ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DIGITAL TWIN-ENABLED INTELLIGENT ATTACK DETECTION MECHANISMS FOR AUTONOMOUS NETWORKS

M.Sc. THESIS

Yağmur YİĞİT

Department of Computer Engineering

Computer Engineering Programme

JUNE 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DIGITAL TWIN-ENABLED INTELLIGENT ATTACK DETECTION MECHANISMS FOR AUTONOMOUS NETWORKS

M.Sc. THESIS

Yağmur YİĞİT
(504201542)

Department of Computer Engineering

Computer Engineering Programme

Thesis Advisor: Prof. Dr. Berk CANBERK

JUNE 2023

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

OTONOM AĞLAR İÇİN DİJİTAL İKİZ DESTEKLİ AKILLI SALDIRI TESPİT MEKANİZMALARI

YÜKSEK LİSANS TEZİ

Yağmur YİĞİT
(504201542)

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Berk CANBERK

HAZİRAN 2023

Yağmur YİĞİT, a M.Sc. student of ITU Graduate School student ID 504201542 successfully defended the thesis entitled “DIGITAL TWIN-ENABLED INTELLIGENT ATTACK DETECTION MECHANISMS FOR AUTONOMOUS NETWORKS”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Prof. Dr. Berk CANBERK** ....
Istanbul Technical University

Jury Members : **Assoc. Prof. Nazım Kemal ÜRE** ....
Istanbul Technical University

Prof. Dr. Fatih ALAGÖZ ....
Boğaziçi University

....

Date of Submission : **16 May 2023**
Date of Defense : **19 June 2023**

To my family,

FOREWORD

As a completion of my master's degree in computer engineering at Istanbul Technical University (ITU), I've put forward this thesis. My studies have enlightened me in many areas of cyber security, diving further into digital-twin-enabled cyber-secure network solutions for next-generation networks. I would like to extend my greatest appreciation to my professor Berk Canberk for having taught me all that I've learned. His vision has changed my life. Thanks to his direction, I've been able to understand where I could better apply my knowledge, which has led me to transition from industry into academia. I'd like to give special thanks to my professors Trung Q. Duong, Sema Oktuğ, Gökhan Seçinti, and colleagues at BCRG for their support throughout my academic life. Without them, I wouldn't be who I am today.

Despite the challenges faced, I've been able to achieve great solutions for network issues in many areas. During my research on digital twin networks, connection and transition issues between layers were quickly resolved thanks to my professor's support and guidance, working together tirelessly.

I must also mention how my professor's inspiration and motivation pushed me to think outside the box. Thanks to him, some of my other accomplishments even include getting research published, including one international journal paper in the IEEE Communications Standards Magazine, International conference papers in the 17th International Conference on Distributed Computing in Sensor Systems (DCOSS), my air pollution research conference paper IEEE Global Communications Conference (GLOBECOM) Workshops 2022, and for digital twin-enabled secure port networks research in the IEEE International Conference on Communications (ICC) 2023. Moreover, my two patent applications are currently under review, and I have just submitted my two journal papers yet. Without his endless support, I couldn't have gotten it done.

I also need to extend my deepest gratitude to all my family and friends who have stood by me during my countless sleepless nights of research and given patience and psychological support.

Finally, thanks to the scholarship programmes Google DeepMind, ITU-Turkcell, and ITU-BTS Group for funding my research during my studies. This thesis was supported by the UK Royal Society International Exchange Grant IES\R3\213169.

May 2023 Yağmur YİĞİT
(Research Assistant)

TABLE OF CONTENTS

ABBREVIATIONS

DT	:	Digital Twin
DDoS	:	Distributed Denial of Service
ISP	:	Internet Service Provider
YANG	:	Yet Another Next Generation
AutoFS	:	Automated Feature Selection
AutoCM	:	Automated Classification Method
TwinCoNet	:	DT-enabled detection system for autonomous ISP core networks
TwinPot	:	DT-assisted honeypot
AI	:	Artificial Intelligence
IoT	:	Internet of Things
IAPH	:	International Association of Ports and Harbors
CPS	:	Cyber-Physical Systems
IETF	:	Internet Engineering Task Force
ML	:	Machine Learning
MLP	:	Multilayer Perceptrons
SDN	:	Software-Defined Networking
EM	:	Expectation-Maximization
DeT	:	Decision Trees
NB	:	Naive Bayes
SVM	:	Support Vector Machines
RF	:	Random Forests
FS	:	Feature Selection
NETCONF	:	Network Configuration Protocol
RPCs	:	Remote Procedure Calls
NETMOD	:	NETCONF Data Modeling Language
KPIs	:	Key Performance Indicators
RFE	:	Recursive Feature Elimination
BFE	:	Backward Feature Elimination
ANOVA	:	Analysis of Variance
ReLU	:	Rectified Linear Unit
ADT	:	Microsoft Azure Digital Twin
DTDL	:	Digital Twin Definition Language
PS	:	Proposed Solution
SMOTE	:	Synthetic Minority Oversampling Technique
D1 or D2	:	Dataset 1 or Dataset 2
DNN	:	Deep Neural Network
LSTM	:	Long Short Term Memory
KNN	:	K-Nearest Neighbors
RNN	:	Recurrent Neural Networks
LR	:	Logistic Regression
CNN	:	Convolutional Neural Network
MANSIM	:	Maneuvering Simulation Laboratory

SYMBOLS

ε	: Threshold of the detection time
$\varnothing$	: Selected method
D_{perf}	: Detection performance
c_{attack}	: The number of DDoS attacks detected
C_{max}	: The maximum average number of DDoS attacks
Sys_{pre}	: System precision
SV	: Sensitivity
F_N **or** FN	: False negative
T_P **or** TP	: True positive
F_P *or* FP	: False positive
F_{meas}	: F-measure
λ_i	: The weighted sum of precision and recall of i_{th} method
ν_i	: The determination time of i_{th} classification method
t_{end}	: The finishing time
t_{start}	: The starting time
γ	: The reliability of the classification method

LIST OF TABLES

Page

LIST OF FIGURES

DIGITAL TWIN-ENABLED INTELLIGENT ATTACK DETECTION MECHANISMS FOR AUTONOMOUS NETWORKS

SUMMARY

In parallel with the increasing number of internet users and technological advancements, there has been a corresponding surge in cyberattacks targeted at internet services. Consequently, safeguarding services from cyber threats has become crucial for organizations. Digital-twin (DT) technology has gained considerable attention due to its numerous advantages in various industries, including real-time monitoring and control in manufacturing, risk assessment in industry, and predictive maintenance in the aerospace sector. It is expected to have a significant influence in developing "self-X" capabilities and "zero-touch" operations and maintenance in 6G networks. DT allows for performance testing, real-time network monitoring, quick simulation, and optimization, which maximize its potential in the network domain. Furthermore, it enhances the cost-efficiency of evaluation, prediction, and optimization processes compared to physical systems. However, despite its many benefits, the DT potential is yet to be widely explored for network anomaly detection.

Over the last few years, the importance of next-generation networks has received increasing recognition due to the growing demand for efficiency and the large volume of data. With the advent of intelligent infrastructure and facilities, the authorities have recently focused on cyber-security, making it a primary concern. While traditional security solutions can help safeguard the systems from harmful entities, they do not provide enough transparency for security researchers to learn about attackers' behaviors. Distributed denial of service (DDoS) attack solutions currently cannot manage vast amounts of aggregated data rates. Therefore, they are unsuited to the core networks of Internet service providers (ISPs). In the same way, the current intrusion detection mechanisms are insufficient to detect external attacks, making it challenging to handle them effectively for seaport networks.

This thesis proposes two intelligent detection mechanisms for ISP core and seaport networks to solve attack detection problems and provide autonomous network characteristics. The first mechanism is TwinCoNet, a DT-enabled detection system for autonomous ISP core networks that uses online learning and the YANG paradigm. TwinCoNet is designed to handle highly aggregated data rates and uses an Automated Feature Selection module (AutoFS) to identify the most suitable features for each router, enabling independent operation for each router. The second mechanism is TwinPot, a DT-assisted honeypot designed to handle external attacks in smart seaports. The proposed intelligent attack detection mechanism uses DT technology for internal attacks and the Automated Classification Method (AutoCM) to categorize various forms of attacks.

The performance of both mechanisms is tested using extensive datasets. Results from the first mechanism reveal that the proposed system effectively identifies the attacks, adjusts the feature selection technique, and estimates the attack within fifteen minutes of it commencing, with an accurate categorization rate of ninety-seven percent. On the other hand, the second mechanism findings demonstrate that our approach successfully detects simultaneous internal and external attacks on the system and changes the classification technique. Overall, this study highlights and demonstrates the potential of DT technology in enhancing cyber-security solutions for critical infrastructures such as ISP core networks and seaports.

OTONOM AĞLAR İÇİN DİJİTAL İKİZ DESTEKLİ AKILLI SALDIRI TESPİT MEKANİZMALARI

ÖZET

Son yıllarda internet erişimi daha kolay ve kullanışlı hale gelmesi ve hızlır bir şekilde gelişen teknolojik ilerlemelerle birlikte, internet kullanıcılarının sayısı da önemli ölçüde artmıştır. Bu trendin, giyilebilir teknolojiler, akıllı evler ve akıllı limanlar gibi daha fazla cihazın internete bağlanmasıyla devam etmesi beklenmektedir. Buna paralel olarak, internet hizmetlerine yönelik siber saldırıların sayısı da giderek artmaktadır. Bu nedenle, siber tehditlerden ilgili hizmetleri korumak, kuruluşlar için oldukça büyük bir önem teşkil etmektedir. Bugünlerde organizasyonlar ya da işletmeler verimliliği artırmak için oldukça fazla çaba sarf etmektedirler. Bu çaba, olumsuz çevresel etkileri azaltmak için endüstri genelinde artan bir istekle birlikte gelmektedir. Bu nedenle, bu ihtiyacı karşılamak için otonom ağlara olan talep günden güne artmaktadır. Otonom ağlarda, akıllı saldırı tespit mekanizmaları saldırıların tespiti ve sorun oluşması durumunda hızlı bir şekilde yedek planlar oluşturup uygulama yeteneğinin sağlanabilmesi için oldukça önemlidir. Ancak, mevcut çalışmaların 5G veya 6G ağlarında kullanılan yapay zeka modellerinin potansiyel saldırı vektörlerinin değerlendirilmesi konusunda sınırlamaları olduğundan dolayı otonom ağlar için akıllı saldırı tespiti mekanizmalarının analiz yetenekleri araştırma ve endüstri topluluğu için önemlidir.

Dijital ikiz (DT) teknolojisi, imalat sektöründe gerçek zamanlı izleme ve kontrol, endüstride risk değerlendirmesi ve havacılık sektöründe tahmine dayalı bakım gibi çeşitli sektörlerde birçok avantajı sayesin son zamanlarda oldukça dikkat çekmektedir ve bu alandaki çalışmalarda hız kazanmaktadır. DT teknolojisinin, 6G ağlarında "self-X" yeteneklerinin ve "zero-touch" işlemlerinin ve bakımının geliştirilmesinde önemli bir rol oynaması beklenmektedir. DT, ağ izleme, performans testi, optimizasyon ve hızlı simülasyon için gerçek zamanlı bir ortam sağlar ve bu sayede ağ alanında potansiyelini en üst düzeye çıkarır. Ayrıca, fiziksel sistemlere kıyasla değerlendirme, tahmin ve optimizasyon süreçlerinin maliyet etkinliğini arttırır. Bahsedilenler gibi birçok faydasına rağmen, DT'nin ağ anormalliklerini tespit etmek için potansiyeli henüz geniş çapta araştırılmamıştır.

Son yıllarda, işletmelerin verimliliklerini artırmak ve büyük veri hacimlerini yönetmek için yeni nesil ağlara olan ihtiyaç giderek artmaktadır. Ayrıca, akıllı altyapı ve tesislerin gelişmesiyle birlikte, son zamanlarda siber güvenlik önemli bir konu haline gelmiştir. Bu artan talepler, geleneksel güvenlik çözümlerinin yetersiz kaldığı birçok sorunu da beraberinde getirmiştir. Örneğin, geleneksel güvenlik çözümleri sistemleri zararlı unsurlardan koruyabilirken, güvenlik araştırmacılarının saldırganların davranışlarını öğrenmeleri için yeterli şeffaflığı sağlayamamaktadır.

İnternet Servis Sağlayıcılarının (ISP) en önemli önceliği, içerik sağlayıcıları ve son kullanıcılar arasında yüksek hızda, kayıpsız bağlantı ve kesintisiz erişim sağlamaktır. ISP'ler ağlarında büyük miktarda trafiği yönetirler ve dağıtılmış hizmet reddi (DDoS) saldırıları, kritik ağ altyapılarına ve hizmetlerine yöneliktir. DDoS saldırısı, birçok kompromize edilmiş makinenin hedef sunucuyu hedef alarak, erişmeye çalışan meşru kullanıcıların erişimini engelleyerek gerçekleşen bir güvenlik açığı biçimidir. DDoS saldırılarının dağıtık yapısı, bunları izlemeyi veya karşı etkiyi zorlaştırır. Bu saldırılar çeşitli şekillerde gelir ve farklı desenler benimseyebilir, bu da onları tespit etmeyi zorlaştırır. Ek olarak, ağ güvenlik açıklarını kullanarak yazılım hizmetleri için talepler oluşturabilirler, bu da saldırıların önlenmesini daha da zorlaştırır. Gerçek zamanlı tespit ve önlem zorluğuna rağmen, DDoS saldırılarının tespiti önemlidir çünkü ciddi sonuçları olabilir. Mevcut DDoS çözümleri veri merkezleri veya kenar ağları ile sınırlıdır, yani sadece kenar yönlendiricilerde veya veri merkezinde çalışırlar ve diğer yönlendiricileri korumazlar. Bir ISS, DDoS çözümünü kenar ağı cihazlarının çoğunda etkinleştirse bile, hala birkaç makine risk altındadır. Sonuç olarak, veri merkezi ve kenar ağı çözümleri tüm ağ boyunca DDoS saldırılarını tespit edip önleyemediğinden yetersizdir. Mevcut DDoS çözümleri yüksek miktardaki toplam veri akışlarını yönetme konusunda yetersiz olduğundan ISP çekirdek ağlarındaki saldırı tespiti için uygun değillerdir. Bu nedenle, ISP'lerin kendi ağlarında DDoS saldırılarına karşı korunmaları için daha güçlü ve etkili bir çözüm geliştirilmesi gerekmektedir. Bu çözümler, ağın her seviyesindeki yönlendiricileri korumalıdır. ISP'ler, müşterilerine daha iyi hizmet vermek ve ağlarını DDoS saldırılarından korumak için yenilikçi yeni nesil teknolojilerin ağlarında kullanımını yaygınlaştırmalıdırlar.

Benzer şekilde, son yıllarda verimlilik talebinin artması ve giderek artan yük hacmi gibi zorluklarla karşı karşıya kalan deniz limanlarında, gelecek nesil liman fikri daha belirgin hale geldi. Ancak bu evrim, yeni nesil teknolojilerin daha da geliştirilmesiyle hız kazandı. Gelecek nesil limanlar, akıllı çözümleri deniz limanı ortamlarında kullanmaya odaklanan akıllı limanlar olarak da adlandırılmaktadır. Akıllı limanlar, üretkenliği ve güvenliği arttırmak ve giderleri azaltmak için yenilikçi teknolojileri kullanmaktadır. Bugünün akıllı limanları, bağlantı için kablolu ağları kullanırken, gelecekteki akıllı limanların kablosuz teknolojileri kullanması ön görülmektedir. Bu sayede, veri toplama ve iletimi daha hızlı ve verimli hale gelirken lojistik süreçler daha da optimize edilebilecek ve güvenlik önlemleri daha da artırılacaktır. Deniz limanları, hem mevcut cihazlarından hem de harici cihazlardan aynı anda saldırıya uğrayabilirler. Mevcut saldırı algılama mekanizmaları dış saldırıları tespit etmek için yetersizdir, bu da mevcut çözümlerin liman ağlarını etkili bir şekilde ele almalarını zorlaştırmaktadır. Liman ağlarında akıllı saldırı tespitinin önemi, siber saldırıları hızlı bir şekilde tespit etmek ve azaltmak için olduğu kadar iş sürekliliğini sağlamak için de büyük bir öneme sahiptir. Bu nedenle, deniz limanlarındaki siber güvenlik açıklarını kapatmak için yeni nesil teknolojilerin kullanımı önem kazanmaktadır. Akıllı saldırı tespiti mekanizmaları, verimli ve öz-yönetimli ağlar sağlanarak otonom ağ hedeflerine ulaşılmasına yardımcı olabilmektedirler. Saldırıların hızla tespit edilmesi ve müdahale edilmesi, bir limanın etkinliğini ve verimliliğini artırabilir, operasyonel kesintileri önleyebilir ve hatta potansiyel bir güvenlik riskini ortadan kaldırabilir. Bu nedenle, akıllı saldırı tespit sistemleri, liman ağlarının güvenliği için önemli bir konudur. Bu sistemler en son teknolojik geliştirmeleri kullanarak, saldırıları hızlı bir şekilde tespit edip, saldırıların nedenlerini analiz edebilir ve hatta öngörebilirler. Bu sayede, liman ağlarında oluşabilecek saldırılar etkili bir şekilde engellenebilir.

