2015). Since then, IDSs have become a key area of computer security research and development. Intrusion detection is a common cybersecurity technique tasked with exposing malicious actions of hosts and/or networks in environments (Liang, Ma, Sadiq, & Yeung, 2019; Young, Zambreno, Olufowobi, & Bloom, 2019). In case an intrusion is been detected, the classical IDS raises an alarm.

Recently, IDS faces significant major challenges from several perspectives: the massive size of network traffic data, high dimensional training dataset, frequent alterations in environments, changing nature of intrusions and the need for real-time detections. Responding to these challenges, a tremendous amount of detection methods and strategies have been proposed, with machine learning (ML) methods being one of the most common. However, evaluation and benchmarking of these methods is problematic due to multiple criteria issues (Amudha, Karthik, & Sivakumari, 2013; G. Kumar, 2014)(Magán-Carrión, Urda, Díaz-Cano, & Dorronsoro, 2020).

Two basic sets of criteria, namely, security and performance, are commonly used to evaluate and benchmark the IDS classifiers. In the security aspect, the classifier is capable of producing the desired results. The security of ML-based IDS can be explored through measuring criteria of the accuracy (Azmi & Pishgoo, 2013), true positive rate (Azmi & Pishgoo, 2013), true negative rate (Azmi & Pishgoo, 2013), false positive rate (Saracino, Sgandurra, Dini, & Martinelli, 2018), false negative rate (Donkal & Verma, 2018), negative predicted value (Azmi & Pishgoo, 2013), positive predicted value (Carvalho, Abrão, de Souza Mendes, & Proença Jr, 2018), f-score (Donkal & Verma, 2018), incorrectly classified instances (Panigrahi & Borah, 2018), the area under curve AUC (Carvalho et al., 2018), mean absolute error (Panigrahi & Borah, 2018), root mean squared error (Panigrahi & Borah, 2018), relative absolute error (Panigrahi & Borah, 2018), root relative squared error (Panigrahi & Borah, 2018) and Matthews correlation coefficient (Azmi & Pishgoo, 2013). By contrast, the performance aspect means producing the desired results using minimum resources in the computing environments (i.e. time and space). The performance of ML-based IDS can be explored by calculating the criteria of testing time (Rahman, Ahmed, & Kaiser, 2016), training time (Kabir & Hu, 2014), CPU time (Mehetrey, Shahriari, & Moh, 2016), CPU usage (Saracino et al., 2018) and memory usage (Midi, Rullo, Mudgerikar, & Bertino, 2017). Effective and efficient evaluation of ML-based IDSs to select the best IDS classifier are critical and important processes because they improve the security of the computer networks and information systems in addition to providing comprehensive information to developers, potentially aiding in manufacturing effective and efficient versions of IDSs.

Multi-criteria decision-making (MCDM) methods widely contribute to IDS and have been adopted in its different parts, such as (1) data analysis and detection part as in (El-Alfy & Al-Obeidat, 2014), (Yan, Gong, & Deng, 2016), (KP, 2019), (Saraeian & Shirazi, 2020), (Sharma & Kaul, 2018), (2) response part for deciding the suitable response as in (Iannucci & Abdelwahed, 2016), (Shameli-Sendi, Louafi, He, & Cheriet, 2016), (Singh & Kaushik, 2018), (Singh & Kaushik, 2019) and (3) in different IDS related evaluations, such as comparing different security controls involving IDS as in (Lv, Zhou, & Wang, 2011), evaluation of different IDS architectures context as in (Zbakh, Elmahdi, Cherkaoui, & Enniari, 2015) and evaluation and ranking for different IDS ML classifiers as in (Ahmad, Abdullah, & Alghamdi, 2010), (Peng, Kou, Wang, & Shi, 2011), (Robinson & Thomas, 2015), (Panigrahi & Borah, 2018) and (Patsariya & Singh, 2019).

