Advanced Sciences and Technologies for Security Applications

Hamid Jahankhani
David V. Kilpin
Stefan Kendzierskyj *Editors*

Blockchain and Other Emerging Technologies for Digital Business Strategies

Springer

Advanced Sciences and Technologies for Security Applications

Indexed by SCOPUS

The series Advanced Sciences and Technologies for Security Applications comprises interdisciplinary research covering the theory, foundations and domain-specific topics pertaining to security. Publications within the series are peer-reviewed monographs and edited works in the areas of:

- biological and chemical threat recognition and detection (e.g., biosensors, aerosols, forensics)
- crisis and disaster management
- terrorism
- cyber security and secure information systems (e.g., encryption, optical and photonic systems)
- traditional and non-traditional security
- energy, food and resource security
- economic security and securitization (including associated infrastructures)
- transnational crime
- human security and health security
- social, political and psychological aspects of security
- recognition and identification (e.g., optical imaging, biometrics, authentication and verification)
- smart surveillance systems
- applications of theoretical frameworks and methodologies (e.g., grounded theory, complexity, network sciences, modelling and simulation)

Together, the high-quality contributions to this series provide a cross-disciplinary overview of forefront research endeavours aiming to make the world a safer place.

The editors encourage prospective authors to correspond with them in advance of submitting a manuscript. Submission of manuscripts should be made to the Editor-in-Chief or one of the Editors.

More information about this series at https://link.springer.com/bookseries/5540

Hamid Jahankhani · David V. Kilpin ·
Stefan Kendzierskyj
Editors

Blockchain and Other Emerging Technologies for Digital Business Strategies

Springer

Editors
Hamid Jahankhani
Northumbria University London Campus
London, UK

David V. Kilpin
Rushden, UK

Stefan Kendzierskyj
Cyfortis
Worcester Park, UK

ISSN 1613-5113 ISSN 2363-9466 (electronic)
Advanced Sciences and Technologies for Security Applications
ISBN 978-3-030-98224-9 ISBN 978-3-030-98225-6 (eBook)
https://doi.org/10.1007/978-3-030-98225-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Cybersecurity Challenges in Small and Medium Enterprise (SMEs)

Hamid Jahankhani, Lakshmi N. K. Meda, and Mehrdad Samadi

Abstract Over the past decade, digitalization has played a greater role in improving the business stature of small business by opening new venues, extending their reach, and thereby improving value producing opportunities. It also brought the small businesses additional responsibilities of having to deal with the security risks and threats which are ever-present on the digital platform. This chapter aims to review and analyse the cybersecurity risks in small businesses due to the adaption of new digital technologies. The chapter discusses the security risks and challenges dealt by small business, the key constraints in implementing a security program to comply with the legal and regulatory requirements, and how the cloud technology can answer many of those challenges.

Keywords Information governance · SME · Cybersecurity · Security information and event management (SIEM) · Cloud computing · Risk management

1 Introduction

This review chapter is designed to examine the process of SMEs' compliance with cyber security policies.

A small or medium enterprise (*SME*) is a company under independent ownership and operation, and which is not viewed as a leading firm in its sector (SBA, 2015). The American Small Business Administration uses a measure based on the value of the company and size of the workforce to categories the scale of a firm in a given industrial sector SBA (2016), U.S. Censés Bureau (2013).

H. Jahankhani (✉)
Northumbria University London Campus, London, UK
e-mail: Hamid.jahankhani@northumbria.ac.uk

L. N. K. Meda · M. Samadi
Northumbria University London, London, UK

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_1

Fig. 1 Small Business security stature [3]

Information technology has increased the potential of small businesses and created vast wealth-building opportunities. The advent of various new technologies, platforms, applications had catapulted small businesses to the forefront of industrialisation and created excellent markets to showcase their business prowess. Although small businesses have been leading from the front in implementing various new technologies to improve their business, they often fell behind to improve the information security measures relative to these new technologies [1]. SMEs often lacked the resources and security strategies to counter the ever-aggressive cyber criminals. Threats such as phishing, malware, spam, ransomware etc., often led to devastation for small businesses [2]. Figure 1 shows the current security stature of small businesses.

Cybersecurity is generally offered the least priority by small business when conducting day-to-day business. According to Lurey [4], 66% of decision makers in SME's believe an attack to be unlikely, which is in contrary to the reality where 67% experienced a cyberattack. An estimated 1.6 million small businesses fall prey to cyberattacks every year (Nominet UK, 2021). The importance of security can be understated as an attack can be devastating to a small business evident by the fact that 60% of SME's are out of business within six months of an attack [5]. With the usage of smartphones, tablets and IOT devices within SMEs, securing the information and assets is of more importance than ever.

Many a times, small business owners do not have a great understanding of the information risks and information security obligations from an industrial and regulatory standard point of view. In a survey by Osborne [6] on small businesses, one of the key issues has been the lack of knowledge and resources in implementing and handling security processes. The lack of strong cybersecurity solution is evident by the following picture which shows the lack of appropriate security measures in small businesses.

Armenia et al. [7] believes that small enterprises, while not equipped with the necessary skills and knowledge to streamline the cybersecurity process, are still required to handle cybersecurity issues owing to the increased number of cyberthreats. It essential for business owners to look at cybersecurity as an integral part of business and to implement the required security policies, guidelines, and procedures and to streamline them with all the core functions of the business.

Another key constraint for small businesses is the lack of infrastructure, human resources, and the budget to implement a comprehensive IT security system. A survey by UK government [8] revealed that while small business owners are concerned about the cybersecurity risks, many lack the resources to implement a cybersecurity solution, and with merely 33% providing any sort of security training. In a similar study, 81% stated that additional resources and time would benefit to address cybersecurity problems [9]. With 52% of SMBs agreeing to the lack of IT staff within the business [10], and the lack of appropriate resources [11], cyberthreats are an inevitability for small businesses.

One of the latest technologies offering a solution to cybersecurity challenges within small business is the Cloud. Cloud is a multi-discipline technology that offers various types of solutions to small businesses and creates a level playing field to succeed in the business world. Cloud is simply described as computing capabilities accessible over a communication medium, typically, the internet. According to Moskowitz [12], small businesses are more attracted towards cloud solutions. Cloud based computing services, applications and infrastructure solutions benefits the budget and resource constraint small businesses with an ideal way to implement IT solutions. With many "pay-as-you-go or even free solutions" [13], cloud has become an affordable and viable technology for small businesses. Several cloud-based solutions in data storage, hosting services, high-performance infrastructure services, analytical applications, real-time data solutions, support services, e-learning solutions etc., enables the small businesses to focus their resources on business management instead of on procuring infrastructure and people to maintain an effective IT infrastructure.

According to recent studies, the potential of cloud solutions is evident by the growth of SMEs and the fostering business practices in small businesses. Assante et al. [14] suggest that by using cost-effective and integrated cloud solutions, small businesses can enter the markets without any large investments into IT infrastructure. The adoption of cloud-based security solutions can oversee a small business implement the required cyberdefenses strategies to comply with the legal, regulatory requirements and to mitigate any cybersecurity threats. Cloud security solutions provide with the necessary flexibility and scalability to a small business while being cost-effective. Many security solutions such as Cisco Umbrella, Fire Eye's ETP etc., use cloud technology to provide efficient, robust, and up-to-date comprehensive security to any small business without the need for any additional hardware. With most of the cloud security providers offering solutions with multiple security services such as storage, data sharing, backup, hosting etc. as a single setup solution, it makes perfect sense for small businesses to identify an appropriate solution befitting their security needs and to secure their business.

2 Implications for Social Change

SMEs face many challenges, not least their growing need to keep from harm themselves against future cyber-attack.

Gartner [15] noted the growing threat posed by cybercrime, and predicted that the cost to businesses from these attacks would rise to around $6 trillion by 2021. Due to the increase in the number of cyber-attacks, spending on ICT security rose to over $86.4 billion in 2017. The geographical areas of Asia and South America have the highest numbers of recorded cyber-attacks. The highest ranked country by percentage of ICT devices or systems attacked was India (12.87%), followed by Taiwan (9.21%), Malaysia (8.01%) and South Korea (5.56%). Those states with the lowest percentage include Denmark (0.65%), Czech Republic (0.55%), and Finland (0.34%).

Between 2006 and 2020, the USA recorded 156 significant cyber-attacks, which was higher than the total combined number of attacks in the UK, India and Germany (Visual Capitalist, 2021). A cyber-attack is considered "significant" when a government organization, or a company in the defence or high-tech industries is involved and the financial impact exceeds $1 million, Carmen Ang (2021). The consequences for cyber criminals who are apprehended are severe. For example, a convicted cybercriminal in the USA can expect to be given a 20- year prison sentence if the attack involves a government institution and the attack puts national security at risk. However, cybercrime continues to grow around the world, in spite of the identified consequences involved if caught, with particular countries apparently being more at risk of attack than others.

SMEs in general provide approximately 80% of all jobs World Economic Forum (2010), which demonstrates the key importance of such companies to the state economic growth. Despite their importance, many SMEs have failed to adopt adequate advanced cybersecurity systems, and thus their risk of suffering cyber-attack has increased due to the fact that they have automated their ICT processes over recent years Serianu (2017). While larger scale organizations have invested in adequate security measures due to their higher perceived level of risk, SMEs are left at greater risk because they cannot always afford the necessary protective measures, exposing them to greater damage when an attack occurs. SMEs in general face several challenges, especially the high costs of cybersecurity systems, limited funds, and inadequately trained staff. As part of overall business risks, many SMEs have incorporated the issue of cybercrime as part of their risk exposure, with many investing in antivirus systems as part of their security package. However, SMEs lag behind their larger corporate counterparts when it comes to the adoption of cybersecurity systems [16]. Consequently, SMEs need to develop their resilience to cyber-attacks by devising a clear cyber security plan including procedures and instructions for staff. The concept of cyberspace is now an essential feature of contemporary life.

Throughout the 1980s, banks in particular faced cyber threats from those working inside the industry, as well as from the first instances of malware, which was able to replicate itself in ICT systems. Incidents of cybercrime grew rapidly during the

1990s as the Internet became increasingly widely used. During the early 2000s, cyber-attacks against large websites were repulsed for the first time, while other forms of cyber-attack emerged, notably the theft of individual users' data and identity. In the second half of the 2000s, script injection and scripting across sites were employed for the first time, followed by the development of Tor software as well as online black-market platforms such as "Silk Road" in the second decade of the 2000s.

In the 2010s, the amount of cybercrime on the Internet has increased dramatically, particularly that involving ransomware and Nanded [17]. The inherent weaknesses of a company's security system mean that it is at greater risk of cyber-attack, and adequate allocation of both resources and technology are essential to ensure sufficient protection and security Antonescu and Birău [18].

3 Information Security Risk Management

In 2016, it was predicted that by 2021 cybercrime would cost organizations around the world around $6 trillion per year. Moreover, a report entitled Sputnik and produced by Atlas VPN revealed that cybercrime had cost organizations around the world over $1 trillion by 2020. A report by McAfee [19] showed that 20% of the participating organizations had no cyber security plans, consequently many such organizations were vulnerable to cyber-attacks. The impact of cybercrime includes, not only financial losses, but also loss or destruction of data, decreased productivity, the disclose of individuals' personal data, corporate financial data, with the resulting loss of reputation to the organization concerned Cybersecurity Ventures [20]. The issues of cybercrime are likely to increase as the numbers of Internet users are predicted to rise to 6 billion by 2022 and 7.5 billion by 2030 Cybersecurity Ventures [20].

Cybersecurity has become an essential security policy concern all around the world as ICT has become increasingly widely used and e-commerce has grown in scale [21]. It was noted by Ponsard et al. [22] that, due to the rapid increased use of ICT, cyber security is largely not addressed by organizations, while emerging new threats place unprotected organizations at higher risk of cyber-attack. Simultaneously, Watkins [23] found that the cost of implementing cyber security systems has risen due to the rise in the frequency and complex nature of attacks against industrial targets. A study by Klaper and Hovy [24] concluded that governments suffer especially damaging consequences from cyber-attacks, while individual users can learn to protect themselves against the worst consequences if such attacks happens. However, many kinds of sensitive data becomes vulnerable to cyber-attack whenever an individual's or organization's computer system is connected to the Internet. In this way, cybercrime can cause disruption to businesses, and makes customers feel concerned about their own data, leading to damage to the company's reputation.

For businesses in the SME sector, managing cybersecurity remains a significant issue due to the fact that their resources are more limited than established organizations, while they face a similar level of threat [25]. Hayes and Bodhani [26] noted that

SMEs are more vulnerable to cyber-attack largely because their data is under-valued by their management.

The management of ICT security in SMEs can be considered a branch of specific business computing systems which focuses on issues of security arising from information technology Polkowski and Dysarz [27].

According to Twisdale [28], cybersecurity remains a threat for SMEs and any related companies, and the fact that bigger scale firms have begun to address this issue means that SMEs are at even greater risk of cyber-attack. However, SMEs can still make themselves resilient in terms of cyber risk compared to bigger companies even though they have more limited available resources. Henson and Garfield [29] also note the divergence between cyber security practice in SMEs and their larger counterparts, with SMEs still having a great deal to put in place to catch up with the latter.

The management of cyber risk relates to any processes that have been implemented in order to limit the risks posed by cyber-attacks. Cyber-risks are defined as those risks arising from cyber threats. The main cyber risk management factors adopted by SMEs are as follows.

3.1 Information Technology Capability and Investment

Businesses frequently do not recognize the importance of cybersecurity, and the main reasons for this include the issues of time, price and the complex nature of security systems Henson and Garfield [29]. However, studies have indicated that the majority of organizations conduct a range of assessments to check their level of cyber security, including audits, penetration tests and vulnerability checks. Despite this, most firms are unaware of the level of investment they are making in this area, therefore there is room for improving resource allocation when it comes to cyber security issues Research data (2019). Findings by Hills and Atkinson [30] indicated that the resources of many SMEs and their levels of investment in cybersecurity are limited, despite the fact that such businesses increasingly rely on ICT systems and the availability of these [31]. The perception of the benefits of investment in cyber security is limited to that of security according to Kluitenberg [32], while a study by Fielder et al. [33] concluded that due to the shifting nature of cyber threats, investment in security is still a challenge for many companies, even those who have adequate resources at their disposal.

3.2 Management Attitude Toward Security

Studies indicate that most organizations have not yet fully developed their cybersecurity systems, which in many cases are operated by an internal employee as one of their additional duties Research data (2019). However, a nominated person to focus

on this role is essential if an organization is to develop and implement an effective cybersecurity system. Sadok and Bednar [25] note that SMEs should develop their security system by involving their employees and other insider stakeholders in both analysing the risks they face and developing their cyber security policy. In this way, the entire organization backed by support from managers can be fully involved in the development of a cybersecurity procedure Ponsard et al. [22].

3.3 Training and Awareness

According to Topping [34], awareness is an essential element of ICT security systems, although SMEs have not always viewed themselves as being at risk of cyber-attack, and often see security as being expensive and technically complex.

Aldawood and Skinner [35] found that the lack of awareness among employees is a more serious issue than companies' technical weaknesses, and consequently organizations need to focus on training staff about cybersecurity issues. Organizations' main information system users are at high risk from cyber threats, leading to both economic and data losses for individuals, corporate and governmental bodies Nilsen et al. [36], consequently SMEs should prioritise cybersecurity training for their employees Valli et al. [37].

4 Does Cloud Answer the Cybersecurity Challenges in SMEs?

Kurpjuhn [38] believes that the cybersecurity threats to a SME is as significant as a large organisation. The author states the SMEs prioritisation of business resources towards growth rather than on the security and strongly believes that cybersecurity should be a priority regardless of the business size. The authors discuss the SMEs inclination towards Cloud solutions but expresses concerns regarding the Cloud provider security breaches with examples such as Dropbox, iCloud etc.

In a study of Cybersecurity implications in small businesses, Tam et al. [39] states that most of the traditional security solutions currently in use are biased towards large businesses "in terms of scale, cost and usage" and are unsuitable for small businesses, and that more research is needed towards cybersecurity in small businesses. The authors state that the business landscape in SME is different to large organisations in terms of technical requirements and mentions the lack of technical knowledge in business owners. The authors state the potential of Cloud to traditional cybersecurity products and specifies examples such as common network scanning tools which only work in local infrastructure and not on a cloud. While the research is carried out on small businesses within Australia, the research work is true to small businesses all over the world.

In a case study of SMEs offering internet-based services, Lindström et al. [40] stressed the importance of data collection, storage, and communication. In a network of various connected devices, distributed functions, and automations systems, the real challenge according to the authors is not having cybersecurity procedures and functions in the base plan. The authors felt that "cybersecurity should not add unnecessary additional work" to businesses but current security solutions are not adaptable to small businesses which also deters the implementation of a cybersecurity solutions in a small business.

A different perspective of cybersecurity is expressed by Lloyd [41] in which he believes that advantages of cybersecurity should be discussed lot more than the business data breaches, disruption, penalties etc. The author believes that by having a robust cybersecurity programme, small businesses can create products that can achieve substantial business growth. The author clearly states that implementing the best business practices and achieving cyber resilience can be daunting but not insurmountable and is the key for small businesses to thrive in the digital world.

The importance and the benefits of Cloud platforms for SMEs is affirmed by Sultan [42] as far as in 2011. The author stated that virtualisation and grid computing, the fundamental technologies behind Cloud computing, delivers the SMEs with benefits of implementing new technologies into the business and to improve the efficiency of IT resources without any significant investment. The author stressed on the fact that the pay-as-you-go structure of most cloud solutions would be an attractive proposition to SMEs and start-ups. The author did arise concerns of relinquishing control, security and privacy issues, reliability issues etc., but felt that the benefits outweigh the concerns and Cloud can offer better security in real-time when compared to next-to-none security solutions used mostly in small businesses.

The concept of utilising Cloud as a solution for IT security issues and as a reliable way of secure communication was discussed by Zelenay et al. [43]. The authors strongly believed that the lack of resources, infrastructure and knowledge often compromises the information security in a small business. The authors believe that Cloud can become a revolutionary change in the world of IT services and will boost the role of IT in developing businesses. The benefits of Cloud to Traditional IT services were compared and the authors detail the significance of Cloud in terms of costs, speed, security, performance, scalability, and productivity. A list of some sample cloud services offering excellent data security were mentioned.

In an analysis of Information and Communication Technologies (ICT) in small businesses in Tanzania and Poland, Nycz et al. [44] states that the implementation of cloud security solution requires careful planning in choosing a reputable cloud service provider and the required services to maximize the returns for the investment by the business. The authors state the unstable business environment in small businesses due to the business owners not paying enough attention to the application of security codes and practices. The authors developed a hybrid model to secure ICT in Private Cloud Computing and tested the model successfully in a small IT company to prove the efficiency.

Assante et al. [14] pointed out the business improvements and technical efficiency as key prospects for small businesses along with the benefits of low costs, efficient

data movement, agility, and scalability. The authors analysed data from a survey among SMEs from seven different nations and stressed the importance of utilising Cloud technologies to acquire the relevant cybersecurity strategies to stay ahead of competition. The authors briefly stated about the challenges such as reliability, lack of control and security as barriers that SMEs might face while adopting cloud technologies.

A Service Oriented Architecture (SOA) based novel framework for cloud migration is proposed by Nussbaumer and Liu [45]. The authors felt the need for a business to understand its processes and requirements before migrating to the cloud. The authors proposed that the key requirements for small business cloud solution are cost, flexibility, performance, security, reliability, service and support. Even though the authors tested the framework using a business scenario, the developed framework did not include an economical perspective, hence the authors believe that the framework should be complemented with additional work on the economical feasibility before the complete effectiveness can be measured.

One of the key difficulties often faced by the small business is the ability to understand the security requirements of the business and the lack of knowledge to categorize the requirements in a systematic way which is key in choosing a proper cloud security solution. Godfrin [46] developed a cloud search model to identify the security and legal requirements for a business by providing answers to some simple questions as an input to the search. The search accesses a repository to identify a suitable solution which can be contemplated by the business according to their resources and feasibility. The authors developed a repository of several cloud services which are suggested after identifying the functional, non-functional, and legal requirements of a business. The system is sampled by IT professionals and works as it is supposed to, according to the authors.

The regulatory obligations of Cloud providers were discussed by Lovrek et al. [47], The authors discussed about the various types of cloud service architectures and the security concerns of consumers that arise due to the data processing that happens at a different location and believes that a regulatory framework for cloud services will determine the acceptability of the service. The authors stressed the requirement for a balance between strict regulation and complete freedom while keeping the key service conditions of protecting consumer data, quality of service, and possibility for cloud provider change as prime objectives of cloud providers.

5 Cybersecurity Frameworks

5.1 NIST Cybersecurity Framework

The American National Institute of Standards and Technology has created a framework consisting of a core set of cybersecurity policies, stages of implementation and overall profile which can assist many organisations, including SMEs in improving

Fig. 2 Lack of appropriate security measures [49]

their cybersecurity systems NIST [48]. This framework allows for flexibility and thus it can be adopted by companies who are just beginning to build their cyber security response, or by larger scale companies with similar scale budgets NIST [48].

The framework consists of five modules as follows:

- The Identity module leads to the development of a cross-organisational approach regarding awareness and management of cybersecurity risks.
- The Protect module focuses on the infrastructure needed to develop and implement adequate security systems.
- The Detect module helps the organisation to develop adequate means of detecting the origin of cyber-attacks in a short time.
- The Respond module puts in place adequate responses to cyber-attacks which have already been identified.
- The Recovery module allows for future resilience plans to be developed as shown in Fig. 2 (Infused Innovations 2019).

5.2 The Center for Internet Security (CIS) Critical Security Controls

A set of specific controls which can be adopted to prevent cyber-attacks has been developed by a combination of national security, law enforcement, forensic and incident response organizations, and these controls can be used by firms when first developing their cyber security response. Moreover, Gerberding [50] notes that this set of controls is not as comprehensive as alternative cybersecurity frameworks.

5.3 COSO Enterprise Risk Management Framework

The Committee of Sponsoring Organizations of the Treadway Commission (COSO) adopted a framework which enables organizations to develop risk management strategies which consider the constantly changing need to respond to evolving cyber security risks COSO (2018). The framework consists of three concepts, namely the objectives, components and structure of the organisation. Organisations can use the control components to both identify their cyber profile and defend themselves effectively against cyber-attacks Deloitte (2015).

Control environment: identifies the management's assessment of cyber risk within the organization and their degree of awareness of steps already taken to manage this risk. Risk assessment: assesses how much the organization and its stakeholders understand the impact of how any potential cyber-attack might affect their objectives in terms of the operation, reporting and compliance. Control activities: identifies the level of internal risk management which has been written and implemented by the company to address cyber risks, according to their own internal standards. Information and Communication: this involves both internal and external methods of communication and how these are adequately secure in order to minimize cyber risk. Monitoring Activities, identifies existing systems in place for monitoring cyber risk, the controls devised and how any weaknesses are dealt with.

5.4 ISO/IEC 27001

The ISO standard offers organisations a template allowing them to set up, implement, operate, monitor, maintain and improve their Information Security Management System (ISMS). Many SMEs seek to acquire ISO27001 in place of enhance their security of ICT system, while most organizations acquire ISO17001 which shows they comply with a set of regulations and business rules which refer to the security of data.

Acquiring ISO/IEC 27001 requires the organization to complete six steps in a cross-organizational collaborative approach, as shown below:

1. Define security policy.
2. Define and scope the Information Security Management System.
3. Carry out internal risk assessment.
4. Manage identified Risk.
5. Determine the control objectives and controls for implementation.
6. Prepare applicable security statement.

5.5 Italian Cybersecurity Framework

A specific framework has been devised for the Italian market, based on the NIST framework and consisting of the core, profile and implementation stages. It enables businesses to compare their own practices against a baseline set of cybersecurity risks and procedures so that improvements can be made.

Marco and De Luca [51] note that this framework is applicable to a wide range of businesses, including SMEs and based on their individual size and characteristics. Both NIST and the government of Italy were involved in the development of this framework, but it can be applied to businesses in other locations.

5.6 Control Objectives for Information and Related Technologies (COBIT)

The adoption of COBIT offers organisations a framework within which they can develop clear ICT security policies and ensure good practice in this area. It consists of a range of ICT controls leading to the achievement of a well-defined organisational framework for use with ICT management, based on a set of regulations which must be complied with ISACA (2018). There are four key areas which COBIT covers, as described below:

1. **Plan and Organise:** focuses on how technology is used by the organisation and any improvements which may be required to ensure its goals and objectives are met.
2. **Acquire and Implement:** identifies the organization's ICT requirements in terms of obtaining, installing and maintaining the technology based on the company's existing business processes.
3. **Delivery and support:** develop a strategy to manage delivery services.
4. **Monitor and evaluate:** devises an evaluation process to ensure that the ICT system being used by the organization continues to meet the original design objectives and to ensure compliance with any identified regulations.

5.7 Information Technology Infrastructure Library (ITIL) Framework

This framework helps organisations to identify their security needs and put an integrated ICT security strategy in place based on their own requirements. The process begins with a set of non-specific guidelines which organisations can use to devise their own strategic plan, and which is not specific to any particular industry or technology.

Once these baseline requirements have been identified, companies can plan, implement and evaluate their security strategy. The guidelines consist of the following elements:

1. **Service Strategy**: examines the organization's capabilities and identifies how security can be managed in terms of designing, planning and implementing the strategy.
2. **Service Design**: focuses on service management needs in terms of designing developing the necessary strategy.
3. **Service Transition**: develops further strategies to improve capability when new or amended services are put into operational practice.
4. **Service Operation**: focuses on measuring the effectiveness of any support service involved in the strategy in order to guarantee value for the organization and the service provider.
5. **Continual Service Improvement**: develops plans to create and maintain value for customers.

5.8 Information Security for Small and Medium Sized Enterprises (ISSA) 5173 (UK)

In 2011, ISSA 5173 working group sets out recommendations on information security controls for small and medium enterprises. SME are targets of vulnerability to cyber-attacks. They are easy targets for cybercriminals due to limited resources, knowledge and infrastructure. Because SMEs have a fragile structure, they may be destroyed after a cyber-attack and data breach. Larger companies can protect themselves against attacks depending on their budget, but they also depend on the security of SMEs. Cybercriminals try to pave the way for larger companies through SMEs that have lower security. As a result, the security of SMEs is one of the major socio-economic challenges. The document published by ISSA, aimed at helping to provide an appropriate level of security for SMESs.

6 The Most Common Types of Cyber-Attacks in Businesses

According to Borna news Borna News (2021), the rapid rise of ransom cyber- attacks shows that cyber risk is not limited to one sector or industry in the economy and is becoming a financial and security threat worldwide. Cases of cyber-attacks are expected to become more widespread and complex, as the risk of hackers being caught is relatively low and the benefits of these illegal operations are high. According to the Fitch Ratings Borna News (2021), the increase and intensification of attacks is a negative factor in the credit rating of companies.

During 2020, the number of cyberattacks aimed at extortion worldwide increased by 485%, and ransomware attack was one-quarter of the total cyberattacks that took

place in the previous year. On the other hand, it is estimated that the cost of these operations was $20 billion. The number of extortionate cyberattacks in which the victim company is threatened with disclosure of stolen data is also on the rise; 77% of all attacks happened in the first three months of this year.

The increasing number of such attacks has increased the costs imposed on victims. According to the Xavier Institute Borna News (2021), the average ransom demanded in cyber-attacks was $220,000 in the first three months of this year. This shows a growth of 43% compared to the fourth quarter of 2019. New cases of cyber-attacks could add to the international efforts of governments and the private sector to prepare for such attacks.

Companies that do not have advanced and up-to-date security networks and systems are more vulnerable to cyber-attacks than other companies. However, the risks of falling victim to a cyber-attack are much higher for large and influential organisations. Hackers attack all sectors of the economy, but some parts are more attractive to them than others.

Speciality service companies such as small law firms and financial services companies are attractive targets for hackers because of their high vulnerability. Cyber-attacks against schools and health care providers doubled last year to 2354 Borna News (2021). The critical point is that the ransom does not guarantee hackers will give the stolen files to the victim or avoid publishing them, and paying a ransom can put a financial services company at greater financial risk.

Governments and corporations must believe in the growing threat posed by cyber-attacks and take the necessary measures to counter them. Cyber-attacks of any kind have a significant impact on companies and the economy, although they cannot be stopped entirely. However, their risk must be mitigated, and the motivation of hackers must also be reduced as much as possible.

The biggest unforeseen problem for small business owners in the last decade were tax matters, while today, with the growth of businesses, it can be said that the threats posed by cyber-attacks have replaced this. In its 2019 Global Risk Report, the World Economic Forum lists data misuse and cyber threats as the fourth and fifth most serious risks facing businesses around the world, respectively.

Whether a small business or a technology giant company, every company has vulnerabilities that hackers can exploit in today's world. On the other hand, hackers can target any business, regardless of its size. Start-ups and small and medium-sized businesses are more vulnerable to cyber-attacks than other businesses due to their cohesive and small structure.

According to the Forbes website Myba (2020), ransomware attacks, especially those involving RYUK which targets companies and institutions, was set to increase by 300% in 2020, with most of these attacks focusing on small businesses in the United States. Smaller firms are more vulnerable because they do not have sufficient resources to build a robust cyber security infrastructure and need to pay more attention to security guidelines and protocols.

Cybercrime has changed dramatically over the past few years and has spread to such an extent that a small or medium-sized enterprise alone does not have the knowledge and ability to deal with it. Today, hackers also start a business and sell

their anti-cyber tools and expertise in various packages and services to lower-level hackers, encouraging the growth and spread of attacks. Phishing attacks, denial of service (DoS), ransomware, and malware are common attacks that SMEs should be aware of.

- *Phishing Attacks*: A type of social engineering-based cyberattack in which hackers use disguise methods, such as fake bank portal sites or site payment pages, to access and steal users' information.
- *DDoS attacks*: Attempts to permanently or temporarily interrupt a company's online services, including sites and mobile applications.
- *Ransomware attacks*: In this type of attack, all or part of the company's critical data is stolen or encrypted by complex algorithms. The hacker then demands large sums of money to release and hand over the data key.
- *Malware Attacks*: In this type of attack, hackers inject malicious software into the operating system in various ways, thereby providing a platform for ransom or damage.

7 Security Information and Event Management (SIEM)

According to Raja et al. [52], SIEM (Security Information and Event Management) is known as the combination of different Security Event Management and Security Information Management. The major role of SIEM is to collects data and information of events from different devices and arrange them into a common format. All the events are gathered and collected for analysing the behaviour and functioning of the entire system. The analysis and monitoring of the single sources of the event can help in detecting the events of attacks such as Probe/DoS. This method is utilized for analysing the flood attacks of SYN TCP through the application of the RETE algorithm on the network. In this system, an alert alarm will sound in the case of a TCP SYN attack commitment.

Vielberth et al. [53], stated that most large and medium business organizations are utilizing the SIEM in their Centres of Security Operations so that they can get a higher level of awareness regarding cyber security practices. The utilization of SIEM will let the organizations collect and analyze the different information related to security measures in a centralized manner. This will help them in enhancing the process of threat detection and increase the reaction time to any of accidents.

As per the research conducted by Corcoran (2018), it has been mentioned that attack frequency has increased and as a result different business organizations are getting targeted by new and different attackers, hacking with more sophisticated attacking tools.

The attacks are being done on various levels in the organization's OSI models. So according to Corcoran (2018), the different business organizations need to implement the multiple layers and protocols of cyber security in order to avoid the issues and address them more effectively. But this implementation does also have a limitation that all the data and information collected in this can be mixed and it will

become difficult to correlate the different facts for deriving better results. In order to improve this, the organization need to encourage the use of Security Information and Event Management in their cyber security practices. The implementation of Security Information and Event Management will help the different business organizations in segregating all the information and filter it according to the utilization for different purposes. The utilization of Security Information and Event Management will facilitate the organization with centralized storage space for all the information related to the cyber security of the organization, for improving the different systems that are required in the organization for enhancing the security of information. Security Information and Event Management will help the business organization to defend the different vulnerabilities and attacks on their data and information. In this research, the main emphasis is provided on the benefits that are gained by the organization on adopting the SIEM and its implementation process for achieving the appropriate and effective results. Moreover, a proper explanation of the different recommendations that can help the organization in improving the process will also be provided through this research.

8 Conclusions

The importance of information security in any business cannot be understated as security risks can become threatening to the existence of the business itself if not dealt accordingly. The security risks to small businesses are no less than a large organisation and the consequences are even more damaging to a small business. The lack of required knowledge, infrastructure, security personnel, and resources often turn small businesses highly vulnerable to cyberattacks. Cloud based security solutions provides the cash-stricken small businesses with the necessary comprehensive security services to handle day-to-day threats and to secure the business from outside threats. The key characteristics of cloud security solutions such as flexibility, scalability, cost-effective, less technical, and ease of deployment often prove to be a great combination to secure a small business. The compliance of cloud solutions in accordance with GDPR, Data Protection Act, and other regulatory standards will be carried out.

References

1. Help Net Security (2021) What are the most common cybersecurity challenges SMEs face today?—Help Net Security. Help Net Security. Available at: https://www.helpnetsecurity.com/2021/07/07/smes-cybersecurity-challenges/. Accessed 5 Aug 2021
2. Witts J (2021) The top 5 biggest cyber security threats that small businesses face and how to stop them. Expert Insights. Expert Insights. Available at: https://expertinsights.com/insights/the-top-5-biggest-cyber-security-threats-that-small-businesses-face-and-how-to-stop-them/. Accessed 9 Aug 2021

3. Yazbeck E (2021) When it comes to Cybersecurity, the small and medium business community needs to do better. SMC Consulting. Available at: https://www.smcconsulting.be/when-it-comes-to-cybersecurity-the-small-and-medium-business-community-needs-to-do-better/. Accessed 15 Aug 2021
4. Lurey C (2019) Cyber mindset exposed: keeper unveils its 2019 SMB cyberthreat study—keeper security blog—cybersecurity news & product updates. Keeper Security Blog. Available at: https://www.keepersecurity.com/blog/2019/07/24/cyber-mindset-exposed-keeper-unveils-its-2019-smb-cyberthreat-study/. Accessed 26 July 2021
5. Galvin J (2018) 60 Percent of small businesses fold within 6 months of a cyber attack. Here's How to Protect Yourself. Inc.com. Available at: https://www.inc.com/joe-galvin/60-percent-of-small-businesses-fold-within-6-months-of-a-cyber-attack-heres-how-to-protect-yourself.html. Accessed 28 July 2021
6. Osborne E (2015) Business versus Technology: sources of the perceived lack of cyber security in SMEs (Working Paper). Oxford University Research Archive, p 10. Available at: https://ora.ox.ac.uk/objects/uuid:4363144b-5667-4fdd-8cd3-b8e35436107e/download_file?file_format=pdf&safe_filename=01-15.pdf&type_of_work=Working+paper. Accessed 6 Aug 2021
7. Armenia S, Angelini M, Nonino F, Palombi G, Schlitzer M (2021) A dynamic simulation approach to support the evaluation of cyber risks and security investments in SMEs. Decis Support Syst 147:113580. https://doi.org/10.1016/j.dss.2021.113580. Accessed 8 Aug 2021
8. UK government (2020) https://www.gov.uk/government/statistics/cyber-securitybreaches-survey-2020/cyber-security-breaches-survey-2020
9. Gough O (2016) Majority of businesses neglecting cybersecurity due to lack of resources. Small Business. Available at: https://smallbusiness.co.uk/majority-businesses-neglecting-cybersecurity-2535173/. Accessed 10 Aug 2021
10. Umawing J (2019) SMBs lack resources to defend against cyberattacks, plus pay more in the aftermath—Malwarebytes Labs. Malwarebytes Labs. Available at: https://blog.malwarebytes.com/business-2/2019/10/smbs-lack-resources-to-defend-against-cyberattacks-plus-pay-more-in-the-aftermath/. Accessed 9 Aug 2021
11. Benz M, Chatterjee D (2020) Calculated risk? A cybersecurity evaluation tool for SMEs. Bus Horizons 63(4):531–540. https://doi.org/10.1016/j.bushor.2020.03.010. Accessed 7 Aug 2021
12. Moskowitz S (2017) The small and medium-sized enterprise (SME). Cybercrime and Business, pp 45–68. https://doi.org/10.1016/B978-0-12-800353-4.00004-X. Accessed 6 Aug 2021
13. Ricci R, Battaglia D, Neirotti P (2021) External knowledge search, opportunity recognition and industry 4.0 adoption in SMEs. Int J Prod Econ 240:108234. https://doi.org/10.1016/j.ijpe.2021.108234. Accessed 12 Aug 2021
14. Assante D, Castro M, Hamburg I, Martin S (2016) The use of cloud computing in SMEs. Procedia Comput Sci 83:1207–1212. https://doi.org/10.1016/j.procs.2016.04.250. Accessed 10 Aug 2021
15. Gartner (2017) Business impact of security incidents and evolving regulations driving market growth
16. Verbano C, Venturini K (2013) Managing risks in SMEs: a literature review and research agenda. J Technol Manag Innov 8(3):186–197. https://doi.org/10.4067/S0718-27242013000400017
17. Pathak PB, Nanded YM (2016) A dangerous trend of cybercrime: ransomware growing challenge. Int J Adv Res Comput Eng Technol 5(2):371–373
18. Antonescu M, Birău R (2015) Financial and non-financial implications of cybercrimes in emerging countries. Procedia Econ Finance 32:618–621
19. McAfee (2018) Economic impact of cybercrime—no slowing Dow
20. Cyber Security Ventures (2017) 2017 Cybercrime Report
21. Kaur S, Sharma S, Singh A (2015) Cyber security: attacks, implications and legitimations across the globe. Int J Comput Appl 114(6)
22. Ponsard C, Grandclaudon J, Dallons G (2018) Towards a cyber security label for SMEs: a European perspective. In: ICISSP, pp 426–431

23. Watkins B (2014) The impact of cyber attacks on the private sector.no. August, 1-1. Whetten DA (1989) What constitutes a theoretical contribution? Acad Manage Rev 14(4):490–495. The framework outlines 7 points which you can use to evaluate your research work.
24. Klaper D, Hovy E (2014) A taxonomy and a knowledge portal for cybersecurity. In: Proceedings of the 15th annual international conference on digital government research. ACM, pp 79–85
25. Sadok M, Bednar PM (2016) Information security management in SMEs: Beyond the IT Challenges. In: HAISA, pp 209–219
26. Hayes J, Bodhani A (2013) Cyber security: small firms under fire (Information Technology Professionalism). Eng Technol 8(6):80–83
27. Polkowski Z, Dysarz J (2017) It security management in small and medium enterprises. Sci Bull-Econ Sci 16(3):134–148
28. Twisdale JA (2018) Exploring SME vulnerabilities to cyber-criminal activities through employee behavior and internet access (Doctoral dissertation, Walden University)
29. Henson R, Garfield J (2016) What attitude changes are needed to cause SMEs to take a strategic approach to information security? Athens J Bus Econ 2(3):303–318
30. Hills M, Atkinson L (2016) Towards cyber-resilient & sustainable SMES: the case study of added value from a large IT reseller
31. Santos-Olmo A, Sánchez L, Caballero I, Camacho S, Fernandez-Medina E (2016) The importance of the security culture in SMEs as regards the correct management of the security of their assets. Future Internet 8(3):30
32. Kluitenberg H (2014) Security risk management in it small and medium enterprises. In: Proceedings of 20th Twente student conference on IT
33. Fielder A, König S, Panaousis E, Schauer S, Rass S (2018) Risk assessment uncertainties in cybersecurity investments. Games 9(2):34
34. Topping C (2017) The role of awareness in adoption of government cyber security initiatives: a study of SMEs in the UK
35. Aldawood H, Skinner G (2018) Educating and raising awareness on cyber security social engineering: a literature review. In: 2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE). IEEE, pp 62–68
36. Nilsen R, Levy Y, Terrell S, Beyer D (2017) A developmental study on assessing the cybersecurity competency of organizational information system users
37. Valli C, Martinus IC, Johnstone MN (2014) Small to medium enterprise cyber security awareness: an initial survey of Western Australian business
38. Kurpjuhn T (2015) The SME security challenge. Comput Fraud Secur 2015(3):5–7. https://doi.org/10.1016/S1361-3723(15)30017-8. Accessed 2 Aug 2021
39. Tam T, Rao A, Hall J (2021) The good, the bad and the missing: a narrative review of cybersecurity implications for Australian small businesses. Comput Secur 109:102385. https://doi.org/10.1016/j.cose.2021.102385. Accessed 2 Aug 2021
40. Lindström J, Eliasson J, Hermansson A, Blomstedt F, Kyösti P (2018) Cybersecurity level in IPS 2: a case study of two industrial internet-based SME offerings. Procedia CIRP 73:222–227. https://doi.org/10.1016/j.procir.2018.03.302. Accessed 11 Aug 2021
41. Lloyd G (2020) The business benefits of cyber security for SMEs. Comput Fraud Secur 2020(2):14–17. https://doi.org/10.1016/S1361-3723(20)30019-1. Accessed 18 Aug 2021
42. Sultan N (2011) Reaching for the "cloud": How SMEs can manage. Int J Inf Manage 31(3):272–278. https://doi.org/10.1016/j.ijinfomgt.2010.08.001. Accessed 6 Aug 2021
43. Zelenay J, Balco P, Greguš M (2019) Cloud technologies—solution for secure communication and collaboration. Procedia Comput Sci 151:567–574. https://doi.org/10.1016/j.procs.2019.04.076. Accessed 4 Aug 2021
44. Nycz M, Martin MJ, Polkowski Z (2015) In: 2015 7th International conference on electronics, computers and artificial intelligence (ECAI). IEEE, Bucharest. https://doi.org/10.1109/ECAI.2015.7301182. Accessed 19 Aug 2021
45. Nussbaumer N, Liu X (2013) Cloud migration for SMEs in a service oriented approach. In: 2013 IEEE 37th annual computer software and applications conference workshops. IEEE. https://doi.org/10.1109/COMPSACW.2013.71. Accessed 16 Aug 2021

46. Godfrin (2016) Legal requirements and identifying data security for cloud service. In: 2016 Second international conference on science technology engineering and management (ICONSTEM). Chennai: IEEE. https://doi.org/10.1109/ICONSTEM.2016.7560948. Accessed 19 Aug 2021
47. Lovrek I, Lovrić T, Lucic DL (2012) Regulatory aspects of cloud computing. In: SoftCOM 2012, 20th international conference on software, telecommunications and computer networks. IEEE. Available at: https://ieeexplore.ieee.org/document/6347661/authors#authors. Accessed 11 Aug 2021
48. NIST (2018) Framework for Improving Critical Infrastructure Cybersecurity
49. Owen-Jackson C (2021) How to protect your small business from cyber-threats. Secure Futures. Available at: https://www.kaspersky.com/blog/secure-futures-magazine/small-business-cybersecurity/29177/. Accessed 25 Aug 2021
50. Gerberding K (2017) NIST, CIS/SANS 20, ISO 27001—simplifying security control assessment
51. Marco B, De Luca R (2015) Financial distress and earnings manipulation: evidence from Italian SMEs. J Acc Finance. Available at SSRN: https://ssrn.com/abstract=2596295
52. Raja MSN, Vasudevan AR (2017) Rule generation for TCP SYN flood attack in SIEM Environment. Procedia Comput Sci 115:580–587. https://doi.org/10.1016/j.procs.2017.09.117
53. Vielberth M, Pernul G (2018) A security information and event management pattern. In: 12th Latin American conference on pattern languages of programs, vol 1, no 1, pp 1–12

Artificial Intelligence Based Malicious Traffic Detection

Lakshmi N. K. Meda and Hamid Jahankhani

Abstract Cyberattacks have become a nightmare for businesses, often having to spend time and resources identifying one or mitigating another. The current research is an effort to develop an artificial intelligence-based security solution that can meet the SME demands of providing a security solution capable of detecting cyberattacks in real-time before they eventually become a crisis for the business. The proposed solution uses multiple layers of deep neural networks using ReLu activation function and Adam algorithm as an optimizer to provide the detection capabilities. While Wireshark provides the model with powerful network monitoring capabilities, Weka fulfils the data pre-processing role to provide the AI module with a clean and structured dataset. The solution is tested for its ability to study the network patterns and capability to distinguish between the regular and malicious traffic. The proposed model is tested using two different datasets, a dataset created in a virtual lab environment, and an IoT-23 dataset. The performance of the proposed AI model is tested on the metrics of 'accuracy' and 'loss'. The model performed well in distinguishing the network traffic on both the datasets. The model will provide the required augmentation capabilities for SMEs to better handle cyber threats.

Keywords Deep neural networks · Rectified linear unit · ReLu · AI · Cyberattack · Adam algorithm · Weka

1 Introduction

Cyberattacks are a buzzword in the world of IT, and to maintain a safe and secure IT infrastructure is often complex, challenging, and requires time, resources, and expert knowledge, all of which are at a premium for Small and Medium Enterprise's (SME). With a natural shortage of resources and dedicated IT personnel, SMEs typically lack

L. N. K. Meda
Northumbria University, London, UK

H. Jahankhani (✉)
Northumbria University London Campus, London, UK
e-mail: Hamid.jahankhani@northumbria.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_2

continuous network monitoring leading to a delayed response to security threats. A clear reverse trend is seen in the recent times, as IT Security became a focus area for SMEs. More and more businesses are investing resources in the security domain, but with the current market flooded with variety of security tools, applications, and solutions to deal with various concepts of security, implementing a robust security solution is often unfeasible. To maintain an enterprise grade security, SMEs had to invest in multiple security solutions and still fails to identify an attack.

A network architecture in a business typically involves multiple end user devices surrounding one or more server machines and network devices. The security function is often handled by the individual end-user devices with the network devices often using the pre-configured or default settings. While some businesses use a complete security solution that handles all the devices within the network, most of the networks lack continuous monitoring and had to rely on the alerts produced by the security products. Without a dedicated security personnel monitoring the network activities, these networks infrastructures are prone to malicious attacks. The damage caused by these attacks have more impact on the small and medium businesses than large organisations with 60% of SMEs often close their business within 6 months following an attack [15]. The research aims to design an AI model capable of predicting network attacks by continuously monitoring the network for anomalies. By utilising the deep learning capabilities of AI, the proposed model monitors the network activity and provide an alert when the traffic matches any previously learnt patterns.

2 Literature Review

Artificial Intelligence has the potential to become a revolution in solving problems for a variety of sectors such as healthcare, transport, agriculture, marketing, banking, finance etc. According to Marr [21], the influence of AI in augmenting human capabilities can be seen in all facets of human life. “Artificial Intelligence (AI) is going to change the world more than anything in the history of mankind. More than electricity”, says Dr. Kai-Fu Lee [28].

The importance of artificial intelligence in the field of Cybersecurity has been discussed by Kurpjuhn [19] who emphasises the importance of adding ‘intelligent security capabilities’ to a business’s security infrastructure to counter the ‘proliferating malware and cyberthreats’. The author strongly believes that understanding the vital components that makeup a reliable and robust security solution is key to effectively integrate AI into business IT security rather than mere automation of the security tasks. The author explained about how traditional sandboxing can be boosted by adding AI, to continually analyze and learn from the network traffic, thereby providing robustness to the entire security infrastructure. The author ascertains that importance of up-to-date information that needs to be consistently fed to the security system to maintain ground against zero-day threats which could only be possible by using AI based security solutions. While the author did not provide any practical implementation, he strongly maintained a strong belief that the efficiency

of security solutions in future will largely depend upon the effective integration of artificial intelligence.

The efficiency and need for the use of AI in cybersecurity is discussed by Hofstetter et al. [13]. The authors proposed a hybrid approach of using AI and machine learning to detect malicious activities within a business network. Unlike a fully automated AI solution, the hybrid system adds human input to provide the necessary insights, which according to the authors, not only improves the accuracy and performance but also adds a 'policing aspect' which helps with accountability. A two-phase approach is discussed with (i) phase one dedicated to developing a data-tree with the application of several learning models and human input and (ii) online testing phase two in which the system will provide with a suitable detection plan for a business based on the data-tree. Even though the authors solution is novel, it requires a constant input from human experts to train the system and the accuracy of the system entirely depends on the data-tree which is based on the learning models.

According to Chan et al. [8], the three main benefits of AI in cybersecurity are the detection of false positives, predictive analysis, and the development of an immune system. With the use of Neural Networks and Expert Systems, the authors states that a sophisticated system that can function like human brain while finding solutions from learning models and past data, can be developed. The authors discussed the case studies of Illumio, Blue Talon identify the key requirements of data protection and access and to explain the benefits of using artificial intelligence in such business scenarios. While the AI based cybersecurity systems currently being used are not entirely failsafe with potential ethical concerns, lack of expertise in unsupervised learning etc., the authors state that use of AI in cybersecurity will attain a state where they will be widely accepted and will be more applicable and accessible too.

More thorough research on the capabilities of AI in cybersecurity was put forward by Zeadally et al. [36]. The authors attributed the dire need for improving the current cybersecurity solutions owing to the lack of cyber governance skills, harness the potential of new technologies and fragmented cybersecurity frameworks. The authors discussed the traditional cyberthreats and the legacy security solutions being used to mitigate those threats. The concept of machine learning techniques such as Naïve Bayes, Decision trees, K-nearest neighbours (k-NN), Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), Self-Organising Maps (SOMs) etc. were discussed in-depth about their suitability for use in cybersecurity. The appropriateness and application of the above techniques at various domains such as Internet, Botnets, IOTs, critical infrastructure etc. were critically discussed to explain how the potential of AI in cybersecurity. The authors believe that with further advances in the technology, AI can provide innovative solutions to detect and mitigate sophisticated cyberthreats. The authors statement that AI based machines thinking humanly in future is not overstated.

A framework capable of automatically analysing the patterns of a cyberattack to identify a potential threat to critical infrastructure is proposed as far back as in 2008 by Flammini et al. [11]. The DETECT (Decision Triggering Event Composer and Tracker) framework, developed by the authors allows for triggering a focussed and fully automated response when a threat is detected. By training the system using the

experts knowledge base, improved probability of detection (POD) and situational awareness can be achieved while reducing false positive rates. The authors used Java programming to develop Event Trees in Scenario GUI to detect a threat and trigger a predefined automated response and proved the efficacy of the proposed system using a subway terrorist attack example. While the system appears to be efficient, the practicality in terms of critical infrastructure solely depends on the amount of data that can be analysed at a given point of time and the depth of the security knowledge base that is fed into the framework.

The impact of using Automated Decision Systems (ADS) in cybersecurity was analysed in-depth by Chamberlain et al. [7]. They analysed the ADS systems based on the decision, autonomy, and impact levels. The authors provided comparisons between autonomous systems, AI, and other new technologies such as data science, and states that while AI based systems works well with specialized tasks, they do need human insights and skillset to identify the broader impact and risk. According to the authors, the current level of AI decision systems works primarily based on the analytical parameters, while humans can provide the necessary intuition while making critical decisions. The authors used Stacey's complexity model to identify the relationship between AI and autonomous systems, decisions by humans and machines and considers that AI can augment humans but not completely replace them in the decision making.

A contrasting research of how Artificial Intelligence and Machine Learning can be used both offensively and defensively is carried out by Kamoun et al. [16]. The authors briefly discussed the use of AI/MLS as a defensive mechanism for intrusion detection, network traffic identification, threat detection, digital forensics, botnet and spam detection etc., while the main focus was placed on the offensive uses of AI. By using a System-Fault-Risk (SFR) framework, the authors categorized the cyberattacks as Probe, Scan, Spoof, Flood, Misdirect, Execute and Bypass, and discussed how AI/MLS systems can be used to carry out offensive activities. The authors expressed recommendation that the misuse of AI/MLS models should be proactively anticipated and be considered while developing future solutions. The research discussed the potential of AI systems being used in distinctive sides but lack any practical implementations or recommendations.

Sapavath et al. [30] developed an AI based model to detect cyberthreats and conducted practical tests and evaluations in a virtualized wireless network. The authors used Bayesian network model to develop the test network and used three different data sets to train the model to ensure accuracy of the threat detection. The transmitting power of a user device is taken into consideration to differentiate a malicious user from legitimate users. Whenever the transmitted power is higher than the predefined threshold, the system will generate an alarm and will trigger automated actions to investigate the activity from the specific user. The parameters of Accuracy Score, Recall, Precision, MCC, False-positive rate and G-mean were used to analyze the results. The authors conducted multiple tests and achieved an accuracy of 99.8% in detecting a cyberattack and the proposed model. They compared the results with Deep Neural Network (DNN) and Random Forest models and concluded that their proposed model performed better and used less learning time and achieved better

accuracy levels. The proposed model presents a great start but the performance in a real network needs to be evaluated along with other input parameters alongside the transmitting power.

Thaseen et al. [33] proposed a network traffic classification mechanism without performing decryption of data packets. The proposed approach captures network traffic using Wireshark tool to create a dataset and performs data pre-processing using Weka tool to extract the classification factors necessary for the application of machine learning algorithms. By implementing multiple machine learning algorithms such as NB (naïve bayes), RF (random forest), KNN (KNearest Neighbors) and SVM (support vector machine) the model is tested for predicting and analysis of malicious traffic. The results proved that RF algorithm has the best accuracy in identifying a packet as one of a normal, malicious, normal encrypted or malicious encrypted packet. The authors future work consists of using deep learning models to improve the general performance and classification of network traffic.

One of the comprehensive reviews of using Artificial Intelligence (Deep Learning Methods) in cybersecurity is carried out by Macas and Wu [20]. The authors stated that Deep Learning (DL) models can be highly effective in solving complex cybersecurity problems and set out to review various Deep Learning models while proposing a DL framework of their own. The authors discussed how DL techniques such as RNNs, CNNs, DBNs, DBM, AEs, and GANs can aid in intrusion detection, malware analysis, IoT defence systems etc. The authors proposed a framework with a series of steps involving data collection, pre-processing, feature extraction, evaluation, training and validation, and model selection. The authors state that the input data quality and feature selection play a key role in developing predictors which is vital for machine learning. By dividing the data set into training, validation and a test set, multiple DL models can be tested with the data set with multiple algorithms and specific parameters. The validation set accuracy will help in identifying the ideal model for the network. The chosen model will then be trained using all the available data with the best possible parameters, and will be periodically evaluated to ensure the system can predict zero day attacks. The authors did a commendable work in evaluating various DL models but left several key aspects of data collection, feature extraction, model suitability etc. to the experts who plan to use the framework.

Another key research in identifying the information required to explain the decisions made by AI systems is conducted by Jaigirdar et al. [14]. The authors stressed the importance of identifying the factors that led the AI system to reach a decision, to ensure the system's transparency and accountability. The authors tried to answer four key questions of (i) What information is required to understand a decision made by the AI system, (ii) possibilities of checking solutions accuracy and the decision-making steps to ensure transparency, (iii) adding security-based evidence to detect 'misrepresentation, deliberate bias, safety-mismatch or backdoor', (iv) possibilities of adding parameters to ensure and justify ethical and policy issues. The authors felt the need for inducing a governance and policy factors into AI based systems to ensure complete transparency of the decision making. The authors used PROV-DM data model as a base and developed a 'Six Ws' framework to identify the attributes required to present explainable AI properties. The authors put the framework into

action by using a sample scenario of bank loan processing to find the justification to the decision made by the AI based decision making system. The authors stressed the need for linking the input parameters to the output decisions and deriving an explanation as to why a particular decision is made. The authors came to a conclusion that it is essential to maintain the 'transparency and explainability' in AI systems along with the addition of security and legal attributes and the proposed framework could act as a steppingstone.

An open-ended review and research on identifying the capabilities of intelligent attack detection techniques using artificial intelligence is conducted by Aljabri et al. [3]. The authors started by identifying several classic network attack classification techniques such as port-based, payload-based, Deep packet inspection (DPI), behavioural techniques, rationale-based, Bag of Flow (BoF) etc., all of which depended primarily on the database of pre-defined attack signatures to detect an attack and their shortcomings due to lack of intelligence. The authors prime the importance of intelligent techniques such as machine learning (ML) and Deep learning (DL) and their statistical capabilities in learning and analysing network traffic, thereby identifying network anomalies at a much faster rate compared to non-intelligent techniques. According to the authors, the ability to learn attack scenarios and patterns from a wide variety of sources to train these intelligent systems makes them promising for the future. The authors analysed the current research in Intelligent systems for network attack mitigation using wide variety of techniques based on logistic regression (LR), random forest (RF), decision trees (DT), ensemble of DT, support vector machine (SVM), naïve Bayes (NB), K-nearest neighbor (KNN), K-means clustering etc. The authors identified the trend of using the up-to-date and advanced Machine Learning and Deep Learning techniques of artificial neural network (ANN), recurrent neural network (RNN), convolutional neural network (CNN) and deep neural network (DNN), in network traffic analysis and threat mitigation. The authors categorised all the current research according to the threat type and the possible futuristic solutions. Another important part of the research by Aljabri et al. was to identify the current research trends based on the various datasets available for researchers. The authors stressed the complexity of identifying the right technique along with the right dataset to develop a highly accurate model and agreed that finding a model that works for all types of threats could be a silver bullet if at all one could develop one.

A flow-data based approach to detect and classify malicious network traffic is proposed by Abuadlla et al. [1]. The ease of capturing data from any network device and the scalability offered by aggregated traffic metrics, biased the authors to use flow-data for their proposed model. In a critical anomaly detection phase, malicious traffic is differentiated from normal network traffic and in stage two, the detection and classification module identifies the attack characteristics and classifies the type of attack. To test the proposed IDS system two different neural network methods are chosen, MLFF (Multilayer Feedforward neural network) and RBFN (Radial Basis Function Network). By using multiple training algorithms (Radial Basis Function net, Levenberg-Marquardt and Resilient Backpropagation) the authors were able to achieve a detection rate of 94.2% in anomaly detection phase and a detection rate of 99.42% in classification phase. The model was able to achieve 100% in detecting

DoS and land attacks, and 99.9% in port-scan attacks which shows better performance compared to similar models using large datasets. Although falling behind the multifunction feedforward neural networks in its classification abilities, RBFN neural networks were more suited to real-time networks owing to its simple architecture and hybrid learning capabilities. The authors plan to continue the research of developing a more accurate model with minimal features and less training time.

Yuan et al. [35] sought the use of advanced deep learning techniques to counter DDoS, one of the most harmful network attacks. Called DeepDefense, the authors developed a recurrent neural network learn patterns from the network traffic and to automatically extract high-level features to achieve powerful representation and trace network attack activities. Using the ISCX2012 dataset, the authors selected multiple numerical, Boolean and text fields and used binary, BoW conversions techniques to extract the required features. By designing a Bidirectional Recurrent Neural network and comparing the selected traffic fields against attack vectors the authors were able to differentiate malicious traffic. The authors tested the DeepDefense approach using different RNN models (LSTM, CNNLSTM, GRU, 3LSTM) and concluded that the DeepDefense approach is highly capable in learning from historic network traffic, effective in detecting DDoS and offers better performance in terms of generalization compared to shallow ML models. The use of CNN, RNN, Long Short-Term Memory Neural Network (LSTM) and Gated Recurrent Unit Neural Network (GRU) proved beneficial while training large datasets and helped reduce the error rates by 39 as much as.69% in some scenarios.

Mohammad et al. [24] conducted several experiments using artificial neural networks to identify phishing attacks. Identifying a phishing attack is extremely complicated given the dynamic nature of the websites and hence requires a model that necessitates a constant improvement in the prediction capabilities. The authors automated the network structuring process and conducted experiments that showed resilience against noisy data and fault tolerance while achieving high accuracy. The authors used 17 different features such as ip address, long URL, server form handler, DNS record, age of domain etc., to represent the neurons from a dataset of 1400 phishing and legitimate websites. The research is based on the principle that Phishing detection is a classification problem, and the selected fields were given the values of either “phishy” or “legitimate”. The authors used a neural network with one hidden layer called multi-layered perceptron, in which the number of neurons can be changed to adapt to the complicated relationship between the input and output variables. Although a bit complex, the authors firmly believe that their model will automate the network restructuring process with fewer user inputs and can be adapted to any future updates.

A focus on mitigating zero-day threats using machine learning is carried out by Beaver et al. [6] who states that the ability of machine learning tools to accommodate complex and huge data sizes facilitates a strong combination of data analysis and human augmentation. In the work, they used AdaBoost (adaptive boost) ensemble learner to reliably differentiate malicious network traffic, in a simulation with network settings that mimics a real-time operational network. The model was tested with four levels of decision-making: (1) top-level wrapper placing a cap on training data’s

false-positive rates, a maximum false-positive rate was imposed to 0.01; (2) internal level consisting of AdaBoost ensemble with 1000 rounds of boosting; (3) internal level implementing a decision tree, and (4) an anomaly detection algorithm to identify whether the traffic is normal. Over the 18 experimental runs, the system was able to achieve 94% malicious traffic detection rate and 1.8% false-positive rates. The model was able to detect 89% of the attacks on which the system was not trained which proves the ability to detect zero-day threats. In future, the authors plan to introduce more parallelism to the entire architecture and thereby be able to reach real-time machine learning NIDS capability at 1Gbps.

In a similar research focussing on identifying zero-day botnet attacks in real-time, Ahmed et al. [2] evaluated the accuracy of DNNs (Deep neural network) and determined the capabilities of DNNs by applying the algorithm to a CTU-13 dataset. The proposal was divided into two parts, the first part using a feed-forward back propagation ANN (artificial neural network) and the second part using a DNN, to compare the botnet detection capabilities between deep learning models and traditional machine learning models. By running several experiments with input consisting of multiple feature set, an accuracy of over 99.6% were achieved by using deep learning ANN model with a total loss of 0.54. The accuracy of the deep learning model in detecting a botnet attack is higher than other machine learning models using SVM, Decision Tree and NB. The authors would like to examine the efficacy of the model on alternate datasets and plans to apply deep learning model to detecting other threat types such as DDoS in a future study.

Chou et al. [9] used deep learning algorithms using opensource software tools to develop a system capable of classifying malicious traffic from normal traffic. The research used a deep neural network (DNN) forward propagation algorithm on an NSL-KDD dataset to classify DoS and probing attacks. The model was able to achieve an accuracy of 98.99% in detecting DoS attacks and 97.65% in detecting a probing attack. While the model's performance was as expected on some attacks, it performed poorly in accurately classifying attacks such as U2R (User to Root) and R2L (Remote to Local), which the authors believe is down to monotonous nature of the training dataset causing over-learning. For future, the authors aim to improve upon the training characteristics of the model to overcome the current limitations.

Dutta et al. [10] discussed the issues in anomaly detection in Intrusion Detection Systems using traditional ML models and proposes a stacked generalization approach to achieve reliable classification of outliers. The proposed method utilizes deep models such as DNN, LSTM (Long Short-Term Memory) and meta-classifier (logistic regression) to improve the anomaly detection. The proposed model involves two stages: (1) a data pre-processing stage utilizing a Deep Sparse AutoEncoder (DSAE) which uses a sigmoid instead of neural activation functions; (2) classifier modelling using stacked ensemble (DNN, LSTM) to eliminate bias towards a particular dataset. An assessment of the proposed model is against multiple heterogeneous datasets (IoT-23, LITNET-2020 and NetML-2020) resulted in improved performance compared to other techniques such as RF (Random Forest) and SVM (Support Vector Machine). When validated using pre-specified datasets, the authors found significant improvement in evaluation metrics and can provide the required

accuracy in detecting anomalous behaviour in a network. The authors plan to carry out experiments on sophisticated datasets use advanced computational methods such as Apache Spark to enhance the scalability to cater large network traffic data.

A novel deep learning self-taught learning (STL) based intrusion detection system is proposed by Al-Qatf et al. [4] which helps with dimensionality reduction and feature learning. The model utilizes sparse autoencoder algorithm to improve unsupervised feature learning which reduces the training and testing times and boosts the prediction accuracy. The model is tested using NSL-KDD dataset, the non-numerical features of the dataset are encoded using 1-n system to suit the proposed model before being normalized to map all features. The STL based model proved to be extremely accurate in classifying malicious traffic with significant improvement is training and testing times. In a direct comparison against shallow classification techniques such as J48, naïve Bayesian, RF and SVM, the proposed model showed higher accuracy rate particularly under two-category (normal and attack) and five-category (normal and five attacks) classification. In future research the authors plan to use multiple stages of STl with a hybrid model for better feature representation.

To counter the high false-positive rates which are often accompanied with high accuracy rates in traditional models, considering the spatial and temporal features in the data, Wu and Guo [34] proposed a hierarchical neural network LuNet. LuNet consisted of several layers of convolutional neural networks (CNN) and recurrent neural networks (RNN) which learn in sync from the training data. The CNN + RNN synergy along with increased learning granularity can be utilized to effectively extract spatial and temporal features. CNN often aims at spatial features while RNN targets temporal features and over multiple levels feature extraction becomes spatial oriented. The authors overcame the challenge by introducing a LuNet block which combines both the CNN and RNN blocks at each level thereby retaining all the necessary features. The model is tested on two different non-redundant datasets NSL-KDD and UNSW-NB15 and the design maintained significantly lower false-positives while offering high validation accuracies and detection rates. LuNet makes use of cross-validation scheme to tackle the imbalanced distribution of NSL-KDD dataset. The performance of LuNet is categorized as: (1) Binary Classification in which LuNet distinguishes normal traffic with attack traffic; (2) Multi-Class Classification in which LuNet classifies the traffic as normal or belonging to one type of attack provided in the dataset. In Binary Classification, detection rates of 99.42% and 98.18%, accuracy rates of 99.24% and 97.4% and false-positive rates of 0.53% and 3.96% were achieved for NSL-KDD and UNSW-NB15 datasets. In Multi-Class Classification, the accuracy rates of LuNet averages at 99.05% and 84.98%, detection rates at 98.84% and 95.96% and false-positive rates of 0.65% and 1.89% for the two different datasets. The inefficiency of LuNet in attack classification to some attacks such as backdoors and worms is primarily down to the lack of sufficient samples in the training data and will become a part of authors future work.

The competency of deep learning techniques in detecting cyberattacks was brought to mobile cloud environment by Nguyen et al. [26]. A novel framework was proposed leveraging deep learning algorithms to train a neural network which can detect cyberattacks with high accuracy. All the user requests in the network are

sent through an attack detection module which classifies the traffic and forwards any suspicious packets to the security control module which then verifies the suspicious packet to take appropriate action of either allowing the packet or to block it. By using feature analysis and dimension reduction in the deep learning model, the required features to train the model are extracted. The model is tested on three different datasets NSL-KDD, UNSW-NB15 and KDDcup 1999 to evaluate the performance of the model on the metrics of accuracy, precision and recall, and compare them to other machine learning algorithms such as K-means, RF (random forest classifier), Gaussian Naïve Bayes, Multilayer Perceptron (MLP) etc. The proposed model was able to achieve best performance metrics compared to all other models with an accuracy rate of 90.99%, 95.84 and 97.11% for three datasets. The other evaluation parameters of precision and recall also achieved optimal metrics which demonstrates the stability, robustness, and flexibility of the proposed model. The authors aim to take the model real-time, testing the accuracy on real devices to evaluate the detection times and the power consumption rates.

Mohammad and Alsmadi [23] emphasized the importance of feature selection as a critical component to any AI classification model regardless of the internal algorithm. The efficiency of the selected feature set determines the performance of the classification model; hence the authors proposed an innovative algorithm for feature selection called HW (the Highest Wins). HW uses a statistical approach of measuring the distance between the observed and expected probability values, which is similar to other feature selection algorithms such as X2 (chi-square) and IG (information gain), but is more robust, simple and easy to comprehend. To test the generalization ability of the new method, ten datasets with varying input features were used, and the results showed significant improvement in reducing the dimensionality over other classification models. The evaluation metrics of recall, accuracy, precision and F1 score showed better results using features selected by HW technique, while class imbalance was still an issue similar to X2 and IG. In a second experiment, two versions of NSL-KDD datasets, binary and multiclass, were used. The experiment resulted in performance boost not only in all the evaluation metrics but also in the classification time and number of rules produced. The authors left the class imbalance issue for future work along with identifying advanced search techniques to improve the process of feature selection.

3 Research Methodology

The aim for the research is to identify the feasibility of using artificial intelligence based intrusion detection system in a real-time network. There has been ample amount of research happening in the past few years to develop a fully functional AI model capable of predicting cyberattacks. While most of the research efforts were successful in a way or other, a fully functional model is yet to be designed, or at least is not released publicly as an open source. Most of the researchers focussed on finding the right AI model for implementation across network architectures of

diverse magnitude, which is why the past models are tested using various generalized datasets to identify the threat detection capabilities. There was no research dedicated upon the implementation of AI tailored to the requirements of small and medium scale businesses (SMEs). The current research tries to bridge the gap by introducing an AI component to the traditional security practices used in SMEs as a way to better detect malicious threats. The research tries to follow and utilize some of the methodologies and practices used by other researchers to reach up to speed with the current trends in machine learning implementations. Artificial Intelligence is a broad field of science that comes in various flavours and subsets such as machine learning, natural language processing, expert systems etc. The current research uses a Neural Network (NN) to develop a threat detecting model to cater the needs of an SME.

The research follows a quantitative approach of measuring the performance of the proposed system using multiple evaluation metrics. The research works on the principle that the deep learning model can be integrated to the current security topology in an SME network and is fed with data that flows through the network. The proposed model continuously evaluates the network traffic and triggers an alert whenever the traffic pattern matches a pre-learnt pattern or when the accuracy levels reach a pre-defined threshold.

The research relies on the fact that malicious traffic often exhibits features that are often unique from other traffic and can be identified by carefully choosing the feature set from any given dataset. For example, features such as packet entropy values from source ip address, average arrival time, source bytes can provide the required information to identify a DoS attack [18]. Feature selection plays an important role in a machine learning model in accurately detecting an attack.

3.1 Neural Networks

Neural networks are computing networks consisting of artificial neurons simulating the neurons in human brain that are capable of processing and learning from large amount of data. The arrangement and interconnection between these neurons determine the characteristics of the neural network and its logical problem solving abilities. Neural networks can learn from complex and nonlinear data and can be trained to classify data, identify relationships and patterns, generalize and reason, generate predictions, etc., and the efficiency and accuracy of a neural network relies on the training data that is input to the model.

A typical neural network consists of an input layer, one or more hidden layers, and an output layer. These layers consist of a large number of nodes which are interconnected and have a threshold and weight value. An activation function is used to change the output value beyond the threshold which activates the node and data is passed on to the next level. Abstraction is key as the input data is passed on to several hidden layers where the data processing is performed based on the weight function before passing the data to the output layer. A learning function defines the weight

value which can be increased or decreased to achieve the desired output. Various algorithms exist to understand and correlate the communication between these layers such as Feedforward neural network, convolutional neural network (CNN), Recurrent neural network (RNN) etc.

A simple feedforward neural network technique is used for the current research consisting of multiple hidden layers which are adjusted according to the quantity of the input data.

- Model Topology

The research makes use of multiple tools and technologies to develop and implement an artificial intelligence model capable of detecting malicious threats in an SME environment. The key phases in the whole model are:

- Network monitoring

As the name infers, the network monitoring phase involves data monitoring and capturing. Network monitoring is a critical component in a typical IT network to monitor the network performance and fault tolerance. The research assumes that the network has sensors or monitoring points located at key locations such as servers, firewalls, routers etc. within the network to capture traffic in real-time. The captured data is then fed to the proposed AI model for training and analysis.

- Data pre-processing

A critical aspect of the proposed AI model is data pre-processing which is the process of converting raw data to clean data and is often the first step in working with data. According to Press [29], data scientists often spend around 80% of their time collecting and organising the data. The data collected from the network often unstructured and consists of huge amount of information such as files, audio, video, scripts etc., and machine learning models do not have the required capabilities to understand and process the data. Data pre-processing is critical as incorrect formatting or cleaning often causes more harm than good, and well-organised data is more valuable than the powerful algorithms of machine learning [22].

There are three key aspects of data pre-processing which are:

1. Data cleaning

Raw data can be incomplete, unorganised, and hugely complex for the machine learning models to process. The amount of missing, incomplete and noisy data often adds up to the processing time and accuracy of the proposed model. The data is cleaned of unwanted text, symbols, duplicates, missing data, blanks, etc. in the data cleaning stages.

2. Data transformation

The traffic data captured from the network consists of several fields such as ip addresses, data and time stamps, protocols, packets sizes, number of packets, etc.

There will be large columns of data that are often divided as numerical or categorical. The performance and accuracy of a machine learning model depends entirely on the feature selection, features that are important for the machine learning model. The selected features often dictate the algorithm accuracy, and the data processing time.

The research uses multiple features such as source address, destination address, protocol, source and destination ports, label, label detail and threat as its dataset features.

3. Data reduction

Typical data processing consisting of lot of features costs huge amount of computing resources and time. Data reduction reduces the amount of amount of data in a dataset while maintaining the integrity of the data, i.e., keeping a minimum set of features required for data analysis. Data reduction can be achieved by feature selection and extraction, removing non-essential features, deriving new features by combining existing ones etc. Dimensionality reduction using PCA (Principal component analysis) is one of a key technique where new features are derived from large set of variables.

The current research uses the following principal features: source address, source port, protocol, label, and label_detail.

3.2 Building and Training of the Model

The final part of the proposed model is the area of actual building, training, and testing of the AI model. This part splits the pre-processed data into three types of data: training, validating, and testing data. The model identifies patterns, generates insights, and learn from the training data and validates its knowledge using the validation data. The entire skillset is then tested against the testing data to predict the accuracy level of the model.

In the current research there are two different datasets used for evaluation, a network traffic dataset created in a virtual environment consisting of a server and two other machines, one of them acting as an attacker, and the second dataset is an IoT-23 dataset [12], used to validate the performance and efficiency of the proposed model.

3.3 Tools and Virtual Environments

The research makes use of various open-source tools to design and implement an AI model capable of predicting malicious attacks. The virtual environment is designed using virtual machine program Virtualbox with multiple virtual machines using different flavours of Linux. The test network is as shown in the figure below. The

network consists of a server that is configured to act as a DNS and DHCP roles. The attacker machine is used to carry out malicious tasks while network activity is captured using Wireshark software. For the research simple scan attacks using nmap are performed (Fig. 1).

Some of the other key tools are detailed below.

- TensorFlow/TensorBoard

TensorFlow is an open-source ecosystem of tools, libraries and resources for building and deploying machine learning applications. With a comprehensive set of low-level and high-level APIs (Application Programming Interface), capability of running multiple CPUs or GPUs or mobile platforms, TensorFlow is highly scalable and is an end-to-end platform for developing deep neural network models. TensorFlow is typically programmed using Python, but can be used with a variety of programming languages such as Java, C, C++, R, Matlab etc. TensorFlow provides several pre-trained models and datasets for multiple platforms and these models are suitable for production use which makes it very appealing and well used in various sectors such as healthcare, automobile, image, and face recognition systems, virtual assistant etc. [32].

TensorBoard provides additional capabilities to TensorFlow by providing the visualisations, graphs and measurements needed during a machine learning application workflow. The evaluation metrics used in TensorFlow such as accuracy, loss etc. can be tracked efficiently using TensorBoard.

- Wireshark

Wireshark is a network monitoring and analysis tool that provides customizable in-depth monitoring, auditing, and performance capabilities to organisations. Although being an open source allowing unrestricted use, Wireshark is adaptable, highly stateful, and provides extensive high-fidelity logs for network measurement and analysis. Unlike other monitoring tools, Wireshark offers powerful filters to identify the network issues with ability to use in either graphical mode or TTY mode using

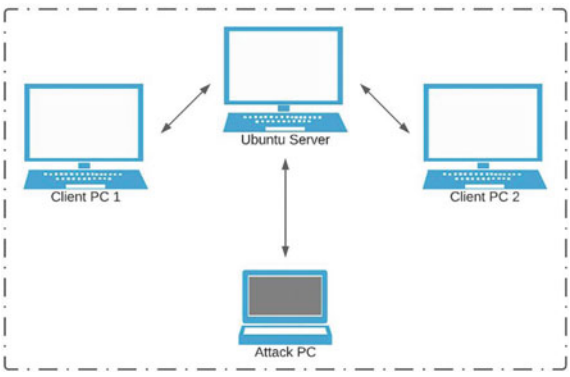


Fig. 1 Virtual lab setup

Tshark utility. Wireshark also provides support with decryption for many protocols such as WPA/WPA2, WEP, IPsec, SSL/TLS etc. The network can be monitored irrespective of the end-devices using wired (Ethernet), wireless (802.11), Bluetooth, frame relay etc.

- IoT-23 Dataset

With a goal of offering large and real dataset for researchers to develop machine learning applications, IoT-23 dataset was captured at Stratosphere Laboratory in Czech Republic. With support from cybersecurity software firm Avast, network traffic from three real IoT devices (Smart LED lamp, Intelligent personal assistant, and a Smart door lock) was captured over a period of two years. The dataset was divided into 23 captures with twenty captures of malicious network traffic and three captures of genuine network traffic. By executing specific malware samples in a controlled environment, the malicious activities were captured, categorised, and labelled accordingly. IoT-23 is a comprehensive dataset with 20 scenarios of data offering more than 280 million flows of malicious data traffic belonging to multiple attack categories such as horizontal port scans, DDoS, Okiru malware etc. With three benign traffic captures providing authentic traffic data to test, IoT-23 presents researchers with excellent resources to conduct machine learning experiments.

3.4 Ethical/Social/Legal Issues

Artificial intelligence is one of the most promising technologies deeply embed with human life and is widely used in various fields such as healthcare, finance, banking, automobiles etc. While AI definitely provide answers to several critical questions of this generation, the ethical, social and legal issues surrounding the use of AI needs more transparency.

- Ethical Issues

The role of ethics in AI is discussed by Anderson and Anderson [5] that AI machines should follow an ethical principle set while making decisions about the actions and procedures to follow, and such as system would find more acceptance than the one without. Ethical issues with AI based systems are often classified depending on whether the systems are considered as subjects or objects [25]. When considering AI as a tool some of the key concerns are data privacy, manipulation, bias, transparency, employment, and autonomy issues, and when AI is considered as a subject has issues with machine ethics and moral agency. Some of other ethical concerns includes misuse, questionable design, and unintended consequences which might determine the effectiveness of AI models in real-time scenarios.