Bu çalışma, saldırı tespit sorunlarını çözmek ve otonom ağ özellikleri sağlamak için ISP çekirdek ve liman ağlarında kullanılmak üzere iki akıllı tespit mekanizması önermektedir. İlk mekanizma, otonom ISP çekirdek ağları için DT destekli bir tespit sistemi olan TwinCoNet'tir. Bu sistem, online öğrenme ve YANG paradigmasını kullanmaktadır ve yüksek düzeyde toplanan verileri hızlı bir şekilde yönetmek için tasarlanmıştır. Ayrıca, her yönlendirici için en uygun özellikleri belirlemek için Otomatik Özellik Seçimi modülü (AutoFS) kullanmaktadır. Böylece her yönlendiricinin bağımsız çalışmasını sağlamaktadır. İkinci mekanizma, akıllı limanlarda dış saldırıları yönetmek için tasarlanan bir DT destekli bal küpü tuzağı (honeypot) olan TwinPot'tur. Bu sistem, iç saldırılar için DT teknolojisini kullanarak farklı saldırı türlerini sınıflandırmak için Otomatik Sınıflandırma Yöntemini (AutoCM) kullanmaktadır.

Ayrıca bu tez çalışmasında, önerilen her iki mekanizmanın performansı geniş veri kümeleri kullanılarak test edilmiştir. İlk mekanizma sonuçları, önerilen çözümün DDoS saldırılarını etkili bir şekilde tanımladığını, özellik seçim yönteminin ilgili durumlarda başarı ile güncellendiğini ve gerçek sınıflandırma oranı yüzde doksan yedi olarak DDoS saldırısı başladıktan yaklaşık on beş dakika sonra saldırıyı tahmin ettiğini göstermektedir. Diğer yandan, ikinci mekanizma bulguları, yaklaşık olarak eşzamanlı gerçekleşen iç ve dış saldırıların başarıyla tespit edildiğini ve ilgili durumlarda sınıflandırma yönteminin başarıyla değiştirildiğini göstermektedir.

Bu tez çalışması, DT teknolojisinin ISP çekirdek ağları ve liman ağları gibi kritik altyapılarda siber güvenlik çözümleri geliştirilmesindeki potansiyelini ve başarısını ortaya koymaktadır. Bu teknoloji sayesinde, ciddi saldırılara karşı koruma sağlayan yenilikçi çözümler üretmek mümkün olacaktır. Bu çözümler, kritik altyapılarda daha güvenli bir siber ortam yaratılmasına yardımcı olacak ve siber saldırılardan kaynaklanan potansiyel zararları minimize edecektir. Ayrıca, bu çalışma, diğer sektörlerde de DT teknolojisinin kullanım potansiyelini göstermektedir ve bu teknolojinin siber güvenlik alanında önemli bir rol oynayabileceğini vurgulamaktadır.

1. INTRODUCTION

As the Internet has become more accessible and convenient, the amount of Internet users has significantly increased. According to recent statistics, global internet users surpassed 5.3 million in 2022 [2]. Fig. 1.1 shows the number of internet users worldwide. This trend is expected to continue as more devices such as wearables, smart homes, and smart seaports become internet-enabled. However, as internet users and gadgets increase, the cyberattack count aimed at online services also rises. It has become a pressing issue for organizations to protect their services from cyber threats.

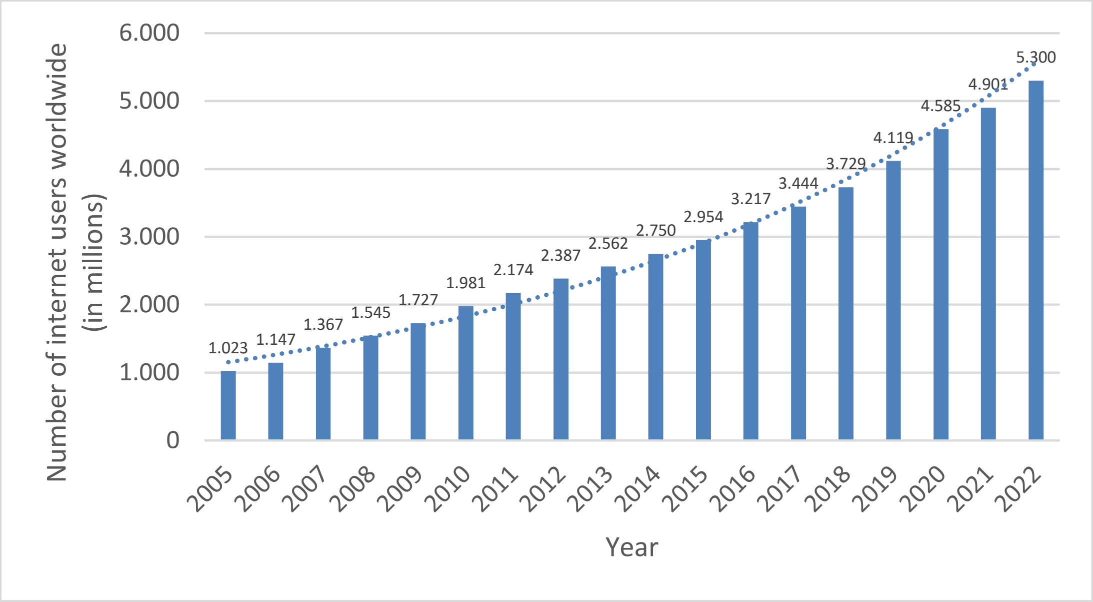


Figure 1.1 : The global number of internet users 2005-2022 [2].

The sectors, such as seaport transportation, are rapidly digitizing themself in line with fast-growing technologies like Artificial Intelligence (AI) and the Internet of Things (IoT). Nowadays, businesses are striving more than ever to enhance productivity. This drive is accompanied by a growing industry-wide desire to reduce negative environmental impacts. Consequently, the demand for autonomous networks is also rising to address this need.

In autonomous networks, the attack detection mechanisms are crucial to ensure their resilience and ability to implement contingency plans in the event of problems. However, existing research on the use of AI algorithms in 5G or 6G networks is limited in assessing potential attack vectors due to the limited use cases considered. It is important for the academic and industry sectors to examine the use of AI models more thoroughly regarding attack detection for autonomous networks [6]. The cost of a single DDoS attack can exceed 1.6 million U.S. dollars, and the average cost of cybercrime was 13.0 million U.S. dollars in 2018, with DDoS attacks costing 1,721,285 U.S. dollars [7]. Moreover, the global cybersecurity market is expected to be valued at 403 billion U.S. dollars by 2027, worth 176.5 billion U.S. dollars in 2020 [8]. These statistics highlight the importance of effective attack detection in autonomous networks and the need for robust cybersecurity measures.

The concept of digital twin (DT) has gained immense popularity owing to its advantages across various domains, such as risk analysis in manufacturing, industrial real-time remote monitoring and control, aerospace predictive maintenance, and more. The DT technology is expected to be crucial in developing “zero-touch” operations, self-optimization, self-configuration, and self-management capabilities of 6G networks [9] [10]. DT can be effectively utilized in real-time or near-real-time network monitoring, optimization, performance testing, and rapid simulation. Additionally, DT offers a cost-effective way to evaluate, predict, and optimize processes compared to physical systems. However, despite all of DT’s advantages, its application for network anomaly detection remains insufficiently understood [11] [12].

The importance of intelligent attack detection regarding Internet service providers (ISP) and seaport networks is explained as follows:

Internet Service Providers Networks: Ensuring high-speed, lossless connectivity and availability between content providers and end-users is the top priority for Internet Service Providers (ISPs), and they handle a massive amount of traffic on their networks. However, the critical network infrastructure and services of ISPs are vulnerable to distributed denial of service (DDoS) attacks. A DDoS attack is a security breach involving several infected machines targeting a victim server, rendering it inaccessible to genuine users attempting to gain access to it. The distributed nature of DDoS attacks makes them challenging to trace or counteract them. These attacks come in various forms and can adopt different patterns, making them challenging to identify. Additionally, they can exploit network vulnerabilities and generate requests for software services, making them even more challenging to mitigate. Despite the difficulty in real-time detection and mitigation, it is crucial to detect DDoS attacks as they can have severe consequences. The DDoS protection offered by the market's current DDoS solutions is limited to edge networks or data centers, which means that other routers are not protected because they only function in their restricted area. Only a few of the ISP's edge network devices remain vulnerable even if the DDoS solution is enabled on most of them. Consequently, these solutions are inadequate as they cannot detect and prevent DDoS attacks throughout the entire network.

DDoS attacks significantly spiked in 2020, resulting from pandemic-related circumstances, and a twenty-two percent rise compared to pre-pandemic years [13] [14]. According to the DDoS Attack Report of Kaspersky, there was an increase in the share of attacks lasting between 5 to 139 hours in the last quarter of 2020, as illustrated in Fig. 1.2. Moreover, the duration of the longest DDoS attack exceeded 302 hours, and the average attack duration was less than 240 minutes [3]. As of 2020, ISP servers' average hourly downtime costs range from 301,000 to 400,000 dollars [15]. As can be seen from these figures, these attacks cause significant financial losses. In 2020, several ISPs, including FDN, K-net, Bouygues Télécom, and SFR in France, EDP in Belgium, and Signet, Delta, Online.nl, FreedomNet, Tweak.nl, and Caiway in the Netherlands, suffered from DDoS attacks [16]. In May 2021, Irish internet service providers also suffered from DDoS attacks [17]. Although these attacks lasted for less than a day, they caused disruption in the services of the ISPs.

As can be understood from the aforementioned, time is a significantly essential metric in attack detection, and intelligent attack detection solutions are necessary for autonomous ISP networks.

Seaport Networks: Seaports and their facilities are vital in the maritime industry and national economies since they handle a significant portion of the goods transported

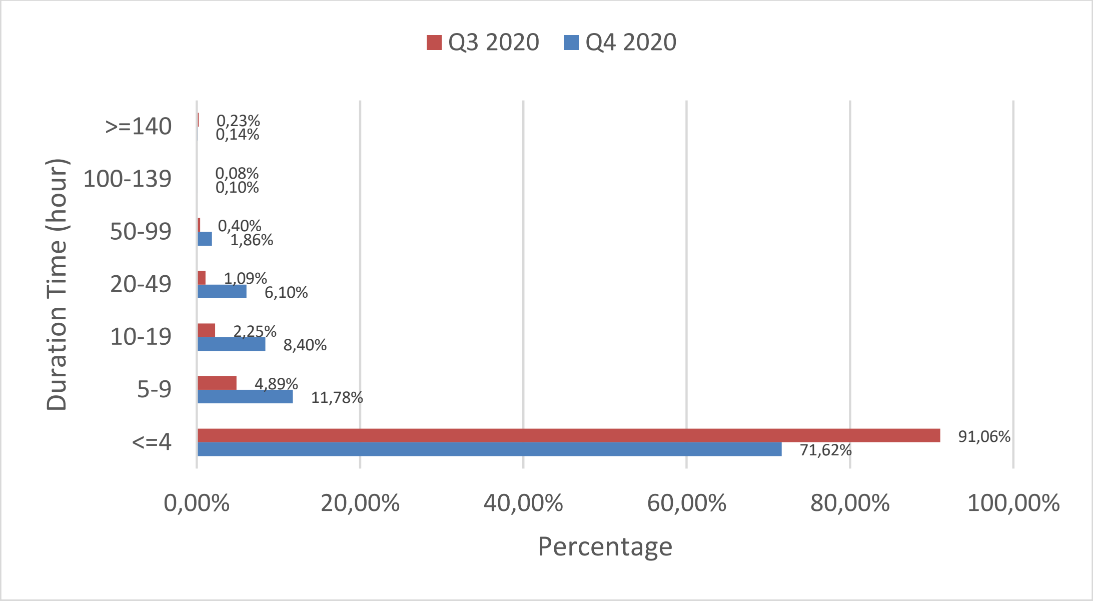


Figure 1.2 : Distribution of DDoS attacks by duration [3].

worldwide [18]. Digitalization is thus becoming more and more crucial to ensure operational effectiveness and quantifiable cost reductions at seaports. In the past decade, there has been a growing need to increase efficiency and handle the increasing volume of goods in seaports. This has resulted in the emergence of smart ports, also known as next-generation ports, that utilize intelligent solutions in port operations. Markets and Markets' Tech Report predicts that the smart port market will grow rapidly at a rate of 24.3%, reaching a total worth of $5.7 billion by 2027 from an estimated $1.9 billion in 2022 [19].

The increased importance of intelligent infrastructure and facilities has brought cybersecurity to the forefront of seaport and maritime authorities' concerns. It is a top priority for most ports to ensure the security of their operations. From February 2020 to May 2020, the International Association of Ports and Harbors (IAPH) reported a fourfold increase in cyberattacks against the marine sector worldwide. Over a more extended period, attacks against operational systems increased by an astonishing 900% between 2017 and 2020 [20]. For instance, two container terminals' operations at the Port of Rotterdam were rendered inoperable by a ransomware attack in June 2017, despite this port having fully automated all of its activities in accordance with the smart port concept [21]. In May 2020, a cyberattack on Iran's Shahid Rajaee Port took down the infrastructure, clogged up the roads, and stopped cargo operations, forcing manual work and reducing efficiency [20]. This highlights the importance of intelligent attack detection in seaport networks, which can quickly detect and mitigate cyber-attacks to reduce downtime and ensure operational continuity. Additionally, intelligent attack detection can help achieve the goals of an autonomous network by enabling efficient and self-managing networks.

1.1 Motivation of Thesis

The increasing dependence on digital technologies has led to a rise in cyber threats that pose significant risks to critical infrastructure, including ISP and seaport networks. The need for intelligent attack detection mechanisms in these domains is paramount to prevent cyber-attacks, mitigate their effects, and protect sensitive data. This section reviews the need for intelligent attack detection in ISP and seaport networks, highlighting the challenges and limitations of traditional security measures.

1.1.1 The need of intelligent attack detection for Internet service providers networks

The detection of DDoS attacks is crucial for ISPs to protect their networks and infrastructure. However, due to the high bandwidth of their core network, it is challenging to detect attacks in real time. As a result, detecting DDoS in an ISP's core network is unessential to them. Traditional offline learning approaches are unsuitable for the core network's big volume and diversified data. Thus, online learning models are needed to ensure stable performance. Additionally, DT's synchronized monitoring and management abilities allow for real-time monitoring and control of the core network. Contrary to conventional DDoS monitoring services, DT-assisted core networks may enable services directly connecting to network operational data instead of updated data. Additionally, network automation using Yet Another Next Generation (YANG) model can minimize the core network complexity and increase efficiency. By selecting only the necessary characteristics, the YANG model can simplify the core network and enable it to operate more efficiently. Overall, the need for intelligent attack detection is crucial for ISPs to effectively manage their networks and protect their infrastructure and services from DDoS attacks.

1.1.2 The need of smart attack detection for seaport networks

Traditional solutions such as firewalls and antivirus software are insufficient to enhance the security of IoT and Cyber-Physical Systems (CPS). Transparent security solutions are needed for security researchers to learn about the behavior of attackers. Honeypots can serve as potential solutions as they provide valuable insights into attackers' behavior. Honeypots are security tools that set virtual traps to attract and potentially compromise attackers [22]. They can be either physical or virtual, but virtual honeypots require higher fidelity to be more convincing to attackers. To increase the complexity and simulation fidelity of honeypots, DT technology can be utilized, making them more attractive to attackers. Seaports are vulnerable to attacks from both existing and external devices. Current intrusion detection mechanisms are inadequate to detect external attacks, making it essential to have more advanced systems capable of handling attacks at the desired level. Therefore, it is crucial to have a comprehensive security strategy that can detect and mitigate attacks from all sources. Intelligent attack detection mechanisms can play a vital role in this strategy by continuously monitoring the network for potential threats and providing real-time alerts to the related personnel.