MCDM is a method that deals with decisions involving the selection of the most suitable alternative from a group of alternatives in accordance with a group of criteria or attributes (Antunes & Henriques, 2016). MCDM is widely used in several fields for different applications. MCDM finds and ranks appropriate solutions to choose the suitable alternative (Aruldoss, Lakshmi, & Venkatesan, 2013). MCDM applications involve energy management (J.-J. Wang, Jing, Zhang, & Zhao, 2009), energy planning (Haralambopoulos & Polatidis, 2003), transportation (Qu & Chen, 2008), geographical information systems (Gbanie, Tengbe, Momoh, Medo, & Kabba, 2013) (Ligmann-Zielinska & Jankowski, 2012), resource and budget allocation (Phillips & e Costa, 2007) and medicine (O. S. Albahri et al., 2021).

At present, evaluation and benchmarking of IDS classifiers in terms of security and performance is incomplete and scattered. Some studies have discussed the evaluation and benchmarking of IDS classifiers, but they are limited to one of the evaluation aspects while ignoring the others (Robinson & Thomas, 2015), (Patsariya & Singh, 2019). Moreover, they ignored the importance of each criteria weight or give them the same importance (Robinson & Thomas, 2015), (Panigrahi & Borah, 2018), (Patsariya & Singh, 2019). Giving equal weights criteria cancels the different importance of criteria. Although IDS classifiers development has received considerable research attention, IDS classifier evaluation and benchmarking are limited and require more consideration (Magán-Carrión et al., 2020). Thus, according to the huge number of developed supervised ML-based IDS studies, this field continuously faces major challenges in the evaluation of ML classifiers used in IDSs considering the discussed evaluation criteria of security and performance (Robinson & Thomas, 2015), (Patsariya & Singh, 2019) (Arshad et al., 2020).

1.3 Research Problem

An IDS is a crucial technique used to protect a network against attacks. Recently, existing IDSs have been approached by several ML classifiers to maintain advances in IDS research and to increase detection rate, decreasing false alarm rate and decreasing processing costs (Hodo, Bellekens, Hamilton, Tachtatzis, & Atkinson, 2017).

Variations have been found in using the evaluation criteria in both aspects (i.e. security and performance) (Milenkoski, Vieira, Kounev, Avritzer, & Payne, 2015). No consensus has been reached as regards which of those criteria are the most suitable in the evaluation of IDS classifiers because the current studies depend on a single incomplete aspect (Kumar, 2014). In each study, the evaluations have been achieved with reference to some security criteria, and most studies ignored the performance. The basis used in selecting the evaluation criteria is unclear (Novaković, Veljović, Ilić, Papić, & Milica, 2017). To the best of our knowledge, no exclusive study has presented a reliable benchmarking solution of IDS classifiers based on standardised evaluation (Magán-Carrión et al., 2020).

A comprehensive evaluation and benchmarking for IDS classifiers in terms of security and performance is considered a major challenge due to multiple difficulties, which are categorised into two consecutive parts: standardisation and MCDM issues. **First**, several evaluation criteria must be considered, and the most suitable ones based on standardisation procedure due to several criteria have been proposed in the literature in both aspects (i.e. security and performance) (Milenkoski et al., 2015). **Second**, the evaluation criteria do not have the same importance; thus, properly considering the increasing significance/importance of some standardised criteria and reducing others in the evaluation of IDS classifiers (W. Guo, Chen, Cai, Wang, & Tian, 2017). Thus, assigning the proper weight for each criterion needs to be achieved. **Third**, another important issue is criteria conflict/tradeoff. According to a survey conducted by (Tavallaee, 2011), the most widely used metrics by the intrusion detection research community are true positive rate (TPR) and false positive rate (FPR) along with the receiver operating characteristic (ROC).

On the basis of the values of these metrics, determining better IDS classifiers among others is very difficult, especially when the issue of conflict/tradeoff is encountered (Ahmad et al., 2010). Another example is that precision and recall are critical factors for better performance in classifying skewed datasets, and a conflict/tradeoff exists among the precision and recall . **Fourth**, some alternatives have defeated others and could be selected as the best ones according to some criteria, whereas other alternatives can be prioritised over the previous alternatives according to different criteria (Fessi, Benabdallah, Boudriga, & Hamdi, 2014) (Yan et al., 2016) (Khan & Baig, 2010). According to the discussed issues, the evaluation and benchmarking of IDS classifiers falls under standardisation and complex MCDM problems (Hassan, Gumaei, Alsanad, Alrubaian, & Fortino, 2020). Figure 1.1 illustrates the problem configuration.