- Social Issues

One of the most delicate and complex questions with artificial intelligence is the application of AI with human like decision-making abilities. Human intelligence often holds reasoning and responsibility at its core and the decisions often include social implications in mind. The social functions applicable to humans such as responsibility, accountability, predictability, incorruptibility transparency etc. needs to be integrated into the already complex AI algorithms and even then, proper governance is required to oversee the implementation. An AI system should be able to gain public trust on its performance and handling which largely depends on maintaining a clear view of responsibility/liability and transparency regarding the accountability. According to Ouchchy et al. [27], providing accurate information access to the public through value statements, factsheets often improve the public trust factor for an enhanced AI adoption.

- Legal Issues

The adaption of AI into the society always begs the question of rationale behind the decision-making process followed by a particular algorithm. The implementation of AI also brings various legal issues surrounding the data protection and privacy, which would have huge implications on the society. A lack of transparency to inspection is a key problem in AI implementations and is often questioned in the face of law.

While the proposed system doesn't have any social, ethical and legal issues might still be applicable with respect to the huge amount of data collected and fed into the model. The information collected from the network will be transformed into binary values to avoid any potential misuse and incorrect representation. Biased feature set selection might become an issue if the design is not governed properly.

3.5 Limitations

The research to develop AI based threat model to detect network attacks is performed completely on virtual environment and, while the performance is as expected and in line with the requirements, the performance when integrated with real networks needs more experimentation and might require considerable changes and testing before it can realize the true capabilities. The use of IoT-23 dataset provides the evaluation metrics to measure the performance even when the AI model is not being used in real-time.

4 Data Analysis and Critical Discussion

The project aims to build an AI tool suitable for SMEs that can detect anomalies in the network traffic. The proposed system can be integrated into a real network system

with data input fed from multiple locations in the network. The tool continuously analyses the network traffic and by using past knowledge learnt from training and testing, predicts whether the traffic is malicious and generates alerts for the security experts. The proposed AI system can be manually monitored or configured to operate automatically according to the recommendations of the AI system.

The performance of the system can be characterized on two important criteria: (a) efficiency of data pre-processing and (b) accuracy and loss metrics of the AI module. There are multiple parameters in each phase of the model that will provide authenticity that the performance of the proposed model is in line with the targeted requirements of the given network. It is equally important to analyse and understand the logical process of the AI model to better gauge the efficiency of the proposed model. The factors that the system is based on are detailed in the following sections.

4.1 Data Pre-processing

The proposed model uses TensorFlow machine learning tool to analyze the data from a csv file, uses its analytical capabilities to measure and compare the new data against the data that was used to train the AI system. Data pre-processing is a key process for the AI model to be effective as the model works entirely based on the data that was fed into the system. The type and quality of data, number of features, size of the dataset etc., forms the core of the data processing stage which in turn affects the AI model's resourcefulness.

- Data Capturing

In the data pre-processing stage, the network traffic is captured using Wireshark network scanning tool. Wireshark allows for traffic analysis at the connection level and can be configured to monitor for activity based on the protocol such as ICMP (Internet Control Message Protocol), TCP (Transmission Control Protocol), UDP (User Datagram Protocol) etc. The network designed in the virtual environment has a server and multiple client machines and the Wireshark tool is run on the attack machine.

The AI tool TensorFlow reads data from files with extension of CSV (Comma Separated Values) and hence the data captured by the Wireshark tool is saved as a CSV file enabling the AI model to extract the data contained in the file.

Data Formatting and Organizing

The pcap files that are captured holds various fields of data such as source and destination addresses, protocols, time stamps, etc. The details of a pcap file converted into csv are as shown in Fig. 2.

The same information for a data file from an IoT-23 dataset is as shown in Fig. 3. The IoT-23 dataset shown is already cleaned, organised, and labelled. As seen in the figure below, the dataset contains many featured columns which can provide

22	23.01194323	fe80::480d:8efc:4951:3ec4	ff02::c	SSDP	208	M-SEARCH * HTTP/1.1	
23	23.98793857	192.168.56.26	192.168.56.22	ICMP	98	Echo (ping) request id=0x6d79, seq=2/512, ttl=64 (reply in 24)	
24	23.98882465	192.168.56.22	192.168.56.26	ICMP	98	Echo (ping) reply id=0x6d79, seq=2/512, ttl=64 (request in 23)	
25	23.98909203	192.168.56.26	192.168.56.22	DNS	86	Standard query 0x8995 PTR 22.56.168.192.in-addr.arpa	
26	23.99017251	192.168.56.22	192.168.56.26	DNS	133	Standard query response 0x8995 PTR 22.56.168.192.in-addr.arpa PTR mail.testlab.co PTR as1.testlab.co	
27	24.98967964	192.168.56.26	192.168.56.22	ICMP	98	Echo (ping) request id=0x6d79, seq=3/768, ttl=64 (reply in 28)	
28	24.99028872	192.168.56.22	192.168.56.26	ICMP	98	Echo (ping) reply id=0x6d79, seq=3/768, ttl=64 (request in 27)	
29	24.99057869	192.168.56.26	192.168.56.22	DNS	86	Standard query 0x2ba7 PTR 22.56.168.192.in-addr.arpa	
30	24.99134663	192.168.56.22	192.168.56.26	DNS	133	Standard query response 0x2ba7 PTR 22.56.168.192.in-addr.arpa PTR as1.testlab.co PTR mail.testlab.co	
31	25.99228491	192.168.56.26	192.168.56.22	ICMP	98	Echo (ping) request id=0x6d79, seq=4/1024, ttl=64 (reply in 32)	
32	25.99784925	192.168.56.22	192.168.56.26	ICMP	98	Echo (ping) reply id=0x6d79, seq=4/1024, ttl=64 (request in 31)	
33	25.99809629	192.168.56.26	192.168.56.22	DNS	86	Standard query 0x7651 PTR 22.56.168.192.in-addr.arpa	
34	25.9998853	192.168.56.22	192.168.56.26	DNS	133	Standard query response 0x7651 PTR 22.56.168.192.in-addr.arpa PTR as1.testlab.co PTR mail.testlab.co	
35	27.01488557	fe80::480d:8efc:4951:3ec4	ff02::c	SSDP	208	M-SEARCH * HTTP/1.1	
36	28.01348909	PcsCompu_1f:30:76	PcsCompu_44:a6:86	ARP	42	Who has 192.168.56.22? Tell 192.168.56.26	
37	28.01416588	PcsCompu_44:a6:86	PcsCompu_1f:30:76	ARP	60	192.168.56.22 is at 08:00:27:44:a6:86	
38	28.05169825	PcsCompu_44:a6:86	PcsCompu_1f:30:76	ARP	60	Who has 192.168.56.26? Tell 192.168.56.22	
39	28.05172729	PcsCompu_1f:30:76	PcsCompu_44:a6:86	ARP	42	192.168.56.26 is at 08:00:27:1f:30:76	

Fig. 2 Traffic capture in VM dataset

#fields	ts	uid	id.orig_h	id.orig_p	id.resp_h	id.resp_p	proto	service	duration	orig_bytes	resp_bytes	conn_state	local_orig	local_resp	missed_bytes	history	orig_pkts	orig_ip_bytes	resp_pkts	resp_ip_bytes	tunnel_parents	label	detailed-label
#types	time	string	addr	port	addr	port	enum	string	interval	count	count	string	bool	bool	count	string	count	count	count	count	set[string]	string	string
1.53E+09	CbJLGOox	192.168.1(	123	81.2.254.2	123	udp	-	0.00549	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CF3cnH4jD	192.168.1(	123	147.231.1(	123	udp	-	0.001741	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CigmSt3br	192.168.1(	123	31.31.74.3	123	udp	-	0.004495	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	Cav32mAc	192.168.1(	123	147.251.4(	123	udp	-	0.000988	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CIvPFAAD	192.168.1(	123	147.231.1(	123	udp	-	0.001487	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CsayIRCZT	192.168.1(	123	37.157.19(	123	udp	-	0.004739	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CapItr2s2r	192.168.1(	123	147.231.1(	123	udp	-	0.001743	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CqXbOUQ	192.168.1(	123	31.31.74.3	123	udp	-	0.007737	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CTCAcUsI	192.168.1(	123	147.231.1(	123	udp	-	0.001493	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	CHtZjyvKa	192.168.1(	123	80.79.25.1	123	udp	-	0.002248	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-
1.53E+09	Cuz5ANgJ	192.168.1(	123	37.157.19(	123	udp	-	0.004997	48	48	SF	-	-	0	Dd	1	76	1	76	(empty)	Benign	-

Fig. 3 Traffic capture in IoT-23 dataset

additional information regarding the network but are not appropriate for the current research.

The csv file after removing unwanted columns and changing the labels for the IoT-23 dataset is shown in Fig. 4. The total number of columns are reduced and the label set is completely redone to provide a clear understanding and ease of use while working further on the dataset.

- **Feature Selection**

The csv file consists of multiple fields of data organised using the tools described above. The csv files are structured, labelled and tab-organized and ready for the AI model to be trained upon. The final part of the data pre-processing is the feature selection which is the key for the AI model to identify malicious activity accurately. The features must be relevant, fit the purpose and is often considered the core of the machine learning [31].

time	source_address	source_port	destination_address	destination_port	protocol	orig_pa	label	label_detail	threat
CfHl9r3XN	192.168.100.103	34243	147.7.65.203	49560	tcp	180	Benign -		0
CwxSC63a	192.168.100.103	43763	154.8.94.65	14336	udp	40	Benign -		0
CUmrqr4s	192.168.100.103	51524	65.127.233.163	23	tcp	180	Malicious	PartOfAHorizontalPortScan	1
CH98aB3s	192.168.100.103	56305	63.150.16.171	23	tcp	60	Malicious	PartOfAHorizontalPortScan	1
C3GBTkIN	192.168.100.103	41101	111.40.23.49	23	tcp	60	Malicious	PartOfAHorizontalPortScan	1
CC6vK11d	192.168.100.103	43763	196.170.198.141	11764	udp	40	Benign -		0
CDe43c1P	192.168.100.103	60905	131.174.215.147	23	tcp	180	Malicious	PartOfAHorizontalPortScan	1
CJaDcG3M	192.168.100.103	44301	91.42.47.63	23	tcp	60	Malicious	PartOfAHorizontalPortScan	1
CMBrup3E	192.168.100.103	50244	120.210.108.200	23	tcp	60	Malicious	PartOfAHorizontalPortScan	1
C7USrA15	192.168.100.103	34840	145.164.35.6	21288	tcp	60	Benign -		0
Cwm6z43	192.168.100.103	43763	118.112.204.139	2441	udp	40	Benign -		0

Fig. 4 Modified traffic capture csv file in IoT-23 dataset

The feature selection process reduces the unnecessary columns in the dataset which have little or no impact on the model's output and are often considered as noise. The feature selection can be automated using feature selection algorithms such as filter methods, wrapper methods, and embedded methods. A good pre-processed dataset improves accuracy, reduces overfitting and training time.

The feature or attribute selection for the current model is carried out using Weka, a tool which has the tools for data preparation for AI models. The csv file generated in the previous section is loaded into the Weka tool which allows for several machine learning algorithms to be implemented on the data. The 'ClassifierAttributeEval' algorithm for attribute classifier with 'Ranker' search method is used to identify the best feature set in the dataset. The process can be seen in Fig. 5.

As seen in the figure above the attributes that makes a difference in the AI model's performance are identified according to their ranking order. As identified above, the label_detail holds the utmost value which identifies whether the given packet belongs to a malicious traffic or genuine network traffic. Weka allows for several algorithms to be tested on the data before concluding on a suitable algorithm for the given network requirement, and for the current model the 'ClassifierAttributeEval' algorithm provides the best result.

```
=== Run information ===

Evaluator:     weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --
Search:        weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1
Relation:      1-1dataset-weka.filters.unsupervised.attribute.Remove-R1,7
Instances:     1008748
Attributes:    8
               source_address
               source_port
               destination_address
               destination_port
               protocol
               label
               label_detail
               threat
Evaluation mode:    evaluate on all training data


=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (numeric): 8 threat):
        Classifier feature evaluator

        Using  Wrapper Subset Evaluator
        Learning scheme: weka.classifiers.rules.ZeroR
        Scheme options:
        Subset evaluation: RMSE
        Number of folds for accuracy estimation: 5


Ranked attributes:
  0   7 label_detail
  0   3 destination_address
  0   2 source_port
  0   4 destination_port
  0   6 label
  0   5 protocol
  0   1 source_address

Selected attributes: 7,3,2,4,6,5,1 : 7
```

Fig. 5 ClassifierAttributeEval algorithm using Weka

4.2 *Artificial Intelligence System*

During the data pre-processing phase, the network traffic data is collected, organized structurally into tab-separated CSV (Comma Separated Values) files. Depending on the amount of network traffic, the data is fed into the TensorFlow as a single csv file or multiple csv files. The system reads the data, carries out the functions defined in the TensorFlow code, and calculates the accuracy which is the measure against the pre-learnt knowledge from the training data.

- TensorFlow Code

TensorFlow allows for the creation of AI models with several layers of neural networks and is capable of processing large amounts of data within these neural network layers. The entire process is extremely resource intensive and requires the use of dedicated computing infrastructure capable of running extremely complex parallel processes. The current model is hence compiled using Google Colab, a cloud environment that gives developers access to high-end computing resources.

The code used for TensorFlow AI model is programmed using Python language and is as shown in the picture in Fig. 6. The initial part of the code deals with importing the required python libraries that are needed for running the TensorFlow model. The next part deals with mounting a cloud drive which stores all the csv files generated during the pre-processing stage. The model is capable and coded to learn from a single csv file or from multiple csv data files. The features set from the csv files are then organised and categorised according to the selected feature set.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from tensorflow import feature_column
from tensorflow.keras import layers
from sklearn.model_selection import train_test_split
from pandas.api.types import CategoricalDtype
#Use Pandas to create a dataframe
#In windows to get file from path other than same run directory see:
#https://stackoverflow.com/questions/16952632/read-a-csv-into-pandas-from-f-drive-on-windows-7
dataframe = pd.read_csv("/home/lakshmi/Desktop/CSV/combineddataset2.csv")
#print(dataframe.head())
#show dataframe details for column types
#print(dataframe.info())
#print(pd.unique(dataframe['user']))
#https://pbpython.com/categorical-encoding.html
dataframe["source_address"] = dataframe["source_address"].astype('category')
dataframe["source_port"] = dataframe["source_port"].astype('category')
dataframe["label"] = dataframe["label"].astype('category')
dataframe["source_address_cat"] = dataframe["source_address"].cat.codes
dataframe["source_port_cat"] = dataframe["source_port"].cat.codes
dataframe["label_cat"] = dataframe["label"].cat.codes
#print(dataframe.info())
#print(dataframe.head())
#save dataframe with new columns for future datmapping
dataframe.to_csv('dataframe-export-allcolumns.csv')
#remove old columns
del dataframe["source_address"]
del dataframe["source_port"]
del dataframe["label"]
#restore original names of columns
dataframe.rename(columns={"source_address_cat": "source_address", "source_port_cat": "source_port", "label_cat": "label"}, inplac
print(dataframe.head())
print(dataframe.info())
```

Fig. 6 Python code for VM dataset, Part-1

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
The tensorboard extension is already loaded. To reload it, use:
  %reload_ext tensorboard
File names: ['/content/drive/My Drive/AI/Dataset/1-1dataset.csv', '/content/drive/My Drive/AI/Dataset/3-1dataset.csv', '/content/drive/My
      time  source_address  ...  destination_port  orig_packets
0  2179639            7194  ...             49480            68
1  3114058            7194  ...             14308             1
2  1626229            7194  ...                23            68
3   907023            7194  ...                23             6
4   172311            7194  ...                23             6

[5 rows x 10 columns]
<class 'pandas.core.frame.DataFrame'>
Int64Index: 3274347 entries, 0 to 1374
Data columns (total 10 columns):
 #   Column               Dtype
---  ------               -----
 0   time                 int32
 1   source_address       int16
 2   destination_address  int32
 3   protocol             int8
 4   label                int8
 5   label_detail         int8
 6   threat               int8
 7   source_port          int32
 8   destination_port     int32
 9   orig_packets         int16
dtypes: int16(2), int32(4), int8(4)
memory usage: 99.9 MB
None
```

Fig. 7 Python Code output for VM dataset

The output for the TensorFlow code above is shown in Fig. 7. The code shows the feature imported by the program from the csv data files along with the column types and other information in the data file.

The second part of the code is when the artificial intelligence is put to work on the data imported from the csv data files. The code is detailed in the figure below which starts with splitting the large amount of data into multiple categories of training, testing and validation. The key for this model is the 'Threat' column, which identifies the traffic as benign or malicious and learns the characteristics of such threat type by looking at data from other feature columns in the same row. The dataframe works by reading line by line data and studying the type of information contained in the fields for the corresponding benign or malicious traffic type (Fig. 8).

The last part of the code builds, compiles, and trains a sequential AI model using TensorFlow functions. An activation function called ReLU (Rectified Linear Unit) is used to create a network with multiple layers, the details of the layers are shown below in the figure. The model uses a stochastic optimizer algorithm 'Adam' [17] to test the model in 'accuracy' and 'loss' metrics. A stream of continuous data is input to the system and higher value of accuracy determines that the network traffic data is malicious. The model is run for five iterations or epochs to determine the accuracy levels.

Figure 9 shows the neural network layers for the proposed AI model.

```
#Split the dataframe into train, validation, and test
train, test = train_test_split(dataframe, test_size=0.2)
train, val = train_test_split(train, test_size=0.2)
print(len(train), 'train examples')
print(len(val), 'validation examples')
print(len(test), 'test examples')
#Create an input pipeline using tf.data
# A utility method to create a tf.data dataset from a Pandas Dataframe
def df_to_dataset(dataframe, shuffle=True, batch_size=32):
  dataframe = dataframe.copy()
  labels = dataframe.pop('threat')
  ds = tf.data.Dataset.from_tensor_slices((dict(dataframe), labels))
  if shuffle:
    ds = ds.shuffle(buffer_size=len(dataframe))
  ds = ds.batch(batch_size)
  return ds
#choose columns needed for calculations (features)
feature_columns = []
for header in ["source_address", "field_type", "label"]:
    feature_columns.append(feature_column.numeric_column(header))
#create feature layer
feature_layer = tf.keras.layers.DenseFeatures(feature_columns)
#set batch size pipeline
batch_size = 32
train_ds = df_to_dataset(train, batch_size=batch_size)
val_ds = df_to_dataset(val, shuffle=False, batch_size=batch_size)
test_ds = df_to_dataset(test, shuffle=False, batch_size=batch_size)
#create compile and train model
model = tf.keras.Sequential([
  feature_layer,
  layers.Dense(128, activation='relu'),
  layers.Dense(128, activation='relu'),
  layers.Dense(1)
])
model.compile(optimizer='adam',
              loss=tf.keras.losses.BinaryCrossentropy(from_logits=True),
              metrics=['accuracy'])
model.fit(train_ds,
          validation_data=val_ds,
          epochs=5)
loss, accuracy = model.evaluate(test_ds)
print("Accuracy", accuracy)
```

Fig. 8 Python Code for VM Dataset, Part-2

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_features (DenseFeatur  multiple                 0
 es)

 dense (Dense)               multiple                  768

 dense_1 (Dense)             multiple                  16512

 dense_2 (Dense)             multiple                  129

=================================================================
Total params: 17,409
Trainable params: 17,409
Non-trainable params: 0
```

Fig. 9 Neural network layers

4.3 AI Model Analysis

The AI model proposed in this research is tested on a network simulation in a virtual environment and the model's performance is verified using alternate IoT-23 dataset. The results of the AI system are discussed below:

Analysis of AI model using VM dataset

The dataset from the virtual machine setup is tested on the AI system and the output is as shown in Fig. 10.

The dataset contains a total of 1,008,750 flows of data which are captured while performing malicious networks acts from the attack machine. The dataset is split into 645,600 samples of training data, 161,400 examples of validation data and 201,750 flows of data acting as test examples. The AI model uses the training data to generate insights and acquire knowledge which is validated using the validation samples. The system tests the knowledge acquired using the test dataset before calculating the accuracy and loss parameters.

The proposed system reached an accuracy level of 46.74% with an average loss of 0.69% over 5 epochs. The low accuracy level can be attributed to the similar

```
Consider rewriting this model with the Functional API.
20175/20175 [==============================] - 97s 4ms/step - loss: 736507.6250 - accuracy: 0.4996 - val_loss: 0.6905 - val_acc
uracy: 0.4633
Epoch 2/5
20175/20175 [==============================] - 102s 4ms/step - loss: 0.6908 - accuracy: 0.4650 - val_loss: 0.6905 - val_accurac
y: 0.4633
Epoch 3/5
20175/20175 [==============================] - 108s 5ms/step - loss: 0.6908 - accuracy: 0.4650 - val_loss: 0.6905 - val_accurac
y: 0.4633
Epoch 4/5
20175/20175 [==============================] - 108s 5ms/step - loss: 0.6908 - accuracy: 0.4650 - val_loss: 0.6905 - val_accurac
y: 0.4633
Epoch 5/5
20175/20175 [==============================] - 106s 4ms/step - loss: 0.6907 - accuracy: 0.4650 - val_loss: 0.6904 - val_accurac
y: 0.4633
6305/6305 [==============================] - 22s 3ms/step - loss: 0.6911 - accuracy: 0.4674
Accuracy 0.4674398899078369
```

Fig. 10 AI Model performance for VM dataset

numbers of malicious and benign traffic packets. The dataset consists of a smaller number of threat samples compared to regular network traffic which affected the accuracy levels. Better accuracy levels are possible when the model is trained with more malicious samples with diverse training samples and attack versions.

The AI model performed satisfactorily for the given virtual environment, but the accuracy levels are nowhere close to make any sound decisions in terms of the malicious nature of the network traffic, especially in a real-time network. The model is hence tested using another dataset captured using real devices with diverse attacks and over several years, which is IoT-23 dataset. The working of the mode using a diverse and large dataset such as IoT-23 is discussed below.

4.4 Analysis of AI Model Using IoT-23 Dataset

The IoT-23 dataset consists of traffic captures from 3 real IoT hardware devices connected to a network with access to the internet. The data is captured in 20 different scenarios using 20 different malware samples, and the entire dataset is clearly labelled with the malware types used within the respective scenarios. The IoT-23 dataset used in the current research makes use of 8 scenarios of data consisting of millions of data traffic flows.

The IoT-23 dataset is input to the proposed AI model to test the performance of the model and to ascertain its capabilities in predicting malicious traffic. Figures 11 and 12 shows the python code changed to adapt to the multiple csv files from the IoT-23 dataset. Unlike the AI model that is run in a virtual environment for the first part of the performance analysis, the AI model is run on a Google Colab cloud environment due to the processing power required to handle the large number of data flow samples in the IoT-23 dataset.

Figure 13 shows the computational graph that shows the data flow through the layers in the AI model. The graph shows how the optimizer and other metrics are used.

The performance of the AI model on the IoT-23 dataset is shown in Fig. 14.

As seen in the output window, the proposed TensorFlow model achieved excellent results in terms of accuracy and loss. When given a large dataset with enough samples to train and test, the model consistently showed accuracy levels of over 99% with loss values confined to 0.0093%. The model performed exceptionally showing the capabilities of the proposed model when given a large dataset with enough data samples.

Fig. 11 Updated TensorFlow Code for IoT-23 Dataset, Part 1

```
import numpy as np
import pandas as pd
import tensorflow as tf
import glob
from tensorflow import feature_column
from tensorflow.keras import layers
from sklearn.model_selection import train_test_split
from pandas.api.types import CategoricalDtype
from tensorflow.keras.callbacks import TensorBoard
from google.colab import drive
drive.mount('/content/drive')

#Use Pandas to create a dataframe
path = "/content/drive/My Drive/AI/Dataset/"
files = glob.glob(path + "/*.csv")
print('File names:', files)
dataframe = pd.DataFrame()
content = []

# To check all the csv files in the path
for filename in files:
    df = pd.read_csv(filename, index_col=None)
    content.append(df)

dataframe = pd.concat(content)
```

4.5 Performance Analysis Using TensorBoard

The artificial intelligence model's performance is analysed further to identify the performance concentrations over the processing period. To perform additional analysis, it is important to track and evaluate model, the process flow, hence advanced analytical view of the whole is essential. TensorFlow comes with an add-on web application called TensorBoard which provides the necessary visualisations to keep track of the model's metric such as learning rate, loss etc.

TensorBoard is an external process and hence needs to be called by the TensorFlow machine learning process. Once called, TensorBoard runs simultaneously along with TensorFlow process and logs all the events, LogDirs, execution summaries, process values etc., which are then presented in a variety of ways such as distributions, graphs, histograms etc. The python code for the proposed AI model is modified to accommodate the TensorBoard calling process.

The following figure shows the Scalars in TensorBoard which shows the key metrics of accuracy and loss and their performance over each of the epochs (iterations). As seen in the figure the trend of validation data is upwards which proves that the model is learning well, but at epoch 3 and 9, the model is slightly overfitting which needs to be weight adjusted. Typically, an AI model is run for several hundreds of epochs as the machine learning curve increases by running the model for more iterations which holds true to the current model. The training and validation are expected to reach a steady upward curve by running the algorithm for more epochs.

```
#print(dataframe.head())
#show dataframe details for column types
#print(dataframe.info())
#print(pd.unique(dataframe['user']))
#https://phpython.com/categorical-encoding.html
dataframe["time"] = dataframe["time"].astype('category')
dataframe["source_address"] = dataframe["source_address"].astype('category')
dataframe["destination_address"] = dataframe["destination_address"].astype('category')
dataframe["protocol"] = dataframe["protocol"].astype('category')
dataframe["label"] = dataframe["label"].astype('category')
dataframe["label_detail"] = dataframe["label_detail"].astype('category')
dataframe["threat"] = dataframe["threat"].astype('category')
dataframe["source_port"] = dataframe["source_port"].astype('category')
dataframe["destination_port"] = dataframe["destination_port"].astype('category')
dataframe["orig_packets"] = dataframe["orig_packets"].astype('category')


dataframe["time_cat"] = dataframe["time"].cat.codes
dataframe["source_address_cat"] = dataframe["source_address"].cat.codes
dataframe["destination_address_cat"] = dataframe["destination_address"].cat.codes
dataframe["protocol_cat"] = dataframe["protocol"].cat.codes
dataframe["label_cat"] = dataframe["label"].cat.codes
dataframe["label_detail_cat"] = dataframe["label_detail"].cat.codes
dataframe["threat_cat"] = dataframe["threat"].cat.codes
dataframe["source_port_cat"] = dataframe["source_port"].cat.codes
dataframe["destination_port_cat"] = dataframe["destination_port"].cat.codes
dataframe["orig_packets_cat"] = dataframe["orig_packets"].cat.codes
#dataframe["label_detail"] = dataframe["label_detail"].astype('string')
#print(dataframe.info())
#print(dataframe.head())
#save dataframe with new columns for future datmapping
dataframe.to_csv('dataframe-export-allcolumns.csv')
#remove old columns
del dataframe["time"]
del dataframe["source_address"]
del dataframe["destination_address"]
del dataframe["protocol"]
del dataframe["label"]
del dataframe["label_detail"]
del dataframe["source_port"]
del dataframe["destination_port"]
del dataframe["threat"]
del dataframe["orig_packets"]
#restore original names of columns
dataframe.rename(columns={"source_address_cat": "source_address", "destination_address_cat": "destination_address",
                          "time_cat": "time", "protocol_cat": "protocol", "label_detail_cat": "label_detail", "label_cat":
                          "label", "threat_cat": "threat", "source_port_cat": "source_port", "destination_port_cat":
                          "destination_port", "orig_packets_cat": "orig_packets"}, inplace=True)
print(dataframe.head())
print(dataframe.info())
#save dataframe cleaned up
dataframe.to_csv('dataframe-export-int-cleaned.csv')

#dataframe = np.asarray(dataframe).astype(np.float32)
#dataframe = dataframe.astype(np.float32)
#tf.convert_to_tensor(dataframe, dtype=tf.int32)
#Split the dataframe into train, validation, and test
train, test = train_test_split(dataframe, test_size=0.2)
train, val = train_test_split(train, test_size=0.2)
print(len(train), 'train examples')
print(len(val), 'validation examples')
print(len(test), 'test examples')
#Create an input pipeline using tf.data
# A utility method to create a tf.data dataset from a Pandas Dataframe
def df_to_dataset(dataframe, shuffle=True, batch_size=32):
  dataframe = dataframe.copy()
  labels = dataframe.pop('threat')
  ds = tf.data.Dataset.from_tensor_slices((dict(dataframe), labels))
  if shuffle:
    ds = ds.shuffle(buffer_size=len(dataframe))
  ds = ds.batch(batch_size)
  return ds
#choose columns needed for calculations (features)
feature_columns = []
for header in ["source_address", "source_port", "protocol", "label", "label_detail"]:
    feature_columns.append(feature_column.numeric_column(header))
#create feature layer
feature_layer = tf.keras.layers.DenseFeatures(feature_columns)
#set batch size pipeline
batch_size = 32
train_ds = df_to_dataset(train, batch_size=batch_size)
val_ds = df_to_dataset(val, shuffle=False, batch_size=batch_size)
test_ds = df_to_dataset(test, shuffle=False, batch_size=batch_size)

#create tensorboard callback
tensorboard_callback = [TensorBoard(
    log_dir="logs",
    histogram_freq=1,
    write_graph=True,
    write_images=False,
    update_freq="epoch",
)]

#create compile and train model
model = tf.keras.Sequential([
  feature_layer,
  layers.Dense(128, activation='relu'),
  layers.Dense(128, activation='relu'),
  layers.Dense(1)
])
model.compile(optimizer='adam',
              loss=tf.keras.losses.BinaryCrossentropy(from_logits=True),
              metrics=['accuracy'])
model.fit(train_ds,
          validation_data=val_ds,
          epochs=10, callbacks=[tensorboard_callback])
loss, accuracy = model.evaluate(test_ds)
print("Accuracy", accuracy)
```

Fig. 12 Updated TensorFlow Code for IoT-23 dataset, Part 2

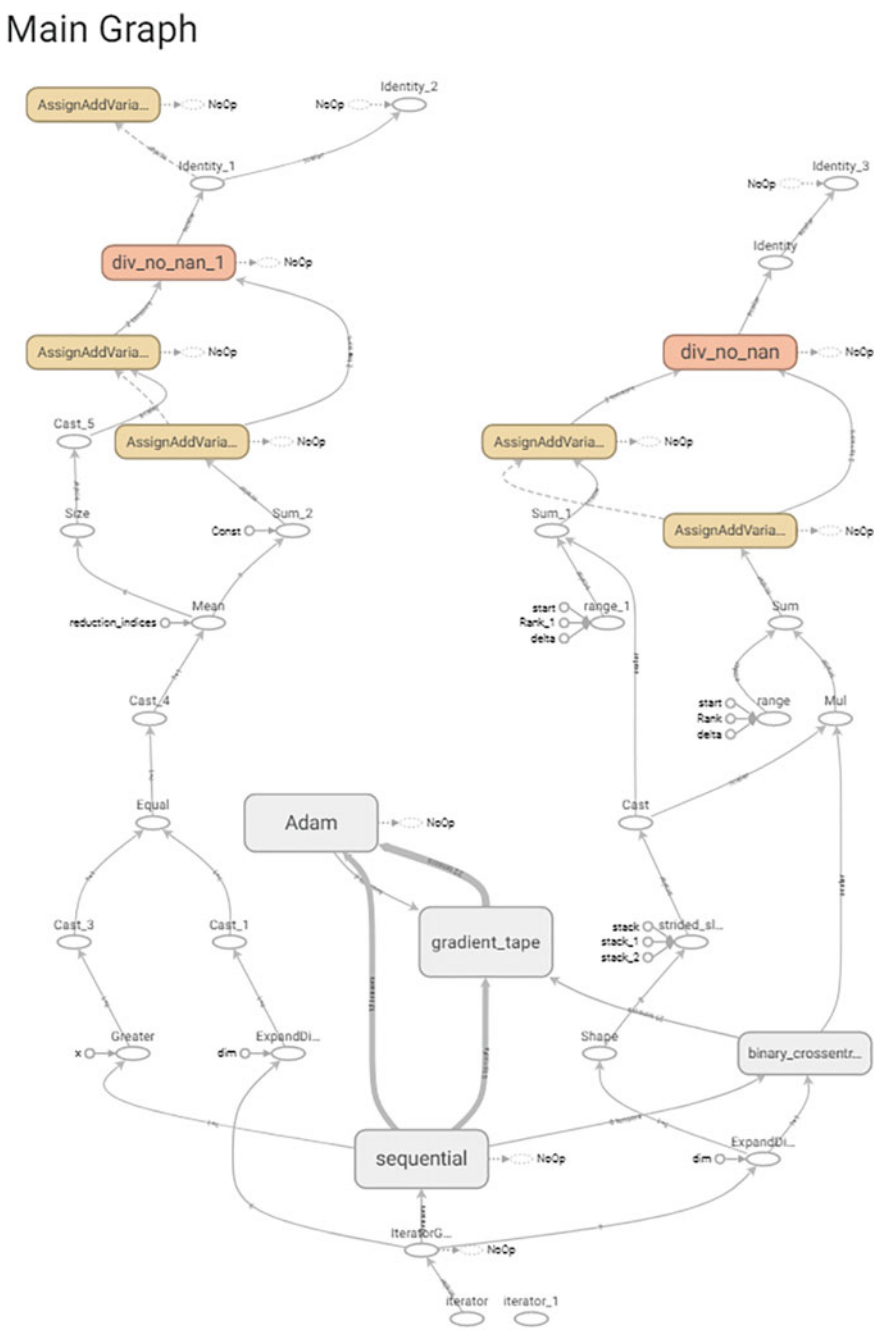


Fig. 13 TensorBoard Graph showing AI code flow