1.2 Contributions of Thesis

In this study, we propose two intelligent detection mechanisms for ISP core and seaport networks to solve attack detection problems and provide one of the autonomous network characteristics. Additionally, our suggested solutions are consistent with the digital twin network paradigm of the Internet Engineering Task Force (IETF) [23].

1.2.1 Internet service providers core networks

We propose TwinCoNet, a DT-enabled detection system that uses online learning and the YANG paradigm for autonomous ISP core networks. A YANG model imports into the system only those features that are wanted. Thanks to its learning abilities, DT replicates the network entities and offers a smart detection technique.

The classification technique selection is essential for attack detection. We analyzed several machine learning (ML) methods taking into account the training and testing intervals. We discovered that the TwinCoNet architecture performs better when multilayer perceptrons (MLP) are used. In order to filter out and identify the most suitable features, we recommend an Automated Feature Selection (AutoFS) module.

TwinCoNet's main contributions can be summarised as follows:

- We introduced a smart detection mechanism called TwinCoNet that utilizes DT for autonomous ISP core networks.
- An online learning technique is implemented to ensure stable performance by continuously updating the model with sequential data and making predictions for future data at each step. Additionally, we suggest a labeling algorithm for labeling unlabeled data.
- By implementing the AutoFS module and YANG model, we minimize the core network complexity and identify the salient features for each router, enabling our system to operate independently for each router.

1.2.2 Smart seaport networks

The seaports are vulnerable to attacks from both their internal and external devices. The current intrusion detection systems cannot detect external attacks, and therefore, the security systems are not as effective as desired. To address this issue, we propose a DT-assisted honeypot named TwinPot, specifically designed to handle external attacks in smart seaports. TwinPot provides a more comprehensive analysis of attacks and contributes to enhancing the cybersecurity mechanisms of smart seaports. Additionally, we propose an intelligent attack detection mechanism to handle various attack types using DT technology in the current seaport entities. Our proposed network-aware intelligent attack detection mechanism employs an Automated Classification Method (AutoCM) to identify and classify different attack types.

The main contributions of the TwinPot architecture are as follows:

- We introduced TwinPot, which combines DT and honeypot technologies to analyze the behaviors of attackers in external entity attacks on smart seaports and enhance their security mechanisms.
- We suggest an intelligent attack detection mechanism that is network-aware and utilizes DT technology for smart seaports to address various attack types. Additionally, we implement the Automated Classification Method (AutoCM) to handle different types of attacks.

2. LITERATURE SURVEY

This section provides a thorough analysis of related works in several areas, including attack detection, DT, honeypot, and maritime technologies. This review aims to highlight the existing works in these areas and identify the research gaps that need to be addressed.

2.1 Current Attack Detection Solutions

Even though there have been many studies conducted on detecting attacks, most of these studies have concentrated on data center solutions and offline learning strategies.

Table 2.1 : The summary of current works in different domains [1].

Work	Method	Performance (%)	Dataset	Domain
[24]	Entropy Algorithm	80 (Detection Rate)	Produced Dataset	SDN
[25]	DeT, NB, RF, SVM, EM	74.1-99.4 (F1-Score)	CICDDoS2017	IoT
[26]	Isolation Forest	96.01 (F1-Score)	NSL-KDD	Fog Computing
[27]	Bagging, Boosting, Stacking	92.2-93.4 (Accuracy)	CICDDoS2019	Smart Grid
[28]	MLP	98.18 (F1-Score)	CICDDoS2019	Intrusion Detection System

In Table 2.1, we have summarized various DDoS detection approaches used in different fields. To secure communication endpoints, the StateSec, presented by J. Boite *et al.*, DDoS attack recognition and mitigation technique based on state-driven Software-Defined Networking (SDN) [24]. According to simulation results, this method is more effective than sFlow regarding control layer utilization. Z. K. Maseer *et al.* examined a number of ML methods, including expectation-maximization (EM), decision trees (DeT), naive Bayes (NB), support vector machines (SVM), and random forests (RF) [25]. In terms of overall accuracy and runtime performance, experimental findings indicated that DeT and RF models outperform the competition. Some researchers have utilized an autoencoder, which is a neural network with multiple layers, for the purpose of feature selection (FS) in detecting DDoS attacks. After performing FS using the autoencoder, a classification algorithm is used for the actual detection of the attacks [26] [28]. T. T. Khoei *et al.* ran a study to compare the performance of three ensemble-learning methods detecting anomalies in a smart grid. The study found that the stacking-based learning approach performed better than the other two techniques [27].

We investigated the relationship between the number of parameters used in DDoS detection and its accuracy by analyzing existing studies that utilized machine learning methods. Table 2.2 presents some of the studies that have investigated to determine the correlation between the number of parameters employed for DDoS detection and the accuracy achieved. However, upon scrutinizing these studies, we could not establish a direct relationship between the number of parameters and the detection accuracy. Based on our analysis of previous works, we assumed that ten parameters are adequate to detect DDoS attacks with high accuracy. Therefore, we opted to use ten parameters in this study as part of the FS process.

After analyzing the studies, we apprehend that parameter selection should make after examining the effect of parameter number on accuracy. Hence, some studies have employed feature selection methods for attack detection [38] [39]. A semi-supervised weighted k-means detection technique utilizing a hybrid FS approach was presented by Y. Gu *et al.* [40]. The majority voting methodology, which has a significant computational expense, combines the characteristics of seven feature selection techniques into an ensemble framework that is suggested in another work [41]. A flow grouping method was also proposed in one study, based on the shared behaviors of virtual machines to detect DDoS attacks in a data center [42]. Additionally, another method was developed that uses a system made up of flow monitoring and traffic filtering to counter DDoS attacks coming from a data center [43]. S. Garg *et al.* suggested a hybrid processing technique to identify network anomalies in cloud data

Table 2.2 : The existing attack detection works.

Work	Method	Number of Parameters	Accuracy (%)
[29]	one class K-NN classifier	32	83
[24]	entropy based algorithm	4	80
[30]	artificial neural network	5	82,1 - 95,6
[31]	multivariate data analysis	2	97,53 - 98,6
[32]	naive bayes, decision tree, and logistic regression	6	99,3
[33]	SVM	3	99,53
[34]	SVM and multilayer perceptron	2	98,42 - 99,39
[35]	isolation forest and one class SVM	4	93,33
[36]	artificial neural network	5	98
[37]	decision tree and naive bayes	11	98 - 99

center networks [44]. In software-defined sensor networks, we introduced an AutoML framework to identify network anomalies in a network-aware manner [45].

2.2 Current Digital Twin and Honeypot Works

Only a limited number of studies have employed DT for anomaly detection, with most of them concentrating on constructing the DT system instead of detecting anomalies. A study by A. Saad *et al.* introduced an IoT-based DT system for microgrids to improve their resistance to attacks [46]. They gave the DT's mathematical concept and execution. The outcomes demonstrated that their structure effectively minimizes the attacks. In addition, numerous studies have presented techniques to assign labels to unlabelled data. One such method introduced a modified version of the label propagation [47]. The findings showed that the proposed technique significantly increased fault classification accuracy. In our previous work, we performed an experiment to assess how well the YANG model worked between the DT and service layers in the context of smart cities' air monitoring [4]. Our objective was to enhance the data collection performance. Another study presented a DT-based structure that handles synchronization and communication issues with IPv6 architecture, resulting in more accurate net-zero waste services [48].

Despite their potential to detect malicious activities and collect attack samples, honeypots are not widely utilized in IoT environments. J. Franco *et al.* provided a comprehensive evaluation and classification of existing honeypot research in their study [22]. They also highlighted the open research challenges in this area, suggesting that there is a need for more honeypot research to protect innovative environments like ports and buildings from attacks. W. Zhang *et al.* presented a hybrid honeynet that is capable of detecting and capturing malicious activities and samples from networks [49]. The honeynet control center can send commands and files to any physical node within the honeynet. The proposed approach was implemented and proved to be successful in preventing multiple unidentified malicious attacks. A honeypot framework for the IoT that includes high and low interaction components was introduced by B. Wang *et al.* [50]. By analyzing the malware binaries captured by the proposed framework, they discovered several types of malware that were controlled by at least eleven attack groups.

2.3 Current Maritime Works

The maritime industry places significant importance on security. F. Akpan *et al.* highlighted the significance of security issues in the maritime industry and discussed [51]. They identified potential vulnerabilities in ship systems and suggested possible solutions to address the challenges posed by security threats. Another study emphasized the significance of safeguarding vital infrastructures while encouraging emerging new technologies [52]. This study pointed out that the increased connectivity in smart ports leads to the development of a distinctive ecosystem and security risks. In a thorough literature review, G. D'Amico *et al.* identified the recurring themes in logistics initiatives for smart and sustainable port towns [53]. They emphasize the demand for sustainable and smart logistics with digital technologies for data analysis and port cities in their suggested multidimensional framework, which incorporates already-existing enabling areas for sustainable logistics. A different study presented an overview of security challenges for IoT systems in seaports [54]. The authors identified potential vulnerabilities of IoT systems when implemented in seaports, as well as their benefits, and provided a review of existing IoT-enabled smart seaports.

The use of DT in seaport environments is gaining momentum, and although there are several works in this domain, the research is still in its early stages. A DT scheme for next-generation ports was presented by H. Li *et al.* [55]. They also introduced SingaPort.NET as a set of DTs developed using the proposed scheme to tackle various modeling, simulation, and optimization challenges faced by the marine logistics industry. R. Du *et al.* explored potential use cases for 5G-enabled smart ports that leverage the power of DT [56]. These use cases involved various aspects of port operations, including automated rubber-tired gantry and remote-controlled ship-to-shore cranes. They anticipated that in the future, cutting-edge ports, there would be no wired connections at all. Another research study proposed a new method that utilizes a DT for inland river management by integrating 3D video fusion technology [57]. The study revealed that combining 3D video technology with DT can enhance emergency reaction and river monitoring by increasing efficiency and reducing the supervisor's workload.

Despite the numerous studies discussed above on the use of DT and related technologies in various domains, none of them specifically focused on DDoS detection using online learning and the YANG model in ISP core networks or on the combination of honeypot and DT technologies for attack detection and behavioral analysis in smart seaports. These are important areas of research that require further exploration, especially as the use of IoT and other connected technologies in seaports and other critical infrastructure continues to grow, increasing the risk of cyber attacks. Future research could investigate the potential of using DT and related technologies to improve DDoS detection in ISP core networks, as well as to enhance the security of intelligent seaports through the integration of honeypot and DT technologies for attack identification and behavioral analysis.

3. TWINCONET: DT-ENABLED DETECTION ARCHITECTURE FOR AUTONOMOUS ISP CORE NETWORKS

This chapter provides a comprehensive explanation of the proposed architecture, which is enabled by DT technology for autonomous ISP core networks.

3.1 TwinCoNet System Model

The primary objective of the TwinCoNet is the detection of DDoS attacks using DT in the core networks of ISPs The proposed TwinCoNet architecture works on each router separately and operates as follows:

- A digital representation of physical items and synchronization between layers are generated.
- DT gathers all necessary data from the core network routers thanks to the real-twin deployment data collection system with linked sensors.
- Through a YANG model, the system imports data from digitized items in DT.
- The system FS method receives only ninety-two attributes.
- The ten most relevant attributes inputted into the system MLP is then determined using the system FS technique.
- The system MLP assesses whether a DDoS attack has taken place.
- The performance metrics threshold values are then examined.
- The features utilized by the AutoFS module are changed if one of the performance measures falls below its threshold values.
- Within the AutoFS module, a thousand samples are randomly chosen for the current data in each of the five FS methods. Each FS method implements its algorithm to identify the ten most pertinent features.

- The data is labeled by the proposed labeling method.
- Afterward, the MLP utilizes the labeled data for both training and testing purposes.
- The optimal FS technique and MLP model are selected for the given data by the final FS algorithm.
- The system FS technique and MLP model are then modified in accordance with the AutoFS module's findings.
- The system uses the current FS technique and MLP model if the performance metrics meet or exceed their threshold values.

Fig. 3.1 depicts the proposed TwinCoNet architecture. All the decisions required are made in the brain part of the proposed system.

3.2 YANG Data Model

The Network Configuration Protocol (NETCONF), along with NETCONF remote procedure calls (RPCs), and notifications, may handle models of configuration and state data that are created using the modeling language YANG [58]. The IETF's NETMOD (NETCONF Data Modeling Language) working group created the YANG data modeling language, which was later specified in RFC 6020 in October 2010. A data hierarchy structure comprising configuration, status information, RPCs, and alerts, may be created for use in NETCONF-based activities using YANG. As a result, every piece of information sent between a NETCONF client and server may be fully described.

Data structures in YANG are structured into modules and submodules, which can be included from submodules or imported from other external modules. The hierarchy is augmentable, which means that data nodes may be added to the hierarchy established in one module by another. Using a conventional modular language, YANG defines a data structure with a tree topology. It can construct complicated reusable data structures by grouping and has an inherently scalable data type. A few key YANG model elements are the hierarchical configuration data models, data modularity

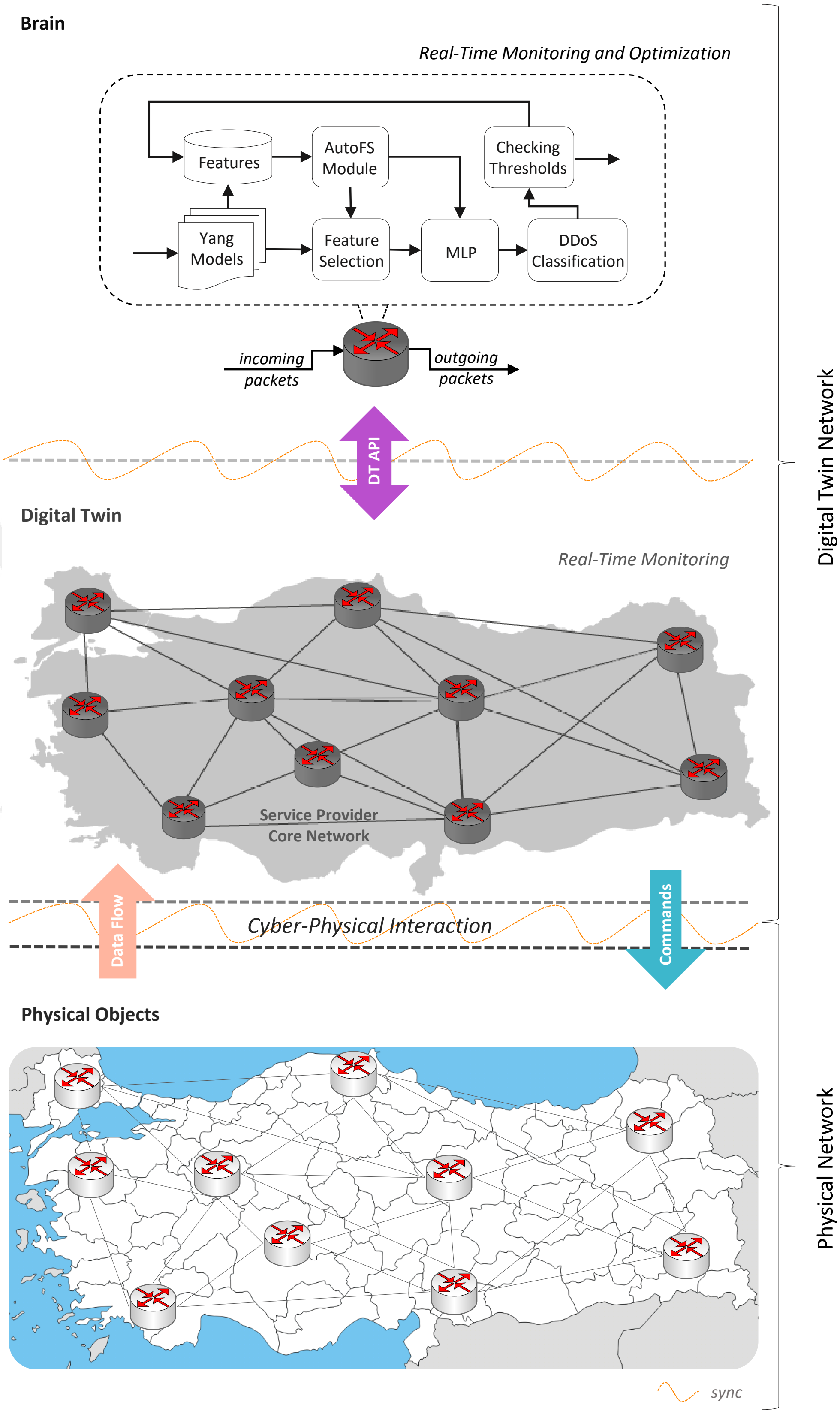


Figure 3.1 : The proposed TwinCoNet architecture [1].

through modules and submodules, reusable types and groups (structured types), and extension through augmentation techniques [59].