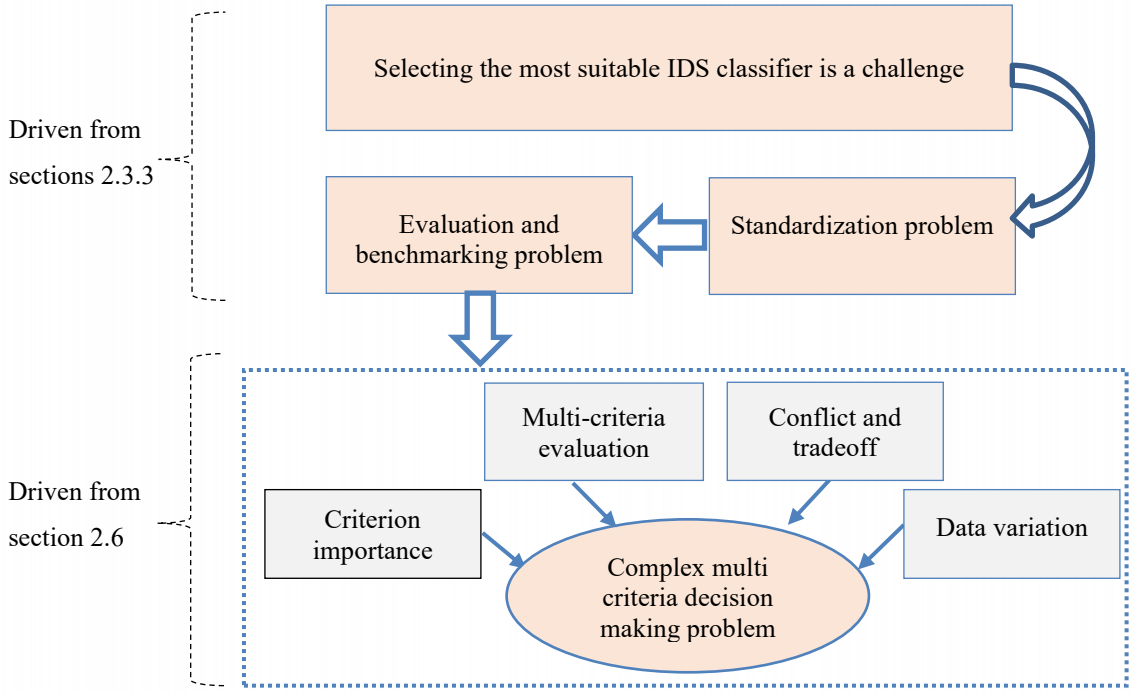

Figure 1.1. Problem configuration

1.4 Research Questions

This section presents the research questions that will be addressed in this research:

1. What are the existing evaluation criteria for IDS classifiers in terms of security and performance?
2. What are the important criteria to be used for evaluation and benchmarking IDS classifiers?
3. What are the requirements needed to construct a standard evaluation and benchmarking framework for IDS classifiers?
4. How to develop an integrated platform which includes evaluation criteria identification and criteria weighting for IDS classifiers evaluation and benchmarking?
5. To what extent are the results of the proposed framework valid?

1.5 Research Objectives

1. To investigate the existing evaluation criteria in terms of security and performance for IDS classifiers and highlight the weaknesses.
2. To determine the most important evaluation criteria in terms of security and performance aspects for IDS classifiers based on the fuzzy Delphi method (FDM).
3. To formulate a decision matrix based on crossover in-between determined evaluation criteria and IDS classifiers.
4. To develop a benchmarking framework for IDS classifier-based formulated decision matrix using MCDM methods.
5. To validate and evaluate the proposed benchmarking framework.

1.6 Relationship between Research Objectives, Research Questions and Research problem

The research questions are sketched to provide the direction and focusing of the research and the research objectives provide answers to the research questions. Table 1.1 presents the research questions, and they are answered by the research objectives. The table also determines what part of the research problem will be solved when each research objective is achieved.