```
2095581 train examples
523896 validation examples
654870 test examples
Epoch 1/10
WARNING:tensorflow:Layers in a Sequential model should only have a single input tensor, but we receive a <class 'dict'> input: {'time': <tf.Tensor 'Iterator(
Consider rewriting this model with the Functional API.
WARNING:tensorflow:Layers in a Sequential model should only have a single input tensor, but we receive a <class 'dict'> input: {'time': <tf.Tensor 'Iterator(
Consider rewriting this model with the Functional API.
65487/65487 [==============================] - ETA: 0s - loss: 1.3995 - accuracy: 0.9748WARNING:tensorflow:Layers in a Sequential model should only have a si
Consider rewriting this model with the Functional API.
65487/65487 [==============================] - 250s 4ms/step - loss: 1.3995 - accuracy: 0.9748 - val_loss: 0.0028 - val_accuracy: 0.9984
Epoch 2/10
65487/65487 [==============================] - 272s 4ms/step - loss: 0.0150 - accuracy: 0.9967 - val_loss: 3.2167e-04 - val_accuracy: 1.0000
Epoch 3/10
65487/65487 [==============================] - 271s 4ms/step - loss: 0.0136 - accuracy: 0.9971 - val_loss: 8.6603e-04 - val_accuracy: 1.0000
Epoch 4/10
65487/65487 [==============================] - 246s 4ms/step - loss: 0.0914 - accuracy: 0.9596 - val_loss: 0.0147 - val_accuracy: 0.9954
Epoch 5/10
65487/65487 [==============================] - 260s 4ms/step - loss: 0.0251 - accuracy: 0.9924 - val_loss: 0.0107 - val_accuracy: 0.9980
Epoch 6/10
65487/65487 [==============================] - 271s 4ms/step - loss: 0.0114 - accuracy: 0.9972 - val_loss: 0.0029 - val_accuracy: 0.9991
Epoch 7/10
65487/65487 [==============================] - 251s 4ms/step - loss: 0.0154 - accuracy: 0.9949 - val_loss: -1.3915e-04 - val_accuracy: 1.0000
Epoch 8/10
65487/65487 [==============================] - 243s 4ms/step - loss: 0.0094 - accuracy: 0.9983 - val_loss: 0.0083 - val_accuracy: 0.9980
Epoch 9/10
65487/65487 [==============================] - 270s 4ms/step - loss: 0.0088 - accuracy: 0.9981 - val_loss: -1.6733e-04 - val_accuracy: 1.0000
Epoch 10/10
65487/65487 [==============================] - 262s 4ms/step - loss: 0.0526 - accuracy: 0.9796 - val_loss: 0.0102 - val_accuracy: 0.9980
20465/20465 [==============================] - 53s 3ms/step - loss: 0.0093 - accuracy: 0.9982
Accuracy 0.998204231262207
Model: "sequential"
```

Fig. 14 TensorFlow Output for IoT-23 dataset

Figure 15 also shows the loss metric over each epoch. As the model is run multiple times the loss value decreased substantially within 4 epochs before settling at the bottom with rates close to 0.006%, which is close to the bare minimum of lose values.

Figures 16 and 17 shows the TensorBoard distributions and histograms i.e., the statistical distributions of the tensors over time. The distributions identify the weight changes in the model over the entire processing period.

Data Analysis

The aim of the current research is to develop a malicious traffic identifying Artificial Intelligence tool that can be integrated to a real network. The proposed AI model produced the results that look promising in achieving the goal of accurately identifying malicious traffic in a network. While the current model is yet to be tested on a real network, the accuracy levels achieved using the current model does look assuring for implementing in a real network.

The key phase of the AI model is the data pre-processing during which the network traffic captured is cleaned, optimized, and structured to befit machine learning model's input requirements. The current dataset was captured with more than 25 columns of data but was reduced to 8 featured columns holding the key attributes of source and destination address, source and destination port, protocol, label and label detail and threat. To verify the attributes and their effectiveness, the Weka tool is run with multiple algorithms such as 'CorrelationAttributeEval', 'GainRatioAttributeEval','WrapperSubsetEval' etc. before finalising on 'ClassifierAttributeEval' algorithm which is better suited to the current AI model design. The attributes are ranked giving the AI model a head start in identifying the key features for the AI model to generate the best possible evaluation metrics.

The AI model achieved low accuracy while implemented on a dataset created using a virtual lab environment. The VM data set contains multiple fields with the necessary features such as source and destination addresses, protocols etc., and contains traffic for port scanning, information gathering attacks. While the dataset is fairly large, the

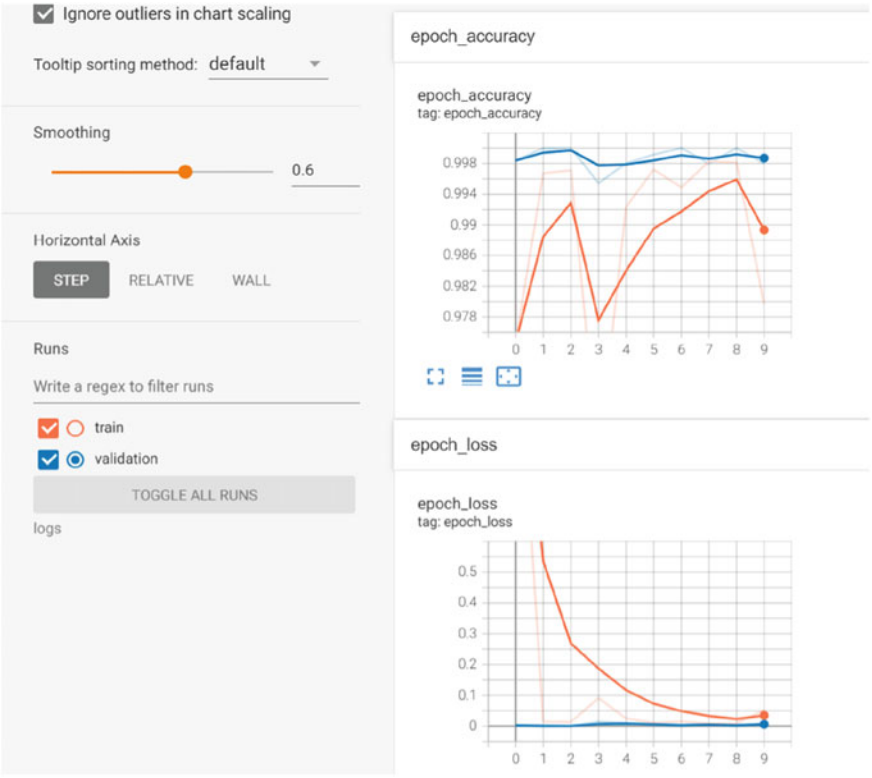


Fig. 15 AI Model performance analysis using TensorBoard Scalars

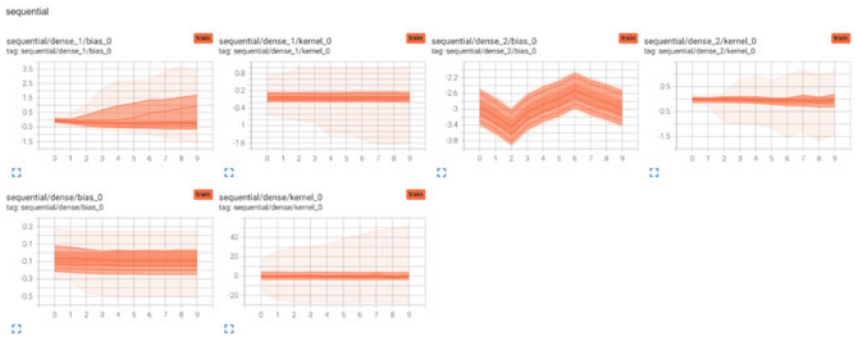


Fig. 16 AI Model performance analysis using TensorBoard Distributions

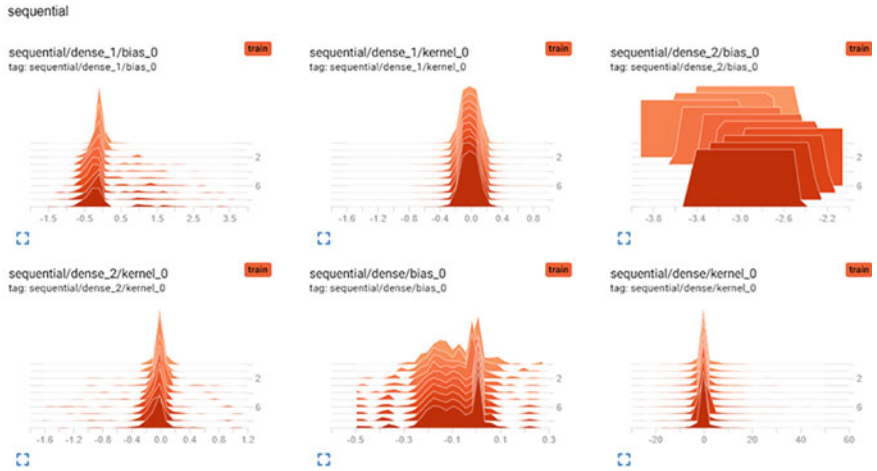


Fig. 17 AI Model performance analysis using TensorBoard Histogram

dataset needed more traffic flows with malicious traffic for the model to learn from and probably more variation in the attack types as well. A similar dataset IoT-23 has large number of traffic flow with the model able to learn from more than 745,000 flows of data with most of it being malicious. With large number of flows to test and validate the system, the AI model was able to achieve accuracies of over 99% while the accuracy was limited to just 46% for the VM dataset. A similar trend continues with the loss values where the VM dataset posted values of 0.69% while the IoT-23 dataset which features more columns of data, variations in data resulted in very little data loss value of 0.0093%.

The AI model proposed in this research will improve the effectiveness of traffic monitoring in a network and would suit the security requirements of SMEs. The model handles the process of network monitoring and can automatically take appropriate response actions relevant to a particular scenario. The traffic in a network will be continuously analysed by the AI model, and the anomalous traffic will be identified leading to an alert or appropriate action. The lack of dedicated security analysts is neutralised with the AI model's ability to handle automation of alerts and appropriate actions. The use of AI model answers the security challenges in an SME environment by placing emphasis on continuous monitoring and analysis for malicious patterns before a threat can be executed. By setting optimum threshold levels in malicious traffic resemblance with the trained data, a threat can be neutralised at the roots before transforming to a large-scale attack leading to business losses.

The performance metrics of 'accuracy' and 'loss' defines the model's suitability in an SME environment. With well-structured dataset an accurate AI model can be designed, and the performance of the current research model ascertains the same. The functions of ReLU (Rectified Linear Unit) and Adam optimizer can be adjusted to any network's requirement to achieve better results.

5 Conclusions and Future Work

The research aims to support and improve security standards in SMEs (Small and Medium Enterprise) by developing an artificial intelligence based malicious traffic detection system. SMEs often lack enough knowledge, infrastructure, and resources to implement a security system capable of actively analysing the network traffic for malicious activity. A system capable of automatically identifying threats and acting upon them simultaneously is unusual and out of bounds for SMEs, which is the motivating factor for the current research.

The project is set with a goal of developing an AI based tool capable of detecting threats in an SME network which is achieved by training an AI model to learn from training data consisting of attack traffic and regular traffic. Once the model has gained enough knowledge, it can analyse the network traffic for anomalies and provide alerts or appropriate response actions to mitigate the threat from becoming a reality.

The proposed AI model recorded an average performance with an accuracy level of 46% when tested using a dataset created in a virtual lab environment. To compare the model through various scenarios, further tests were carried out using alternate dataset, the IoT-23. The model performed exceptionally well with the evaluation parameters of accuracy achieving consistently more than 99% underlining the proposed model's performance. The sub-par performance with the primary dataset can be attributed to the lack of large number of malicious data flows within the VM dataset which limited the learning factor for the proposed model. The test using the IoT-23 dataset produced more commendable performance by topping the evaluation metrics of 99% accuracy and 0.0068% loss. The model indicates that the AI model provides the security boost necessary for SMEs to thwart any potential cyberattacks.

While the proposed model potentially brings invaluable augmenting capabilities in an SME environment, there are few hurdles that needs to be overcome to realize the true potential of AI system. The primary issue is with integrating the AI system to the current network infrastructure which might require additional infrastructure to aid the processing requirement of artificial intelligence. An effective integration mechanism of the AI model needs to be defined according to the network architecture with adequate security measures placed for the model itself as it processes every bit of information in the entire network. Another important factor would be the knowledge requirement to data pre-processing which directly affects the performance of the model. An incorrect feature extraction can lead to more damage as the system is based on the key features of the dataset. While there still exists some unknowns in the whole research due to the lack of real-time implementation, it is expected to provide additional security capabilities to an existing network architecture and could perfectly augment in making an informed decision about security incidents.

- Future Work

While the current model performed well in the experimental tests carried out in this research, the implementation in a real network is desirable which will be a part of future work. The AI model's learning patterns needs more analysis by running the

model for more iterations (100 epochs at the least). Future work will be a complete end-to-end fully integrated tool capable of analysing network traffic on the fly.

References

1. Abuadlla Y, Kvascev G, Gajin S, Jovanovic Z (2014) Flow-based anomaly intrusion detection system using two neural network stages. Comput Sci Inf Syst 11(2):601–622. Available at: http://www.doiserbia.nb.rs/img/doi/1820-0214/2014/1820-02141400035A.pdf. Accessed 21 Dec 2021
2. Ahmed A, Jabbar W, Sadiq A, Patel H (2020) Deep learning-based classification model for botnet attack detection. J Ambient Intell Hum Comput. https://doi.org/10.1007/s12652-020-01848-9. Accessed 18 Dec 2021
3. Aljabri M, Aljameel S, Mohammad R, Almotiri S, Mirza S, Anis F, Aboulnour M, Alomari D, Alhamed D, Altamimi H (2021) Intelligent techniques for detecting network attacks: review and research directions. Sensors 21(21):7070. https://doi.org/10.3390/s21217070. Accessed 17 Dec 2021
4. Al-Qatf M, Lasheng Y, Al-Habib M, Al-Sabahi K (2018) Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. IEEE Access 6:52843–52856. https://doi.org/10.1109/ACCESS.2018.2869577. Accessed 19 Dec 2021
5. Anderson M, Anderson S (2006) Guest editors' introduction: machine ethics. IEEE Intell Syst 21(4):10–11. https://doi.org/10.1109/MIS.2006.70. Accessed 28 Dec 2021
6. Beaver J, Symons C, Gillen R (2012) A learning system for discriminating variants of malicious network traffic. In: 8th Annual cyber security and information intelligence research workshop. Association for Computing Machinery, New York. https://doi.org/10.1145/2459976.2460003. Accessed 18 Dec 2021
7. Chamberlain L, Davis L, Stanley M, Gattoni B (2020) Automated decision systems for cyber-security and infrastructure security. In: 2020 IEEE security and privacy workshops (SPW). IEEE, San Francisco, pp 196–201. Available at: https://doi.org/10.1109/SPW50608.2020.00048. Accessed 18 Oct 2021
8. Chan L, Morgan I, Simon H, Alshabanat F, Ober D, Gentry J, Min D, Cao R (2019) Survey of AI in cybersecurity for information technology management. In: 2019 IEEE technology & engineering management conference (TEMSCON). IEEE, Atlanta, pp 1–8. https://doi.org/10.1109/TEMSCON.2019.8813605. Accessed 18 Oct 2021
9. Chou L, Tseng C, Lai M, Chen W, Chen K, Yen C, Ou T, Tsai W, Chiu Y (2018) Classification of malicious traffic using tensorflow machine learning. In: 2018 International conference on information and communication technology convergence (ICTC). IEEE, Jeju, pp 186–190. https://doi.org/10.1109/ICTC.2018.8539685. Accessed 19 Dec 2021
10. Dutta V, Choraś M, Pawlicki M, Kozik R (2020) A deep learning ensemble for network anomaly and cyber-attack detection. Sensors 20(16):4583. https://doi.org/10.3390/s20164583. Accessed 9 Dec 2021
11. Flammini F, Gaglione A, Mazzocca N, Pragliola C (2008) DETECT: a novel framework for the detection of attacks to critical infrastructures. In: Martorell et al (eds) Safety, reliability and risk analysis: theory, methods and applications, pp 105–112
12. García S, Grill M, Stiborek J, Zunino A (2014) An empirical comparison of botnet detection methods. Comput Secur 45:100–123. https://doi.org/10.1016/j.cose.2014.05.011. Accessed 25 Dec 2021
13. Hofstetter M, Riedl R, Gees T, Koumpis A, Schaberreiter T (2020) Applications of AI in cybersecurity. In: 2020 Second international conference on transdisciplinary AI (TransAI). IEEE, pp 138–141. https://doi.org/10.1109/TransAI49837.2020.00031. Accessed 18 Oct 2021
14. Jaigirdar F, Rudolph C, Oliver G, Watts D, Bain C (2020) What information is required for explainable AI?: a provenance-based research agenda and future challenges. In: 2020 IEEE

6th international conference on collaboration and internet computing (CIC). IEEE, Atlanta, pp 177–183. https://doi.org/10.1109/CIC50333.2020.00030. Accessed 20 Oct 2021
15. Johnson R (2019) 60 Percent of small companies close within 6 months of being hacked. Cybercrime Magazine. Available at: https://cybersecurityventures.com/60-percent-of-small-companies-close-within-6-months-of-being-hacked/. Accessed 9 Jan 2022
16. Kamoun F, Iqbal F, Esseghir M, Baker T (2020) AI and machine learning: a mixed blessing for cybersecurity. In: 2020 International symposium on networks, computers and communications (ISNCC). IEEE, Montreal. https://doi.org/10.1109/ISNCC49221.2020.9297323. Accessed 20 Oct 2021
17. Kingma D, Ba J (2015) Adam: a method for stochastic optimization. In: 3rd International conference for learning representations. Cornell University, San Diego. Available at: https://arxiv.org/abs/1412.6980. Accessed 4 Jan 2022
18. Kurihara K, Katagishi K (2014) A simple detection method for DoS attacks based on IP packets entropy values. In: 2014 Ninth Asia joint conference on information security. IEEE, Wuhan. https://doi.org/10.1109/AsiaJCIS.2014.20. Accessed 25 Dec 2021
19. Kurpjuhn T (2019) Demystifying the role of AI for better network security. Network Secur 2019(8):14–17. Available at: https://www.sciencedirect.com/science/article/pii/S1353485819300972. Accessed 17 Oct 2021
20. Macas M, Wu C (2020) Review: deep learning methods for cybersecurity and intrusion detection systems. In: 2020 IEEE Latin-American conference on communications (LATINCOM). IEEE, Santo Domingo, pp 1–6. https://doi.org/10.1109/LATINCOM50620.2020.9282324. Accessed 18 Oct 2021
21. Marr B (2021) What is the importance of artificial intelligence (AI)|Bernard Marr. Bernard Marr. Available at: https://bernardmarr.com/what-is-the-importance-of-artificial-intelligence-ai/. Accessed 17 Oct 2021
22. Mesevage T (2021) What is data preprocessing & what are the steps involved? [Blog] MonkeyLearn. Available at: https://monkeylearn.com/blog/data-preprocessing/. Accessed 25 Dec 2021
23. Mohammad R, Alsmadi M (2021) Intrusion detection using Highest Wins feature selection algorithm. Neural Comput Appl 33(16):9805–9816. https://doi.org/10.1007/s00521-021-05745-w. Accessed 22 Dec 2021
24. Mohammad R, Thabtah F, McCluskey L (2013) Predicting phishing websites based on self-structuring neural network. Neural Comput Appl 25(2):443–458. https://doi.org/10.1007/s00521-013-1490-z. Accessed 17 Dec 2021
25. Müller V (2020) Ethics of artificial intelligence and robotics, 1st edn. Metaphysics Research Lab, Stanford University, Stanford
26. Nguyen K, Hoang D, Niyato D, Wang P, Nguyen D, Dutkiewicz E (2018) Cyberattack detection in mobile cloud computing: a deep learning approach. In: 2018 IEEE wireless communications and networking conference (WCNC). IEEE, Barcelona. https://doi.org/10.1109/WCNC.2018.8376973. Accessed 22 Dec 2021
27. Ouchchy L, Coin A, Dubljević V (2020) AI in the headlines: the portrayal of the ethical issues of artificial intelligence in the media. AI Soc 35(4):927–936. https://doi.org/10.1007/s00146-020-00965-5. Accessed 28 Dec 2021
28. Pelley S (2019) Facial and emotional recognition; how one man is advancing artificial intelligence. Cbsnews.com. Available at: https://www.cbsnews.com/news/60-minutes-ai-facial-and-emotional-recognition-how-one-man-is-advancing-artificial-intelligence/. Accessed 17 Oct 2021
29. Press G (2016) Cleaning big data: most time-consuming, least enjoyable data science task, survey says. Forbes. Available at: https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#58fdfc6f637d. Accessed 25 Dec 2021
30. Sapavath N, Muhati E, Rawat D (2021) Prediction and detection of cyberattacks using AI model in virtualized wireless networks. In: 2021 8th IEEE international conference on cyber security and cloud computing (CSCloud)/2021 7th IEEE international conference on edge computing

and scalable cloud (EdgeCom). IEEE, Washington, DC, pp 97–102. https://doi.org/10.1109/CSCloud-EdgeCom52276.2021.00027. Accessed 19 Oct 2021
31. Shaikh R (2018) Feature selection techniques in machine learning with python. Medium. Available at: https://towardsdatascience.com/feature-selection-techniques-in-machine-learning-with-python-f24e7da3f36e. Accessed 7 Jan 2022
32. Svenkatsai123 (2021) Why tensorflow is so popular—tensorflow features. GeeksforGeeks. Available at: https://www.geeksforgeeks.org/why-tensorflow-is-so-popular-tensorflow-features/?ref=rp. Accessed 25 Dec 2021
33. Thaseen I, Poorva B, Ushasree P (2021) Network intrusion detection using machine learning techniques. In: 2020 International conference on emerging trends in information technology and engineering (ic-ETITE). IEEE, Vellore. https://doi.org/10.1109/ic-ETITE47903.2020.148. Accessed 21 Dec 2021
34. Wu P, Guo H (2021) LuNet: a deep neural network for network intrusion detection. In: 2019 IEEE symposium series on computational intelligence (SSCI). IEEE, Xiamen, pp 617–624. https://doi.org/10.1109/SSCI44817.2019.9003126. Accessed 19 Dec 2021
35. Yuan X, Li C, Li X (2017) DeepDefense: identifying DDoS attack via deep learning. In: 2017 IEEE international conference on smart computing (SMARTCOMP). IEEE, Hong Kong, pp 1–8. https://doi.org/10.1109/SMARTCOMP.2017.7946998. Accessed 17 Dec 2021
36. Zeadally S, Adi E, Baig Z, Khan I (2020) Harnessing artificial intelligence capabilities to improve cybersecurity. IEEE Access 8:23817–23837. https://doi.org/10.1109/ACCESS.2020.2968045. Accessed 18 Oct 2021

Video Camera in the Ambient Assisted Living System. Health Versus Privacy

The Results of the Online Questionnaire for the Healthcare Stakeholders

David Josef Herzog

Abstract Significant growth of the ageing population segment brings the necessity of additional healthcare investment. Besides somatic disorders, part of the older patient group suffers from Mild Cognitive Impairment and dementia. According to the World Health Organization, currently 55 million people worldwide suffer from dementia only. The projection is 75 million in 2030 and 132 million by 2050 (WHO, 2021). Mild Cognitive Impairment is often the first stage of dementia. Most of the patients with MCI and dementia are home-based. Ambient Assisted Living can improve the wellbeing of patients and their relatives without considerably raising the price tag for healthcare. In the current work, the questionnaire was created for healthcare stakeholders in order to conceptualize potential AAL for MCI patients. In this paper, the role of a video observation in AAL is analyzed with help of non-parametric respondents' group comparison.

Keywords Ambient assisted living · Mild cognitive impairment · Smart home · Video camera · Privacy · Vital signs · ADL

1 Introduction

The smart home is a home-based system of systems, which consists of monitoring sensors, connected together as a net to the analytical and automated appliances, with local and distant control of indoors management and environment. The smart home concept encompasses several utilitarian dimensions: in-house automated systems with control and monitoring; communication; health monitoring; entertainment [33]. Smart homes can be integrated with a smart IoT environment and permanent ubiquitous health monitoring under the aegis of Artificial Intelligence [3]. The medical aspects of AAL are divided into the supervision and monitoring part and support part. There are numerous AAL healthcare support systems, which can be subdivided into

D. J. Herzog (✉)
University Fernando Pessoa, Praça 9 de Abril, 349 Porto, Portugal
e-mail: 37491@ufp.edu.pt

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_3

several groups in accordance with the needs of patients, who have systemic impairments, or its narrow type, built for patients with a singular pathology. Comparative analysis tends to be more technical than medical [28]. For a similar reason, important medical parameters are often treated together with less important ones. From the healthcare point of view, home-based patients need to be monitored with a focus on the critical parameters. The most important parameters are vital signs, such as heart rate, blood pressure, breath rate, body temperature. The top causes of death include heart diseases, vascular damage, e.g. stroke, and respiratory diseases [59]. Most of the health conditions can be observed with help of different sensors, which give necessary information immediately. They help to control the condition of the person, health dynamic, follow up of the medication and treatment procedures, subject to biological compatibility [34]. In the case of Mild Cognitive Impairment, besides possible somatic dysfunctions, moderate memory, cognitive and psychiatric impairments create additional requirements for the AAL system [9].

2 Medical Data Registration in AAL

Vital signs and behavioural patterns can be registered with help of sensors. The data is collected via sensors' network and transmitted with help of middleware to the analytical tools. An alert is set on in the case of an emergency. The long-term problems can be followed up and general condition and level of representational independence evaluated.

2.1 Vital Signs

2.1.1 Heart Rate

Heart function is central to wellbeing. The resting heart rate of a healthy adult person is normally regular and has 60–80 beats per minute. It can vary with physical and emotional load, medication intake or underlying medical conditions. Heartbeat rate, regularity, volume, peripheral signs of blood perfusion give important information about the health status. Ischemic heart disease, arrhythmia, heart valve pathologies, vascular diseases are diagnosed and controlled with permanent checks of these characteristics. The heart rate can be measured by non-invasive Smart Wearable Sensing devices (SWS). Sensors or SWS are placed on the heart area, major arteries, peripheral arteries. They are pulse meters, electrocardiogram (ECG) sensors or SWSs, echo cardiac sensors. Blood perfusion in different body parts can be assessed by electroplethysmography and photoplethysmography techniques [35]. ECG, besides the heart rate and its rhythmicity, gives information about potential changes in the myocardial integrity and cardiac conduction condition. In some cases, sensors are seamlessly implemented as a part of smart shirts with textile-integrated non-invasive

magnetic sensors [55]. They often are worn as wristbands or similar wearable items. There are ways to assess heart rate distantly, for example with the help of Doppler radar [32].

2.1.2 Breath Rate

Respiratory diseases are the next major factor of pathology and death. The breathing rate of a healthy adult person is normally relatively regular, with 16–20 cycles of inhalations and exhalations per minute. The BR can be measured by wearable on-body sensors, wearable seamless shirt-integrated sensors, wearable breath analysis sensors [29]. The last ones can measure exhaled CO_2 to evaluate breath effectiveness. Sensors can be used to control exhaled acetone to control glucose metabolism for patients with diabetes [41]. Blood oxygenation is often measured by peripheral photoplethysmography and can be combined with pulse rhythm measurement and tissue perfusion level monitoring. Exist different wearable types of sensors for permanent use, designed as earbuds, finger rings or wristwatches [53]. There are also methods of distant respiratory rate monitoring with help of infrared Doppler sensors by the Kinect [37].

2.1.3 Blood Pressure

The normal arterial blood pressure of a healthy adult person is 110–140 mmHg systolic and 70–90 mmHg diastolic. The BP directly reflects myocardial function, heart valves integrity and functionality, and indirectly neuro-humoral heart rate, vascular tonus and blood volume regulation. The blood pressure can be measured by wearable pressure sensors, placed on the skin above the underlying subcutaneous artery. Usual places are: wrists, biceps, ankles. The measurement can be done with help of an inflated cuff, cuff-less pressure sensors, cutaneous tension sensors, photoplethysmographic sensors, measurement of pulse wave transit time, by combining two sensors along with the blood flow [63]. Some researchers propose ultrasound sensors [61]. Invasive methods are used to measure blood pressure in the main blood vessels to control their integrity [25]. Sometimes other bodily liquids require pressure measurements.

2.1.4 Body Temperature

Surface Body Temperature (BT) of healthy adult person usually homeostatically fixed around 36.6 °C if measured on the skin or in the oral cavity. It is an important parameter of metabolism. Core body temperature is higher and achieves a level of 38 °C. Temperature is measured through the contact body wearable sensors or distantly, with help of infrared sensors. Wearables are designed in different forms as bracelets, watches, jewelry, smart clothes. Non-contact infrared sensors are used less

often for body temperature measurement. However, systems based on the temperature detection proposed for the indirect cardiac rate measurement [19] or breath rate measurement [4].

2.1.5 Physical Activity

The normal gait as a physical process is divided into several phases, which repeat cyclically. The gait cycle usually is comprised of eight phases. It can be structured as two big sequential phases for the right and left leg, with stance taking 62% of the time and swing 38%. Each phase is then subdivided into four stages. One of the most important parameters is walking speed. In the metastudy, speed measurements are checked in 40 studies on more than 23,000 adults in different countries [10]. Normal walking speed is around 1.2–1.4 m/s, while pathological is supposed to be lower than 0.6 m/s [18]. Abnormal walking may reflect musculoskeletal pathology, neurological dysfunction, skin pathology or more general abnormality. Some researchers propose walking speed to be the sixth vital sign (the fifth is Body Mass Index, BMI). There are numerous ways to measure convenience in-home walking speed. Stationary sensors are based on a Doppler effect or electromagnetic tracking system, wide area pressure sensors, furniture pressure sensors, video and audio sensors. Wearable inertial sensors include accelerometers, gyroscopes, electromyographic sensors, pressure sensors, goniometers [54]. Accelerometers can be used to measure acceleration-deceleration and start/stop time, because they may change in some pathological conditions. More complex activities than walking are also routinely registered in most AAL systems. Utilities usage [17] or mounted sensors, signaling about refrigerator usage, doors and windows operation, other activities are usually monitored with help of different sensors and ontological models [6, 22]. Walking speed can be predicting sign for the future health condition [38], as well as Activities of Daily Living and Instrumental Activities of Daily Living (ADL and IADL) [43, 48]. ADL is estimated by the level of independence with: bathing, dressing, toileting, transferring, continence, feeding.

2.2 AAL for MCI and Dementia Patients

2.2.1 Diagnostic Aspects

Standard intelligence is generally reflected by IQ. The normal IQ is 85–115 (100 ± 15), ± 1 SD. MCI is diagnosed, when a permanent general IQ decline from a previously normal level is lower than 85 and higher than 70. Patients with dementia have stable IQ lower than 70. There are multiple methods of intelligence tests, various types of intelligence and diagnostic is non-trivial, but for simplicity IQ level is relevant enough, with more than 50% cases of MCI and dementia constituted by Alzheimer Disease (AD). There are many more causes for MCI and dementia: pseudobulbar affect, Parkinson's disease, frontotemporal lobar degeneration, Lewy body disease,

vascular diseases, traumatic brain injury, substance/medication use, HIV infection, prion diseases, Huntington's disease [42]. Dementia as a condition has specific modes of behaviour. Patients have problems memorizing necessary information and have difficulties performing everyday tasks. Reminders have to be more persistent, patient and avoid provoking a negative emotional reaction. Unfinished tasks, like open doors, gas stoves, the water supply may have adverse results. Misuse of objects creates danger for patients, their close relatives, carers, neighbors and visitors. Wandering without a clear objective, especially in an environment with obstacles, stairs, windows without protection is potentially harmful to dementia patients.

2.2.2 ADL

While MCI can be transitory between normal cognition and dementia, this pathology may stay for years and condition, in some cases, can improve. MCI is not easily diagnosed. ADL of patients with MCI is found to be lower than that of healthy old people. It can be connected to general activity during the day and to walking speed as well [20]. Comparative analysis of data sets also shows the difference in IADL. It can help to detect cognitive decline early [40]. Special service-oriented application (SOA) AAL platform "DemaWare" is created to address part of these issues, but partially based on obtrusive camera wearing for complex activity recognition [51].

In the European AD automated diagnostic project, Dem@Care patients' movement data is collected from the wearable ankle-mounted accelerometer. Additional data adaptation by creating more day and week time domains improves automated diagnostic [8]. Memory is one of the functions, which often suffers profoundly in dementia and MCI. It creates multiple problems, especially with repetitive tasks. Some AAL components are built to compensate for the loss of the function. There are attempts to create systems (HERMES) with the ability to remind about daily tasks, free time use [14]. The addition of smart objects, smart pillboxes, electronic calendars, smart white goods with reminders creates a better environment for patients with memory loss.

2.2.3 Spatial Movement

The connection between ADL tasks and cognitive impairment is well known and often reported [36]. Moreover, a strong positive correlation between quantitative gait characteristics and dementia is found in several studies. Mostly affected are step velocity and step length. A number of daily bounds (sessions, rounds) negatively correlates with cognitive status [27]. In some studies proposed prediction of the mental status change, based on the walking features, as speed, angular velocity and balance [30]. Other researchers found only a spatial correlation between gait and cognition for healthy old people [56]. However, in a major longitudinal study of 2938 mentally healthy participants, of which 2233 participants were reassessed and 226 developed dementia. Future decline correlated with walking speed. It is

also proposed that diminished mental processing speed plays a crucial role in lower walking speed. One standard deviation in walking speed shows a potential increase in the possibility of future dementia [60].

2.2.4 Sleep Abnormalities in MCI

While sleep deficit or disturbed rest often have negative impact on mental abilities, there are signs of the influence of cognitive dysfunction or conditions, leading to it, on the rest/activity patterns and sleep architecture. Sleep and wake pattern is often disturbed in MCI patients [16]. Ability to register patient activities in AAL during day and night are clearly demonstrated [52]. These findings can be supported on the level of EEG registration. These pathological changes can be predictive in the case of MCI and correlate with deterioration. Specific signs during non-REM sleep phase show future MCI in aging patients [52].

2.2.5 Mental Health AAL Applications

Monitoring of mental health cases in the AAL system can be divided into two big groups. One deals with psychiatric emergencies, such as suicide, psychotic events, major depression, alcoholic or drug-induced events. In every such case, the patient is potentially dangerous for himself through self-harm or self-neglect or can be dangerous to other people. The other type is long-term supervision, which can deal with emergencies, but mainly intended to be diagnostic and supportive in the case of a chronic condition with potential for physical and mental deterioration. Behaviour detection is based on the complex events analysis, activity time ratio, daily activity rates, complex event processing (CEP) and pattern recognition against existing pre-collected sets [58]. Other systems are focused on the RFID of GPS objects usage [21]. For emergency cases, different prediction models are based on the sensors combinations and behavioural data sets [1]. Other systems propose a connection with previous patients' records for better diagnostics [2]. Identification and prediction of abnormal behaviour with the help of neural networks (NN) are proposed by another team [26].

2.3 Video Observation in AAL

There is a number of observation methods in the AAL. Video cameras are one of the frequently chosen types of sensors. While they are very useful for communication, operating cameras for monitoring in AAL is supposed to be invasive and raise concerns about patients' privacy. There are intermediate solutions when the image is reproduced as an abstract imitation. However, this method of representation is relevant to other positioning systems as well. In addition, video stream demands higher

requirements for the data transfer and, subsequently, higher energy consumption. Several types of video observation are suggested. The usual RGB or RGB-D sensors are part of several AAL projects. Infrared and thermal cameras are another type. Optical sensors of different types are also utilized in AAL. All types are used for fall prediction and report, general well-being and medical observation, ADL measurement, communication and abuse prevention. The cameras can be static, movable, wearable, used as a group or in combination with other sensors [45].

As multysensor ubiquitous system can be costly and laborious to implement, video cameras represents reasonable price and effectiveness. The wearable camera, despite some specific advantages, e.g. demonstration the personal view and reflecting ADL and spatial activity, is restricted by lower compliance, certain level of user inconvenience and battery life limitations [12]. Privacy concerns are also high [5].

Laser-based optical systems can be utilized for positioning or fall-report. However, there are some difficulties to adopt security technologies for the AAL needs.

3 AAL Evaluation Methods

Existing and planned AAL systems have to be evaluated, validated and tested. Several methods are used for surveys and assessments. There are also theoretical methods, modeling in silico and practice, prototyping, live-in lab experiments and dry runs. After the start of the practical use data from the system and stakeholders is routinely collected and assessed with help of analytical tools.

3.1 Conceptual Stage

Health, WHO Quality of Life, WHOQOL survey and questionnaires are used for patients [39]. Healthcare specialists and caregivers formulate medical and social requirements and then the information is presented to technical specialists for conceptual validation, reference design and prototyping. The opinion of the social institutes and caretakers is also taken into account. At this stage, initial questionnaires are presented to stakeholders. The choice of every element in the system is based on the cross-section of requirements, from the skeleton to the user interface in later stages. Architecture is more influenced by technical standards.

3.2 Model

An initial phase demands the formulation of functional requirements, based on stakeholders needs. ADL, IADL with help of Prototype is envisaged with help of modeling, scenarios creation, personas and simulations. Data flow evaluation and model quality

control are used, with analysis of acquisition, transmission and usage [23]. In the first stage, sensors are chosen and calibrated. In the second stage ways and periodicity of transmission are analyzed, as well as security. In the third stage storage, backup and potential recovery are envisaged. Personas are used and portraits of potential users are generated. Father, more narrowly focused questionnaires can be utilized together with expert reviews.

3.3 Prototype

In the next stage, the prototype is tested for usability, functionality and interoperability by special tools or in living labs [13]. Experiments and questionnaires are instruments on this stage, as well as reviews [44]. While technical and instrumental measures are more objective and based on external metrics, questionnaires tend to be more subjective instruments and require different instruments of analysis. Both approaches have to be balanced in every case. For example, in living labs, there are different approaches to the information presented for the actors or patients. Some can be informed about testing, and others are instructed afterwards. In every case, the nature of questionnaires and surveys may differ.

3.4 Impact Assessment

Wider impact assessments include social, financial, industrial and political impacts, as demonstrated in the "Learnings from the 2019 and 2020 AAL Impact Assessment Final report." by http://www.aal-europe.eu and Technopolis group. There are three main types of frameworks of impact evaluation: Re-EIM, MAST and UTAUT. [31]. Re-EIM is an acronym of "Reach", "Effectiveness", "Adoption", "Implementation". Reach speaks about the type and size of focus groups, inclusion and exclusion criteria. Effectiveness measures all effects, including positive and negative impacts. Adoption calculates the number of stakeholders, who adopted the scheme or system. Implementation registers social, financial, administrative and other costs. Maintenance is a measure of long-term adoption, level of institutionalization or routine practice change. MAST is a Model for Assessment of Telemedicine. It is a multidomain approach to healthcare IT systems, which include Ambient Assisted Living. It is divided into three stages: a preliminary assessment, multidisciplinary assessment and transferability assessment. At every stage multifaceted analysis of social, administrative, financial, ethical and other aspects is done. Safety, effectiveness, maturity, possibility to be practically adapted are surveyed. The UTAUT stays for the Unified Theory of Acceptance and Use of Technology. It is a model framework, which consists of four elements: performance expectancy, effort 31 expectancy, social influence and facilitating factors [57]. There are also price value, hedonic motivation and habits in the extended model. The behavioural intention in the model is also

influenced by age, gender and experience. All factors lead to user behaviour. Every model can be used separately, partially or in full, with extensions, in combination with other models or provide elements fora specially constructed framework.

3.5 Questionnaires

One of the widely employed methods is a questionnaire. Economical and policy institutions create a significant impact on the AAL provisions. At the same time stakeholders: technical service providers, medical service providers, e.g. institutions and workers, end-users, such as patients and family carers, are the most important immediate players in the field of AAL. Current, prospective and retrospective assessment of stakeholder opinion through the survey is an important tool, addressing various aspects of the AAL. The level of acceptance, satisfaction, informed opinion or professional view is significant in the design and exploitation of AAL systems. Questionnaires can be subjective report tools but include objective information e.g. technical or biomedical parametric qualitative and quantitative data for comparison [7]. Objective information can be obtained by other means than questionnaire. Psycho-social factors, such as subjective acceptance, readiness to learn new technologies or to be involved in services with extensive AAL components are also important. Results can be presented as qualitative data, but scaled questions and frequency tables allow formal non-parametric statistical analysis.

3.5.1 Questionnaire Framework

The general framework depends on the questionnaire objectives and weights, attributed to certain metrics and variables. It is planned on the stages of conceptualization and questionnaire design [11]. Extensive literature review leads to the general understanding of the necessity, goals and the type of respondents the survey is targeting. While the concept influences every part of the questionnaire and every change in it, the nature of the expected category of interviewees is quite clearly split between the general population sample and the expert group. It affects the length of the survey and cognitive load, required from the respondents. The design depends on questionnaire structure, complexity, types of questions, wording, instructions, types of feedback and ways of administration. When the questionnaire is completed, it is tested, reevaluated, adjusted and implemented for data collection.

3.5.2 Types of Questionnaires

There are several types of questionnaires [46]. They depend on the research goal, focus group, type of questions, length and depth. Questions can be more qualitative or quantitative, open and closed, dichotomous, with simple dual answer options, or

multiple options, factual or opinion-based. Scaled questions of several Likert types are also often used to measure level or degree. Complex questions can be designed with internal subquestions and mixed options. Butteries of questions and specific batches can be arranged in sections or be spread randomly. Questionnaires are used in a direct interview, by mail, phone, online application, mobile app. The obtained information is often analyzed with the help of statistical instruments. The separate complex research class is formed by multifaceted surveys, designed with axes for several stakeholders. AAL4ALL project [15], run as an interdisciplinary, academic and industrial scheme. This project, for example, is created with a goal to answer questions about applicability, affordability and necessity to provide AAL as part of the communal healthcare program. It includes three major groups of respondents: patients as end-users, informal caregivers and healthcare and social care providers. Another complex approach is to present the same type of questions to different stakeholders in iterations, known as Delphi Survey, named after a well-known historical oracle. Questions are iteratively updated by answers and re-presented to the "oracle" panel of experts [49]. Results are scaled, which helps to rank importance inside of the questions' groups.

3.6 Questionnaire. Statistical Analysis

3.6.1 Reliability

Several well-established tests are applied for examination of internal consistency and reliability of the questionnaires. Split-half methods of different complexity are usually employed. Tabled results of the Likert scale responses undergone specific procedures. Cronbach's Alpha (tau equivalent), Revelle's beta, McDonald's omega, Guttman's lambda are described below. Test-retest reliability is checked by Cohen's kappa [47].

Cronbach's Alpha

Tau-equivalent or Cronbach's alpha is a measure of covariance between elements of the questions group. This parameter counts "dimensions" of the questionnaire and their interrelation with the help of the covariance matrix. Every respondent result is compared with the entire count of each observation. The higher number of "dimensions" and a stronger correlation between them gives higher results for alpha. Cronbach's alpha results are considered valid in the range of 0.8–0.9, with variations up and down. Alpha below 0.5–0.65 is considered to be a sign of low reliability, while higher than 0.9 shows redundancy or a high number of "dimensions".

$$\alpha = \frac{N\overline{c}}{\overline{v} + (N-1)\overline{c}},$$

where N is set power, $\overline{c}$ is average covariance for every element, $\overline{v}$ is average variance. Kuder–Richardson Formula 20 (KR-20) $\frac{n}{(n-1)} \times \left(1 - \frac{(\Sigma x \times y)}{v}\right)$ is a variant of alpha for binary items in dichotomous questions, where n *i*s a size of the sample, v is variability, x is the proportion of respondents answering positively, y—the proportion of respondents, answering negatively. KR-21 is used for questions with a close rate for questions, where m is mean count for the test:

$$\frac{n}{(n-1)} \times \left(1 - \frac{m \times (n-m)}{(n \times v)}\right).$$

Revelle's B

Beta is minimum or lowest split-half type test estimate of internal reliability.

$$\beta_{x_1} = \frac{c_{dp_1} - c_{dp_2} \times c_{p_1,p_2}}{1 - c^2_{p_1,p_2}},$$

where c is correlation/covariance, p_1 and p_2 are predictors and d is dependable variable. Beta is supposed to be more conservative estimator, than alpha—the later has tendency to "overshoot".

McDonald's Omega

Omega as a parameter is similar to alpha. Confirmatory Factor Analysis (CFA) of factor F for n variables X_n are connected by load l_n and influenced my the error e_n.

$$\omega = \frac{\left(\sum l_n\right)^2}{\left(\sum l_n\right)^2 + \left(\sum \sigma_{e_n}\right)^2}$$

The additional level of factor analysis is added. The covariance matrix of results is obliquely rotated and then so-called Schmid-Leiman or S-L second transform is performed.

$$C_m \approx \sum \left(FSS^T + D_m^2\right)$$

where C_m is square $p \times p$ correlation matrix, F is factors matrix $n \times n$, S is $n \times p$ matrix, D_m^2 is $n \times n$ diagonal matrix. F is transformed to create F_m second-order correlation matrix.

McDonald's omega is supposed to be a more reliable coefficient than Cronbach's alpha. Levels of 0.7–0.95 show reliability of the results.

Guttman's λ2

Coefficient lambda is similar to alpha and tau-equivalent of reliability. It comes in several grades, starting from lambda 1. The difference is that for alpha is used more random algorithm, while lambda has a lower level of randomness. Covariance between sums of items and average variances are included into the formula:

$$\lambda_2 = \left(1 - \frac{\sum C_i}{C_x}\right) + \sqrt{\left(\frac{n \times \sum \sum C_{i,j}^2}{C_x^2}\right)},$$

where $C_{i}^{2},_{j}$ is covariance between results.

Lambda can be employed for more complex tasks. Lambda is usually higher than Cronbach's alpha. Values for reliable test are 0.8–0.95.

3.6.2 Correlation. Non-parametric Methods

There is a significant difference between parametric and non-parametric analysis. Usual statistical analysis is based on mean, variance, standard deviation, analysis of probability distribution and ANOVA, analysis of variance. Parametric methods often consider continuous data. Non-parametric data does not have a usual tendency for normal distribution and often discrete ranks. Nominal data is presented by nominal categories, while ordered data is also scaled. In non-parametric methods, most important measurements are mode, median, quartiles and interquartile range (IQR) [50]. Sets of data can be compared between each other to trace independent or common sources of results [24].

Mann-Whitney-Wilcoxon Test

Mann-Whitney-Wilcoxon Test checks the equality of two ordinal sets of data. Sets can be of unequal size. MWW test calculates "unbiased" U parameter. It checks the equality of distribution and the supposed independence of sets.

$$U = N_x N_y + N_x \frac{(N_x + 1)}{2} - \sum R_x$$

where N_x is set X, N_y is set Y and R is the sum of ranks. Precision of the test is lower with significant difference between sets there is a possibility for type II error in this case.

Kruskal–Wallis Test

K-W test or one-way rank analysis of variance (ANOVA), calculates H parameter to test mutual dependency of data sets. KW test is designed for two or more sets. The size of data sets can be unequal, because the calculation does not involve paired comparison.

$$H = \frac{12}{n(n+1)} \sum_{x=1}^{m} \frac{R_x^2}{n_x} - 3(n+1)$$

where n is certain data set power, m is number of groups, R_x is rank of x and x is number of the data set.

Two or more samples are compared. Big differences between sets can cause type I error, giving false positive results.

Spearman's Rho

Spearman's Rho correlation coefficient is a rank analogue of Pearson coefficient. When Pearson coefficient is applied for continuous data, Spearman's Rho can be used for non-parametric ordinal data. Two sets of the same size, for example answers on two questions, are compared pairwise.

$$\rho = 1 - \frac{6 \times \sum \left(R_x - R_y\right)^2}{n\left(n^2 + 1\right)}$$

where n is number of results.

4 Research

4.1 Method

Medical and social requirements for the AAL are formulated on the conceptual stage. There are several ways to find answers, theoretical and practical. Any route gives only partial vision. The needs of caretakers and healthcare stakeholders are collected by questionnaires and expert suggestions. The process can be iterative, mixed and include detailed recommendations. The best approach is to try to encompass all these raised problems in one research to weight and compare information between subquestions. In the conditions of limited research complex questionnaire for healthcare stakeholders is the easiest way to obtain necessary preliminary answers. Web-based

questionnaire is easy to deliver worldwide. In current research Google Forms-based questionnaire was used.

4.2 Reliability

Questionnaire was tested on several runs before wide implementation. Reliability of the questionnaire is checked in Jasp 0.14.0.0—for scaled questions. Responses on 50 questions have McDonald's = 0.899, Cronbach's = 0.911, Guttman's 2 = 0.921. ω α λ The highest values in the questionnaire are: for Cronbach's is 0.920; for McDonald's is α ω 0.934; for Guttman's 2 is 0.932 (Table). Values above 0.9 may reflect (a) redundancy of the λ test—there are specially added questions in some dimensions to recheck values of the responses (b) multidimensionality of the test. 15 questions with opposite scales and negative results in the table were excluded from analysis. In this section is presented simple analysis and comparative description analysis between 76 sections and questions without group results comparison. In some cases Spearman's Rho pairwise correlation test is performed.

4.3 Focus Group

The Ambient Assisted Living system design has to be based on the opinion of the main stakeholders: healthcare professionals, technical stakeholders, administrative stakeholders and patients. Every opinion group is important, and the opinion has to be assessed appropriately.

Healthcare professionals represent a specific cross-section of society with a skilled understanding of patient's needs in specific conditions. Years of focused training and practice give a wealth of information about the needs and problems of home-based patients. Still, there is a range of possible opinions, dictated by the professional view, personal experience and wide scope of technical, social and organizational knowledge.

This study is based on a complex questionnaire. The questionnaire is presented to the healthcare workers, mainly medical doctors. In order to achieve the best possible combination, heterogeneous groups of medical professionals from different countries are included. In order to obtain as much and as wide information as possible and to keep the sample big enough despite complexity of the questionnaire all specialists with finished medical education or clinical psychology diploma were considered.

The respondents were reached via web of personal contacts and with help of social media. The main reason was to eliminate subjective element of self-report about profession and professional experience.

More than three hundred medical specialists were contacted in the USA, Canada, UK, Netherlands, Germany, Switzerland, Sweden, Greece, Israel, Armenia, Ukraine, Belarus and Russian Federation and asked to answer the questionnaire. Around 120

Table 1 Age and gender structure

	Number	Minimal-maximal	Mean, years	Median	Mode
Age	60	21–63	49.9	50	50
Gender: F	29	21–60	49.0	50	50
Gender: M	31	42–63	50.7	50	49

agreed to participate, of whom 60 answered all questions. Those who did not finish the questionnaire named several reasons for it: unknown topic, the length and complex nature of the questionnaire, heavy workload and shortage of time because of the COVID-19 pandemic. Country name was removed from the questionnaire for reason of required anonymity. However, there was no informally registered difference in the approach of specialists, depending on the country of practice or residence. Age and gender were collected for statistical necessities.

4.3.1 Age and Gender Structure

Age and gender were collected for statistical necessities. The age is from 21 to 63, with average age 49.9 years (Table 1).

4.3.2 Medical Profession

There are 60 respondents. Of those who answered, there are 41 medical doctors, 10 nurses, 4 paramedics, 2 dental medicine doctors, 3 clinical psychologists. Some information is available about medical doctors' specialization. Limitations arose from wide options of the question about the medical profession, so some doctors did not mention their specialization. The additional matter is a possibility to have more than one profession and report only one, often the most recent. Physician, MD—14. Psychiatrist, narcologist—9. Neurologist—3. Geriatric consultant—2. ONT consultant—1. Gynaecologist—2. Surgeon—1. Family doctor—1. Anaesthesiologist—4. Haematologist—1. Paediatrician—1. Dentist, DMD—2. Urologist—1. Traumatologist, Orthopaedist—1. There was no option to learn nurses' specializations. The age distribution by major professional groups is presented in Table 2.

4.4 Health Versus Privacy. Results

Data is obtained from answers to multiple-choice questions, matrix questions and scaled item-by-item questions. Healthcare and IT experience is projected on questions about medical aspects of technology implementation. Results are assessed with the help of descriptive and analytical statistics. Google Forms provide not only an

Table 2 Age distribution for professional groups

	Age				
	DMD	MD	Psychologist	Nurse	Paramedic
Valid	2	41	3	10	4
Missing	0	0	0	0	0
Mean	49.500	50.951	48.667	45.900	50.250
Median	49.500	50.000	52.000	49.500	50.500
Mode [a]	48.000	49.000	40.000	50.000	47.000
Variance	4.500	37.298	57.333	100.100	6.250
Range	3.000	32.000	14.000	35.000	6.000
Minimum	48.000	31.000	40.000	21.000	47.000
Maximum	51.000	63.000	54.000	56.000	53.000

[a] More than one mode exists, only the first is reported

easy way for questionnaire implementation but also a preliminary analytical structure. Numerical data and percentages are presented for scales and frequency tables, graphs and histograms provide visual information. All data can be extracted as an XML file. Excel and analytical software help to analyze data. JASP package is used for data analysis.

I this chapter are provided only descriptive statistics for statements about video camera in AAL. All group comparative analysis is not shown.

4.4.1 Descriptive Statistics

Sensors in AAL, Answers' Frequency Table

Table 3 with more than one possible answer per question per option.

Discussion: There is a tendency is recognition of the prominent role of video cameras in the AAL system—for security (68%) and for abuse prevention (75%). At the same time the invasive nature of video registration is acknowledged. 48% of the respondents think that video sensors can be switched on only at the time of the emergency. Still, wearable sensors are perceived as more invasive (47%) than video cameras (33%). 72% of the interviewees marked video camera as second best for communication. Microphone received 82% for the communicative purposes. Microphones are recognised as second best for the security and abuse prevention.

72% of the respondents believe that video camera is best sensor in the AAL system for the MCI patients. 68% support the importance of the microphones in the system. Preference of the video camera as one of the most important sensors, even if it can be switched on 24 h 7 days a week, raises question about privacy of the patient.

Table 3 Sensors in AAL, answers' frequency table

	Video camera	Microphone	Infrared positioning sensor	Mechanical pressure sensors	Wearable sensors	Temperature, air sensors
Most important sensors in AAL	38 63.3%	35 58.3%	38 63.3%	33 55.0%	49 81.7%	14 23.3%
The best combination of sensors in AAL	38 63.3%	34 56.7%	41 68.3%	35 58.3%	48 80.0%	18 30.0%
These sensors are not necessary for AAL	18 30.0%	14 23.3%	8 13.3%	10 16.7%	4 6.7%	35 58.3%
These sensors are too invasive to be switched on 24 h a day/7 days a week	20 33.3%	14 23.3%	12 20.0%	11 18.3%	28 46.7%	6 10.0%
These sensors can be switched on only in emergency	29 48.3%	21 35.0%	15 25.0%	14 23.3%	10 16.7%	21 35.0%
These sensors can be used for communication	43 71.7%	49 81.7%	6 10.0%	4 6.7%	7 11.7%	3 5.0%
These sensors are most informative	27 45.0%	21 35.0%	28 46.7%	28 46.7%	44 73.3%	10 16.7%
These sensors are least informative	8 13.3%	16 26.7%	13 21.7%	12 20.0%	8 13.3%	33 55.0%
These sensors are important for security	41 68.3%	29 48.3%	23 38.3%	11 18.3%	19 31.7%	11 18.3%
These sensors can help to prevent abuse	45 75.0%	30 50.0%	14 23.3%	9 15.0%	17 28.3%	7 11.7%

Sensors in the Ambient Assisted Living System for Patients with Mild Cognitive Impairment. Statements and Results

A. "The video camera is the best sensor in the AAL system for patients with Mild Cognitive Impairment".
B. "Microphones are very important in the AAL system for patients with Mild Cognitive Impairment".

Table 4 Descriptive statistics by statements, "Sensors in AAL"

Statement	Likert 6–10 (%)	Likert 8–10 (%)	Likert 10 (%)	Mean	Median	Mode	Quartiles	IQR
A	43; 71.7	26; 43.3	9; 15.0	6.6	7	7	5; 7; 9	4
B	41; 68.3	27; 45.0	8; 13.3	6.6	7	8	5; 7; 8	3
C	39; 65.0	26; 43.3	11; 18.3	6.5	7	10	4; 5; 7	4.5
D	38; 63.3	25; 41.7	11; 18.3	6	7	10	3; 7; 9	6
E	35; 58.3	23; 38.3	6; 10.0	6	6	1	3.5; 6; 8.5	5

C. "The video camera and microphones are too invasive to be used 24 hours a day/7 days a week as AAL sensors for patients with Mild Cognitive Impairment".
D. "Video cameras and microphones can be used in AAL only for emergency".
E. "Video camera and microphone can be used in AAL for patients with Mild Cognitive Impairment for communication only" (Table 4).

Discussion: Vitally important sensors have to be part of AAL for MCI patients according to most of the opinions. This includes wearable sensors, door and windows' sensors, smart water and gas leak sensors (These statements are not shown here). However, there is less consensus about invasive ones', such as video camera and microphone, or sensors, controlling positioning and gestures, and smart electricity sensors. While nearly two thirds of respondents generally support use of all these sensors, there is a disagreement. Support for video camera and microphone use is quite significant. There is a vision of necessity to use it not only for communication, even though it might be switched on 24/7.

Privacy of the Patient with Mild Cognitive Impairment in a Home Equipped with the Ambient Assisted Living System. Statements and Results

A. "Patient's health is more important than privacy issues in AAL".
B. "Privacy is more important than the patient's health in AAL".
C. "Privacy issues in AAL can harm the mental health of patients with Mild Cognitive Impairment".
D. "Patients with paranoid thoughts are not advised to live in a home with the AAL system".
E. "Emotionally sensitive patients are not advised to live in a home with the AAL system".
F. "AAL system is not more invasive than traditional healthcare" (Table 5).

Discussion: There is a tendency to put health problems before privacy problems, even though a significant part of respondents disagree with this less balanced, by their opinion, view. There is also no consensus about the danger of AAL systems for emotionally sensitive patients or those with paranoid thoughts. At the same time more than half of respondents see AAL system more invasive, than traditional healthcare

Table 5 Descriptive statistics by statements, privacy of patients with MCI in AAL

Statement	Likert 6–10 (%)	Likert 8–10 (%)	Likert 10 (%)	Mean	Median	Mode	Quartiles	IQR
A	41; 68.3	27; 45.0	15; 25.0	7	7	10	5; 7; 9.5	4.5
B	13; 21.7	7; 11.7	3; 5.0	4.1	3.5	3	2; 3.5; 5	3
C	35; 58.3	26; 43.3	14; 23.3	6.5	7	10	5; 7; 9	4
D	36; 60.0	18; 30.0	5; 8.3	6	6	8	5; 6; 8	3
E	32; 53.3	15; 25.0	3; 5.0	5.7	6	5	5; 6; 7.5	2.5
F	21; 35.0	9; 15.0	2; 3.3	4.8	5	5	3; 5; 7	4
G	36; 60.0	20; 33.3	6; 10.0	6	6	8	4.5; 6; 8	3.5

system. There is correlation between answers to questions A and G. Spearman's Rho $r_s = 0.34036$. *P*-value is 0.00779. There is also correlation between answers to questions E and G. Spearman's Rho $r_s = -0.25529$. *P*-value is 0.049.

5 Conclusion

The decision about the necessity for inclusion of video sensors into the general AAL and AAL for MCI patients design depends at the same time on the healthcare needs and technical solutions and feasibility.

For medical reporting, communication, abuse control and security video cameras are supposed to be suitable by 72%—they believe it is the best sensor. Microphones are supported by 68%. Still, 65% of the respondents agree that video cameras and microphones are too invasive to be switched on 24/7, and 63% think they can be turned on in the case of an emergency. Only 58% agree with the statement that video cameras have to be used for communication only.

The solution to the privacy problem can be technological. The use of another type of sensors, ways of the information presented and switching on only in the case of emergency are practical ways to lower the intrusive nature of observation and to improve patient's privacy. Infrared motion registration is seen as important by 77% of the respondents. 72% think gesture recognition is necessary for the AAL system for MCI patients. 72% accept the necessity of pressure sensors, mounted on the furniture for positioning. The complex manner of information collection gives an opportunity to compensate for the less obtrusive way of non-permanent video camera use.

References

1. Alam MGR, Abedin SF, Al Ameen M, Hong CS (2016) Web of objects based ambient assisted living framework for emergency psychiatric state prediction. Sensors 16(9):1431

2. Alam MGR, Kim SS, Abedin SF, Bairaggi AK, Talukder A, Hong CS (2015) Prediction of psychiatric mental states for emergency telepsychiatry. In: Proceedings of the Korean society of information science and technology, pp 1139–1141
3. Amin SU, Hossain MS, Muhammad G, Alhussein M, Rahman MA (2019) Cognitive smart healthcare for pathology detection and monitoring. IEEE Access 7:10745–10753
4. Andre N, Druart S, Gerard P, Pampin R, Moreno-Hagelsieb L, Kezai T, Francis LA, Flandre D, Raskin JP (2009) Miniaturized wireless sensing system for real-time breath activity recording. IEEE Sens J 10(1):178–184
5. Arning K, Ziefle M (2015) "Get that camera out of my house!" conjoint measurement of preferences for video-based healthcare monitoring systems in private and public places. In: International conference on smart homes and health telematics. Springer, Cham, pp 152–164
6. Augustyniak P, Ślusarczyk G (2018) Graph-based representation of behavior in detection and prediction of daily living activities. Comput Biol Med 95:261–270
7. Bethlehem J (2009) Applied survey methods: a statistical perspective, vol 558. Wiley, New York
8. Bian C, Khan SS, Mihailidis A (2018) Infusing domain knowledge to improve the detection of Alzheimer's disease from everyday motion behaviour. In: Canadian conference on artificial intelligence, pp 181–193. Springer, Cham
9. Blackman S, Matlo C, Bobrovitskiy C, Waldoch A, Fang ML, Jackson P, Mihailidis A, Nygård L, Astell A, Sixsmith A (2016) Ambient assisted living technologies for aging well: a scoping review. J Intell Syst 25(1):55–69
10. Bohannon RW, Andrews AW (2011) Normal walking speed: a descriptive meta-analysis. Physiotherapy 97(3):182–189
11. Brancato G, Macchia S, Murgia M, Signore M, Simeoni G, Blanke K, Hoffmeyer-Zlotnik J (2006) Handbook of recommended practices for questionnaire development and testing in the European statistical system. European statistical system
12. Cardinaux F, Bhowmik D, Abhayaratne C, Hawley MS (2011) Video based technology for ambient assisted living: a review of the literature. J Ambient Intell Smart Environ 3(3):253–269
13. Colomer JBM, Salvi D, Cabrera-Umpierrez MF, Arredondo MT, Abril P, Jimenez-Mixco V, García-Betances R, Fioravanti A, Pastorino M, Cancela J, Medrano A (2014) Experience in evaluating AAL solutions in living labs. Sensors 14(4):7277–7311
14. Costa R, Novais P, Costa Â, Neves J (2009) Memory support in ambient assisted living. In: Working conference on virtual enterprises. Springer, Berlin, pp 745–752
15. Cunha D, Trevisan G, Samagaio F, Ferreira L, Sousay F, Ferreira-Alves J, Simões R (2013) Ambient assisted living technology: comparative perspectives of users and caregivers. In: 2013 IEEE 15th international conference on e-health networking, applications and services (Healthcom 2013). IEEE, pp 41–45
16. Djonlagic I, Aeschbach D, Harrison SL, Dean D, Yaffe K, Ancoli-Israel S, Stone K, Redline S (2019) Associations between quantitative sleep EEG and subsequent cognitive decline in older women. J Sleep Res 28(3):e12666
17. Fell M, Kennard H, Huebner G, Nicolson M, Elam S, Shipworth D (2017) Energising health: a review of the health and care applications of smart meter data. SMART Energy GB, London, UK
18. Fritz S, Lusardi M (2009) White paper: "walking speed: the sixth vital sign." J Geriatric Phys Therapy 32(2):2–5
19. Garbey M, Sun N, Merla A, Pavlidis I (2007) Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. IEEE Trans Biomed Eng 54(8):1418–1426
20. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA (2008) Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimers Dement 4(6):395–405
21. Hodges MR, Kirsch NL, Newman MW, Pollack ME (2010) Automatic assessment of cognitive impairment through electronic observation of object usage. In: International conference on pervasive computing. Springer, Berlin, pp 192–209

22. Ihianle IK, Naeem U, Islam S, Tawil AR (2018) A hybrid approach to recognising activities of daily living from object use in the home environment. In: Informatics, vol 5, no 1. Multidisciplinary Digital Publishing Institute, p 6
23. Kara M, Lamouchi O, Ramdane-Cherif A (2017) A quality model for the evaluation AAL systems. Procedia Comput Sci 113:392–399
24. Kvam PH, Vidakovic B (2007) Nonparametric statistics with applications to science and engineering, vol 653. Wiley, New York
25. Lee S, Shi Q, Lee C (2019) From flexible electronics technology in the era of IoT and artificial intelligence toward future implanted body sensor networks. APL Mater 7(3):031302
26. Lotfi A, Langensiepen C, Mahmoud SM, Akhlaghinia MJ (2012) Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. J Ambient Intell Humaniz Comput 3(3):205–218
27. Mc Ardle R, Morris R, Hickey A, Del Din S, Koychev I, Gunn RN, Lawson J, Zamboni G, Ridha B, Sahakian BJ, Rowe JB (2018) Gait in mild Alzheimer's disease: feasibility of multicenter measurement in the clinic and home with body-worn sensors: a pilot study. J Alzheimers Dis 63(1):331–341
28. Memon M, Wagner SR, Pedersen CF, Beevi FHA, Hansen FO (2014) Ambient assisted living healthcare frameworks, platforms, standards, and quality attributes. Sensors 14(3):4312–4341
29. Mitchell E, Coyle S, O'Connor NE, Diamond D, Ward T (2010) Breathing feedback system with wearable textile sensors. In: 2010 International conference on body sensor networks. IEEE, pp 56–61
30. Mulas I, Putzu V, Asoni G, Viale D, Mameli I, Pau M (2020) Clinical assessment of gait and functional mobility in Italian healthy and cognitively impaired older persons using wearable inertial sensors. Aging clinical and experimental research, pp 1–12
31. Østensen E, Svagård I, Fossberg AB, Moen A (2014) Evaluation of ambient assisted living interventions-which tool to choose? In: Nursing informatics, pp 160–166
32. Otake Y, Kobayashi T, Hakozaki Y, Matsui T (2021) Non-contact heart rate variability monitoring using Doppler radars located beneath bed mattress: a case report. Eur Heart J Case Rep 5(8), p.ytab273. https://doi.org/10.1093/ehjcr/ytab273
33. Pal D, Triyason T, Funilkul S, Chutimaskul W (2018) Smart homes and quality of life for the elderly: perspective of competing models. IEEE Access 6:8109–8122
34. Patel S, Park H, Bonato P, Chan L, Rodgers M (2012) A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 9(1):1–17
35. Pantelopoulos A, Bourbakis NG (2009) A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans Syst Man Cybern Part C (Appl Rev) 40(1):1–12
36. Pereira FS, Yassuda MS, Oliveira AM, Forlenza OV (2008) Executive dysfunction correlates with impaired functional status in older adults with varying degrees of cognitive impairment. Int Psychogeriatr 20(6):1104–1115
37. Procházka A, Schätz M, Vyšata O, Vališ M (2016) Microsoft kinect visual and depth sensors for breathing and heart rate analysis. Sensors 16(7):996
38. Purser JL, Weinberger M, Cohen HJ, Pieper CF, Morey MC, Li T, Williams GR, Lapuerta P (2005) Walking speed predicts health status and hospital costs for frail elderly male veterans. J Rehabil Res Dev 42(4)
39. Queirós A, Dias A, Silva AG, Rocha NP (2017) Ambient assisted living and health-related outcomes—a systematic literature review. In: Informatics, vol 4, no 3. Multidisciplinary Digital Publishing Institute, p 19
40. Riboni D, Bettini C, Civitarese G, Janjua ZH, Bulgari V (2015) From lab to life: fine-grained behavior monitoring in the elderly's home. In: 2015 IEEE international conference on pervasive computing and communication workshops (PerCom Workshops). IEEE, pp 342–347
41. Righettoni M, Tricoli A, Gass S, Schmid A, Amann A, Pratsinis SE (2012) Breath acetone monitoring by portable Si: WO3 gas sensors. Anal Chim Acta 738:69–75
42. Sadock BJ (2020) Kaplan & Sadock's synopsis of psychiatry: behavioral sciences/clinical psychiatry

43. Salguero AG, Espinilla M, Delatorre P, Medina J (2018) Using ontologies for the online recognition of activities of daily living. Sensors 18(4):1202
44. Salvi D, Montalva Colomer JB, Arredondo MT, Prazak-Aram B, Mayer C (2015) A framework for evaluating ambient assisted living technologies and the experience of the universAAL project. J Ambient Intell Smart Environ 7(3):329–352
45. Sanchez-Comas A, Synnes K, Hallberg J (2020) Hardware for recognition of human activities: A review of smart home and AAL related technologies. Sensors 20(15):4227
46. Saris WE, Gallhofer IN (2014) Design, evaluation, and analysis of questionnaires for survey research. Wiley, New York
47. Sideridis G, Saddaawi A, Al-Harbi K (2018) Internal consistency reliability in measurement: aggregate and multilevel approaches. J Mod Appl Stat Methods 17(1):15
48. Snyder CW, Dorsey ER, Atreja A (2018) The best digital biomarkers papers of 2017. Digital Biomark 2(2):64–73
49. Spitalewsky K, Rochon J, Ganzinger M, Knaup P (2013) Potential and requirements of IT for ambient assisted living technologies. Methods Inf Med 52(03):231–238
50. Sprent P, Smeeton NC (2016) Applied nonparametric statistical methods. CRC Press
51. Stavropoulos TG, Meditskos G, Kontopoulos E, Kompatsiaris I (2014) The DemaWare service-oriented AAL platform for people with dementia. In: AI-AM/NetMed@ ECAI, pp 11–15
52. Taillard J, Sagaspe P, Berthomier C, Brandewinder M, Amieva H, Dartigues JF, Rainfray M, Harston S, Micoulaud-Franchi JA, Philip P (2019) Non-REM sleep characteristics predict early cognitive impairment in an aging population. Front Neurol 10:197
53. Tamura T, Maeda Y, Sekine M, Yoshida M (2014) Wearable photoplethysmographic sensors—past and present. Electronics 3(2):282–302
54. Tao W, Liu T, Zheng R, Feng H (2012) Gait analysis using wearable sensors. Sensors 12(2):2255–2283
55. Teichmann D, Kuhn A, Leonhardt S, Walter M (2014) The MAIN shirt: A textile-integrated magnetic induction sensor array. Sensors 14(1):1039–1056
56. Valkanova V, Esser P, Demnitz N, Sexton CE, Zsoldos E, Mahmood A, Griffanti L, Kivimäki M, Singh-Manoux A, Dawes H, Ebmeier KP (2018) Association between gait and cognition in an elderly population based sample. Gait Posture 65:240–245
57. Venkatesh V, Thong JY, Xu X (2012) Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 157–178
58. Veronese F, Masciadri A, Comai S, Matteucci M, Salice F (2018) Behavior drift detection based on anomalies identification in home living quantitative indicators. Technologies 6(1):16
59. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, Abbasi-Kangevari M, Abbastabar H, Abd-Allah F, Abdelalim A, Abdollahi M (2020) Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 396(10258):1204–1222
60. Welmer AK, Rizzuto D, Qiu C, Caracciolo B, Laukka EJ (2014) Walking speed, processing speed, and dementia: a population-based longitudinal study. J Gerontol Ser A: Biomed Sci Med Sci 69(12):1503–1510
61. Weber S, Scharfschwerdt P, Schauer T, Seel T, Kertzscher U, Affeld K (2013) Continuous wrist blood pressure measurement with ultrasound. Biomed Eng/Biomed Tech 58(SI-1-Track-E):000010151520134124
62. World Health Organization (2021) Global status report on the public health response to dementia
63. Yilmaz T, Foster R, Hao Y (2010) Detecting vital signs with wearable wireless sensors. Sensors 10(12):10837–10862

An Examination of How the Interaction Between Senior IT Managers and C-Level Executives Impacts on Cyber Resilience When Undertaking a Digital Transformation Project

Leia Mills and John McCarthy

Abstract Digital Transformation is a ubiquitous phrase that has become even more prolific in use following a recent world-wide pandemic that has forced many organisations to reconsider their processes, their business model, and their ways of working. However, whilst many organisations have embarked on a digital transformation, cyber resilience is not always considered, causing many transformations to lead to greater vulnerabilities and a higher exposure to cyber security breaches. The aim of this research is to develop a deeper understanding of the phrase 'digital transformations', who predominantly leads on them and if Cyber Resilience is a vital part of these transformation programmes. The research also considers whether the relationship between the C-Level and the IT Team has any bearing on whether Cyber Resilience is included in these transformation projects. The researcher has established that strong leadership was required to make certain that cyber resilience was included in digital transformation projects. In addition to this it must be seen and accepted that Cyber Resilience is an organizational issue for there to be a real impact. The relationship between the board and their view of Cyber Resilience was key to any sort of successful culture acceptance within their organisations. This acceptance also made a difference as to whether it was included in the business strategy. Even where the relationship between the board and the IT team was good, cyber resilience could be missed out. This was due to the lack of real understanding at board level. Therefore, cyber can never just be led by the IT function and should not be seen just as a technical issue.

Keywords Cyber resilience · Digital transformation · Pandemic · Cybersecurity · Cybercrime · Risk management · SME

L. Mills (✉)
Northumbria University London, London, UK
e-mail: leiamills@icloud.com

J. McCarthy
Oxford Systems, Oxford, UK
e-mail: John.mccarthy@oxfordsystems.co.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_4

1 Introduction

This research study aims to set forth the current understanding of both the terms Digital Transformation and Cyber Resilience.

The study also provides an organisation with the knowledge of how best to undertake a digital transformation that considers the organisation's Cyber Security posture, considers the relationship between the C-Level and the IT team and examines whether this relationship has any direct correlation on whether Cyber Resilience is included in the transformation projects that organisations undertake.

The terms Digital Transformation and Cyber Resilience do not often go hand in hand. Even the phrase Digital Transformation can conjure up many different connotations and meanings, while Cyber Resilience is often seen as the remit of IT teams, something to consider once a digital project is ready for roll-out across an organisation.

The pandemic (COVID-19) has forced many organisations into a de-facto transformation, with many of their employees working remotely, pushing them to make their networks accessible from different locations.

However, Cyber Security was rarely considered at the outset of these transformations. Many organisations reacted quickly to the pandemic, mobilising their workforces to work remotely, yet are only considering securing their networks now. Many are still utilising the same infrastructure they had when on-site working was the business model, relying on existing defences to protect them against increasingly sophisticated cyber threats [42].

According to Vial et al. [39], a digital transformation is "a process where digital technologies create disruptions triggering strategic responses from" an organisation [39]. Liere-Netheler et al [23], states that the "transformation affects the operation value creation process, enables new ways of doing business and leads to fundamental change in organisations" [23]. The paper continues that digital transformation does not just affect the workforce, but also affects the organisation as a whole. Governance, policies, and internal frameworks all need reviewing when a digital transformation is undertaken.

According to Hess et al. [15] "no sector or organisation is immune to the effects of digital transformation" and as a result, digital transformation has become a high priority for executive leadership. Nearly 90% of business leaders in the UK and the USA expect "IT and digital technologies to make an increasing strategic contribution to their overall business" [15].

In 2012, the World Economic Forum highlighted that cyber-resilience was not only a growing area of importance for businesses, but also a concept that was gaining greater traction. Bjorck [6], defined Cyber Resilience as "the ability to continuously deliver the intended outcome despite adverse cyber events" [6].

Society's increased reliance on cyber connectedness is threatened by ever-evolving cyber threats, with attack targets diversifying across both individuals and businesses, often causing unpredictability and risk for business. Therefore, cyber

resilience can also refer to "the ability of the system to prepare, absorb, recover and adapt to adverse effects" [24].

According to the latest Ponemon report, published in June 2020 [33], whilst "digital transformation is widely accepted as critical, the rapid adoption of it is creating significant vulnerabilities for most organisations". Further analysis carried out by Ponemon states that this may be due to conflicting priorities between IT teams and the C-Level, with only 16% of participants confirming that IT and the C-Level were fully aligned (Fig. 1).

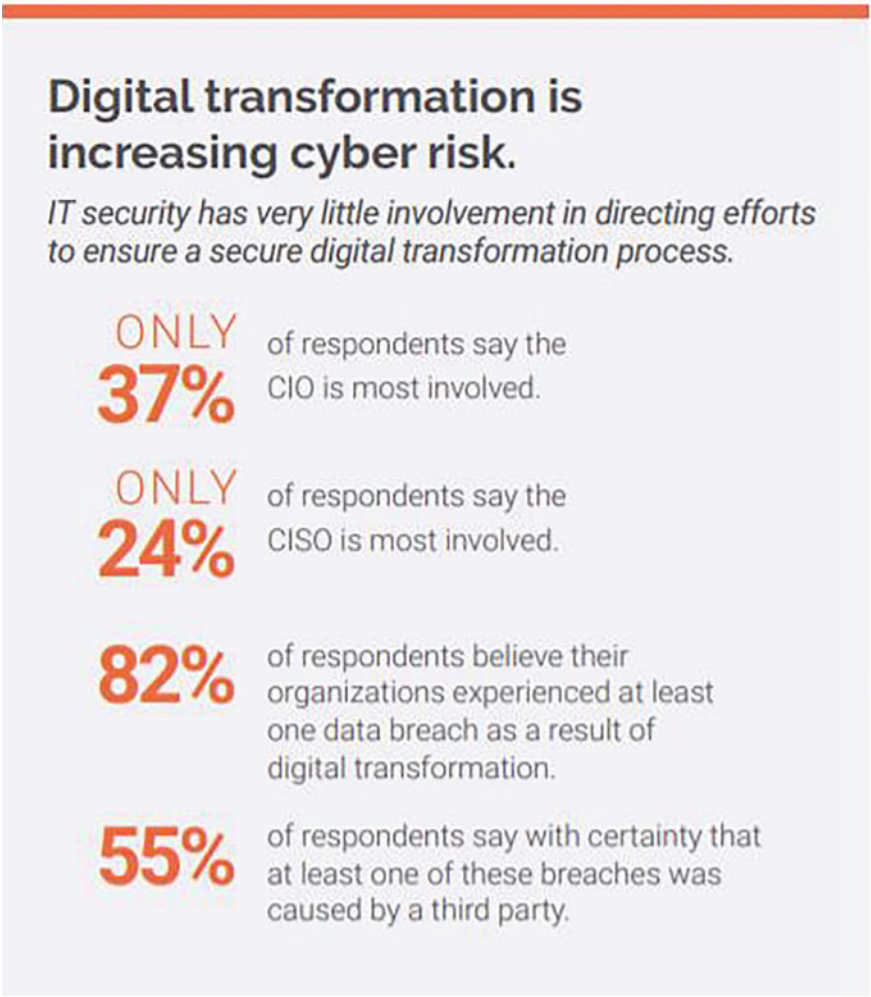


Fig. 1 Digital transformation is increasing cyber-risk [33]

While digital transformation presents a range of opportunities to companies, it also brings challenges. Some of the most prevalent are the evolving security risks resulting from an expansion of digital presence.

Who within an organisation, therefore, should take the lead on Digital Transformations, disruptive as they are? How can organisations ensure Cyber Resilience is a core part of the Digital Transformation of their organisation? Does the relationship between the IT team and the C-Level correlate in any way to whether Cyber Resilience is included in the Digital Transformation?

2 Literature Review

This literature review critically assesses the existing body of work discussing the topics of Digital Transformation and Cyber Resilience, and how industry and academia comment on the way organisations operate with respect to both these concepts.

A precise definition of a 'Digital Transformation' is still contested. In their industry paper, RSM state that 'adding technology to an existing service is not' a digital transformation. Mhlungu et al. points out that, depending on whose view you are taking, what is understood by a digital transformation differs, from an adoption of cloud technologies to upgrading a payment platform [29]. Others view digital transformation as moving from analog to digital, but Osmundsen et al. argues against this viewpoint and states that it moves beyond just a technical process, that it is a fundamental change to the way an organisation does business [32].

Kaplan et al. states that having unclear definitions does not help and that often leaders, when discussing 'digital', believe it to be a term that means their IT function gets an upgrade. The paper goes on to acknowledge that, due to fuzzy definitions, organisations will struggle to 'connect digital' to their business [18]. Organisations need to understand that it is not something they can plug and play but an ongoing process.

In fact, the majority of the literature agrees that Digital Transformation is a disruptor, fundamentally altering businesses and causing explosive growth [32, 39]. However, any dissonance between the business desire to digitally transform and to maintain cyber security, causes tension. Misalignment between IT teams and cybersecurity teams can lead to, not only a poorly executed digital transformation, but one that leaves the organisation open to cyber-attack [18].

Duc and Chirumamilla in their paper Identifying Security Risks of Digital Transformation highlight that organisations recognise the important role cybersecurity plays but that many have limited knowledge in 'identifying and managing risks of cybersecurity'. Further underpinning this, Nominet, in their industry paper, shows the place of cyber security in the age of Digital Transformation, stating that 'cyber security is seen as the single biggest risk when it comes to digital transformation' and that within their research they discovered that only 34% of those they surveyed had included cyber security as part of their transformation strategy. Changes in thinking

need to occur so that any transformation is reviewed against the whole organisation's security posture. But to do this, organisations need to be asking questions at the planning stage, questions that examine the impact of the transformation, not just in terms of opportunities, but also whether the security will hold up [31].

The Ponemon Institute report on Digital Transformation and Cyber Risk also discusses the misalignment between teams, focusing on the IT team and the C-Level, stating that 'conflicting priorities' between the IT team and the C-Level opens up vulnerabilities, with only 16% of their participants reporting an alignment between IT teams and the C-Level when looking to achieve a 'secure digital transformation process' [33].

The researcher chose, for the purposes of this research study, to define Digital Transformation according to both Vial and Liere Netheler, as set out in the background section above. Far more than simply a digital project improving specific individual business processes, Digital Transformation is a wider systemic change, a disruptive process. It fundamentally changes the ways businesses operate, both internally and externally—thus it requires strategic-level responses and direction to ensure success.

The literature also extensively discusses why digital transformations fail. It is important to understand the lessons that can be learned from failed attempts. PRINCE2 project management practitioners often evaluate previous lessons learnt before undertaking a new project. This approach is evidenced in the literature.

Failure to implement cyber resilience at the start of the digital transformation, is not the only obstacle to success. Primary themes include:

- Working in silos
- Lack of clear responsibility for transformation
- Rushed transformations
- Over-arching cyber security concerns
- Issues in leadership and governance
- Failed cultural shifts within an organisation

Boulton cites teams not being aligned and silos forming, along with the lack of clear strategic direction and leadership, as prevalent reasons for failed digital transformations. Silos create vertical structures in an organisation, whereas successful digital transformations need to cut across these silos, as part of a wider systemic change. Where this does not happen, all the elements required for a successful digital transformation are not present [7]. Silos also work against Cyber Resilience. Duc and Chirumamilla comments that Cyber Resilience is also recognised as a 'significant cross-cutting concern that influences various aspects of digital transformation'.

Matt el al explains in their paper Digital Transformation Strategies that often the disconnect between IT Team strategies and business strategies can cause problems. According to the paper, IT strategies often focus on the management of the IT infrastructure, whereas business strategies often look at the transformation of the products or processes or ways to enhance customer experience. Wilson agrees that a lack of leadership buy-in can lead to departments working as silos, making it harder for those working in cyber security to understand how the business is impacted by

threats [41]. Success in a digital transformation requires a close relationship between business strategies and IT strategies since a digital transformation encompasses both of these strategies. This close relationship should also be subject to reassessment as the digital transformation evolves, invoking an iterative process that reviews the alignment at each stage [27]. The presence of discrete silos, resistant to systemic changes, imperils success.

Matt el al also points out that often 'there is no clear answer to which senior manager should be in charge of a digital transformation' [27]. Lack of strong leadership and a failure to have a clear strategy seriously affects the outcome of an organisation's digital transformation.

Some companies rush towards their digital transformation, leaving out vital components such as user training, cyber resilience, and an acceptance that the culture of the business also needs to shift, all of which would help towards the successful implementation of their digital transformation strategy [1]. As the organisation BCG states, only 30% of companies successfully navigate a digital transformation, with most organisations viewing it as a 'work in progress'. In fact, BCG argues that transformations only succeed when they are incremental [2].

The cyber resilience question is discussed in GBM's Security white-paper, which argues that whilst Digital Transformation is becoming a priority for many industries, the changes can 'bring anxiety to an organisation's cybersecurity team and IT professionals' [30]. The paper continues by suggesting six technical pillars of cyber security that allow organisations to undergo effective transformations. In relation to this paper, the first four are the most pertinent:

- Align security strategy to digital transformation.
- Build collaborative Teams.
- Create a blueprint or architecture to implement a security strategy.
- Governance [30]

Cyber security needs to be recognised as a 'cross cutting concern' so that its influence is realised across all aspects of a digital transformation [14]. Mhlungu et al. states that a successful digital transformation needs a leadership that purposefully looks for security systems to mitigate risks [29]. They argue that governance should be deemed a 'critical level' when it comes to organisations succeeding in their transformation projects, and that failing to have effective governance leads to silos in working practices. They also suggest 'intra-organisational collaboration' with purposeful leadership [29].

Interestingly, Tsen et al. further underpins the need for governance and leadership related to Cyber Resilience. This paper states that 'sound' governance directly related to cyber resilience has to come from the top, with executive management 'setting the structure for cyber resilience [38].

Osmundsen et al. goes further and states that to effect a successful digital transformation, 'the organisation as a whole' needs to be part of the change, forming a 'supportive culture' where IT and business initiative are joined up [32]. This cultural change impacts all areas of the business. Doukidis et al. states that there are four pillars of a Digital Transformation: transforming the customer experience, business

processes, business model and organisational transformation, and argues that often failure to achieve across these domains occurs because there is poor cyber security [11].

One way to address the issue of a culture of poor cyber security is covered by Lee et al., pointing towards a digital culture that trains the right people and through this culture builds trust, so that information security enables proper management of security issues [22]. Organisations that have 'situational awareness', allowing them to understand the threat landscape they operate in, will be able to identify the vulnerabilities they will face and need to address. Policies that help organisations to control certain behaviours, and openly educate their staff, allow cyber security to become part of the culture [38].

Loonam et al. points to fostering trust within the organisation to bring about a cultural shift that promotes an organisation-wide understanding that makes cyber resilience part of any digital transformation. By effectively promoting this cultural shift towards cyber resilience being a business-wide issue, a fresh mind-set will follow. This fresh outlook will allow the business to work towards a desired organisational capability, which allows cyber-resilience to embed itself into 'their systems, structures and processes' aligning the security strategy with the business strategy. Senior management thereby starts looking at not just what the business does, but a greater understanding of how things are carried out that 'supports a more cyber-resilient organisation' [26].

The principles found in Tiirmaa-Klaar's paper regarding national cyber resilience can also be mapped to organisations, where he states that to have 'well-coordinated' cyber security systems, there has to be a leadership structure that makes the required decisions. Organisations should move away from viewing cyber as a technical problem that just requires a technical solution but views it as an 'integral part of an organisation's risk management and decision-making process', thereby leading to a security-minded leadership [37]. Loonam et al. agrees that the leadership of an organisation must have a vital role in 'maintaining the security of their organisation', with cyber security issues being 'every executives job' [26]. Bailey et al. further highlight that understanding the issues behind cyber resilience and digital transformation is 'quite different from effectively addressing it'. Therefore, they point out that cybersecurity is a 'CEO-level' issue due to the high-stake challenges posed by becoming cyber resilient [4].

The final cultural shift is focused on the C-level members. Senior management must be cognisant of the effects on all aspects of their business, strategic and operational to be able to deliver a successful digital transformation [25]. An interesting study by the MIT Sloan Management review pointed out that it is not just a transformation of an organisation that needs to be considered when embarking on a digital transformation, but the leaders themselves. That leaders must transform themselves to effectively lead a digital transformation, by fully believing in the organisation's values, creating a shared vision, and becoming digitally savvy themselves [35]. An educational drive of all levels in an organisation, but especially the board, in establishing a knowledge base of the risks associated with digital transformation projects should be carried out [33].

In summary, it is not just the relationship between C-Level and IT that needs to be taken into consideration when assessing the success of cyber resilience being embedded into an organisation's digital transformation. The strength of the organisation's leadership will also need to be assessed, along with an understanding of their view of Cyber Resilience and whether this is merely surface knowledge, for example, an understanding of the terms, or whether it is practitioner-level understanding. If the latter, the importance of cyber security and cyber resilience can be embedded at such a level that the board can make sure that the digital transformation is carried out in the light of required security needs.

A strong leadership function sets the tone for the whole organisation, engendering trust, openness, and transparency. It can also promote a culture that embraces cyber resilience as part of daily operations, educating staff at all levels. All the elements discussed in the literature review identify themes that need to be addressed to support an organisation to achieve a Digital Transformation with Cyber Resilience firmly embedded throughout the process.

3 Research Methodology

This section discusses the background to how research can be carried out, along with the methodology and design that has been used in this study including the choices that the researcher made to carry out their own research.

The purpose of this research is to understand participants perceptions about the relationship between IT and the C-Level when it came to Cyber Resilience being considered when an organisation embarked on a Digital Transformation.

Based on the research and the literature review, the overarching aim was to be able to provide organisations with guidelines to undertake a digital transformation, supporting a successful transformation project that takes full account of the company's cyber security posture.

3.1 Research Design

The research for this study was carried out in two phases. The first being a survey and the second being semi-structured interviews. This survey allowed the researcher to capture data from a wide range of people who felt they had been involved in a digital transformation. It also tested the participant's understanding of the terms Digital Transformation and Cyber Resilience against the view's held by both academic papers and industry papers.

The second phase of research was to gain further understanding of both what the participants meant when they answered 'yes' to having been part of a digital transformation and to examine comments on the relationship between the C-Level and IT. It gave the researcher the opportunity to examine further themes, that had surfaced

through the literature review such as whether Cyber Resilience is seen as a technical issue and the actual role of the leadership function in a Digital Transformation.

The semi-structured interview questions were developed to further illuminate some of the points raised by the participants, that would help with meeting the objectives of the research.

The importance of iteration is that it allows for repeatable processes to be carried out, each time building on previous knowledge and learning. The first iteration of research carried out was the survey, the second iteration, based on the evaluation and analysis process, was the semi-structured interviews.

When describing the people that have taken part in the study, the researcher has chosen to call them participants following Leavy's discussion on how different words, depending on the research are used to describe people, with respondent or subject often used in quantitative research and participant in qualitative [21].

3.2 Data Collection Method

Part 1: Survey

The initial survey was conducted through Microsoft Forms and those volunteers willing to be participants was sent the link via their email or through a personal message on LinkedIn. The interviews were conducted via Microsoft Teams video calling platform. The researcher found this to be an appropriate way of conducting the interviews since the country was in a worldwide pandemic, and therefore face-to-face interviews were not possible. The participants were based across a wide geographical area, with one participant being based in Turkey, which also meant that using a collaboration platform was suitable. The platform allows for the interview to be recorded and generates a transcript that can then be edited and checked back to the original recording. In addition the Teams platform sits within Microsoft's 365 platform and therefore data is encrypted at rest, making it a secure way of storing recordings and is in line with the ethical considerations of research, discussed in more detail below.

Part 2: Semi-structured interviews

The interviews were conducted online and therefore in an informal setting, since all occupants were at home, with the researcher first establishing a rapport with each participant, making sure that it was still convenient to proceed, confirming that the technical element was working, and the situation was conducive to an interview.

The questions written by the researcher were based on the objectives of the study, and the information gathered from both the literature review and the pilot survey. The questions were designed to not only understand more about the digital transformation the participants had already confirmed they had been part of, but also the level of appreciation of cyber resilience, and to map the participants views on the information found in academic and industry papers.

In line with the researcher's objectives, a particular point of interest was to understand how Cyber Resilience featured in participants Digital Transformations, and of course to further understand the relationship between the IT team and the C-Level. As the interviews were semi-structured, there were occasions when the participant made a comment, which prompted an unscripted question from the researcher to be asked to understand further. This enabled the interview to become more of a conversation making both interviewer and interviewee relax and investigate the ideas together. This method of flexibility around the interviewing process is in line with the qualitative methodology of research. As King et al. wrote, qualitative interviews use an 'interview guide'. In this instance, the researcher had a list of questions that they wanted to cover but was flexible enough to adjust the order of the questions, or change the wording to fit the interviewee, or even to follow a thought that led in an 'unanticipated direction' [19].

3.3 Research Philosophy

"Research philosophy is an important part of research methodology" [36]. Others define a research philosophy as "a strategy or architectural design by which the researcher maps out an approach to problem finding or problem solving" [16]. Thakurta and Chetty further explain that there are three philosophical approaches: Ontology—based on the nature of reality, Epistemology, which includes Positivism (data collection and hypothesis development), Realism (what is experienced by our senses and what is experienced by our sensations) and Interpretivism (differences between humans from a social viewpoint) and finally Axiology which looks at judgements, aesthetics and ethics [36].

In addition to choosing a research philosophy, there is also the subject of a research methodology. Often the most quoted methodologies are Qualitative and Quantitative. In aiding the choice of which methodology to follow when carrying out research, Leavy suggests that the researcher questions what it is they are trying to achieve and once the goal is achieved, how it is executed [21].

Quantitative research

Quantitative research is often based on statistics and the processing of numerical data, it enables the researcher to analyse and process large volumes of data to evaluate a hypothesis or test a theory. It also enables data to be quantified regardless of any personal opinions or feelings and allows easier comparison of data [5]. Agreeing with the impersonal nature this method can allow, Leavy reiterates that it can enable neutrality and objectivity. Enabling the outcomes to be related to the process of proving, disproving, or giving credence to established theories, whilst seeking to explain or evaluate [21].

Qualitative research

Qualitative research can often be viewed as a broad term. It 'focuses on how individuals and groups view and understand the world' [28]. It is also used to 'unpack the meanings' that people give to situations. It focuses on 'people's subjective experiences' to enable the researcher to create a 'depth of understanding' [21]. Basias and Pollalis describes it as an approach that provides answers to the 'what, how, when and where' type of questions, thus equipping the researcher to translate concepts and to draw final conclusions based on their own observations [5].

3.4 Research Approach

As discussed in the preceding section, quantitative research is less personal and more statistical in its outcome, as this study focuses on people's perceptions in comparison to formally presented definitions a qualitative approach was more appropriate. It provides 'rich detail' about the thoughts of individuals and supports the 'how' question, rather than the 'how many' question [13].

Therefore, throughout this study, the researcher adopted the qualitative methodology for extracting, reviewing and analysing the data. In tandem with this, the research philosophy of 'interpretivism' was taken, allowing the researcher to allow for differences in view when talking to people of the same organisation, the same sector, or having the same job title. Interpretivism supports the philosophy that 'all the stakeholders' approach is different, and they act according to their interpretation' [36].

The researcher started the initial collection of evidence through a pilot survey of 12 questions designed to get an understanding of participants views on both the topic of digital transformation and cyber resilience. A sample survey, in this research setting, is about 'determining the diversity' of people's views of a topic. This approach is often recommended when exploring meanings and experiences [17]. It also provides the researcher with the ability to reassess the methodology chosen for the study, by reviewing the responses from the participants. When determining the amount of people that the survey needed to reach, Fawcett et al. points out that there is no 'magic number' when conducting a qualitative research project [13]. The survey request was publicised on LinkedIn. The original request asked that people had taken part in or were going to be taking part in a Digital Transformation. 31 results were captured.

After the initial survey was sent out and the responses collected, an analysis was performed on the data, which is laid out in this study under Data Analysis.

Following the analysis, further research was required to delve deeper into some of the themes that transpired from both the analysis and the literature review. This approach fits well with Wadsworth's cyclical research process, which determines that analysis leads to new actions. In this case, the qualitative nature of the feedback led to semi-structured interviews which allowed for the investigation, at a more granular level, some of the responses. Conversing with people 'enables them to share their experiences and understandings' lending itself to the philosophy of interpretivism, where the realities of people's experiences form a body of meaning that is based on

different interpretations. Qualitative research has an exploratory character, which acts as an aid to gaining further understanding of participants feelings, thoughts, and perspectives [19].

3.5 Population and Sampling

For the initial survey, participants were recruited through a request via LinkedIn. As King et al. wrote, the major pitfall of recruiting in this way, does mean that often 'the sample is highly self-selecting'. To try and reduce this, the researcher attempted a 'purposive sampling frame' to get those required to come forward [19]. In this instance the idea was to get as many participants as possible who had been part of a Digital Transformation in the previous 24 months regardless of sector. This was to enable the researcher to get an overview of those involved in these types of projects to capture what the participants understood by the terms Digital Transformation, Cyber Resilience and also their perception of the relationship between the C-Level and the IT Team. The researcher also used opportunities when attending roundtable events to solicit the help of IT Directors, CTO's and CIO's.

For the semi-structured interviews, the researcher chose interviewees from the original survey sampling. The decision to choose the interviewees for the second part of the research was influenced by the literature review that the researcher had carried out which had suggested that strong leadership and a clear understanding of the business was an influential marker to enable a successful digital transformation that included cyber resilience. Therefore, those that had indicated leadership struggles in their original responses were chosen.

To repeat the earlier comments, the leadership of an organisation must have a vital role in 'maintaining the security of their organisation', with cyber security issues being 'every executives job'. With senior management being cognisant of the effects on all aspects of their business, strategic and operational to be able to deliver a successful digital transformation [25]. With this is mind, Participant 8 was chosen, since they clearly indicated that the relationship between the C-Level and IT team was strong, explaining that:

> we brief our employees on WHY we are doing security related tasks.

Participant 18 and 27 were chosen to further explore the comments they made about the relationship between the C-Level and the IT team and Participant 31 was chosen since they stated no cyber resilience had been included in the digital transformation project, they had been part of.

3.6 Data Analysis

The original survey was analysed based on the similarities and differences of the participants responses so that themes could be drawn and then discussed. The original survey also provided the structure for the interview questions which were designed to deepen the understanding of the comments made by the chosen participants.

The interviews were recorded, and a transcript was produced. The researcher then went back over the recording with the transcript to edit it so that it was correct, as automated transcript capture produces errors as not all speech can be captured correctly. Once this was done the researcher went through each script to look for similar lines of thought, contrasting ideas and any themes.

3.7 Ethical Considerations

When conducting research with individuals, whether they are representing themselves or an organisation, this aspect is paramount to making sure that both participant and researcher are clear on the terms of engagement. Informed consent allows participants to understand the nature of the research project and is an important aspect of any research that is carried out [34].

The original survey was sent to people who had agreed to be a volunteer and was emailed to them or sent via a personal message, following their agreement. Each participant for this first stage was advised that their responses would be anonymised and that a name was captured only for the researcher's information to allow for follow-up.

The interviews that were carried out, were based on the researcher hand-picking participants from the survey round. This involved the researcher emailing the participant to ask them if they would be happy to be involved in an informal interview further exploring the themes of digital transformation and cyber resilience. When the participant confirmed they were happy to continue, a set of the interview questions were emailed to them along with the terms of the interview, including the fact it would be recorded. Again, at this stage the participant was advised they could change their mind at any time.

Both the survey and interview results, including the data analysis, is anonymised and does not give any indication of any individual. This anonymisation protects those taking part in the study and ensures their confidentiality (Ali and Kelly 2012).

4 Data Analysis

Part 1: Survey

The initial collection of research was received using Microsoft Forms to create a survey of 12 questions designed to capture information about participants' initial thoughts on Digital Transformation, Cyber Resilience and the relationship between the IT team and the C-Level. The questions were:

- What is your sector and your position in the company?
- What do you understand by the term Digital Transformation?
- What do you understand by the term Cyber Resilience?
- Who do you feel should take responsibility for a Digital Transformation?
- Who do you feel should take responsibility for Cyber Resilience?
- Has your company undertaken a Digital Transformation in the last 24 months?
- Who led on it?
- Was Cyber Security/Resilience considered as part of the project?
- Is the IT strategy aligned with your company's strategic drivers?
- What is the relationship between the C-level and IT team at your company?

The researcher contacted a wide variety of people to ask them to participate in the survey. 31 people responded, from 11 different sectors. The breakdown of the roles detailed in Table 1.

Digital Transformation

As laid out in the background section of this study, the definition chosen for the meaning of digital transformation was defined using two sources. The first from Vial

Table 1 Breakdown of participants

Sector	Number of participants	IT teams	C-level	Other
Catering	1			1
Education	15	6	4	5
Events	1			1
Insurance	1			1
IT services	1		1	
Network systems	2		2	
NHS	1			1
Professional services	5	1	1	4
Retail	1			1
Telecommunications	2			2
Utilities	1			1
TOTAL	31	7	8	17

where he defines it as "a process where digital technologies create disruptions triggering strategic responses from" an organisation [39]. Liere-Netheler et al. defined Digital Transformation where "transformation affects the operation value creation process, enables new ways of doing business and leads to fundamental change in organisations" [23].

Whilst 3 out of the 31 participants recorded that they had not been part of a digital transformation, all participants provided a definition of the term.

When comparing the responses with those of Vial and Liere-Netheler, similarities occurred between both chosen definitions with the breakdown being that more participants had greater synergies with Liere-Netheler's definition. 18 out of the 31 records had similarities with Liere-Netheler, whilst only 4 out of the 31 participants were in line with Vial.

1 record had elements of both Liere-Netheler and Vial and 8 records did not relate to either definition.

When comparing the commonalities between the responses when referring to the definition of Digital Transformation, 22 out of the 31 participants used the term digital technologies, or adoption of digital technologies when referring to digital transformation. With 7 out of the 31 leaning towards a change to the whole organisation.

Two themes emerged:

1. Changes to technology using digital technologies
2. Changes to the entire organisation.

When it came to whose responsibility the execution of a Digital Transformation should be, the responses were as follows:

- 10 out of 29 said that responsibility for Digital Transformation should be the responsibility of everyone.
- 14 out of 31 said that the responsibility for Digital Transformation should be the responsibility of Executive Manager/C-Site
- 2 out of 31 said IT teams
- 5 out of the 31 said that Digital Transformation didn't sit just in one area, either a change team, or a specific business area outside of IT

In terms of those who led on the Digital Transformation, the data showed the following:

- 11 out of 31 said it should be Executive Management/C-Level
- 4 out of 31 said the IT Team
- 2 out of 31 said collaboration between IT and C-Level
- 11 out of 31 indicated 'Other'

Under the other category comments:

- Digital company—so role of the founders and those who showed greatest success in using digital elements (Education)
- HR and Operations (Event Management)

- Cross-organisation Digital Transformation Group (Education)
- Digital Team (Education)
- Different areas of the business (Education)
- Dean of Academic Development (Education)
- Business partners (Education)
- Digital Team (Retail)
- Student experience teams, Marketing/Comms (Professional Services)
- Whoever owns the process (Education)
- Digital team (Education)

From the data captured, the most commonresponse to who led on the Digital Transformation are the leaders of the business. Digital Teams, which are perceived as separate to IT teams, were repeatedly stated under 'other' category, with five respondents citing this area as taking the lead. Those who responded with Digital Teams as the leaders came from the education sector (4 participants) with one participant from the Retail sector.

4.1 Cyber Resilience

Bjorck defined Cyber Resilience as "the ability to continuously deliver the intended outcome despite adverse cyber events" [6]. Whilst Linkov and Kott stated that cyber resilience can also often refer to "the ability of the system to prepare, absorb, recover and adapt to adverse effects" [24].

In a review of the responses 8 of the 31 participants had similarities with Bjork, whilst 12 had similarities with Linkov and Kott. 1 record had similarities with both references. 6 of the participants had no clear definition of Cyber Resilience, with the 4 remaining having only a partial understanding.

When comparing the commonalities between the responses 16 out of the 31 participants linked Cyber Resilience with the idea of withstanding a cyber threat/cyber attack with the remaining answers linked it to continuous deliver. Therefore, two themes emerged:

1. Withstanding cyber threats
2. Continuous delivery of the business model

When it came to understanding whose responsibility it was, in relation to Cyber Resilience, the results showed the following:

- 11 out of 31 said that the responsibility for Cyber Resilience was the responsibility of everyone.
- 18 out of 31 said that the responsibility for Cyber Resilience was the purview of the Exec Management/C-Level
- 2 participants said IT

4.1.1 *Digital Transformation and Cyber Resilience*

What was the link between those responsible for Digital Transformation and Cyber Resilience? Of those who had been part of a digital transformation in the last 24 months, 18 participants indicated that those responsible for Digital Transformation should also be responsible for Cyber Resilience.

The remaining 10 responders saw different people responsible for these elements.

For those who saw the board as being the leaders of the digital transformation, when it came to cyber resilience, they saw it as the job for everyone to take responsibility for.

For those who saw everyone responsible for digital transformation, they felt the board should be responsible for Cyber Resilience.

For the person who felt IT should be responsible for Digital Transformation, they stated that the ultimate responsibility for Cyber Resilience should be 'senior management should have responsibility that the business is equipped and safe'.

The four participants that had other areas, such as change teams or cross team functions responsible for Digital Transformation, two of them felt Cyber Resilience was the remit of everyone, one said that it should be at 'Enterprise level—security leaders' and one said IT should 'take the main workload and input the processes in place', but also felt that 'all stakeholders should be aware of the processes in place to ensure everything continues to function correctly'.

In all of the responses, where they had seen a Digital Transformation, they all stated that Cyber Resilience was considered, except for one participant.

The researcher inferred that their current role was active when the Digital Transformation was taking place, therefore analysed which participants were part of the team who led on the Digital Transformation.

Out of the 7 C-Level level participants, four of them said that the C-Level should lead on a Digital Transformation, with one of them saying 'Everyone' should be involved. The other two felt it should be either 'cross department teams' led by a senior manager or the 'person accountable for the business process that is being transformed'. However, when asking those same set of people who should lead on Cyber Resilience, five of them said it should be C-Level level and two of them felt it should be everyone. Out of these 7 participants, only 3 who felt it should be C-Level responsibility to lead on the Digital Transformation confirmed it was the C-Level that did lead on it, that they were also accountable for Cyber Resilience and that Cyber Resilience was included in the transformation project.

Interesting two of the participants work for the same company in the Network Systems sector the third participant is a CISO at a University. When commenting on the relationship between the IT team and the C-Level, the CISO at the university explained that 'we have the CFO as the lead on the exec board for IT and Cyber to ensure organisational needs are met'. From the researcher's own experience, it is often the case that where there is no C-Level IT technical representative at board level, the CFO acts as the conduit.

The two participants who work for the same company both commented that it was 'strong' and 'very close', however interestingly the CTO went further and explained: 'We are agile enough to be able to ensure the IT team are deploying the technologies and our processes get updated as well as employee training are included. We brief our employees on WHY we are doing security related tasks. This is very important as it puts initiatives in context.'

From the initial survey results, the researcher was able to see that there was no one method or team involved in undertaking a digital transformation, and whilst most people answered in the affirmative that Cyber Resilience was included in the project, when comparing their definition to that of Bjorck and Linkov and Kott, not all articulated it correctly. It was also clear from the analysis that IT teams were not always seen as being the drivers for Cyber Resilience. Therefore, the researcher felt it was appropriate to move to a more in-depth discussion with some of the participants to get a deeper understanding of some of the realities behind their survey responses.

Part 2: Semi-structured interviews

Following the research cycle and based on the data collected from the pilot survey, the researcher reviewed the current position. Reflecting on the information collected and the analysis carried out, the next action in the process was to gain a deeper understanding of some of the comments made in the survey. To do this, new questions were written which were mapped against two of the objectives, and the researcher selected five participants from the original pilot survey based on some of their answers to the survey questions and based on the objectives of this study.

In keeping with the method of qualitative research, the questions were provided as a guide, so that the researcher could flex the questions to fit the flow of the interview, rather than rigidly sticking to a script [19].

Participant 31 was chosen as, in the original survey, they remarked that cyber resilience had not been included in the digital transformation project they had been part of, and Nominet had stated that "cyber security is seen as the single biggest risk when it comes to digital transformation [31]. Participant 8 was chosen due to the small nature of the organisation that they work in and the comment made about strong leadership, which was in line with BCG's comment that 'leadership commitment' was needed as a success criteria. Participants 18 and 27 were chosen due to the comments made about fractious relationships between the IT team and the C-Level. The Ponemon report references this as a stumbling block when it comes to Cyber Resilience highlighting that C-Level and IT Teams have conflicting priorities which can lead to vulnerabilities [33]. Participant 24 was chosen as they referenced a digital transformation being a cultural transformation which fed into Loonam et al.'s thought that fostering a culture of trust enables cyber resilience to become 'part of employee and team behaviour" [26].

The Digital Transformation project, triggers and Cyber Resilience

When asked about what transformation projects they were referring to in the survey, the triggers and whether cyber resilience was included, the responses were varied:

- According to Participant 31, whose role it is to support companies undertaking digital transformation projects, the triggers for the projects he was referring to was directly correlated to the Corona Virus Pandemic, which meant that companies were dealing with their workforce suddenly having to work from home, meaning there was a change in the way people connected and communicated.

Although saying that Cyber Resilience had not been included in the transformation project, the participant explained that:

> "the organisation had lots of branches operating individually, they had a MPLS in place which gave them connectivity, but the IT team didn't know what one team was doing from the other" so the SD-WAN that was put in gave them visibility and included a layer of security as well. "They had an idea that they needed to be more secure because they were all acting independently."

However, it was when they introduced the vendor who was installing the SD-Wan, then that was when security was introduced.

> that drove the project forward really because they were more focused on control of their users outside of the environment … Security was a project for them, but it wasn't their first priority until they really actually got an understanding that the SD Wan technology, the networking technology almost introduced them to that level of security. It was not necessarily an afterthought, but they were more focused on something else and then they realised 'actually we can now just bake it in'.

The trigger for participant 18 was changes to the accessibility polices, which caused the marketing department to push for an upgrade to their website. However, it wasn't just a website change, but also a process change, which caused disruption to their current processes since manual tasks were moved to automated tasks. When asked about Cyber Resilience Participant 19 responded:

> we built in plans to do a pen test …. We needed to not just test the new website but test our API to make sure because it was pulling data through from an internal database, that there was no way someone could form a request to that API and then access the data. It was built with defence in depth in mind to minimise whatever we could.

Participant 8 related how their business growing meant they had to scale their functions quickly and to help this moved a lot of their applications to the cloud. Due to this they had to think about the security of this. Implementing multi-factor authentication against each of the applications was really important in making sure they were secure, including running a tool that looked at where credentials had been shared.

> It actually looked at all the applications that the individuals in your company have allowed data to be passed through to, and it's amazing, right? 'Cause you will find things like this employee over there they've allowed their Google data to basically communicate to their smart watch, right? You think that's a normal thing for an individual to do, but that's their business account. Sending stuff to and from their smart watch and when you look at it like that, it's really quite interesting, and I think a lot of the users are kind of just yeah to all of this stuff without necessarily understanding, realizing, or knowing that all of this data is shared amongst all these cloud vendors. So I think, yeah, that for me is it is a key thing about all of this cyber security stuff is people

The transformation that participant 27 referred to was digitising the student on-boarding experience which was triggered by seeing the emergence of SSAS technologies and how other companies were doing things better by moving to digital options.

When asked how cyber resilience was included in these projects, the response was:

> We have an independent cyber security team and part of our governance process means they have to be brought in at an early stage.

With Participant 24, the reason behind the Digital Transformation was seeing other companies getting ahead.

> The market also forced us because other companies were beginning to use different technologies and we didn't want to stay behind them, so that was one of the reasons.

With regards to cyber resilience, hired consultants were brought in to help with the education of the staff and a training effort was carried out across the whole organisation. However, despite wanting to make sure people were included in the changes, they did not like it at all:

> It was a really good training plan, but we did it wrong. Everybody hated us, because they spent two whole days doing it, so they postponed all the training to the end of the year. It was too much for them all at once, so yeah, they hated us!

5 Methodologies and Frameworks

There are many frameworks that companies could employ to help support their journey towards Cyber Resilience, examples such as the ISO series and the NIST framework, however when asked whether a specific framework or methodology was used, Participants 31 and 24 did not use any framework. Participant 18 said that they relied on their own experiences and "just make it do what it needs to do in a secure way."

Participant 8 also responded in the negative, but said they did have Cyber Essentials, although the comment made was:

> we are cyber essentials. We're working through Cyber Essentials Plus, but that's more coming in from the angle that we just need the badge, it's almost embarrassing if you don't have the badge, because you just get asked right? So, we are just going through that to tick some boxes.

Only participant 27 confirmed that his organisation did adhere to frameworks, but despite being the IT Director, did not actually know what they were:

> They do but to be honest, I couldn't tell you what it was. So they are all members of the team that have come from that industry. You know they are cyber people so they're quite robust and yeah, disciplined in the way they do stuff and they follow specific methodologies for specific circumstances.

5.1 Cyber Resilience—A Business Problem or an IT Problem?

In exploring the question of the relationship between the IT team and the C-Level, the researcher wanted to examine where Cyber Resilience sat in the organisation and whether it was part of the business strategy and therefore part of the organisations culture, or whether it was viewed as a technical issue that IT would deal with. Wilson [41] wrote that cyber security in an organisation cannot be seen as only a technical issue because as organisations embark on Digital Transformations, it will assume an 'even greater importance as consumer-driven and people-driven experiences' are connected with technology. He further states that it is "as much a human behaviour and business culture problem" [41]. Interestingly the participants agreed in all but one case that cyber is still seen as a technical issue, rather than an organisational concern, therefore businesses still have some way to go on their cyber resilience journey.

Participant 31 felt that the experience they had was that the board only got involved in the sign-off from a contractual agreement, there was no push from them to find out whether cyber was included within the project. They felt that the:

> CTO is closer to the board so much more focused on the business whereas IT Team they are doing the day-to-day job, know exactly what their [users] need, they know exactly how their users work, the systems work and they know the technology that they need to go after …. There is a disconnect from a board perspective. There's always an awareness from up above, whether they understand it, is a different thing. Scare tactic messaging is definitely working its way into the board.

Participant 18 who acts as an IT Director felt that he was employed to be the expert on all things IT, including cyber. There was no separate cyber security strategy and whilst there was a cyber security group that met monthly, it was made up entirely of member of the IT team. When asked whether cyber resilience was viewed as a technical problem the answer was 'yes'.

> Security is viewed as a technical problem, although it's slowly changing. Took two attempts to get them to make information security training mandatory. It was point blank rejected first time around. I think like any organisation, business need always triumphs, security need does not at the moment.

Participant 8 highlighted the evolution at their company:

> A couple of years ago Cyber Strategy was definitely a bolt-on. But now it's very much aligned to business strategy. It's still in draft, but it does correlate strongly to our overall business strategy—they are separate, but they are linked. As we grew, and got more people, more resource, we were able to carve up tasks, it was evolution of business, rather than a lightbulb moment. Takes a lot of effort to do, but what's the risk of not doing it right, if you don't write this stuff down and you don't look at it and implement it ... and someone gets a bit of malware in your organisation, it's gonna end badly.

Further comment was made about examples he had seen when working with other organisations:

> You see it every day on the news …. And all the big global companies … declaring on a public basis … and I still think in business owners minds they just think 'oh yes, security is important, but do you know what, it ain't going affect me … it's still amazing what people think …. Or oh we're okay, because we've got a firewall.

In complete contrast participant 27 confirmed that at their organisation it was very much seen as a business issue, with users being educated and trained,

> we have a culture and an acceptance that this is the world we live in.

However, despite this, the culture is not embraced by all:

> the board does not like it—they suffer it, they don't play nicely together, they wish rather that it went away.

When asked what the trigger had been for them setting it out for the organisation the response was:

> Only for self-preservation and avoidance of embarrassment, not because they value it, because they don't want to get holed up in front of some select committee at Parliament to explain what on earth went wrong. It has only been taken seriously in this last round after GDPR came in and after the ICO got all their powers. That's when it really started being taken seriously and they said we'd better do something about this.

This is backed up by the GDPR Associates' report regarding the new powers that the ICO have, where there is an anticipation that the ICO 'will come down hard on businesses' that are unable to show that they have 'sufficient control and protection over the data they own' [14].

Participant 24 had a similar story:

> Getting buy was always hard because, especially the chairman of the board, he was really considered on the cost base and he always wanted to see the business value in a short time. But you know in Cyber it is not easy to measure the business value is such a short time. So, I spent most of my time trying to convince him. Too much time

In the end it was stories on the news that helped move things forward:

> Politicians speaking on the news and they begin to hear that and then it's okay something is coming and may miss something, so then they begin to support you.

6 Critical Discussion and Recommendations

A critical analysis of the research yielded some interesting results. All of the participants spoke about a digital transformation that partially included cyber resilience. Participant 8 confirmed it was brought in as the project developed, with four of the other participants confirming it was part of the transformation to varying degrees. Only one participant could be considered to evidence a robust organisational drive to make sure cyber resilience was part of any transformation, and this was down to having a separate cyber security team. It was therefore clear that Cyber Resilience

was not embedded in Digital Transformation at an organisational level, and this was primarily due to the differing appreciations of Cyber Resilience between the IT Team and the C-level.

Each of the participants confirmed that it was 'scare tactics' during briefings to the C-level by IT teams that prompted a change of attitude to Cyber Resilience at C-level. 'Scare tactics' by external parties have also forced C-level executives to think differently, with one participant confirming they only had 'cyber insurance' because the CFO had been scared at a CFO forum by anecdotes shared by others.

In other cases, an actual cyber security breach made C-level teams recognise the importance of cyber resilience. Participant 18 stated that when the Vice-Chancellors PA clicked on something she should not have, and therefore was brought face to face with the reality of their lack of cyber security and cyber training, a change in attitude at the C-level became evident.

Predominantly cyber security is still seen very much as a technical issue, despite much of the literature, both academic and industry papers, pointing to strong leadership being the deciding factor, not only in whether a digital transformation is successful but also whether the organisation has a strong cyber resilience culture. Many industry articles state that 'scare tactics' do not work due to the C-level not really understanding the extent of the cyber risks their organisations face [20].

However, the interview evidence contradicts this, demonstrating that where there has been a direct impact on an organisation, such as threat of a fine, perhaps from the ICO, changes at C-level have become evident. Furthermore, a real cyber breach such as a successful phishing email against an executive PA, has led to changes at the C-level.

In all but one instance, the cyber resilience maturity level is low, and comments were made that the C-level does not generally want to engage with the issue seriously. In the exceptional case, the C-level engaged, which led to the awareness of the centrality of cyber resilience permeating throughout the organisation. The relationship between the CIO and the security/IT Team was strong, balancing what users need against security requirements. It was also being pushed down from the C-level, embedded in their practices with training programmes and strong policies to back it up along with clear consequences for non-compliance. "When the executive team leads from the front, it sets the tone for the rest of the organization" [9].

When referring to those companies that become aware of cyber breaches at other companies, and do not respond with changes to their own security posture, [9] remarks that despite there being a lot of high-profile cases in the news, it often is not representative of what companies feel they will face, and therefore creates a feeling of unreality. The key, therefore, to improving cyber security posture is to make the risks very real to the C-level, exposing them to the wide range of potential risks through scenario planning. This will enable the C-level to perceive cyber risk, and therefore, in turn, the need for cyber resilience, in a context that they can understand rather than something they only see on the news [9].

The example of participant 27 shows how successful an organisation can be when they embrace cyber security with a holistic point of view, embedding it into all practises and policies. The organisation that participant 27 works for has found that

having a separate security team, brought in at the start of any project, ensures that an organisation is doing all they can to make themselves secure and robust. Policies that include consequences remind staff of the seriousness of any cyber breach, whilst training supports staff within the organisation to understand their role and responsibilities as part of the cyber-defences of their organisation. This, of course, resonates with the oft-observed maxim that 'people can be both the strength and weakness of an organisation'. In fact, regardless of how much cyber security and resilience is embedded into a digital transformation, should the organisation fail to enculturate cyber resilience into their staff, there will be serious questions about the effectiveness of their security and resilience posture.

Participant responses reinforced themes discovered in the literature review, leading to the following conclusion: the relationship between IT teams and the C-Level is important, and will have significant impact on whether Cyber Resilience is embedded within digital transformation. This reality has far-reaching implications.

These research results demonstrate that if the C-Level does not feel it is important to embed Cyber Resilience within the Digital Transformation, then they tend to ignore it. Examples from the primary research include one instance where the C-level executive management team flatly refused, twice, information security training for the organisation's staff until an incident directly affected the organisation. Clearly then, the key to embedding Cyber Resilience within Digital Transformation, is to embark on a program designed to elucidate the risks of cyber negligence, the behaviour and culture changes needed across the organisation, and the benefits of systemic cyber resilience. This must be targeted at the C-Level first since this serves as a vector to cascade resilient practices across the organisation. The literature concurs with this, "It is very important to embed positive cyber security behaviours, which can result to thinking becoming a habit, and a part of an organisation's cyber security culture" [3].

The following therefore are recommendations for organisations seeking to embed Cyber Resilience within the C-Level, leading to the conditions for successful Digital Transformations:

- Align the organisations priorities so that the business strategy includes Cyber Resilience as a key strategic pillar, supported by leadership governance.
- Ensure the C-Level has a clear understanding of what cyber resilience is, what it means to the organisation in terms of potential risks and benefits, and how it can be embedded in every individual digital project, thereby creating an organisational blueprint.
- Training pitched at C-Level members, for them to individually understand the impact of NOT having cyber resilience embedded within the organisation. The training should be contextual to both their role and to organisation's business activities.
- Encourage the C-Level to stop seeing cyber resilience as a technical IT issue, underlining it as an organisational issue.
- Have clear roles and responsibilities with a matrix style approach to transformation, to avoid silos and promote cross-organisational working.

7 Conclusions

The aim of this research was to develop a deeper understanding of the term 'digital transformation', who predominantly leads on them and if Cyber Resilience is a vital part of these transformation programmes. The research also considered whether the relationship between the C-Level and the IT Team has any bearing on whether Cyber Resilience is included in these transformation projects. To achieve this, objectives were set.

In this study, a comprehensive literature review was carried out and documented based on both academic and industry papers that commented on both Digital Transformation and Cyber Resilience, and how both have been approached using a variety of methods. The literature review showed that the precise definition of Digital Transformation is contested, and that Cyber Resilience is still not considered a vital part of the transformation process.

A pilot survey and semi-structured interviews were undertaken to collect primary research which enabled this researcher to examine what transformation projects had been carried out, the role of cyber resilience in them, and how teams and processes were enacted to underpin the cyber aspect. Each organisation had a different approach and different teams headed up the transformation projects. All of the individuals interviewed, except one, confirmed that Cyber Resilience is still seen very much as a technical issue rather than as a key pillar of the business strategy.

After examining both the literature and the primary research, it can be concluded that the C-Level's understanding of Cyber Resilience, and the importance that is attached to it from their viewpoint, has a large impact on whether Cyber Resilience is embedded into Digital Transformation projects. Both the literature and the primary research carried out in this study also confirmed that the C-Level plays an important role in making sure cyber resilience is part of the organizational culture, strengthening the enculturation of Cyber Resilience across the organisation, which increases the possibility of Cyber Resilience being successfully embedded in specific transformation projects. Where there is a lack of knowledge at the C-Level level, a dissonance occurs, between the IT Team's appreciation of the importance and need for Cyber Resilience, and the C-Level's appreciation for Cyber Resilience, and the necessary actions to support resilience across the organisation.

Following both the literature review and the primary research recommendations were made in line with the finding that the C-Level is crucially responsible for promoting a culture of Cyber Resilience across the organisation, which is the key to embedding cyber resilience into a digital transformation. The recommendations can be used to apprise the C-level of the importance of cyber resilience.

References

1. Assouline PGV (2018) Guide to digital transformation. [Online] Available at: https://www.infoq.com/articles/Digital-Transformation-Guide-1/. Accessed 19 Jan 2021

2. BCG (2021) Accelerating digital transformation. [Online] Available at: https://www.bcg.com/capabilities/digital-technology-data/digital-transformation/overview. Accessed 18 Jan 2021
3. Bada M, Sasse AM, Nurse JRC (2018) Cyber security awareness campaigns: why do they fail to change behaviour
4. Bailey T, Kaplan J, Rezek C (2014) Why senior leaders are the front line against cyberattacks. [Online] Available at: https://www.mckinsey.com.br/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Why%20senior%20leaders%20are%20the%20front%20line%20against%20cyberattacks/Why%20senior%20leaders%20are%20the%20front%20line%20against%20cyberattacks.pdf. Accessed 05 Feb 2021
5. Basias N, Pollalis Y (2018) Quantitative and qualitative research in business and technology: justifying a suitable research methodology. Rev Integr Bus Econ Res 7(1):91
6. Bjorck F (2015) Cyber resilience—fundamentals for a definition. [Online] Available at: https://www.researchgate.net/profile/Janis_Stirna/publication/283102782_Cyber_Resilience_-_Fundamentals_for_a_Definition/links/56824da408ae1e63f1eed116.pdf. Accessed 20 August 2020
7. Boulton C (2020) 10 reasons why digital transformations fail. [Online] Available at: https://www.cio.com/article/3248946/12-reasons-why-digital-transformations-fail.html. Accessed 1 Feb 2021
8. Davenport T, Westerman G (2018) Why so many high-profile digital transformations fail. [Online] Available at: https://hbr.org/2018/03/why-so-many-high-profile-digital-transformations-fail. Accessed 20 Jan 2021
9. Dennis P (2015) Helping the C-suite assess cyber risk. Risk Manage 62(8):28–29
10. Doukidis G, Spinellis D, Ebert C (2020) Digital transformation—a primer for practitioners. IEEE Softw 37(05):13–21
11. Doukidis G, Spinellis D, Ebert C (2020) Cybersecurity: linchpin of the digital enterprise. [Online] Available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Risk/Our%20Insights/Cybersecurity%20Linchpin%20of%20the%20digital%20enterprise/Cybersecurity-Linchpin-of-the-digital-enterprise.ashx. Accessed 1 Feb 2021
12. Duc AN, Chirumamilla A (2019) Identifying Security Risks of Digital Transformation - An Engineering Perspective. Norway, ResearchGate
13. Fawcett SE, Waller MA, Miller JW, Schwieterman MA, Hazen BT, Overstreet RE (2014) A trail guide to publishing success: tips on writing influential conceptual, qualitative, and survey research. J Bus Logistics [online]. 35 (1):1–16 Available at: https://onlinelibrary.wiley.com/doi/epdf/10.1111/jbl.12039. Accessed18th Sept 2020
14. GDPR Associates (2019) How has GDPR affected cyber security since it became law? [Online] Available at: https://www.gdpr.associates/how-has-gdpr-affected-cyber-security-since-it-became-law/. Accessed 26 Feb 2021
15. Hess T, Matt C, Benlian A, Wiesbock F (2016) Options for formulating a digital transformation strategy. MIS quarterly executive [Online] 15(2):123–139
16. Jamshed S (2014) Qualitative research method-interviewing and observation. J Basic Clin Pharm [Online] 5(4):87–88. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4194943/. Accessed 18 Sept 2020
17. Jansen H (2010) The Logic of Qualitative Survey Research and its Position in the Field of Social Research Methods. In: Forum Qualitative Sozialforschung / Forum: Qualitative Social Research 11(2). https://doi.org/10.17169/fqs-11.2.1450
18. Kaplan J, Richter W, Ware D (2019) Cybersecurity: Linchpin of the digital enterprise. [Online] Available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Risk/Our%20Insights/Cybersecurity%20Linchpin%20of%20the%20digital%20enterprise/Cybersecurity-Linchpin-of-thedigital-enterprise.ashx. Accessed 1 Feb 2021
19. King N, Horrocks C, Brooks J (2018) Interviews in qualitative research, 2nd edn. Sage
20. Kingpin (2021) Cybersecurity marketing—ditch the scare tactics to get boardroom buy in. [Online] Available at: https://kingpincomms.com/insight/cybersecurity-marketing-ditch-the-scare-tactics-to-get-boardroom-buy-in. Accessed 1 March 2021

21. Leavy P (2017) Research design: quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches. Guilford Publications, New York
22. Lee MX, Lee YC, Chou CJ (2017) Essential implications of the digital transformation in industry 4.0. J Sci Ind Res 76(3):465–467
23. Liere-Netheler K, Packmohr S, Vogelsang K (2018) Drivers of digital transformation in manufacturing in Hawaii international conference on system sciences, pp 3926–3935. Waikoloa Beach, HI. Available: https://www.researchgate.net/publication/323381114_Drivers_of_Digital_Transformation_in_Manufacturing. Accessed 19 August 2020
24. Linkov I, Kott A (2018) Fundamental Concepts of Cyber Resilience Introduction and Overview. [pdf] Arxiv Available at: https://arxiv.org/ftp/arxiv/papers/1806/1806.02852.pdf. Accessed 30 Aug 2020
25. Loonam J, Eaves S, Kumar V, Parry G (2018) Towards digital transformation: lessons learned from traditional organisations. Strateg Chang 27(2):101–109
26. Loonam J, Zwelgelaar J, Kuman V, Booth C (2020) Cyber-resiliency for digital enterprises: a strategic leadership perspective. IEEE Trans Eng Manage 1–14
27. Matt C, Hess T, Benlian A (2015) Digital transformation strategies. Bus Inf Syst Eng 57(1):339–343
28. McCarthy J (2019) An Examination of the Impact of E-Business Evolution Within Small and Micro Businesses [pdf]
29. Mhlungu M, Chen J, Alkema P (2019) The underlying factors of a successful organisational digital transformation. SA J Inf Manage 21(1):10
30. Nofel H (2019) The unspoken truth, the role of cybersecurity in breaking the digital transformation deadlock. [Online] Available at: https://gbmme.com/wp-content/uploads/2019/10/GBM-Security-Whitepaper-2019.pdf. Accessed 15 Jan 2021
31. Nominet (2021) Cyber security in the age of digital transformation. [Online] Available at: https://media.nominet.uk/wp-content/uploads/2019/07/Cyber-Security-in-the-Age-of-Digital-Transformation.pdf. Accessed 22 Jan 2021
32. Osmundsen K, Iden J, Bygstad B (2018) Digital transformation: drives, success factors, and implications. Mediterranean, AIS Electronic Library
33. Ponemon Institute (2020) Digital transformation and cyber risk what you need to know to stay safe. [Pdf] CyberGRX Available at: https://get.cybergrx.com/ponemon-report-digital-transformation-2020/. Accessed 15 August 2020
34. Sage publications the ethics of social research. [Pdf] Available at: http://www.sagepub.com/sites/default/files/upm-binaries/34088_Chapter4.pdf. Accessed 21 Sept 2020
35. Schrage M, Pring B, Kiron D, Dickerson D (2021) Leadership's digital transformation. [Online] Available at: https://sloanreview.mit.edu/projects/leaderships-digital-transformation/. Accessed 2 Oct 2021
36. Thakurta SG, Chetty P (2015) Understanding research philosophy. [Online] Project Guru. Available at: https://www.projectguru.in/research-philosophy/. Accessed 20 Sept 2020
37. Tiirmaa-Klaar H (2016) Building national cyber resilience and protecting critical information infrastructure. J Cyber Policy 1(1):94–106
38. Tsen E, Ko RKL, Slapnica S (2020) Organisational cyber resilience and its influence on cyber attack outcomes. University of Queensland, Queensland
39. Vial G (2019) Understanding digital transformation: a review and research agenda. J Strateg Inf Syst [online] 28(2):118–144. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0963868717302196. Accessed 20 August 2020
40. Wadsworth Y (1998) Action research international: what is participatory action research? [Online] Available at: http://www.aral.com.au/ari/p-ywadsworth98.html. Accessed 4 March 2021
41. Wilson S (2020) The pandemic, the acceleration of digital transformation and the impact on cyber security. Comput Fraud Secur 2020(12):13–15
42. Wilson M (2020) Digital transformation cybersecurity, IoT and you. [Online] Available at: https://rsmpartners.com/documents/Digital/RSM-Executive-Briefing-Papers/Digital-Transformation-Cybersecurity-IoT-&-You.pdf. Accessed 23 Jan 2021

Digital Twin Technologies, Architecture, and Applications: A Comprehensive Systematic Review and Bibliometric Analysis

Rosemary Ofosu, Amin Hosseinian-Far, and Dilshad Sarwar

Abstract Digital Twins, as a suite of technologies is progressively developing significant momentum in several fields of study. Various research works have been conducted outlining the concept, the underlying technologies, general and context-specific architectures, and applications. This study has been undertaken to identify relevant research areas, key authors, publishers, and geographical distribution of publications on digital twins through a systematic review and bibliometric analysis, to inform the trajectories of future research in the field. A keyword-based search for journals was first conducted in Web of Science Core Collection to obtain documents relevant for this study, and a systematic review was performed in accordance with the PRISMA guidelines. A bibliometric analysis was then performed on the extracted data using the VOSviewer software. The Tableau software and Microsoft Excel were also used to analyse and visualise some of insights derived from the analysis.

Keywords Digital twin · DT · Architecture · Reference model · Digital twin application

1 Introduction

David Gelernter, an American computer scientist from the Yale University predicted digital twins back in the 1990s; He referred to the notion as "Mirror World" and described it as a technological voodoo figurine which would allow the world to be seen more profoundly, through the use of massive open software masterpieces [1]. The concept, according to Grieves and Vickers [2], gained recognition in the year 2002 when the University of Michigan presented the development of a Product Life-cycle Management (PLM) centre to industry. The conceptual idea for PLM had all

R. Ofosu · A. Hosseinian-Far (✉) · D. Sarwar
Department of Business Systems and Operations, University of Northampton, Northampton NN1 5PH, UK
e-mail: Amin.Hosseinianfar@northampton.ac.uk

D. Sarwar
e-mail: Dilshad.Sarwar@northampton.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_5

the features of a digital twin: the tangible asset, the virtual asset, link for data flow from tangible asset to virtual asset, and vice versa. Digital twin can be described as the virtual delineation of a physical asset using data and stimulators for optimisation, real-time predictions, observing, and management, for superior decision making [3]. According to Maria [4], the key components of digital twins are the model of the physical asset, the timeseries data captured with sensors from the asset, the unique identifiers that connect the physical asset and the virtual model, and its monitoring capability using cameras, sensors, etc. The implementation of digital twin facilitates cost reductions, improved product design [5], and predictive and preventive maintenance [6]. In the bid to implement digital twin, various architecture and reference models have been proposed. As per Talkhestani et al. [7] and Aheleroff et al. [8], there has not been a clear holistic reference architecture for digital twin implementation: Different digital twin architectures have been proposed in different situations. Currently, digital twin is being applied in numerous sectors: Manufacturing, Healthcare, Production, Education, City Management, and many others.

There has been a number of bibliometric analyses in several aspects of digital twins: Ante [9] performed a bibliometric analysis on digital twin technology in smart manufacturing and industry 4.0, Radanliev et al. [10] evaluated artificial intelligence and IoT cyber-physical systems in industry 4.0, and a bibliometric review was completed for digital twin enabled smart industrial systems by Ciano et al. [11], among others. This study however sought to comprehensively assess and analyse trends in overall digital twin research in terms of publication volumes, geographical dissemination, key authors, major journals, prominent publishers, countries and organisations, and research areas to determine the current and future trajectories for digital twin research. This was achieved by identifying and reviewing existing literature in relation to digital twins, its architecture, and applications, selecting keywords and conducting a systematic review, and adopting, and applying appropriate analysis tools for conducting a thorough bibliometric analysis, and discussing the findings. This helped to provide better insight on the state of existing literature in digital twins, and to inform future directions of research within the field.

2 Related Work

2.1 DT Definitions

Digital Twin (DT), which has been referred to as a fundamental enabler to digital transformation by Kritzinger et al. [12], has been defined by several researchers after its inception by Michael Grieves in 2002. In [2], digital twin was described as a holistic virtual view achieved by stripping information from a physical asset. Majumdar et al. [13] characterised digital twin as a fundamental paradigm that will include measurable data of material level attributes with high-level sensitivity. Wright and Davidson [14] defined digital twin as a workable virtual model of a physical

object. The concept was similarly described by Rosen et al. [15] as accurate models of the present state of a process and its conduct in collaborating with its ecosystem in the real world. Madni [16] referred to the concept as a dynamic virtual model of a service, process, or system. Other researchers, including Barricelli et al. [17], Jones et al. [18] and Schleich et al. [19], in their definitions, related digital twin to computer-based models that mirror a physical system. This explains how the virtual model stimulates, emulates, and mirrors the physical asset using information assessed from the physical asset. On the other hand, Alam and Saddik [20] and Schroeder et al. [21] referred to digital twin as "Part of a Cyber-Physical System" which may be defined as a group of physical units which have virtual components as their digital version, that work together with a virtual reality via a communication medium.

While one may be tempted to think that a digital twin is only a simulation or model, Grieves [22], Kritzinger et al. [12] and Negri et al. [23] argue that it goes beyond that: A digital twin is an intelligent model which can evolve, and it follows the lifecycle of the physical asset. It allows predictions of future system failures and defects and facilitates simulation in order to test new configurations and facilitate predictive and preventive maintenance. The mirroring process is enabled by the harmonisation and continuous communication between the physical asset, its immediate environs, and its digital twin.

2.2 DT Characteristics

These authors in their definitions of digital twins had these characteristics in common: the physical object, the virtual model, and its interactions.

2.2.1 The Physical Object

This is the real-world artefact from which a digital replica is created. It may be an equipment, a component of an equipment, a process, or a living organism. Researchers often use specific terms such as 'product', 'component', 'system', etc. to refer to these physical objects. As they are real-world objects, they are not usually characterised by the word 'physical'. In order to generalise and encompass all forms of physical objects, literature presents the use of some terms, including 'Physical Asset' [24, 25], 'Physical Entity' [17, 26], and 'Physical System' [27, 28]. For the purpose of this study, all three terms will be used interchangeably. The physical entity has its own surrounding environment or real-world space within which it exists which includes all the parameters that may have an impact on the physical entity. This real-world space in which the physical entity exists has been named in literature as 'Physical Environment' [18, 29]. The physical environment includes the location, infrastructure, and technologies available, the time, and the status of the physical entity, among others.

2.2.2 The Virtual Entity

According to Bauer et al. [30], there are several kinds of virtual representations of physical assets: databases, 3D models, social media accounts, avatars, etc. However, the virtual entity is a controlled virtual representation of the physical asset that is precise on both a micro and macro level. Just as with the physical asset, the virtual entity is referred to by a number of terms for specificity, 'cyber', 'model', etc. For generalisation purposes [18] proposed the use of 'Virtual Entity'. Equally, the virtual entity has a surrounding environment which is a mirror the real-world environment. This has been popularly referred to as 'Virtual Environment' by researchers, including Grieves [22] and Toivonen et al. [31]. Lohtander et al. [32] described the virtual environment as parallel environment. This is because it precisely reflects the procedures and actions of the physical environment.

2.2.3 The Interaction Between the Physical Entity and the Virtual Entity

There is an endless connection between the physical entity and virtual entity in a digital twin. Barricelli et al. [17] explained that data is continually exchanged and revised as a result of real-time data uploads and big data analytics. Through this connection the state of the physical entity is conveyed to and mirrored by the virtual entity. Similarly, as explained by information flow and processes from the virtual entity is transmitted to and displayed by the physical entity. The physical entity and the virtual entity complement each other by facilitating data collection, storage and analysis from the entities and surrounding environment [33]. The data, processes and information that flow between the physical entity and virtual entity are known as parameters and it is a two-way flow that can influence both entities.

2.3 Digital Twin Technologies

There are new digital technologies springing up each day that support digital transformation. As part of this digital transformation, a technology that creates a virtual prototype of a physical entity has been introduced and is known as a digital twin technology [34]. Digital twins render unique visibility into physical systems and processes to be able to spot bottlenecks, be innovative and to restructure operations. As such, Aho [35], termed digital twin as a facilitating technology for smart lifecycle management. According to Qi et al. [36], in the bid to create a digital twin of a physical entity, various digitalisation technologies have been employed. These digitalisation technologies include Internet of Things, Artificial Intelligence, Machine Learning, big data analytics, and cloud computing. The digitalisation technologies facilitate the converging of the physical and virtual entities of a digital twin.

2.3.1 Internet of Things

Internet of Things (IoT) is considered as one of the key enablers of digital twin. According to Nord et al. [37], IoT has no standard definition, although it has been defined by several researchers over the years. Lee and Lee [38] defined it as a network of machines that can interact with each other. Correspondingly, Ornes [39] characterised IoT as a connection of devices that keep growing and is able to capture and distribute data. Daya et al. [40] gave a more detailed definition by including features of connectivity, its nature that facilitates storage and sharing of data, and the communication among devices. They defined IoT as a network of real-world objects that are connected digitally to sense, observe, and collaborate in order to enable information sharing. These definitions refer to devices that are connected together and interact. The two components that these connected devices have in common are sensors to collect data and a means to analyse and communicate the collected data in real-time. This real-time analysis is key to digital twins as the sensors attached to the connected devices collect data and feed it to the digital twin instantaneously [41]. Using this data, new concepts and logic are developed and tested on the digital twin.

2.3.2 Artificial Intelligence

Artificial Intelligence (AI) has been described by Amisha et al. [42] as the ability of computers and other technologies to mimic intelligent actions and critical thinking equivalent to a human. Despite the importance of intelligent systems, Teng and Gong [43] argue that, without a learning ability it cannot be truly referred to as an intelligent system. The method through which a system is able to acquire knowledge on its own is known as Machine Learning (ML). Artificial intelligence and machine learning have become leading problem-solving methods, as drastic improvements in the capability and use of advanced analytical tools have changed the extraction of useful insights from big data. According to Dilmegani [44], digital twins benefit from artificial intelligence and machine learning, since artificial intelligence and machine learning algorithms enable the development of some digital twins as well as the processing of big data gathered from digital twins. Machine learning frameworks facilitate the development of systems that can make independent decisions and make accurate predictions about future conditions using real-time data [45]. Through machine learning, the processes become more intelligent, leading to more accurate management and analysis of complex data for better predictions.

2.3.3 Big Data Analytics

The large amounts of data generated from several different sources at high speed has necessitated big data analytics [46]. Vassakis et al. [47] defined big data analytics as the collection, storage, assessment, and visualisation of large data sets to ascertain valuable insights to promote innovation, and transform businesses, as well as

economies. Arunachalam et al. [48] explained that big data analytics is multidisciplinary and is characterised by its capability of managing data with four qualities: Volume (the size of the data), Velocity (the speed at which the data is produced), Variety (the format in which the data comes), and Veracity (the consistency of the data). Other researchers also characterise big data by the 3Vs (Volume, Velocity and Variety) [46, 49] or the 5Vs (Volume, Velocity, Variety, Veracity and Value) [50, 51]. According to Frankenfield [52], most of the methods and processes involved in big data analytics are automated to evaluate data in order to identify critical patterns and metrics for better understanding. Big data analytics adopts different aspects of numerous scientific disciplines such as artificial intelligence, machine learning, statistics, system theory, and many others. As argued by IEEE Big Data [53], big data analytics acts as an enabler for digital twins such that it allows creators of the digital twins to swiftly identify new development opportunities, make a diagnosis and rectify complications before they get out of control.

2.3.4 Cloud Computing

Cloud computing is one of the fastest emerging technologies in computing. National Institute of Standards and Technology (NIST) defined cloud computing as a model for facilitating accessible, on-demand system access to shared configurable computing resources such as servers, storage, applications, and networks with little or no service management interference. As explained by Hosseinian-Far et al. [54], the emergence of cloud computing is to address the need for businesses to collect and store huge amounts of quality data from numerous sources. Creating and managing a digital twin requires intensive computing and storage of data. According to Kumar et al. [55], cloud computing model consists of cloud provider, one who provides cloud services, a cloud consumer, one who gets the cloud service from the cloud provider; a cloud carrier, one who provides connectivity between the cloud provider and the cloud consumer; a cloud broker, one who interacts between the cloud provider and the cloud consumer to facilitate the business transaction; and the cloud auditor, one who independently assesses the cloud services, information systems, and performance and security operations. They explained further that the formalised service models are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). Dohrmann et al. [45] expounded that, through the use of software-as-a-service (SaaS) solutions, which is a cloud-based software provision method for users, the cost of processing and storing the large amounts of data involved in digital twins continuous to decrease. It enables the developers of digital twins to acquire their needed computing resources as and when required at affordable costs. The computing power and resources necessary for real-time running simulations and forecasting of big data has been made widely and readily available by cloud computing.

2.4 DT Architecture

Generally, a digital twin architecture consists of a physical entity, its virtual entity, and a communication mechanism between the physical and virtual entities. As per Khan et al. [56], there is no single openly accepted architecture for digital twins. As such, several digital twin architectures have been proposed for different settings. The proposed digital twin architectures include 5C architecture, COGNITWIN, Intelligent digital twin, Six-Layer architecture, among others [57].

2.4.1 The 5-Layer Model of Cyber-Physical Systems

As the affordability and availability of computer networks, data acquisition systems and sensors are continually increasing, the use of high-tech methods has become a major force in several industries. These high-tech methods have led to the generation and use of large amounts of data, known as big data, accessed through these sensors and networked systems. According to Greer et al. [58] the technologies used to manage the networked systems between a physical entity and its computational abilities is known as a Cyber-Physical System (CPS). Lee et al. [59] proposed a 5-layer Cyber-Physical System known as the 5C architecture which gives a procedural guideline for creating and implementing a Cyber-Physical System for manufacturing applications. The five levels of the 5C architecture includes 'Connection', where effective sensors are selected, and precise and consistent data is acquired from manufacturing systems. Due to the availability of different types of data, a unified and tether-free technique for acquiring and transferring the data to the main server is useful. Also, it is crucial to select the most appropriate sensors for this level, 'Conversions', where important information is inferred from the raw data acquired. At this level, the architecture draws awareness to the health value and estimated useful life left for the machines involved through health management applications and prognostics; 'Cyber', which is the vital information centre such that it receives information from every connected device to form a network. Due to the massive information received on this level, the previous, current, and future performance of each device in the connection can be assessed to facilitate self-comparison among devices; 'Cognition', where infographics are used to transfer the vital information acquired to users. Here, decisions can be made on priority functions due to the availability of information on device statuses and comparative information; and 'Configuration', which is the response from the virtual space to the physical space and serves as an executive control to enable devices to self-configure and self-adapt. The corrective and preventive decisions undertaken in the Cognition Level is applied here (Fig. 1).

According to Lee et al. [59], employing a 5-Layer Cyber-Physical System in factories offer numerous benefits to production lines which consists of elements such as sensors, machinery, and production system. For the elements, the moment the sensory data from vital components has been transformed into useful information, a cyber-twin of the components will be liable for securing time machine

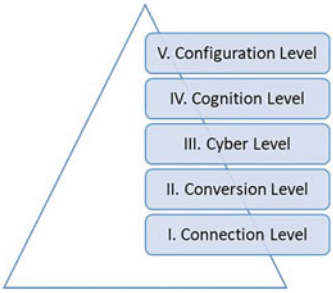


Fig. 1 An image depicting the 5 layer cyber-physical system/5C architecture adopted from [59] and illustrated by author

records and integrating future measures to provide self-awareness and self-forecast. Subsequently, more complex machine data would be collected to the elements information to observe the status and create the cyber-twin of each device. These cyber-twins provide the extra self-assessment ability. Then with the production system, accumulated information from elements and device level information provide self-customisability and self-operability to the factory. At this point, the level of knowledge available guarantees a near zero production downtime and facilitates production and inventory planning for optimisation. Ahmadi et al. [60] asserts that the 5-Layer Cyber-Physical System facilitates engineering of better devices by leveraging performance data, remote device management, and operation optimisation, among other.

2.4.2 Six-Layer Digital Twin Architecture

A Six-Layer Digital Twin Architecture was created by Redelinghuys et al. [61] to facilitate interaction between the physical and digital entities, as well as the digital entity and the external world, targeted at circumstances where the products of several suppliers are utilised in the physical entity and the rest of the digital twin. According to Redelinghuys et al. [62], the Six-Layer digital twin architecture was inspired by the 5C architecture of Lee et al. [59]. The first and second layers comprise of the physical entity. The first layer consists of physical devices such as sensors which exchange signals with the local controller. The local controller is located at the second layer which is the data source for the physical entity. The third layer provides a communication interface which is supplier neutral between the physical entity and the other layers of the architecture. Aside communicating with layer 4, it is able to directly log data to the cloud in layer 5 and at times layer six [62]. The fourth layer, also known as the IoT Gateway processes the data from the third layer to obtain useful insights for the upper layers. The fourth layer interfaces with the cloud data source, and the local data source. As such, adding a graphical user interface to this layer, where some of the major digital twin operations can be managed, is appropriate. The fifth layer is a cloud-based information repository for the physical and digital twin. As different stakeholders may have different information needs, numerous repositories are seen at this level. Holding the repositories in the cloud improves ease of access and

use, and connectedness to the digital twin [61]. The sixth layer serves as a dashboard which connects the user to real-time historic information about the physical entity. It is equipped with emulation and simulation software such as Siemens Tecnomatix Plant Simulation which allows a user to interface with this layer.

2.4.3 COGNITWIN

Abburu et al. [63] presented an abstract architecture of the Cognitive Twin Toolbox (COGNITWIN) focusing particularly on the process industry. Three stages of twins were established: a Digital Twin, which makes use of only isolated prototypes of the physical system; a Hybrid Twin which has the ability to interrelate with its prototypes; and a Cognitive Twin which uses protracted prototypes that include proficient knowledge for problem-solving and to handle unfamiliar circumstances. The toolbox suggests five layers: Model Management Layer, Data Ingestion and Preparation Layer, Service Management Layer, Twin Management Layer, and a User Interaction Layer.

The required model types are almost parallel to the distinct semantic models in the reference framework for digital twin proposed by Josifovska et al. [64] and contains first-order theory models based on the underlying physics, empirical models, etc. The Service Management Layer is liable for managing services, like registering and planning. Two types of services are distinguished. Data-driven and model-based driven services are the two types of services recognised. The Twin Management Layer controls the composition of the digital twin, especially, the management problem as a result of changes in the performance of the physical system. The toolbox also presents a User Interaction Layer where clients can delve into the COGNITWIN.

2.5 Digital Twin Applications

As a result of artificial intelligence, machine learning, Internet of Things, Big Data Analytics, and cloud computing working together, digital twins has become more detailed and predictive, thereby facilitating valuable applications. Digital twins have been applied in several sectors, ranging from healthcare to manufacturing. Some application cases emerging from literature are discussed as follows.

2.5.1 Manufacturing

According to Fuller et al. [65], the greatest reason for applying digital twins in the manufacturing sector is finding ways to track and monitor products while saving time and money. Other drivers for the application of digital twins in manufacturing include the need to gain competitive advantage, requirements for production flexibility on the market, the desire to follow a worldwide movement, the need to achieve