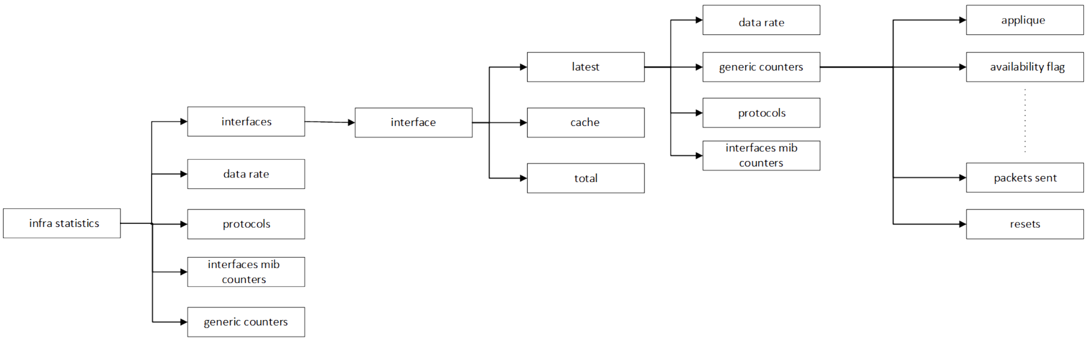


Figure 3.2 : An exemplary YANG Model design.

Using the YANG paradigm in the network has several advantages. Some of them include decreased operating costs by cutting down on legacy network engineering expenditures, faster service rollout made possible by automation, enhanced serviceability, quicker diagnosis and repair times, and more [60]. YANG facilitates shifting from traditional command-line-interface-based management to a more efficient data model-driven approach. On account of the large volume of sensor data in the core network routers compared to other routers, importing all data into the system can significantly reduce the system's performance. Minimizing the amount of data being transferred to the system is critical to address this issue. Hence, we utilize a YANG model, which allows us to achieve this goal by reducing the amount of sensor data being imported into the system. We can selectively pull only the necessary data using YANG's modular structure, which includes modules and submodules. This approach is aligned with the principles of model-driven network automation. Fig. 3.2 shows a sample illustration of the YANG model.

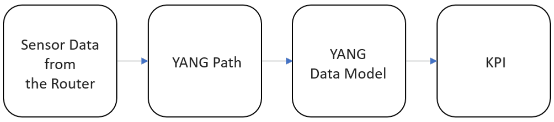


Figure 3.3 : The relationship between the YANG terms [4].

Two key performance indicators (KPIs), KPI-1 and KPI-2, have been established for detecting DDoS attacks. KPI-1 comprises 37 sensors, while KPI-2 contains 55 sensors. Data from these KPIs is obtained using a YANG model, which offers model-driven network automation and allows data flexibility through modules and submodules. The suggested system gets data from all 92 sensors in the core network routers using YANG Paths, which lowers the volume of data delivered to the system and increases its effectiveness. According to the hierarchy's structure, YANG Paths are used to collect sensor data from routers, which then create YANG models, which in turn form KPIs. The relationship between the mentioned terms is depicted more clearly in Fig. 3.3. The sensors of the KPIs can be seen in Fig. 3.4.

3.3 Automated Feature Selection (AutoFS) Module

We assessed various FS methods and narrowed our selection to five that we deemed most appropriate for the proposed system. The AutoFS module identifies the optimal FS method for the system from among these five methods: Chi-square, Recursive Feature Elimination (RFE), Backward Feature Elimination (BFE), Analysis of Variance (ANOVA) F-value Selection, and Fisher Score. To account for the dynamic nature of network data and its online availability, we employ various techniques for FS techniques to yield a range of results for different types of data to address this variability [61] [62]. Chi-square is a statistical technique that assesses the relationship between two categorical variables in discrete data. It helps to identify and select relevant features while excluding unnecessary features. RFE algorithm ranks features based on their importance and continuously eliminates a small number of elements per loop in order to locate the best-performing feature subset. It performs a greedy search to achieve its goal. BFE is a feature selection method that initiates with the entire set of features and gradually removes them one by one based on a feature scoring technique until an optimal subset of features is identified. ANOVA F-value Selection is a statistical method that aims to detect the impact of a feature on the class variable. This technique relies on f-tests to statistically evaluate the equality of means. The Fisher Score evaluates the importance of each feature independently based on their scores and subsequently ranks them according to their significance. After examining these methods, we observed that their performance varies depending on the network traffic. Specifically, ANOVA F-value Selection and RFE methods were found to be the most appropriate for real-time traffic, while BFE and Fisher Score methods were

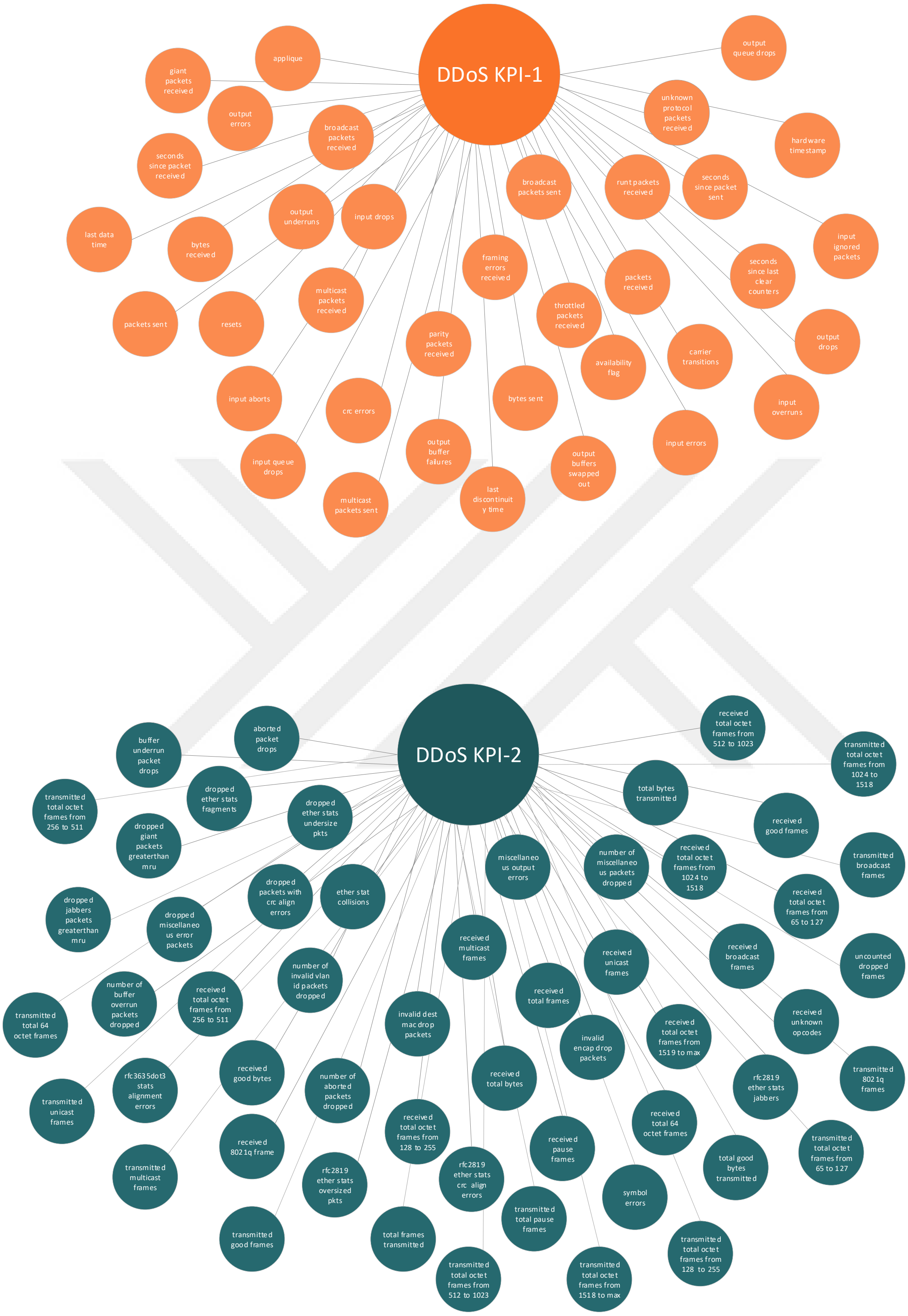

DDoS KPI-1
applique
giant packets received
output errors
seconds since packet received
broadcast packets received
last data time
bytes received
output underruns
input drops
packets sent
resets
multicast packets received
input aborts
crc errors
input queue drops
multicast packets sent
parity packets received
output buffer failures
last discontinuity time
framing errors received
bytes sent
broadcast packets sent
throttled packets received
output buffers swapped out
runt packets received
packets received
availability flag
carrier transitions
input errors
output queue drops
unknown protocol packets received
hardware timestamp
seconds since packet sent
seconds since last clear counters
input ignored packets
output drops
input overruns
DDoS KPI-2
aborted packet drops
buffer underrun packet drops
transmitted total octet frames from 256 to 511
dropped ether stats fragments
dropped giant packets greaterthan mru
dropped ether stats undersize pkts
dropped jabbers packets greaterthan mru
dropped miscellaneous error packets
dropped packets with crc align errors
ether stat collisions
transmitted total 64 octet frames
number of buffer overrun packets dropped
received total octet frames from 256 to 511
number of invalid vlan id packets dropped
transmitted unicast frames
rfc3635dot3 stats alignment errors
received good bytes
number of aborted packets dropped
transmitted multicast frames
received 802.1q frame
rfc2819 ether stats oversized pkts
transmitted good frames
invalid dest mac drop packets
received multicast frames
received total octet frames from 128 to 255
total frames transmitted
rfc2819 ether stats crc align errors
transmitted total octet frames from 512 to 1023
received total bytes
miscellaneous output errors
received total frames
received pause frames
transmitted total pause frames
transmitted total octet frames from 1518 to max
number of miscellaneous packets dropped
received unicast frames
invalid encap drop packets
symbol errors
received total octet frames from 1024 to 1518
received total octet frames from 1519 to max
received total 64 octet frames
transmitted total octet frames from 128 to 255
received total octet frames from 512 to 1023
total bytes transmitted
received good frames
received total octet frames from 65 to 127
received broadcast frames
received unknown opcodes
rfc2819 ether stats jabbers
total good bytes transmitted
transmitted total octet frames from 1024 to 1518
transmitted broadcast frames
uncounted dropped frames
transmitted 802.1q frames
transmitted total octet frames from 65 to 127

Figure 3.4 : KPIs for DDoS Attack detection.

found to be better suited for non-real-time traffic. Additionally, the Chi-square method is well suited for best-effort traffic.

The AutoFS module is responsible for selecting the optimal FS method and MLP model for the system when one of the performance metrics falls below its threshold values. It randomly selects a thousand samples from the current data, which are then input into the five FS methods. We determined the ideal number of parameters to be ten after analyzing the number of parameters utilized for DDoS detection in the literature. The FS methods choose the ten most suitable features based on their algorithms. However, since our data is unlabeled, we need to assign labels. Therefore, we have developed a labeling algorithm for this purpose.

3.3.1 Labeling algorithm for online learning

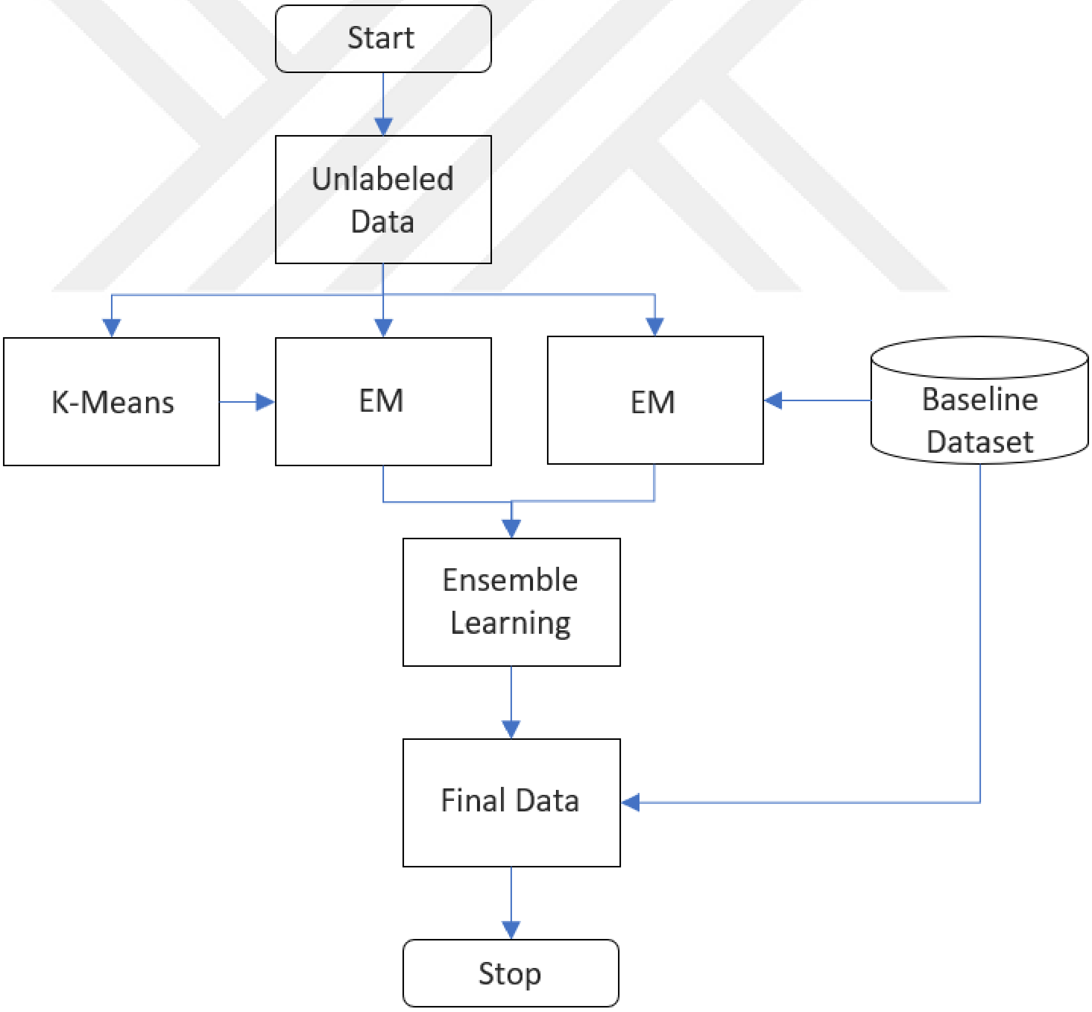


Figure 3.5 : The flow of the proposed labeling algorithm.

Fig. 3.5 depicts the sequence steps of the labeling algorithm. The proposed labeling algorithm operates by utilizing an ensemble learning algorithm that merges the K-Means and EM algorithms, and its steps are as follows:

- The input to the algorithm is unlabeled data.
- Initially, set K to two for K-Means clustering, which divides the data into two groups. This step helps to determine the range of EM initial values, improving the EM algorithm's stability and speed of convergence.
- Next, the EM algorithm assigns probabilistic weighted labels to the unlabeled data.
- Since DDoS attacks are rare, we built a baseline dataset of one thousand samples, including sixty-five percent of 'DDoS Attack' samples, to correctly predict labels.
- The second EM algorithm utilizes both the baseline and unlabeled data to find the local maximum likelihood estimation of parameters in its probabilistic models.
- The outputs of the first and second EM algorithms are inputs for ensemble learning. The data's final labels are then decided.
- Finally, the output of the ensemble learning and the baseline data are combined.

We obtained a labeled dataset consisting of two thousand samples with ten features. MLPs are trained and tested using the labeled dataset. Subsequently, recall and detection time values of the five techniques are obtained and sent to the final FS algorithm.

3.3.2 Final feature selection algorithm

This algorithm, which is shown in Fig. 3.6, selects the optimal FS method by maximizing recall and optimizing detection time for the system. Subsequently, the system FS technique and MLP model weights are changed based on the optimal FS method.

```
Require: $N$ is 5, set of methods $M_i=(M_1, M_2,..., M_N)$ includes set of recall $r_i=(r_1, r_2,..., r_N)$, and set of detection time $t_i=(t_1, t_2,...,t_N)$, threshold of the detection time $(\mathcal{E})$
Ensure: Selected method $(\emptyset)$
if $t_i \leq \mathcal{E}$ then
    while N ← i do
        find maximum $r_i$ ;
        $\emptyset \leftarrow r_i$ ;
    end
end
while N ← i do
    if $r_i = \emptyset$ then
        count++;
    end
end
if count ≥ 2 then
    while N ← i do
        find minimum $t_i$ ;
        $\emptyset \leftarrow t_i$ ;
    end
end
return $\emptyset$
```

Figure 3.6 : The pseudocode of the final feature selection algorithm.