Table 1.1

Link among research questions, research objectives and research problem

Research Questions	Research Objectives	Research problem mapping Specific Problem	General problem
1. What are the existing evaluation criteria for IDS classifiers in terms of security and performance?	1. To investigate the existing evaluation criteria in terms of security and performance for IDS's classifiers and highlight the weaknesses.		Evaluation and benchmarking problem
2. What are the important criteria to be used for evaluation and benchmarking IDS classifiers?	2. To determine the most important evaluation criteria in terms of security and performance groups for IDS's classifiers based on Fuzzy Delphi.	Standarization issue	
3. What are the requirements needed to construct a standard evaluation and benchmarking framework for IDS classifiers?	3. To formulate a decision matrix based on crossover in-between determined evaluation criteria and IDS classifiers.	1. Multi Evaluation criteria problems.	
4. Is there any integrated platform include evaluation criteria identification and criteria weighting for IDS classifiers evaluation and benchmarking?	4. To develop a benchmarking framework for IDS classifiers based formulated decision matrix using MCDM methods.	2. Trade off criteria and Conflicting criteria. 3. Importance of criteria. 4. Data variation	
5. To what extent are the results of the proposed framework valid?	5. To validate and evaluate the proposed benchmarking framework.		

1.7 Research Scope

This research has a cross-domain nature, thereby primarily focusing on expert systems, evaluation and benchmarking. The research is designed to solve the IDS classifier selection problem, as illustrated in Figure 1.2.

Different research methods are involved in the present study because the problem is classified as an inter-disciplinary problem. Experimental analysis is the selected research method used to select, evaluate and adopt a suitable multi-criterion scoring in IDS classification algorithms.