process transparency, and safety concerns, among others [66]. As manufacturing processes are increasingly becoming digital, it is opening up opportunities for smart manufacturing. Qi and Tao [67] explained that digital twins in manufacturing help to bridge the gap between the physical and virtual processes such that the use of IoT for the collection of real-time data in large volumes, based on cloud computing allows manufacturers to identify bottlenecks in their processes, trace to the root-cause and find the best possible solution. This ensures that manufacturing processes are efficient and more competitive. In accord, Xu et al. [68] also explained that digital twins provide a new concept for fault diagnosis such that issues in manufacturing processes which cannot be traced to its root-cause and assessed physical, can be evaluated on the virtual twin, with the appropriate what-if analysis to find the best solution.

2.5.2 Healthcare

Digital twins in healthcare is classified under digital health technology. Here, the physical entity may be living, may be in the form of wearable devices, software for diagnosis, medical devices, or drug development. Philips [69] explored digital twins of a human heart. They focused on how clinicians could confidently assess disease states of the heart, determine treatment, and guide therapies enabled by anatomical intelligence. Liu et al. [70] developed a healthcare system referred to as cloud healthcare which is based on digital twin healthcare (CloudDIGITAL TWINH). It offers important insights into a design of a setup that consists of the patient, the physician, the digital twin, and the technical implementation of the digital twin prototype. With the digital twin healthcare potential solutions can be assessed in virtual environments, such as drug experimentations, preparations and simulations for surgeries, staff scheduling, etc. Also, Orcajo [71] described wearables with sensors that feed real-time data to the cloud healthcare to enable patient monitoring and help develop model for the early detection of symptoms, diagnosis of diseases at its early stage and assess the effectiveness of treatment. Several drugs may also be tested on the digital twin patient to select the best drug for the situation, given the patient's medical records and conditions.

2.5.3 Smart Cities

According to Kosowatz [72], digital twins are being employed in the planning of cities to facilitate planning and prediction. The digital twin of a city is expected to reflect accurately and affect the laid down procedures used to operate and manage the city. As explained by Khajavi et al. [73], the virtual mirror of a city with all the constituents of the city represented on the virtual entity, provides an opportunity to improve operability and city planning. The major areas identified by Kumar et al. [74] as vital for developing smart cities are physical infrastructure, planning, information technology infrastructure, and smart solutions such as tourism services, tragedy management, etc. Smart cities are to improve residents' lives, promote security and environmental

efficiency, through centrally regulated and supervised technical infrastructure [75]. As asserted by Farsi et al. [76], digital twin technologies are important for city development and efficiency as it facilitates monitoring, the ability to clearly visualise, detecting and predicting concerns in real-time. Having the ability to view all nooks and crannies, and systems within a city virtually and applying what-if scenarios by leveraging on IoT technologies, artificial intelligence, and other technologies, to see what effects it will have on the city and its residents gives an opportunity to improve traffic flow, enhance energy efficiency, and improve security, among others. Some examples of smart cities given by Kosowatz [77] include Singapore, which has integrated smart technologies into housing via a framework that takes into account, buildings, living, environment and planning to be able to analyse solar diffusion, wind flow, best sites for new constructions, etc.; Dubai, which uses artificial intelligence to monitor bus drivers in order to reduce car accidents caused by exhaustion, and also having self-governing police stations where residents can make reports and pay for fines without dealing with a human being; and London, which desires to achieve a connected London by installing 5G cells 200 m apart, using drones to identify places where cellular antennas can be installed, and fitting lampposts with sensors and charging ports for electric vehicles.

2.5.4 Education

Digital twin is becoming a new tool for education. According to Hinduja et al. [78], it allows rapid teaching and learning of new concepts. Using digital twins in education increases flexibility such that equipment that are too expensive for schools to afford and processes that are too slow to study physically may be accessed by just the click of a computer mouse. As explained by Hinduja et al. [78], digital twins in education encourages creativity and facilitates the simulation of complex experiments. In support, Sepasgozar [79] also reasoned that digital twin in education improves creativity as there is the opportunity to have an active learning experience rather than a passive one. An example is students at Aarhus School of Marine and Technical Engineering using 3D replicas of automation systems to program automated production lines [16].

2.5.5 Transportation

As the populations increase, cities face the challenge of effective management of transport flows. There are usually control centres to ensure normal traffic flows on road networks and digitalising these centres is a key step to effective management. Rudskoy et al. [80] explained that digitalised control centres are achieved by implementing Intelligent Transport Systems. One of the major elements of the Intelligent Transport System is digital twins that make use of mathematical modelling methods to assess transport networks, identify issues and propose viable solutions. Access to the transport networks enables easy management of all aspects of the networks

and the information related to it. According to Zhaohui et al. [81], digital twin in transportation ensures proper management of infrastructure, virtual assessments, and experiments to curb transportation issues.

3 Materials and Methods

This was quantitative research such that it employed quantitative methods for the study, that is, a systematic review and a bibliometric analysis. A systematic review was performed to obtain appropriate studies in digital twins for the research. A systematic review provides a summary of existing studies upon which informed judgements and recommendations can be made. A bibliometric analysis was also conducted to investigate the level of publications, the most researched subject areas, the most influential authors, and publishers, as well as the geographical distribution as used by Okumus et al. [82].

3.1 Identification of Sources and Data Collection

In identifying data sources, three major research databases were considered: Web of Science, Scopus, and Google Scholar. While some studies including Bramer et al. [83] concluded that a single data source is not adequate for bibliometric analysis and systematic review, others such as Rice et al. [84] argued that a single database is sufficient as other databases have no impact on the results. A single database was selected based on the latter argument.

The data relevant to this study was retrieved from Web of Science Core Collection, an index of high-quality peer-reviewed publications currently managed by Clarivate Analytics. In order to identify all applicable publications, no lower limit was set, and the search was extended through to May 2021. The keyword string below was used to produce the preliminary database of publications in Web of Science.

> "Digital Twin*" OR "DT" AND "Digital Twin* + Architecture" AND "Digital Twin* + Application" AND "Digital Twin* + Reference Model"

This keyword search yielded 2826 results. The quick filters on Web of Science were then used to filter the results on broad categories such as document type, the search results were limited to only journals; language, only the articles written in English were selected, and documents published later than May 2021 were excluded. This reduced the results to 1490 documents. The titles and abstracts of the publications were then manually assessed. The final dataset for this study after excluding additional 552 publications included a total of 938 publications. This search procedure was informed by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines for performing systematic reviews (Fig. 2).

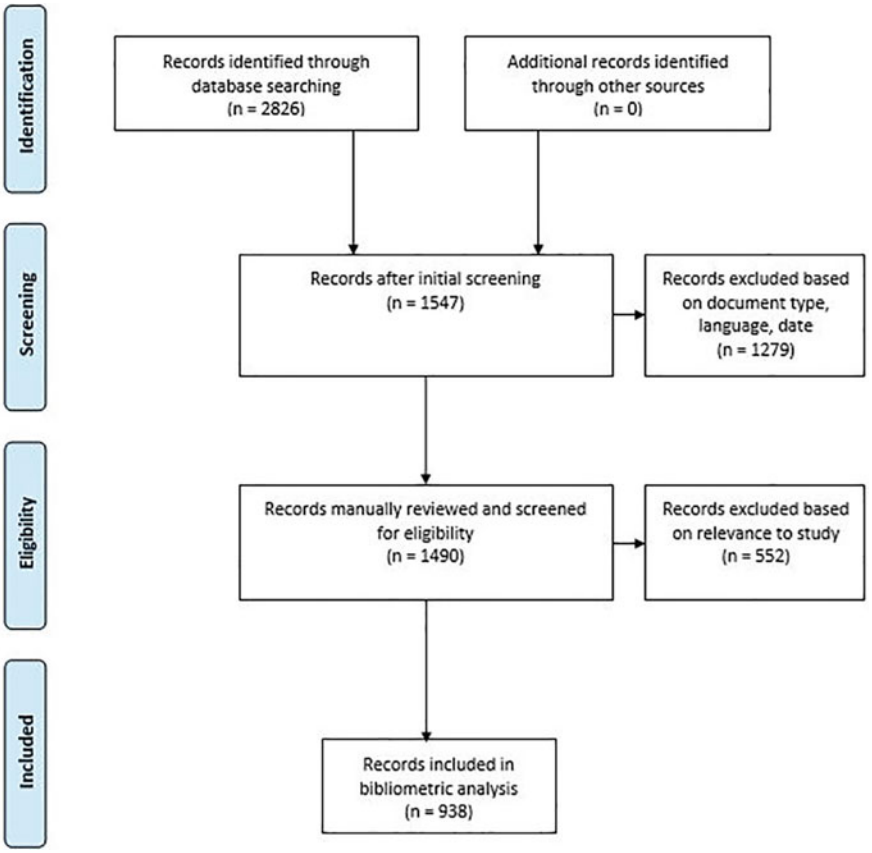


Fig. 2 PRISMA flow diagram describing the collection of digital twin technology, architecture, and application documents from Web of Science [85]

3.2 Data Analysis Methods

The data consisting of 938 documents was exported from Web of Science as both an Excel file and a Tab Delimited file. This is because the various analysis tools to be used could only work with one of these file types. The file consisted of meta data for each document, including name of authors, affiliated organisations and countries, title of article, source, abstract, publishers and publication locations, publication years, and other citation data. In order to use the VOSviewer software for the bibliometric analysis, a thesaurus file had to be created to filter the data, as used by Hallinger and Chatpinyakoop [86]. The thesaurus file had a 'label' column and a 'replace by' column, where same words expressed in a slightly different manner could be represented in singular form. For instance, some articles wrote 'industry 4.0' as 'industry 4.0' and without the thesaurus file to correct this during the keyword

analysis, VOSviewer was giving results for 'industry 4.0', 'industry 4' and '0'. Also, names such as 'Tao, Fei' represented as 'Tao, F' were being treated as different names. The quantitative analysis made use of descriptive methods, co-authorships, citation and co-citation analysis, and co-occurrence analysis. Descriptive methods such as the use of pie charts, bar charts, tree map charts and maps were used to present basic features such as publication growth trend, publication outlets and publication geographical distribution, and research areas. This was done with Excel and Tableau software programs.

VOSviewer, a software tool for structuring and visualising bibliometric systems such as journals and other publications was used to perform a bibliometric analysis. Jan van Eck and Waltman [87] explain that VOSviewer has the ability to interpret and display large bibliometric maps in a comprehensible manner. Using the VOSviewer software, co-authorship analyses have been performed, a keyword co-occurrence has also been conducted to identified top concepts that have appeared across several of the research in relation to digital twins, and a citation and co-citation for authors and articles has been undertaken to identify key authors and key articles in digital twin research. According to Zupic and Čater [88], the co-citation analysis complements the citation analysis as it captures a broader literature base by basing its analysis on the referenced list. As such, it is not uncommon for items to appear in a citation analysis list and not appear on a co-citation analysis list, especially if the citation analysis is based on just a single database.

4 Findings and Discussion

4.1 Findings

The results are presented in line with the deductive approach followed by Karakus et al. [89]. It begins from more general findings and flows down to more specific results. It begins by giving details of the data such as the publications per year, research areas, and publications per publisher. From there, the findings on co-authorship follows as per the countries, organisations, and authors. The results on co-authorships are then presented according to countries, authors, and documents, after which the co-occurrences of author keywords are shown. Citation and co-citation of authors and documents are also presented. This makes it easier to understand as it gives broad knowledge about the data and findings before moving on to more specific outcomes.

4.1.1 Publication Trend

Out of the 938 documents retrieved from Web of Science, 1 was published as far back as 2004, 1 in 2014, 20 in 2017, 41 in 2018, 137 in 2019, 380 in 2020 and 310

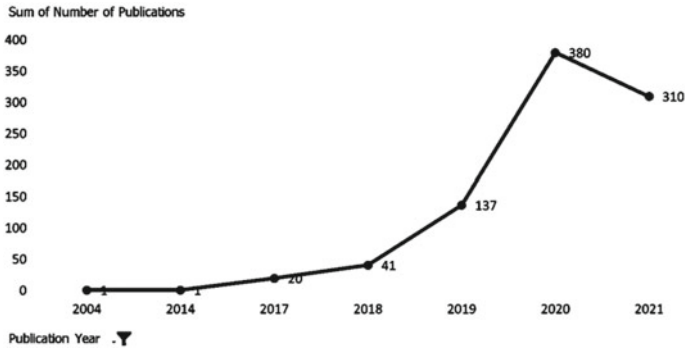


Fig. 3 Growth trajectory of publications on digital twin technology, architecture, and applications (n = 890)

from January 1 to May 31, 2021. 48 documents were early access journals released before May 31, 2021. Although the date for the search had no lower limits, there was no publication date older than 2004. As seen in Fig. 3, there has been an upwards increasing trend in the publications on digital twin technology, its architecture, and applications because it recently started gaining thrust and the benefits derived from this concept has aroused the interest of industrial and research populations over the years [90].

This growth trend indicates that about 88% of the literature in digital twins have been produced in the past three years (2019–2021). This exponential growth is due to the realisation by industries and researchers the fascinating and substantial opportunities digital twin technologies offer [27, 28]. These benefits of digital twins are projected to drive research over the years to come.

4.1.2 Publication Outlets

The 938 documents used for this study were published across 387 journals. The top 20 journals published 412 out of the 938, which makes 43.9% of the total number of journals. IEEE Access, which is a multidisciplinary peer-reviewed open access journal of the Institute of Electrical and Electronics Engineering (IEEE) was the most productive channel, publishing 63 journals in total. The next two were Applied Sciences (Basel, Switzerland) and Journal of Manufacturing Systems which published 46 and 44 documents respectively. The remaining journals published below 30 documents each out of the data used as shown in Fig. 4.

The publishers of these documents were also evaluated and ranked using the number of publications. A total of 125 publishers were identified in relation to the 938 documents assessed. Some publishers had just 1 publication and it would be cumbersome to include all 125 publishers. As such, the top 20 publishers with the highest number of publications were selected and presented in Table 1. From the table, the publisher with the highest number of publications is MDPI (Multidisciplinary

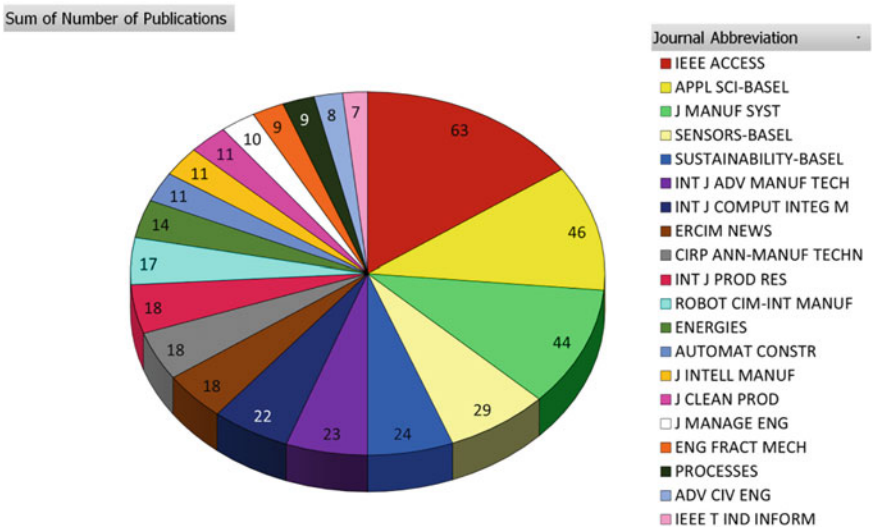


Fig. 4 Top 20 journals in the publication of articles on digital twin technology, architecture, and applications ranked by number of articles in Web of Science until May 31, 2021

Digital Publishing Institute), a publisher of open access scientific journals, with 172 publications. The next was IEEE (Institute of Electrical and Electronics Engineers) Publishing with 115 publications. Elsevier Science Ltd, Elsevier and Pergamon-Elsevier Science Ltd came next with 81, 79, and 60 publications respectively. Elsevier and Springer appeared a number of times in the top 20 but from different publication cities as shown in Table 1.

The countries of the publishers were also assessed and represented in the form of a map in Fig. 5 to show which countries where strongest in terms of publishing of articles on digital twin technology, its architecture, and applications. From the map, it is reflected that the United Kingdom had the most publications with 294 publications. United States of America (USA), Switzerland, and Netherlands followed with 221, 179, and 102 in that order.

4.1.3 Research Areas

The study was not limited to any specific research areas in order to ascertain meaningful representations in all fields in relation to digital twins as explained by Ross and Zaidi [91]. As such, 155 research areas were identified from the data extracted. The research area with the highest number of publications was Engineering, with 141 publications, which was almost twice as many publications as that of the next research area: Computer Science, Engineering and Telecommunications with 72 publications. The gap between the first two research areas was 69 publications, showing a clear advancement in research of digital twins in the field of engineering. Due to the large

Table 1 Top 20 publishers in the publication of research on digital twin technology, architecture, and applications ranked by number of publications until May 31, 2021

Publisher	Number of publications	Publication city	Country
MDPI	172	Basel	Switzerland
IEEE—Inst Electrical Electronics Engineers Inc.	115	Piscataway	USA
Elsevier Sci Ltd	81	Oxford	England
Elsevier	79	Amsterdam	Netherlands
Pergamon-Elsevier Science Ltd	60	Oxford	England
Taylor & Francis Ltd	57	Abingdon	England
Springer	25	New York	USA
Springer London Ltd	25	London	England
Springer Heidelberg	25	Heidelberg	Germany
Wiley	18	Hoboken	USA
European Research Consortium Informatics and Mathematics	18	Sophia Antipolis Cedex	France
Emerald Group Publishing Ltd	14	Bingley	England
Hindawi Ltd	13	London	England
ASCE—Amer Soc Civil Engineers	11	Reston	USA
ASME	10	New York	USA
Sage Publications Ltd	9	London	England
Wiley-Hindawi	9	London	England
Frontiers Media Sa	7	Lausanne	Switzerland
Walter De Gruyter GMBH	6	Berlin	Germany
Elsevier Science Inc.	6	New York	USA

number of research areas identified, only the top 20 were represented in the Tree Chart below (Fig. 6).

4.1.4 Co-authorship Analysis

According to [92], co-authorship analysis is the assessment of the collaboration relationships between two or more research publications in a specified. It may be in terms of authors of the publications, its affiliated organisations, and/or countries.

The co-authorship of countries consists of countries, which are represented by nodes and links which connect the nodes in the form of co-authorships. There is a

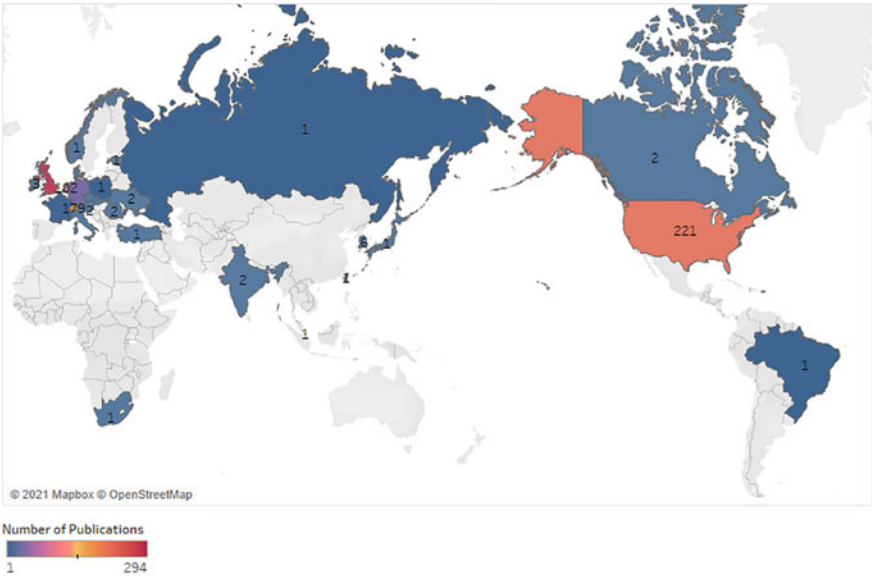


Fig. 5 Geographical distribution of countries of publishers of research on digital twin technology, architecture, and applications

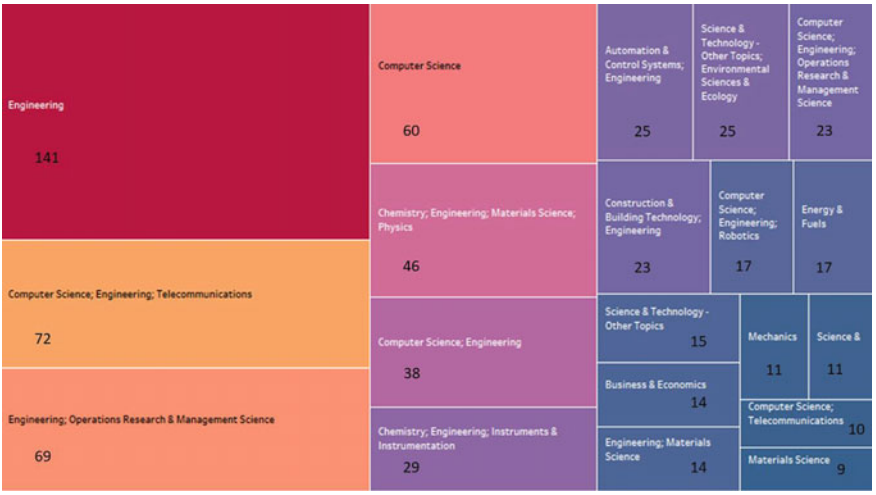


Fig. 6 A tree map chart presenting the top 20 research areas studied by the identified documents on digital twin research

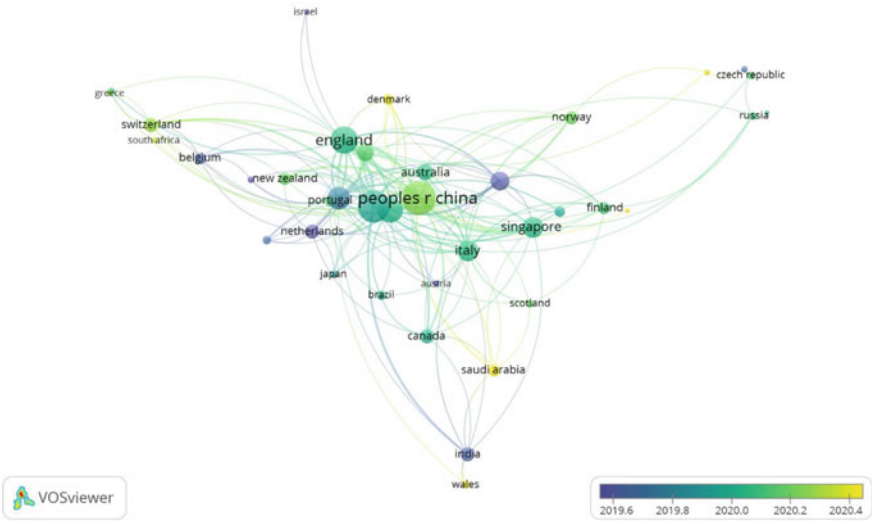


Fig. 7 Country co-authorship overlay visualisation map of the literature on digital twin technology, architecture, and applications (n = 70 countries in the co-authorship network; threshold of 5 documents per country; display of 39 countries)

link between two countries if they have co-authored at least one document and the size of the nodes here are proportional to the total link strength of the country. For the purpose of this study, the minimum number of documents of a country was set to 5, and out of 70 countries associated with the publications, 39 met the threshold. The co-authorship of countries is presented with an overlay visualisation in Fig. 7. Out of these 39 countries, the People's Republic of China had the most co-authorships with 228 documents, 3495 citations, and a total link strength of 100. United States of America (USA) came in second with 139 documents, 1595 citations, and 89 as its total link strength. The details of the next 3 countries are specified in (X, Y, Z) format, made to represent number of documents, number of citations, and total link strength. The countries that followed are England with (81, 634, 63), Germany with (106, 930, 57), and France with (47, 620, 44). These were ranked as the top 5 countries for co-authorship based on their total link strength. The colour of the nodes in Fig. 7 show how recent publications from the represented countries have been.

The co-authorship of organisations shows the relationship patterns between organisations related to co-authored documents. In this case, the organisations are represented by the nodes and the connections between the nodes are shown by the links. Setting the minimum number of documents of an organisation to 10, 18 organisations out of the 1159 organisations associated with the publications in the data met the threshold. Out of these 18 organisations, the largest set of connected nodes consisted of 11 organisations as shown in the network visualisation in Fig. 8. The total number of documents, citations and total link strengths for each organisation was calculated. The organisation with the highest total strength was University of Hong Kong with

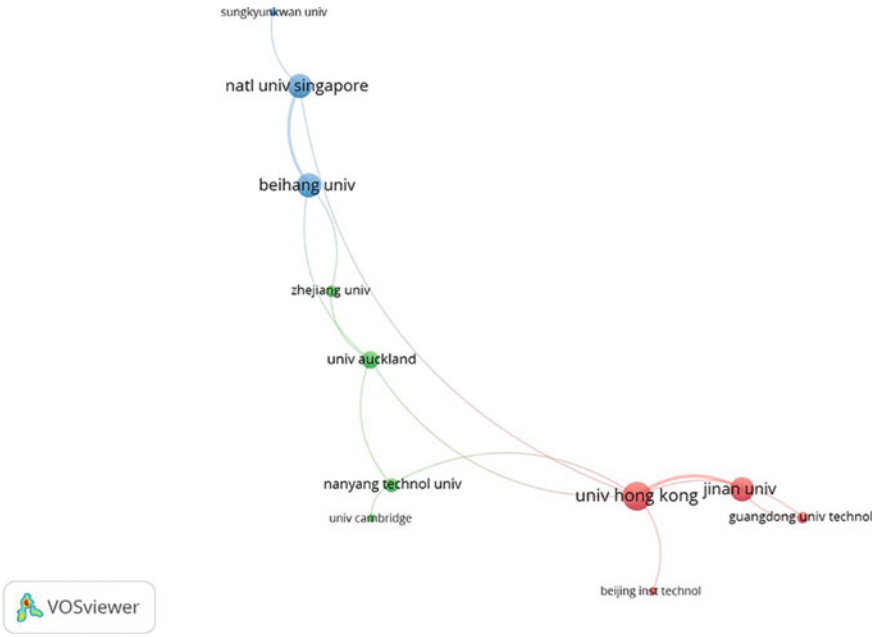


Fig. 8 Organisation co-authorship network visualisation map of the literature on digital twin technology, architecture, and applications (n = 1159 organisations in the co-authorship network; threshold of 10 documents per organisation; display of 11 organisations

15 documents, 92 citations, and 12 for total link strength. This was followed by Beihang University with 23 documents, 1364 citations and 8 for total link strength. The other organisations are presented in the format (X, Y, Z), representing number of documents, citations, and total link strength respectively. Jinan University followed with (12, 27, 8), National University of Singapore with (14, 738, 8), University of Auckland with (11, 324, 4), Nanyang Technological University with (12, 217, 3), Guangdong University of Technology with (11, 530, 2), Zhejiang University with (10, 46, 2), Beijing Institute of Technology with (10, 167, 1), Sungkyunkwan University with (11, 107, 1), and then University of Cambridge with (17, 230, 1). The organisations stated were the top 11 organisations in terms of total link strength which constituted the largest set of connected nodes.

Author co-authorships of the publications are presented with a density visualisation in Fig. 9. For inclusion purposes, the minimum number of documents of an author was set to 5. Of the 3410 authors identified in the dataset, 25 met the threshold. In Fig. 9, the density visualisation was weighted by the total link strength and the portions turning yellow signify a larger total link strength. Using the total link strengths, the top 10 authors with strongest co-authorships are presented below. The author with the strongest co-authorship was Qiang Liu with 8 documents, 475 citations and total link strength of 29. Xin Chen and Jiewu Leng followed, each with 7 documents, 461 citations, and a total link strength of 29. As shown in Fig. 9, these 3

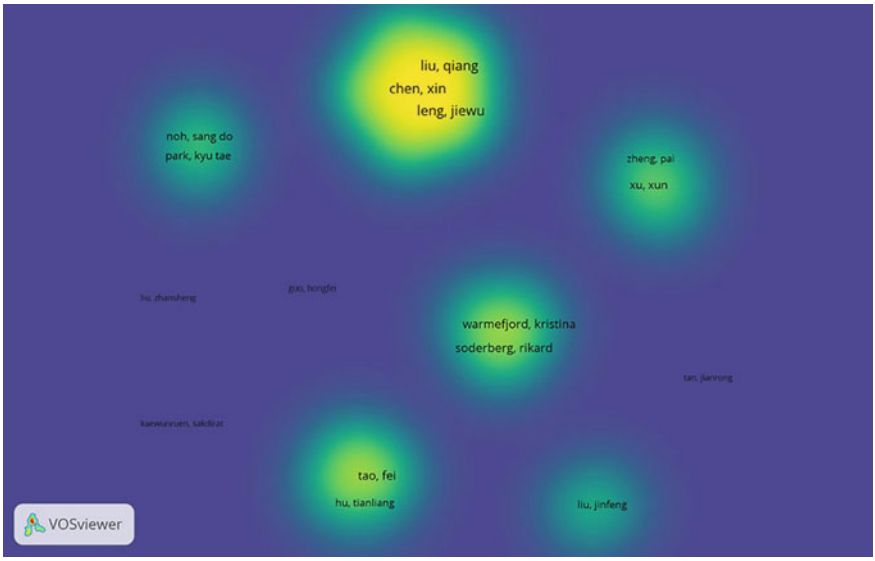


Fig. 9 Author co-authorship density visualisation of the literature on digital twin technology, architecture, and applications (n = 3410 authors in the co-authorship network; threshold of 5 documents per author; display of 16 authors)

authors have the brightest and biggest portion of yellow in the density visualisation. The specifics of the remaining of the top authors in co-authorships are presented in the format (X, Y, Z). Ding Zhang followed with (5, 331, 23), then Douxi Yan with (5, 240, 22), Fei Tao with (11, 1266, 12), Rikard Soderberg with (8, 197, 12), Kristina Warmefjord with (7, 187, 11), A. Y. C. Nee with (6, 653, 10), and Lars Lindkvist with (5, 181, 9).

4.1.5 Co-occurrence of Author Keywords

Keywords are important words or phrases of an article, that usually represent its main content [93]. Co-occurrence of these keywords exhibit their interconnectedness based on their combined presence in articles. The co-occurrence of author keywords for this study are shown in Fig. 10 with a network visualisation. For inclusion purposes, the minimum number of occurrences of keywords was set to 10. Of the 3110 keywords identified, 25 met the threshold. The total number of occurrences and total link strength for the keywords were calculated and the most frequent keyword was Digital Twin, with 542 occurrences and a total link strength of 429. The next keyword was Industry 4.0 with 89 occurrences and total link strength of 137, and then Internet of Things with 60 occurrences and a total link strength of 111. The occurrences and total link strength of the remaining keywords are presented in the format (X, Y) in that order. The next keyword was Cyber-Physical Systems with (49, 77), then Smart

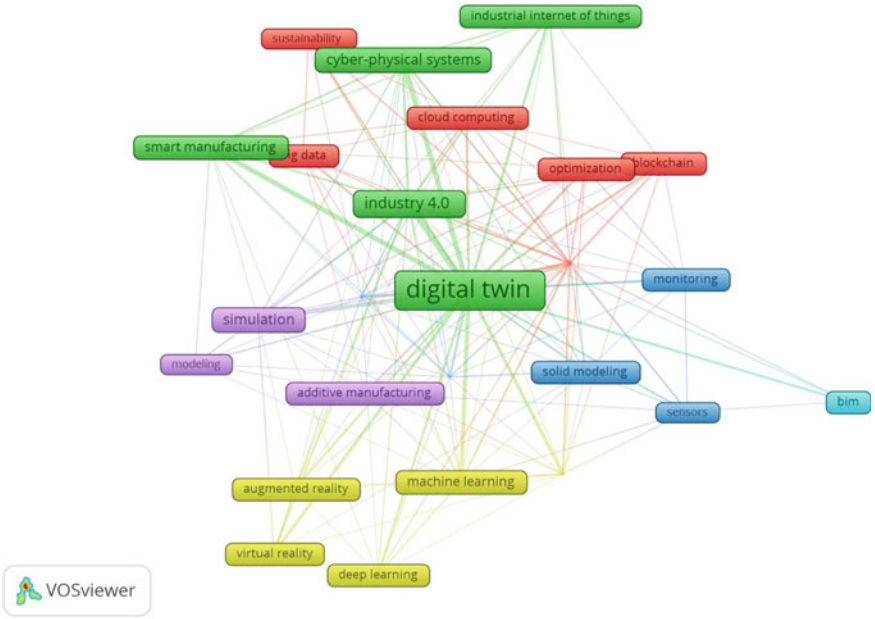


Fig. 10 Co-occurrence of author keywords network visualisation map for literature in digital twin technology, architecture, and application until May 31, 2021 (n = 3110 keywords; threshold of 10 occurrences; display of 21 keywords)

Manufacturing with (40, 63), Simulation with (46, 58), Machine Learning with (43, 58), Artificial Intelligence with (27, 45), Manufacturing (18, 37), and then Virtual Reality with (22, 32). These are the top 10 keywords based on the occurrences and total link strength.

Considering Fig. 10, the colours of the nodes represented as frames signify different clusters in which the keywords are regularly linked to each other. The cluster with the green colour consists of Digital twins, Industry 4.0, Cyber-Physical Systems, Industrial Internet of Things, and Smart Manufacturing, four of which are among the top 5 keywords.

4.1.6 Citation and Co-citation Analyses

Citation and co-citation analysis were employed in order to identify the most influential authors publishing on digital twin technology, its architecture, and applications. The citation analysis, using the Web of Science citation of authors, highlights the most prominent authors such as Fei Tao with 1266 citations, A. Y. C. Nee with 653 citations, Qinglin Qi with 617 citations, and the others, making up the top 20 as shown in Table 2.

Table 2 Top 20 most cited authors in research on digital twin technology, architecture, and applications until May 31, 2021

Rank	Author	Number of documents	Web of science citations	Citations per document
1	Fei Tao	11	1266	115
2	A. Y. C. Nee	6	653	109
3	Qinglin Qi	5	617	123
4	Meng Zhang	4	524	131
5	Qiang Liu	8	475	59
6	Xin Chen	7	461	66
7	Jiewu Leng	7	461	66
8	Ang Liu	4	414	104
9	Hao Zhang	5	394	79
10	Ding Zhang	5	331	66
11	Nabil Anwer	4	325	81
12	Xun Xu	8	308	39
13	Benjamin Schleich	4	282	71
14	Sandro Wartzack	2	276	138
15	Luc Mathieu	1	276	276
16	Morteza Ghobakhloo	1	244	244
17	Douxi Yan	5	240	48
18	He Zhan	1	222	222
19	T. DebRoy	3	220	73
20	Abdulmotaleb El Saddik	4	210	53

The outcome of the author co-citation analysis is presented in Table 3. From the table, the top 20 most influential authors identified in terms of co-citations include 5 of the most influential authors identified using direct citations: Fei Tao, Qinglin Qi, Benjamin Schleich, Jiewu Leng, and Hao Zhang. Despite this not being an unusual occurrence [94], it should be noted that Web of Science data includes only the first author of a cited document; other authors are not considered in a co-citation analysis of cited authors, and this could be the explanation for the considerable difference.

A citation analysis was performed to complement the identification of the job journals in digital twin research. Based on the number of citations, the top 20 journals were selected. Out of the top 20 most cited journals in Table 4, 17 of the journals are among the top 20 journals ranked according to number of publications in Fig. 4. IEEE Access came out as the most prominent journal in both cases.

Similarly, a citation analysis was conducted to identify the most influential articles in digital twin literature. For inclusion, the minimum number of citations per document was set to 80 citations. Out of the 938 documents, 20 met the threshold.

Table 3 Top 20 most co-cited authors in research on digital twin technology, architecture, and applications until May 31, 2021

Rank	Author	Co-citations	Total link strength
1	[a]Fei Tao	821	4768
2	Michael Grieves	281	1841
3	[a]Qinglin Qi	161	1280
4	Jay Lee	161	1033
5	[a]Benjamin Schleich	138	969
6	[a]Jiewu Leng	137	1050
7	Elisa Negri	123	912
8	Thomas H.-J. Uhlemann	117	939
9	Stefan Boschert	117	863
10	Roland Rosen	110	880
11	Edward Glaessgen	106	803
12	Yuqian Lu	102	742
13	Rikard Soderberg	99	725
14	Kazi Masudul Alam	80	635
15	Eric J. Tuegel	79	659
16	Werner Kritzinger	78	571
17	[a]Hao Zhang	76	707
18	Yingfeng Zhang	67	578
19	Michael Schluse	65	572
20	Cunbo Zhuang	64	585

[a] Indicates that the author also appeared in Table 2

Majority of the top 20 most cited articles in Table 5 were authored by the most influential authors presented in Tables 2 and 3 such as Fei Tao, Qinglin Qi, Benjamin Schleich, and Michael Schluse among others. The Web of Science citation counts of the topmost cited articles on digital twins are in moderation, as compared to citation counts of other digital technologies.

The articles were grouped into topical clusters and differentiated with colours as shown in Fig. 11. Cluster 1 is made up of 6 red coloured frames. It is made up of Alam and Saddik [20], Soderberg [98], Zhang et al. [104], Liu et al. [70], Leng et al. [29], and Ding et al. [107]. These articles concentrated on digital twin architecture and digital twin design methodologies. Cluster 2, consisting of 5 green coloured frames focused on digital twins in manufacturing to achieve smart manufacturing. This cluster consisted of Qi and Tao [67], Schleich et al. [19], Ghobakhloo [95], Tao et al. [100], and Tao et al. [26], three of which are the top 3 most cited articles. Cluster 3 consists of 4 blue frames: Tao et al. [97, 99], Knapp et al. [106], and Schluse et al. [108]. These articles focused on emerging and fast-growing technologies in digital

Table 4 Top 20 most cited journals in the publication of research on digital twin technology, architecture, and applications until May 31, 2021

Rank	Journal	Total number of publications	Total number of citations
1	[a]IEEE Access	63	1489
2	[a]CIRP Annals—Manufacturing Technology	18	699
3	[a]International Journal of Production Research	18	566
4	[a]Journal of Manufacturing Systems	44	539
5	[a]Robotics and Computer-Integrated Manufacturing	17	353
6	[a]IEEE Transactions on Industrial Informatics	7	351
7	[a]Journal of Ambient Intelligence and Humanized Computing	7	280
8	[a]The International Journal of Advanced Manufacturing	23	245
9	[a]International Journal of Computer Integrated Manufacturing	22	245
10	[a]Journal of Cleaner Production	11	219
11	[a]Sustainability-Basel	24	169
12	[a]Applied Sciences-Basel	46	166
13	Computers in Industry	5	158
14	[a]Journal of Management in Engineering	10	118
15	[a]Journal of Intelligent Manufacturing	11	110
16	[a]Automation in Construction	11	107
17	[a]Sensors-Basel	29	87
18	[a]Engineering Fracture Mechanics	9	52
19	IEEE Transactions on Power Electronics	5	46
20	[a]Energies	14	39

[a] Indicates that journal appeared in Fig. 4

Table 5 Top 20 most cited articles on digital twin technology, architecture, and applications until May 31, 2021

Rank	Article	Article title	Times cited, WoS core
1	Qi and Tao [67]	Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison	288
2	Schleich et al. [19]	Shaping the digital twin for design and production engineering	276
3	Ghobakhloo [95]	The future of manufacturing industry: a strategic roadmap toward Industry 4.0	244
4	Tao and Zhang [96]	Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing	244
5	Tao et al. [97]	Digital twin in industry: state-of-the-art	222
6	Alam and Saddik [20]	C2PS: a digital twin architecture reference model for the cloud-based cyber-physical systems	197
7	Soderberg [98]	Toward a digital twin for real-time geometry assurance in individualized production	150
8	Tao et al. [99]	Digital twin-driven product design framework	138
9	Tao et al. [100]	Digital twin driven prognostics and health management for complex equipment	128
10	Zhuang et al. [101]	Digital twin-based smart production management and control framework for the complex product assembly shop-floor	127
11	Bolton et al. [102]	Customer experience challenges: bringing together digital, physical, and social realms	116
12	Zheng et al. [103]	A systematic design approach for service innovation of smart product-service systems	115

(continued)

Table 5 (continued)

Rank	Article	Article title	Times cited, WoS core
13	Zhang et al. [104]	A digital twin-based approach for designing and multi-objective optimization of hollow glass production line	112
14	Liu et al. [70]	Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system	109
15	Ivanov and Dolgui [105]	A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0	107
16	Tao et al. [26]	Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0: correlation and comparison	103
17	Leng et al. [29]	Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop	103
18	Knapp et al. [106]	Building blocks for a digital twin of additive manufacturing	103
19	Ding et al. [107]	Defining a digital twin-based cyber-physical production system for autonomous manufacturing in smart shop floors	101
20	Schluse et al. [108]	Experimentable digital twins-streamlining simulation-based systems engineering for industry 4.0	87

twins. Cluster 4, making up of Tao and Zhang [96] and Zhuang et al. [101] in yellow frames focused on digital twin shop floors.

As the document citation analysis conducted was from only one database, Web of Science, a document co-citation analysis was conducted to obtain a broader perspective of the documents that have been influential in the development of digital twin literature. It was noteworthy that the topmost co-cited document in Table 6 did not appear in the most cited publications in Table 5. Further perusal found that the article was filtered out as a Proceedings Paper during the initial screening of the dataset. There is a reasonable level of overlapping between the top 20 documents with the most citations in Table 5 and the top 20 most co-cited documents in Table 6. There were 9 documents that appeared on both lists: Schleich et al. [19], Tao and Zhang

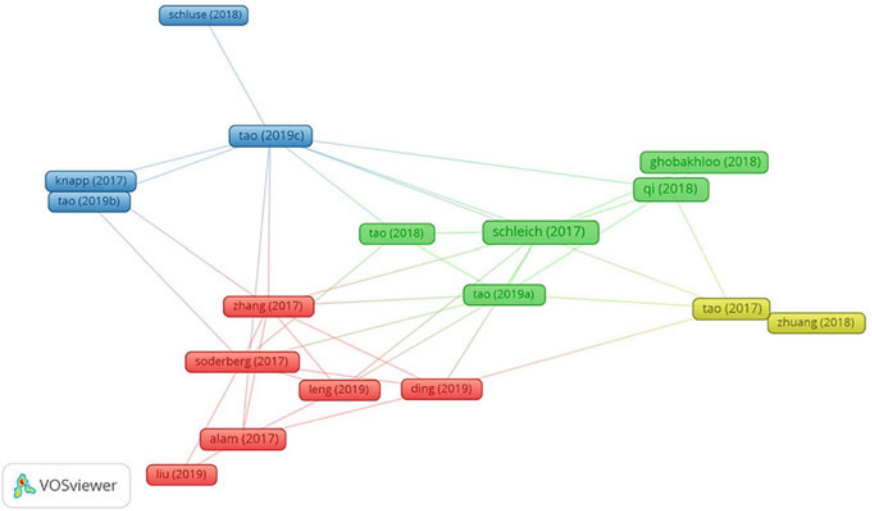


Fig. 11 Network visualisation map of the top 20 most cited articles on digital twin technology, architecture, and applications until May 31, 2021

[96], Tao et al. [99], Qi and Tao [67], Soderberg [98], Alam and Saddik [20], Tao et al. [97], Tao et al. [100], Zhang et al. [104]. Other documents appearing on the most co-cited documents list is a demonstration of the ability of co-citation analysis to identify influential documents without restriction of a particular database used.

It should be noted that the articles are represented by the surname of the first authors in this study only because that it how it was presented by the VOSviewer visualisation.

4.2 Discussion

In this study, the publications on digital twin technology, digital twin architecture and digital twin architecture were retrieved from Web of Science database. This dataset was filtered, analysed, and visualised using descriptive methods and bibliometric methods. Excel, Tableau and the VOSviewer software were used to evaluate and visualise the data. In this circumstance, the publication trend, publication outlets, research areas, co-authorship of countries, organisations, and authors, co-occurrence of author keywords, and citation and co-citation of authors were analysed and presented.

For the publication trend, there is an upward trend in the number of publications on digital twin technology, its architecture and application over the years. The number of publications gained momentum at about 2017 when researchers and industries began to get more curious about the possibilities of digital twins. According to Datta

Table 6 Top 20 most co-cited articles on digital twin technology, architecture, and applications until May 31, 2021

Rank	Article	Article title	Citations
1	Tao et al. [100]	Digital twin-driven product design, manufacturing, and service with big data	176
2	[a]Schleich et al. [19]	Shaping the digital twin for design and production engineering	118
3	Grieves and Vickers [2]	Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems	107
4	Rosen et al. [15]	About the importance of autonomy and digital twins for the future of manufacturing	106
5	[a]Tao and Zhang [96]	Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing	102
6	[a]Tao et al. [97]	Digital twin in industry: state-of-the-art	100
7	[a]Qi and Tao [67]	Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison	99
8	Glaessgen and Stargel [109]	The digital twin paradigm for future NASA and US air force vehicles	88
9	Negri et al. [23]	A review of the roles of digital twin in cps-based production systems	88
10	[a]Soderberg [98]	Toward a digital twin for real-time geometry assurance in individualized production	79
11	Kritzinger et al. [12]	Digital twin in manufacturing: a categorical literature review and classification	77
12	[a]Alam and Vickers [20]	C2PS: a digital twin architecture reference model for the cloud-based cyber-physical systems	74
13	Boschert and Rosen [110]	Digital twin—the simulation aspect	74
14	Uhlemann et al. [111]	The digital twin: realizing the cyber-physical production system for industry 4.0	74
15	[a]Tao et al. [99]	Digital twin-driven product design framework	69
16	[a]Tao et al. [100]	Digital twin driven prognostics and health management for complex equipment	66
17	Lee and Lee [38]	A cyber-physical systems architecture for industry 4.0-based manufacturing systems	65
18	Grieves [22]	Digital twin: manufacturing excellence through virtual factory replication	63
19	Lu et al. [112]	Digital twin-driven smart manufacturing: connotation, reference model, applications, and research issues	62

(continued)

Table 6 (continued)

Rank	Article	Article title	Citations
20	[a]Zhang et al. [104]	A digital twin-based approach for designing and multi-objective optimization of hollow glass production line	62

[a] Indicate that the article appeared in Table 5

[113], digital twins is gaining thrust because its possibilities are endless, and it may offer real-time precision. Considering that January to May 31, 2021, have as many publications as the whole of 2020 is an indication that the research and applications of digital twins is growing and will continue to grow speedily.

Considering publication outlets, the journal analysis performed concluded that the research on digital twins, its architecture and applications are being published in good quality multi-disciplinary journals. These journals specialise in engineering, manufacturing, medicine and health, science and technology, robotics, computing, environment, culture, economics, and social sustainability among others. The first journal was IEEE Access. Applied Sciences came in second, then Journal of Manufacturing Systems, then Sensors, and Sustainability, making up the top 5 journals. Using the direct citation analysis in Table 4 it was found that the top 20 most influential journals publishing articles on digital twins, its architecture, and applications in terms of citations and the top 20 in terms of number of publications Fig. 4 are in correspondence and are mutually reinforcing. The 3 journals which were not part of the top 20 most influential journals in terms of citations were part of the top 25. Also, the publishers of the journals identified in the studies were ranked in terms of the number of documents published. It was found that the major publishers such as MDPI, IEEE, Elsevier, Taylor & Francis Ltd, Springer, Wiley, etc. are highly ranked publishers. From Fig. 5 showing the map of publisher countries, majority of the publications were done in the United Kingdom, USA, Switzerland, and Netherlands. This however does not imply that these countries are the leading countries in relation to research on digital twins.

Several research areas in the form of clusters were identified in the analysis. As no specific research areas were given, the ones provided by Web of Science is what was used. Engineering was the most researched field as it had the most documents. Hartmann and Auweraer [114] explained that due to the incremental nature of the complexity of engineering design methods, research and development efforts are being made regularly to find easier and more efficient ways, and the research and employment of digital twins is a forward leap. Several of the other research areas identified had Engineering as part of the cluster. The other top research areas sciences such as Computer Science, Management Science, Chemistry, Material Science, among others. As concluded by Ante [9], digital twin is being considered in several scientific subjects, and this goes to support the claim. Despite Business and Economics making it to the list of top research areas, it was only social science

among the top 50 research areas, and it had few publications as compared to the other sciences.

Co-authorship analysis of countries, organisations and authors was performed. Country co-authorships indicated that China had the most co-authorships with a total link strength of 100, 228 documents out of the 938, and 3495 citations. Over the years, China has become a drive to acknowledge in digital technologies. According to Wang et al. [115], it is among the first 3 countries for venture capital investment in digital technologies such as 3D printing, artificial intelligence, and virtual reality. It is therefore not startling that majority of the research on digital twins, its architecture and applications is from China, and the country is the strongest at making collaborations with other countries in this field. United States of America (USA), England, Germany, and France, all which are technologically advanced countries, followed in that order. Considering the rankings by Wood [116], all top 5 countries for co-authorships in digital twin research are among the top 20 most innovative and research-enthusiastic countries. Aside identifying the top organisations and authors in co-authorships of digital twin research, the organisation and author co-authorship analysis confirmed the results of the country co-authorship analysis. Nine out of the eleven organisations shown in Fig. 8 are in China. The topmost organisation, University of Hong Kong, is under the Special Administrative Region of China. Beihang University, Jinan University, and National University of Singapore are all institutions in China. The top 5 authors in co-authorship are also from China, even though they may not be currently living in China: Qiang Liu, Xin Chen, Jiewu Leng, Ding Zhang, and Douxi Yan. The findings suggest that research on digital twins are poorly dispersed among countries worldwide. As these countries, organisations and authors are the most influential in co-authorships for this research, it will be prudent for upcoming or not so prominent countries, organisations, and others in terms of research on digital twins to seek for research collaborations with them for a wider audience and a higher impact.

The outcome of the co-occurrence of author keywords signify that researchers have mainly studied these concepts related to the main concept of digital twins: industry 4.0, cyber-physical systems, smart manufacturing, internet of things, industrial internet of things, cloud computing, big data, simulation, augmented reality, virtual reality, machine learning, deep learning, additive manufacturing, sensors, monitoring, optimisation, sustainability, BIM (Building information modelling), and blockchain. Digital twins, being at the core of industry 4.0, incorporates digital technologies like internet of things, machine learning, cloud computing, big data, etc. to create cyber-physical systems, virtual realities, and augmented realities via simulations and modelling. Digital twins facilitate with real-time monitoring for optimisation and sustainability. Also, application of digital twins in manufacturing, known as smart manufacturing, is noticeably high as put forth by [65].

An author citation analysis was undertaken to rank the most influential authors in the research of digital twins. The top 20 most influential authors are presented in Table 2 with Fei Tao being ranked number 1. Among the top authors presented in terms of citations, majority, including Fei Tao, Qiang Liu, Xin Chin, Xu Xun, Douxi Yan, and A. Y. C. Nee were part of the top authors in co-authorships in Fig. 9.

An author co-citation analysis was undertaken to complement the author citation analysis, and although the results were not greatly overlapping, Fei Tao remained the most influential author, while Jiewu Leng, Qinglin Qi, Hao Zhang, and Benjamin Schleich made it as part of the top 20 authors for co-citation (Refer to Table 3). The results also highlighted a major gender bias in the research of digital twins. This goes to support the assertion by García-González et al. [117] that there is a gender inequality in research. The results of the citation analysis for journals was in conformity with the results of the journal ranking using number of publications. Although there were changes in the rankings, 17 of the journals from the list of top 20 journals using number of publications in Fig. 4 appeared in the list of top 20 journals using direct citations in Table 4.

From the document citation and co-citation analysis, the most influential articles in digital twin research were identified in Tables 5 and 6. The articles are grouped into clusters that focused on smart manufacturing, digital twin architecture and design methods, emerging technologies in digital twins such as product design and additive manufacturing, and digital twin shop floors. These articles were authored by majority of the top authors identified in Tables 2 and 3 and were also published in the most influential journals.

5 Conclusion and Recommendation

5.1 Conclusion

In the past few years, digital twin research has increased significantly. The concept has impacted several sectors such as manufacturing, production, computing, engineering, etc. According to the analysis, it can be concluded that, digital twin research will continue to increase in the years to come, especially in the engineering, computing, and other applied sciences. However, there is more room for further research in the business field, and other social sciences. Also, the outcome of this research is a wakeup call for countries and organisations trailing in digital twin research, and researchers, particularly females to invest more in digital twin research, since studies have shown that digital twin is a worthwhile digital technology with great prospects for effectiveness, efficiency, and optimisation.

5.2 Limitations

The limitations of this study include the use of only one index for the identification of publications for the bibliometric analysis. Although Web of Science is a major index for bibliometric analysis, there is the possibility of exclusion of key documents in the research of digital twins as these studies may not be available in the index.

Also, since this is a quantitative study, it is not possible to identify the full impact of publications as a qualitative study would have. For instance, an author may not be part of the top authors in terms of number of documents but may be very influential in his field of study.

5.3 Future Works

It is recommended that in future, several indices be combined to replicate this study for a broader spectrum of publications. It is also recommended that a qualitative study of the research in digital twins be performed, as a qualitative study of the research publications in digital twins may provide further information on the outcome of this study. A bibliometric analysis for digital twin research can also be performed in the different research areas in order to identify the key authors, publishers, documents, and journals in these fields of study.

References

1. Gelernter DH (1991) Mirror worlds: or the day software puts the universe in a shoebox … how it will happen and what it will mean. Oxford University Press, New York
2. Grieves M, Vickers J (2017) Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In: Kahlen FJ, Flumerfelt S, Alves A (eds) Transdisciplinary perspectives on complex systems, pp 85–113. Spinger, Cham
3. Rasheed A, San O, Kvamsdal T (2020) Digital twin: values, challenges, and enablers from a modelling perspective. IEEE Access 8:21980–22012
4. Maria G (2020) Software advice. [Online] Available at: https://www.softwareadvice.com/resources/what-is-digital-twin-technology/. Accessed 25 May 2021
5. Crawford M (2021) The American society of mechanical engineers (ASME). [Online] Available at: https://www.asme.org/topics-resources/content/7-digital-twin-applications-for-manufacturing. Accessed 11 Sept 2021
6. Daneshkhah A, Hosseinian-Far A, Chatrabgoun O (2017) Sustainable maintenance strategy under uncertainty in the lifetime distribution of deteriorating assets. Strategic engineering for cloud computing and big data analytics. Springer, Cham, pp 29–50
7. Talkhestani BA, Jung T, Lindemann B, Sahlab N, Jazdi N, Schloegl W, Weyrich M (2019) An architecture of an intelligent digital twin in a cyber-physical production system. at-Automatisierungstechnik 67(9):762–782
8. Aheleroff S, Xu X, Zhong RY, Lu Y (2021) Digital twin as a service (DTaaS) in industry 4.0: an architecture reference model. Adv Eng Inform 47:101225
9. Ante L (2021) Digital twin technology for smart manufacturing and industry 4.0: a bibliometric analysis of the intellectual structure of the research discourse. Manuf Lett 96–102
10. Radanliev P et al (2021) Digital twins: artificial intelligence and the IoT cyber-physical systems in industry 4.0. Int J Intell Robot Appl
11. Ciano MP, Pozzi R, Rossi T, Strozzi F (2020) Digital twin-enabled smart industrial systems: a bibliometric review. International J Comput Integr Manuf 1–19
12. Kritzinger W et al (2018) Digital twin in manufacturing: a categorical literature review and classification. IFAC-PapersOnLine 51(11):1016–1022

13. Majumdar PK, Haider MF, Reifsnider K (2013) Multi-physics response of structural composites and framework for modeling using material geometry. In: 54th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, p 1577
14. Wright L, Davidson S (2020) How to tell the difference between a model and a digital twin. Adv Model Simul Eng Sci 7(13)
15. Rosen R, von Wichert G, Lo G, Bettenhausen KD (2015) About the importance of autonomy and digital twins for the future of manufacturing. IFAC-PapersOnLine 567–572
16. Madni AM (2019) Exploiting digital twin technology to teach engineering fundamentals and afford real-world learning opportunities. American Society for Engineering Education, Tampa
17. Barricelli BR, Casiraghi E, Fogli D (2019) A survey on digital twin: definitions, characteristics, applications, and design implications. IEEE Access 7:167653–167671
18. Jones D et al (2020) Characterising the digital twin: a systematic literature review. CIRP J Manuf Sci Technol 29(Part A):36–52
19. Schleich B, Anwer N, Mathieu L, Wartzack S (2017) Shaping the digital twin for design and production engineering. CIRP Ann 66(1):141–144
20. Alam KM, Saddik AE (2017) C2PS: a digital twin architecture reference model for the cloud-based cyber-physical systems. IEEE Access 5:2050–2062
21. Schroeder GN, Steinmetz C, Pereira CE, Espindola DB (2016) Digital twin data modeling with automationML and a communication methodology for data exchange. IFAC-PapersOnLine 49(30):12–16
22. Grieves M (2014) Digital twin: manufacturing excellence through virtual factory replication. Michael W. Grieves, LLC
23. Negri E, Fumagalli L, Macchi M (2017) A review of the roles of digital twin in CPS-based production systems. Proc Manuf 11:939–948
24. Huynh BH, Akhtar H, Sett MK (2019) A universal methodology to create digital twins for serial and parallel manipulators. In: 2019 IEEE international conference on systems, man and cybernetics (SMC), pp 3104–3109. Barl, Italy
25. Paripooranan CS, Vivek DC, Abishek R, Karthik S (2020) An implementation of AR enabled digital twins for 3-D printing. In: 2020 IEEE international symposium of smart electronic systems (iSES), Chennai, India
26. Tao F, Qi Q, Wang L, Nee AYC (2019) Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: correlation and comparison. Engineering 5(4):653–661
27. Ketzler B et al (2020) Digital twins for cities: a state of the art review. Built Environ 46(4):547–573
28. Ketzler B et al (2020) Digital twins for cities: a state of the art review. Built Environ 46(4):547–573
29. Leng J et al (2019) Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop. J Ambient Intell Humaniz Comput 10:1155–1166
30. Bauer M et al (2013) IoT reference model. Enabling things to talk. Springer, Berlin, Heidelberg, pp 113–162
31. Toivonen V, Lanz M, Nylund H, Nieminen H (2018) The FMS training center—a versatile learning environment for engineering education. Proc Manuf 23:135–140
32. Lohtander M et al (2018) Micro manufacturing unit and the corresponding 3D-model for the digital twin. Proc manuf 25:55–61
33. Al-Ali AR et al. (2020) Digital twin conceptual model within the context of internet of things. Future Internet 12(163)
34. Lawton G (2021) VentureBeat. [Online] Available at: https://venturebeat.com/2021/03/12/why-accenture-lists-digital-twins-as-top-five-technology-trend-in-2021/[Accessed 13 July 2021]
35. Aho P (2020) An open source digital twin framework. Master's Thesis, s.l., Tampere University
36. Qi Q et al (2021) Enabling technologies and tools for digital twin. J Manuf Syst 58(Part B):3–21

37. Nord JH, Koohang A, Paliszkiewicz J (2019) The internet of things: review and theoretical framework. Expert Syst Appl 133:97–108
38. Lee I, Lee K (2015) The internet of things (IoT): applications, investments, and challenges for enterprises. Bus Horiz 58(4):431–440
39. Ornes S (2016) Core concept: the internet of things and the explosion of interconnectivity. Proc Natl Acad Sci 113(40):11059–11060
40. Ben-Daya M, Hassini E, Bahroun Z (2019) Internet of things and supply chain management: a literature review. Int J Prod Res 57(15–16):4719–4742
41. Dave A (2020) iotforall.com. [Online] Available at: https://www.iotforall.com/how-digital-twins-accelerate-the-growth-of-iot. Accessed 24 July 2021
42. Amisha, Malik P, Pathania M, Rathaur VK (2019) Overview of artificial intelligence in medicine. J Family Med Prim Care 8(7):2328–2331
43. Teng X, Gong Y (2018) Research on application of machine learning in data mining. IOP Conf Ser Mater Sci Eng 392(6)
44. Dilmegani C (2021) AI multiple. [Online] Available at: https://research.aimultiple.com/digital-twins/. Accessed 24 July 2021
45. Dohrmann K, Gesing B, Ward J (2019) Digital twins in logistics: a DHL perspective on the impact of digital twins on the logistics industry. DHL Customer Solutions & Innovation, Troisdorf
46. Esmaeilbeigi M et al (2020) A low cost and highly accurate technique for big data spatial-temporal interpolation. Appl Numer Math 153:492–502
47. Vassakis K, Petrakis E, Kopanakis I (2018) Big data analytics: applications, prospects and challenges. In: Skourletopoulos G et al (eds) Mobile big data. Lecture notes on data engineering and communications technologies, p 1. Springer, Cham
48. Arunachalam D, Kumar N, Kawalek JP (2018) Understanding big data analytics capabilities in supply chain management: unravelling the issues, challenges and implications for practice. Transp Res Part E Logistics Transp Rev 114:416–436
49. Zerhari B, Lahcen AA, Mouline S (2015) Big data clustering: algorithms and challenges. In: Proceeding of international conference on big data, cloud and applications (BDCA'15)
50. Chen H-M, Kazman R, Haziyev S, Hrytsay O (2015) Big data system development: an embedded case study with a global outsourcing firm. Florence, Italy, IEEE
51. Polat S, Esen F, Bilgic E (2019) Analysis of the 5Vs of big data in virtual travel organizations. In: big data and knowledge sharing in virtual organizations, pp 43–70. IGI Global
52. Frankenfield J (2020) Investopedia. [Online] Available at: https://www.investopedia.com/terms/d/data-analytics.asp. Accessed 30 Dec 2020
53. IEEE Big Data (2018) Big data for digital twins. Seattle, IEEE
54. Hosseinian-Far A, Ramachandran M, Slack CL (2018) Emerging trends in cloud computing, big data, fog computing, IoT and smart living. Technology for smart futures. Springer, Cham, pp 29–40
55. Kumar PR, Raj HP, Jelciana P (2018) Exploring data security issues and solutions in cloud computing. Proc Comput Sci 125:691–697
56. Khan A et al (2020) Towards smart manufacturing using spiral digital twin framework and twinchain. IEEE Trans Ind Inform
57. Steindl G et al (2020) Generic digital twin architecture for industrial energy systems. Appl Sci 10(24)
58. Greer C, Burns M, Wollman D, Griffor E (2019) Cyber-physical systems and internet of things. National Institute of Standards and Technology (NIST) Special Publication 1900-202
59. Lee J, Bagheri B, Kao H-A (2015) A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf Lett 3:18–23
60. Ahmadi A, Cherifi C, Cheutet V, Ouzrout Y (2017) A review of CPS 5 components architecture for manufacturing based on standards. Colombo, Sri Lanka
61. Redelinghuys A, Basson A, Kruger K (2019) A six-layer digital twin architecture for a manufacturing cell. In: Borangiu T, Trentesaux D, Thomas A, Cavalieri S (eds) Service orientation in holonic and multi-agent manufacturing. SOHOMA 2018 studies in computational intelligence, pp 412–423. Springer, Cham

62. Redelinghuys AJH, Basson AH, Kruger K (2020) A six-layer architecture for the digital twin: a manufacturing case study implementation. J Intell Manuf 31:1383–1402
63. Abburu S et al. (2020) COGNITWIN—hybrid and cognitive digital twins for the process industry, pp 1–8. Cardiff, UK
64. Josifovska K, Yigitbas E, Engels G (2019) Reference framework for digital twins within cyber-physical systems, pp 25–31. Montreal, QC, Canada
65. Fuller A, Fan Z, Day C, Barlow C (2020) Digital twin: enabling technologies, challenges and open research. IEEE Access 8:108952–108971
66. Neto AA, Deschamps F, Silva ER, Lima EP (2020) Digital twins in manufacturing: an assessment of drivers, enablers and barriers to implementation. Proc CIRP 93:210–215
67. Qi Q, Tao F (2018) Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. IEEE Access 6:3585–3593
68. Xu Y, Sun Y, Liu X, Zheng Y (2019) A digital-twin-assisted fault diagnosis using deep transfer learning. IEEE Access 7:19991–19999
69. Philips (2018) philips.co.uk. [Online] Available at: https://www.philips.co.uk/healthcare/resources/feature-detail/ultrasound-heartmodel? Accessed 25 July 2021
70. Liu Y et al (2019) A novel cloud-based framework for the elderly healthcare services using digital twin. IEEE Access 7:49088–49101
71. Orcajo EM (2021) LinkedIn. [Online] Available at: https://www.linkedin.com/pulse/6-digital-twin-applications-healthcare-revolution-enrique/. Accessed 25 July 2021
72. Kosowatz J (2021) The American society of mechanical engineers (ASME). [Online] Available at: https://www.asme.org/topics-resources/content/smart-cities-look-for-digital-twins. Accessed 25 July 2021
73. Khajavi SH et al (2019) Digital twin: vision, benefits, boundaries, and creation for buildings. IEEE Access 7:147406–147419
74. Kumar H, Singh MK, Gupta MP, Madaan J (2020) Moving towards smart cities: solutions that lead to the smart city transformation framework. Technol Forecast Soc Chang 153
75. Niaros V, Kostakis V, Drechsler W (2017) Making (in) the smart city: the emergence of makerspaces. Telematics Inform 34(7):1143–1152
76. Farsi M, Daneshkhah A, Hosseinian-Far A, Jahankhani H (2020) Digital twin technologies and smart cities. Springer, Cham
77. Kosowatz J (2020) The American society of mechanical engineers. [Online] Available at: https://www.asme.org/topics-resources/content/top-10-growing-smart-cities. Accessed 26 July 2021
78. Hinduja H, Kekkar S, Chourasia S, Chakrapani HB (2020) Industry 4.0: digital twin and its industrial applications. Int J Sci Eng Technol 8(4)
79. Sepasgozar SM (2020) Digital twin and web-based virtual gaming technologies for online education: a case of construction management and engineering. Appl Sci 10(13)
80. Rudskoy A, Ilin I, Prokhorov A (2021) Digital twins in the intelligent transport systems. Transp Res Proc 54:927–935
81. Zhaohui W et al (2021) Review on the construction and application of digital twins in transportation scenes. J Syst Simul 33(2):295–305
82. Okumus F et al (2019) A bibliometric analysis of lodging-context research from 1990 to 2016. J Hospitality Tourism Res 43(2):210–225
83. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH (2017) Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Syst Rev 6(245)
84. Rice DB et al (2016) Are MEDLINE searches sufficient for systematic reviews and meta-analyses of the diagnostic accuracy of depression screening tools? A review of meta-analyses. J Psychosom Res 87:7–13
85. Moher D et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7)
86. Hallinger P, Chatpinyakoop C (2019) A bibliometric review of research on higher education for sustainable development, 1998–2018. Sustainability 11(2401)

87. Jan van Eck N, Waltman L (2018) VOSviewer Manual. Universiteit Leiden, Leiden, Netherlands
88. Zupic I, Čater T (2015) Bibliometric methods in management and organisation. Organ Res Methods 18:429–472
89. Karakus M, Ersozlu A, Clark AC (2019) Augmented reality research in education: a bibliometric study. EURASIA J Math Sci Technol Educ 15(10)
90. Pires F et al (2019) Digital twin in industry 4.0: technologies, applications and challenges, pp 721–726. IEEE
91. Ross PT, Zaidi NLB (2019) Limited by our limitations. Perspect Med Educ 8(4):261–264
92. Jalal SK (2019) Co-authorship and co-occurrences analysis using BibliometrixR package: a case study of India and Bangladesh. Ann Libr Inf Stud 66:57–64
93. Chen X et al (2016) Mapping the research trends by co-word analysis based on keywords from funded project. Proc Comput Sci 91:547–555
94. Philip H, Chatpinyakoop C (2019) A bibliometric review of research on higher education for sustainable development, 1998–2018. Sustainability 11(2401)
95. Ghobakhloo M (2018) The future of manufacturing industry: a strategic roadmap toward Industry 4.0. J Manuf Technol Manage 29(6):910–936
96. Tao F, Zhang M (2017) Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. IEEE Access 5:20418–20427
97. Tao F, Zhang H, Liu A, Nee AYC (2019) Digital twin in industry: state-of-the-art. IEEE Trans Indus Inform 15(4):2405–2415
98. Söderberg R, Wärmefjord K, Carlson JS, Lindkvist L (2017) Toward a digital twin for real-time geometry assurance in individualized production. CIRP Annals 66(1):137–140
99. Tao F et al (2019) Digital twin-driven product design framework. Int J Prod Res 5(12):3935–3953
100. Tao F, Zhang M, Liu Y, Nee AY (2018) Digital twin driven prognostics and health management for complex equipment. CIRP Ann 6(1):169–172
101. Zhuang C, Liu J, Xiong H (2018) Digital twin-based smart production management and control framework for the complex product assembly shop-floor. International J Adv Manuf Technol 96(1):1149–1163
102. Bolton RN et al (2018) Customer experience challenges: bringing together digital, physical and social realms. J Service Manage 29(5):776–808
103. Zheng P, Lin TJ, Chen CH, Xu X (2018) A systematic design approach for service innovation of smart product-service systems. J Clean Prod 201:657–667
104. Zhang X, Chen H, Wang W, de Pablos PO (2016) What is the role of IT in innovation? A bibliometric analysis of research development in IT innovation. Behav Inf Technol 35(12):1130–1143
105. Ivanov D, Dolgui A (2021) A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. Prod Plann Control 32(9):775–788
106. Knapp GL, Mukherjee T, Zuback JS, Wei HL, Palmer TA, De A, DebRoy TJAM (2017) Building blocks for a digital twin of additive manufacturing. Acta Materialia 135:390–399
107. Ding K, Chan FT, Zhang X, Zhou G, Zhang F (2019) Defining a digital twin-based cyber-physical production system for autonomous manufacturing in smart shop floors. Int J Prod Res 57(20):6315–6334
108. Schluse M, Priggemeyer M, Atorf L, Rossmann J (2018) Experimentable digital twins—Streamlining simulation-based systems engineering for industry 4.0. IEEE Trans Indus Inform 14(4):1722–1731
109. Glaessgen E, Stargel D (2012). The digital twin paradigm for future NASA and US Air Force vehicles. In: 53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA, 1818p
110. Boschert S, Rosen R (2016) Digital twin—the simulation aspect. Mechatronic futures. Springer, Cham, pp 59–74

111. Uhlemann THJ, Lehmann C, Steinhilper R (2017) The digital twin: Realizing the cyber-physical production system for industry 4.0. Procedia Cirp 61:335–340
112. Lu Y, Liu C, Kevin I, Wang K, Huang H, Xu X (2020) Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. Robot Comp Integr Manuf 61:101837
113. Datta SPA (2017) Emergence of digital twins. J Innov Manage 5:14–34
114. Hartmann D, der Auweraer HV (2021) Digital twins. Progress in industrial mathematics: success stories. Springer, Cham, pp 3–17
115. Wang KW et al (2017) Digital China: powering the economy to global competitiveness. McKingsley and Company, New York
116. Wood T (2021) Visual capitalist. [Online] Available at: https://www.visualcapitalist.com/national-innovation-the-most-innovative-countries-by-income/. Accessed 9 Sept 2021
117. García-González J, Forcén P, Jimenez-Sanchez M (2019) Men and women differ in their perception of gender bias in research institutions. PLoS ONE 14(12)

Emerging Technologies: Blockchain and Smart Contracts

Aristeidis Davelis, Usman Javed Butt, Gemma Pendlebury, and Khaled El Hussein

Abstract This chapter begins by briefly covering the history of Blockchain and introducing its core elements, continuing to explain the fundamentals of Blockchain technology and Smart Contracts. A discussion is made on nodes, consensus mechanisms, digital signatures and cryptographic hashes, types of blockchains, Ethereum, and Smart Contracts benefits. After that, it explores the distributed ledger technology (DLT) and blockchain as a subset of DLT in greater detail, discussing the benefits and challenges of DLT with Blockchain. It then presents us with some of the interesting use cases of Blockchain within various industries including financial, healthcare, manufacturing, and agriculture. Furthermore, it provides a roadmap for successfully implementing Blockchain in modern business, with recommendations on preparation, design and planning, implementation, and review. Finally, it explores future trends in DLT, blockchain and Smart Contracts.

Keywords Blockchain · Smart contracts · Distributed ledger technology · Bitcoin · Ethereum · Consensus · Blockchain use cases · Blockchain implementation · Blockchain benefits and challenges · Future of blockchain

1 Introduction

The emergence of Bitcoin and various other cryptocurrencies over the first two decades of the twenty-first century, gave birth to the first implementation of Blockchain. In its conventional form, Blockchain can be perceived as a set of immutable and tamper-proof data, arranged in cryptographically linked discrete sets

A. Davelis · U. J. Butt (✉) · G. Pendlebury · K. E. Hussein
Northumbria University Engineering and Environment, London, UK
e-mail: usman.butt@northumbria.ac.uk

A. Davelis
e-mail: aris.davelis@northumbria.ac.uk

K. E. Hussein
e-mail: khaled.el-hussein@northumbria.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_6

Table 1 The value at stake from blockchain varies across industries adapted from: Carson et al. [3] (McKinsey & Company)

Industry	Revenue potential	Cost reduction potential
Agriculture	Average-High	High
Automotive	High	Average
Financial	Average-High	High
Healthcare	High	High
Insurance	Average	High
Real estate	High	High
Public	High	High
Media and communications	High	Average
Transport and logistics	Low-Average	Average-High
Utilities	Average-High	High

of information called blocks [1]. Data on the Blockchain is shared among participants via a distributed ledger, similar to a database stored in a decentralised system [2].

Growing beyond cryptocurrency, Blockchain found applications in various sectors with huge potential impact, due to its traceability, trust, immutability and decentralisation properties. Additionally, Blockchain technology quickly evolved with the inclusion of "Smart Contracts", which allow for programmable instructions that can be set to automatically trigger when specific conditions are met [3]. These benefits provide Blockchain with the ability to increase revenues and reduce costs in multiple industries. Some of the most impacted industries in terms of benefiting from Blockchain can be seen on Table 1.

1.1 From Conception to Implementation

Although Blockchain is considered an emerging technology, the concept of Blockchain is far from new. The idea of a distributed, secure, trusted system which would replace the need for mutual trust between varying parties, was introduced as early as 1979 [4]. Chaum's suggestions followed the basic principles behind Distributed Ledger Technology (DLT), with the later-conceived Blockchain being a specific type of DLT.

Later in 1990, the use of linked cryptographic hash functions had been suggested as a means to providing a "digital document time-stamping service", which could work under a "distributed trust" environment [5]. The proposal aimed to digitally sign and timestamp documents, in order to ensure they have not been tampered with. This essentially described the fundamental idea behind Blockchain. Very soon afterwards, cryptographer Nick Szabo introduced the idea of Smart Contracts, as a

means of securely and reliably automating contract execution over a network. Under this concept, a contract's clauses could be "embedded" in hardware or software components and make the breach of the contract nearly impossible or prohibitively expensive [6].

It wasn't until over a decade later however that the Blockchain concept started materialising. In 2008, a cryptographer using the alias "Satoshi Nakamoto" published a paper on a distributed digital currency system called "Bitcoin", utilising a timestamp server and cryptographic hash functions. These were used to link immutable "blocks" of transactional and other data in the form of a chain, which were verified by a "Proof-of-Work" consensus process [7]. Bitcoin was quickly implemented and the first Bitcoin transaction took place in 2009, giving birth to multiple cryptocurrencies over the following years.

Finally, the Ethereum cryptocurrency and blockchain implementation actually incorporated Smart Contracts upon its release, as mentioned in its white paper in 2014. In Ethereum, as Buterin [8] describes, Smart Contracts could be perceived as programmable entities on top of the blockchain, similar to "cryptographic 'boxes' that contain value which only unlocks when certain conditions are met".

Since the emergence of Ethereum, Blockchain has been undergoing constant development and improvement, evolving beyond cryptocurrency and establishing itself in multiple business applications in numerous sectors as mentioned earlier.

1.2 Core Elements

Some of the main concepts of Blockchain and DLT are presented in summary on Table 2, and will be discussed in greater detail in the following sections.

Table 2 Core elements of blockchain and DLT

Concept	Description
Decentralisation	The fact that no single entity or institution controls the operation of Blockchain [9]
Blocks	Discrete linked sets of transactions (or other information) stored on the Blockchain
Distributed Ledger	The record of all transactions, stored and shared across all devices on the Blockchain network
Consensus	The process of verifying new transactions [2], in order to agree on the addition of new blocks and the new state of the ledger
Nodes	Devices participating in the Blockchain network and the consensus process, holding copies of the ledger
Miners	The nodes competing to add new blocks to a cryptocurrency blockchain, and be awarded cryptocurrency if successful [10]
Smart contracts	Programmable instructions stored on the chain, that can be set to automatically trigger when specific conditions are met [3]

2 The Fundamentals of Blockchain

The Bitcoin cryptocurrency could be considered the first significant practical implementation of blockchain technology. It is an open-source, public, global, peer- to-peer (P2P) Distributed Ledger Technology, where users request transactions which are placed on the Bitcoin network, while being authenticated using their personal digital signature. Sets of financial transactions are pooled and then recorded in a 'block' that is added to a chain of similar blocks, in chronological order. Each block contains a cryptographic hash of the previous block's header information, to link the blocks in a chain. This is performed by decentralised Bitcoin nodes (miners) who hold a record of the blockchain ledger [7]. The Bitcoin nodes act collectively as a decentralised time-stamp server that uses Proof of Work (PoW) computation for the consensus mechanism, proving the chronological order of transactions and allowing blocks to be added to the chain once tested and confirmed by other nodes in the network [11].

2.1 Node Architecture

Blockchain is a type of DLT, which as mentioned earlier is a term for any shared databases, which are distributed amongst various parties. It is important to note however, that not all DLTs use blockchain technology.

In their purest implementations, blockchain systems are both distributed and decentralised. Nodes on a distributed system communicate with each other directly to fulfil a shared aim as a single platform; node coordination and fault tolerance however remains a concern. To protect themselves against issues caused by faulty nodes, blockchain systems utilise consensus mechanisms with varying fault tolerance levels. While distributed systems can still be controlled by a central authority, blockchain's decentralised concept means that no one central authority controls the network; it is instead controlled by the computations of the distributed nodes through the consensus model. This formation without a centralised controller is known as peer-to-peer (P2P) architecture [12] (Fig. 1).

2.2 Consensus Types

Blockchain uses consensus mechanisms, for nodes on the network to reach agreement on the blockchain state at any given time. This is challenging, as the algorithms must handle security of transmissions, synchronisation, failure, performance, and malicious nodes, as Chicarino et al. note [13]. Consequently however, consensus is critical to blockchain security [14].

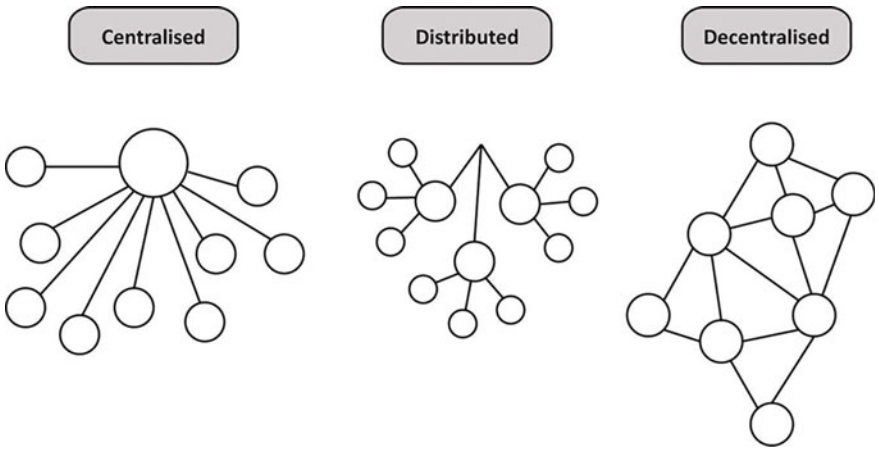


Fig. 1 Comparison of centralised, distributed and decentralised systems. Adapted from: Bashir [12]

There are many consensus mechanisms. The Bitcoin consensus type is called Proof of Work (PoW) and it uses the SHA-256 hash algorithm. The computing power required to complete PoW directly correlates with the economic value gained, aiming to disincentivise attackers. In Bitcoin's PoW consensus, a computed block hash must result in a number lower than the latest network target, in order for the hash to be valid. The average number of attempts to achieve this is the difficulty rating, and network hash rate is the times per second it takes to produce a valid hash. The target is changed over time depending on the nodes' processing power, to maintain a block generation time of approximately 10 min [15].

PoW is problematic due to the computing power it requires leading to a high consumption of energy and potentially negative environmental impact. Various consensus mechanism alternatives have been proposed as a solution to this issue. As Shibata [11] mentions, one alternative for cryptocurrency is Proof of Stake (PoS) which consumes less energy than PoW. With PoS, miners must prove their ownership of the currency to validate in consensus, working on the assumption that people who own more of the currency would be less likely to compromise the system. Some PoS solutions combine stake size with currency age or randomisation, to mitigate unfairness in the likes of the richest owner dominating the blockchain.

Delegated Proof of Stake (DPoS) is similar to PoS, but stakeholders elect the nodes delegated to produce blocks.

Another popular consensus mechanism is Practical Byzantine Fault Tolerance (PBFT), which is designed to tolerate byzantine faults (unknown percentage of faulty nodes) and malicious nodes, requiring every node to be visible to the network. It is managed by a primary node which is selected to sequence the transaction phases; pre- prepared, prepared, and commit. Each phase transitions to the next when at least two thirds of nodes have agreed [16].

2.3 Digital Signatures and Cryptographic Hash

In Bitcoin, digital signatures are used to prove that a transaction was legitimately initiated by a specific user. When a user requests a transaction, they must sign it with their private key to encrypt the transaction before it is broadcast to the network. Then the mining nodes validate the signature's authenticity using the public key, in order to confirm that the funds belong to the user who made the request, thus ensuring non-repudiation and data origin authentication. Bitcoin uses the Elliptical Curve Digital Signature Algorithm (ECDSA) to generate the key pairs, which is based on the standard Digital Signature Algorithm (DSA). As Raj [17] notes, the ECDSA uses a mathematical equation to create points on a graph curve, which are then used with a randomly generated number to calculate a private and public key pair.

In general, cryptographic hashes are integral to the Bitcoin blockchain. A cryptographic hash function converts input data, to output a corresponding string of fixed length known as a 'hash' or 'message digest'. To complete PoW and add a block to the chain, Bitcoin miners must use computing power to solve a complex mathematical puzzle, in the form of discovering the value of a pseudo-random generated number called a nonce (number only used once). As Bitcoin [18] states, the nonce is a random number which miners need to guess to produce the correct hash output and add the block to the chain. The nonce is then passed through a SHA- 256 cryptographic hash function in combination with transaction data to produce the result. Once complete, the result is broadcast to the other miners who verify whether it is correct, and if it is, the new block is added.

Whenever a new block is added to the chain, it also contains a link to the previous block using a cryptographic hash of the previous block header. This makes it almost impossible to tamper with, as changing one block would mean all previous blocks would need to be altered. Bitcoin blockchain headers contain the block version, the previous block header's (SHA256) hash (which ensures previous block headers cannot be altered without altering this), the Merkle root (SHA256) hash (computed from the block's transactions and thus ensuring they cannot be altered without altering the header), a timestamp, nBits (a value that the header must be equal to or less than), and the verified nonce. As Maleh et al. [19] point out, the distributed users' need for consensus, alongside the cryptography, protects blockchain against malicious alterations of the chain.

2.4 Types of Blockchains

There are different types of blockchain networks, each with its individual attributes:

Public (or permissionless) blockchains are transparent and open to the public to use, without a central authority governing them. The system is governed by the

blockchain network itself, with Bitcoin and most common cryptocurrencies being a widely known example.

Private (or permissioned) blockchains, where access to the system is restricted by a centralised authority and network transactions are not transparent to unauthorised parties [20]. These blockchains are common in corporate environments.

Consortium blockchains, which are controlled by a group of organisations [21]. As Maleh et al. [19] note, in this type of blockchain the network nodes are predetermined and controlled by the consortium. This essentially creates a semi- private blockchain concept, where the system is divided into a private section accessed by known entities and a public part available to anyone.

Another type of blockchain, based on architecture rather than access permissions and governance, is the sidechain (or pegged sidechain). As Bashir [12] highlights, sidechain refers to a blockchain linked to another blockchain using a two-way peg, to allow assets to be moved across the two chains (Fig. 2).

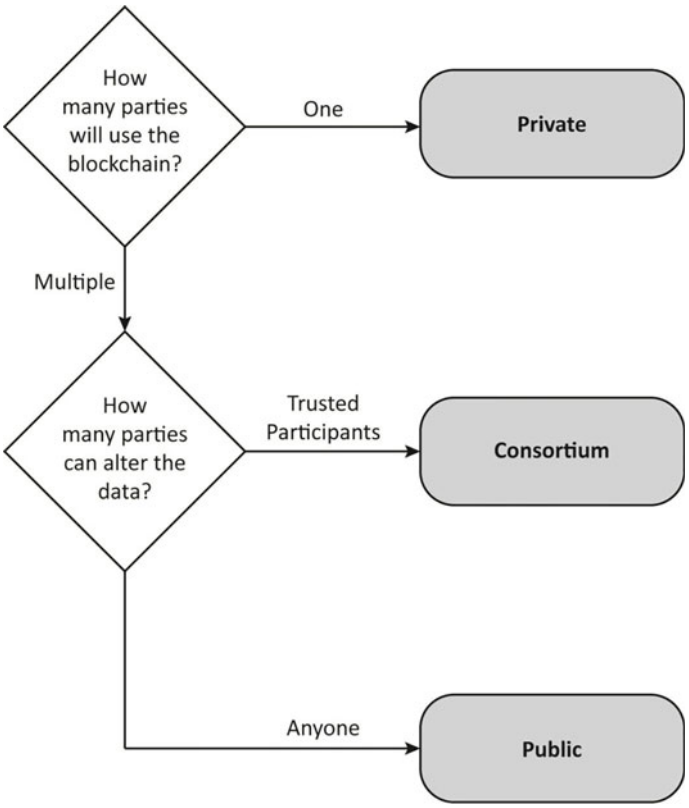


Fig. 2 Types of blockchains. Adapted from: Maleh et al. [19]

2.5 Ethereum and Smart Contracts

As Staples et al. [22] note, Smart Contracts store pre-programmed instructions that automatically execute transactions on the blockchain, when certain criteria are met. This negates the need for a centralised trusted third party, which would otherwise be required to perform clearing, settlement, or funding from issuers. Smart Contracts are programmed with predefined functions that determine what is the correct output based on input, in the execution of an agreement between multiple parties. The contract is hosted on the blockchain and after development is complete, the final code is stored on the blockchain to be invoked by relevant transactions. There are multiple platforms that facilitate Smart Contracts (e.g., Ethereum, Hyperledger Fabric, and NXT). Ethereum is the most popular platform implementing Smart Contracts; programmers can use Solidity to write decentralised apps (DApps) which are compiled using the Turing-complete runtime environment of the Ethereum Virtual Machine (EVM), and funded through a payment of Ethereum called 'gas' [23]. As Khan et al. [24] mention, the amount of 'gas' depends on the complexity of the contract. Furthermore, as Petrov [25] states, because the contract can be programmed to span the entire process of the agreement on the blockchain, all transactions are secure, immutable, and clearly auditable.

Smart Contracts can cover the entire agreement or accompany a traditional contract, and just handle the execution of certain actions, like the movement of funds. Like any software code, the parameters and functions of Smart Contracts need to be clearly defined, and often relatively simple requests are actioned. For example, if one rule of a contract between parties X and Y is honoured by party X, that triggers the contract to transfer some agreed funds from party Y to party X. However, developers are also using Smart Contracts to invoke other Smart Contracts which increases the potential complexity. Even so, Smart Contracts are not suitable for subjective decision making on contracts where there is any ambiguity involved, but they are very suitable for paying funds when certain events take place, or forfeiting fund as a penalty for commitments not being met. As Levi and Lipton [26] highlight, because the Smart Contract runs automatically on the blockchain, an intermediary (e.g. a judiciary or escrow holders) is not required to action or enforce the contract. Moreover, the enforcement of Smart Contracts is anonymous because there is no centralised control over the blockchain [27].

There are several variations of Smart Contracts; some have terms which are primarily written in natural language, but can use code to automatically perform the actions required by the agreement. Other contracts are only written in code, without any natural language describing the terms. Finally, hybrid contracts are a combination of the two to a greater or lesser extent. For example, the terms are mainly code based, but some provisions are written in natural language [28].

3 Beyond Blockchain

3.1 Distributed Ledger Technology

Distributed Ledger Technology (DLT) emerged before the development of blockchain and Bitcoin. As mentioned, early development of DLT was identified in the works of Bayer et al. [29] and Haber and Stornetta [5], wherein the chain of cryptographically linked data blocks was developed for secure and efficient handling of the digital data with the inclusion of hashing functions and Merkle Trees. In addition, DLT can be considered as an umbrella term wherein multiple systems operate without any central authority or operator. Furthermore, blockchain technology can be considered as a subset of the DLT environment, because of the use of hash-linked data blocks.

The definition of DLT as per Natarajan et al. [30] is a broad category of shared ledgers, which are defined for sharing records amongst multiple parties. On the other hand, the European Central Bank defines DLT as a technology that provides the benefit of storing and accessing information associated with specific assets to the users and holders within a shared database [31]. Regardless of definition specifics, as Rozario and Thomas [32] state, the use of DLT reduces the dependency on a central validation system, which consequently provides the users with the ability to settle their transactions independent of information of securities and cash within the distributed system.

There is no central coordinator existing in a DLT which operates in the traditional form. Figure 3 depicts that DLT provides multiple controls to different parties, making the system more preferable as compared to any other distributed or centralised database.

DLT also includes linking of cryptographic hashes for developing tamper evidence, and the result sharing helps in building an authoritative aspect of the records. It is also important to understand the concept of Ledger within DLT because

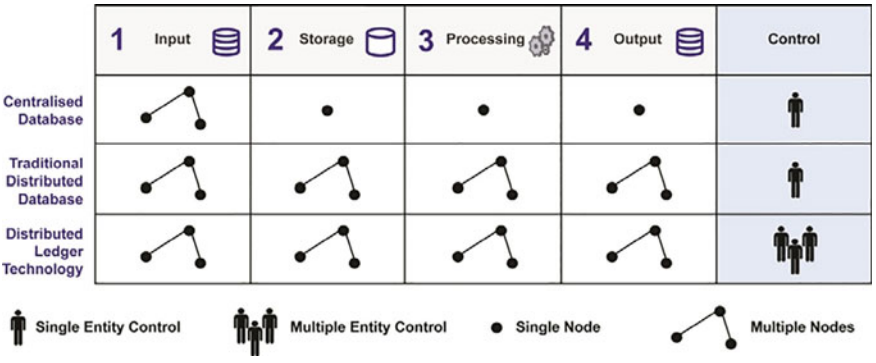


Fig. 3 Centralised databases and DLT

it is one of the most significant components of the system. The ledger would hold different meanings as per the application of DLT system, but as Hewa et al. [33] state, it is mainly developed through the data collected by an individual node, and the data that is commonly held by the majority of the nodes (Fig. 4).

Another significant aspect of the DLT system is the private key which is often referred to while making authorised transactions in the system [34]. This private key is a cryptographic sign of the initiator, which helps in changing the record state and fulfils the transaction instructions [35]. As Alharby et al. [36] highlight, private keys are very significant because they provide validation and guarantee that the transaction has been initiated by a true holder, proving there is no compromise with the safety of the ledger or the overall system.

DLT systems are also composed of multiple actors, which are interacting in the system. These actors are grouped into four significant categories as depicted on Fig. 5.

DLT systems are based on three core layers; the Network, the Protocol and the Data layer, which include sets of components and processes that enable the overall system to function [37]. The system view of the DLT is depicted on Fig. 6. This demonstrates that the foundation of DLT is based on the protocol layers, wherein the

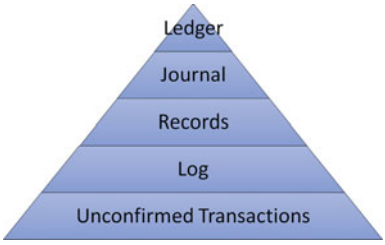


Fig. 4 Transactions to records—Development of a ledger in DLT

Fig. 5 Actors of DLT

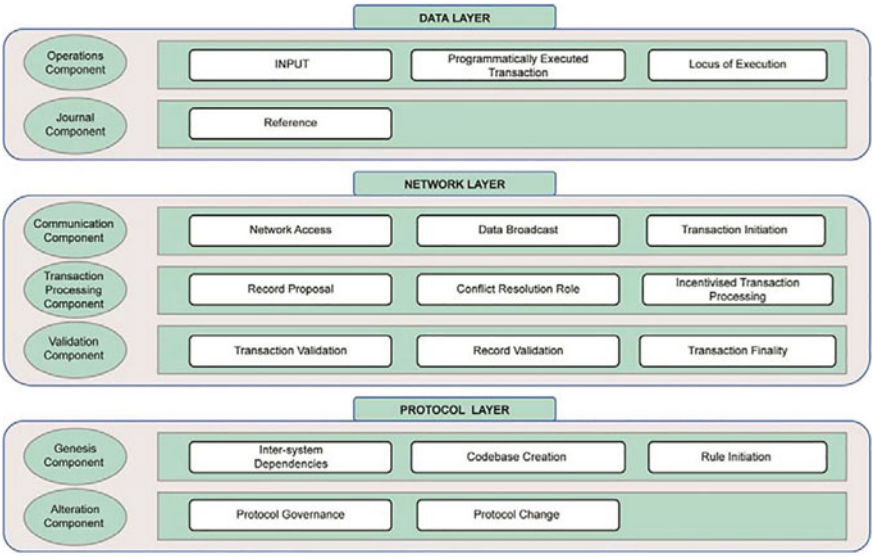


Fig. 6 DLT system layer view. Adapted from: Rauchs [39]

formal rules and the governing codes are established. The two components of the Protocol layer, Genesis and Alteration, are responsible for the network launch and the evolution of the system on time respectively [38].

Secondly, the Network layer is based on the implementation of the protocols that are defined in the foundational layer, and contains the identification of the ways through which data would be shared and updated in the network and the ledger. This layer also has three significant components. These are the Communications, Transaction Processing and Validation components, which enable better incorporation of the authorisation in the records [31]. Finally, the third layer of the DLT system is the Data layer, where the Operations and the Journal components are found. These are governing the flow of data in this layer, along with the storage of records as per the references and engagement of the nodes [40].

3.2 Benefits of DLT and Blockchain

DLT and blockchain systems provide the benefit of independent validation, wherein each participant in the system has the independency to verify the transaction state and maintain the integrity of the system [41]. Furthermore, they offer the benefit of shared recordkeeping, since multiple parties are collectively maintaining, creating and updating the records as per their level of authority.

Additionally, the system in the form of blockchain is capable of developing tamper evidence and censorship resistance, which indicates that a single party performs

the transactions or processes unilaterally [42]. The censorship resistance can also appear with the blocking or censoring of the transactions, or the change in the rules of the overall system within blockchain. Moreover, the decentralisation feature of DLT systems can be enabled at the desired degree at the different layers, as per the requirement of the system actors [43]. This implies that there is higher tamper resistance along with higher censorship resistance and low trust requirement, which makes the DLT a more desirable and reliable system as compared to fully centralised ones.