3.3.3 Multilayer perceptron

MLP is a neural network that feeds forward and includes multiple hidden layers where the output of each hidden layer feeds into the input of the subsequent ones. We analyzed various MLP structures and found that the most effective MLP consisted of five layers such as one input layer, three hidden layers, and one output layer. In addition, we tested various activation functions and determined that Rectified Linear Unit (ReLU) was the most effective for the hidden layers, while softmax was optimal for the output layer. Hence, the MLP that we used had the aforementioned parameters. Additionally, to avoid overfitting the MLP, we implemented the dropout technique.

We adopted online learning for the core network since network data is continuously flowing and variable, unlike offline learning, where the network data is fixed. Our system monitors the threshold of performance metrics and updates the MLP model's weights if any of them fall below the threshold values. This indicates that the model's performance is suboptimal and needs to be improved. The AutoFS module facilitates the weight updates of our system's MLP, ensuring stable performance. In this way, we achieve online learning by dynamically adjusting the MLP's weights based on the network's current status. MLP classifies network data as 'DDoS' and 'Not DDoS.' It can be seen in equation 3.1.

$$output = \begin{cases} 1, & \text{DDoS} \\ 0, & \text{Not DDoS} \end{cases} \tag{3.1}$$

3.4 Performance Evaluation

This section includes details about the simulation environment, followed by the simulation results. Four metrics are used to assess the performance of our proposed system: detection performance, system precision, F-measure, and sensitivity.

Detection performance (D_{perf}) of DDoS attacks, which is given by equation 3.2, depends on the count of attacks and the maximum average number of attacks, which is fixed and weekly updated.

$$D_{perf} = \left[\frac{(c_{attack}/C_{max}) \cdot 100}{t} \right] \tag{3.2}$$

where c_{attack} is the number of DDoS attacks detected, C_{max} is the maximum average number of DDoS attacks, and t is the time. System precision (Sys_{pre}) and sensitivity (SV) are defined by the following formulas:

$$Sys_{pre} = \left[\frac{T_P}{(T_P + F_P)} \right] \tag{3.3}$$

$$SV = \left[\frac{T_P}{(T_P + F_N)} \right] \tag{3.4}$$

where F_N is false negative, T_P is true positive, and F_P is false positive.

The fourth metric which is F-measure (F_{meas}) is calculated by the following formula:

$$F_{meas} = \left[\frac{(Sys_{pre} \cdot SV)}{((Sys_{pre} + SV)} \cdot 2 \right] \tag{3.5}$$

3.4.1 Simulation details

To generate twin graphs of physical objects, we utilized the platform-as-a-service technology Microsoft Azure DT (ADT) [63]. ADT has a number of capabilities, including input and output plugins, real-time representation, the DT Definition Language (DTDL), an open modeling language, and real-time modeling. We developed digital models that reflect actual physical things using DTDL, specifically core network routers. These models pinpoint the entities' semantic connections and are connected in a graph that reflects their interactions. To test our system, we created a proof-of-concept model, and an illustration of the twin graph using ADT is shown in Fig. 3.7. An external output plugin that streams data from the ADT platform can be used to store and analyze the data. We transferred data to the brain using the DT API after building the twin graphs and models with predefined interfaces and relationships between elements. The proposed solution's (PS) brain part was designed as a microservice-based system.

For the purpose of assessing PS as the input plugin for ADT, we used two different datasets. The first dataset, referred to as D1, is the CICDDoS2019 dataset which mimics actual data and includes benign and the latest DDoS attacks [64]. It consists of eighty-five features. The second dataset, named D2, is the ToN IoT dataset, developed to measure the effectiveness and accuracy of AI-driven cybersecurity tools for future IoTs and industrial IoTs [65]. It includes forty-three features.

Upon analyzing the datasets, we discovered that the datasets were not balanced. Specifically, D1 consisted of 5,159,863 samples of DDoS attacks and only 1,502 samples of Not DDoS attacks, while D2 contained 300,000 samples of DDoS attacks and 161,043 samples of Not DDoS attacks. To address this imbalance, we applied random undersampling by twenty percent on D1 to decrease the number of Not DDoS samples, followed by synthetic minority oversampling technique (SMOTE). For D2, we used the near-miss undersampling method to reduce the number of Not DDoS samples. These steps helped us achieve more balanced datasets. We combined both datasets and used stratified ten-fold cross-validation to randomly split the data while

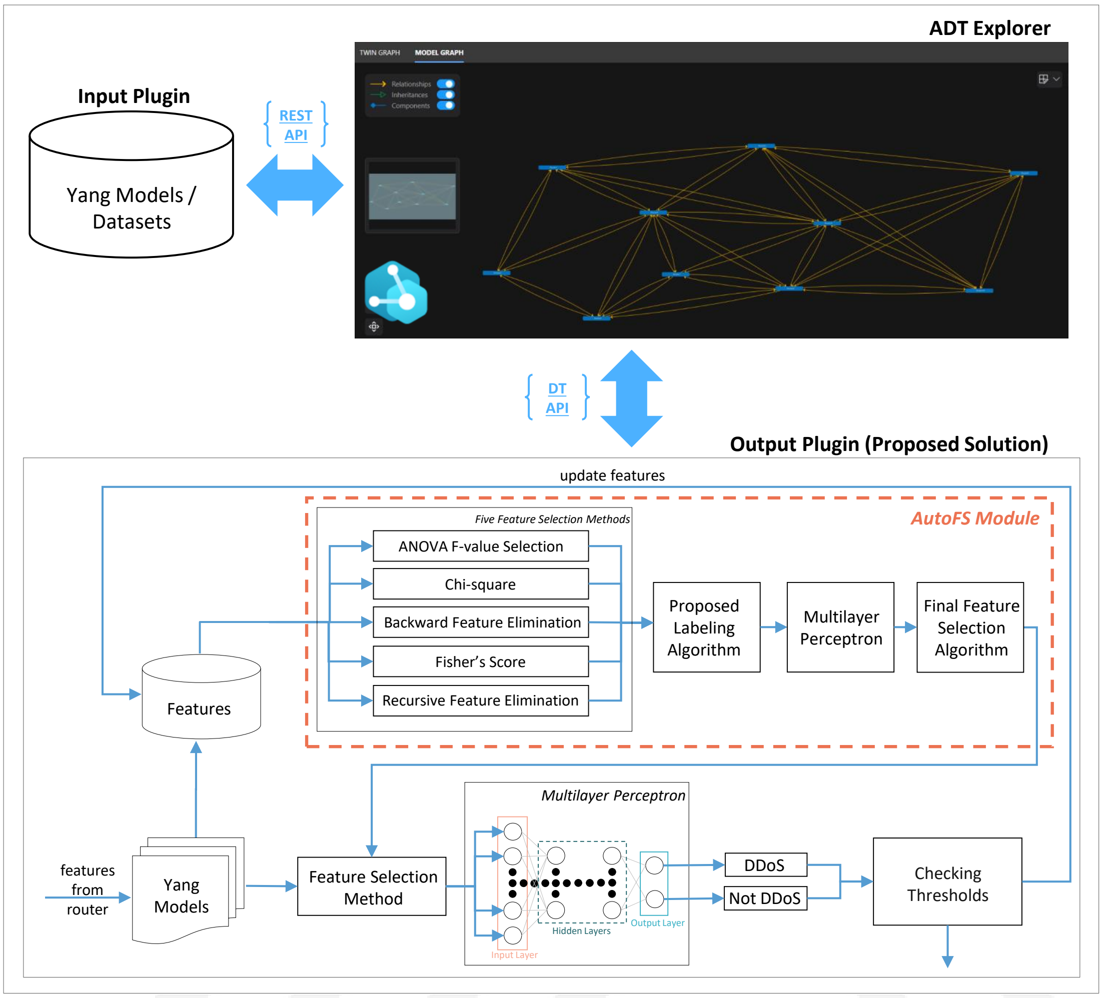


Figure 3.7 : The simulation architecture visualization of the proposed system.

preserving each subset's class distribution. Furthermore, we created a baseline dataset to evaluate the performance of the AutoFS module.

3.4.2 Simulation results

We evaluated PS against multiple algorithms, including a basic deep neural network (DNN), long short-term memory (LSTM), which is a deep learning framework, and RF. The DNN architecture used in the study consisted of five hidden layers, with each layer having a different number of neurons. The number of neurons in each layer was as follows: 8191, 4096, 2048, 1024, and 512. The LSTM architecture consists of four connected layers, with each layer having a specific number of neurons: 256, 128, 64, and 32. The performance measures of the methods were evaluated by computing the

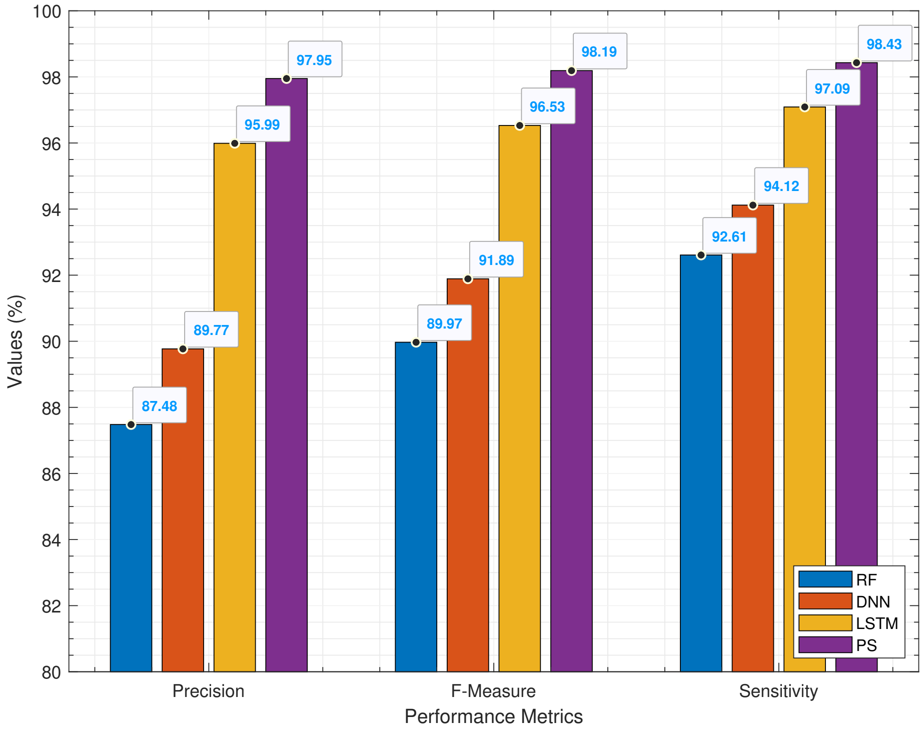


Figure 3.8 : The precision, F-measure, and sensitivity comparison [1].

weighted average of system precision, F-measure, and sensitivity values. The DDoS class was given a weight of sixty percent, while the Not DDoS class was given a weight of forty percent.

The performance results of the methods are presented in Fig. 3.8. PS outperforms the other methods due to its ability to update the model through online learning. We then evaluated the effectiveness of both the AutoFS module and the online learning approach. The findings indicate that PS achieves a true classification rate of 97% by successfully updating the feature selection technique and the MLP weights.

3.4.2.1 Detection performance analysis

During our collaboration with our industry partner, BTS Group, we analyzed the DDoS solutions of four companies and anonymized them as Solutions 1-4. Fig. 3.9 illustrates these solutions' detection rates and latencies, where the data center is shown on the right side of the green dashed line, and the core network is shown on the left. Based on the results, we observed that these solutions detect DDoS attacks approximately 100 minutes after the attack starts, which begins at 60 minutes and ends at 290 minutes. Hence, we concluded that their detection rates and latency are inadequate to ensure the overall network performance.

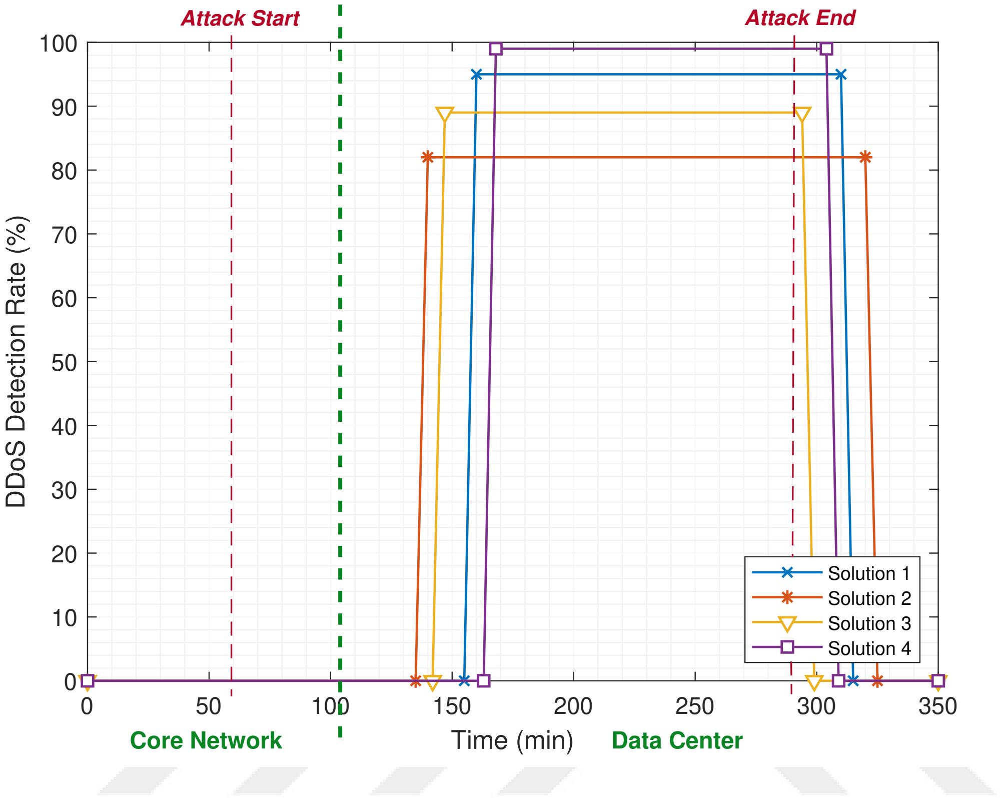


Figure 3.9 : The detection efficiency of existing solutions [1].

Based on our analysis, PS is capable of estimating the onset of a DDoS attack in about 15 minutes. Furthermore, it has a ninety-seven percent detection rate. The detection performance of PS is shown in Fig. 3.10.

When Fig. 3.9 and Fig. 3.10 are examined, it is evident that PS outperforms the previous solutions regarding latency and detection rate.

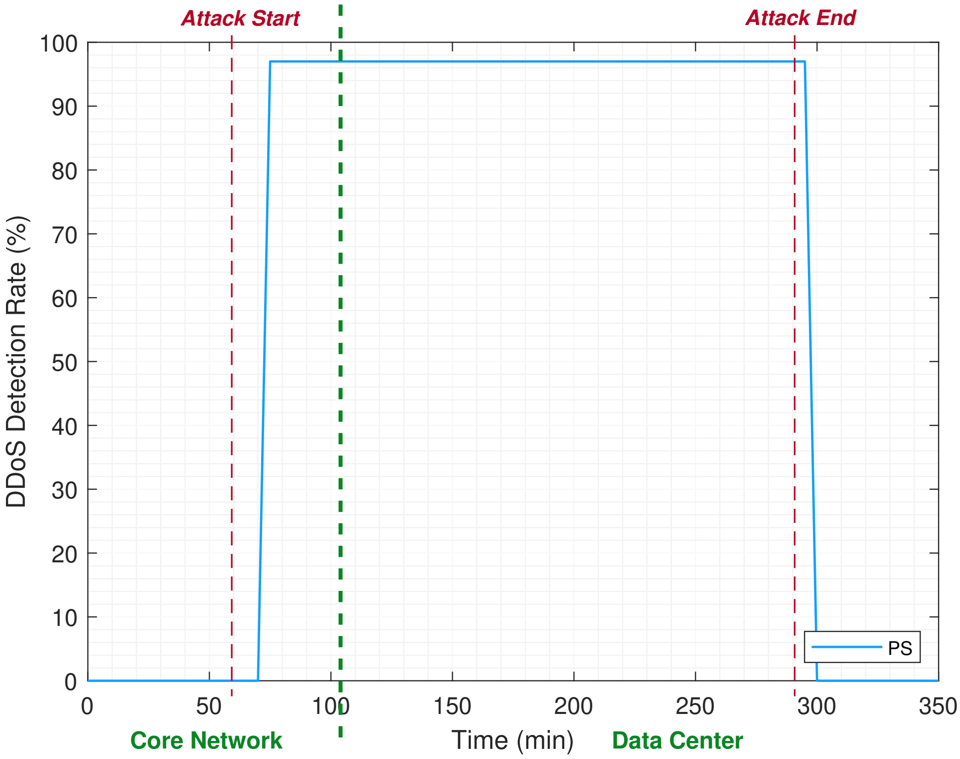


Figure 3.10 : The TwinCoNet architecture detection performance [1].