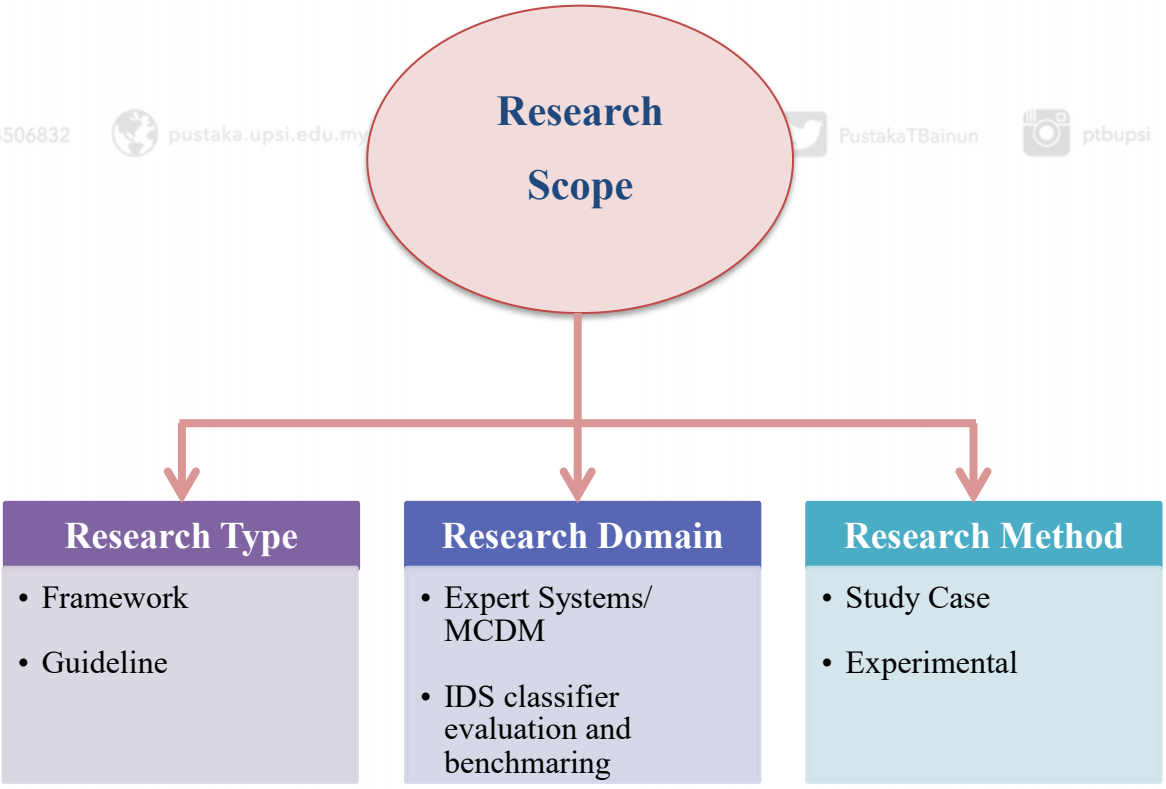

Figure 1.2. Scope of the study

1.8 Significance of the study

Developing a standardised evaluation and benchmarking framework for IDS classifiers has practical and theoretical significance:

1.8.1 Practical significance of the study

Practically, through the proposed standardisation and benchmarking framework for IDS classifiers, developers of IDS products will choose the classification model that was evaluated in a more reliable and complete perspective (i.e. performance and security) (Hassan et al., 2020). Consequently, they will be able to provide precise and reliable IDSs, resulting in more secure computer network environments and more safe and reliable networking and digital services. These changes will improve the performance of computer networks and provide customers with more confidence (Herrera-Semenets, Bustio-Martínez, Hernández-León, & van den Berg, 2021; Katkar, Shukla, Shaikh, & Dange, 2021).

1.8.2 Theoretical importance the study

This study contributes through adopting the systematic literature review approach to provide an overview of existing IDS classifiers evaluation and benchmarking approach in terms of security and performance and to highlight the trends of research on this topic. This study also contributes to filling the lack of research in this research area. The proposed taxonomy of the related literature in this study can bring several benefits as well, including imposing organisation on the mass of publications; sorting out the different studies into a

meaningful, manageable and coherent layout and providing researchers with important insights into the subject field in several ways. The importance of the proposed taxonomy lies in its outlining the potential directions of research in the field, revealing research gaps and mapping the literature on IDS and classification into distinct categories, emphasising the weak and strong features in terms of research coverage. In addition, this study provides a guide to the most important criteria that should be adopted to evaluate IDS classifiers.

1.9 Organisation of Thesis

This study is composed of six chapters. The structure of the study is illustrated in Figure 1.3.

A brief review of the six chapters is as follows.

Chapter 1: Introduction. This chapter presents the research background and research problem and proposes the research questions and objectives. It also explains the relationship between research objectives, research questions and research problems. This chapter illustrates the research scope and the significance of the study.

Chapter 2: Literature Review. A systematic review analysis for the literature that combines IDS and MCDM is presented here. Critical analysis for the reviewed articles is also presented. The requirements for IDS classifiers evaluation are proposed. The chapter ends with challenges and open issues of IDS evaluation and the recommended solutions.

Chapter 3 – Research Methodology. The research methodology is described in detail here. It consists of in five main phases: preliminary, standardisation of IDS evaluation criteria, decision matrix construction, evaluation framework development and validation

and evaluation of the proposed framework. This chapter will illustrate in detail how the five research objectives will be achieved during the various phases.

Chapter 4 – Results and Discussion, Validation and Evaluation. This chapter presents and discusses the results of the most suitable evaluation criteria set for IDS classifier evaluation. Subsequently, this chapter presents the constructed decision matrix. The results of the developed MCDM evaluation framework are presented, and how the proposed solution solves the problems outlined in the problem statements is explained. Then, the chapter presents the validation and evaluation process of the proposed framework.

Chapter 5: Conclusion and Future Work. This chapter provides the conclusion, which is followed by the highlights, the summary of research contributions, the limitations and a discussion of future work.

1.10 Chapter Summary

This chapter presents the background of the study. It describes the concept of IDS and classifiers, as well as the criteria that have been used in the evaluation process. The most vital point of this study's background is the criteria used to evaluate the IDS classifiers and the relative importance of each criterion. This discussion is followed by detailed explanations of the problem statement, the research objectives and scope and the study's significance. Lastly, the organisation of the thesis is presented.