Apart from the aforementioned benefits, a DLT system or blockchain provides flexibility and freedom in deploying applications with a customised perspective. The distributed ledger has opaqueness and high-performance capabilities which enable efficient use of time and resources over the computers [44]. On top of this, the security of the distributed ledger is quite high because there is no compromise of stored data and easier access achieved through everyday practice.

The DLT system implemented in the form of blockchain would also be providing benefits like cost reduction within the organisation, due to easier reporting and auditing. The institutions would be able to perform multiple tasks through the use of blockchain which provides the benefit of speed and efficiency. Transactions are performed in seconds or less, which would be contributing to better management of time and efficiency requirements.

3.3 Challenges of DLT and Blockchain

The adversarial environment is considered as a strong indicator for the DLT system to have malicious actors as part of the system [45]. This environment depicts that the system is not used for the purpose it was intended for and there is exploitation of the consensus in the system, due to unauthorised authentication and disrupted network transferred transactions [46]. In other words, the intrusion in the system through unauthorised access will be affecting the DLT system's transactions.

The DLT systems are subjected to the challenge of compromise when the private keys are stolen and there is no security available for the transactions that are initiated with authorised attempts [47]. This provides an opportunity to the hackers to steal the data through engagement in the transactions without the knowledge of the true owners and authorised parties in the DLT.

The organisations with slow and cumbersome technological developments can experience the challenge of lack of scalability. Additionally, the difficulty in integration of the blockchain will be affecting the overall performance and complexity of the system. Additionally, the scalability and transaction speed of public blockchain systems is quite low, as compared to the private and consortium blockchains, making them less reliable. The governance of the private and consortium blockchains is also stronger compared to public ones [48]. This indicates that there is higher susceptibility to the environment, regarding the functioning of blockchains and DLTs without centralised administration.

Finally, contradictory to the advantages and benefits explored for DLT and blockchain, there is a technological challenge due to DLT still being the early stage of development. There are significant concerns prevailing in terms of the resilience and robustness which can affect the functioning of large software and the volumetric transactions [49].

4 Use Case Examples

A few use case examples based on real life applications of Blockchain and Smart Contracts are discussed in this section. This list is in no way exhaustive, aiming to provide a better view of how blockchain has already started adding value to a multitude of sectors, industries and markets.

4.1 Bitcoin and Cryptocurrencies

Bitcoin cryptocurrency could be considered the original implementation of blockchain, and has been recognised as a legitimate means of payment for goods and services in certain industries. As Raj [17] points out, cryptocurrencies differ from traditional currency, in that their creation and transference is decentralised and does not require an intermediary. Along these lines, Bitcoin was designed to solve the previous digital payment systems' weaknesses (including double spending) and remove the need for trusted intermediaries to mediate financial disputes, by creating a peer-to-peer system which relies on cryptographic proof of immutably recorded transactions taking place and their chronological order, rather than trust [7].

Blockchain has been described as the Internet of Value and Bitcoin became the first unit of value therein. It has proven blockchain's usefulness in transferring, and accounting for, value on a decentralised ledger across a peer-to-peer network. The first real world payment using Bitcoin was in 2010; since then Bitcoin's success has been followed by the creation of a multitude of alternative cryptocurrencies. A prominent example is Ethereum. Ethereum was developed by Vitalik Buterin, whose fascination with Bitcoin led him to start a project in 2013 to develop a blockchain platform (with a built-in cryptocurrency—Ether) on which developers can build applications [20].

4.2 Smart Contract Implementations

As a digital alternative to traditional contracts detailing agreements between parties, Smart Contracts are useful in multiple industries, including logistics, shipping supply chain, insurance and charities. Especially with regards to cross-organisational agreements, they can increase profit by accelerating processes, reducing real-time tracking

costs and enhancing cross-border payments. Crowdsourcing can also benefit from Smart Contracts, for the purpose of raising funds for a project through small donations from a large number of people. Additionally, charitable organisations could benefit in showing how contributions from various sources are handled so that performance data could be audited.

Due to their transparency, Smart Contracts can also be used for proving the provenance of data, managing access and sharing, as a reliable trust-less alternative to trusting a centralised intermediary for data handling. Apart from data provenance, other uses involve product tracing, with asset tracking from food to vehicles being typical examples. Another use case example for Smart Contracts is device management, in order to ensure synchronisation, authentication and data integrity for devices deployed on a decentralised network, rivalling the traditional client–server model. This is particularly useful in the case of IoT devices.

Furthermore, as Khan et al. [24] observe, non-fungible tokens (NFT) are a way to prove a unique asset's ownership using Smart Contracts on Ethereum. NFTs can be used to record ownership for real-estate, collectables, and art. The concept of NFT extends to anything non-fungible as the name suggests, meaning unique and not simply interchangeable with something else, unlike currency which is a fungible item. NFTs create a way to represent uniqueness, scarcity, and proof of ownership in digital form. The ownership is on public record, so it can be easily verified, and items can be exchanged in ways that could be difficult with physical assets without an intermediary. Similar to creating limited print runs of artwork, NFT creators can decide the scarcity by creating limited or numbered replicas, each with a unique ID. NFTs can also be programmed to stipulate if the creator is to receive royalties if the asset is sold on [50].

Another promising application of Smart Contracts is Decentralised Finance (DeFi), which refers to financial services built on blockchains like Ethereum, driven by Smart Contracts. An example is Compound, which is an open-source platform for lending and borrowing cryptocurrency built on Ethereum using Smart Contracts. As Leshner and Hayes [51] note, Compound allows users to earn interest on their cryptocurrency and tokenise assets using native tokens ('cTokens'), which are created from Ethereum ERC-20 tokens. Another example is Stellar, which is an open source blockchain-based system for trading multiple currencies affordably and efficiently. Like Ethereum 'gas', Stellar requires payment to use services in the form of the cryptocurrency 'lumens' [52].

4.3 Financial Services

One key aspect of blockchain is that it can be relatively cost-efficient to maintain financial services on it. This is a quality that could aid poorer jurisdictions, in extending financial services to previously excluded members of society, thus promoting financial inclusivity. By sharing transaction information via a blockchain ledger, blockchain can assist with reducing the cost and time of reconciliation,

which currently suffers from lack of visibility between reconciling organisations. The immutability of blockchain adds an extra layer of security against altering records and makes transactions easily auditable [53].

Under the context of Financial Services, blockchain can also be used for workflow automation with the possibility to script workflows across organisations, using Smart Contracts to validate data and action workflow steps. This could be also combined with an off chain 'oracle' for data inputs to the workflows through a third-party service. In addition, blockchain can increase the security and efficiency of systems where information needs to be shared between financial institutions about customers, providing them with a unique ID and immutably stored information that they can choose to share (e.g. customer identity verification). This also applies to any document sharing, which can be added to the chain with a hash signature for verification and can be linked to transactions for auditing purposes. Furthermore, because of its decentralised architecture, blockchain is well suited to building financial exchanges, offering settlement time and cost reduction while increasing transparency. As Roy [53] mentions, other financial applications of Blockchain include trade finance, cross-border remittance, secure IPOs and asset tracking.

4.4 Anti-Money Laundering

Regulated organisations and financial institutions currently must comply with anti-money laundering (AML) laws by conducting Know-Your-Customer (KYC) checks, which verify an individual's identity at the start of the business relationship and then performing Customer Due Diligence (CDD) throughout the business relationship to ensure they are not facilitating financial crime. The KYC process in itself can be time consuming and disjointed. FATF [54] also points out that financial institutions experience issues with financial inclusion, integration, legacy systems, data sharing and accuracy, when implementing AML solutions.

KYC systems built with blockchain technology can enhance this process by providing a distributed source for customer verification on a consortium blockchain, shared privately among financial institutions. Such systems utilise blockchain's immutability, transparency, and distribution to pool data in a single solution from multiple legacy systems. As Bashir [12] points out, this in turn results in reducing costs and the complexity associated with the KYC process. As an example to this aspect of the technology, ConsenSys and the Codefi product suite is an already established anti- money laundering solution built on blockchain. Codefi is designed to assist organisations with KYC and KYT blockchain monitoring and security auditing tools for smart contracts on Ethereum [55].

4.5 Healthcare and Pharmaceuticals

Blockchain has the potential to address issues currently experienced in the healthcare industry, surrounding control, integration, complexity and auditability. It could help increase availability and trust, decrease expenditure, and protect privacy.

As Bashir [12] mentions, a strong example use case in healthcare is stopping the distribution of counterfeit medicine. Counterfeit medicines can contain incorrect levels of active ingredients with unpredictable effects on the body. According to the World Health Organisation, 0.2 of 1 million deaths are caused by drug tampering and research has shown that 10–15% of global medicines are counterfeited. Studies have shown that blockchain technology can be used to authenticate medicine and enhance the monitoring of their production and distribution, to combat counterfeiting and tampering. One example is using blockchain to establish medicine provenance, tracking each medicine from manufacturer to distributor, then from carrier to hospital or pharmacy. The distributors are registered with a unique ID in the blockchain, and each delivery is recorded, time and location stamped, using their key to sign for verification. Furthermore, sensors are attached to each package that monitor factors like humidity and temperature to ensure the medicine is not spoiled. At the destination, QR codes can be scanned by the consumer to reveal the provenance of the medicine.

In a separate example application, Quzmar et al. [56] discuss that researchers devised a system that recorded information including hospital, administering doctor, drug dosage, dispensing pharmacy, and specific patient data—in addition to the delivery information which could be shared among hospitals which required the data. It used Smart Contracts to define the access permissions for security and confidentiality. As Wilson [57] mentions, British hospitals have already utilised the distributed ledger technology Everyware, developed by Hedera, to track and monitor COVID-19 vaccines. Similarly, researchers are investigating the use of blockchain for COVID-19 vaccine and contact tracing. One obvious challenge is maintaining the balance between transparency and privacy, although as Ricci et al. [58] propose, cryptographic techniques could be utilised to protect highly sensitive data while maintaining auditability.

4.6 Supply Chain

As Hewett et al. [59] discuss, the COVID-19 pandemic highlighted issues with the reliability of existing supply-chain processes for tracking, financing, authenticating, and delivering goods. Blockchain can revolutionise supply chain and trade flows systems and is a major area of development. Warren et al. [60] highlight that trustless and decentralised platforms can maintain a reliable and shared source of data for stakeholder coordination and the supply chain networks' operations.

Blockchain's transparency of transactions and privacy through encryption makes it suitable for tracking the provenance of products, to provide reassurance to purchasers that they are genuinely ethically sourced. Products can be tracked from the source onwards to provide the purchaser with a complete picture, through historical records, of the product's journey thus far. One example is the Everledger platform, which is used to prove the provenance of diamonds and that they are from 'conflict free' sources. Utilising blockchain technology, the diamond is issued with an identifying number and traced, so that its history is transparent to any buyer who wants to verify the source and route that the diamond took, up to the point of purchase. Hargrave and Karnoupakis [20] point out that these benefits can also be applied to other luxury items, as well like designer goods.

While organisations are developing permissioned blockchains for purposes like trade finance platforms, system interoperability is currently limited. Interoperability is a major challenge and opportunity for development to meet the needs of global supply chain requirements, in order to facilitate the exchange of information between blockchain systems globally. As Hewett et al. [59] mention, this is crucial for systems to effectively track the transport of goods, globally across systems, with visibility for all parties involved (manufacturers, logistics, retail etc.). Khan et al. [24] note another example along these lines, with a UK firm who have partnered with developers to create a blockchain system that tracks halal livestock through the process, thus improving traceability and providing verification of halal meat.

4.7 Agriculture

Agriculture efficiencies can result from the use of blockchain for 'smart' agriculture. Distributed ledger technology and blockchain can be used to solve the issue of needing a dispersed single platform for the storage and sharing of agricultural data regarding climate and land that affects production. Transaction transparency can be used to tackle the issue of trust in food safety, providing a reliable source of safety information that cannot be tampered with once verified and added on the blockchain. This would assist in making issues traceable back to the root cause.

Furthermore, blockchain can assist the agricultural supply chain, recording real-time data and aid in balancing the transactions of farmers and markets, while helping supply meet demand. For example, as Dong and Fu [61] mention, blockchain is currently being used in China to aid collaboration across the industry. Scattered small scale farms are joining forces to share their production information, using the transparency and distribution of blockchain data, with Smart Contracts to automate management processes. This allows farmers to operate as a larger enterprise, improve precision, decision making, resolve management issues and share the provenance of products with the consumer.

5 Implementation in Modern Businesses

5.1 Discovery and Preparedness Assessment

Before designing, procuring, and implementing a blockchain solution, organisations must determine whether blockchain is a requirement and if the business is prepared to adopt it. They need to engage stakeholders to assess the suitability of blockchain, and detail any potential use cases in resolving problems that affect the business, while also identifying potential risks and assessing the proposed solution's Return on Investment (ROI). This stage also entails considering the solution's intrinsic qualities, and assessing if its implementation will help the organisation with cost savings and risk reduction (Fig. 7).

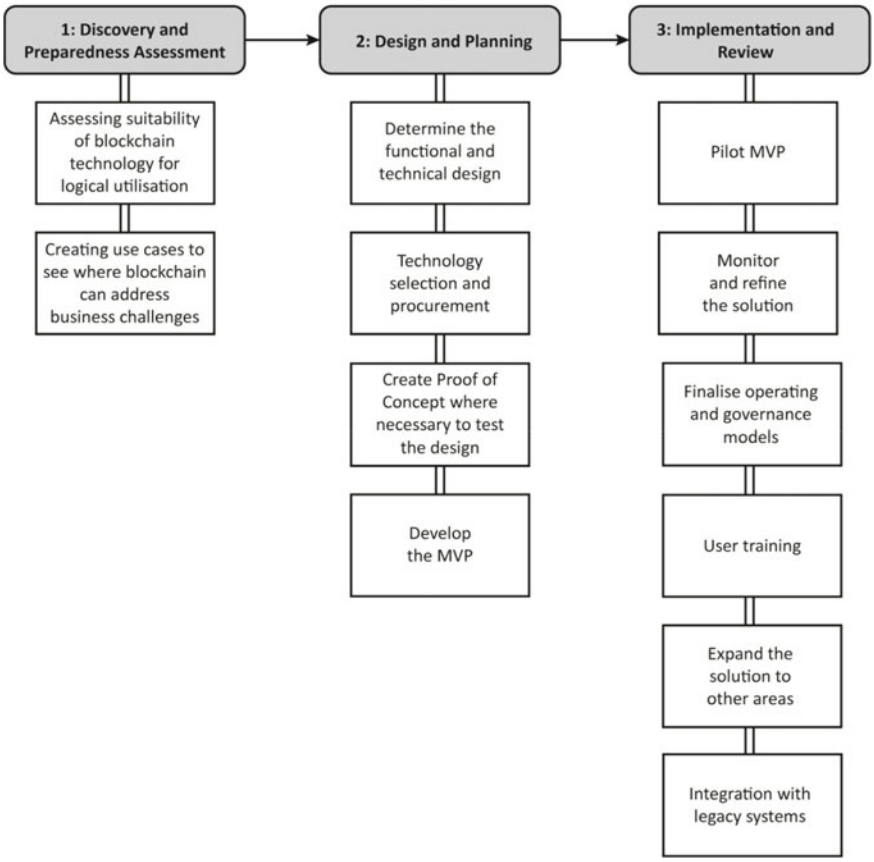


Fig. 7 Blockchain implementation roadmap

Once suitability has been determined, a Project Management Office (PMO) should establish a project team to organise the project and drive its phases forward, while monitoring milestones and risks, and coordinating all parties involved. Having the right team and skill sets is essential to a successful implementation, and to this end blockchain Subject Matter Experts (SMEs) should be engaged from the early stages, so that they can provide knowledgeable recommendations. Product and Business Owners need to monitor the development progress to ensure the solution is meeting their requirements, while existing enterprise architects and other technical staff should also be consulted and trained to address any skills gaps on implementation and maintaining the system. At the implementation stage, operational users need to be trained to interact with the technology, as ACT-IAC [62] suggests.

Organisations should also consider the adoption rate in the market they operate in, how this might impact their adoption strategy and if a blockchain solution is in line with their long-term business strategy. They need to evaluate their current technological environment and assess the complexity of integrating a blockchain solution in their ecosystem (including involving any participating partners), whether it will be an addition to existing systems, or whether it will fully or partially replace them. This evaluation also needs to consider the resulting impact on existing service, as well as the overall impact on the wider business and its processes, and what new policies and governance controls would need to be put in place. Governance and legal responsibilities need to be considered, and how they are affected if the solution needs to span multiple organisations in a consortium. Furthermore, the level of shared responsibility and trust that will be afforded to each involved party needs to be determined [63].

Figure 8 shows an example process flow for determining if there is a strong reason for an enterprise to adopt blockchain. As Knott [64] proposes, blockchain is most beneficial when a problem requires a decentralised, distributed solution that can transparently, securely, and immutably account for the history of assets and transactions. Blockchain will ultimately provide a permanent record that is easily audited and can prove the provenance of an asset; as a solution it is particularly suited to securing information between trusted parties, while allowing multiple parties to maintain a single source of truth for data sharing without an intermediary.

At this discovery stage, detailed technical planning is not required, but the higher-level technical aspects of blockchain should be considered to help assess its suitability and potential risks. By outlining the use cases in their organisation, the decision makers can consider the technical specifications of blockchain and its suitability. They should consider the type of blockchain, how accessible the data will be and if data is to be stored on or off the blockchain. What data validation requirements are there for adding data to the blockchain and if a 3rd party or notary will need to be involved to ensure data integrity. The size of the network, distribution and number of nodes should also be considered and how this will influence which consensus mechanism to adopt. Furthermore, the technical performance and scalability requirements for the solution need to be examined at this stage, such as transaction speeds and the amount of data to be processed.

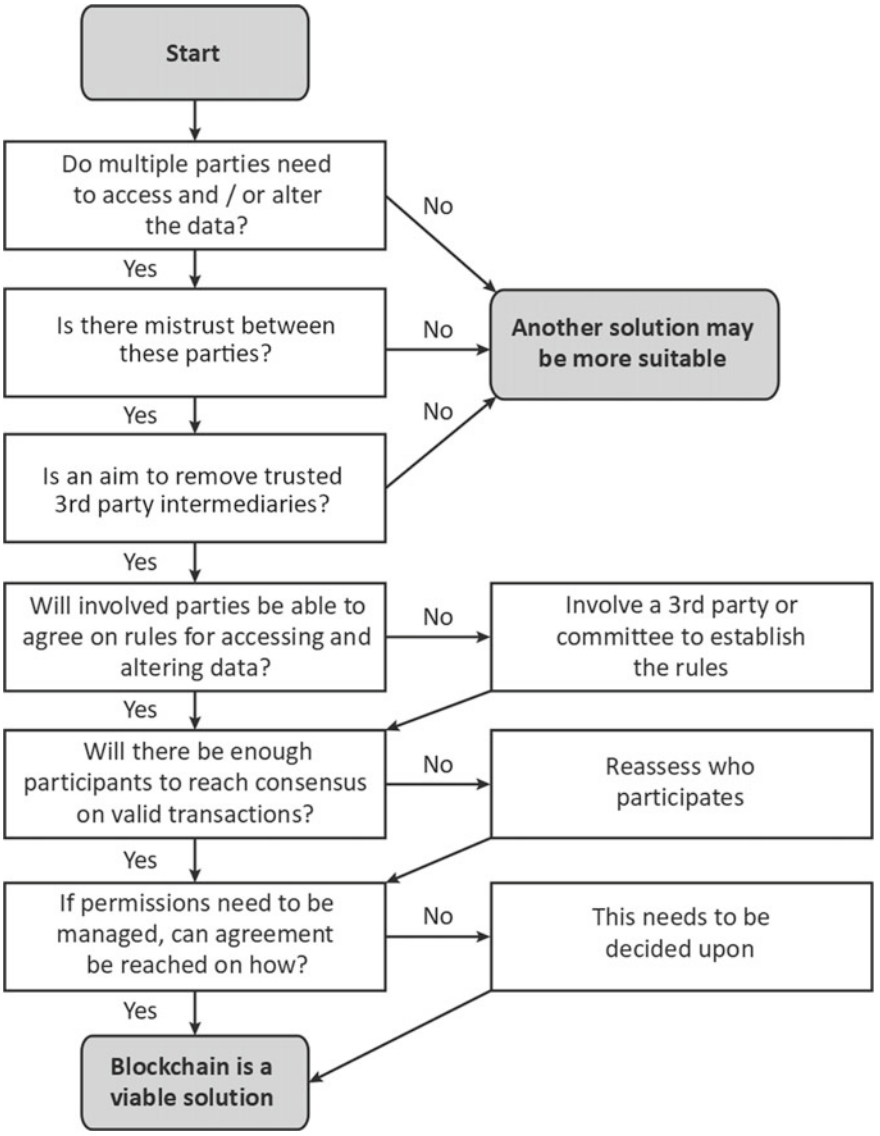


Fig. 8 Assessing blockchain's viability. Adapted from: Knott [64]

5.2 Design and Planning

Planning a Minimal Viable Product (MVP) is the recommended way to trial the blockchain solution on a smaller scale, within an area of the organisation that has been assessed for its suitability with a strong use case. Overall, a MVP can quickly

demonstrate the ROI a blockchain solution could bring, by testing it on a smaller scale and allowing for improvements while minimising turmoil for the business. It is also important however, to take a long-range approach when planning out future expansion and integrations with legacy systems and new technology [62]. This can be accomplished by adopting a phased approach, addressing different areas and needs within the organisation, and using a modular framework for blockchain projects that can be easily adapted for future developments as new standards emerge. Keeping the possibility of future interoperability with other systems in mind is beneficial. On top of this, as Hargrave and Karnoupakis [20] suggest, it is also important to consider the user experience when designing the solution, endeavouring to create a user focused design with simple interfaces for enthusiastic uptake and ease of adoption.

The planning phase should not only plan the technical details, but also map out the management of the solution's implementation, ongoing development, and maintenance. A change management plan should be of high priority, as it shapes the solution while users trial it and later use it, impacting the user engagement, adoption and ROI. It should also be adapted in consideration of specific blockchain qualities like immutability and automated Smart Contract execution. As Mpinyuri [65] suggests, a cross-functional team should also be appointed to deal with legal, compliance, and governance challenges (like privacy and jurisdictional responsibility), to work across all organisations involved with the solution. Other considerations during this phase include who is responsible for each aspect of the system, how the system will be governed, what the risks are, and what controls will be in place to mitigate them.

Before procurement, the blockchain technology specifications should be determined by assessing the business and technical requirements for the solution, considering the use cases identified in the discovery phase. Basic blockchain elements should be reviewed; for example, how transactions are performed, validated, and change the system. Furthermore, as ACT-IAC [62] states, the block composition and encryption mechanisms, the type of consensus and how any anomalies are handled, and how data is communicated between nodes should be examined. Among the technical considerations, it should be determined whether the network will be permissioned (private or consortium), permissionless (public) or a hybrid. Other standard operating requirements include the infrastructure platform (private, Software-As-A-Service/SaaS, Blockchain-As-A-Service/BaaS etc.), node location, user interfaces, reliability, and maintenance. Bashir [12] also points out that the organisation could use BaaS where blockchain services are managed externally and provide a platform ready for organisations to build their own Distributed Applications (DApps) onto. If an existing blockchain platform will be used, like Ethereum, any costs incurred (e.g., 'gas') will need to be taken into account [62]. In addition, choosing the consensus mechanism that best meets the performance requirements of the solution is very important, since it can impact transaction volumes, throughput and scalability. Moreover, as Staples et al. [22] note, another consideration is tied to whether all data will be stored on the chain (which can be more secure) or whether there will be use of indicators that point to a database located elsewhere (which can lead to better performance).

As a final consideration in this phase, the security specifications of the blockchain itself concerning encryption, consensus and data transport need to be defined. Organisations should research the different consensus methods (Proof of Work, Proof of Stake etc.) in the context of their network design, analysing the security risks related to each, in order to balance security with performance needs without compromising the system. Alongside that, user access must be considered (digital signatures, key management, authorisation, and authentication) to help preserve data integrity and mitigate unauthorised access, especially since this can be a major area of vulnerability with blockchain solutions. As Maleh et al. [19] note, any integrations with other systems, or applications, like Smart Contracts, will also need to be scrutinised for security risks.

5.3 Implementation and Review

In this phase, the MVP is delivered, and the design, operational and governance models are refined. If the MVP has then proven to be successful and beneficial to the business, the solution is expanded and integrated with other systems as required. By this implementation stage, the organisation should confirm the solution design and other elements with a clear project plan which schedules activities, resources, budgets, and risks. This will enable the implementation of the selected deployment model, in order to create the core blockchain solution. It is recommended that the project goals and scope are finalised to identify the functional requirements that are being addressed, alongside the technical design and security controls, including technical details relating to the selected architecture and governance rules. Nevertheless, the project would benefit from taking an Agile and iterative approach to testing and quickly reworking any part of the solution that is not effective. As Welfare [66] suggests, it is important to recognise that the implementation is a continuous process of improvement, which will most likely require much revision.

The initial planning regarding the solution participants needs to be refined during the implementation phase and altered, particularly if more organisations join the project during this phase. Deciding who is authorised to access specific data and how they are authenticated to access the system is also of utmost importance, and the onboarding and off-boarding strategy for participating organisations should be refined, to ensure verification of entrants and mitigate security issues. Overall, security should be regularly monitored to ensure any vulnerabilities are detected early. Additionally, operational processes must also be defined, together with adequate training plans for all participants to support them in using the technology.

Another activity at the implementation stage is integrating the blockchain with other applications and databases. Extensive testing should be performed with any integrated systems that could experience negative consequences from the blockchain implementation or vice versa. As Koteska et al. [67] underline, these components should be identified, and a test strategy should be formulated for testing with the original use-cases and functional requirements in mind.

Expanding the solution beyond the initial MVP presents its own challenges, especially when increasing the number of parties involved. Consensus on governance matters should be regularly reviewed by an appointed steering committee and adapted, if necessary, as the network of collaborating organisations grows [66]. Deloitte [68] also recommends planning the wider roll-out and any technical requirement changes that come with operating at scale. This includes slowly increasing the number of participants on the Minimum Viable Ecosystem (MVE) by one partner at a time, with ongoing improvements made as size increases. Finally, a legal operational model is essential, with policies and guidelines on governance, including matters like technical updates, ownership rights, onboarding, legal requirements etc. Ultimately, clear governance principles which define the participants' roles and responsibilities are important for the success of a wider roll-out between multiple parties.

6 Future Trends

The evolution of DLTs and the increasing use of blockchain have increased the future potential for digital currencies and the transactions that take place in the digital environment. Future developments and trends in the context of Smart Contracts include research in the direction of formal verification, Smart Contract-based organisational management and the Layer 2 of Smart Contracts [69]. As Fox et al. [70] state, it is identified that Smart Contracts with formal verification will increase the confidence associated with the technology and overall user acceptance.

On the other hand, the rapid developments in technology associated with the Internet of Things and Artificial Intelligence, will lead to changes in the infrastructure associated with implementation of Smart Contracts. As Tarr [71] notes, there will be improvements in the conceptual frameworks and the fundamental theories associated with the use of blockchain and Smart Contracts due to the influence of AI, as well as the changing dynamics of organisational management. Due to their programmatic nature, Smart Contracts also bear the potential of transforming DLT systems and applications in the future, leading to the emergence of more complex blockchain architectures.

As Farahaniet et al. [72] suggest, DLT systems are expected to aid the development of governance frameworks in the future, which would be functional without the compromise of freedom and independence, as acquired via their decentralised attributes. Additionally, McCorry et al. [73] observe that future developments in blockchain would help in increasing performance and speed, and improving interoperability over multiple platforms. Finally, the implementations and research conducted in the context of DLT and blockchain, would aid in increasing the quality of system audits and reduce the expectation gap that exists between regulatory bodies, auditors and end users.

7 Conclusion

As discussed, the potential of blockchain and Smart Contract implementations is huge for a large number of industries, although they still emerge from their early development stages. The gradually increased adoption of blockchain projects as enterprise-level solutions have put the technology to the test, and will demonstrate its efficiency over the coming years. Furthermore, the integration of Smart Contracts in blockchain applications has enabled the technology to reach new levels of customisation and complexity, adapting to bespoke business needs, and meeting the elaborate demands of today's fast-paced industry.

At the end of the day, the ability to automate, secure and fortify the way transactions work in the online world is starting to be regarded as a necessity rather than an idealistic scenario. This is especially urgent in a digital environment where fraud, mistrust and lack of traceability have evolved into serious issues affecting individuals, organisations and governments on a daily basis.

References

1. Puthal D, Mohanty SP, Kougianos E, Das G (2020) When do we need the blockchain? IEEE Consum Electron Mag 10(2):53–56
2. Tasatanattakool P, Techapanupreeda C (2018) Blockchain: challenges and applications. In: 2018 international conference on information networking (ICOIN), IEEE, pp 473–475
3. Carson B, Romanelli G, Walsh P, Zhumaev A (2018) Blockchain beyond the hype: What is the strategic business value. McKinsey & Company 1–13
4. Chaum DL (1979) Computer systems established, maintained and trusted by mutually suspicious groups. University of California, Electronics Research Laboratory
5. Haber S, Stornetta WS (1990) How to time-stamp a digital document. In: Conference on the theory and application of cryptography. Springer, Berlin, pp 437–455
6. Szabo N (1997) Formalizing and securing relationships on public networks. First Monday 2(9)
7. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. [Online] Available at: https://www.debr.io/article/21260.pdf. [Accessed 31 July 2021]
8. Buterin V (2014) A next-generation smart contract and decentralized application platform. Ethereum White Paper 3(37)
9. Kolb J, AbdelBaky M, Katz RH, Culler DE (2020) Core concepts, challenges, and future directions in blockchain: a centralized tutorial. ACM Comput Surv (CSUR) 53(1):1–39
10. Crosby M, Pattanayak P, Verma S, Kalyanaraman V (2016) Blockchain technology: Beyond bitcoin. Appl Innovation 2(6–10):71
11. Shibata N (2019) Proof-of-search: combining blockchain consensus formation with solving optimization problems. IEEE Access 7:172994–173006
12. Bashir I (2020) Mastering Blockchain—third edition. 3rd ed. s.l.:Packt Publishing
13. Chicarino V, Albuquerque C, Jesus E, Rocha A (2020) On the detection of selfish mining and stalker attacks in blockchain networks. Annal Telecommun, pp 1–10
14. Ren W et al (2020) A flexible method to defend against computationally resourceful miners in blockchain proof of work. Inf Sci 507:161–171
15. Lantz L, Cawrey D (2020) Mastering blockchain. s.l.: O'Reilly Media, Inc
16. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: Architecture, consensus, and future trends. In: 2017 IEEE international congress on big data (BigData congress). IEEE, pp 557–564

17. Raj K (2019) Foundations of blockchain. Packt Publishing, s.l.
18. Bitcoin (2021) Block chain. [Online] Available at: https://developer.bitcoin.org/reference/block_chain.html [Accessed 31 July 2021]
19. Maleh Y, Shojafar M, Alazab M, Romdhani I (2020) Blockchain for cybersecurity and privacy: architectures, challenges, and applications. Taylor & Francis Group, ProQuest Ebook Central, s.l.
20. Hargrave SJ, Karnoupakis E (2019) What Is Blockchain?, s.l.: O'Reilly Media, Inc. As seen on: https://www.oreilly.com/library/view/what-is-blockchain/9781098114749/
21. Taskinsoy J (2019) Blockchain: a misunderstood digital revolution. Things you need to know about blockchain. SSRN Electron J
22. Staples M et al. (2017) Risks and opportunities for systems using blockchain and smart contracts. Data61 (CSIRO), Sydney
23. Ethereum (2021) Dapps [Online] Available at: https://ethereum.org/en/dapps/. [Accessed 19 Aug 2021]
24. Khan SN, Loukil F, Ghedira-Guegan C, Benkhelifa E, Bani-Hani A (2021) Blockchain smart contracts: Applications, challenges, and future trends. Peer- to-peer Networking Appl 1–25
25. Petrov D (2020) Blockchain Ecosystem in the Financial Services Industry. FAIMA Bus Manage J 8(1):19–31
26. Levi SD, Lipton AB (2018) An introduction to smart contracts and their potential and inherent limitations. [Online] Available at: https://corpgov.law.harvard.edu/2018/05/26/an-introduction-to-smart-contracts-and-their-potential-and-inherent-limitations. [Accessed 19 Sept 2021]
27. Kannengießer N, Lins S, Dehling T, Sunyaev A (2019) What does not fit can be made to fit! Trade-offs in distributed ledger technology designs. In: Proceedings of the 52nd Hawaii international conference on system sciences
28. Law Commission (2020) Smart contracts. [Online] Available at: https://www.lawcom.gov.uk/project/smart-contracts/. [Accessed 19 Sept 2021]
29. Bayer D, Haber S, Stornetta WS (1992) Improving the efficiency and reliability of digital timestamping. In: Capocelli R, De Santis A, Vaccaro U (eds) Sequences II. Springer, New York
30. Natarajan H, Krause S, Gradstein H (2017) Distributed ledger technology and blockchain. World Bank Group
31. Pinna A, Ruttenberg W (2016) Distributed ledger technologies in securities post-trading revolution or evolution?. ECB Occasional Paper No. 172. Available at: https://www.ecb.europa.eu/pub/pdf/scpops/ecbop172.en.pdf
32. Rozario AM, Thomas C (2019) Reengineering the audit with blockchain and smart contracts. J Emerg Technol Account 16(1):21–35
33. Hewa TM, Hu Y, Liyanage M, Kanhare S, Ylianttila M (2021) Survey on blockchain based smart contracts: technical aspects and future research. IEEE Access
34. Schulz KA, Gstrein OJ, Zwitter AJ (2020) Exploring the governance and implementation of sustainable development initiatives through blockchain technology. Futures 122:102611
35. Peters GW, Panayi E (2016) Understanding modern banking ledgers through blockchain technologies: future of transaction processing and smart contracts on the internet of money. In: Banking beyond banks and money. Springer, Cham, pp 239–278
36. Alharby M, Aldweesh A, van Moorsel A (2018) Blockchain-based smart contracts: A systematic mapping study of academic research (2018). In: 2018 International conference on cloud computing, big data and blockchain (ICCBB). IEEE, pp 1–6
37. Chowdhury MJM, Ferdous MS, Biswas K, Chowdhury N, Kayes ASM, Alazab M, Watters P (2019) A comparative analysis of distributed ledger technology platforms. IEEE Access 7:167930–167943
38. Kannengießer N, Lins S, Dehling T, Sunyaev A (2020) Trade-offs between distributed ledger technology characteristics. ACM Comput Surv (CSUR) 53(2):1–37
39. Rauchs M (2022) Ep. 71 – DLT Conceptual Framework - Insureblocks. [Online] Available at: https://insureblocks.com/ep-71-dlt-conceptualframework/. [Accessed 21 Mar 2022]

40. Cong LW, He Z (2018) Blockchain disruption and smart contracts. NBER working paper series. Working paper 24399. Available at: http://www.nber.org/papers/w24399.pdf
41. Ølnes S, Ubacht J, Janssen M (2017) Blockchain in government: Benefits and implications of distributed ledger technology for information sharing. Gov Inf Q 34:355–364
42. Okada H, Yamasaki S, Bracamonte V (2017) Proposed classification of blockchains based on authority and incentive dimensions. In: 2017 19th international conference on advanced communication technology (icact). IEEE, pp 593–597
43. Badr NG (2019) Blockchain or distributed ledger technology what is in it for the healthcare industry? In: KMIS, pp 277–284
44. Kuo TT, Kim HE, Ohno-Machado L (2017) Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 24(6):1211–1220
45. Janowicz K, Regalia B, Hitzler P, Mai G, Delbecque S, Fröhlich M, Martinent P, Lazarus T (2018) On the prospects of blockchain and distributed ledger technologies for open science and academic publishing. Semant web 9(5):545–555
46. Wang S, Ouyang L, Yuan Y, Ni X, Han X, Wang FY (2019) Blockchain—enabled smart contracts: architecture, applications, and future trends. IEEE Trans Syst Man Cybern Syst 49(11):2266–2277
47. Rauchs M, Glidden A, Gordon B, Pieters GC, Recanatini M, Rostand F, Vagneur K, Zhang BZ (2018) Distributed ledger technology systems: a conceptual framework. Available at SSRN 3230013
48. Deshpande A, Stewart K, Lepetit L, Gunashekar S (2017) Distributed ledger technologies/blockchain: challenges, opportunities and the prospects for standards. Overview Rep Brit Stand Inst (BSI) 40:40
49. Benedict G (2019) Challenges of DLT-enabled scalable governance and the role of standards. J ICT Standard 195–208
50. Ethereum (2021) Non-fungible tokens (NFT). [Online] Available at: https://ethereum.org/en/nft/ [Accessed 12 09 2021]
51. Leshner R, Hayes G (2019) Introduction. [Online]. Available at: https://compound.finance/documents/Compound.Whitepaper.pdf. [Accessed 19 Sept 2021]
52. Stellar (2021) Intro to stellar. [Online] Available at: https://www.stellar.org/learn/intro-to-stellar. [Accessed 19 Sept 2021]
53. Roy I (2020) Blockchain Development for Finance Projects. Packt Publishing, s.l.
54. FATF (2021) Opportunities and challenges of new technologies for AML/CFT, s.l.: FATF
55. Consensys (2021) CODEFI COMPLIANCE: AML-CFT compliance for ethereum—Powered digital assets. [Online]. Available at: https://consensys.net/codefi/compliance/. [Accessed 22 Aug 2021]
56. Quzmar A, Qatawneh M, Al-Maaitah S (2021) Reducing counterfeit drugs with blockchains: A survey. s.l., 2021 Int Conf Inf Technol (ICIT), pp 143–148
57. Wilson T (2021) British hospitals use blockchain to track COVID-19 vaccines. [Online] Available at: https://www.reuters.com/article/uk-health-coronavirus- blockchain-idUSKBN29O0RW. [Accessed 11 Sept 2021]
58. Ricci L, Maesa DDF, Favenza A, Ferro E (2021) Blockchains for COVID-19 contact tracing and vaccine support: a systematic review. IEEE Access 9:37936–37950
59. Hewett N, van Gogh M, Palinczki L (2020) Inclusive deployment of blockchain for supply chains: Part 6–A framework for blockchain interoperability. In: Geneva, Switzerland: world economic forum
60. Warren S, Wolff C, Hewett N (2019) Inclusive deployment of blockchain for supply chains: Part 1—Introduction. World Economic Forum, Geneva
61. Dong N, Fu J (2021) Development path of smart agriculture based on blockchain. s.l., In: 2021 IEEE Asia-pacific conference on image processing, electronics and computers (IPEC), pp 208–211
62. ACT-IAC (2021) Blockchain playbook. [Online] Available at: https://blockchain-working-group.github.io/blockchain-playbook [Accessed 12 10 2021]
63. PWC (2016) Blockchain: key questions for your business. PWC, s.l.

64. Knott N (2019) Is your business ready for blockchain? [Online] Available at: https://bankingblog.accenture.com/does-your-fs-business-need-blockchain-how-to-get-started. [Accessed 13 Oct 2021]
65. Mpinyuri EB (2019) Beyond cryptocurrencies: financial applications of blockchain technology, University of Johannesburg, s.l.
66. Welfare A (2019) *Commercializing Blockchain.* s.l.: Wiley
67. Koteska B, Karafiloski E, Mishev A (2017) Blockchain implementation quality challenges: a literature review. Belgrade, SQAMIA 2017: 6th workshop of software quality, analysis, monitoring, improvement, and applications
68. Deloitte (2019) Business blockchains, s.l.: Deloitte
69. Byström H (2019) Blockchains, real-time accounting, and the future of credit risk modeling. Ledger 4
70. Fox MB, Glosten LR, Greene EF, Guan SS (2021) Distributed ledger technology and the securities markets of the future: a stakeholder survey. Columbia Bus Law Rev 2021(2)
71. Tarr JA (2018) Distributed ledger technology, blockchain and insurance: opportunities, risks and challenges. Insur Law J 29(3):254–268
72. Farahani B, Firouzi F, Luecking M (2021) The convergence of IoT and distributed ledger technologies (DLT): Opportunities, challenges, and solutions. J Netw Comput Appl 177:102936
73. McCorry P, Shahandashti SF, Hao F (2017) A smart contract for boardroom voting with maximum voter privacy. In: International conference on financial cryptography and data security. Springer, Cham, pp 357–375

An Investigation into How Smartphones Can Be Secured Against MiTM Attacks: Financial Sector

David Steiner-Otoo and Hamid Jahankhani

Abstract MiTM attack aims to violate data in transmission through the air medium in a wireless network; MITM exploits compromise data confidentiality and integrity and are conceivably the most productive types of cyberattacks utilised today. The increasing use of personal devices like smartphones connecting to the internet via Wi-Fi has made wireless attacks on users more crucial. The cyber adversary becomes a "middleman" between two targets to intercept private communication, decrypt traffic, and exploit valuable information like bank details and credit cards. The new WPA3 protocol security features such as 256-bit encryption, OWE (Opportunistic Wireless Encryption), Simultaneous Authentication of Equals (SAE), and disallowing outdated legacy protocols provides risk mitigation against attacks. However, vulnerabilities in WPA3 have been reported whereby a device can be downgraded from WPA3 to WPA2, which opens the system up for DoS and MiTM attacks. This research investigates Wi-Fi-based exploits against the ecosystem of smartphones in the financial sector. Aircrack-ng and Ettercap are open-source tools accessible through the Kali Linux framework. These tools are utilised to demonstrate simulated DoS and MiTM attacks to explain the reported WPA3 vulnerabilities.

Keywords 802.11i · Blockchain · COVID-19 · Encryption · WPA2 · WPA3 · Kali linux · Penetration testing · Self-sovereign identity · Wi-Fi · Wi-Fi 6 · Wi-Fi 7 · WLAN

1 Introduction

Wi-Fi stands for Wireless Fidelity with the generic name of IEEE 802.11, and suffixes are added to represent improved versions of Wi-Fi. Recently launched Wi-Fi 6 (and Wi-Fi 6E) or 802.11ax is the current release, and Wi-Fi 7 or labelled 802.11be is

D. Steiner-Otoo
Northumbria University London, London, UK

H. Jahankhani (✉)
Northumbria University London Campus, London, UK
e-mail: Hamid.jahankhani@northumbria.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_7

projected to be released in 2024 [1]. Wi-Fi is a necessity today and in finance, and the technology extends to robotics, the Internet of Things—IoT, and the Internet of Everything—IoE [2]. The global economic value generated using Wi-Fi is estimated to grow from $3.3 trillion to $4.9 trillion in 2025, while mobile data is anticipated to rise from approximately 51 billion Gigabytes in 2020 to 226 billion Gigabytes in 2026 on a month-to-month basis [3]. However, there is a growing concern about wireless security vulnerabilities.

Man-in-The-Middle or MiTM (sometimes abbreviated in the literature as MIM, MiM, MITMA) attacks have different names such as machine-in-the-middle [4], monkey-in-the-middle [5], person-in-the-middle [6], or Man-in-the-PC/Phone (MITPC/Phone attack [7]. MiTM exploits aim to violate confidentiality and integrity of data in transmission through a wireless (Wi-Fi network, mitm exploits are conceivably the most productive types of cyberattacks utilised today [8].

The importance of smartphones in daily life, including financial transactions, social media and culture, cannot be understated. Smartphone usage permeates and impacts every demography in Great Britain and globally. However, users have limited knowledge in mitigating risk against hackers, while application developers do not always consider and implement the appropriate security checks during development.

2 Literature Review

The genesis of banking and financial institutions can be traced as far back to ancient civilisation in Kemet (in northern Africa, present-day Egypt) before 4000 BCE, which has some of the "oldest recorded civilisation" that in turn influenced the advancement of later societies and cultures in ancient Asia, Greece, and the Roman Empire [9]. In evaluating the evolution of money and monetary institutions, religion and finance have a direct correlation and significance, the early banks started "in the temples consecrated to the ancient gods" [10]. As was in ancient civilisations, and followed by the Romans, the religious temples such as the temple of Jerusalem and Apollo at Delphi, worship edifices functioned as the first banks or financial institutions (Innes 1913, cited in [11]. Labate [10] describes the early temples as the initial repositories (i.e., banks) where money and treasures of wealthy Romans were deposited in the basements of numerous temples. The temples were involved in banking activities like lending based on their good names and reputation, the priests acted as modern-day banking officers who monitored deposits and loans. The temples were secure because the buildings were regularly inhabited by faithful worshipers and ecclesiastics and constantly guarded by soldiers [10]. This can be analogised as the equivalent of modern-day financial institutions' cybersecurity tools and techniques to secure against theft, financial loss and data attacks. In essence, the priests acted as the temple/bank's Chief Financial Officer (CFO) and Chief Technology Officer (CTO), the patrolling soldiers were the firewalls and intrusion detections systems, while the devout worshipers unknowingly acted as the early form of threat intelligence gatherers—all being risk mitigations against attacks.

Financial institutions have evolved over the centuries from the traditional to contemporary FINTECHs, together with technological advances. Data, an intangible commodity, comes into the equation, so security becomes imperative and more challenging to achieve nearly 100%. At the core of most digital transactions is the reliance on protection to mitigate against cyberattacks such as man-in-the-middle exploits. Thakor [12] describes Fintech as "the use of technology to provide new and improved financial services". This includes innovations in payment services in cryptocurrencies and the role played by Blockchain-assisted intelligent contracts. The goal of financial innovations integrated with technological advances is to lower financial services costs or risks, improve digital security for the consumer, and improve social welfare [13]. The most significant disruption and innovation by Fintech are with cryptocurrency payment systems like Bitcoin, which are digital and virtual currencies stored in electronic/digital wallets in cyberspace that allows peer-to-peer transactions independent of traditional financial banks. Cryptocurrency transactions rely on decentralised control, security and verification methods based on cryptography-based distributed digital ledger technology, the Blockchain, that supplants the conventional banks [12].

2.1 MiTM Attacks

Man-in-The-Middle exploits is a significant security concern whereby threat actors target data in transmission between two legitimate endpoints to compromise the data integrity and confidentiality [14]. The malicious third party can intercept, read, modify or control the communication traffic. MiTM attacks require a communication channel, the popularly used are radio frequency and Wi-Fi, Bluetooth, GSM (Global System for Mobiles), NFC or Near Field Communication [15]. Mobile devices are prone to such attacks [16] when in the process of securing connectivity with an access point or a server. The review of existing literature shows an abundance of research journals on MiTM attacks in healthcare services, transportation, and retail sectors, but a limited number of articles in the financial industry. Financial institutions hold a large quantum of sensitive data, when exposed, this can cause harm to the UK and global economic security and personal interests [17]. Financial institutions are compelled and legally obligated to report security and data breaches to the ICO [18] to satisfy relevant legislations—data protection regulations (GDPR) and the Data Protection Act (2018) [19]. According to the UK's Financial Conduct Authority, FCA [20], cyberattacks against banks in Britain have risen from five in 2014 to forty-nine in 2017. However, banks are reluctant to report such attacks for fear of bad publicity and punishment from regulators. According to Carnegie [21], in January 2021, American Express and the Reserve Bank of New Zealand suffered a cybersecurity attack resulting in a data breach, in March 2021, Wall Street was targeted in New Capital Call cyber fraud scheme, also in March 2021, the American insurance company CNA was hit by a cyberattack. The limited number of MiTM attack research papers in the financial sector is mainly because research experiments

must be conducted in laboratories, which are often not ideal environments and not representative of fully functioning financial institutions.

2.2 *Security Vulnerabilities and Attacks in Mobile Banking and Trading Apps*

A study by Zheng et al. [22] analysed security vulnerabilities in Android OS based mission-critical smartphone apps such as mobile trading and banking apps. The study examined application repackaging attacks whereby a legitimate Android app is reverse-engineered, malicious program codes inserted and rebuilt as a new application. The study found that ineffective security mitigation measures were the main reasons malicious repackaged apps are easily uploaded in Android markets like Google Play, Amazon Appstore, and other app markets. A report by Ciscomag [23] suggested that more than 50% of mobile banking apps were vulnerable to data theft and fraud because of "inadequate security layers". Android OS is an open-source model, making malicious tools and applications easier access to data and information on users' smartphone apps. The authors found that anti-malware tools use signature-based or static analysis methods which evade obfuscation, allowing hackers to adapt by using metamorphism and polymorphism to evade anti-malware countermeasures. Due to inadequate security, 76% of banking apps have vulnerabilities that can be hacked without accessing the physical device, and 33% can be attacked without having administrative privileges [24]. The authors proposed user education as an essential step to protect mobile banking apps against hackers. However, users are more interested in the app functionality and user experience (UI) and do not see themselves as security experts. Another attack prevention approach was using a trusted agency guaranteeing the developer identity and genuineness of the application by inserting "an assurance signature" into the application package so that users can make better-informed decisions when installing apps on their smartphones.

X-Force Exchange by IBM [25] is a cloud-based TI open-source platform that allows users to quickly research current global security threats, share and act on threat intelligence supported by human and ML generated intelligence. More mature and advanced threat intelligence tools are currently available on the market, such as Kaspersky [26] Lab, which collects data from worldwide sources to give in-depth insights into cyber threats targeting financial institutions and revealing potential evidence of cyberattacks [27]. Insights' External Threat Protection (ETP) analysed vulnerability that is "engineered to discover, examine and mitigate cyber risks" and patch critical vulnerabilities [28]. However, TI has limitations due to the overwhelming quantity of available data, the challenges security teams face in identifying the most relevant data, and difficulty making valuable use of them. In some instances, the available intelligence (i.e., data) is out of date. The timeliness of data is essential in understanding strategies, tactics, and motivation of threat actors to protect against intrusive attacks and zero-day exploits. According to OWASP [29], the top ten mobile

risks are improper platform usage, insecure data storage, insecure communication, insecure authentication, insufficient cryptography, insecure authorisation, client code quality, code tampering, reverse engineering, and extraneous functionality.

2.3 Android and iOS—Security Compromise Issues and Analysis

Authors Garg and Baliyan [30] conducted a study on Android and iOS to ascertain security vulnerabilities between the two OSs. This comparative qualitative study included an analysis of the security model, system architecture, encryption mechanism, and app permissions. It listed the most common flaws in both platforms and presented a vulnerabilities assessment of the two OSs. The journal discussed malware attacks on Android and iOS and suggested future research and app development to prevent growing cyberattacks on the platforms. MiTM, DoS/DDoS, SYN flooding attacks are the most common attacks. The authors collected data from CVEDetails, a security vulnerability data source containing listings of publicly reported computer security vulnerabilities and the severity of flaws [31]. However, the US Department of Homeland Security (DHS) and CISA (cisa.gov, n.d.) maintain and sponsor a separate computer vulnerabilities database known as NVD, which also allows searches by but not limited to OS, vulnerability type, product name, severity, and impact. Although both CVE and NVD databases are synchronised, the authors could have missed high other risk vulnerabilities reported in NVD but not available in CVE during their study, hence limiting the study's accuracy. The study showed that the overall number of vulnerabilities in both platforms decreased between 2017 and 2019 because of improved detection rates due to the use of ML and DL algorithms. However, there were 61% more vulnerabilities with Android compared to 39% of iOS platforms.

2.4 Cyberattacks During COVID-19 Pandemic

Cyberattacks during the SARS-CoV-2 or COVID-19 [32] pandemic period in 2020 saw a considerable surge in attacks against financial institutions, individuals, and organisations.

2.4.1 Analysis of Cyberattacks During Pandemic

Lallie et al. [33] used mixed methods to present a timeline of events and analysis into the SARS-CoV-2 pandemic in the context of cyberattacks and cybercrimes that has witnessed a massive surge [34] compared to previous periods. The authors highlight cyberattacks types and persistency experienced in the UK and worldwide from

the onset of the global pandemic in 2020, which witnessed increased cybersecurity challenges ever recorded by citizenry and industry [35]. Hiscox [36] contends that cybercrime is growing in severity and frequency primarily due to inherent risks in centralised identity systems. This review focuses on the different types of cyber-attacks occurring during the pandemic and the impact on people. However, the review will not dwell on the timeline aspects of the attacks due to the unreliability of the timelines, limitations, and inaccuracies because the URLs on which they were reported could have been updated multiple times. The analysis of the cyberattacks was examined in the context of global and UK specific events and attacks to show how threat actors had developed advanced and sophisticated modii Operandi in the cyberattack offensive during the pandemic. During the rapid spread of COVID-19 worldwide in 2020, a significant increase in cyberattacks and cybercrime campaigns was perpetuated in the technology-driven society. Some attacks were indiscriminate, and others targeted. Coronavirus-themed scams that impersonated public authorities, fraud—especially financial fraud, and offering COVID-19 cures were reported. The COVID-19 pandemic cybercrime landscape included DDoS, ransomware, phishing, data harvesting malware like banking Trojans, and malicious domains. Statista3 [37] reported that the malware "Dridex" was the most prevalent banking trojan accounting for 26% of trojans during 2020.

The authors found that the most significant cybersecurity scams targeted the public at large, and millions of ordinary individuals were forced to work from home and suffer from raised anxiety and stress levels, and financial worries. Cybercriminals exploited the people's fears and uncertainty that had come about due to the "unstable social and economic situation" because of COVID-19 [38]. Furthermore, experiences of people working en-masse from home revealed the lack of preparation by both software vendors in terms of their product security. Organisations rapidly deployed remote networks and systems, enabling staff to perform tasks from home without the necessary attention to security vulnerabilities when VPNs could have been deployed, for example. January to April 2020 saw 907,000 spam messages, 737 malware exploits, and 48,000 malicious URLs related to COVID-19 reported by just one Interpol private sector partner. However, Interpol [38] contends that the most significant shift in cybercrime from small businesses and individuals has been an attack on critical national infrastructures like healthcare services [39], requiring new levels of oversight and security [40]. Unlike traditional attacks, advanced persistent threat (APT) groups build highly customised malware that is very targeted to increase the chances of success and achieve maximum impact [41], which were responsible for major critical infrastructure cyber exploits.

The UK NCSC, USA NSA, and Canada's CSE attributed the APT APT29, also called Cozy Bear or the Dukes, to the Russian Intelligence Services cyber espionage group, which targeted COVID-19 vaccine developments [42]. Lallie et al. use the UK's CPS categorisation of cybercrime guidelines, categorising cybercrime into two categories, namely, "cyber-dependent" and "cyber-enabled" crimes Cyber-enabled crimes include financial fraud, phishing, pharming, and extortion. In contrast, a cyber-dependent crime includes denial of service, hacking, and malware [43]. Definitions

of cyber-enabled and cyber-dependent crimes, including cybersecurity by default, are provided in the footnote.

In taking the UK as a case study to analyse the pandemic related cyber-crimes, the authors demonstrated direct correlations, meaning the association between news and policy announcements (such as the UK government hardship fund announcement in 24/03/2020 supporting the citizenry and economy) and associated cybercrime campaigns. The authors reported that by the 7th of May 2020, an excess of 160,000 suspect emails was reported to the NCSC [44], and £4.6 m was lost to coronavirus related scams affecting 11,260 victims of smishing or phishing campaigns. The 43 different types of cyberattacks investigated were categorised,86% involved phishing/smishing attacks; malware accounted for 65%; financial fraud was 34%; extortion was 15%, and pharming accounted for 13%.

COVID-19 and related cybercrimes impacted individuals' data and assets, the workforce. It presented challenges to information governance and regulatory compliance, social-economic structures, and how people communicated and lived[1].

Securing the individual's personal and sensitive information became a severe problem, such as the theft of a person's digital identity through the hacking and unauthorised access to PII or personally identifiable information (including name, national insurance number, and credit card details) via MiTM exploits, data breaches, and identity theft. According to the GDPR law (ICO n.d.), personal data breaches include unauthorised access to personal data transmitted. User's sensitive digital identity and information reside with service providers and centralised systems, and in most cases, users lack control over their digital identity and data flow [45]. The use of AI technology solutions and Self-Sovereign Identity (SSI) identity management system (IDM) offers a decentralised digital identity approach, a better security solution, and is more likely to enable the user to take back control of their digital identity and footprint. This reduces the risk of data breaches during data in transmission MiTM attacks while not depending on one trusted third party or external sources.

2.4.2 Cybersecurity Attack Vectors, Methods and Technics During Pandemic

Susukailo et al. [46] use qualitative analysis to describe cyberattack vectors, methods, and technics deployed by hackers during the global pandemic in 2020. It identifies the most frequent targets for hackers and the tactics used during cyberattacks. The authors review the cyber security challenges, possible countermeasures to improve the security situation, and cyber security controls to mitigate risk against the attack vectors analysed. The author contends that financial gain (arising from the COVID-19 financial crises) is the ultimate motivation of hackers, which is a necessary aspect

[1] Cyber-dependent crime is an offence, "*that can only be committed using a computer, computer networks or other form of information communications technology (ICT)*". Cyber-enabled crimes *are, "traditional crimes, which can be increased in their scale or reach by use of computers, computer networks or other forms of information communications technology (ICT)" (McGuire and Dowling 2013).*

of the attack vectors. Attack vectors were categorised into three groups: (i) social engineering attacks, (ii) interruption of critical business functions attacks, and (iii) critical infrastructure attacks. This review looks at the aspects of vulnerabilities involving social engineering. Social engineering exploits were the most prevalent attack vector, which preyed on an individual's compassion or fear and the need to find information online or in newspapers to protect themselves against the coronavirus. Hackers created numerous fake charitable websites deceiving people to earn money or infecting their computers with malware. Azourlt (also known as Puff-Stealer and Rultazo) was a common and popular stealer-type malware, used fake coronavirus phishing emails and online maps to steal the victims banking information, including credit card details and passwords, as well as cryptocurrency [47]. The Azourlt malware's delivery mode was through MS Office document, made simple by the hackers manipulating people's fears.

By opening the file, the malware exploited the CVE-2017–11882 MS Office Equation Editor vulnerability to download the malicious executable; the malicious executable file proceeds to make computer registry changes to run when the system starts. At system startup, the malware launches itself and steals personal data, and the malware deletes itself after a 3-s timeout. Social engineering techniques included the creation of fake online shops to sell COVID-19 related medical supplies and medicines with the sole purpose of attackers being financial gain. Susukailo et al. argue that the primary control to apply to deter social engineering attacks is information assurance strategies such as end-user training or awareness sessions with examples containing fake pandemic online resources. A secondary control against such attacks argued by the authors is the enablement of email malware scannings and phishing detection modules such as those found in Office 365 [48], G Suite (Google.com n.d.), and VirusTotal (Virustotal.com n.d.) browser extension.

2.5 Blockchain Technology

Fartitchou et al. [49] define blockchain (BC) as "a decentralised distributed database technology secured by means of cryptographic algorithms", the append-only ledger database cannot be altered. The BC works in a P2P (Peer-to-Peer) system, with each node in the blockchain system having a duplicate of the blockchain. Additionally, records of transactions and timestamps are made simultaneously and distributed and do not involve a "trusted" 3rd party entity or jurisdiction (Singh et al. 2019, cited in [49]). The security and performance behind BC are due to the cryptographic algorithms like RSA, Rivest-Shamir-Adleman, [50] and ECDSA (Elliptic Curve Digital Signature Algorithm) (Johnson et al. 2001, cited in [49]), and proof-work (PoW) and proof-of-state (PoS) consensus protocols. Notwithstanding advanced and integrated security mechanisms, BC technology has weaknesses and have "certain vulnerabilities" to attacks [49]. According to Orcutt [51], hackers have stolen about $2 billion work of cryptocurrencies from trading platforms since 2017, for example,

$1.1 million was taken from Ethereum Classic and $450 million bitcoins stolen from MtGox [49].

2.5.1 Blockchain and Self-Sovereign Identity Systems to Address Cyberattacks

Researchers Bandara et al. [52] proposed a blockchain and SSI based digital identity platform called "Casper" to address inherent problems with centralised identity systems such as cyberattacks and data breaches. However, a single definition of digital identity presents complexities in proving who the person says they are in the digital realm. It also offers legal, social and economic issues that have yet to be standardised or established and opens up favourable opportunities for a hacker to impersonate the individual [53]. Bitcoin has influenced the SSI evolution due to its underlying Distributed Ledger Technology (Dunphy and Petitcolas 2018, cited in, [54]). The majority of current identity platforms utilise centralised data storage architecture such as central servers and cloud storage, which have intrinsic security, data privacy, and user control issues. Stockburger et al. [55] contend that data is unprotected and insecure without digital identity. Casper integrates blockchain and SSI-based approaches and is an Android and iOS based mobile identity wallet app. The actual user/customer identities were contained in the individual's smartphone wallet app.

The proof of the user identities is contained in a blockchain-based decentralised storage system as SSI proof. SSI negates the requirement for central trusted authority [56]. The Casper platform's SSI-based system gives a Zero Knowledge Proof (ZKP) mechanism in verifying the identity information. The Casper platform is adaptable and can be used in banking, healthcare, government agencies, and businesses. Casper is intended to ensure security, decentralised and ZKP verifiable identity by utilising blockchain and SSI-based approaches. Zero-knowledge proof is a complex protocol incorporating encryption techniques. The prover convinces the other party, the verifier, of the truthfulness of an assertion or statement without the disclosure of other specifics than the statement itself [57].

For methodology, the researchers' use case for the Casper project was the implementation of an inter-bank Know Your Customer (KYC) for banking clientele. Customer identity or decentralised identity (DID) was embedded in the QR code of the mobile wallet. For the Casper project, all the user's personal data was stored in the user's mobile device hardware based on the SSI model; cryptographic DID proofs and other information were stored in blockchain storage. The researchers demonstrated that customers could prove their identity and be able to share their data with other banks, organisations, hospitals, and other entities when they used the mobile wallet. Furthermore, other entities were able to verify customers' identities using ZKP and to verify credentials; the researchers provided a mobile and web-based app for admin staff such as bank officers and healthcare service admins. The researchers' findings proved that the use of blockchain and SSI enabled DID systems coupled with iOS/Android mobile identity wallet (to capture and verify user's identity proofs)

addressed the threats and challenges in centralised identity systems. It also offered greater data privacy, confidentiality, integrity, and authentication while providing authorisation features. Even though blockchain technology seems ideal concerning SSI, Bokkem et al. [54] argue that there are limitations, for example, when the users lose the private/public key pair, the identity proofing process needs to start from the beginning to re-establish their digital identity.

2.5.2 Hyperledger Framework—MiTM Exploits in a Blockchain-Based Identity System

Bhattacharya et al. [58] examined scenarios whereby Personally Identifiable Information (PII) or personal data can be disclosed through credential exchanges between SSIs, risking MiTM exploits in a blockchain-based identity system like Hyperledger Indy. Hyperledger Indy is part of the Hyperledger framework (including Hyperledger Fabric, Cello, Iroha, Explorer and Composer), comprising open-source tools involving different organisations to build robust business-driven blockchain-based enterprise solutions [59]. The authors analysed the risk of MiTM attack that could takeover between two unknown peers DID connections in the initial setup process. An essential aspect of SSI systems is the unique relationships among peers in which an identity holder can form a relationship with another identity holder. Therefore, unless the two peers can satisfy each other about the authenticity of the peer ID connection, each party must verify the other when a new connection is established by using "verifiable credentials" (Deventer et al. 2020, cited in [58]). However, if a hacker can proxy a request/response between the two entities, then the authentication process between the entities fails.

Bhattacharya et al. proposed a mechanism to detect and mitigate the risk of MiTM attacks between peer SSIs. This involved an agency of self-signing features utilising the sender's private key peer ID, which will guarantee that the party generating the message and delivering it is the actual sender; and the use of unique DIDs and keys, which can only be resolved by the two parties in the relationship with each other. A mismatching signature alerts the receiving party that the message was not originating from the original transmitter. At this point, any peer connection is terminated to stop PII and data breaches. Additionally, Bhattacharya et al. proposed a quantitative model that computes a reputation score for credential issuers, enabling a quantitative confidence level value for the issuer. This aids in eliminating privacy and security concerns when there is a communication with a new peer that presents verifiable credentials that the issuer issued. The limitation of this study is that there was no comparison with other SSI ecosystems; it would have been worthwhile to present comparative analysis, however brief, with at least one other SSI system to ascertain how MiTM risk mitigations are handled. Furthermore, the authors did not propose best practices on building trust between DIDs, also did not suggest what minimum data would be required to complete a task to prevent the accumulation of private data by an attacker or even by legitimate parties.

2.6 *Using Artificial Intelligence Mitigation Predicated ML Techniques Against Attacks*

Zhang et al. [60] investigated and analysed DDoS attack detection and prevention using artificial intelligence mitigation predicated ML techniques. This work presented a detailed survey on the current advancements in detecting attacks using machine learning algorithms (Random Forest tree) plus Naïve Bayes. It provided recommendations on AL methods to be utilised to detect and prevent DDoS attacks. Typically, AI techniques include ML, natural language processing, and speech recognition [61]. The authors contend that the average size of packets, pack size variance, number of packets, number of bytes, bit rate, packet rate, and time interval are features that can be used in detecting DDoS attacks. However, Anandshree et al. [62] have argued that detecting DDoS attacks are complex because legitimate data packets are not distinguishable from illegitimate packets. And Yuan et al. [63] have suggested that AI/ML defences are more advantageous as countermeasures against DDoS attacks than other antidotes such as Blockchain risk mitigation techniques.

Zhang et al. use of AI techniques offer substantially higher accuracy in identifying and averting DDoS attacks. Applying Naïve Bayes in ML classifications provides about 97% accuracy. Adding a Random Forest tree or Gaussian Naïve Bayes with the data obtained produces at least 99% accuracy in detecting DDoS attacks. Substantially, automatically detecting packets from DDoS exploits becomes the primary mechanism for risk mitigations. Verisign [64] DDoS trends claim that:

- The top three industries targeted were the financial industry, IT Services/SAAS/Cloud, and the Telecom sectors.
- The financial sector represented 57% of mitigation activity, the highest routinely targeted industry; IT Services/SAAS/Cloud experienced 26% had the second-highest amount of DDoS attack; the Telecom sector represented 17% of mitigation activity.
- 58% of DDoS attacks mitigated by Verisign used at least two different attack types.
- User Datagram Protocol (UDP) floods accounted for 50% of DDoS exploits.
- The second highest frequent attack vector or 26%, were TCP-based attacks in the quarter.

Support Vector Machine (SVM) and Artificial Neuron Network are other ML algorithms applied to the DDoS defence anomaly detection phase. The authors recommend using Naïve Bayes and random forest trees to be used in classifying regular traffic and pernicious traffic for better performance. Furthermore, the authors recommend combining ML algorithms to detect DDoS exploits; these have a "better accuracy and performance".

2.7 *New Security Features in Wi-Fi 6 WPA3 and Enhanced WPA2 Security*

The enhanced security in WPA2 and the adoption of new security features in Wi-Fi 6 and WPA3 (Wi-Fi Protected Access 3) (Wi-Fi [65] introduced in June 2018 has been mandated for use in devices connecting to wireless networks to make data in transmission security more robust. The goal of WPA3 certification is securing home Wi-Fi networks, whilst enterprise wireless networks use EAP-pwd to authenticate users. Both the WPA3 certification and EAP-pwd use the Dragonfly handshake to give forward secrecy and protection against dictionary attacks [66]. The new WPA3 protocol could be a significant disruptor in MiTM attacks.

University of Leuven, Belgium, KU Leuven, researcher Vanhoef [67], discovered the KRACK or **K**ey **R**einstallation **A**tta**ck**s vulnerability in the WPA2 protocol. KRACK attack exploits the 4-way handshake protocol used in the WPA2 cryptographic mechanism when a device such as a smartphone is joining a wireless network. Threat actors can steal victims' data such as login credentials and credit card information when in transmission over WI-FI networks using the KRACK exploit. WPA3 aims to improve cyber security in the networks. Table 1 shows Common Vulnerability and Exposures (CVEs) attacks identified through specific instantiations of KRACK attacks; each CVE ID illustrates a specific WPA2 KRACK vulnerability.

The new announcements are: (i) new security specifications in the WPA3 protocol and (ii) enhancements to WPA2 security specifications.

2.7.1 WPA2 Enhancements

WPA2 enhancements include:

Table 1 CVEs identified through KRACK vulnerabilities

- CVE-2017-13077: Reinstallation of the pairwise encryption key (PTK-TK) in the 4-way handshake.
- CVE-2017-13078: Reinstallation of the group key (GTK) in the 4-way handshake.
- CVE-2017-13079: Reinstallation of the integrity group key (IGTK) in the 4-way handshake.
- CVE-2017-13080: Reinstallation of the group key (GTK) in the group key handshake.
- CVE-2017-13081: Reinstallation of the integrity group key (IGTK) in the group key handshake.
- CVE-2017-13082: Accepting a retransmitted Fast BSS Transition (FT) Reassociation Request and reinstalling the pairwise encryption key (PTK-TK) while processing it.
- CVE-2017-13084: Reinstallation of the STK key in the PeerKey handshake.
- CVE-2017-13086: reinstallation of the Tunneled Direct-Link Setup (TDLS) PeerKey (TPK) key in the TDLS handshake.
- CVE-2017-13087: reinstallation of the group key (GTK) when processing a Wireless Network Management (WNM) Sleep Mode Response frame.
- CVE-2017-13088: reinstallation of the integrity group key (IGTK) when processing a Wireless Network Management (WNM) Sleep Mode Response frame.

Source Vanhoef [67]

1.6.2.1 Improved authentication, configuration requirements, and encryption.
1.6.2.2 Mandatory use of PMF—Protected Management Frames.

- Management frames are used in initiating and terminating Wi-Fi connectivity, management frames transmitted are not encrypted, and its integrity is not verified without PMF
- With PMF, network management traffic integrity is ensured
- Protects against (i) eavesdropping (ii) replay (iii) forging od management action frames
- Protects against DoS/DDoS traffic-based attacks that use de-authentication and disassociation frames to remove a client from a network whereby the client is forced to authenticate again, a tactic used in MiTM attacks such as smartphones.

2.7.2 WPA3 Wi-Fi Security Features

WPA3 augmentations provide:

2.6.3.1 more robust encryption with mandated 256-bit encryption, compliance with CNSA approved cypher suite requirements [68]. The overall effect is to enable 192-bit encryption security for Wi-Fi networks.

2.6.3.2 new OWE (Opportunistic Wireless Encryption) protects against eavesdropping; this replaces open, unencrypted networks and allows hackers to read and modify users' traffic. OWE enables individualised user encryption on public networks such as airports and cafes to defend against brute force or dictionary password attacks on networks relying on password-based authentication.

2.6.3.3 SAE or Simultaneous Authentication of Equals is the new powerful password-based authentication method replacing PSK (Pre-Shared Key) mode, which is susceptible to passive and active brute force attacks. SAE limits the number of guesses an attacker makes – this currently stands at a "rate of 4000,000 possible passwords per second". SAE adds to the user experience, which does not change [68].

2.6.3.4 Device Provisioning Protocol (DPP), a mechanism in provisioning IoT appliances with limited or no user interface in a trusted network.

2.6.3.5 will disallow outdated legacy protocols.

2.7.3 Reported WPA3 Vulnerabilities

Notwithstanding the improved security features in the new WPA3 protocol, it is not perfect; recently, some vulnerabilities have been reported:

- Denial of Service/MiTM attack
  - Fragmentation and Aggregation attack

- Downgrading Attack
  - Exploits backward compatibility
  - Exploits dragonfly handshake
- Side-Channel Attacks
  - Timing-based
  - Cache-based

3 Methodologies and Frameworks

Due to the increase in cyberattacks and the requirement for security appraisal and risk mitigation strategies, a few methodologies and frameworks have been developed to aid in a structured approach to cybersecurity research. These include NIST 800–115, OSSTM, PTES, OWASP, and MSF. The following frameworks are adaptations from research by Shanley [69].

The NIST SP 800–115 document is a technical guide to information security testing and is adaptable for assessment; the guide aids entities/organisations to develop their own information security (IS) methodology. It was developed for US federal government agencies; however, it is freely available for use by the private sector [70]. Unlike OSSTMM, NIST SP 800–115 does not focus on penetration testing alone but as part of a general process that focuses on the identification of vulnerabilities through repeatable, detailed planning and execution assessments, followed by conducting analysis. Like OSSTMM, NIST SP 800–115 does not suggest tools for cybersecurity tests, although it lists some tools that can be used, and assumes that the security professional has the requisite skills and knowledge to conduct penetration tests.

OSSTM is a security approach utilised in evaluating operational security and analysis. It is an open-source license and an audit methodology designed to be a "consistent and repeatable measurement of security at the operational level" developed by ISECOM [71]. Tests are partitioned into five channels: these channels test (i) data/information controls, (ii) mobile devices, wireless devices, and physical security access controls, (iii) human interactions and personal security awareness levels, (iv) social engineering and fraud control levels (v) telecommunications and computer networks, (vi) physical security access, and (vii) buildings and perimeters [72]. OSSTMM is for penetration testing to satisfy regulatory requirements [73]. OSSTMM recommends best practices, guidelines, and trust metrics for assessing risks and attack surfaces, it does not recommend what tools to use because it assumes that security professionals will have adequate knowledge of techniques and tools to perform the tasks in the modules [71].

Penetration Testing Execution Standard (PTES) is a penetration testing standard providing guidelines for the entire scope of pen testing activities in seven main sections covering (i) "pre-engagement interactions, (ii) intelligence gathering,

(iii) threat modelling, (iv) vulnerability assessment, (V) exploitation, (vi) post-exploitation, and (vii) reporting" [74]. Faircloth [75] suggests that pen testing of wireless networks includes the same methodologies used in testing individual systems. Like the Open Web Application Security Project, OWASP, PTES is a community standard, which aims to improve web applications via the provision of tools, guidelines, and reports [76]. PTES does not give technical guidelines on the process of conducting a pen test. Instead, the process is described at a conceptual level. The PTES standard has technical guidelines which include specifications of specific tools and instructions on how to use them [75]. Like OSSTM, PTES assumes that the security professional will have some knowledge of techniques and tools of pen-testing. However, unlike OSSTM and NIST SP 800–115, PTES attempts to remedy the shortcomings by providing methods, guidelines, tools, and techniques in a single document.

The Open Web Application Security Project (OWASP) Foundation (OWASP n.d.) is an international technical not-for-profit organisation aiming to improve security in software focusing on research, testing, tools and resources, methodologies, education, and training. OWASO research updates information on the latest prevalent vulnerabilities for web applications [77]. The OWASP Testing Guide or OTG is a framework for web applications, software development security, web application security testing methodology, which explains "how to test for evidence of vulnerabilities within the application due to deficiencies with identified security controls", and reporting (OWASP n.d.). The OTG offers a web application testing methodology more focused on security relating to the software development stages instead of identifying vulnerabilities after the software is developed and released to the public. Testing includes white box and black box testing. The OTG is divided into three main sections (i) the OWASP testing framework, (ii) web application security testing introduction and objectives and (iii) reporting, with each section having further detailed sub-divisions. The Application Threat Modelling is provided by the OWASP guide, which is used in application testing security flaws during the design of the application (OWASP1 n.d.). OWASP WebGoat is an insecure J2EE application developed to educate pen testers on web application security [78]. OTG is mainly suited for web applications only. Unlike OSSTMM, OTG has a strong focus on the security of web applications during the Software Development Life Cycle (SDLC) with recommended tools for the security professional. The OWASP To 10 is a security risk awareness document and a de facto industry application security standard. Furthermore, in testing application technical security controls, the OWASO's ASVS or Application Security Verification Standard is the standard applied (OWASP2 n.d.).

Metasploit Framework (MSF) is the world-leading pent testing solution; it is a modular penetration testing platform that enables testers to "write, test, and execute exploit code" written in Ruby programming language. Metasploit can be run as a stand-alone or from Kali Linus. MSF is a multitude of tools providing the environs for pen-testing and vulnerabilities development. It consists of multiple tools for enumerating networks, investigating security vulnerabilities, attack executions, and detecting evasion (Rapid7 n.d.). Metasploit was initially developed in 2003 as an open-source

license by HD More. It was acquired in 2009 by Rapid7 company providing vulnerability management solutions, and it oversees development and funding [79]. The elements of MSF that can be leveraged are the virtual or isolated working environment, MSF ability to launch exploits, its database, and the Meterpreter payload. The exploit database is a repository storing all attacks that MSF can launch. Once an exploit is selected and brought to the foreground, it can be customised and then launched, the attack result is displaced when an attack is completed.

Metasploit Framework is run manually in the command-line for developers and security researchers. It is used extensively in pen testing with exploits of more than 1650 and with features such as import of network scan. In contrast, MSF Pro is a commercial version with advanced features such as a web interface, automation, integration via remote APIs, network discovery, and website application evaluations for OWASP vulnerabilities—and much more (Rapid71 n.d.).

The primary modules in MSF are stored under directory: /usr/share/metasploit-framework/modules/[80]:

(a) Exploit—modules that use payloads
(b) Auxiliary—modules include port scanners, sniffers, fuzzers, etc.
(c) Payloads—consists of code that runs remotely
(d) Encoders—ensure that payloads make it to the destination intact
(e) Nops—keeps payload size consistent across exploit attempts.

Armitage is a Java-based GUI for MSF and is accessible by multiple parties for collaboration within a pen testing team [81]. Unlike NIST 800–115, OSSTMM, PTES, and OWASP, MSF provides a suite of tools to provide practical pen testing solutions for which security professionals can take advantage.