4. TWINPOT: DT-ASSISTED HONEYPOT ARCHITECTURE FOR AUTONOMOUS SMART SEAPORTS

This chapter thoroughly explains our proposed architecture for autonomous smart seaports, which utilizes DT technology to assist honeypots.

4.1 TwinPot System Model

TwinPot architecture aims detection of external attacks on private networks. Our proposed system comprises two networks, as DT network and the TwinPot network. The DT network consists of the physical, DT, and service layers, while the TwinPot network comprises TwinPot and its service layer. Fig. 4.1 depicts our proposed system architecture for private networks to reach their autonomous goals. We investigated the seaport network as a private network.

In the proposed system, internal and external traffic are the two subcategories of the seaport network traffic. The traffic generated by the seaport's current assets is called 'internal traffic.' On the other hand, traffic from sources other than the existing entities of the seaport, such as IoT devices on incoming cargo ships, is labeled as 'external traffic.' We suggest an intelligent detection method to identify potential attacks originating from the existing seaport assets.

The entities and their relationships in DT are duplicated into TwinPot to draw the attention of attackers away from the main assets. When an attack is detected in TwinPot, its behavioral analysis is conducted in the TwinPot service layer, and the results are sent to the DT service layer to enhance the attack detection mechanism in the actual system. When no attack is detected, incoming traffic is directed to the DT service layer.

TwinPot is a unique solution providing highly sophisticated honeypot entities that can deter potential attackers from targeting the actual assets of the network. To achieve this, we use graph representation similar to DT and periodically transfer data from DT

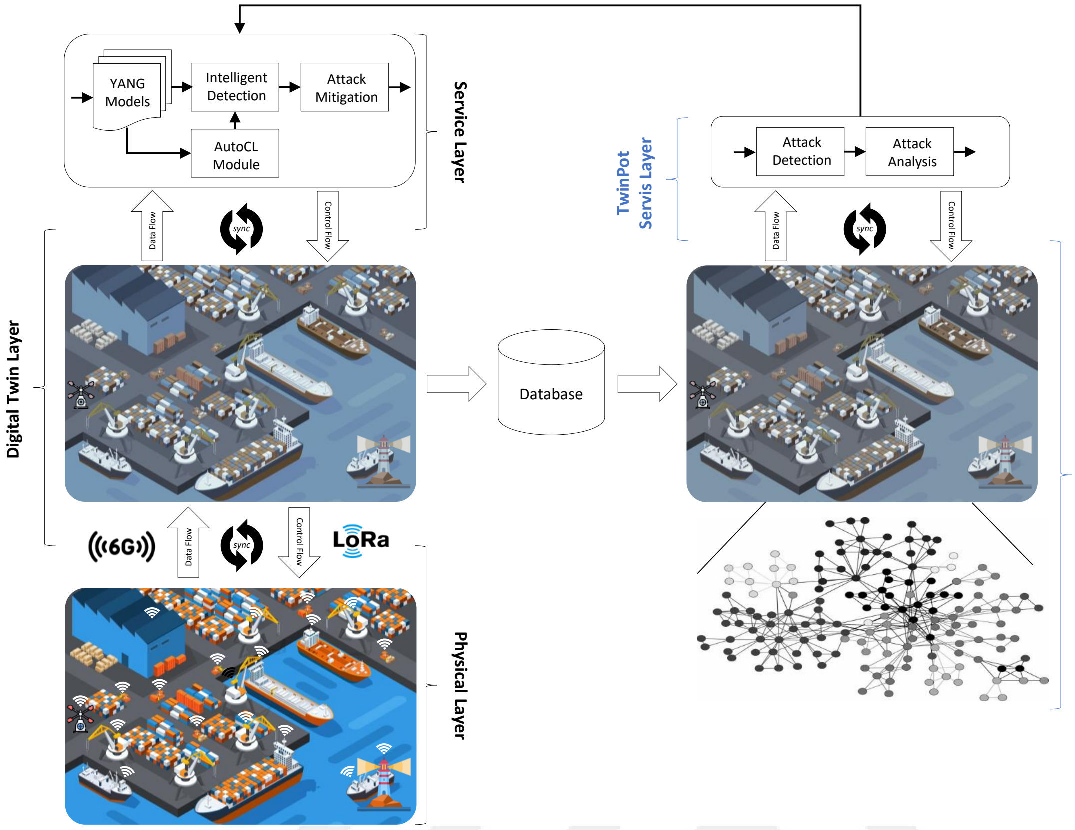


Figure 4.1 : The proposed TwinPot architecture [5].

to a database. An automatic parsing algorithm then extracts the data and removes any sensitive information using YANG models sent to TwinPot entities. By periodically updating data in the database, TwinPot entities can provide a realistic sense of the network to external attackers. Thus, the TwinPot entities are able to effectively deter external attackers from targeting real entities.

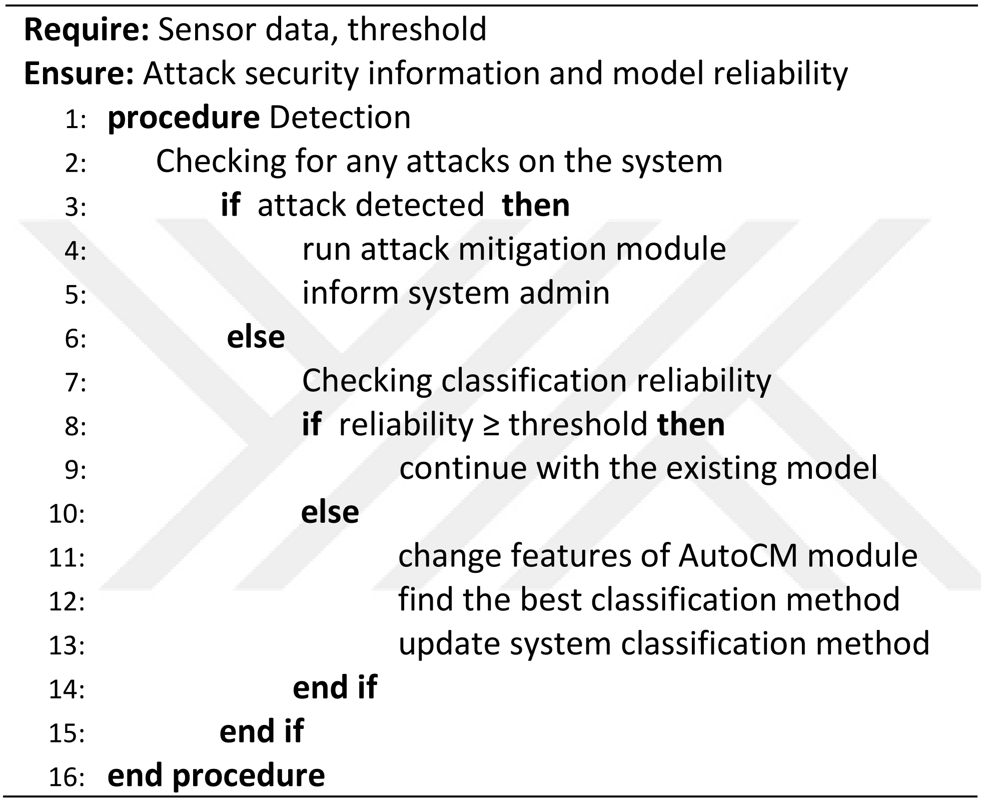
When TwinPot detects an attack, it performs a behavioral analysis of the attack and identifies its characteristics. The analysis is then sent as a report to the DT service layer to improve the network's defense mechanism against similar attacks in the future. Meanwhile, legitimate external traffic is directed to the DT service layer for normal processing. By using TwinPot in combination with DT, we can improve the network's security and gain valuable insights into potential attack vectors and the effectiveness of our defense mechanisms.

4.2 Intelligent Attack Detection Mechanism

The proposed attack detection system focuses on detecting attacks on internal seaport assets. The proposed system comprises three layers: the physical layer consisting of physical entities; the digital twin layer, including digital counterparts of those physical entities; and the service layer facilitating the attack detection mechanism. To achieve intelligent attack detection, Fig. 4.2 provides the pseudocode, which outlines the steps involved in detecting attacks. The algorithm analyzes the behavior of the entities and checks for anomalies that indicate potential attacks. Moreover, it defines a set of thresholds and rules to accurately distinguish between normal and malicious behavior. Our entity-based intelligent detection method operates as follows:

- Digital twins are created to mirror the physical entities.
- The physical layer and the DT layer are synchronized.
- The twins utilize a networked sensor-based data acquisition system to collect data from the physical layer's entities.
- After collecting data from physical entities in the physical layer, the twins transfer the data to the service layer through YANG models.
- The service layer implements the most suitable classification method to identify potential attacks.
- If an attack is identified by the system, the attack mitigation module takes action to mitigate the attack and sends a notification to the system administrator to inform them of the attack.
- The system checks the reliability of the classification method if no attack is detected.
- If the classification model's reliability is equal to or above the threshold, indicating that it can accurately distinguish between attacks and normal traffic, the system continues to operate using the existing model.

- If the reliability of the classification method drops below a predetermined threshold, the system updates the AutoCM features.
- The AutoCM module selects the most suitable classification technique for the system out of a selection of ten different techniques.
- Once the AutoCM identifies the best classification method among the ten techniques, it updates the system's classification method to the chosen technique.



```
Require: Sensor data, threshold
Ensure: Attack security information and model reliability
procedure Detection
    Checking for any attacks on the system
        if attack detected then
            run attack mitigation module
            inform system admin
        else
            Checking classification reliability
            if reliability $\geq$ threshold then
                continue with the existing model
            else
                change features of AutoCM module
                find the best classification method
                update system classification method
            end if
        end if
end procedure
```

Figure 4.2 : The intelligent attack detection mechanism' pseudocode [5].

4.2.1 AutoCM module

The AutoCM module is designed to choose the optimal classification method from a set of ten distinct techniques for the system. To address various attack types in smart maritime ports, we suggest AutoCM, which consists of ten distinct classification techniques, including LSTM, RF, K-Nearest Neighbors (KNN), DeT, NB, Recurrent Neural Networks (RNN), Logistic Regression (LR), Convolutional Neural Network (CNN), MLP, and SVM. Each classification method yields distinct outcomes for diverse attack types.

The proposed AutoCM operates in the following manner:

- AutoCM receives input data consisting of a thousand features imported using YANG models.
- To label the unlabelled data, we utilized the labeling algorithm in subsection 3.3.1, and we also used a baseline dataset similar to the TwinCoNet architecture.
- The unlabelled data is labeled using the labeling algorithm, and an additional one thousand samples are added to the dataset from the baseline dataset.
- Two thousand labeled data samples are used to train and evaluate the models of ten classification methods.
- Subsequently, the final algorithm selects the most suitable classification method for the system based on the performance evaluation of the ten classification algorithms trained and tested using the two thousand labeled data samples.

After training and testing the models, the performance metrics of each model are sent to the final method algorithm.

4.2.2 Final method algorithm

This algorithm evaluates the effectiveness of each classification technique and selects the most optimal one. Then, the system's classification method is updated based on the best-performing technique.

In the final method algorithm, false negative (FN) and false positive (FP) values are considered, which occur when the model estimated the positive and negative class inaccurately, respectively. The determination time is also taken into account. Our optimization objective for this algorithm is as follows:

$$arg\,max\,(\alpha_i\lambda_i + \beta_i\nu_i),\ i \in [1,10] \tag{4.1}$$

In equation 4.1, λ_i represents weighted sum of precision and recall of i_{th} method, and ν_i is determination time of i_{th} classification method. Among the ten techniques, we find the best classification method that has the maximum λ_i and optimal v_i values. We calculate λ_i according to the following formula:

$$\lambda_i = (0.6)\frac{TP}{TP+FN} + (0.4)\frac{TP}{TP+FP},\ i \in [1,10] \tag{4.2}$$

In Equation 4.2, TP presents the true positive, the value that occurs when the model correctly identifies the positive class, of i_{th} classification method. The attack determination time is another critical metric for attack detection systems. Equation 4.3 shows the determination time formula, in which finishing time is (t_{end}) and starting time is (t_{start}) of i_{th} classification method.

$$v_i = t_i^{end} - t_i^{start},\ i \in [1,10] \tag{4.3}$$

In addition, we have defined the reliability of the classification method using the following formula:

$$\gamma = \{1 - \frac{FN}{TP+FN}\} \tag{4.4}$$

In Equation 4.4, γ stands for the reliability of the classification method. Since FN is a significant metric for data division, we define the method's reliability based on FN.

4.3 Performance Evaluation

This section provides information on the simulation environment in detail. The simulation results are then presented after the simulation is performed.

4.3.1 Simulation details

We used the ADT platform in this simulation to create twin graphs of physical objects as in the previous simulation. To generate maritime data, we utilized MANSIM v2.01 (Maneuvering Simulation Laboratory), which is a maritime simulation code that creates time-dependent ship motion data for surface vessels [66]. The software can produce time-dependent ship motion data for surface vessels with 3DOF and 4DOF simulations. It can simulate ships with twin-screw/twin-rudder and single-screw/single-rudder configurations [67]. The software is available for free and can be accessed online [68]. Based on the math model, there are different coefficients and parameters.

The MANSIM code necessitates various inputs, including details of the ship and its environment, hydrodynamic coefficients, and solver parameters. The first three factors pertain to the ship's physical characteristics and surroundings, while the fourth determines the specifics of the simulation. Furthermore, the code allows for controlled movements and includes an extra input for control parameters. Depending on the chosen mode, standard maneuvering tests, or free maneuver, the propeller rotation rate and rudder angle can be input or output values. Once the inputs are identified, the code returns the ship's accelerations, velocities, and position at each time step. We employed these parameters to model ship assets in the seaport.

In addition, we utilized two datasets to portray IoT assets within the seaport and assess the effectiveness of our approach in detecting potential cyber-attacks. The first dataset we utilized to evaluate the effectiveness of our solution in detecting attacks is called Edge-IIoTset. It is a recent dataset designed for the IoT and IIoT and consists of fourteen attacks on connectivity protocols [69]. We employed the internal traffic from this dataset to assess the effectiveness of our system. We utilized the SDSN dataset we developed in our previous study for external traffic [45]. This dataset includes nine distinct UDP DDoS attack datasets generated by varying traffic payloads and heterogeneous traffic speeds.

4.3.2 Simulation results

We compare the DNN method, which was determined to be the best-performing algorithm in the work of M. A. Ferrag *et. al* [69], with our PS. To evaluate the detection of internal attacks, we utilized 5,000 instances of twelve attack traffic samples and 15,000 normal traffic samples from the Edge-IIoTset, along with 20,000 data samples generated from MANSIM, to form a comprehensive dataset for internal attacks in the smart seaport. The sensitivity metric was used to evaluate performance, which is the ratio of correctly identified attack samples to all samples that were supposed to be attacked. Fig. 4.3 illustrates the performance results of our proposed solution for internal attacks.

To evaluate external attacks, we utilized 2000 samples of attack traffic and 6000 samples of normal traffic from the SDSN dataset. These samples were combined with 8000 MANSIM data samples to create a comprehensive dataset for external attacks in a smart seaport.

The comparison of performance for external attacks is illustrated in Fig. 4.4. The findings indicated that our proposed solution outperformed the DNN algorithm in detecting both internal and external attacks.

To evaluate the system's performance under simultaneous internal and external attacks, we sent a fixed number of both types of attacks to the system and compared the detection rates of our PS and the DNN algorithm. We defined the detection rate on a scale of 0 to 1, where "0" represents out-of-detection, and "1" indicates the detection of all attacks. We analyzed the performance of the system when subjected to simultaneous attacks for all possible combinations of six internal and six external attacks. The detection rate of simultaneous attacks is illustrated in Fig. 4.5. Our proposed solution outperformed the DNN algorithm in terms of detecting simultaneous internal and external attacks, indicating that it is more robust. Overall, our proposed approach successfully identifies both internal and external attacks.