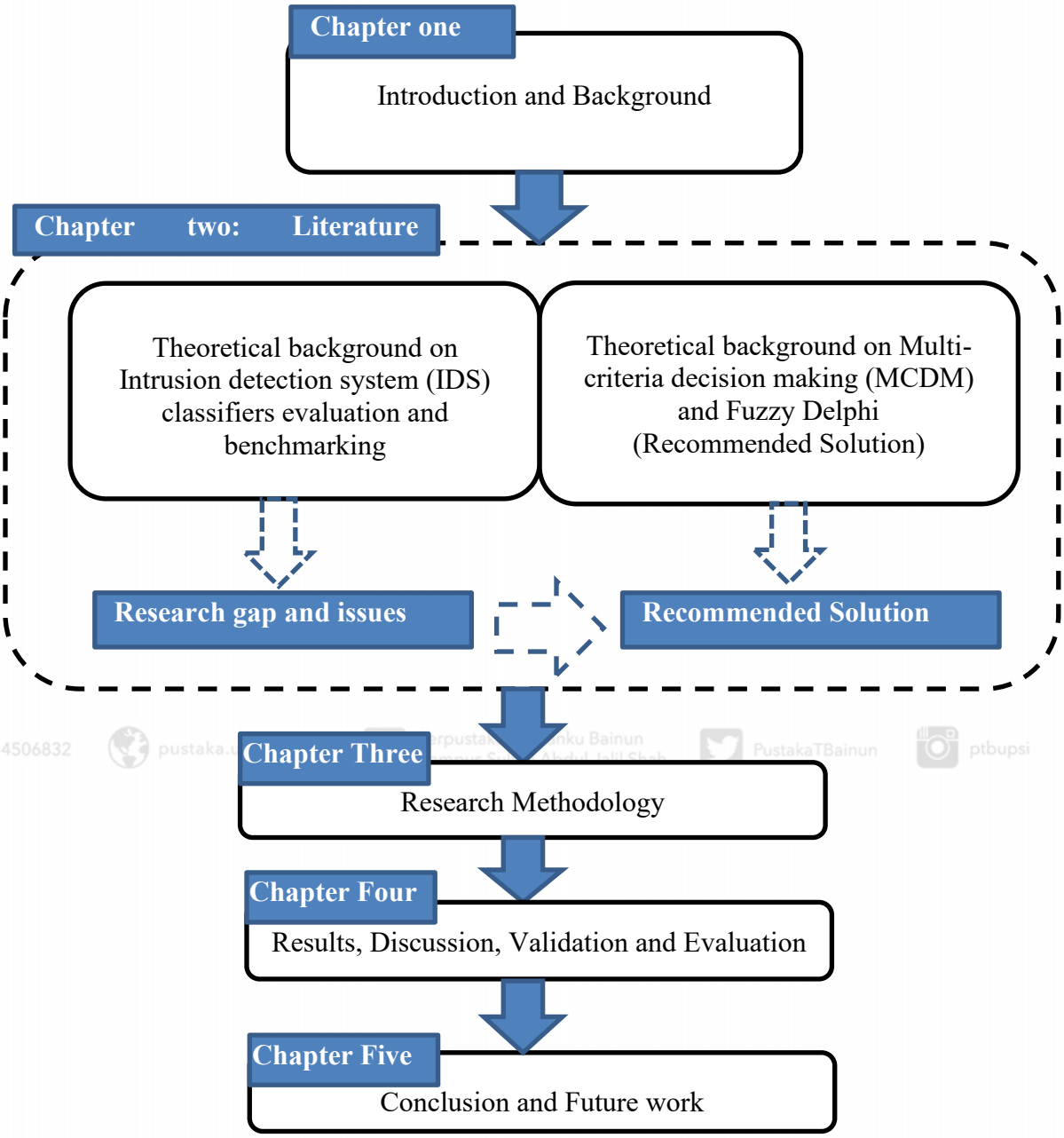

Figure 1.3. Organization of the thesis