The main advantage of MSF is its modularity that allows combinations of exploits with any payload; this acts as a motivation for pen testers and exploits coders. A significant disadvantage and limitation of MSF are that most of the exploit in the MSF system is Windows platform-based, probably because many applications have been developed for the Windows operating system, which is prevalent globally.

Kali Linux framework is a leading open-source and advanced penetration testing software; it is used for advanced information security tasks, ethical hacking, uncovering vulnerabilities, assessing network security, reverse engineering, and computer forensics; it is Debian-based Linux distribution (Kali.org n.d.). Penetration testing can be defined as "the operational process of analysing or evaluating the security of a computer system or network" (Arkin et al. 2005, cited in [82]. Kali contains over 600 tools and is used by ethical hackers to test their security skills. This includes the Aircrack-ng suite of tools used for demonstrating attack scenarios in this project. The pen testing conducted throughout this project is termed ethical or white-hat hacking; this is legal [83].

3.1 Research Design

The research design will attempt to answer the research questions and be presented in three phases. Securing communications (i.e., data) is critical in WLANs as data is communicated through the air medium.

Phase 1 will review the downgrade of WPA3 to WPA 2, leading to DoS/DDoS and MiTM attacks.

Phase 2 will demonstrate a DoS/DDoS attack on a private smartphone through Aircrack-ng; this is a penetration testing technique. Bacudio et al. [84] contend that pen testing is a sequence of events conducted to "identify and exploit security vulnerabilities", this confirms the ineffectiveness and effectiveness of implemented measures regarding security.

Phase 3 will extend upon Phase 2 and demonstrate an Ettercap-based MiTM attack via ARP Poisoning; this is also pen-testing.

The materials utilised in this project consists of the following:

- MacBook Pro (Retina, 13-inch)
  - Processor—2.6 GHz Dual-Core Intel Core i5
  - Memory—8 GB 1600 MHz DDR3
- Wireless Adapter Card: ALFA NETWORK (AWUS036NHA)
  - Monitor/injection mode support
  - 802.11b/g/n protocols support
  - Supports 150Mbps 2.4 GHz wireless access
- Wireless router
- Oracle VM VirtualBox (n.d.) (virtual environment).

3.2 Data Analysis

Data analysis will comprise interpreting the outcome obtained from the pen testing outlined in the research design, how this relates to WPA3 newly discovered vulnerabilities and degradation to WPA2, and the consequences thereof.

A few researchers, such as Vanhoef and Ronen [85], have performed systematic analysis into the recently released and enhanced security in the WPA3 protocol. The researchers found severe vulnerabilities, including downgrade, denial of service, MiTM and side-channel attacks on WPA3.

3.3 MiTM Attack Demonstrations

3.3.1 Phase 1: Review of Dragonfly Degradation and "Dragonblood" Exploit

Recent research by Vanhoef and Ronen [66] on the newly launched WPA3 protocol demonstrated security vulnerabilities, which the researchers termed "Dragonblood". The Dragonblood vulnerability directly correlates with the ability to degrade the Dragonfly handshake mechanism of WPA3 to WPA2 and subsequent MiTM attacks. However, the WPA3 Dragonblood vulnerability does not form part of the demonstrations in this project because this was not part of the research proposal. Furthermore, time constraints and availability of WPA3 devices and materials would not have been readily available at the onset of this project. Detailed discussions of the Dragonfly mechanism and related Dragonblood exploit are presented in Chap. 4—Data Analysis and Discussions.

3.3.2 Phase 2: DoS/DDoS Attack on WPA2

This practical will capture a WPA2 4-way handshake between an AP and a client (smartphone) using the Aircrack-ng suite of tools in the Kali framework. An attempt will be made to use brute force in cracking or breaching the password; this will be for pen testing purposes.

3.3.3 Staying Anonymous During Pen-Testing: Spoofing MAC Address

During pen-testing, it is paramount to be anonymous; this can be achieved by changing the MAC address, anonymity avoids detection. The MAC or Media, Access Control address, is unique to every device's NIC or Ethernet network interface card; the MAC address is 48 bits long [86]. Prior to performing the attack scenarios, the MAC address is changed, this is also known as "spoofing" the MAC address. The change is not permanent but temporary and exists in RAM only. GNU MAC Changer or Macchanger is a Kali tool used for MAC address manipulation in network interfaces; the MAC address is randomised, as illustrated in Fig. 1 (Kali.org n.d.). When the MacBook is restarted, the original MAC address is restored.

Alternatively, the MacBook Terminal tool and Linux command can be used to spoof the MAC address as follows:

(i) Obtain the MAC address of the machine with the command:

ifconfig or ifconfig en0 | grep ether

(ii) Generate Hexadecimal MAC number—Fig. 2.
(iii) (a) disconnect from wi-fi then connect to wi-fi but not AP/router.

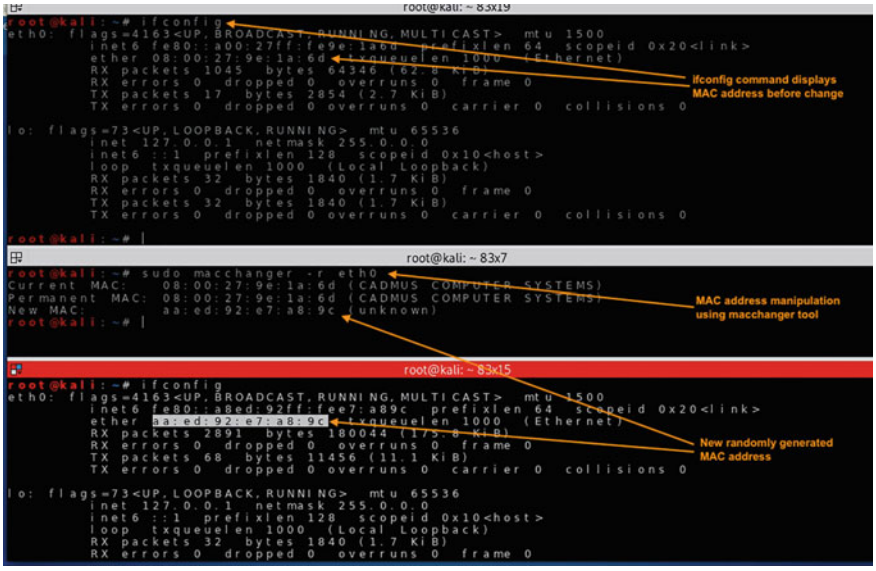


Fig. 1 Changing MAC address with Macchanger Linux command

Fig. 2 An alternative method to generate a random Hexadecimal MAC number

(b) disconnect from VPN.

(iv) Followed by commands:

sudo --login

ifconfig en0 ether 06:28:66:ae:cd:45 ← from new MAC address generated in (ii)

Step 1: Update Kali.

Use the command: *sudo apt update*—Fig. 3.

Then upgrade to the latest Kali version with the command: *sudo apt full-upgrade -y* (Fig. 4).

Fig. 3 Updating Kali in the virtual environment

Fig. 4 Upgrading Kali

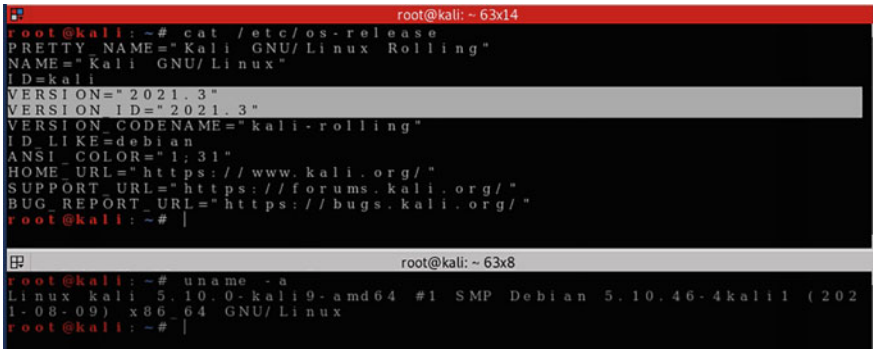


Fig. 5 Kali and Linux versions

Step 2: check for Kali and Linux versions—Fig. 5.
Terminal horizontal split screen shows:

(a) Kali version in use: *cat /etc/os-release.*
(b) Linux version: *uname -a.*

Step 3: the wireless interface details—Fig. 6.

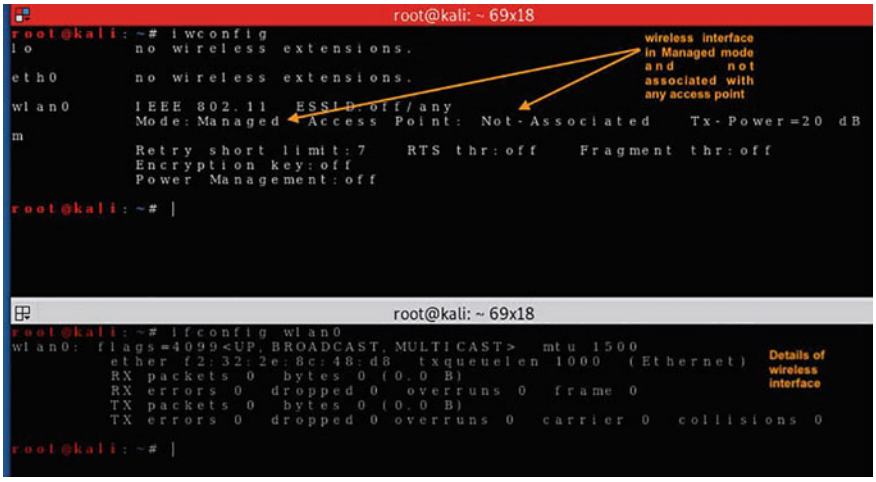


Fig. 6 wireless interface details

Figure 7 screenshot shows the green light of the ALPHA [87] wireless adaptor device, which is switched "on".

Figure 8: Terminal command: *airmon-ng start wlan0*—command puts the wireless interface in "Monitor" mode for the purpose of packet capture from surrounding APs

Step 4: kill processes that might interfere with Monitor mode—Fig. 9

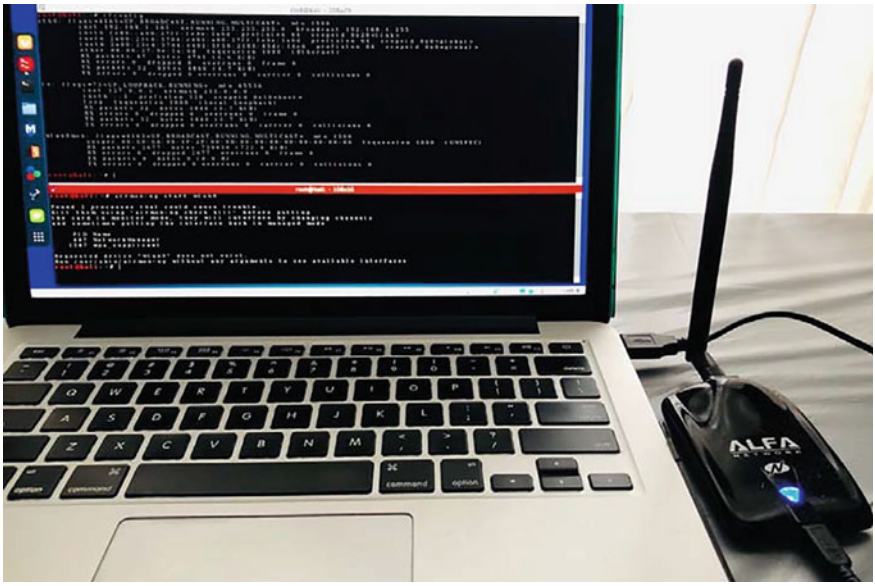


Fig. 7 Wireless adapter WLAN0 set to monitor mode to sniff data packets

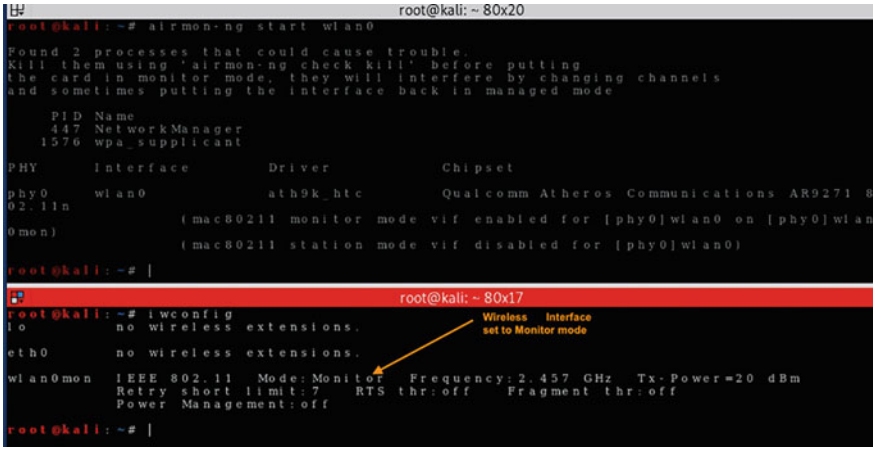


Fig. 8 Wireless interface in monitor mode

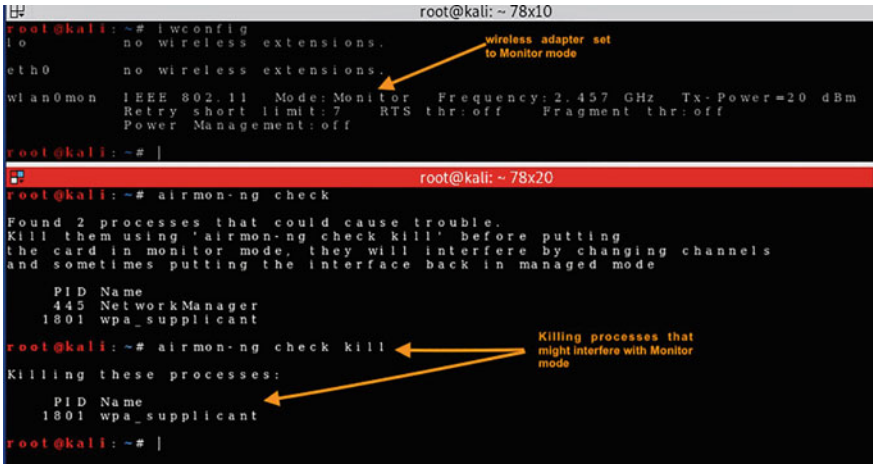


Fig. 9 Processes interfering with monitor mode

Step 5: command *airodump-ng:* capturing/sniffing available networks/APs in the vicinity and grabbing packets using the interface. This process is also called "channel hopping"; by hopping multiple channels to detect APs or routers within range, as shown in Fig. 10.

BSSID is the MAC address of the target network; ESSID is the name of wireless networks within range; PWR is the signal strength; CH is the channel; Beacon is the access point/network broadcasting its presence; ENC is the encryption protocol used by the network, e.g., WPA2; #Data is the number of data packets being sent, and AUTH is the authentication used on the network.

Step 6: targeted sniffing on specific AP of interest (BSSID = 18:82:8C:1D:F4:5B) and writing captured data packets to file named "capture"—Fig. 11.

Command: *airodump-ng -c6 -w capture -d 18:82:8C:1D:F4:5B wlan0mon.*

Step 7: DoS/De-authentication attack on AP and station of interest.

The aim is to detach the station from the AP/router so that in the process of re-association with the AP, the 4-way handshake is captured, as shown in Fig. 12.

Command: *aireplay-ng –deauth 0 -a 1x:82:8x:1D:×4:5B -c 5x:xx:96:Bx:8B:3E wlan0mon.*

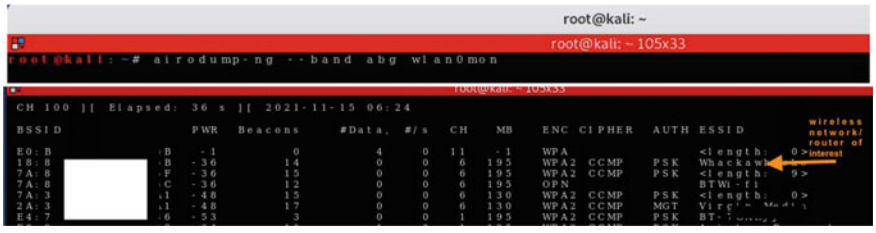


Fig. 10 The screenshot shows AP with details

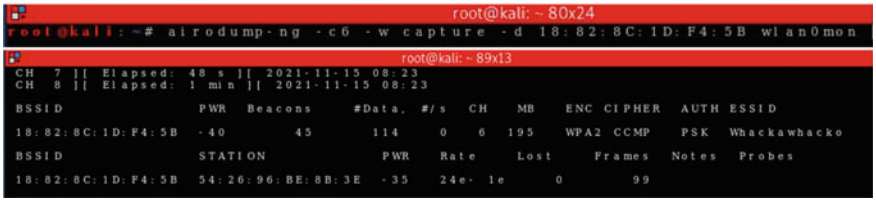


Fig. 11 Packet sniffing on a specified AP and station

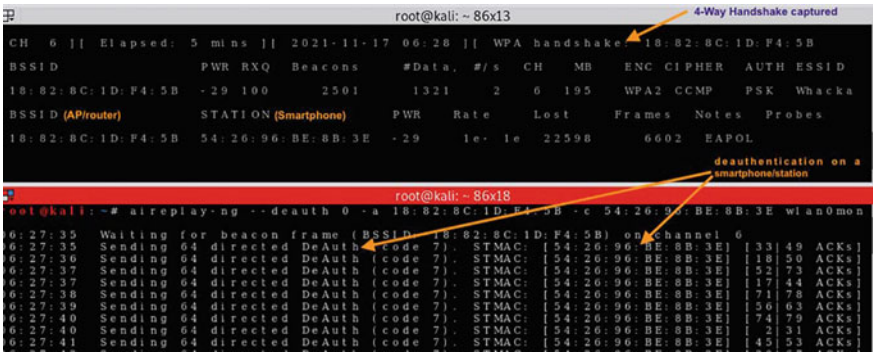


Fig. 12 Deauth attack a station (i.e., smartphone) to capture 4-Way Handshake

Capture file with data—Fig. 13.

Figure 14—stop monitor mode after data capture

A Shell script program simulates a DDoS attack by changing the MAC address and attacking the AP in the program loop (Fig. 15).

Step 8: Start Wireshark analyser to view 4-Way Handshake and other data details.

Step 9: Aircrack-ng—brute force cracking of WPA password (Fig. 16).

Successful cracking of keywords will depend on the password complexity, how comprehensive and extensive the wordlist being used is, and the password not ordinarily found in a dictionary. In this case, the password was not found because it is complex; it is a personalised passphrase comprising of (a) 15 characters long, (b) alphanumeric and (c) special characters.

Fig. 13 Capture file with data

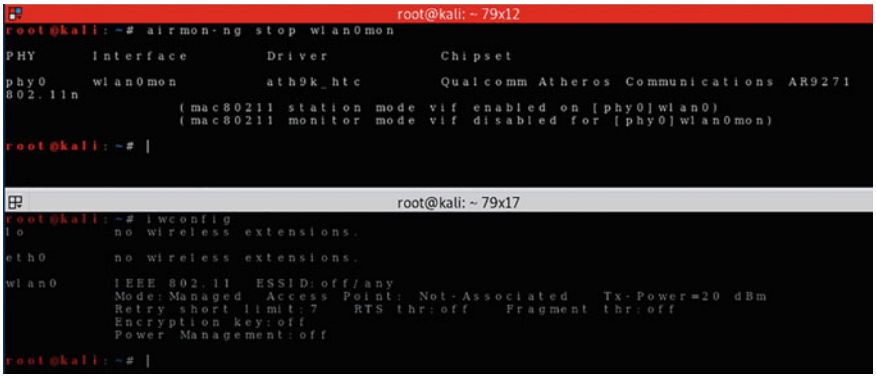


Fig. 14 Stop monitor mode

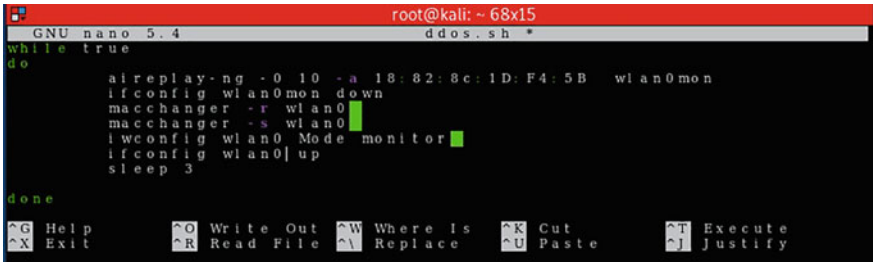


Fig. 15 DDoS shell script program

Fig. 16 Aircrack-ng command to capture password

Figure 17 shows that brute force to crack passwords did not work in this instance.

More advanced and sophisticated tools are available to crack complicated passwords, as shown in Fig. 18. These tools include the GPU Hashcat and Python CUPP tool; both generate brute force attacks. The use of such a sophisticated attack is not within this project's scope.

3.3.4 Phase 3: Ettercap-Based ARP Poisoning MiTM Attack

Ettercap is an open-source tool pre-installed in Kali Linux. In this simulation scenario, address resolution protocol (ARP) poisoning MiTM attack is demonstrated against a Wi-Fi network between a router and a target user, a smartphone. During a regular data transmission over Wi-Fi, messages are routed over the network by associating

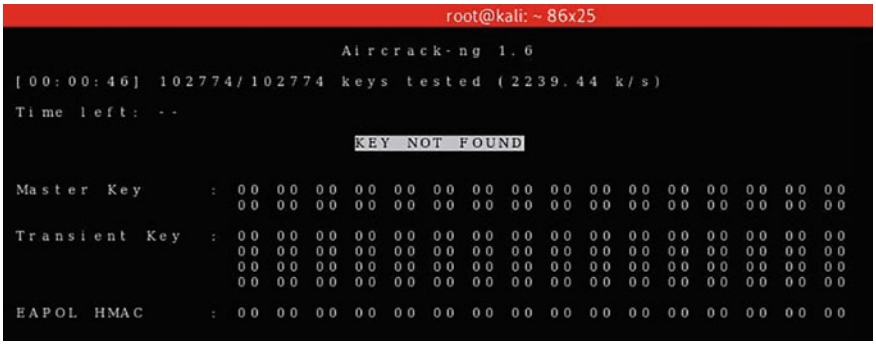


Fig. 17 Password not found in brute force cracking of station/smartphone password

Fig. 18 Strong password

the device MAC address and its IP address; this is done via the ARP. However, this can be "spoofed" to change the data traffic routing whereby messages meant for the target smartphone are transmitted to the hacker instead, allowing the hacker to deny service and man-in-the-middle the smartphone.

Step 1: The Default Gateway (Fig. 19) is determined to be 192.168.1.254 using the command: *netstat -nr.*

St**ep 2**: Enumeration (Fig. 20) to extract machine names and network resources using the command: *nmap -sn.*

The command: *arp-scan -l* can also be used to scan the network for IP addresses with their corresponding MAC address.

Step 3. Ettercap can be run in either command mode or by using the graphical interface. Allow IP forwarding using the command in Fig. 21. Number 1 indicates that ip_forwarding is now enabled.

Step 4. Ettercap MiTM attack in terminal command mode (Fig. 22).

Starting MiTM Ettercap attack manually: *sudo ettercap -T -S -i eth0 -M arp:remote /192.168.1.254// //192.168.1.97//*

Step 5. Ettercap MiTM attack in graphical interface mode (Fig. 23).

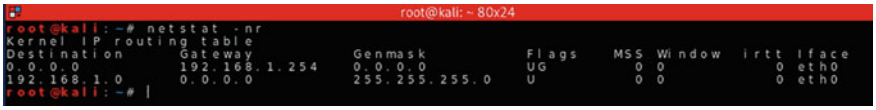


Fig. 19 System default gateway

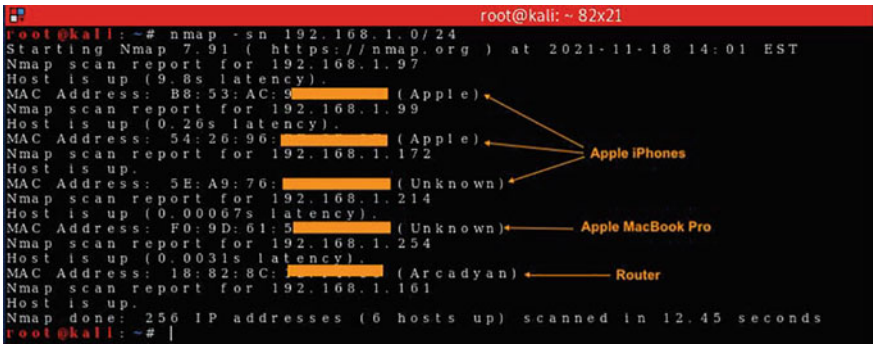


Fig. 20 Enumeration process to extract machine names

Fig. 21 IP forwarding enabled

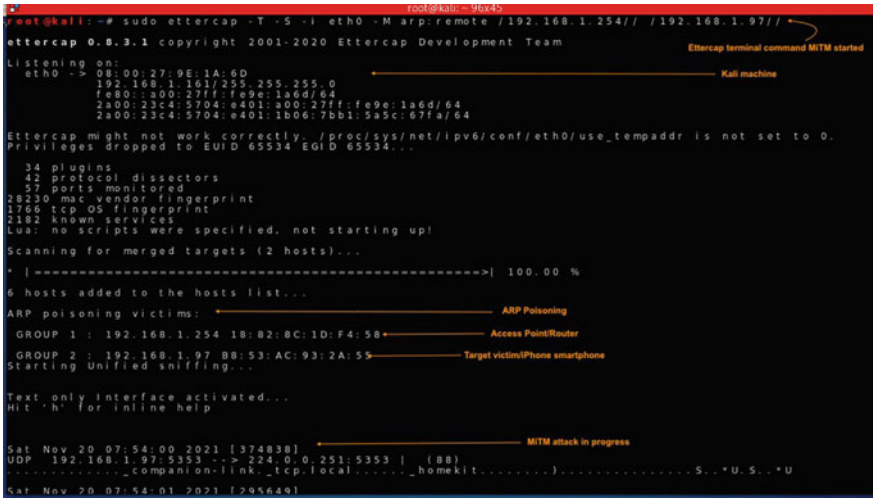


Fig. 22 Ettercap MiTM attack in terminal mode

$$\text{TRGET1} = \text{iPhone (smartphone)}$$
$$\text{TARGET2} = \text{Kali machine}$$

Step 6. MiTM ARP poisoning. in progress—Fig. 24.

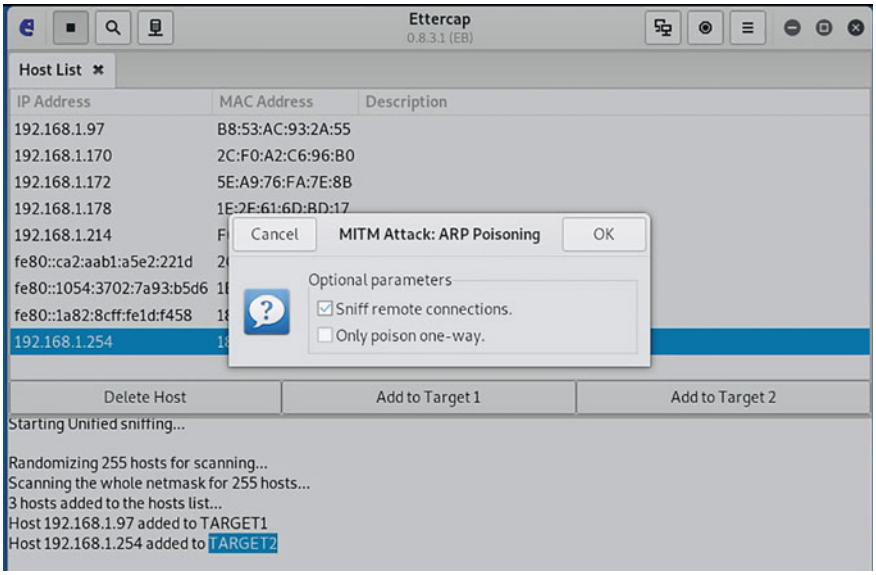


Fig. 23 The selected target for ARP poisoning attack

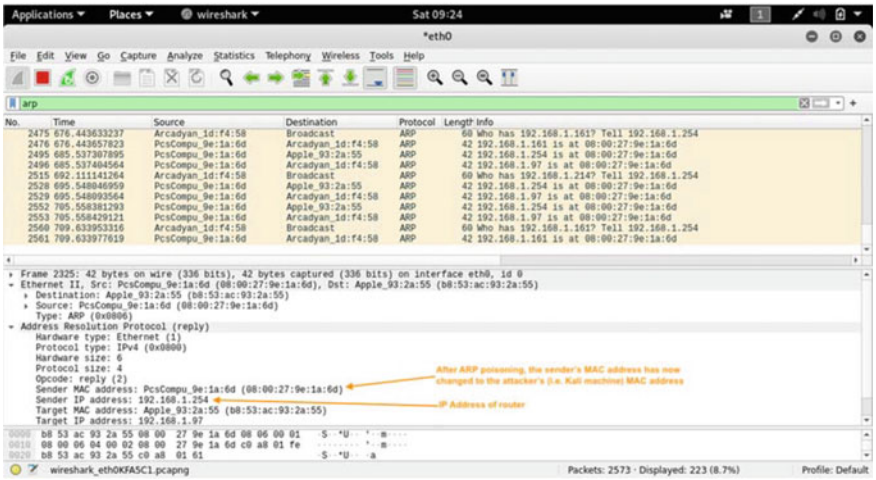


Fig. 24 MiTM ARP poisoning in progress

4 Data Analysis and Discussion

4.1 Explaining the 4-Way Handshake Problem/vulnerability

Understanding the 4-Way Handshake mechanism is critical to the comprehensive appreciation of the WPA3 dragonfly mechanism and the Dragonfly attack, leading to DoS and MiTM attacks. It is commonly known that the 4-way handshake method (as defined in 802.11i) utilised in WPA2-Personal wi-fi networks and applied by all secured Wi-Fi systems in generating a new session key can readily be cracked using a single capture of a data packet as demonstrated in Chap. 3. The weaknesses in the 4-way handshake are demonstratable in the KRACK vulnerability Vanhoef and Piessens [88].

A client such as a smartphone connects to a Wi-Fi network by authentication and association; this is a mutual process. The association stage is a typical connection to Wi-Fi at airports and cafes where no actual authentication occurs; no passwords are needed. This is Open System and Null authentication allowing all clients to authenticate without a password (Wireless [89].

The main elements or keys of interest in the 4-Way handshake are MSK (Master Session Key), PMK (Pairwise Master Key); GMK (Group Master Key); PTK (Pairwise Transit Key); GTK (Group Temporal Key); ANonce, SNonce; and MIC. The actual authentication is conducted during the 4-way handshake and is predicated on the shared secret PMK or Pairwise Master Key. The PMK resides in the client now called the supplicant, and APs now called the authenticator during the handshake. In a personal network, the Pairwise Master Key is generated from a pre-shared password, while for an enterprise, the PMK is generated using 802.1 × authentication. The PTK is generated by combining the PMK, MAC address of the authenticator and supplicant, plus the ANonce (Authenticator Nonce), and SNonce (Supplicant nonce)(Vanhoef and [88].

PTK can be derived as:

$$\text{PTK} = \left(\text{PMK} + \text{MAC}_{(\text{authenticator})} + \text{MAC}_{(\text{supplicant})} + \text{SNonce} + \text{ANonce}\right)$$

When generated, the PTK is divided into three, KCK (Key Confirmation Key), KEK (Key Encryption Key), and TK (Temporal Key). KEK and KCK protect handshake messages, and the TK is utilised in protecting regular data-frames. When WPA2 is used, the 4-way handshake transmits the GTK to the supplicant [88].

In level one, the MSK is generated through 802.1 × and EPA-TLS encryption.

In level two, GMK and PMK keys are generated from the MSK, and PTK and GMK keys are generated from the PMK.

Level three keys are used for data encryption.

After the initial authentication and association, security validation and the 4-way handshake process commence where messages exchanges occur over EAPoL (Extensible Authentication Protocol over LAN).

Message 1 (Fig. 25): The AP sends an EAPOL message containing Anonce, a randomly generated number, to the station to generate the PTK.

Message 2 (Fig. 26): After the creation of the PTK by the station, the station sends out SNonce required by the AP to generate its own PTK for unicast traffic encryption.

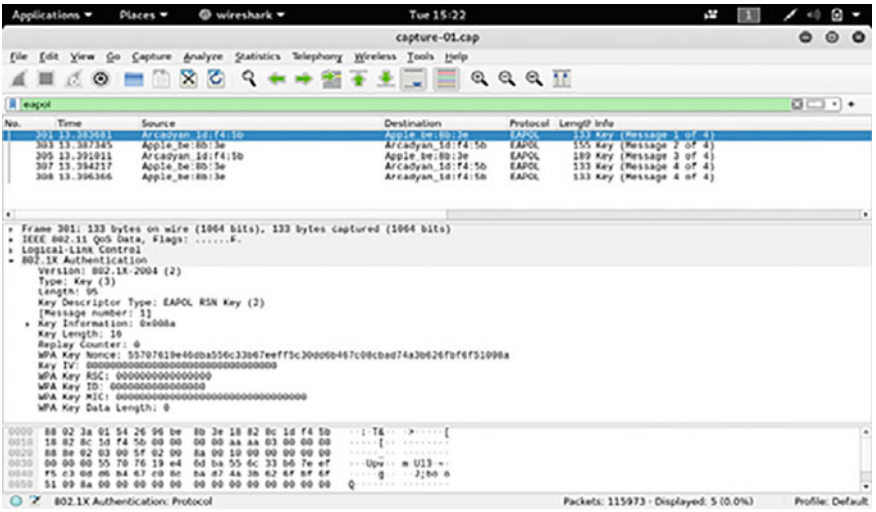


Fig. 25 Wireshark view of message 1 details

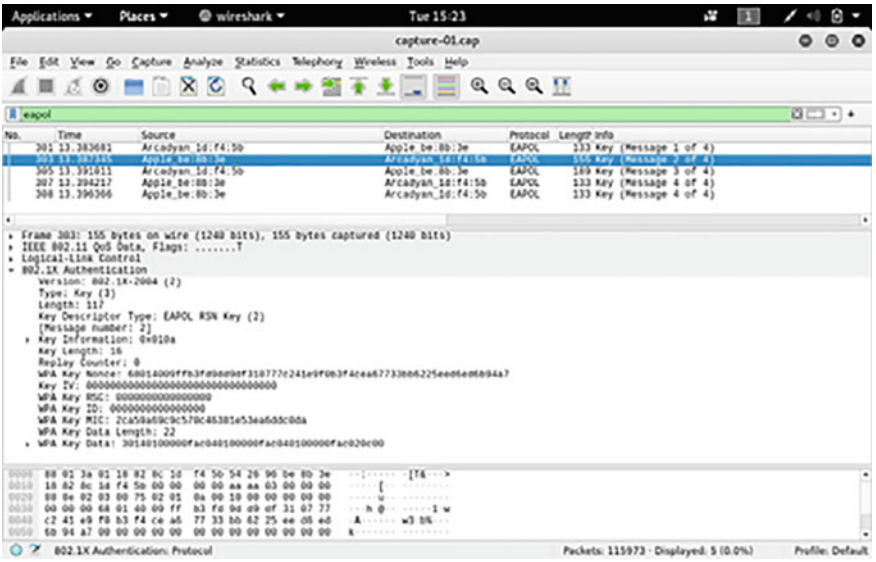


Fig. 26 Wireshark view of message 2 details

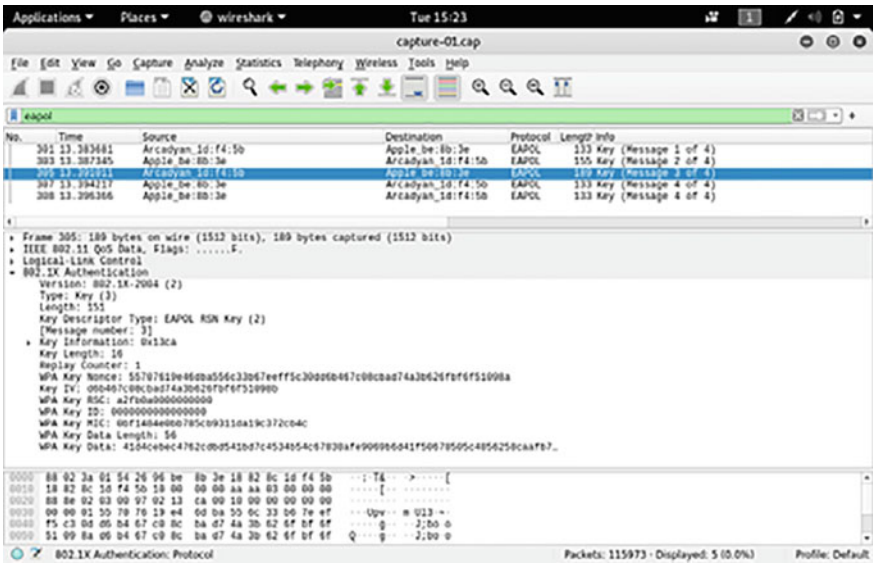


Fig. 27 Wireshark view of message 3 details

In the process, the station sends an EAPOL message containing the message integrity check (MIC) to ensure the AP can check if the message is modified or corrupted.

Message 3 (Fig. 27): AP sends a message to the station containing the GTK.

Message 4 (Fig. 28): Station sends a fourth and final message to AP confirming the installation of keys.

Upon successfully completing the 4-way handshake, the virtual control port is opened to allow the flow of encrypted data, unicast data is encrypted with PTK, and multicast data is encrypted using the GTK.

However, messages can be dropped or lost in transition; the authenticator (AP) retransmits message number 3 if the appropriate acknowledgement response is not received. Potentially, the supplicant may get message number 3 multiple times. Upon receiving message number 3 again, the same session key is reinstalled, thereby resetting the nonce number (the incremental transmit packet number) and the received replay counter used by the data-confidentiality protocol. A hacker can force resets of the nonce by "collecting and replaying retransmissions of message 3". Hence, the protocol in the data confidentiality is violated by forcing nonce reuse in this way. For instance, packets are re-playable, can be decrypted, and or forged [67].

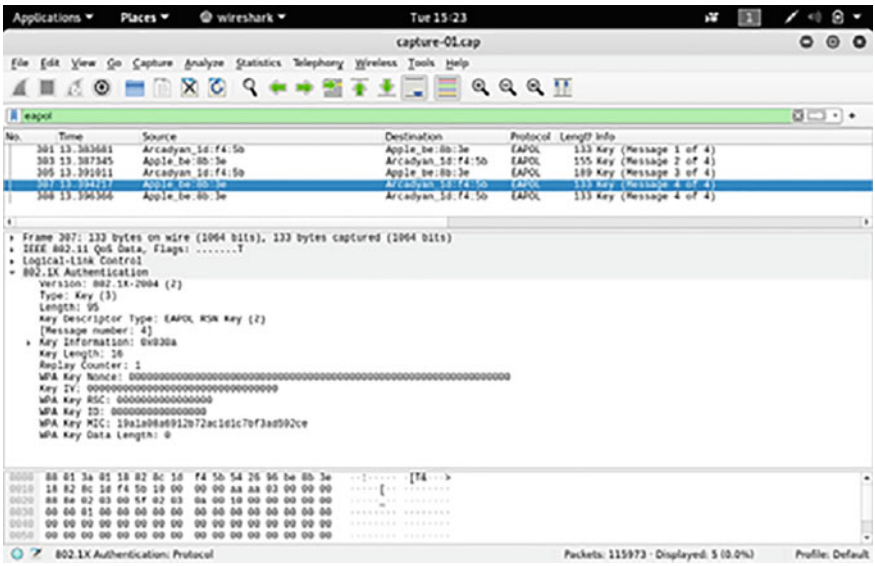


Fig. 28 Wireshark view of message 4 details

4.2 WPA3 Dragonfly Handshake and the Dragonblood Vulnerability

The Dragonfly handshake mechanism and WPA3 design flaws, and the Dragonblood attack are documented by Vanhoef and Ronen [66], the researchers who discovered the vulnerabilities. A complete account of the Dragonfly mechanism and Dragonblood attack is not within the scope of the project proposal. However, this section presents a brief synopsis taken from the research paper by Vanhoef and Ronen [66].

The improved WPA3 Dragonfly-Handshake is intended to make it extremely difficult for hackers to breach the 4-way handshake resistance against offline brute-force dictionary attack; the introduction of WPA3 perfect forward secrecy aids in preventing hackers from decrypting previous traffic following a key breach and thereby making use of Zero-knowledge proofs. The WPA3 vulnerabilities fall into two categories: (i) downgrade attacks against WPA3 enabled devices and (ii) weaknesses in the SAE (Simultaneous Authentication Equals) handshake, also known as the Dragonfly handshake. The adoption of the SAE in WPA3 allows for transition mode connections and compatibility with older devices using WPA2. In this situation, an adversary can modify beacons, making the client think the AP is supporting WPA2 protocol only. By using known WPA2 security attacks like PMKID and KRACK, the attacker can recover the network password. In essence, the hacker forces the device with WPA3 to use WPA2, which negates the KRACK and PMKID countermeasures. This mode of attack is termed the downgrade attack. By this point, the hacker can

adequately capture data to carry out a dictionary attack even though the downgrade attack is detected by the WPA2 4-way handshake.

Further to the SAE compatibility downgrade attacks, another attack against the Dragonfly handshake worth mentioning is the Dragonfly password encoding mechanism side-channel attacks known as the hash-to-curve operation. The hash-to-curve operation has a high overhead, which allows a hacker to exploit the high overhead. This is done by impersonating a client to "impersonate a user and transmit a commit frame, and to deliberately delay the response speed at the access point with subsequent attacks to perform a DOS attack ". The Dragonblood attack is currently the most critical vulnerability in the recently released WPA3 security protocol and requires immediate correction before WPA3 enabled devices become widely available for use.

4.3 Zero-Day Attack

The recently announced Dragonblood attack can be termed a zero-day attack because it is a security vulnerability on the new WPA3 protocol; it will continue to be abused until the vendor patches the exploit [90]. The Window of Vulnerability or WOV is the timeframe the vulnerability is initially made public to the time the security patch is finalised or when the exploitations reduce to insignificance. t_0 equals the time when the first client gets a patch p, t_1 equals the time the last client gets patch p. Given that Δ_{attack} is the time the hacker requires to reverse engineer the patch p and make it a viable exploit, then WOV starts at $t_{0+}\,\Delta_{attack}$ and finishes at t_1 [91].

4.4 Data Analysis and Data Visualization

Data visualisation is visually representing information that communicates information concisely and clearly without being confusing and cluttered; it is a compelling visual to enhance understanding of the phenomenon [92]. Using graphs and charts to illustrate cyberattack patterns and activities instead of reading through several logs, reports and spreadsheets enable the security administrator to pinpoint the severity and scope of cyberattacks expeditiously. Additionally, using DV saves time analysing extensive data and applying faster action [93]. However, using data visualisation highlights the requirement in more robust data governance and data management and the necessity for clear boundaries and data dissemination or transmission, monitoring, and tracking—among individuals with the ability to alter data origination and "write back to the system record through their visual discovery activities" [94]. Practices and privacy attitudes of organisations in the collection of data carry ramifications for data confidentiality, availability, and integrity [95].

The pen testing simulations conducted for this project and the data produced are applicable to the financial services sector. For financial services entities, the CBEST

(Bank of England Penetration Testing Framework) mechanism by The [96] is the primary means to evaluate security safeguards in the financial sector by employing sophisticated threat intelligence coupled with achievable pen-testing simulations. The Annual Cybersecurity Report by Bulletproof [97] suggests that DoS or DDoS attack could cost a lar ge business upwards of $2 million and $120,000 for a small-medium enterprise.

4.5 Comparative Data Analysis Between MiTM DoS and ARP Poisoning Attacks

Data are analysed using the Wireshark Statistics tools. Figure 29 screenshot gives the DoS scenario capture file data like the file name, length, Hash properties and encapsulation; capture suration (start and end time); hardware; interface type and packet size limit; and capture statistics.

The graph in Fig. 30 (obtained through Wireshark interface, Statistics I/O Graphics function) shows typical DoS traffic generated; the peaks in the graph indicate bursts of traffic in 100 ms time intervals. These were created in Phase 2 practical by generating denial of service attacks in the Kali Linux platform. In this case, numerous significant traffic bursts were generated, indicating the many deauthentication attacks during the

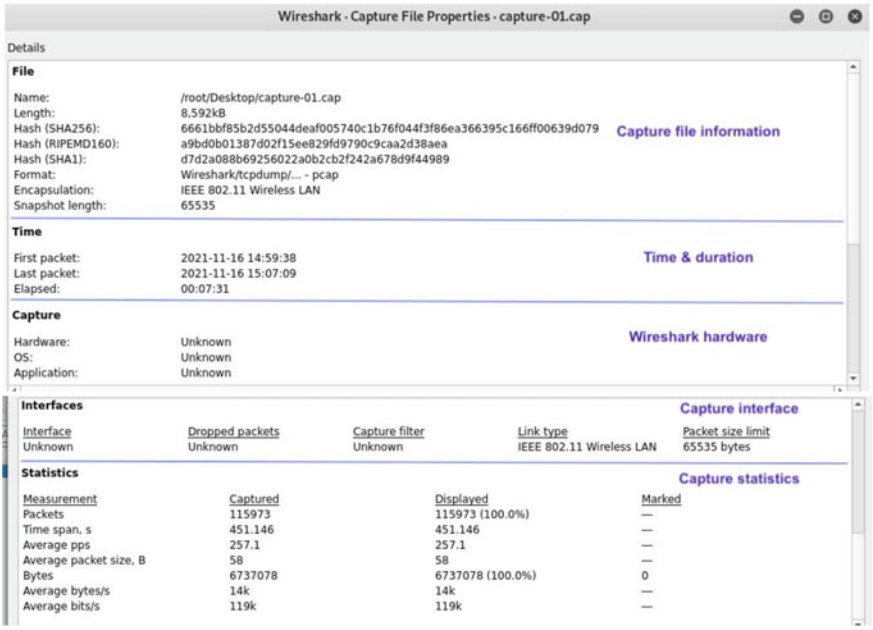


Fig. 29 DoS capture file properties

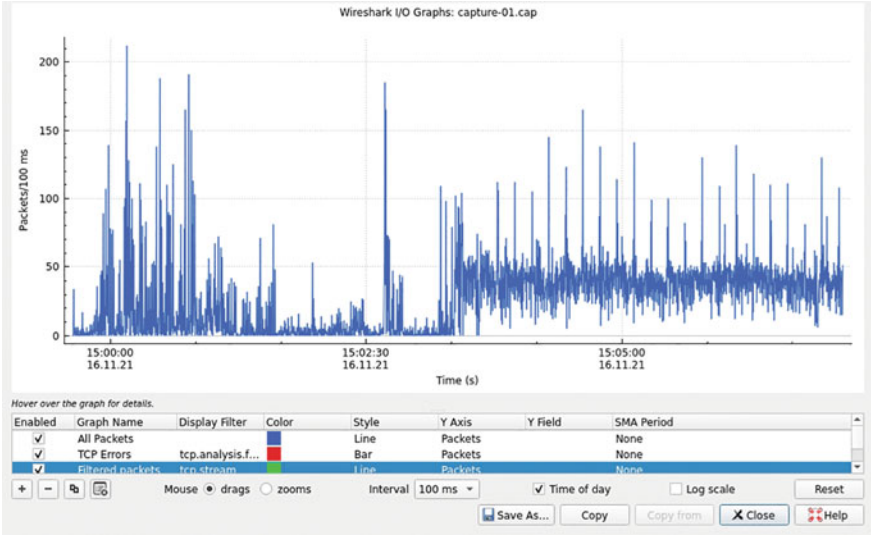


Fig. 30 Wireshark input/output traffic graph during DoS attack

DoS scenario. Cybersecurity professionals can use Wireshark statistics to identify traffic bursts when an attack occurs more quickly. Figure 31 is the shark input/output traffic graph during MiTM ARP poisoning.

Figure 32 is a Wireshark analyser showing the 4-Way Handshake. EAPoL filter is applied to obtain the 4-way handshakes.

Figure 33 is the hierarchy or tree of all captured packets, with each row showing statistical values for each protocol. The first column is the protocol's name, IEEE 802.11 wireless LAN protocol; the second column is %age of protocol packets. The third column is the protocol's total number of packets captured, which in this scenario is 115,973.

Figure 34 is a screenshot indicating a de-authentication packet number 55093 and the relative peak of the graph.

Figure 35 screenshot is the deauthentication details of specific packet number 55093. It gives details such as the wireless LAN protocol (802.11), BSS ID of the AP, and the Apple [98, 99] iPhone under deauthentication attack.

4.6 Security in 802.11ax and 802.11be; 5G and 6G

To what extent the recently released Wi-Fi 6 (and 6E) certification can become a game-changer as a countermeasure against Wi-Fi-based MiTM cyberattacks is too early to determine. Moreover, improvements in Wi-Fi 7 or 802.11be and the standardisation process are already being considered for release in 2024. The new features

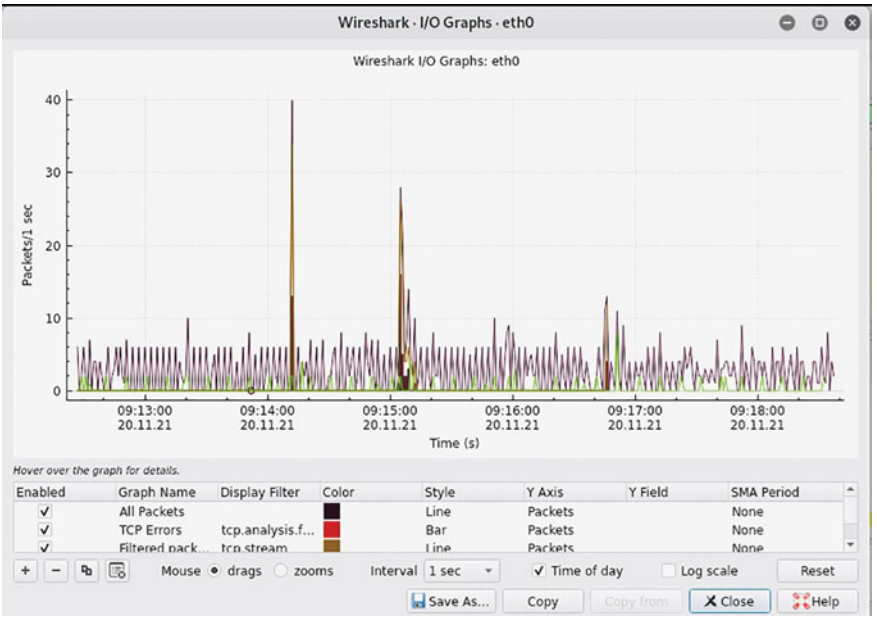


Fig. 31 Wireshark input/output traffic graph during MiTM ARP poisoning

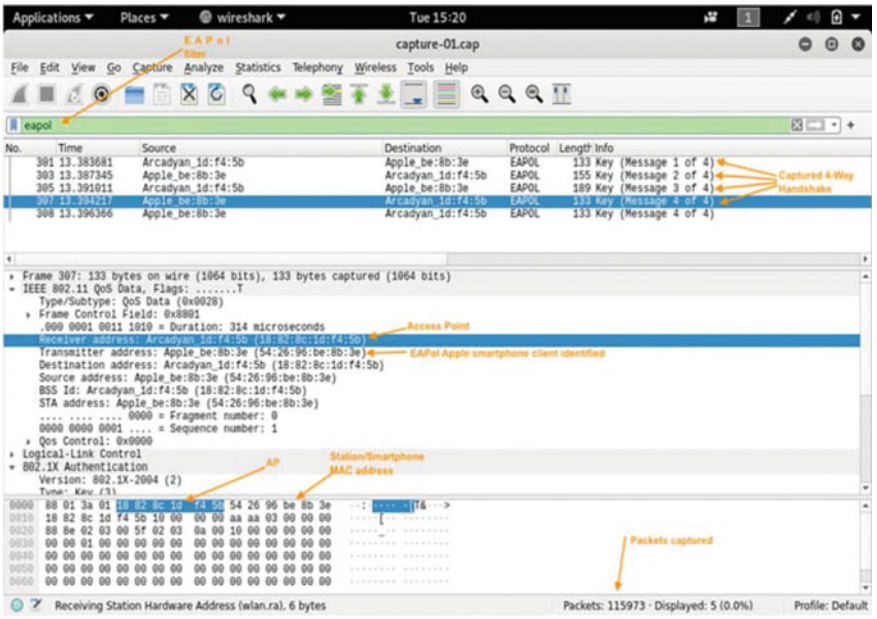


Fig. 32. 4-Way Handshake in Wireshark analyser

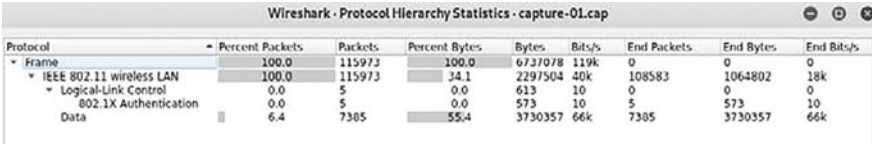


Fig. 33 Protocol hierarchy of captured packets

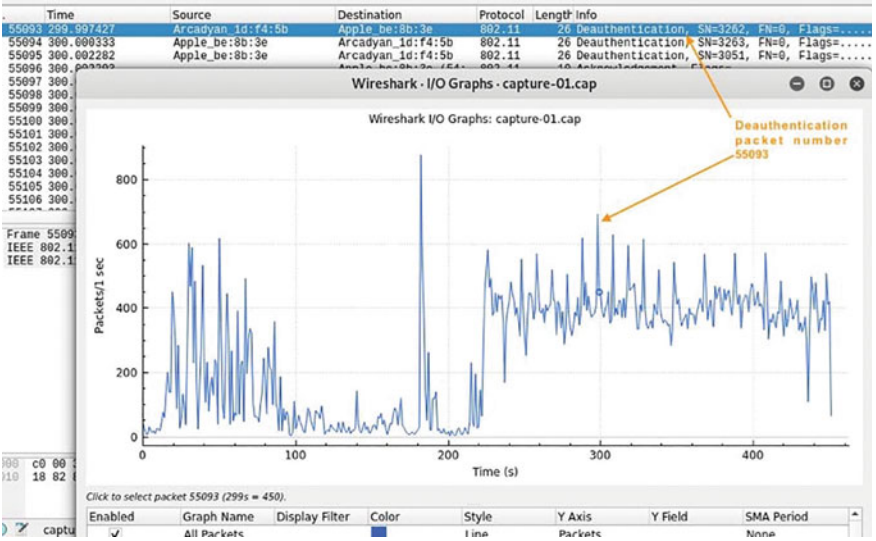


Fig. 34 Specific deauthentication packet

in Wi-Fi 7 aim to revolutionise technologies in areas such as Interactive Robotics, Virtual Reality, and Automated Vehicles. Gulasekaran and Sankaran [100] contend that Wi-Fi 6 principal objective is improved efficiency in the network with multiple access points, different traffic loads and capacity enhancements, and multiple clients. Wi-Fi 6E is the terminology and not a standard, it refers to the spectrum expansion and designation for the use of Wi-Fi 6 into the radio frequency of the 6 GHz band (Cisco, n.d.). The recently released Wi-Fi 6 and 6E will soon be surpassed by Wi-Fi 7, which is expected to deliver Extremely High Throughput (EHT) and is projected for release in 2014 according to the developmental timelines. Wi-Fi 7 aims to improve data speeds of at a minimum of 30 Gb/s per access point, about 4X faster than Wi-Fi 6, efficient operations in and backward compatibility with 2.4, 5, and 6 GHz devices. MIMO or Multi-Input, Multi-Output technology is the ability of the network to multi-task by sending data to many devices simultaneously instead of one at a time. Wi-Fi 7 improvements will include MIMO enhancements by doubling the maximum number of supported SU-MIMO (single-user MIMO) and MU-MIMO (multi-user MIMO) spatial streams per station to 16 [101].

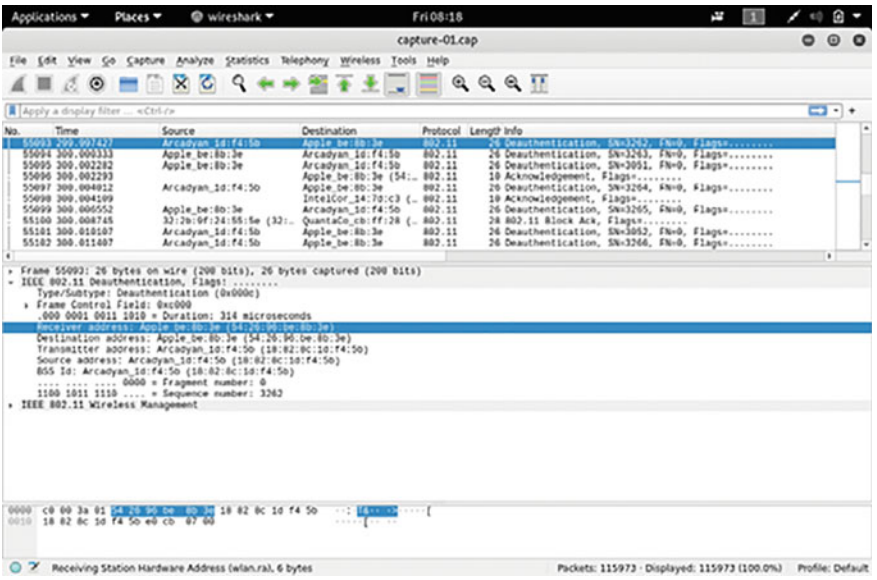


Fig. 35 Deauthentication details of packet number 55093

According to Wang et al. [102], the significant improvement of 5G technology is the facilitation of connecting the rising and challenging numbers of devices like smartphones and IoT connecting to networks. 5G technology will accommodate simultaneous "high-quality services", making networks more dynamic. 6G networks are intended to give low latency (6G radio latency is 0.1 ms or 10% of 5G), higher reliability, efficient and secure transmission services, and will have "AI-empowered" capabilities. Though 5G systems are compliant with IoT, 6G networks will work with IoE and be decentralised with the capacity to make intelligent decisions. However, the large quantum of devices and services can overload and overwhelm networks that can lead to increased vulnerabilities and more cyberattacks such as DoS and MiTM. There are security and privacy issues such as access control, authentication, encryption, and malicious behaviour concerns due to many diverse application scenarios and business types using 5G and 6G networks. Blockchain and machine learning technologies could assist in the prediction of incoming cyberattacks [102].

Security is a critical component of wireless networks, whether in Wi-fi 6 or 7, as data is transmitted in the airwaves. Nonetheless, the prevention of unauthorised data access and tampering will be dependent upon the data confidentiality and integrity mechanisms of the Wi-Fi Protected Access 3 protocol and further future security enhancements.

4.7 Wi-Fi and Human Health

Wi-Fi network traffic has grown over recent years and is projected to increase further, coupled with a considerable increase in the number of smartphones and other devices with Wi-Fi installed accessing the Internet [103]. Whereas only 9% of 55 to 64 years olds used a smartphone in 2012, this rose to 87% by 2020 [104]. Among older adults of 64 + years, smartphones are used for varied social and non-social reasons, and research done in this area suggests that it contributes towards self-control, emotional gain, social influence, self-control, loneliness, and fear of missing out [105]. However, smartphones have been "associated with excessive dependency" and "nomophobia", which is fearing the inability to avail oneself of their smartphone. For many smartphone users, the device has become an extension of the individual and may have become an "addiction" to the smartphone [106].

However, what is generally not discussed is the potential effects of Wi-Fi on human health, although extensive literature exists on the subject. This section does not form part of this original research proposal; however, it is worth discussing, albeit briefly. Research by Pall [107] contends that as Wi-Fi use becomes more and more common, so does the increased exposure to potential Wi-Fi health effects, considering that many individuals could be unsheltered to Wi-Fi fields for 4 to 8 + hours daily. The researcher argues that multiple peer-reviewed scientific research has demonstrated that Wi-Fi engenders sperm/testicular damage, neuropsychiatric effects like electroencephalographic (EEG) changes, cellular DNA damage, oxidative stress, apoptosis, endocrine changes, and calcium overload in human beings as well as in animals. Additionally, the author argues that each of the effects may also be generated by other microwave frequencies or EMF (electromagnetic frequencies). The author suggests that the use of aluminium mesh wire will aid in reflecting the impact of EMFs and, hence lowering the possible effects.

5 Conclusion and Future Work

With the ever-increasing growth in the use of Wi-Fi technology in volume and frequency, especially among smartphone users, so is the urgent need to mitigate risks against cyberattacks involving users' PII and financial data breaches. DoS and MiTM cyberattacks against data in transmission through the air medium cannot be made 100% safe, "these risks cannot be removed entirely" [108]. This paper started with a review of existing literature encompassing a brief history of the evolution of financial institutions, the definition of and what MiTM entails, and the security vulnerabilities and attacks on mobile banking and trading apps. This is followed by cyberattack vectors, methods and technics employed during the COVID-19 pandemic and their successes. Blockchain and self-sovereign identity systems are the novel technologies being employed to address cyberattacks; these are discussed. However, SSI technology is still in its infancy without a universally agreed standard framework or

protocol. The new security features in the Wi-Fi WPA3 protocol and the recently discovered Dragonblood vulnerability is extensively reviewed. The Dragonblood attack is currently the most critical vulnerability in the WPA3 security protocol requiring immediate attention and correction before WPA3 enabled devices become widely available for use.

Research methodologies and philosophical underpinnings are discussed, followed by evaluating the different and popular frameworks available in cybersecurity domain research. The paper posits the Kali Linux in a virtual environment as the most favourable framework to utilise for this project. DoS and MiTM ARP poisoning attack scenarios are demonstrated against iPhone smartphone clients and a British Telecom router (access point) using Aircrack-ng suite of tools and Ettercap software within Kali. The resulting DoS and MiTM attack data are captured in a capture file and used for data analysis. Wireshark Statistics tool combined with cybersecurity data visualisation is utilised as the method to view data captured during attack simulations. Data visualisations provide security experts with the ability to quickly identify malicious threat activity, anomalies, and business threat intelligence. However, data visualisation also has implications for data governance and data management. The new sixth-generation or Wi-Fi 6 (and 6E) based on 802.11ax standard could be a game-changer as a countermeasure against MiTM cyberattacks; this is discussed.

In terms of future work will be beneficial to replicate Dragonblood pen testing attacks on WPA3 systems as discovered by researchers Vanhoef, Piessens and Ronen. This will help ascertain to what extent such attacks can be carried out. More importantly, such future work will aid in establishing the degree of complexity an attacker requires to bypass the enhanced security features in WPA3 and then perform a downgrade attack leading to DoS and MiTM exploits on smartphones. A successful attack on the new WPA3 protocol requiring a high level of sophisticated laboratory setup would imply that WPA3 cannot be easily breached ordinarily by hackers. Therefore, the latest security features in WPA3 are working better than its predecessors, WPA2. Future work will also involve obtaining permission from equity trading platforms (i.e., IG, Interactive Brokers, FinecoBank, and Saxo Markets) for pen-testing smartphone and Wi-Fi 6 DoS and MiTM attack scenarios.

Furthermore, as a critical countermeasure against DOS and MiTM attacks, the implementation and use of data protection protocols like VPN technology in all operating systems and devices using Wi-Fi as the means of communication will exceptionally "provide data confidentiality, integrity, and origin authentication across untrusted networks such as the internet" [109]. All smartphones sold to the public should have VPN software pre-installed in client devices. IPsec (IP Security) and IKE (Internet Key Exchange) VPNs should be incorporated in all systems at the business or organisational level. The adoption of VPNs in conjunction with implementing the new WPA3 and Wi-Fi 6 standards will significantly improve data security

References

1. Khorov E, Levitsky I, Akyildiz IF (2020) Current status and directions of IEE 802.11be, the future Wi-Fi 7. Available online at: https://ieeexplore.ieee.org/document/9090146. [Accessed on: 31st Dec 2021]
2. Chauhan S, Sharma A, Pandey S, Rao KN, Kumar P (2021) IEEE 802.11be: A review on Wi-Fi use case. Available online at: https://ieeexplore.ieee.org/document/9596344. [Accessed on: 29th Nov 2021]
3. Artech House (2021) 'Wi-Fi 6 protocol ad network'—New book release by Artech House. Available online at: https://artechhouse.prowly.com/152110-wi-fi-6-protocol-and-network-new-book-release-by-artech-house. [Accessed on: 22nd Nov 2021]
4. Internet Society (2020) Fact sheet: machine-in-the-middle attacks. Available online at: https://www.internetsociety.org/resources/doc/2020/fact-sheet-machine-in-the-middle-attacks/. [Accessed on: 8th Sept 2021]
5. Conrad E, Misenar S, Feldman J (2014) Chapter 7—Domain 7: operations security. Available online at: https://www.sciencedirect.com/topics/computer-science/session-hijacking. [Accessed on: 8th Sept 2021]
6. Australian Cyber Security Centre (2020) Person-in-the-middle. Available online at: https://www.cyber.gov.au/acsc/view-all-content/glossary/person-middle. [Accessed on: 8th Sept 2021]
7. Choi H, Kwon H, Hur J (2015) A secure OTP algorithm using smartphone application. Available online at: https://ieeexplore.ieee.org/document/7182589. [Accessed on: 29th Sept 2021]
8. Ahmad DRM, Dubrawsky I, Flynn H et al (2002) Session hijacking. Hack proofing your network. Available online at: https://www.sciencedirect.com/topics/computer-science/man-in-the-middle-attack. [Accessed on: 8th Oct 2021]
9. Hilliard AG (2010) Kemet (egypt) historical revision: Implications for cross-cultural evaluation and research in education. Available online at: https://reader.elsevier.com/reader/sd/pii/S0886163389800480?token=ED42893A8B31F4062E6F49A8387994C9AB328CF74C1C52373D896730257ABFE3F7F5F110A4CCA47BD8DAF27324BC7638&originRegion=eu-west-1&originCreation=20211024102646. [Accessed on: 24th Oct 2021]
10. Labate V (2016) Banking in the Roman World. Available online at: https://www.worldhistory.org/article/974/banking-in-the-roman-world/. [Accessed on: 24th Oct 2021]
11. Henry JF (2002) The social origins of money: The case of Egypt. Available online at: https://www.csus.edu/indiv/h/henryjf/PDFS/Egypt.PDF [Accessed on: 24th Oct 2021]
12. Thakor A (2020) Fintech and banking: What do we know? Available online at: https://reader.elsevier.com/reader/sd/pii/S104295731930049X?token=29DD3820FA5696C91B4F5B624E2E6B76CA0311E57B9A71EBF3E38ECC25A6A4078DBBDDAB48B0406850FAE3885608D1E1&originRegion=eu-west-1&originCreation=20211024123857 [Accessed on: 24th Oct 2021]
13. Frame WS, Wall LD, White LJ (2018) Technological change and financial innovations in banking: Some implications for banking. Available online at: https://poseidon01.ssrn.com/delivery.php?ID=936024072000088007124097104085070011037045010029091017099126060018014112005116119102042034085030032003037019022011074028116005112067029071017063124066081006109073068084031076125092004076110106120080012097016104068074003020&EXT=pdf&INDEX=TRUE. [Accessed on: 25th Oct 2021]
14. Bhushan B, Sahoo G, Rai AK (2018) Man-in-the-middle attack in wireless and computer networking—a review. Available online at: https://ieeexplore.ieee.org/document/8344724. [Accessed on: 25th Oct 2021]
15. Malik A, Ahsan AShahadat MMZ, Tsou JC (2019) Understanding man-in-the-middle through a survey of the literature. Available online at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiljOzTzdb0AhWNGewKHTdaA28QFnoECAIQAQ&url=https%3A%2F%2Fpdfs.semanticscholar.org%

2Fc2c7%2F182b3fce4003e4dff71c0ed85e0a34aaf830.pdf&usg=AOvVaw0rUR5MgqsKmLdBetYaI182. [Accessed on: 9th Dec 2021]
16. Oriyano S-P, Shimonski R (2012) Mobile Attacks. Available online at: https://www.sciencedirect.com/topics/computer-science/man-in-the-middle-attack. [Accessed on: 25th Oct 2021]
17. Su Z, Wang H, Wang H, Shi X (2021) A financial data security sharing solution based on blockchain technology and proxy re-encryption technology. Available online at: https://ieeexplore.ieee.org/document/9332363. [Accessed on 25th Oct 2021]
18. ICO (2018) Guide to the general data protection regulation (GDPR). Available online at: https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf. [Accessed on: 26th Oct 2021]
19. Gov.uk (2018) Data Protection Act 2018. Available online at: https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted. [Accessed on: 20th Oct 2021]
20. FCA (2018) Effective global regulation in capital markets. Available online at: https://www.fca.org.uk/news/speeches/effective-global-regulation-capital-markets. [Accessed on: 24th Oct 2021]
21. Carnegie (2021) Timeline of cyber incidents involving financial institutions. Available online at: https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline [Accessed on 24th Oct 2021]
22. Zheng X, Pan L, Yilmaz E (2017) Security analysis of modern mission-critical android mobile applications. Available online at: https://www.researchgate.net/publication/312428641_Security_analysis_of_modern_mission_critical_android_mobile_applications. [Accessed on 23rd Mar 2021]
23. Ciscomag (2020) Half of mobile banking apps are vulnerable to fraud data theft. Available online at: https://cisomag.eccouncil.org/flaws-in-mobile-banking-apps/. [Accessed on: 19th Sept 2020]
24. Coker J (2020) Widespread security vulnerabilities in mobile banking apps. Available online at: https://www.infosecurity-magazine.com/news/security-vulnerabilities-mobile/. [Accessed on: 30th Mar 2021]
25. IBM (2021) IBM X-Force threat intelligence index. Available online at: https://www.ibm.com/security/data-breach/threat-intelligence. [Accessed on: 13th Sept 2021]
26. Kaspersky (2020) Top 7 mobile security threats in 2020. Available online at: https://usa.kaspersky.com/resource-center/threats/top-seven-mobile-security-threats-smart-phones-tablets-and-mobile-internet-devices-what-the-future-has-in-store. [Accessed on: 15th Mar 2020]
27. Kaspersky (2021) Kaspersky threat intelligence. Available online at: https://www.kaspersky.com/enterprise-security/threat-intelligence. [Accessed on: 16th Mar 2021]
28. Insights.com (2021) Insights external threat protection (ETP) suite. Available online at: https://intsights.com/products. [16th Mar 2021]
29. OWASP (2016) OWASP mobile top 10. Available online at: https://owasp.org/www-project-mobile-top-10/. [Accessed on: 7th Dec 2021]
30. Garg S, Baliyan N (2021) Comparative analysis of Android and iOS from security viewpoint. Available online at: https://www.sciencedirect.com/science/article/abs/pii/S1574013721000125. [Accessed on: 13th Mar 2021]
31. CVEDetails (2021) The ultimate security vulnerability data source. Available online at: https://www.cvedetails.com [14th Mar 2021]
32. WHO.int (2021) Naming the coronavirus disease (COVID-19) and the virus that causes it. Available online at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it#:~:text=Official%20names%20have%20been%20announced,%2DCoV%2D2). [Accessed on: 3rd Nov 2021]
33. Lallie HS, Shepherd LA, Nurse JRC, Erola A, Epiphaniou G, Maple C, Bellekens X (2021) Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. Available online at: https://www.sciencedirect.com/science/article/pii/S0167404821000729. [Accessed on: 12th Oct 2021]

34. Sharma R, Sharma N, Mangla M (2021) An analysis and investigation of infostealers attacks during COVID'19: A case study. Available online at: https://ieeexplore.ieee.org/document/9478163. [Accessed on: 16th Oct 2021]
35. WHO (2021) Coronavirus disease (COVID-19) update. Available online at: https://www.who.int/bangladesh/emergencies/coronavirus-disease-(covid-19)-update#:~:text=On%20this%20website%20you%20can,on%2031%20December%202019. [Accessed on: 12th Oct 2021]
36. Hiscox (2021) The Hiscox cyber readiness report 2021. Available online at: https://www.hiscox.co.uk/cyberreadiness. [Accessed on: 13th Oct 2021]
37. Statista3 (2020) Most prevalent banking trojans worldwide in 2020, by type. Available online at: https://www.statista.com/statistics/1238991/top-banking-trojans-worldwide/. [Accessed on: 12th Oct 2020]
38. Interpol (2020) Interpol report shows alarming rate of cyberattacks during COVID-19. Available online at: https://www.interpol.int/en/News-and-Events/News/2020/INTERPOL-report-shows-alarming-rate-of-cyberattacks-during-COVID-19. [Accessed on 12th Oct 2021]
39. WHO (2020) Attacks on health care in the context of COVID-19. Available online at: https://www.who.int/news-room/feature-stories/detail/attacks-on-health-care-in-the-context-of-covid-19. [Accessed on: 12th Oct 2021]
40. Deloitte (2021) CBEST: Putting cyber defences to the test. Available online at: https://www2.deloitte.com/uk/en/pages/financial-services/articles/cbest.html. [Accessed on: 15th Mar 2021]
41. Cole E (2013) The changing threat. Available online at: https://www.sciencedirect.com/topics/computer-science/advanced-persistent-threat. [Accessed on: 12th Oct 2021]
42. NCSC (2020) Advisory: APT29 targets COVID-19 vaccine development. Available online at: https://www.ncsc.gov.uk/files/Advisory-APT29-targets-COVID-19-vaccine-development.pdf. [Accessed on: 12th Oct 2021]
43. CPS (2019) Cybercrime—prosecution guidance. Available online at: https://www.cps.gov.uk/legal-guidance/cybercrime-prosecution-guidance. [Accessed on: 12th Oct 2021]
44. NCSC (2021) Using TLS to protect data. Available online at: https://www.ncsc.gov.uk/guidance/using-tls-to-protect-data. [Accessed on: 24th Sept 2021]
45. Cucko S, Turkanovic M (2021) Decentralized and self-sovereign identity: Systematic mapping study. Available online at: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9558805. [Accessed on: 13th Oct 2021]
46. Susukailo V, Opirskyy I, Vasylyshyn S (2021) Analysis of the attack vectors used by threat actors during the pandemic. Available online at: https://ieeexplore.ieee.org/document/9321897. [Accessed on: 25th Oct 2021]
47. Azourlt (2021) Azourlt. Available online at: https://any.run/malware-trends/azorult. [Accessed on: 17th Oct 2021]
48. Microsoft.com (2021) Microsoft 365.Bring out your best in school, work, and life. Available online at: https://www.microsoft.com/en-us/microsoft-365. [Accessed on: 17th Oct 2021]
49. Fartitchou M, Makkaoui KE, Kannouf N, Allali ZE (2020) Security on blockchain technology. Available online at: https://ieeexplore.ieee.org/document/9199622. [Accessed on: 25th Oct 2021]
50. Rivest RL, Shamir A, Adleman L (1978) A method for obtaining digital signatures and public-key cryptosystems. Available online at: https://web.williams.edu/Mathematics/lg5/302/RSA.pdf. [Accessed on: 25th Oct 2021]
51. Orcutt M (2019) Once hailed as unhackable, blockchains are now getting hacked. Available online at: https://www.technologyreview.com/2019/02/19/239592/once-hailed-as-unhackable-blockchains-are-now-getting-hacked/. [Accessed on: 25th Oct 2021]
52. Bandara E, Liang X, Foytik P, Shetty S, Zoysa KD (2021) A blockchain and self-sovereign identity empowered digital identity program. Available online at: https://ieeexplore.ieee.org/document/9522184. [Accessed on: 13th Oct 2021]
53. NIST (2017) NIST special publication 800–63–3. Digital identity guidelines. Available online at: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-63-3.pdf. [Accessed on 10th Sept 2021]

54. Bokkem DV, Hageman R, Koning G, Nguyen L, Zarin N (2019) Self-sovereign identity solutions: The necessity of Blockchain technology. Available online at: https://arxiv.org/pdf/1904.12816.pdf [Accessed on: 14th Oct 2021.]
55. Stockburger L, Kokosioloulis G, Mukkamala A, Mukkamala RR, Avital M (2021) Blockchain-enabled decentralised identity management: The case of self-sovereign identity in public transport. Available online at: https://www.sciencedirect.com/science/article/pii/S2096720921000099#bib14. [Accessed on: 14th Oct 2021]
56. Stokkink Q, Pouwelse J (2018) Deployment of a blockchain-based self-sovereign identity. Available online at: https://arxiv.org/pdf/1806.01926.pdf. [Accessed on: 14th Oct 2021]
57. Zhang P, Schmidt DC, White J, Dubey A (2019) Chapter seven – Consensus mechanisms and information security technologies. Available online at: https://www.sciencedirect.com/science/article/abs/pii/S0065245819300245 [Accessed on: 14th Oct 2021]
58. Bhattacharya MP, Zavarsky P, Butakov S (2021) Enhancing the security and privacy of self-sovereign identities on Hyperledger Indy blockchain. Available online at: https://ieeexplore.ieee.org/abstract/document/9297357. [Accessed on: 14th Oct 2021]
59. Aggarwal S, Kumar N (2020) Chapter sixteen—Hyperledger. Available online at: https://www.sciencedirect.com/science/article/abs/pii/S0065245820300711. [Accessed on: 15th Oct 2021]
60. Zhang B, Zhang T, Yu Z (2018) DDoS detection and prevention based on artificial intelligence techniques. Available online at: https://ieeexplore.ieee.org/document/8322748/references#references. [Accessed on: 13th Sept 2021]
61. Li X, Zhang T (2017) An exploration on artificial intelligence application: From security, privacy and ethic perspective. Available online at: https://ieeexplore.ieee.org/document/7951949/authors#authors [Accessed on: 13th Sept 2021]
62. Anandshree N, Kh J, De T (2016) Distributed denial of service attack detection *using Naïve Bayes Classifier through info gain feature selection.* Available online at: https://www.researchgate.net/publication/309638524_Distributed_denial_of_service_attack_detection_using_Naive_Bayes_Classifier_through_Info_Gain_Feature_Selection. [Accessed on: 13th Sept 2021]
63. Yuan X, Li, C, Li X (2017) Deep Defense: Identifying DDoS attack via deep learning. Available online at: https://ieeexplore.ieee.org/abstract/document/7946998/authors#authors. [Accessed on: 13th Sept 2021]
64. Verisign (2018) Q1 2018 DDoS trends report: 58 per cent of attacks employed multiple attack types. Available online at: https://blog.verisign.com/security/q1-2018-ddos-trends-report-58-percent-of-attacks-employed-multiple-attack-types/. [Accessed on: 13th Sept 2021]
65. Wi-Fi Alliance (2021) Discover Wi-Fi. security. Available online at: https://www.wi-fi.org/discover-wi-fi/security [Accessed on 24th Oct 2021]
66. Vanhoef M, Ronen E (2019a) Cryptology ePrint Archive: Report 2019/383. Available online at: https://eprint.iacr.org/2019/383. [Accessed on: 13th Nov 2021]
67. Vanhoef M (2017) Key reinstallation attacks. Breaking WPA2 by forcing nonse reuse. Available online at: https://www.krackattacks.com. [Accessed on: 27th Oct 2021]
68. NSA (2018) Cybersecurity report. WPA3 will enhance Wi-Fi security. Available online at: https://media.defense.gov/2019/Jul/16/2002158109/-1/-1/0/CTR-CYBERSECURITY-TECHNICAL-REPORT-WPA3.PDF. [Accessed on: 25th Oct 2021]
69. Shanley A (2010) *Penetration testing frameworks and methodologies:* a comparison and evaluation. Available online at: https://ro.ecu.edu.au/theses_hons/1553. [Accessed on: 21st Sept 2021]
70. NIST (2008) Technical guide to information security testing and assessment. Available online at: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-115.pdf. [Accessed on: 29th Oct 2021]
71. ISECOM (2010) OSSTMM 3. The open-source security testing methodology manual. Available online at: https://www.isecom.org/OSSTMM.3.pdf. [Accessed on: 29th Oct 2021]
72. Rounsavall R (2017) Storage area networking security devices. Available online at: https://www.sciencedirect.com/topics/computer-science/open-source-security-testing-methodology-manual. [Accessed on 29th Oct 2021]

73. Wilhelm T (2010) Methodologies. Available online at: https://www.sciencedirect.com/topics/computer-science/open-source-security-testing-methodology-manual [Accessed on: 29th Oct 2021]
74. Pentest-standard.org (2014) PTES. Available online at: http://www.pentest-standard.org/index.php/Main_Page. [Accessed on: 29th Oct 2021]
75. Faircloth J (2017) Wireless penetration testing. Available online at: https://www.sciencedirect.com/topics/computer-science/penetration-testing [Accessed on: 29th Oct 2021]
76. Knowles W, Baron A, McGarr (2016) The simulated security assessment ecosystem: Does penetration testing need standardisation? Available online at: https://www.sciencedirect.com/science/article/pii/S0167404816300906#fn0055. [Accessed on: 29th Oct 2021]
77. Gantz SD (2014) Audit-related organization, standards, and certifications. Available online at: https://www.sciencedirect.com/topics/computer-science/open-web-application-security-project. [Accessed on: 29th Oct 2021]
78. Kritikos K, Magoutis K, Papoutsakis M, Ioannidi S (2019) A survey on vulnerability assessment tools and databases for cloud-based web applications. Available online at: https://www.sciencedirect.com/science/article/pii/S2590005619300116. [Accessed on: 29th Oct 2021]
79. Holik F, Horalek J, Marik O, Neradova S, Zitta S (2015) Effective penetration testing with Metasploit framework and methodologies. Available online at: https://ieeexplore.ieee.org/abstract/document/7028682 [Accessed on: 29th Oct 2021]
80. Rapid7 (2021) Metasploit modules and locations. Available online at: https://www.offensive-security.com/metasploit-unleashed/modules-and-locations/ [Accessed on: 29th Oct 2021]
81. Offensive Security (2021) Armitage. Available online at: https://www.offensive-security.com/metasploit-unleashed/armitage/. [Accessed on: 29th Oct 2021]
82. Filiol E, Mercaldo F, Santone A (2021) A method for automatic penetration and mitigation: a red hat approach. Available online at: https://www.sciencedirect.com/science/article/pii/S1877050921017063?via%3Dihub. [Accessed on: 6th Dec 2021]
83. Patil S, Jangra A, Bhale M, Raina A, Kulkarni P (2018) Ethical hacking: The need for cyber security. Available online at: https://ieeexplore.ieee.org/document/8391982 [Accessed on: 6th Dec 2021]
84. Bacudio AG, Yuan X, Chu BTB, Jones M (2011) An overview of penetration testing. Available online at: https://www.researchgate.net/publication/274174058_An_Overview_of_Penetration_Testing. [Accessed on: 8th Nov 2021]
85. Vanhoef M, Ronen E (2019) Dragonblood: Analysing the Dragonfly handshake of WPA3 and EAP-wpd. Available online at: https://eprint.iacr.org/2019/383. [Accessed on: 8th Nov 2021]
86. Conrad E, Misenar S, Feldman J (2016) Chapter 5—Domain 4: Communication and network security (designing and protecting network security). Available online at: https://www.sciencedirect.com/science/article/pii/B9780128024379000059. [Accessed on: 7th Dec2021]
87. ALPHA (2020) AWUS036NHA. Available online at: https://www.alfa.com.tw/products/awus036nha?variant=36473966166088. [Accessed on: 13th Nov 2021]
88. Vanhoef M, Piessens F (2017) Key reinstallation attacks: forcing nonce reuse in WPA2. Available online at: https://papers.mathyvanhoef.com/ccs2017.pdf. [Accessed on: 24th Nov 2021]
89. Wireless Hacking (2004) A brief overview of the wireless world. Available online at: https://www.sciencedirect.com/topics/computer-science/key-authentication. [Accessed on: 24th Nov 2021]
90. Al-Rushdan H, Shurman M, Alnabelsi SH, Althebyan Q (2019) Zero-day attack detection and prevention in software-defined networks. Available online at: https://ieeexplore.ieee.org/document/8991124. [Accessed on: 22nd Nov 2021]
91. Johansen HD, Johansen D, Renesse RV (2007) FirePatch: secure and time critical dissemination of software patches. Available online at: https://www.researchgate.net/publication/220722547_FirePatch_Secure_and_Time-Critical_Dissemination_of_Software_Patches. [Accessed on: 22nd Nov 2021]
92. Sherman R (2015) Advanced analytics. Available online at: https://www.sciencedirect.com/topics/computer-science/data-visualization-tool. [Accessed on: 19th Nov 2021]

93. Shealy M (2021) How data visualization helps prevent cyberattacks. Available online at: https://www.klipfolio.com/blog/how-data-visualization-prevents-cyber-attacks. [Accessed on: 2nd Dec 2021]
94. Ryan L (2016) Data visualization as a core competency. Available online at: https://www.sciencedirect.com/topics/computer-science/data-visualization-tool. [Accessed on 19th Nov 2021]
95. Cobb C, Sudar S, Reiter N, Anderson R, Roesner F, Kohno T (2017) Computer security for data collection technologies. Available online at: https://www.sciencedirect.com/science/article/pii/S2352728516300677. [Accessed on: 19th Nov 2021]
96. Bank of England BoE (2021) CBEST threat intelligence-led assessments. Available online at: https://www2.deloitte.com/uk/en/pages/financial-services/articles/cbest.html. [Accessed on: 15th Mar 2021]
97. Bulletproof (2019) Bulletproof annual cyber security report 2019. Available online at: https://www.bulletproof.co.uk/industry-reports/2019.pdf. [Accessed on: 13th Sept 2021]
98. Apple (2007) Apple reinvents the phone with iPhone. Available online at: https://www.apple.com/uk/newsroom/2007/01/09Apple-Reinvents-the-Phone-with-iPhone/. [Accessed on: 8th Sept 2021]
99. Apple (2021) iPhone 12 and iPhone 12 mini. Available online at: https://www.apple.com/iphone/. [Accessed on: 8th Sept 2021]
100. Gulasekaran SR, Sankaran SG (2021) Wi-Fi 6 protocol and network. Artech House, Norwood, MA
101. Garcia-Rodriguez A, Lopez-Perez A, Galati-Giordano L, Geraci G (2021) IEEE 802.11be: Wi-Fi 7 strikes back. Available online at: https://ieeexplore.ieee.org/document/9433521. [Accessed on: 29th Nov 2021]
102. Wang M, Zhu T, Zhang T, Zhang J, Yu S, Zhou W (2020) Security and privacy in 6G networks: new areas and new challenges. Available online at: https://www.sciencedirect.com/science/article/pii/S2352864820302431. [Accessed on: 1st Dec 2021]
103. Cisco (2020) Cisco annual internet report (2018–2023). White paper. Available online at: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. [Accessed on: 13th Sept 2021]
104. Statista (2021) Number of smartphone users worldwide from 2016 to 2021 (in billions). Available online at: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. [Accessed on 13th Mar 2021]
105. Busch PA, Hausvik GI, Ropstad OK, Patterson D (2021) Smartphone usage among older adults. Available online at: https://www.sciencedirect.com/science/article/pii/S0747563221001060 [7th Sept 2021]
106. Fryman S, Romine W (2021) Measuring smartphone dependency and exploration of consequences and comorbidities. Available online at: https://www.sciencedirect.com/science/article/pii/S2451958821000567. [Accessed on 7th Sept 2021]
107. Pall ML (2018) Wi-Fi is an important threat to human health. Available online at: https://www.sciencedirect.com/science/article/pii/S0013935118300355. [Accessed on: 29th Nov 2021]
108. Kovacic S, Dulic E, Sehidic A (2017) Improving the security of access to network resources using the 802.1x standard in wired and wireless environments. Available online at: [Accessed on 16th Oct 2021]
109. Kleidermacher D, Kleidermacher M (2012) Data protection protocols for embedded systems. Available online at: https://www.sciencedirect.com/topics/engineering/virtual-private-networks. [Accessed on: 2nd Dec 2021]

Digital Transformation, Leadership, and Markets

Aysha Kattakath Mulangat Hydros and Umair B. Chaudhry

Abstract In the electronic era, the company environments are evolving into an insecure, composite, and indecisive atmosphere. This quick evolution can be blamed upon the rise of technology, competitive market, and legal and regulatory compliances. Teichert (Acta Univ Agriculturae Silviculturae Mendelianae Brunensis 67:1673–1687, [1]) observe this condition as the reason for forcing the business plans and policies of organizations to adapt to technological innovations. According to Gartner (no date), digital transformation could be termed as the evolution of technology as well as digital optimization including the discovery of advanced digital business models. Gartner also suggests labelling digital transformation as 'digitization' since nationalized establishments use the term digital transformation to mention normal IT practices like usage of online services. Vaughan [2] suggests the main four areas to consider while an organization plan on a digital transformation project as process transformation, business model transformation, domain transformation, cultural/organizational transformation. The benefit of a successful digital transformation according to the author is improved processes, fostering collaborations, broadening service options, and transforming the customer experience.

Keywords Digital transformation · Digital leadership · Artificial Intelligence · Machine Learning · Digital Twin · Governance · Blockchain · Zero trust · Cyber Security · AR · VR · Ransomware

1 Digital Leadership

Successful digital transformation can be achieved by the support from top-level management, incorporating company business strategies into the transition process [3]. Nowadays, the word leadership is a collective and strategical term used to define the work culture, its employees and company environment [4]. Wagire et al. [5] identify the use of maturity models in organizations to evaluate the transition process.

A. K. M. Hydros · U. B. Chaudhry (✉)
Northumbria University London, London, UK
e-mail: umair.chaudhry@northumbria.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_8

This evaluation within the company indicates the dawn of endless innovative digital transformation technologies. The usage of maturity models for such assessments offers a methodical approach for widely accepted standards and advanced productivity [6]. According to Reinhardt [7], a mature digital transformation cannot be achieved through technological innovations alone. Alongside, companies should consider addressing the parallel consequences of transition processes and each level of company executives must be made aware of every aspect of the transition from an organisational point of view and on developing supervisory and personal skills at each level of executives. The author introduced a digital competence management model which can be used by companies to use instead of self-organised methods.

Further, Beresford [8] considers four prototypes that could be considered in this electronic era: teamwork among CEOs, transactional leadership among CEOs to enhance performance, transformational methods through CEOs contributing innovative ideas, and comparing similar practices. Whereas Jensen et al. [9] put forward a combination of transactional, transformational, authentic leadership and experimental conduct to develop a digital leadership model. The author suggests using more than one maturity model for leadership models for enhanced efficiency. The digital leadership conceptual model proposed by Prince and Ann [10] combines authentic, transactional, and transformational leadership perspectives into a single leadership module which can be made use of deriving digital imperatives from it, which in turn paves the way for produce the most appropriate digital transformation. The findings from a study conducted by Mihardjo and Furinto [11] revealed that both digital leadership and innovation management have an impact on long-term competitive advantage and digital disruption, with digital leadership having a bigger impact than innovation management. As per Singh's research [12], digital leadership is based on three pillars: an agile attitude, leading transformation and driving business. Digital leaders must combine two seemingly opposing abilities: the speed with which new value is created and the ability to scale organisations. This 'speed + scale' approach reflects the market's increasing competitiveness. Organizations must frequently explore beyond standard career pathways to uncover such "bilingual" executives Mcmanus [13].

As stated by Kokot et al. [14], the behaviour of executives along with their vision towards attaining digital maturity has drastically changed with the evolution of digital transformation. Researchers identify technology, strategy, humans, and the effort for dominance in the industry as the major factors that label the pyramidal connection between digital transformation, leadership, and maturity. Borowska [15] recommends digital leaders consider implementing policies to employ accomplished and skilled individuals, encourage staff to take part in transition practices by adopting company principles to the evolving market and by efficiently using digital resources. This is backed up by Gartner [16], listing certain successful traits of digital leaders as those who are ready to take on new challenges, who have clear ideas on implementing innovative ideas and using existing solutions, who abstain from business limits, and who are technology enthusiasts. The willingness of digital executives to challenge the organization's existing models and practices is critical to generating a new corporate strategy. Builders, Catalysts, Explorers, and Connectors are four

types of behaviours used by the most successful digital leaders, states Mcmanus [13]. The new core of leadership is digital leadership. These are the skills that any firm intending to succeed in the digital economy—and all leaders seeking to lead the way—must possess.

2 Cyber Security Defence Including Ransomware Attacks for Small and Medium-Sized Businesses in the Digital Economy

Conway and Codkind [17] state that the leaders of large companies can overcome security challenges when compared to small and medium-sized organizations. It is because of the lack of business analyst specialists in the leader's team and their poor decision-making skills. Despite the benefits and opportunities that digital technologies provide, and despite the huge growth in uptake in recent years, many SMEs continue to lag in adoption, and digital adoption disparities between smaller SMEs with 10 to 49 employees have grown over the previous decade [18]. NCSC [19] warns small and middle-sized organizations to be aware of the chance of 1 in 2, stating that they are likely to face a cyber-attack. NIST [20] states that hackers are diligently targeting small organizations who possess' weak cybersecurity because of low budgets. This is proven by Hiscox [21] in a survey that disclose that cyber-attacks on UK companies rose from 40 to 55% between 2018 and 2019. Disruptions caused by such incidents are the same for small companies and large-scale organizations. Often, small-scale industries are assessed as weaker targets considering their small size.

SMEs will need advice, support, and direction from trusted sources to help cement the change, mitigate risks, and maximise the potential of the new tools [18]. A study conducted by Tam et al. [22] utilizes proof from various industries, government sources, and investigation departments to identify the factors affecting the security of small firms. Small business features such as agility, large cohort size, and fragmented IT architecture, according to the authors, can aid with cyber-security. The report concludes by suggesting that legal and policy changes be made to help small businesses become more cyber-resilient. Misleading information, according to Petratos [23], can have a significant impact on the company. According to the author, the increased frequency and complex cyberattacks, electoral tampering, and crises such as the COVID-19 have made it a crucial concern for corporations. Misleading information is used in a variety of cyberattacks. The most common sort of cyberattack against businesses is social engineering [24].

As the usage of IT technology grew, so did the use of social engineering techniques, such as baiting, pretexting, tailgating, and quid pro quo. Nowadays, most cyberattacks incorporate social engineering techniques such as baiting, pretexting, tailgating, and quid pro quo [25]. That's not to suggest small businesses are not concerned about cyber threats; according to the Cyber Security Breaches Survey [26], 78% now

prioritise cyber security. According to the research, only 15% of small businesses have a defined cyber incident management plan, indicating that greater awareness has not translated into action. The majority believe that the major goals of cyber security are risk elimination, data leakage prevention, and limiting malware and hacker risks. Nevertheless, corporate executives must recast it as a growth enabler to fully realise cyber security's potential. If small firms want to take advantage of the potential prospects, they must be able to operate safely in the internet environment. An effective cyber security plan combined with proper risk management may help companies innovate, differentiate, and grow their business. When the focus is solely on the cost, the necessity of strong cyber security might be lost in translation. The focus should be on the business benefits of a strong security posture and the prospect for growth, rather than on the expense [27].

Belitski and Liversage [28] suggest that programs that support the development of e-leadership as a skill should be considered by government entities. They should do so to encourage the development of practices that make technology-based training more accessible to small businesses. Access to less expensive or subsidised technology should become a more concentrated issue, with politicians applying price pressure to data and technology providers to provide a reduced barrier to entry. To make technology more accessible, new pricing and payment systems should be established. Small company initiatives that foster and develop invention and commercialization must continue. Small business e-leaders should take advantage of the support resources provided through online education, such as online open courses. Partnerships that link government support efforts to small enterprises and institutions on their route to commercialization can help achieve this.

3 Information Governance, Security, and e-Governance

Although digital transition helps companies to achieve enhanced efficiency, flexibility and client service, organizations must not leave aside considering governance aspects. As per Jefferies [29] article, such organizations will end up risking security obligations and data privacy. Implementing legal regulation alone does not result in proper IT governance. It comprises a collection of practices performed in nonregional areas along with people having diverse cultures and contrasting backgrounds. According to ISACA [24], the exceptional focus on digitalization happened because of the Covid-19 pandemic. As more employees and organizations moved to work from home system, the need for efficient digital transition became an emergency. A digital implementation gap might be faced by companies who failed to meet their business continuity plans, which in turn might have paved way for a digital strategy execution gap. This gap demonstrates the relevance of IT governance to ensure company sustainability.

Reis et al. [30] state that the IT security team is liable for executing an information security management system (ISMS) by protecting confidentiality, integrity, and availability of every digital asset in an organization and establishing a proper disaster

recovery plan if any breach or loss happens. The authors also recommend Information Security Officers be vigilant and report an incident to ensure security policies and standards compliance. According to Smallwood [31], IG is a total package that is accountable for every aspect of data. IG practices should focus on reducing cyber risks, maintaining total costs, and increasing productivity across all departments. An ongoing effort is needed for maintaining proper IG in the organization.

Additionally, NCSC [32] recommends following risk management standards like ISO 27001 and IEC 62443-2-1:2010 will make organizations have effective security governance. They also urge organizations to have well-structured policies and practices which includes responsibilities and accountabilities of each asset or information system. As no blueprint is available for following cyber security practices, identifying, and maintaining inventories of each asset associated with the organization and anticipating risks associated with them should be followed as a preliminary step towards achieved security goals. A systematic approach should be followed to establish recognized risks and threats are managed and thereby gain a level of confidence through auditing and monitoring. According to NIST [33], developing security awareness among employees and training them to ensure cyber readiness is a business leader's responsibility. Also, having access control policies and procedures in place, maintaining periodic data backups to prevent data loss, and developing a strategic plan to respond and recover from incidents are highlighted as a practice to build a cyber ready culture in an organization.

Even though, there are information security standards and frameworks introduced by several international organizations and bodies, digital transformation in government organizations is not so easy. To overcome this challenge, governments started the e-governance process, using Information and Communication Technology (ICT) integrated with governance practices to provide information and services to citizens. A comparative analysis performed by Nehemia et al. [34] on e-governance and IT governance identifies that the goal of deploying e-governance is classified into the following: providing easy data access to natives, keeping track of public services used by nationals and evaluating them, and support communication between government and related agencies as well as stakeholders. Researchers claim that while e-governance concentrates on using ICTs to serve its citizens, IT governance focuses on ICT standards and policies to help with the selection and implementation of ICT infrastructure and related applications. Organizations can make use of both approaches since each of them have various exclusive goals. Iyer et al. [35] addressed various challenges associated with the implementation of e-governance approaches in a study on the analysis of various models using secure information exchange. Some of the findings include organizational challenges like management change or collaboration disputes among officials, social aspects including IT illiteracy and digital divide, political aspects, and IT and regional challenges.

4 The Role of AR/ VR and Holographic Technology in Corporate Decision Making

Although, experienced managers and human input are no longer used to make business decisions. Augmented Reality (AR) is allowing businesses to use the technology to achieve a strategic advantage, which has turned the decision-making process on its head. Its tremendous growth bodes well for a bright future in enhancing organizational activities. Augmented reality generates "hyper visualisations" by gathering large amounts of data and superimposing clear and understandable 3D modules in the real world. In commercial decision-making, the value of visualising large amounts of complex data cannot be overstated [36]. According to Rokhsaritalemi et al. [37], the design of applications to generate decision-making tools to solve everyday difficulties is made possible by the advancement of new technologies. It's critical to employ appropriate visualisation tactics to build easy user interaction in such an application. Emerging computing technologies like Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) are designed to attain ease of use. VR comes up with a virtual environment in which a user can utilise a VR headset to enter and interact with a virtual environment. Even though VR had an issue with a lack of relationship with actual space, AR technology remedied this problem and introduced a new technique of visualisation that allowed the insertion of virtual content into the real environment. The user can engage with an augmented world created by this technology. Regardless of the relevance of augmented reality, separating the actual and virtual worlds is a difficult task. This issue reduces the amount of user immersion in AR settings. With the advent of the MR environment creating a window between the actual and virtual worlds by fusing them, real-world objects started interacting with a virtual entity to provide the user with practical implementations.

Dwivedi et al. [38] identify that if AR becomes ubiquitous and consumers can perceive accurate virtual AR content continuously, it is expected to have a huge impact on businesses, companies, and society. AR was employed in the operating room to improve surgeon productivity and decision-making [39]. Pantano et al. [40] researched by examining the impact of technological attributes on consumer behaviour, with a focus on the impact of AR systems on customers' online shopping decisions. The study focuses on consumers from Italy and Germany who desire to try on glasses (sunglasses or spectacles) in an e-commerce environment enriched with AR systems like the virtual try-on (smart mirror). Consumers from both countries saw the new system as a powerful instrument for assisting decision-making, with the ability to modify consumer behaviour mostly due to technological qualities like quality, interactivity, response time and quality of information. Buyers were enthusiastic about using virtual try-on to evaluate products like sunglasses and eyeglasses, which typically need physical testing before purchase. Authors hence prove that by satisfying their preferences, the virtual try-on system would take the place of the physical try-on.

After decades of technological obstacles (such as a lack of fast mobile internet, limited mobile processing power, or imprecise sensors), Rauschnabel [41] says that

recent hardware and software developments have enhanced AR's potential as a mass-market technology. For example, Apple's ARKit 3 enables high-quality AR on smartphones, and smart glasses like Microsoft's HoloLens allow users to integrate virtual 3D information straight into their field of vision without having to use their hands (holograms). Predictions and expectations are astounding, pointing to a hybrid future that includes both physical and virtual goods. Leaman [42] tested the potential applications of holograms in the healthcare industry and call holograms 'X-ray vision' which eliminates guesswork while examining a patient. Experts state that as the technology gets more firmly ingrained in business processes, we will witness a growth in MR usage in the next years. We see a rise in sector-specific MR applications, which use out-of-the-box software to tackle business problems in industries such as retail, education, manufacturing, and healthcare [43].

5 Emerging Technologies (Blockchain, Zero Trust and Digital Twin)

5.1 Blockchain Technology

Industry 4.0, or the fourth industrial revolution, could transform and improve a variety of business processes. Advanced manufacturing technologies like advanced robotics, additive manufacturing like 3D printing, augmented reality like virtual reality, simulation, cloud computing and digital platforms, industrial IoT, cyber security, big data analytics, and blockchain are all part of industry 4.0 [44]. Blockchain, according to NIST [45], is a distributed digital database of cryptographically signed transactions organised into blocks. After validation and consensus, each block is cryptographically connected to the one before it (making it tamper obvious). Older blocks become increasingly difficult to change as new blocks are added (creating tamper resistance). New blocks are replicated across network copies of the ledger, and any conflicts are resolved automatically according to predetermined criteria. As per Thuraisingham [46], blockchains primarily deal with transactions, and transactions require data. In various transactions, massive amounts of data must be collected, processed, evaluated, and disseminated.

As Mearian [47] suggests, enterprises will inevitably have to consider if and where blockchain can play a role as part of their digital transformation as they move toward hybrid cloud and cloud-native apps. While distributed ledger technology (DLT) adoption by enterprises is still in its initial phases, with most initiatives being pilots or proofs of concepts, it is evident that blockchain will interrupt businesses. The current blockchain transition envisions a future in which everyone can participate in the transfer of value, which is fuelled by transparency and trust. Business models built on centralised processes are open for competition, from banking services to social networks who will have the ability to redefine commerce by leapfrogging current frameworks. IBM [48] suggests that such businesses may have to illustrate the pros

of blockchain technologies over centralised database applications. Jabbour et al. [49] identify that blockchain technology is used by supply chain stakeholders to improve supply chain performance and enable them digitally. Digitally enabled supply chain management refers to a supply chain that is powered by information technology and makes use of cutting-edge digital technologies and online portals. Sharing correct information, monitoring supply chain operations, strengthening relationship management, and increasing sustainability practices are among the managerial implications outlined by Ebinger and Omondi [50] for the digital transformation of sustainable supply chain management.

Saini [51] states that blockchain has the potential to reduce transaction costs and transform the economy. It is critical to be prepared for the promise of Blockchain as a new development environment. All businesses where data is exchanged across several entities can benefit from blockchain technology. This is backed up by a study conducted by Li et al. [52] on how blockchain technology shapes digital transformation for financial services which proves that the usage of blockchain can help to bridge the gap between digital technology and financial services, enable the deployment of field applications, and create an organic and efficient digital transformation architecture. Additionally, blockchain technology can help logistics companies tackle their most pressing traceability issues, from the merchant's order request to the customer getting the goods. Blockchain technology may be used by logistics companies to track the full activity process of commodities and can keep track of the status of commodities at each link. It can successfully lower enterprise traceability costs, prevent data manipulation, and boost operational efficiency [53]. Despite the substantial disparity in figures, Sosin et al. [54] believe the blockchain technology business will increase. The key drivers of the industry's growth will continue to be an increase in the demand for secure online payments, as well as a desire to cut costs. Researchers demand that the blockchain is beneficial in areas where authorship of all actions and reliable data synchronisation are critical, such as banks, exchanges, insurance firms, certification centres, and so on.

5.2 Zero Trust Security Strategies and Guideline

Even though blockchain technology has shown promise in terms of cybersecurity, and several blockchain security mechanisms have been devised, including access control, user authentication, and transaction security. Integrating blockchain with a zero-trust security framework will provide highly accessible and transparent security measures via a visible blockchain, in which all transactions are visible to restricted operators (Li et al. [53]). As Stevens [55] states, the growing demand for digital transformation as businesses embrace hybrid working has made it difficult for IT professionals to maintain control over their cybersecurity strategy. With the continuous migration to the cloud, the rapid adoption of 'bring your own device' (BYOD) initiatives, and the introduction of hybrid work, zero trust digital transformation goals have accelerated. Enterprises must replace their antiquated network architectures with

zero trust-focused solutions to fully embrace these objectives. This will provide a safe and secure experience for customers and employees while saving significant sums of money in a short period of time [56]. Stevens [55] claims that the implementation of the zero-trust network is a terrific task because hackers can exploit a distant, insecure asset and move laterally into the safe internal system, remote working, and BYOD rules have added substantial hazards to company intranets. This can lead to quick-moving and extremely damaging security breaches.

In the face of a compromised network, NIST [20] defines zero trust (ZT) as a set of concepts and ideas aimed at reducing uncertainty in enforcing precise, least privilege per-request access decisions in information systems and services. Zero trust architecture is a cybersecurity plan for an organisation that incorporates zero trust concepts and includes component interactions, workflow planning, and access controls. As a result of a zero-trust architectural plan, a zero-trust enterprise is the network infrastructure (physical and virtual) and operational procedures that are in place for a company. Building context ensures that a request is trustworthy, which is dependent on good authentication, authorisation, device health, and the value of the data being accessed [57].

NSA [58] outlines the requirements organizations must consider to appropriately confront the modern dynamic threat environment as (i) System monitoring, system management, and defence operations skills that are coordinated and aggressive (ii) Assume that all-important resource requests and network traffic are malicious (iii) Assume that all devices and infrastructure are vulnerable. (iv) Accommodating that all key resource access approvals come with a risk, and being ready to undertake swift damage assessment, control, and recovery activities. NSA [58] also urges organizations to determine the organizational goals as the priority while designing a zero-trust solution. Secondly, focus on protecting important data and considering all the possible ways to control access to such data. Finally, inspecting and logging every traffic before detecting a malicious activity will help organizations to mitigate risks to an extent.

According to Gordon [59], software-defined perimeter (SDP) is the technology that the world will look for in the future when it comes to ensuring secure cloud access. SDP enables policy-based trust levels to be established, allowing users to access apps and resources with relative ease and at a greater level of perfection. SDP also offers direct application access and per-application network segmentation, reducing the attack surface for data centre and cloud applications. Pallais [60] also advises organizations to use robust authentication methods such as two-factor or multi-factor authentication (MFA), which reduces the chance of a data breach by 99.9%. Microsoft has implemented a conditional access policy that is imposed based on the device's compliance status at the time the user attempts to access data as an instant step toward zero-trust security (Fig. 1).

According to Klasnja [61], zero-trust emphasises a data-centric and risk-based approach to access from a security standpoint, which means the security programme must have the right skills to manage the access lifecycle. A programmatic strategy incorporating a mix of processes and technology is required to successfully implement Zero Trust. When used correctly, it aids in the reduction of complexity, the

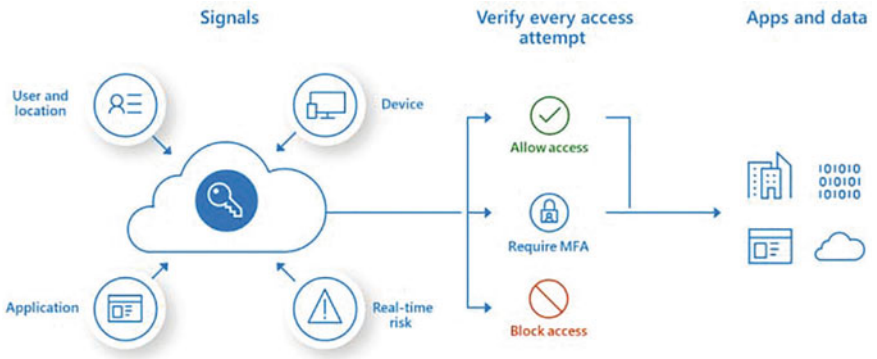


Fig. 1 Conditional Access illustration [60]

reduction of an operational burden, and the eradication of technology debt. It's a context-aware and dynamic security architecture that focuses on applications, identities, and data. Modern cloud-native security solutions extend zero-trust concepts to enable and secure WFA access to apps without requiring public exposure or costly network segmentation. Security, simplicity, and user experience all go hand in hand in this innovative method, providing for seamless access across all hybrid workforce configurations [62].

5.3 Digital Twin

The digital twin, like Zero Trust models, represents a considerable shift from legacy techniques that underpin most of the planning, design, and implementation of existing IT risk and security programmes for both large and small businesses [63]. "Imitations keep the originals alive," says Erol et al. [64]. Considering through the lens of technology, Digital Twins are a clever approximation for maintaining, augmenting, and improving their originals. Digital Twin technology is appreciated in domains such as health, transportation, and energy, and is usually associated with industry 4.0 and engineering. Working in a range of fields, governments want to employ this technology to increase their welfare and technical advancement. According to Moore [65], the digital twin is one of the five technologies projected to have the largest impact on government institutions in the next 5–10 years.

A digital twin, according to IBM [66] is a virtual representation of a physical object that is designed to exactly reflect it. The object will be equipped with a variety of sensors relevant to key areas of functionality, which will generate data on various elements of the physical object's performance. This information is subsequently sent to a processing machine, where it is applied to a digital copy. After being given this information, the virtual model may be used to run simulations, investigate

performance concerns, and suggest changes, all with the purpose of gaining important insights that can later be applied to the original physical thing. Digital twins combine IoT, AI, machine learning, and analytics with a graphic and/or spatial representation to create dynamic digital simulation models that evolve and develop in the same way that their physical counterparts do. Digital twins learn on their own, based on data gathered from professionals with deep topic knowledge and other similar assets. It improves the accuracy of its simulations by incorporating and utilising past data [67].

As stated by Fuldauer [68], testing models in a virtual mode before implementing would be cheaper and less likely to fail in the actual world. Testing and prototyping can greatly increase a city's resilience. Preparing emergency preparations and determining reaction strategies in natural catastrophes such as floods, fires, and earthquakes can assist governments in detecting pollution and landscaping. By having real-time knowledge about any emergency, it can allocate resources, organise operations, and optimise traffic. Emergency responders, like firefighters, can obtain a 3D model of the structure using it in the event of a fire. Firefighters can use augmented reality and artificial intelligence to predict where people are and how the fire will behave. Whereas Zhou et al. [69] discuss the goal of applying digital twin theory to the construction field is to create a virtual information model that can reflect all the data from the building's entire life cycle, which will not only reflect the physical entity information after completion but will also accurately, fully, and immediately reflect the status of building operation for building dynamic management.

Erol et al. [64] conducted research on the impact of the digital twin in the sphere of health. Researchers states that digital twin research has mostly focused on creating digital twins of human organs, studying cell behaviour in the body, and applying suitable medicines and therapies. Furthermore, Digital Twin technology will advance in tandem with the development of sensor, IoT, and machine learning technologies, as well as the development of image systems. Also, the digital twin will be widely used to handle challenges in the management of hospital systems and medical resources, which have grown increasingly apparent because of the COVID-19 epidemic. Similarly, digital twin technology is planned to be used to speed up the vaccination and drug development procedures associated with COVID-19.