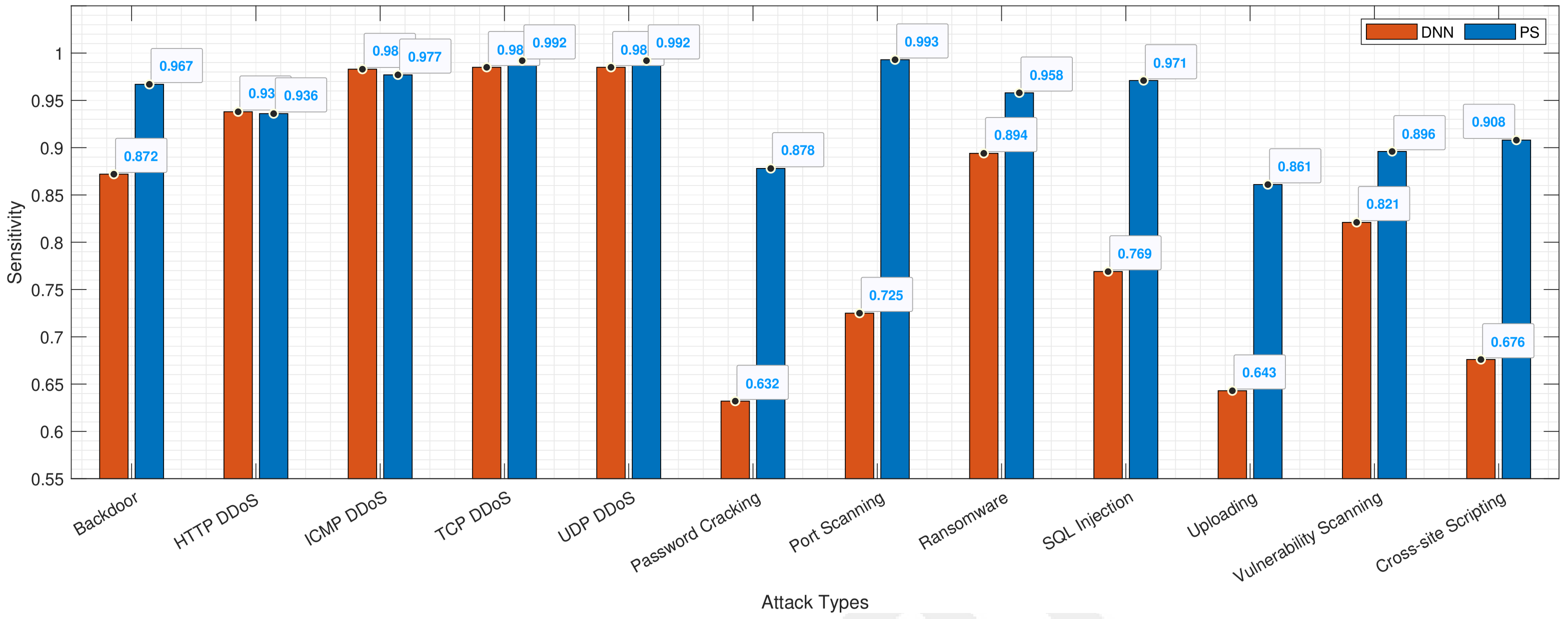


Figure 4.3 : The performance comparison of the internal attacks [5].

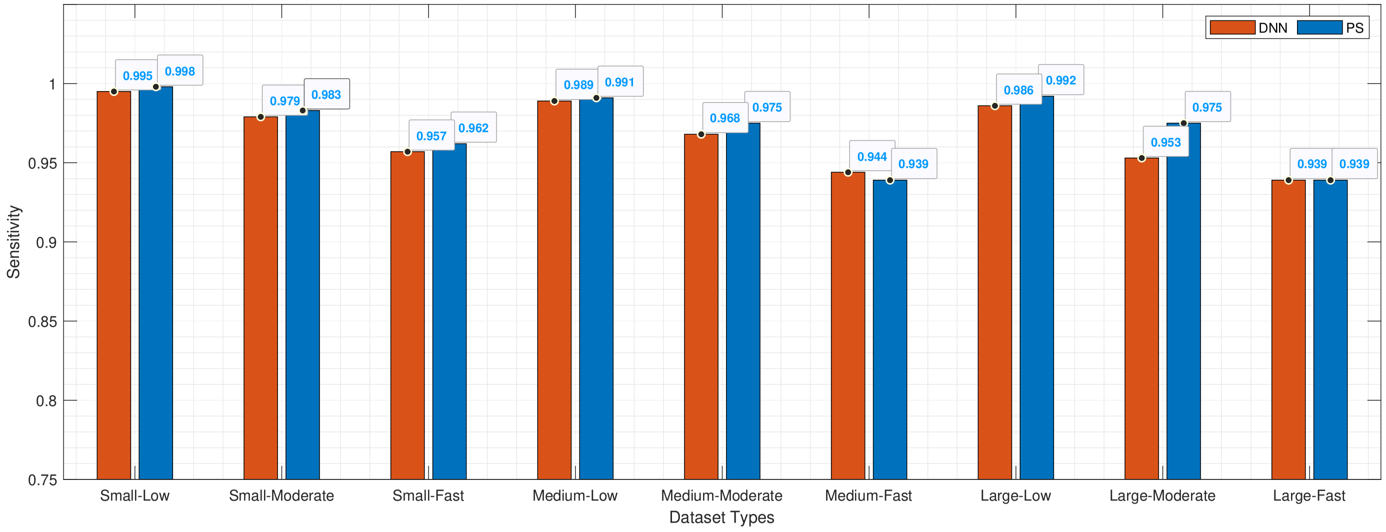


Figure 4.4 : The performance comparison of the external attacks [5].

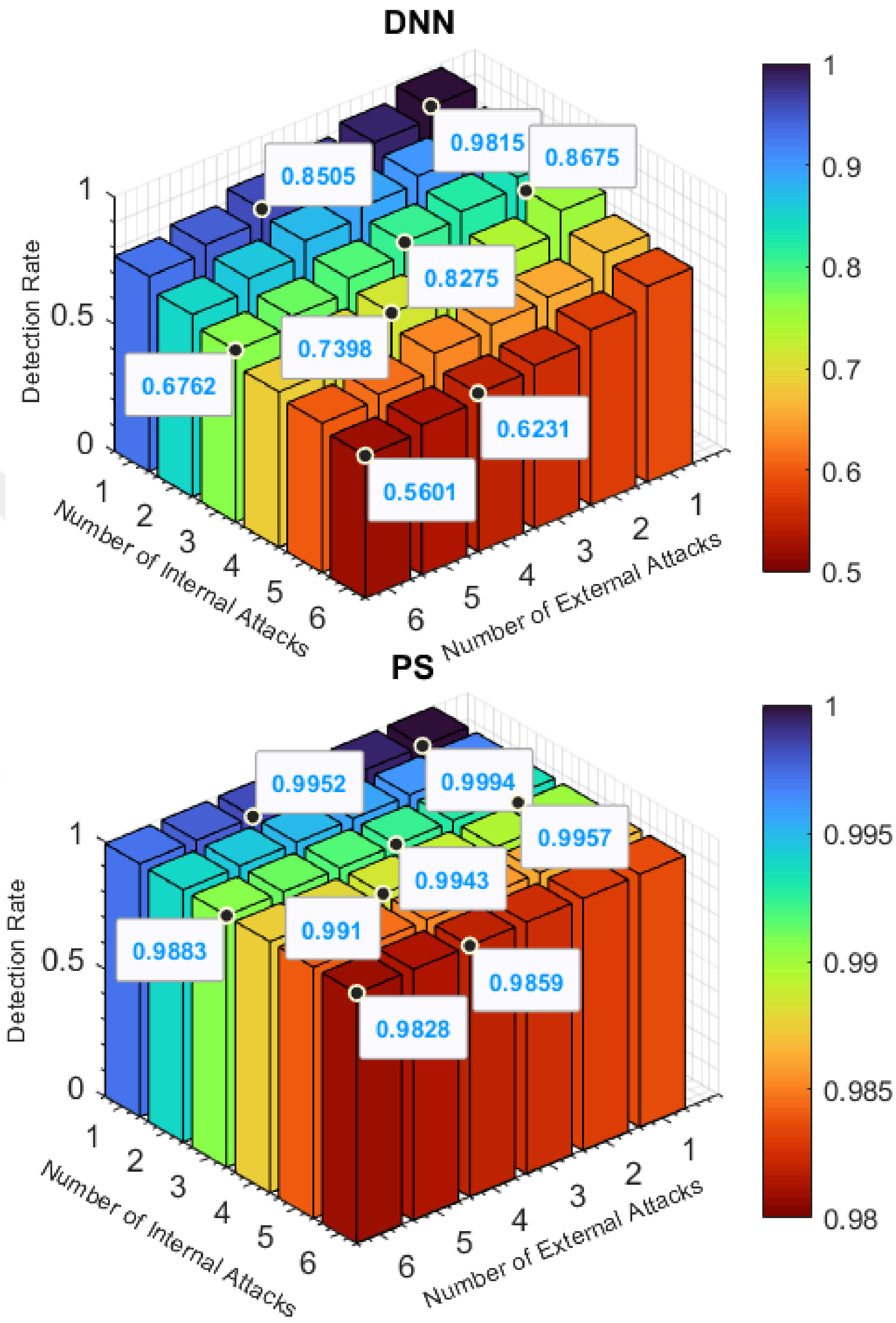


Figure 4.5 : The detection rate comparison of simultaneous attacks [5].

5. CONCLUSIONS

In conclusion, this study highlights the significance of DT technology in enhancing cyber-security solutions for critical infrastructures such as ISP core networks and seaports. We proposed two intelligent detection mechanisms for ISP core and seaport networks to solve attack detection problems and provide autonomous network characteristics. The TwinCoNet mechanism effectively identifies DDoS attacks, adjusts the feature selection method, and estimates the attack within fifteen minutes of it commencing, with an accurate categorization rate of ninety-seven percent. The proposed system ensures that the network can support the model's computational capability by utilizing the necessary data collected from the YANG model and AutoFS module.

Furthermore, we designed the TwinPot system, which is a DT-assisted honeypot for next-generation smart seaports, which divides the seaport network traffic into internal and external traffic. The proposed intelligent attack detection mechanism uses DT technology to handle different attack types for the current entities in seaports. The performance results indicated that the suggested mechanism successfully identifies both internal and external attacks. Additionally, we demonstrated that the proposed system successfully detects internal and external attacks even under simultaneous attacks on the system. Overall, the proposed intelligent detection mechanisms for ISP core and seaport networks are innovative and provide autonomous network characteristics. The performance results of both mechanisms are measured on extensive datasets. The proposed mechanisms show the potential of DT technology in network anomaly detection and cyber-attack detection. The findings of this study highlight the importance of intelligent attack detection to reduce downtime and ensure operational continuity for critical infrastructures.

REFERENCES

[1] **Yigit, Y.**, **Bal, B.**, **Karameseoglu, A.**, **Duong, T.Q. and Canberk, B.** (2022). Digital Twin-Enabled Intelligent DDoS Detection Mechanism for Autonomous Core Networks, *IEEE Communications Standards Magazine*, *6*(3), 38–44.

[2] **Petrosyan, A.** (2023). *Global number of internet users 2005-2022*, Retrieved May 2, 2023, from `https://www.statista.com/statistics/273018/number-of-internet-users-worldwide`.

[3] **Lab, K.** (2020). *DDoS attacks in Q4 2020 Report*, retrieved Jan. 01, 2023, from `https://securelist.com/`.

[4] **Yigit, Y.**, **Huseynov, K.**, **Ahmadi, H. and Canberk, B.** (2022). YA-DA: YAng-Based DAta Model for Fine-Grained IIoT Air Quality Monitoring, *2022 IEEE Globecom Workshops (GC Wkshps)*, pp.438–443.

[5] **Yigit, Y.**, **Kinaci, O.K.**, **Duong, T.Q. and Canberk, B.** (2023). TwinPot: Digital Twin-assisted Honeypot for Cyber-Secure Smart Seaports, *2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, p. .

[6] **Coronado, E.**, **Behravesh, R.**, **Subramanya, T.**, **Fernàndez-Fernàndez, A.**, **Siddiqui, M.S.**, **Costa-Pérez, X. and Riggio, R.** (2022). Zero Touch Management: A Survey of Network Automation Solutions for 5G and 6G Networks, *IEEE Communications Surveys & Tutorials*, *24*(4), 2535–2578.

[7] **Security, A.** (2019). *The Cost of Cybercrime: Ninth Annual Cost of Cybercrime Study*, Retrieved Jan. 7, 2023, from `https://www.accenture.com/`.

[8] **branessenseresearch** (2022). *Cybersecurity Market Size By Solution Forecast 2021-2027*, Retrieved Jan. 05, 2023, from `https://brandessenceresearch.com/technology-and-media/cybersecurity-market-industry-analysis/`.

[9] **Deng, J.**, **Zheng, Q.**, **Liu, G.**, **Bai, J.**, **Tian, K.**, **Sun, C.**, **Yan, Y. and Liu, Y.** (2021). A Digital Twin Approach for Self-optimization of Mobile Networks, *2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp.1–6.

[10] **Dressler, F.** (2007). Self-Organization – Context and Capabilities, *Self-Organization in Sensor and Actor Networks*, John Wiley & Sons, Ltd,

pp.17–31, https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470724460.ch3, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470724460.ch3.

[11] **Wang, Y.**, **Cao, Y. and Wang, F.Y.** (2021). Anomaly Detection in Digital Twin Model, *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, pp.208–211.

[12] **Wu, Y.**, **Zhang, K. and Zhang, Y.** (2021). Digital Twin Networks: A Survey, *IEEE Internet of Things Journal*, *8*(18), 13789–13804.

[13] **Netscout** (2022). *NETSCOUT THREAT INTELLIGENCE REPORT: DDoS in a Time of Pandemic*, Retrieved May 2, 2023, from https://www.netscout.com/threatreport/.

[14] **Imperva** (2020). *DDoS Attacks in the Time of COVID-19*, Retrieved Jan. 2, 2023, from https://www.imperva.com/.

[15] **Statista** (2020). *Average cost per hour of enterprise server downtime worldwide in 2019*, Retrieved Aug. 1, 2022, from https://www.statista.com/statistics/753938/worldwide-enterprise-server-hourly-downtime-cost/#statisticContainer.

[16] **Zdnet** (2020). *European ISPs report mysterious wave of DDoS attacks*, Retrieved Jan. 7, 2023, from https://www.zdnet.com/article/european-isps-report-mysterious-wave-of-ddos-attacks/.

[17] **Independent** (2021). *Irish internet service providers hit by cyber attacks*, Retrieved Jan. 7, 2023, from https://www.independent.ie/business/technology/.

[18] **Ben Farah, M.A.**, **Ukwandu, E.**, **Hindy, H.**, **Brosset, D.**, **Bures, M.**, **Andonovic, I. and Bellekens, X.** (2022). Cyber Security in the Maritime Industry: A Systematic Survey of Recent Advances and Future Trends, *Information*, *13*(1), https://www.mdpi.com/2078-2489/13/1/22.

[19] **Markets and Markets** (2022). *Smart Port Market by Technology, Elements, Throughput Capacity, Port Type - and Region - Global Forecast to 2027*, Retrieved Jan. 12, 2023, from https://www.marketsandmarkets.com/Market-Reports/smart-ports-market-165784113.html.

[20] **of Ports, I.A. and (IAPH), H.** (2021). *IAPH Cybersecurity Guidelines for Ports and Port Facilities, Version 1.0*, Retrieved Jan. 12, 2023, from https://sustainableworldports.org.

[21] **Times, N.N.** (2017). *Rotterdam Port, TNT hit in New Ransomware Attack*, Retrieved Jan. 05, 2023, from https://nltimes.nl/2017/06/28/rotterdam-port-tnt-hit-new-ransomware-attack.

[22] **Franco, J.**, **Aris, A.**, **Canberk, B. and Uluagac, A.S.** (2021). A Survey of Honeypots and Honeynets for Internet of Things, Industrial Internet of Things, and Cyber-Physical Systems, *IEEE Communications Surveys & Tutorials*, *23*(4), 2351–2383.

[23] **Zhou, C.**, **Yang, H.**, **Duan, X.**, **Lopez, D.**, **Pastor, A.**, **Wu, Q.**, **Boucadair, M. and Jacquenet, C.** (2021). *Digital Twin Network: Concepts and Reference Architecture*, Retrieved Dec. 15, 2021, from `https://datatracker.ietf.org/doc/html/draft-zhou-nmrg-digitaltwin-network-concepts-06`.

[24] **Boite, J.**, **Nardin, P.A.**, **Rebecchi, F.**, **Bouet, M. and Conan, V.** (2017). Statesec: Stateful monitoring for DDoS protection in software defined networks, *2017 IEEE Conference on Network Softwarization (NetSoft)*, pp.1–9.

[25] **Maseer, Z.K.**, **Yusof, R.**, **Bahaman, N.**, **Mostafa, S.A. and Foozy, C.F.M.** (2021). Benchmarking of Machine Learning for Anomaly Based Intrusion Detection Systems in the CICIDS2017 Dataset, *IEEE Access*, *9*, 22351–22370.

[26] **Sadaf, K. and Sultana, J.** (2020). Intrusion Detection Based on Autoencoder and Isolation Forest in Fog Computing, *IEEE Access*, *8*, 167059–167068.

[27] **Khoei, T.T.**, **Aissou, G.**, **Hu, W.C. and Kaabouch, N.** (2021). Ensemble Learning Methods for Anomaly Intrusion Detection System in Smart Grid, *2021 IEEE International Conference on Electro Information Technology (EIT)*, pp.129–135.

[28] **Wei, Y.**, **Jang-Jaccard, J.**, **Sabrina, F.**, **Singh, A.**, **Xu, W. and Camtepe, S.** (2021). AE-MLP: A Hybrid Deep Learning Approach for DDoS Detection and Classification, *IEEE Access*, *9*, 146810–146821.

[29] **Trejo, L.A.**, **Ferman, V.**, **Medina-Pérez, M.A.**, **Arredondo Giacinti, F.M.**, **Monroy, R. and Ramirez-Marquez, J.E.** (2019). DNS-ADVP: A Machine Learning Anomaly Detection and Visual Platform to Protect Top-Level Domain Name Servers Against DDoS Attacks, *IEEE Access*, *7*, 116358–116369.

[30] **D. Peraković, M. Periša, I.C. and Husnjak, S.** (2016). Artificial neuron network implementation in detection and classification of DDoS traffic, *2016 24th Telecommunications Forum (TELFOR)*, 1–4.

[31] **Hoque, N.**, **Bhattacharyya, D.K. and Kalita, J.K.** (2015). Botnet in DDoS Attacks: Trends and Challenges, *IEEE Communications Surveys Tutorials*, *17*(4), 2242–2270.

[32] **Zhou, B.**, **Li, J.**, **Ji, Y. and Guizani, M.** (2018). Online Internet Traffic Monitoring and DDoS Attack Detection Using Big Data Frameworks, *2018 14th International Wireless Communications Mobile Computing Conference (IWCMC)*, pp.1507–1512.

[33] **Yu, J.**, **Lee, H.**, **Kim, M.S. and Park, D.** (2008). Traffic flooding attack detection with SNMP MIB using SVM, *Computer Communications*, *31*(17), 4212–4219.

[34] **Preetha, G.**, **Devi, B. and Shalinie, S.** (2014). Autonomous Agent for DDoS Attack Detection and Defense in an Experimental Testbed, *International Journal of Fuzzy Systems*, *16*, 520–528.

[35] **Lyu, M.**, **Gharakheili, H.H.**, **Russell, C. and Sivaraman, V.** (2021). Hierarchical Anomaly-Based Detection of Distributed DNS Attacks on Enterprise Networks, *IEEE Transactions on Network and Service Management*, *18*(1), 1031–1048.

[36] **Saied, A.**, **Overill, R. and Radzik, T.** (2014). Artificial Neural Networks in the Detection of Known and Unknown DDoS Attacks: Proof-of-Concept, *Springer International Publishing*, 309– 320.

[37] **Balkanli, E.**, **Alves, J. and Zincir-Heywood, A.N.** (2014). Supervised learning to detect DDoS attacks, *2014 IEEE Symposium on Computational Intelligence in Cyber Security (CICS)*, pp.1–8.

[38] **Yadav, S. and Subramanian, S.** (2016). Detection of Application Layer DDoS attack by feature learning using Stacked AutoEncoder, *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, pp.361–366.

[39] **Yusof, A.R.**, **Udzir, N.I.**, **Selamat, A.**, **Hamdan, H. and Abdullah, M.T.** (2017). Adaptive feature selection for denial of services (DoS) attack, *2017 IEEE Conference on Application, Information and Network Security (AINS)*, pp.81–84.

[40] **Gu, Y.**, **Li, K.**, **Guo, Z. and Wang, Y.** (2019). Semi-Supervised K-Means DDoS Detection Method Using Hybrid Feature Selection Algorithm, *IEEE Access*, *7*, 64351–64365.

[41] **Das, S.**, **Venugopal, D.**, **Shiva, S. and Sheldon, F.T.** (2020). Empirical Evaluation of the Ensemble Framework for Feature Selection in DDoS Attack, *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, pp.56–61.

[42] **Biswas, R.**, **Kim, S. and Wu, J.** (2021). Sampling Rate Distribution for Flow Monitoring and DDoS Detection in Datacenter, *IEEE Transactions on Information Forensics and Security*, *16*, 2524–2534.

[43] **Biswas, R.**, **Wu, J. and Chen, Y.** (2019). Optimal Monitor Placement Policy Against Distributed Denial-of-Service Attack in Datacenter, *2019 Resilience Week (RWS)*, volume 1, pp.64–70.

[44] **Garg, S.**, **Kaur, K.**, **Kumar, N.**, **Kaddoum, G.**, **Zomaya, A.Y. and Ranjan, R.** (2019). A Hybrid Deep Learning-Based Model for Anomaly Detection in Cloud Datacenter Networks, *IEEE Transactions on Network and Service Management*, *16*(3), 924–935.

[45] **Horsanali, E.**, **Yigit, Y.**, **Secinti, G.**, **Karameseoglu, A. and Canberk, B.** (2021). Network-Aware AutoML Framework for Software-Defined Sensor Networks, *2021 17th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp.451–457.

[46] **Saad, A.**, **Faddel, S.**, **Youssef, T. and Mohammed, O.A.** (2020). On the Implementation of IoT-Based Digital Twin for Networked Microgrids Resiliency Against Cyber Attacks, *IEEE Transactions on Smart Grid*, *11*(6), 5138–5150.

[47] **Xie, Y.** (2020). Modified Label Propagation on Manifold With Applications to Fault Classification, *IEEE Access*, *8*, 97771–97782.

[48] **Ak, E.**, **Duran, K.**, **Dobre, O.A.**, **Duong, T.Q. and Canberk, B.** (2023). T6CONF: Digital Twin Networking Framework for IPv6-Enabled Net-Zero Smart Cities, *IEEE Communications Magazine*, 1–7.

[49] **Zhang, W.**, **Zhang, B.**, **Zhou, Y.**, **He, H. and Ding, Z.** (2020). An IoT Honeynet Based on Multiport Honeypots for Capturing IoT Attacks, *IEEE Internet of Things Journal*, *7*(5), 3991–3999.

[50] **Wang, B.**, **Dou, Y.**, **Sang, Y.**, **Zhang, Y. and Huang, J.** (2020). IoTCMal: Towards A Hybrid IoT Honeypot for Capturing and Analyzing Malware, *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp.1–7.

[51] **Ben Farah, M.A.**, **Ukwandu, E.**, **Hindy, H.**, **Brosset, D.**, **Bures, M.**, **Andonovic, I. and Bellekens, X.** (2022). Cyber Security in the Maritime Industry: A Systematic Survey of Recent Advances and Future Trends, *Information*, *13*(1), `https://www.mdpi.com/2078-2489/13/1/22`.

[52] **de la Peña Zarzuelo, I.** (2021). Cybersecurity in ports and maritime industry: Reasons for raising awareness on this issue, *Transport Policy*, *100*, 1–4, `https://www.sciencedirect.com/science/article/pii/S0967070X20308945`.

[53] **D'Amico, G.**, **Szopik-Depczyńska, K.**, **Dembińska, I. and Ioppolo, G.** (2021). Smart and sustainable logistics of Port cities: A framework for comprehending enabling factors, domains and goals, *Sustainable Cities and Society*, *69*, 102801, `https://www.sciencedirect.com/science/article/pii/S2210670721000937`.

[54] **Jović, M.**, **Tijan, E.**, **Aksentijević, S. and Čišić, D.** (2019). An Overview Of Security Challenges Of Seaport IoT Systems, *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp.1349–1354.

[55] **Li, H.**, **Cao, X.**, **Sharma, P.**, **Lee, L.H. and Chew, E.P.** (2020). Framework of O2DES.NET Digital Twins for Next Generation Ports and Warehouse Solutions, *2020 Winter Simulation Conference (WSC)*, pp.3188–3199.

[56] **Du, R.**, **Mahmood, A. and Auer, G.** (2022). Realizing 5G Smart-Port Use Cases with a Digital Twin, *Ericsson Technology Review*, *2022*(13), 2–11.

[57] **Wu, Z.**, **Ren, C.**, **Wu, X.**, **Wang, L.**, **Zhu, L. and Lv, Z.** (2021). Research on Digital Twin Construction and Safety Management Application of Inland Waterway Based on 3D Video Fusion, *IEEE Access*, *9*, 109144–109156.

[58] **IETF** (2010). *YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)*, Retrieved June 9, 2022, from `https://rfc-editor.org/rfc/rfc6020.txt`.

[59] **Bjorklund, M.** (2016). *The YANG 1.1 Data Modeling Language, RFC 7950*, Retrieved Aug. 12, 2022, from `https://www.rfc-editor.org/rfc/rfc7950.html`.

[60] **Claise, B.**, **Clarke, J. and Lindblad, J.** (2019). *Network Programmability with YANG: The Structure of Network Automation with YANG, NETCONF, RESTCONF, and gNMI*, Addison-Wesley Professional.

[61] **Sun, P.**, **Wang, D.**, **Mok, V.C. and Shi, L.** (2019). Comparison of Feature Selection Methods and Machine Learning Classifiers for Radiomics Analysis in Glioma Grading, *IEEE Access*, *7*, 102010–102020.

[62] **Rachburee, N. and Punlumjeak, W.** (2015). A comparison of feature selection approach between greedy, IG-ratio, Chi-square, and mRMR in educational mining, *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)*, pp.420–424.

[63] **Microsoft** (2023). *Azure Digital Twins Documentation*, Retrieved May 17, 2023, from `https://docs.microsoft.com/en-us/azure/digital-twins/`.

[64] **Sharafaldin, I.**, **Lashkari, A.H.**, **Hakak, S. and Ghorbani, A.A.** (2019). Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy, *2019 International Carnahan Conference on Security Technology (ICCST)*, IEEE, pp.1–8.

[65] **Booij, T.M.**, **Chiscop, I.**, **Meeuwissen, E.**, **Moustafa, N. and den Hartog, F.T.** (2021). ToN_IoT: The role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets, *IEEE Internet of Things Journal*, *9*(1), 485–496.

[66] **Sukas, O.F.**, **Kinaci, O.K. and Bal, S.** (2019). Theoretical background and application of MANSIM for ship maneuvering simulations, *Ocean Engineering*, *192*, 106239, `https://www.sciencedirect.com/science/article/pii/S0029801819304184`.

[67] **Kinaci, O.K.**, **Delen, C.**, **Bitirgen, R.**, **Bayezit, A.**, **Bayezit, I.**, **Ozturk, D. and Gunguder, B.** (2021). Free-running tests for DTC self-propulsion–An investigation of lateral forces due to the rudder and the propeller, *Applied Ocean Research*, *116*, 102877.

[68] **MansimLab** (2021). *Maneuvering Simulation Laboratory Code*, Retrieved Aug 7, 2022, from `https://mansim.org/`.

[69] **Ferrag, M.A.**, **Friha, O.**, **Hamouda, D.**, **Maglaras, L. and Janicke, H.** (2022). Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning, *IEEE Access*, *10*, 40281–40306.

CURRICULUM VITAE

Name SURNAME: Yağmur YİĞİT

EDUCATION:

- **B.Sc.:** 2017, Istanbul Aydın University, Engineering Faculty, Mechatronics Engineering, GPA: 3.79/4.0 (2017 first-class honors in Mechatronics Engineering)

PROFESSIONAL EXPERIENCE AND REWARDS:

- Esteem
  - Reviewer in IEEE Transactions on Vehicular Technology (Q1)
  - Reviewer in ELSEVIER Computer Networks (Q1)
  - Reviewer in IEEE Communications Magazine (Q1)
  - Reviewer in ELSEVIER Computer Communications (Q1)
  - Reviewer in PeerJ Computer Science (Q1)
  - Reviewer in International Conference on Science of Cyber Security (SciSec 2023)
  - Reviewer in IEEE International Black Sea Conference on Communications and Networking (IEEE BlackSeaCom 2023)
  - TPC Member in IEEE Globecom (IoT and Sensor Networks Symposium) 2022
  - Session Chair at the Applications of Machine/Deep Learning in IoT and Sensor Networks Symposium in IEEE Globecom 2022
- Teachings
  - BLG374E - Technical Communication for Computer Engineers at ITU (Spring 2022-2023)
  - BLG212E - Microprocessor Systems at ITU (Fall 2022-2023)
- Work Experience
  - Part Time Researcher - BTS Group (2023-Current)
  - Turkcell Teaching Assistant - ITU (2022-Current)
  - Full Time Researcher - Broadband Communication and Network Automation Research Group (BCRG), ITU (2020-Current)

- 5G R&D Engineer - Netas (2020 Feb-2020 Nov)
- R&D Engineer - General Mobile (2017-2019)

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Y. Yigit**, O.K. Kinaci, T.Q. Duong, B. Canberk, 'TwinPot: Digital Twin-assisted Honeypot for Cyber-Secure Smart Seaports,' *in Workshop on the Evolution of Digital Twin Paradigm in Wireless Communications, IEEE International Conference on Communications (ICC 2023)*, Rome, Italy, May 2023.
- **Y. Yigit**, K. Huseynov, H. Ahmadi, B. Canberk, 'Fine-Grained Air Quality Monitoring using Software Defined YANG Data Model,' *4th Workshop on Future of Wireless Access and Sensing for Industrial IoT (FUTUREIIOT), IEEE Global Communications Conference (IEEE GLOBECOM)*, Rio de Janeiro, Brazil, December 2022.
- **Y. Yigit**, B. Bal, A. Karameseoglu, T.Q. Duong, B. Canberk, 'Digital Twin-Enabled Intelligent DDoS Detection Mechanism for Autonomous Core Networks,' *IEEE Communications Standards Magazine*, September 2022, Scopus Journal Quartile: Q1.
- E. Horsanali, **Y. Yigit**, G. Secinti, A. Karameseoglu, B. Canberk, 'Network-Aware AutoML Framework for Software-Defined Sensor Networks,' *2021 17th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2021, pp. 451-457, doi: 10.1109/DCOSS52077.2021.00076.
- **Y. Yigit**, H. Ahmadi, G. Yurdakul, B. Canberk, Trang Hoang, T.Q. Duong, 'Digi-Infrastructure: Digital Twin-enabled Traffic Shaping with Low-Latency for 6G Smart Cities,' *Submitted for Journal Publication*, 2023.
- **Y. Yigit**, L.D. Nguyen, M. Ozdem, O.K. Kinaci, T. Hoang, B. Canberk, T.Q. Duong, 'TwinPort: 5G Drone-assisted Data Collection with Digital Twin for Smart Seaports,' *Submitted for Journal Publication*, 2023.
- M. Akkoc, G. Yurdakul, **Y. Yigit**, A. Karameseoglu, B. Val, B. Canberk, 'Digital Twin-enabled DDoS Detection Mechanism for Autonomous Core Networks,' *Submitted for Patent, Application Number: PCT/TR2022/051216*, 2022.
- K. Huseynov, G. Yurdakul, **Y. Yigit**, B. Canberk, 'Digital Twin-assisted Air Quality Monitoring System and Method,' *Submitted for Patent, Application Number: PCT/TR2022/051669*, 2022.

OTHER PUBLICATIONS, PRESENTATIONS, AND PATENTS:

- M. Dagli, S. Keskin, **Y. Yigit**, A. Kose, 'Resiliency Analysis of ONOS and Opendaylight SDN Controllers Against Switch and Link Failures,' *2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2020, pp. 149-153, doi: 10.1109/ICRCICN50933.2020.9296200.

SOCIETY MEMBERSHIPS:

- IEEE Student Membership (2021-Current)
- IEEE Young Professionals (2021-Current)
- IEEE Communications Society Membership (2021-Current)
- IEEE Women in Engineering Membership (2023-Current)
- IEEE Computer Society Membership (2023-Current)