Finally, Moller et al. [70] predict that the digital twin will have a substantial impact on intelligent, service-oriented, and green production, and will thus become a prerequisite for the manufacturing industry of the future. As mentioned by Erol et al. [64], with increased application in fields such as health, industry, and smart city management, the maturity level of digital twin technology will rise, and its dissemination will accelerate. Not only in these sectors, but also in piloting, engineering, and medical education, it will become commonplace. The dazzling advantages promised by this technology will be realised in physical form as the number of successful applications grows, and a considerable gain will be made at the point of digitization. It is also evident that the implementation of this structure would result in large gains in emerging countries, where digitalization is critical for advancing the country's development.

6 Ethics, Privacy, Safety and Security of Digital Systems

6.1 Artificial Intelligence and Machine Learning in Management

Society's digitalization pushes the bounds of our talents and opens us to a world of possibilities, but it also tests our moral or ethical boundaries. The expanding usage of technology also implies that augmented and virtual reality, as well as digital platforms, are being used to digitise human-to-human and human-to-organizational interactions. As a result, digitization has penetrated our social-cultural world: we are increasingly doing things online, such as purchasing, transactions, listening to music, contacting friends, acting, and finding a date. Most of the public and political attention is currently focused on privacy issues (particularly personal data protection) and digital security. With the rise of IoT, the primary challenges are the search for digital inviolability of the house and the preservation of privacy [71]. The European Union (EU) passed the General Data Protection Regulation (GDPR), which establishes requirements on any enterprises that target or collect data about EU citizens. Those who break the GDPR's privacy and security regulations will face severe fines, with penalties.

As per Storm and Wolk [72], profiling and automated decision-making are both restricted under the GDPR. Profiling and occasionally automated decision-making are common uses of Artificial Intelligent (AI) systems in relation to persons. This includes fairness, transparency towards individuals, and the right to human intervention. For three reasons, machine learning (ML) using neural networks is fast advancing: (1) Massive increases in data capacity; (2) Massive increases in processing power; (3) Massive improvements in machine learning algorithms and more manual skills to write them. This results in power concentration [73]. The creation, use, and effects of AI are all accompanied by a slew of ethical concerns. These concerns range from the potential impact of AI on citizens' fundamental human rights to the security and use of data collected; from bias and discrimination unintentionally embedded in an AI by a homogeneous group of developers to a lack of public awareness and understanding about the consequences of their choices and usage of any given AI, leading to ill-informed decisions and subsequent harm [74].

According to Anute et al. [75], AI in retail is all about providing personalised offers for your customers, forecasting future trends, and optimising stock and improving supply. However, AI will have a considerably larger negative impact. Issues such as the cc tool provide strong statistics and interactive dashboards to help suppliers gain a better knowledge of frauds and build strategies to combat them. AI applications aren't transparent and can jeopardise privacy. The provision of strong tools to address data privacy for AI systems is a critical first step toward bringing trust into the AI equation [76]. To the detriment of human oversight, an increasing number of functions are being delegated to algorithms, raising concerns about the loss of justice and equitability [77]. Raji et al. [78] propose that a software-development company's algorithmic auditing process could aid in resolving some of the ethical difficulties

addressed. In theory, greater interpretability might be gained by employing simpler algorithms, but this may come at the sacrifice of accuracy.

According to Floridi and Cowls [79], the autonomy dimension, or the ability to refuse to provide AI decision-making power for overwhelming reasons, plays a significant impact. Especially, according to this decide-to-delegate dimension, in which decision making should be retained by humans, practising it when essential and relinquishing it when required. Despite the problems identified and the intentions expressed to overcome them, the development of decision-making algorithms remains a mystery. There have been few attempts to make the algorithms developed public. Attempts to make the process more inclusive, with greater participation from all stakeholders, are also underway. Identifying a relevant pool of social actors may necessitate a significant effort in terms of stakeholder mapping to provide a comprehensive, yet simple, governance in terms of the number of participants and working methods [80].

Unesco [81] recommends adequate data protection frameworks and governance procedures should be implemented in a multi-stakeholder approach at the national or international level, safeguarded by judicial systems, and ensured throughout the life cycle of AI systems. Data protection frameworks and any related mechanisms should be based on international data protection principles and standards for the collection, use, and disclosure of personal data, as well as the exercise of data subjects' rights, while ensuring a legitimate purpose and a valid legal basis for data processing, including informed consent. Also, algorithmic systems necessitate thorough privacy effect analyses, which involve societal and ethical issues as well as creative use of the privacy by design approach. AI actors must hold themselves accountable for the design and execution of AI systems, ensuring that personal information is protected throughout the AI system's life cycle.

6.2 Big Data; Harvesting and Digital Footprint

Like AI technology, big data has become a typical occurrence in almost every industry of the economy. The gathering, use, storage, and sharing of large amounts, diversity, and velocity of data from various sources is referred to as big data. Because it is essential in managing day-to-day operations and making choices, companies deal in large amounts of unstructured, structured, or semi-structured data. It is consequently vital to have a proper big data analytics framework to make educated decisions [82]. When you use the internet, your digital footprint—sometimes known as a digital shadow or an electronic footprint—refers to the trail of data you leave behind [83]. Websites you visit, emails you write, and information you provide online are all part of it. A person's internet actions and gadgets can be tracked using a digital footprint. Internet users either actively or passively leave a digital trace. Information mining in the context of big data focuses on analysing users' access duration, frequency, and browsing habits, dynamically adjusting page structure, and providing personalised information offerings to better satisfy their changing demands. Whereas data

harvesting is a technique for extracting and analysing data from online sources. The development of mining technology provides digital libraries with the technical assistance they need to provide fully intelligent information services [84].

According to Gupta et al. [85], big data has been in the IT industry for a while, but there are still challenges in compiling and evaluating the data. Different sorts of data are stored in different ways by different companies (format). It is overwhelming and difficult to compile, regularise, and omit irregularities without eliminating the information and its worth. Information release without authorization checks, data modifications, and service denial are all examples of security breaches. A study conducted by Zulkarnain et al. [86] illustrates various security and ethical issues related to big data. The authors state that there are enormous privacy risks associated with big data. Data security, customer profiling and the breach of sensitive data are also considered by the researchers as potential risks.

Subsequently, another study performed by Aljehane [82] finds a scarcity of qualified employees capable of managing data analytics in the right manner as the major challenge faced in big data. The author urges experts to be well-versed in the methodologies, tactics, and interventions that should be used in the accurate analysis, storage, sharing, and interpretation of data. The information lifecycle (data provenance, ownership, and classification), the data production and gathering process, and the lack of security measures are all examples of big data-related risks as stated by Moura and Serrao [87]. Maayan [88] point out distributed data as the major big-data security challenge. For speedier analysis, most big data frameworks split data processing duties over multiple systems. Hadoop is a well-known open-source platform for data processing and storage on a large scale. Hadoop lacks consideration for security. The MapReduce mapper can be forced by hackers to display inaccurate lists of values or key pairs, rendering the MapReduce process useless. A system's workload could be lowered by distributed processing, but as the number of computers grows, so will the number of security risks. Big Data security goals are the same as they are for any other data type: to maintain confidentiality, integrity, and availability. Authors also recommend the collection of data from various mediums to recognize identical attacks, applying a standard policy configuration for devices and implementing security measures to mitigate potential attacks as the active security practice [87].

Gupta et al. [85], recommend using an elliptical curve cryptographic algorithm for effective authentication, authorisation prevention, encryption, and audit trails to mitigate such security risks. Even if the information is encrypted, genuine users must be able to access it. This necessitates the implementation of effective access control strategies that enable users to access the appropriate data [89]. The process of securing cryptographic keys against loss or misuse is known as key management. In comparison to dispersed or application-specific key management, centralised key management is more efficient. Secure keys, audit logs, and rules are all accessed at a single point in centralised management systems. For businesses that handle sensitive data, a dependable key management system is critical [88]. Kantarcioglu and Ferrari [89] indicate that blockchains may have significant consequences for big data security and privacy. Highly secured financial transactions, information exchange and

data storage can be obtained using blockchain-based techniques coupled with other cryptographic methods. New insights into evolving data security challenges could be dealt with data stored on blockchains.

6.3 Cybercrimes

Gartner [90] predicts that by 2025, 70% of organizations will move from big data to small and wide data, providing more focus on analytics and making artificial intelligence (AI) less significant. One of Gartner's top data and analytics trends for 2021 is the shift from big data to small and wide data. These developments represent commercial, market, and technological factors that data and analytics leaders cannot afford to overlook. Adlakha et al. [91] claims that the digitalization of data is resulting in an increase in the number of cybercrimes involving the use of a computer, the internet, the World Wide Web, and cyberspace. Consumers, as well as governmental and private businesses, are being targeted by cyber thieves who are growing more skilled. The fundamental cause for the rapid rise in cybercrime is a lack of cyber protection. Given the widespread use of e-mail for both commercial and personal purposes, it's no wonder that phishing e-mails are the most prevalent attack vector for gaining unauthorised access or stealing data from unsuspecting users [92]. Phishing is when cyber thieves deceive their victims by sending them fake emails, text messages, or phone calls. The goal is usually to get you to visit a website that will either download a virus or steal your bank account or other personal information [57].

Abroshan et al. [93] illustrates an example of a common phishing attack process. Firstly, a phishing link or file could be included in the email (or both). Technical phishing protection systems may detect and stop the email before it reaches the user's mailbox. If the technological systems do not prohibit it, the user gets and may open the phishing email. The user clicks on the phishing link in the email or attached file in the second step of the procedure. When the malicious link is clicked, a phishing webpage appears, which helps the attacker to deceive the user and collect sensitive information. The attachment may include embedded malware like viruses or ransomware. It may also contain a false document asking the user to take any immediate action. When sensitive information is shared by the user, or follow the action requested by the attacker, the user's device or account get compromised and a phishing attack takes place.

As per Pathan [94], while various security systems and Intrusion Detection Systems (IDSs) may be used to counteract other sorts of cyber-attacks, phishing cannot be resisted only using such, even if the techniques are advanced. This is because human error is frequently engaged in the process of personal data and information leaking. To guard against phishing attacks, understanding the problem, and controlling online conduct are essential. A study conducted by Abroshan et al. [93] found that a high level of overall risk-taking can raise the likelihood of clicking on a phishing link, and women appear to be more prone to doing so. Researchers recommend future research could concentrate on gender-specific behaviours and other

psychological factors that could influence the chance of being prone to phishing attacks indirectly or directly. An article by Singh et al. [95] focuses on applying deep learning algorithms to examine the URLs of online pages, which is a real-time analysis of the harmful website. The suggested method detects phishing webpages using URLs and a convolutional neural network (CNN).

A paper presented by Athulya and Praveen [96] proposed a model with a hybrid approach to phishing detection in which a random selection of techniques is used to accurately detect real websites without moving on to the next phase. Random characteristics are chosen using a similar method to the keyed intrusion detection system. Phishing is an illicit method of obtaining vulnerable and personal information. According to the findings from Arshey and viji [97], anti-phishing operations can be classified into three groups: detective, preventive, and corrective. Machine Learning algorithms may assess designs that reveal dangerous source's emails by looking at the header piece and a portion of the subject matter in the body of the email. By encrypting and codifying designs, ML prototypes may be taught to sense the proximity of phishing attacks. The emails are parsed using text indexing techniques, then features are selected, and weights are sent through classifiers to identify whether or not the emails are phoney.

According to NCSC [98], most phishing defences rely solely on users' ability to recognise phishing emails. This strategy will only be somewhat successful. Instead, organizations should beef up their defences with more sophisticated safeguards. This will increase company resistance to phishing assaults without interfering with business productivity. Using a multi-layered approach will help to notice and stop a phishing assault before it does harm. Even if certain attacks succeed, it will help organizations to have a recovery plan and reduce the harm. With the advancement of AI, cybercriminals are finding it easier to mount sophisticated phishing assaults. It's vital to keep everyone educated and knowledgeable about cyber hygiene as technology advances. Being prepared is the key to the successful mitigation of cyber-attacks.

7 Conclusions

As discussions and experiments on digital transformation continue, organizations should focus on what new and existing competitors are doing to achieve digital transformation success and decide when to begin the transition. This could be driven by a variety of variables, including widespread adoption of the new technology across the industry, competitors adopting similar trends, and customers demanding it [99]. From a technological and business standpoint, digital transformation may be a difficult and hazardous undertaking. Organizations may have a dual strategy of growing their physical and digital businesses at the same time, but it's vital to maintain the two linked. Raj (no date) recommends organisations establish a direct link between digital transformation and business objectives to help them make reasonable long-term investments without losing short-term profits. Since the pace

of digitization shows no slowing down soon, businesses should consider increasing IT budgets in line with their other strategies. Forrester [100] predicts that rather than following traditional digital transformation strategies, leading businesses will embrace emerging technologies to unleash their employees' creativity and generate innovation that is focused on outcomes rather than simply financial gains. Meanwhile, further studies should consider focusing on the benefits and drawbacks of digital transformation for various business categories and within different technical models.

References

1. Vaughan T (2021) The 4 main areas of digital transformation. https://www.poppulo.com/blog/what-are-the-4-main-areas-of-digital-transformation
2. Teichert R (2019) Digital transformation maturity: A systematic review of literature. Acta Univ Agriculturae Silviculturae Mendelianae Brunensis 67(6):1673–1687
3. Zeike S, Choi KE, Lindert L, Pfaff H (2019) Managers' well-being in the digital era: Is it associated with perceived choice overload and pressure from digitalization? An exploratory study. Int J Environ Res Public Health 16(10):1746. https://doi.org/10.3390/ijerph16101746
4. Larjovuori R-L, Bordi L, Heikkilä-Tammi K (2018) Leadership in the digital business transformation. 212–221. https://doi.org/10.1145/3275116.3275122
5. Wagire et al (2021) Development of maturity model for assessing the implementation of Industry 4.0: learning from theory and practice. Prod Plann Control 32(8):603–622. https://doi.org/10.1080/09537287.2020.1744763
6. Pirola F, Cimini C, Pinto R (2019) Digital readiness assessment of Italian SMEs: A case-study research. J Manuf Technol Manag 1–39
7. Reinhardt K (2018) Integrated model for digital leadership management—key strategies to cope with digital uncertainty from a competence-based perspective. https://doi.org/10.13140/RG.2.2.31720.57605
8. Beresford J (2018) Digital business transformation: an Australian perspective. Gartner 1–17. Retrieved from https://www.gartner.com/doc/3885879/digital-businesstransformation-australia-perspective
9. Gunzel-Jensen F, Rosenberg Hansen J, Felsager Jakobsen ML, Wulff J (2017) A two-pronged approach? Combined leadership approaches and innovative behaviour. Int J Publ Adm 41(12):957–970
10. Prince and Ann K (2018) Digital leadership: transitioning into the digital age. Ph.D. thesis, James Cook University
11. Mihardjo L, Furinto A (2018) The effect of digital leadership and innovation management for incumbent telecommunication company in the digital disruptive era. Int J Eng Technol 7. https://doi.org/10.14419/ijet.v7i2.29.13142
12. Singh T (2020) Defining digital leadership during covid-19. https://hr.economictimes.indiatimes.com/news/industry/defining-digital-leadership-during-covid-19/75735078
13. Mcmanus R (2021) Leadership for the digital revolution. https://www.dukece.com/insights/leadership-digital-revolution/
14. Kokot K, Kokotec IĐ, Čalopa MK (2021) Impact of leadership on digital transformation. In: 2021 IEEE technology and engineering management conference—Europe (TEMSCON-EUR), pp 1–6. https://doi.org/10.1109/TEMSCON-EUR52034.2021.9488620
15. Borowska (2019) Digital leadership for digital transformation. http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.ojs-issn-2082-677X-year-2019-volume-10-issue-4-article-3870
16. Gartner (2020) 7 traits of highly successful digital leaders. https://www.gartner.com/smarterwithgartner/7-traits-of-highly-successful-digital-leaders

17. Conway C, Codkind M (2021) Where digital transformations go wrong in small and midsize companies. https://hbr.org/2021/08/where-digital-transformations-go-wrong-in-small-and-midsize-companies
18. OECD (2021) The digital transformation of SMEs, OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris. https://doi.org/10.1787/bdb9256a-en
19. NCSC (2020) Cyber security small business guide. https://www.ncsc.gov.uk/files/NCSC_A5_Small_Business_Guide_v4_OCT20.pdf
20. NIST (2020) Zero trust architecture. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf
21. Hiscox (2019) Hiscox Cyber Readiness Report 2019. Available at: https://www.hiscox.co.uk/sites/uk/files/documents/2019-04/Hiscox_Cyber_Readiness_Report_2019.pdf
22. Tam T, Rao A, Hall J (2021) The good, the bad and the missing: a narrative review of cybersecurity implications for Australian small businesses. Comput Secur 109:102385. https://doi.org/10.1016/j.cose.2021.102385. ISSN 0167-4048
23. Petratos PN (2021) Misinformation, disinformation, and fake news: cyber risks to business. Bus Horiz 64(6):763–774. ISSN 0007-6813. https://doi.org/10.1016/j.bushor.2021.07.012. (https://www.sciencedirect.com/science/article/pii/S000768132100135X)
24. ISACA (2020) ISACA, state of cybersecurity 2020. Information Systems Audit and Control Association, Schaumburg, IL, Google Scholar
25. ENISA (no date) What is social engineering? https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/what-is-social-engineering
26. Cyber Security Breaches Survey (2019) Department for digital, culture, media and sport. UK Government (3 Apr 2019) www.gov.uk/government/statistics/cybersecurity-breaches-survey-2019. Accessed Nov 2021
27. Lloyd G (2020) The business benefits of cyber security for SMEs. Comput Fraud Secur 2020(2):14–17. https://doi.org/10.1016/S1361-3723(20)30019-1 (https://www.sciencedirect.com/science/article/pii/S1361372320300191). ISSN 1361-3723
28. Belitski M, Liversage B (2019) E-leadership in small and medium-sized enterprises in the developing world. Technol Innov Manag Rev 9(1):64–74. http://doi.org/10.22215/timreview/1212
29. Duncan Jefferies (2021) Digital transformation: Three priorities for governance leaders. https://www.raconteur.net/business-strategy/leadership/digital-transformation-three-priorities-for-governance-leaders/
30. Reis JL, Ferreira R, Sousa RD (2021) Information systems function and governance in portuguese organizations. In: 2021 16th Iberian conference on information systems and technologies (CISTI), pp 1–8. https://doi.org/10.23919/CISTI52073.2021.9476423
31. Smallwood R (2019) Information governance concepts, strategies and best practices. https://www.google.co.uk/books/edition/Information_Governance/z3nADwAAQBAJ?hl=en&gbpv=1&dq=information+governance&pg=PT11&printsec=frontcover
32. NCSC (2019) NCSC CAF GUIDANCE. https://www.ncsc.gov.uk/collection/caf/caf-principles-and-guidance/a1-governance
33. NIST (2019) CISA cyber essentials: the leaders guide. https://www.cisa.gov/sites/default/files/publications/19_1105_cisa_CISA-Cyber-Essentials.pdf
34. Nehemia M, Iyamu T, Shaanika IN (2019) A comparative analysis of e-governance and IT governance. Open Innov 2019:288–296
35. Iyer S, Jajal B, Chauhan DA (2019) Comparative study, analysis and mitigation of e-governance models using secure information exchange. In: 2019 6th international conference on computing for sustainable global development (INDIACom), pp 1232–1235
36. Yasmine D (2021) How can AR contribute to business decision making. https://www.clickz.com/how-can-ar-contribute-to-business-decision-making/264689/
37. Rokhsaritalemi S, Sadeghi-Niaraki A, Choi S-M (2020) A review on mixed reality: current trends, challenges and prospects. Appl Sci 10:636. https://doi.org/10.3390/app10020636
38. Dwivedi YK, Ismagilova E, Hughes DL, Carlson J, Filieri R, Jacobson J, Jain V, Karjaluoto H, Kefi H, Krishen AS, Kumar V, Rahman MM, Rauschnabel PA, Rowley J, Salo J, Tran GA,

Wang Y (2020) Setting the future of digital and social media marketing research: perspectives and research propositions. Int J Inf Manage (2020). https://doi.org/10.1016/j.ijinfomgt.2020.102168. in press Google scholar
39. Tepper OM, Rudy HL, Lefkowitz A, Weimer KA, Marks SM, Stern CS, Garfein ES (2017) Mixed reality with hololens: where virtual reality meets augmented reality in the operating room Plast Reconstr Surg 140(5):1066–1070, 2017 Nov
40. Pantano E, Rese A, Baier D (2017) Enhancing the online decision-making process by using augmented reality: a two country comparison of youth markets. J Retail Consum Serv 38:81–95. https://doi.org/10.1016/j.jretconser.2017.05.011. (https://www.sciencedirect.com/science/article/pii/S096969891730098X). ISSN 0969-6989
41. Rauschnabel PA (2021) Augmented reality is eating the real-world! The substitution of physical products by holograms. Int J Inf Manage 57:102279, https://doi.org/10.1016/j.ijinfomgt.2020.102279. (https://www.sciencedirect.com/science/article/pii/S026840122031478X). ISSN 0268-4012
42. Leaman S (2019) How holographic technology is helping doctors deliver better care. https://www.soprasteria.com/insights/details/how-holographic-technology-is-helping-doctors-deliver-better-care
43. James L (2020) A new holographic reality for business https://eandt.theiet.org/content/articles/2020/07/a-new-holographic-reality-for-business/
44. Ardito L, Petruzzelli AM, Panniello U, Garavelli AC (2019) Towards Industry 4.0: Mapping digital technologies for supply chain management-marketing integration. Bus Process Manag J
45. NIST (2018) Blockchain technology overview. https://nvlpubs.nist.gov/nistpubs/ir/2018/nist.ir.8202.pdf
46. Thuraisingham B (2020) Blockchain technologies and their applications in data science and cyber security. In: 2020 3rd international conference on smart blockchain (SmartBlock), pp1–4. https://doi.org/10.1109/SmartBlock52591.2020.00008
47. Mearian L (2020) How blockchain plays into digital transformation. https://blogs.idc.com/2020/12/14/how-blockchain-plays-into-digital-transformation/
48. IBM (2018) Blockchain….really? Or blockchain…really! https://www.ibm.com/downloads/cas/4QNJLY9Z
49. Jabbour CJC, Fiorini PDC, Ndubisi NO, Queiroz MM, Piato ÉL (2020) Digitally-enabled sustainable supply chains in the 21st century: a review and a research agenda. Sci Total Environ 725(2020):138177. https://doi.org/10.1016/j.scitotenv.2020.138177
50. Ebinger F, Omondi B (2020) Leveraging digital approaches for transparency in sustainable supply chains: a conceptual paper sustain, p 12. https://doi.org/10.3390/su12156129 Google Scholar
51. Saini K (2018) A future's dominant technology blockchain: Digital transformation. In: 2018 international conference on computing, power and communication technologies (GUCON), pp 937–940. https://doi.org/10.1109/GUCON.2018.8675075
52. Li J et al (2020) How can blockchain shape digital transformation: a scientometric analysis and review for financial services. Manag Sci Informatization Econ Innov Dev Conf (MSIEID) 2020:264–267. https://doi.org/10.1109/MSIEID52046.2020.00054
53. Liu H, Islam SMN, Liu X, Wang J (2020) Strategy-oriented digital transformation of logistics enterprises: the roles of artificial intelligence and blockchain. In: 2020 5th international conference on innovative technologies in intelligent systems and industrial applications (CITISIA), pp 1–5. https://doi.org/10.1109/CITISIA50690.2020.9371847
54. Sosin AI, Ivanova OY, Vasilyeva SA (2020) Prospects for implementing blockchain data storage technology as a process of digital transformation of society. In: 2020 international multi-conference on industrial engineering and modern technologies (FarEastCon), pp 1–5. https://doi.org/10.1109/FarEastCon50210.2020.9271270
55. Stevens A (2021) IT leaders on why your organization needs zero trust, with tips on implementation. https://tbtech.co/blockchain/security-and-data/8-it-leaders-on-why-your-organization-needs-zero-trust-with-tips-on-implementation/

56. Security Brief (2021) What to know when considering zero trust digital transformation. https://securitybrief.com.au/story/what-to-know-when-considering-zero-trust-digital-transformation
57. NCSC (2021) Zero trust architecture design principles. https://www.ncsc.gov.uk/collection/zero-trust-architecture/introduction-to-zero-trust
58. NSA (2021) Embracing a zero trust security model https://media.defense.gov/2021/Feb/25/2002588479/-1/-1/0/CSI_EMBRACING_ZT_SECURITY_MODEL_UOO115131-21.PDF
59. Gordon S (2019) A matter of trust. Netw Secur 2019(5):9–11. https://www.sciencedirect.com/science/article/abs/pii/S1353485819300601
60. Pallais D (2019) Why banks are adopting a modern approach to cybersecurity—the Zero Trust model. The Official Microsoft Blog, 18 September. Available at: https://www.microsoft.com/en-us/microsoft-365/blog/2019/09/18/why-banks-adopt-modern-cybersecurity-zero-trust-model/
61. Klasnja V (2020) The intersection of zero trust and digital transformation. https://www.netskope.com/blog/the-intersection-of-zero-trust-and-digital-transformation
62. Zscaler (2021) Securing the future of work with zero trust. https://www.cio.com/article/3636493/securing-the-future-of-work-with-zero-trust.html
63. Goldstein S (2021) Moving to zero trust—A process or a practice? https://www.doublechecksoftware.com/moving-to-zero-trust-a-process-or-a-practice/
64. Erol T, Mendi AF, Doğan D (2020) Digital transformation revolution with digital twin technology. In: 2020 4th international symposium on multidisciplinary studies and innovative technologies (ISMSIT), pp 1–7. https://doi.org/10.1109/ISMSIT50672.2020.9254288
65. Moore S (2019) Top trends from gartner hype cycle for digital government technology 2019—smarter with gartner. Gartner, Oct 2019. [online]. Available: https://www.gartner.com/smarterwithgartner/top-trends-from-gartner-hype-cycle-for-digital-government-technology-2019
66. IBM (n.d.) What is a digital twin? https://www.ibm.com/uk-en/topics/what-is-a-digital-twin
67. Belloni F (2020) Digital twins, a pillar of digital transformation. https://www.techedgegroup.com/blog/digital-twins-a-pillar-of-digital-transformation
68. Fuldauer E (2019) Smarter cities are born with digital twins. https://tomorrow.city/a/smarter-cities-are-born-with-digital-twins
69. Zhou L, An C, Shi J, Lv Z, Liang H (2021) Design and construction integration technology based on digital twin. Power Syst Green Energy Conf (PSGEC) 2021:7–11. https://doi.org/10.1109/PSGEC51302.2021.9541682
70. Möller DPF, Vakilzadian H, Hou W (2021) Intelligent manufacturing with digital twin. IEEE Int Conf Electro Inf Technol (EIT) 2021:413–418. https://doi.org/10.1109/EIT51626.2021.9491874
71. Royakkers L, Timmer J, Kool L et al (2018) Societal and ethical issues of digitization. Ethics Inf Technol 20:127–142. https://doi.org/10.1007/s10676-018-9452-x
72. Storm M, Wolk A (2021) Privacy and the EU's regulation on AI: What's new and what's not? https://www.mofo.com/resources/insights/210422-privacy-eu-regulation-ai.html
73. Green B (2020) Artificial intelligence and ethics: Sixteen challenges and opportunities. https://www.scu.edu/ethics/all-about-ethics/artificial-intelligence-and-ethics-sixteen-challenges-and-opportunities/
74. Bird et al. (2020) The ethics of artificial intelligence: issues and Initiatives. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf
75. Anute N, Paliwal M, Patel M, Kandale N (2021) Impact of artificial intelligence and machine learning on business operations. J Manage Res Anal 8:69–74. https://doi.org/10.18231/j.jmra.2021.015
76. Teich D (2020) Artificial intelligence and data privacy—Turning a risk into a benefit. https://www.forbes.com/sites/davidteich/2020/08/10/artificial-intelligence-and-data-privacy--turning-a-risk-into-a-benefit/?sh=1551b1aa6a95
77. Sareen S, Saltelli A, Rommetveit K (2020) Ethics of quantification: illumination, obfuscation and performative legitimation. Palgrave Commun 6:20. https://doi.org/10.1057/s41599-020-0396-5

78. Raji ID et al. (2020) Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. (Association for Computing Machinery). https://doi.org/10.1145/3351095.3372873
79. Floridi L, Cowls J (2019) A unified framework of five principles for AI in society. Harvard Data Sci Rev 1(1). https://doi.org/10.1162/99608f92.8cd550d1
80. Lo Piano S (2020) Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. Humanit Soc Sci Commun 7:9. https://doi.org/10.1057/s41599-020-0501-9
81. Unesco (2021) Report of the social and human sciences commission (SHS). https://unesdoc.unesco.org/ark:/48223/pf0000379920.page=14
82. Aljehane N (2020) Big data analytics: challenges and opportunities. In: 2020 international conference on computing and information technology (ICCIT-1441). pp 1–4, https://doi.org/10.1109/ICCIT-144147971.2020.9213765
83. Kaspersky (n.d.) What is digital footprint? And how to protect it from hackers. https://www.kaspersky.com/resource-center/definitions/what-is-a-digital-footprint
84. Fangjing Y (2021) Study on library individualized information security under the background of big data. In: 2021 IEEE 6th international conference on big data analytics (ICBDA) 2021, pp 138–142. https://doi.org/10.1109/ICBDA51983.2021.9402989
85. Gupta D, Rani R (2018) A study of big data evolution and research challenges. J Inf Sci 45. https://doi.org/10.1177/0165551518789880
86. Zulkarnain N, Anshari M, Hamdan M, Fithriyah M (2021) Big data in business and ethical challenges. Int Conf Inf Manage Technol (ICIMTech) 2021:298–303. https://doi.org/10.1109/ICIMTech53080.2021.9534963
87. Moura J, Serrao C (2019) Security and privacy issues of big data. https://doi.org/10.4018/978-1-5225-8176-5.ch080
88. Maayan G (2020) Big data security: challenges and solutions. https://www.dataversity.net/big-data-security-challenges-and-solutions/#
89. Kantarcioglu M, Ferrari E (2019) Research challenges at the intersection of big data, security and privacy. https://www.frontiersin.org/articles/10.3389/fdata.2019.00001/full
90. Gartner (2021) Gartner top 10 data and analytics trends for 2021. https://www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021
91. Adlakha R, Sharma S, Rawat A, Sharma K (2019) Cyber security goal's, issue's, categorization and data breaches. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon), pp 397–402. https://doi.org/10.1109/COMITCon.2019.8862245
92. Legg P, Blackman T (2019) Tools and techniques for improving cyber situational awareness of targeted phishing attacks. In: 2019 international conference on cyber situational awareness, data analytics and assessment (cyber SA). pp 1–4, https://doi.org/10.1109/CyberSA.2019.8899406
93. Abroshan H, Devos J, Poels G, Laermans E (2021) Phishing happens beyond technology: the effects of human behaviors and demographics on each step of a phishing process. IEEE Access 9:44928–44949. https://doi.org/10.1109/ACCESS.2021.3066383
94. Pathan AK (2018) Defending against common cyber attacks: Phishing and cross-site scripting. Int Symp Program Syst (ISPS) 2018:1–1. https://doi.org/10.1109/ISPS.2018.8378960
95. Singh S, Singh MP, Pandey R (2020) Phishing detection from URLs using deep learning approach. In: 2020 5th international conference on computing, communication and security (ICCCS), pp 1–4. https://doi.org/10.1109/ICCCS49678.2020.9277459
96. Athulya AA, Praveen K (2020) Towards the detection of phishing attacks. In: 2020 4th international conference on trends in electronics and informatics (ICOEI)(48184), 2020. pp 337–343. https://doi.org/10.1109/ICOEI48184.2020.9142967
97. Arshey M, Angel Viji KS (2021) Thwarting cyber crime and phishing attacks with machine learning: a study. In: 2021 7th international conference on advanced computing and communication systems (ICACCS) 2021, pp 353–357. https://doi.org/10.1109/ICACCS51430.2021.9441925

98. NCSC (2018) Phishing attacks: defending your organization. https://www.ncsc.gov.uk/guidance/phishing
99. Ramesh N, Delen D (2021) Digital transformation: how to beat the 90% failure rate? In: IEEE engineering management review, vol. 49(3), pp 22–25, 1 thirdquarter, Sept 2021. https://doi.org/10.1109/EMR.2021.3070139
100. Forrester (2021) Predictions 2022: this is a year to be bold. https://www.forrester.com/blogs/predictions-2022/

Challenges and Opportunities of Autonomous Cyber Defence (ACyD) Against Cyber Attacks

Michael Oreyomi and Hamid Jahankhani

Abstract The exponential growth of Autonomous Intelligent Malware (AIM) has really changed the landscape in the cyber security fight against that attack of auto generated threats and how these threats needs to be dealt with. This research focuses on the use of Artificial Intelligence (AI), Machine learning (ML) and Deep learning (DL) to mitigate auto-generated cyber attacks which are hard to track and neutralise. Hence, the key issue that the research proposes to address is to evaluate the key challenges in using AI and ML as decision tools against autonomous cyber attacks, and the opportunities it presents. A proposal is posited on the future utilisation of Autonomous Cyber Defence (ACyD) as a tool to neutralise Autonomous intelligent Malware (AIM), embedded into Security Information and Events Management systems (SIEM). ACyD is principally a self-defending and self-healing cyber security system with the sole aim of persistently and autonomously defending all cyber physical systems against cyber attacks. The use of ACyD in cyber defence is a relatively new research area that the author hopes will gain grounds in the future.

Keywords ACyD · Autonomous cyber defence · Artificial intelligence (AI) · Machine learning (ML) · Deep learning (DL) · Cyber attack · Cyber defence · C4ISR · Deep autoencoder (DAE) · Convolutional neural network (CNN)

1 Introduction

As the speed at which data is being processed through the use of ever evolving technologies continues to grow, it has become obvious that cyber defences being deployed cannot keep abreast of the trends of attack techniques being used by attackers if most defences solely rely on human intervention.

M. Oreyomi
Northumbria University London, London, UK

H. Jahankhani (✉)
Northumbria University London Campus, London, UK
e-mail: Hamid.jahankhani@northumbria.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_9

Chen and Wang [1] and Dasgupta [2] argued that "physical devices such as sensors and detectors are not sufficient for monitoring and protection of these infrastructures,hence, there is a need for more sophisticated IT that can model normal behaviours and detect abnormal ones. These cyber defence systems need to be flexible, adaptable and robust, and able to detect a wide variety of threats and make intelligent real-time decisions".

This shift calls for a change of defence strategy and the need to look at implementing automated systems that can track threats and autonomously neutralise them.

Further research indicates that the exponential growth in worldwide cyber attack [3] which has led to serious negative social and economic impact [4], has now called for the use sophisticated and advanced technological solution to address the growing menace of cyber attacks [5].

Huang et al. [3] in their research demonstrated that cyber attacks are rarely localised. Moreover, the alarming rate at which they are increasing is a global issue because attacks traverse beyond geographical boundaries, and cost victims as well as governments about $600 billion dollars per year [6].

As these attacks and costs increases, so does the risk of cyber criminals trying and using sophisticated data exfiltration and attack methods to siphon data from unsuspecting victims and poorly configured systems, and demanding ransom. Data exfiltration takes place when malware and/or a malicious actor carry out an unauthorized data transfer from a compromised computer system. This rising attacks have made many businesses to re-evaluate how they perceive risks.

In 2017, risk practitioners from the PwC were clear in their analysis and cited the fact that the most significant operational risk, is data security [7].

Over the years, many scholars have attempted to define Artificial Intelligence (AI), Machine learning and Deep learning (ML/DL). However, not one specific definition is attributed to AI.

Quite often AI/ML is interchangeably used. However, both terms are actually two consistent paradigms. The term "artificial intelligence "emanated from the 1950s as part of the computer science teachings to demonstrate that computer programmes can be trained to mimic human intelligence. In practice, this means computer programmes can learn, reason and behave like humans, if given the right knowledge. Hence, the computer programs can use its knowledgebase to make decisions without human input.

Machine learning on the other hand, is a specified branch of Artificial Intelligence that translates algorithms to comprehend "phenomena models from experience or instances".

As many other researched have shown that deep learning and machine learning appears to be interchangeably used, the disparity and similarities between the two are worth noting.

Machine learning, deep learning, and neural networks are "all sub-fields of artificial intelligence. However, deep learning is actually a sub-field of machine learning, and neural networks is a sub-field of deep learning" (IBM 2020).

According to Hatcher and Yu [8] DL and ML techniques "are widely used for malware analysis and in finding unforeseen threats because of malicious software". While this aspect of ML is commended, some flaws/limitations have been noted by other researchers.

At the early stages of AI development, Machine learning technology played significant role in addressing many cyber security threats. While it is admitted that ML is quite versatile and powerful in cyber security, the over reliance on feature extraction is a glaring weakness. Golovko [9] in their research posited that developers have to manually compile the various features associated with malware in order to allow ML solution to recognise a threat. This method degrades the efficacy and precision of threat detection. This weakness is associated to the fact that ML only works with pre-defined rules. Hence, undefined rules escape detection and discovery.

Ultimately, the conclusion can be drawn that the performance of ML is measured in terms of the accuracy of the defined feature recognition and extraction.

As a result of this obvious flaw, other methods of detections were being explored. Hence, the need to explore deep neural networks (DNN) which is a sub-domain of ML.

For clarity, AI is referenced in this research to discuss topics that crosses over other areas like ML, DL and related concepts.

For a machine to be considered generally artificial intelligent, it ought to have at least six foundational capabilities. In fact, these six capabilities are required to pass the well known Turing Test proposed by scientist Alan Turing and the Total Turing Test [10]. This six capabilities or disciplines are the ability to understand the natural language spoken or written by another human being, ability to store and process information, ability to reason, ability to learn from new information, ability to see and perceive objects in the environment, and finally, ability to manipulate and move physical objects. While some advanced agents may possess all of the six capabilities, in order to reap the benefits of AI for security, you actually do not need all of them in an intelligent agent. For example, if you are building or deploying an AI based network intrusion detection system, the system does not need the ability to see or physically manipulate objects, but it must have the ability to store, process, and learn from enormous amounts of network logs. On the other hand, an AI powered security surveillance robot watching the entrance of a data centre must have the ability to see and distinguish a person from a wandering deer. That said, at minimum, artificial intelligence systems employed in the field of security possess some common capabilities that make them 'intelligent',they store and process large volumes of data. In the field of security, this data is usually in the form of logs from network devices, work stations, and API calls and so on. They identify patterns in data.

While AI technologies play a key role is cyber defence, they can be attacked and deceived.

One well known method of deceiving AI in adversarial attack environment is the manipulation of training samples which can result in toxic data being fed into the system, and meddling with results. This leads to evasion attack being made easy. Under normal circumstances, adversarial attacks are specifically carried out

to undermine the data integrity of attacked systems, and render the AI defences powerless, by causing classifiers to derive wrong classification. This leads to the evasion of embedded defences.

To avert these attacks, defence methods includes (1) changing training process/input samples, (2) configuring network with more layers and fortifying security settings, (3) Use of external models to classify samples (Akhtar and Mian [11].

1.1 The Import Role and Influence of AI/ML in Cyber Security

While many researchers have put forward many arguments about the use of AI in many areas of our lives, Liu et al. [12] posited that "the application of AI is a valuable factor in facilitating the security aspects in significant areas, like government infrastructure", healthcare and other areas like Automatic Speech Recognition (ASR).

While there is clear evidence of many uses of AI in cyber security solutions, the damaging impact of the effects of cyber attacks lingers on. This is because AI is embedded into every technology that uses microprocessors. Evidence shows that there are microprocessors in nearly everything; pacemakers, smart homes, smart cars etc. Imagine a hacker being able to hack into pacemakers, or control the sensors in smart cars. Imagine how vulnerable that makes everyone feel about privacy. Clearly, there are serious concerns about privacy.

So, how does an artificially intelligent agent actually prevent data breach? First, AI can uncover data patterns that would have been extremely difficult for the human mind to analyse and then classifies the data to make it easy for a security professional to analyse. A more advanced intelligent agent goes one step further. It makes recommendations to the security professional on what actions to take. Lastly, the most advanced form of intelligent agent takes a corrective action on behalf of the security professional to mitigate identified security risk. In the most advanced form, AI could be deployed in threat modelling which is currently being done manually.

The process of threat modelling typically done at the architectural or design stage of development allows cyber security professionals to discover potential threats from an adversary's point of view. Many research have shown that the use of AI in threat modelling is still an area of research and ripe for innovation [13].

2 How Malware Cyber Attacks Occur?

As expected, attackers are persistently fishing for confidential information and "are getting better at identifying loopholes and vulnerabilities in many organisations security" Cyber Threat Intelligence [14].

Well known categories of vulnerabilities are hardware, software, network, platform, management, and technical vulnerabilities.

Hardware vulnerabilities are primarily caused by inadequate physical equipment security, while software, network, platform, and technical vulnerabilities are typically caused by communication and protocol configuration errors, and management vulnerabilities are primarily caused by human and policy errors, among other things.

Many researchers have demonstrated that majority of cyber attacks are motivated by financial gains for the attackers [15]. Hence, to protect systems, sophisticated mechanisms need to be devised to track and eliminate the threats of malware.

2.1 How Malware Attacks Can Be Detected?

Many researches on malware detection have shown that static and dynamic analyses are the two generic techniques primarily implemented by two approaches: signature-based and behaviour-based [16].

Information flow analysis-based approach [17] has been known to detect Android malware. Further research has shown that ML/AI/DL is one of the most promising techniques in detecting mobile malware [18]. However, ML/AI/DL approaches have noticeable weakness in that they are susceptible to adversarial countermeasures by attackers aware of how to avoid detection during an attack.

2.2 Threat Actors and Impacts of Attacks

Set out in Fig. 1 identified by this research are some of the main branches of cyber security applications that has adopted AI techniques to combat cyber attacks and emerging cyber security threats.

According to McAfee threat report of June 2021, the volume of malware threats observed by McAfee Labs averaged 688 threats per minute, an increase of 40 threats per minute (3%) in the first quarter of 2021. Notably, the technology sector suffered a 54% increment is attacks, the education sector attacks increased by 46% and the finance sector malware attack increased by 41%. In another report, IBM reported average cost of data breach to be $3.8 million on average and showed that many organisation have increased the use of automation to detect threats.

3 Literature Review

In their initial research Russell and Norvig [19] advocated that the world need to create “systems that can understand, think, learn, and behave like humans”. In

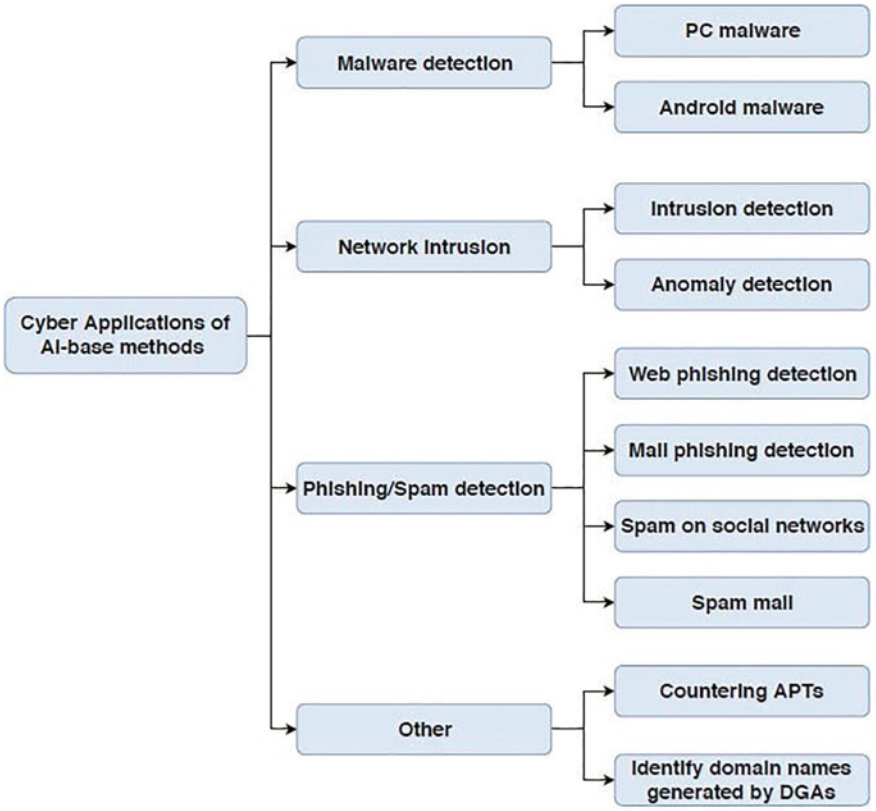


Fig. 1 Main branches of cybersecurity applications adopting AI techniques. *Source* Huang et al. [3]

that research, it was anticipated that machine intelligence will help enhance human intellect to fight cyber attacks which are historically carried out by humans.

Following on from that research, many other researchers have shown that cyber attacks are rarely localised and Huang et al. [3] demonstrated that attacks are exponentially increasing at an alarming rate which can pose a threat to computers and lives beyond geographical boundaries and costing victims of attacks approximately $600 billion dollars annually [6]. Further researches have demonstrated that attacks are now autonomous with little or no human intervention. Effectively, these researches indicate that we are living in an era of autonomous cyber attack agents governed by silent, hard to detect and very dangerous AI/ML/DL agents. The most potent of these are autonomous intelligent malware (AIM). Hence, the focus of this research is to look at the challenges and opportunities in using AI and ML as decision/defence tools against cyber attacks. In particular, the future use of Autonomous Cyber defence (ACyD)—a new and ground breaking research area, in tackling Autonomous intelligent Malware (AIM).

In a recent research, Theron and Kott [5] posited that autonomous malware capabilities and tactics will rapidly evolve and become more dangerous because they will develop the ability to stealthily operate undetected. Hence, the development of autonomous intelligent cyber-defence agents (AICA) using Autonomous cyber defence (ACyD) that are capable of autonomous learning to fight the growing threat of Autonomous Intelligent Malware (AIM), is well overdue.

Furthermore, they argue that the over reliance on humans for cyber security untenable in the foreseeable future and that the need to deploy autonomous intelligent cyber defence agents is imperative, and cited the fact that NATO's IST-152 group has identified the capabilities of AICA to "autonomously plan and execute complex multi-step activities for defeating or degrading sophisticated adversary malware". In addition, "it will have to be capable of adversarial reasoning to battle against a thinking, adaptive malware".

NAT's report established that the purpose of AICA is "to defeat the enemy malware in an environment of potentially disrupted communications where human intervention may not be possible".

In their research, Theron and Kott [5] utilised an unmanned aerial vehicle (UAV) with onboard computers to conduct the study, and it was assumed that the computers were compromised by an enemy malware. It was also assumed that the enemy malware and its capabilities and tactics, techniques, and procedures (TTPs), evolves rapidly. Hence, the cyber defence agent will be capable of autonomous learning. Their research posited that the AICA architecture consists of the following stages: "Sensing and World State Identification, Planning and Action Selection and learning".

In the sensing and world state identification, the agent acquires data from its environment. In the planning and action selection stage, the agent elaborates several actions proposals and in the learning stage, the cyber defence agent utilises its gained knowledge to progressively improve its own efficiency. This tactics is similar to how a commercial SIEM with built-in User behaviour analytics (UEBA) deals with the contents of collected logs before deciding on a course of action.

Some research have argued that future military combats will undoubtedly involve the use of AI-optimized and complex command, control, communications, computers, intelligence, surveillance and reconnaissance (C4ISR) and networks. Conversely, the prevalence of autonomous intelligent things combating with autonomous intelligent things at lightning speed will become standard. The intricate nature and scale of autonomy will invariably set off a new cyber attack tactics which will make AIM become a popular attack tool. If the attacks are successful, it will affect many government infrastructures, disrupt national operations and even disable mission critical operations.

Research has shown that the deployment of Autonomous cyber defence (ACyD) via integrated group of autonomous intelligent cyber defence agents (AICA) can be equipped to neutralise the threats of AIM in any network or computer systems.

While the concept of ACyD is at infancy, Theron and Kott [5] identified a few challenges:

- AICA's reference architecture is at infancy.
- Agents' engineering and certification

- Testability
- Implementation and compatibility technologies
- Autonomous self-engineering and self-assurance

In another study, Kholidy [20] advocated the use of Autonomous Response Controller (ARC), to autonomously protect the cyber physical system (CPS). The research mainly focussed on using the quantitative Hierarchical Risk Correlation Tree (HRCT) that models the paths an attacker can navigate to reach certain goals, and measures the financial risk that the CPS assets face from cyber attacks.

In order to validate ARC in real-time large-scale data, Aurora attack was simulated. The experiments' result shows that the accuracy of ARC outperforms the traditional Static Intrusion Response System (S-IRS) by 43.61%.

Furthermore, the research compared the output of ARC against Suricata intrusion response system, and ARC was able to mitigate the single line to ground (SLG) attacks and recover the CPS to its normal state in 122 s before Suricata does.

To add to this shift in cyber security paradigm, the use of AI and ML to combat cyber attacks as part of a new hypothesis in the use of Artificial intelligence to investigate, detect and neutralise the threat to cyber security autonomously.

Their research suggestion was largely based on the knowledge that the general architecture of Knowledge-based Artificial intelligence (KBAI) consisted of Knowledgebase and inference engine. Hence, they posited embedding human knowledge and decision making expertise in the form of KBAI, into machines and effectively making the machines "expert" systems, typically because inference engine commonly has IF-ELSE rules for "inference from the knowledge base". Hence, the idea is to represent knowledge unambiguously via If-ELSE rules Russell and Norvig [19].

In further research, Russell and Norvig [19] described AI as the systems that replicate human cognitive functions like speech learning, behaviour patterns and problems resolution. However, Kaplan and Haenlein [21] posited that AI is flexible and can learn from external influence and has the ability to adapt to new knowledge. Deep learning (DL) however, as posited by many literatures, is linked to ML concepts and architectures, however, it is positioned higher within the neural layers .

Many researchers have put forward theories of the benefits of DL applications in many branches of medicines and the role it can play in advancing the speed at which medical diagnosis can be carried out.

While is it recognised that DL's cognisant application is useful, its limitations cannot be overlooked in relation to human–machine interaction and translation of data.

Although AI and ML can be interchangeably used, they are in fact two unique hypothesis.

AI is traditionally seen as computer programs that can be trained to become 'intelligent'. In short, this notion implies that algorithms can be trained to become adaptive if equipped with the necessary knowledgebase. For example, they can be trained to take autonomous decisions based on specific rules. In contrast, ML is characterised as a branch of AI that utilises algorithms to comprehend "phenomena models" from experiences and instances.

AI will be used throughout this research to broadly discuss applicable and interrelated areas of DL and ML.

3.1 Malware Identification

The general term used to describe malicious software like viruses and Trojans, is Malware, which has become the most popular attack mechanism today.

The impact of malware attack on the digital economy is enormous and the research into the use of AI to autonomously mitigate threats is exponentially increasing.

In their research, Xu [22] posited the adoption of hardware assisted framework for online detection and elimination based on vitual memory access patterns using logistic regression, random forest classifier, support vector machine, and performing RIPE benchmark suite as part of the experiment to show proof of concept.

In their findings, a true positive rate of 99% with a less than 5% false positive rate was reported by the adopted framework.

In another study by Hashemi et al. [23], operational codes (OpCode), k-nearest neighbors (KNN), and a support vector machine (SVM) were used as ML operational codes to classify malwares. The research concluded that with verifiable empirical data, that the model is far more proficient with high detection rate and low false alarms in malware detection rate.

Building a DL architecture to expand the detection of intelligent malware in a further research, Ye et al. [24] posited the utilisation of AutoEncoder stacked up with multilayer restricted Boltzmann machines (RBMs) to detect unknown malware.

In their findings, it became obvious that heterogeneous DL framework will probably improve overall malware detection in comparison to traditional DL methodologies.

According to the authors, Olalere et al. [25] identifying and evaluating discriminative lexical features of malware URLs allowed them to construct a real-time malware uniform resource locator (URL) classifier in the form of a real-time malware URL classifier. In this case, it manually examined blacklisted malware URLs, which resulted in the identification of 12 distinguishing lexical characteristics. After that, an empirical analysis was carried out on the identified characteristics of both the existing blacklisted malware URLs and the newly collected malware URLs, which revealed that attackers used the same pattern when creating malware URLs as before. Finally, they used an SVM to evaluate the performance and effectiveness of the extracted features, and they were able to achieve an accuracy of 96.95 percent while maintaining a low false negative rate (FNR) of 0.018%.

A more recent trend of research in malware detection focused on mobile malware in general and Android malware in particular. ML and DL were both significant breakthroughs in this area. The raw opcode sequence from a disassembled program was used to classify malware. Novel ML algorithms, namely, rotation forest, for malware identity. An artificial neural network (ANN) and the raw sequences of API method calls were utilized in the research carried out by Karbab et al. [26] to detect

Android malware. A recent study by Wang et al. [27] introduced a hybrid model based on deep autoencoder (DAE) and a convolutional neural network (CNN) to raise the accuracy and efficiency of large-scale Android malware detection.

Interestingly, scientists were fascinated by the use of bio-inspired methods for malware classification. These techniques were mainly used for feature optimization and optimizing the parameter for the classifiers.

However, Li [28] research was eventually able to connect the dot between AI and its use in cyber defence. It was expanded to include the use of AI and Dl as counter attack tools. As much as this research showed the value of AI and DL as defence tools, it did not go far enough to showcase the dangers of other attack methods like data exfiltration, cross site scripting and SQL injection.

Many researches have shown that DL can be infected with malicious data during the training phase. Hence, to increase the impact of corrupt data without being detected, an attacker can introduce virus into the data during the training phase, thereby avoiding detection. Hence, the input is deceptive and the outcome will most likely be deceptive too. In both examples, the attacker can hijack the process and control the data ingestion.

To overcome this known issue, Jiansg et al. [29] posited that the use of Particle Swarm Optimisation (PSO) algorithm be used, to counter the corruption of data and avoiding detection. Their research advocated that more focus be placed on how data is introduced into the DL learning algorithm to ensure the right data is ingested.

Dwivedi et al. [30] in their research to find a solution to DDos attack using malware, suggested the use of grasshopper optimization algorithm (GOA) with machine learning algorithm (GOIDS) to create IDS to discern between normal and attack traffic. GOIDS uses the most relevant dataset from the original IDS dataset and pass selected features to the classifiers, i.e. support vector machine, decision tree, naïve Bayes, and multilayer perception to identify type of attack. The final result from the simulation of using KDD Cup 99, CAIDA 07, and CAIDA08, which are publicly available dataset, shows that GOIDS used in conjunction with decision tree gives high accuracy detection with very low false-positive.

Furthermore, the research of Yusof et al. [31] compliments the work of Dwivedi et al. [30] in that they were able to carry out a comprehensive literature review on the impact of DDoS attack. The result of their observation showed that the machine learning technique was significantly used in the prediction and detection of DDoS attacks.

In their research Al Najada et al. [32] presented a taxonomy for different types of attacks using Deep Learning. Forecasting models were created for each attack independently and then a forecasting model was created for all the attacks using deep learning and distributed random forest considering only a set of attributes to improve the accuracy. The class imbalance case was resolved using the oversampling technique. The developed model could accurately forecast the type of attack or threat.

In addition, Mane et al. [33] carried out work on the latest dataset called Modern DDoS. In their research, they compared the results of six established classification techniques: (1) Random Forest, (2) Naive Bayes, (3) Stochastic Gradient Descent, (4) Decision Trees, (5) Logistic Regression, and (6) K-Nearest Neighbour (KNN)

with the proposed Genetic programming model. They concluded that using genetic programming algorithm produced better results in comparison to other techniques used to detect DDos attacks.

Geetha and Thilagam [34] reviewed the effectiveness of ML and DL algorithms for cyber security and concluded that the solution to be deployed depends on how well the AI is trained to understand the threat.

3.2 Challenges of Detecting AI Cyber Attack

In cyber security, the key challenge is not about deploying a sophisticated endpoint protection but keeping up with the huge amount of security alerts and dealing with them. Humans are simply unable to keep up. As a result, we have no choice but to automate as much of the process as possible.

Among the numerous obstacles that continue to make cyber-attack detection difficult to achieve, the scarcity of training data is by far the most significant. Despite the fact that organisations and businesses rely on well-known network monitoring tools such as Wireshark, millions of people remain vulnerable due to a lack of information about website behaviours and features that could result in an attack. In reality, the majority of attacks are carried out not because threat actors employ sophisticated coding and evasion techniques, but rather because the victims lack the fundamental tools necessary to detect and avoid the attacks. However, machine learning is proving to be a game-changer in our understanding of cyber-attacks, as demonstrated by the research of Alloghani et al. [35], which applied machine learning techniques to Phishing Website data with the goal of comparing five algorithms and providing insight that the general public can use to avoid phishing traps. According to the findings of that study, the Neural Network algorithm is the most effective algorithm [35].

3.2.1 Key Issues

As AI systems learn from data in addition to programmed rules, unanticipated situations that the system has not been trained to handle and uncertainties in human–machine interactions can lead AI systems to display unexpected behaviours that pose safety hazards for its users.

In many AI systems, biases in the data and algorithm have been shown to yield discriminatory and unethical outcomes for different individuals in various domains, such as credit scoring and criminal sentencing. The autonomous nature of AI systems presents issues around the potential loss of human autonomy and control over decision-making, which can yield ethically questionable outcomes in multiple applications such as caregiving and military combat [36].

The scale of adoption of AI threatens to outpace the regulatory responses to address the concerns raised [37].

4 Research Methodology

The main focus of this research is to look at the challenges and opportunities in using AI and ML as decision tools against cyber attacks. Specifically, the research explores the future use of Autonomous cyber defense (ACyD) in tackling Autonomous intelligent Malware (AIM).

The research looks at how autonomous cyber attacks can be detected and suggests countermeasures that can be deployed to abate future attacks.

In this section, attention is drawn to the impact of Autonomous intelligent Malware (AIM), and how to fight its propagation.

In selecting the relevant data to the research topic, qualitative research method was used to collect the data that was used in the research. All non-numerical data used were extracted from archived data on autonomous malware attacks in addition to historical data on autonomous malware attacks using AI/DL and ML.

Several qualitative research methodologies were used to investigate the use of AI to neutralise cyber attacks, focusing on the use of ACyD to detect and nutralise cyber attacks. Case study analysis involved the collection and analysis of case studies related to reported incidences of AIM attacks. Qualitative data were collected from journal articles and other scholarly authoritative sources on the subject.

However, a more detailed research was conducted on how AIM attacks Cyber physical systems (CPS) with the sole aim of compromising the confidentiality, integrity and availability (CIA) of data.

In order to ensure that the research objectives are met, the use of qualitative research sample sizes that are not quantifiable or measurable, (which is the primary advantage), were used.

Within this context, research philosophy used is pragmatism. Research papers, journals and peer reviewed papers were subjected to critical analysis and gaps were examined by using funnelling technique to filter out unnecessary papers/journals. Research is focused on plugging the gaps that have not been addressed in the use of ACyD to combat AIM.

- Research approach
  Throughout the study, the approach used is inductive and based on qualitative data from research materials. For example, Lewis [6] estimated that $600 billion is paid out in ransom every year by victims of malware cyber attack, exemplifies the use of qualitative data to shed light on research area. Further qualitative data are used to buttress established facts.
- Research strategy
  The research strategy is based on grounded theory and utilises data collected from research papers to put forward a proposed framework/solution to combat the threat of AIM.

  Throughout the research, references will be made to the threat of AIM to cyber physical systems (CPS) and this covers all computer system controlled or monitored by computer-based algorithms.

Mono method was used throughout because qualitative data was used throughout the research to put forward a proposed solution.

- Research time horizons
  This is based on longitudinal time horizon as research will be looking at changes that occurred over a period of time. Specifically, the author used researches that are current (2015–2021), with sparse reference to historical papers.

  Finally, the project techniques and procedures are strictly based on collecting secondary data, analysing them and making inferences based on findings.

  Research explored publications (peer reviewed and others) on AI, ML, DL, AIM, ML decision on cyber attack using AIM. Science direct was used as primary data source as well as the IEEE. However, research expanded to other relevant reliable areas, and focused on the most recent researches (2015 onwards), by exploring the gaps.
- Sample Selection
  Although quite a fair number of datasets on AIM attacks and the use of AI/ML/Dl to carry out cybercrimes were collected, data on the key concerns on the impact of these attacks on the society at large and ethical concerns were also addressed.
- Data Analysis
  Clearly, the process of qualitative data analysis included the collection of articles that had previously been published in prestigious journals. Thus, text analysis was employed throughout the research in order to draw conclusions from the data. In order to expatiate points, descriptive datasets were extracted from a variety of sources (journals, peer-reviewed papers, and so on) and turned into case studies.

  Collected data includes current statistics on AIM attacks on cyber physical systems and its impacts on the society at large. In particular, the research focused on the attack on national infrastructure. The use of sample case studies were explored.

  Furthermore, the use of inductive analysis was deployed to understand specific patterns that were essential to gain an insight into, in order to form a holistic perspective on how AIM attack detection and countermeasures were carried out.
- Work done & Algorithms used.
  This research advocates the use of autonomous algorithmic techniques to detect AIM attacks and autonomously neutralise them.

4.1 Research Limitation

The research process carried out to complete the project was quite time consuming. Data gathering and analysis was very intensive as it involves the collection of data from diverse sources. As the data was collected, decision was made on what was relevant to the research topic and what was not needed. Through a selective sifting process, all data that was not needed were discarded.

As qualitative data requires that researchers to explore extended datasets, the data collected was wider than research area and exposed the researcher to extensive knowledge of the area of research. While accessed data are important, data correlation is not always proportionate to causation (Pennebaker et al. 2015). It must be noted that even though causation can be used to support qualitative research attempting to prove causality can be quite a challenge. Therefore, qualitative research can become inconclusive and may require further research which is labour and time consuming.

4.2 Ethical, Economical, Legal, Social and Political Issues

As autonomous Illegal data access, data archiving, remote access and people's attitude to data and privacy are covered in this research, the ethical and social implications are affected.

The research points to the need to address issues like data handling and management, privacy, education and implications of data misuse and access.

This research also considered policy development and its impact on the access, use and storage of data, including privacy when it comes to confidentiality, integrity and availability (CIA).

Legal and regulatory implications were carefully reviewed. Hence, UK legislation like Computer misuse act 1990, malicious communications Act, 2003 and the General Data Protection Regulation (GDPR 2018) which governs data privacy in EU and UK, is referenced. In addition, the penalties imposed on organisations that breach data privacy by Information commission office (ICO) in the UK is highlighted.

Political, Economical and Societal implications of the use of AI/ML/Dl as cyber security defence tools and legislative concerns were also covered.

5 Research Findings and Critical Discussion

As the use of digital systems (mobile phones, computers and others) increases, so does the appetite of attackers to exploit vulnerability in systems. As this appetite grows, so does the greed to continue to demand bigger ransom from victims of cyber attack.

Attack reports from various researches suggest that cyber attacks have moved on from individual attacks to attacks generated by cleverly programmed AI/ML/DL agents.

This rapidly expanding adoption of AI/ML technologies as attack agents, however, has rendered them attractive targets to adversaries who want to manipulate such mechanisms for malevolent purposes Liu et al. [12].

Several researches show that Machine Learning (ML) algorithms, such as Naïve Bayes, convolutional Neural Networks (CNN), LTSM, and Neural Network Support Vector Machine, can be used to detect and neutralise auto generated cyber attacks.

To make many of the these systems to work, the use of the combination of ML/DL etc. (hybrid) detection solution is more effective, because they have built in natural ML, Natural Language Processing (NLP), and Bayesian classifiers. Further research suggests that the detection efficiency is highly dependent on the type of framework deployed.

To combat malware in the cloud, researchers Sun et al. [38] developed a system called Cloud Eyes, which is based on trusted security services for resource-constrained devices and is called Cloud Eyes.

Cloud Eyes demonstrated suspicious bucket cross-filtering for the cloud server, as well as a signature detection mechanism based on the reversible sketch structure, which was provided by retrospective and accurate introductions of malicious signature sections. Cloud Eyes conducted a study in which they implemented a lightweight scanning operator that makes use of the process of signature sections to significantly reduce the scope of accurate matching.

The authors of Gu et al. [39] recommended that feature modelling be performed using a statistical examination strategy in order to extract malware family features, such as the software package feature, consent and application feature, and function call feature, from malware. Additionally, in order to reduce the false-positive rate while simultaneously increasing the detecting capacity of malware variants, a multi-feature detection technique for Android-based systems for detecting and classifying malware is being developed. Furthermore, using block chain technology, it was able to compile a database of malicious Android code that had been distributed.

An anomaly-based detection technique makes use of the information contained in the system's behaviour in order to determine the malignancy of a programme currently under investigation. When it comes to anomaly-based detection, there are two phases to consider: the training (learning) phase and the detection (monitoring) phase. During the training phase, the detec- tor makes an effort to become acquainted with the normal operation of the system. During the training phase, the detector observes and learns the behaviour of the host, the PUI (Perceptual User Interface), or a combination of the two during the experiment. One of the most significant advantages of anomaly-based detection is its ability to detect zero-day cyberattacks. Zero-day attacks are similar to zero-day exploits in that they are difficult to detect by malware detectors.

Its high false alarm rate and the difficulty in determining which features should be learned during the training phase are the two most significant limitations of this

technique, according to the authors. To achieve our primary objective of developing a malware detection system for cloud environments, we have developed an intrusion detection system that makes use of a consolidated WFCM-AANN. Traditional malware detection systems are ineffective in cloud environments because network-based intrusion detection systems (NIDS) are unable to detect encrypted node communication. Furthermore, host-based intrusion detection systems (HIDS) are incapable of detecting the hidden attack trail. This proposed work is divided into two modules, which are the clustering module and the classification module, respectively. The input dataset is grouped into clusters in the clustering module, which makes use of Weighted Fuzzy C-means clustering to accomplish this (MFCM). The cluster centroid is given to the intermittent Auto Associative Neural Network in the classification module, which is used to determine whether or not the information has been intruded into the clusters.

According to them, ARC utilised risk assessment model that uses a Hierarchical Risk Correlation Tree (HRCT) that models the paths an attacker can follow to reach certain goals and calculates the financial risk that the CPS assets face from cyber attacks. ARC also uses a Competitive Markov Decision Process (CMDP) to model the security reciprocal interaction between the protection system and the attacker/adversary as a multi-step, sequential, two player stochastic game in which each player tries to maximize his/her benefit.

The experiments' results depict that the accuracy of ARC outperforms the traditional Static Intrusion Response System (S-IRS) by 43.61%.

Undoubtedly, current literatures recommend the use of anomaly detection algorithms to detect anomalies detected by AI and ML frameworks. However, the combined use of AI, ML, and DL algorithms to detect autonomous malware yields a high detection accuracy.

As gap analysis showed from all the reviews, the use of enhanced Information security event tools is a pragmatic approach in detecting autonomous malware.

While there were some challenges during data collection, the exercise was successfully achieved.

Data was divided into three datasets, namely detection, countermeasures, and AI/DL/ML algorithms for detecting autonomous malware were identified and discussed.

In order to construct a probability generation model, which is referred to as DBN, different layers of the Restricted Boltzmann Machine are used (Deep Belief Networks).

Using Convolutional neural networks (CNNs) and static malware gene sequences (MCSMGS), Meng et al. (2017) presented a modern malware classification model based on static malware gene sequences (MCSMGS) for classification of malware, which they dubbed "malware classifications." First and foremost, the extraction of malware gene sequences from material and informational attributes is carried out. Second, make an attempt to delineate the similarities and associations between each malicious programme in order to determine their delineation.

At the end of the day, a module known as static malware gene sequences-convolution neural networks (SMGS-CNN) was used to improve the accuracy of

malware classification as well as the analysis of the malware gene sequences extracted from the malware samples. They came to the conclusion that the proposed scheme, which was found to be more effective than the SVM model, could achieve classified accuracy of 98 percent. As a second point of distinction from the other CNN solutions, this one extracts information from every layer of the network, in addition to applying a few-shot intrusion detection technique that makes use of a l-nearest neighbour classifier and a linear SVM. If we only have a small training set for a particular class, then few-shot learning may be the most appropriate learning method.

Finally, they made use of the proposed scheme by incorporating it into two well-known public datasets: KDD99 and NSL-KDD. In an experiment, the proposed scheme outperformed the previously used schemes on these datasets, despite the fact that these two datasets are dissimilar and some classes have a smaller training sample than other classes.

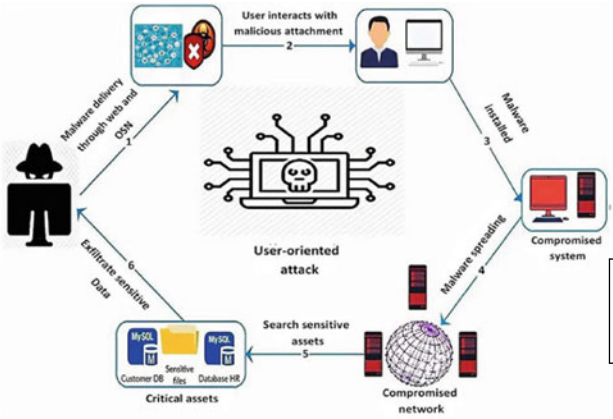
Techniques for detecting autonomous bots were investigated in this study. Attacks are detected using abnormal behavior-based information detection, which makes use of machine learning models. In order to detect and flag messages based on their content and linguistic characteristics, the algorithm employs classification models. Using an algorithmic detection system, the attacker's mood and comprehensibility are captured. To achieve high detection rates, a combination of diverse detection models is used in conjunction with one another.

However, while linguistic algorithms are associated with a high level of autonomous malware attacks, the challenge is in training the algorithm with the appropriate data sets to be effective. For example, in another study, it was discovered that artificial intelligence accuracy can be compromised in order to produce deceptive results by training the AI using the DL algorithm to provide incorrect information. According to Jiansg et al. [29], Particle Swarm Optimisation (PSO) can be used to combat the threat of data poisoning attacks by focusing on how data is introduced to the algorithm during the training phase in order to ensure that high accuracy is maintained.

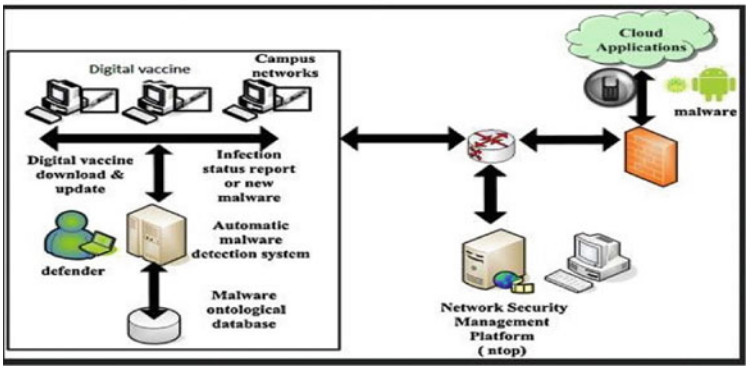
When solving classification problems with an SVM classifier (SVC), the generalisation ability of learning is maximised while also incorporating the Lagrange multiplier optimisation algorithm, resulting in a low number of classification errors being produced.

In general, results obtained through the use of SVM classification algorithms are more accurate than those obtained through the use of other machine learning approaches involving non-optimized search methods, such as artificial neural networks, least squares, k-nearest neighbour, Bayesian probability, and classification and regression trees, particularly when defence systems collect only a limited amount of data for training attack.

5.1 Detecting and Preventing Autonomous Malware



Schematic diagram of Lab set up to detect autonomous malware infection.



See Fig. 2.

5.2 Implementing and Detecting Countermeasures

For the purpose of this research, a commercially available Security Information Events Management (SIEM) tool was installed in a VMware laboratory settings which consisted of virtual networks, application servers, active directory and some datasets to test how the proposed concept would deal with suspicious data. The commercial tool allows the examination of related reports and alerts raised when anomalies are detected. The tool's advanced intelligent engine carries out machine analytics on the data received to seek out anomalies and suspicious data behaviour

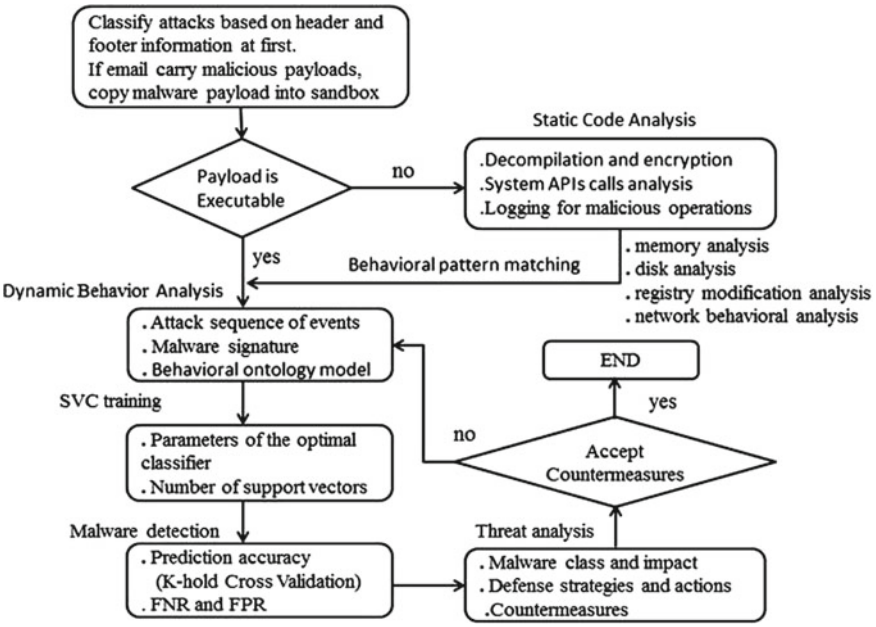


Fig. 2 The execution process of Small Vector Models (SVM) training and malware detection. *Source* https://www.sciencedirect.com/science/article/pii/S0022000014001780#fg0040

patterns. It then categorises these activities into priorities before taken autonomous corrective action.

SIEM's primary role is to collect, digest and analyse data received across the networks and related and connected security systems to identify and access connected applications. It analyses operating system and application logs, identity management, policy compliance and vulnerability management as well as logs from external sources that potentially pose as threat. It monitors user access privileges, directory services and system configuration changes and auditing logs and reviewing incident responses.

The SIEM tool used in this research was able to provide high level security analytics across the isolated lab setup used to demonstrate how it deals with threat.

The section below discuss the proposed methodology and overall concept using the SIEM tool DL and ML tools including the learning algorithm to show how it will deal with autonomous malware attack based on the Mitre Att&ck Framework.

5.3 Mitre ATT&CK

MITRE ATT&CK® is a well known publicly available knowledgebase that documents adversary attack techniques which have been documented over time from field

observations. Using various attack techniques, the vectors aim is to attain a foothold with initial access. The knowledgebase is used to in both private and government sectors to develop threat models and cyber security products that serves the service economy.

Broken down into twelve tactics, the ATT&CK Matrix framework consists of techniques used by attackers to breach cyber security systems.

1. Initial Access—Involves the utilisation of various entry vectors to gain initial access through the use of valid accounts or via remote services.
2. Execution—Tactics involves the running of controlled malicious codes on local or remote system in combination with other tactics to achieve wider goals.
3. Persistence—Tactics involves adversaries ensuring that they keep access to systems open to reduce being cut off to allow them to continue breach.
4. Privilege escalation—Tactics involves gaining higher-level access on the attacked system or network and usually used in conjunction with persistence tactics.
5. Defense evasion—Tactics involves avoiding detection while breaching attacked system.
6. Credential access—Tactic involves theft of user credentials like passwords and user names
7. Discovery—Tactics involves the use of several tactics to gain knowledge about internal systems. Usually, this is called social engineering tactics.
8. Lateral Movement—Tactics involves entering the system to control remote systems.
9. Collection—Tactics is similar to discovery in that attacker use various techniques to gather information about the system before attacking.
10. Command and control—Tactics involves methods used by attackers to control systems under attack.
11. Data exfiltration—Tactics involves data theft from the victim as attacker gains access.
12. Impact—Tactics involves disrupting the attackers victims processes and operation after the systems have been compromised. *Source* https://attack.mitre.org/

5.4 Malicious Autonomous Malware Attack

Configuration was made to detect any suspicious users based on predetermined users pulled from Windows Active Directory system. Suspicious datasets of users that were fed into the system were immediately flagged for attention by the SIEM tool.

Example Process Injection Method: Reflective DLL Injection.

Reflective DLL injection is one of the most commonly used process injection methods deployed by attackers. It involves executing a DLL inside another process by creating a DLL that maps itself into memory when called, rather than relying on Window's API's loader calls. It eludes the storage of DLL on the storages system and

```
powershell -c "Unblock-File %TMP%\Invoke-ReflectivePEInjection.ps1;

Import-Module %TMP%\Invoke-ReflectivePEInjection.ps1;
$PEBytes = [IO.File]::ReadAllBytes('%windir%\System32\calc.exe'); Invoke-
ReflectivePEInjection -PEBytes $PEBytes"
```

Fig. 3 Command used to invoke the reflective injection attack via an AI/DL

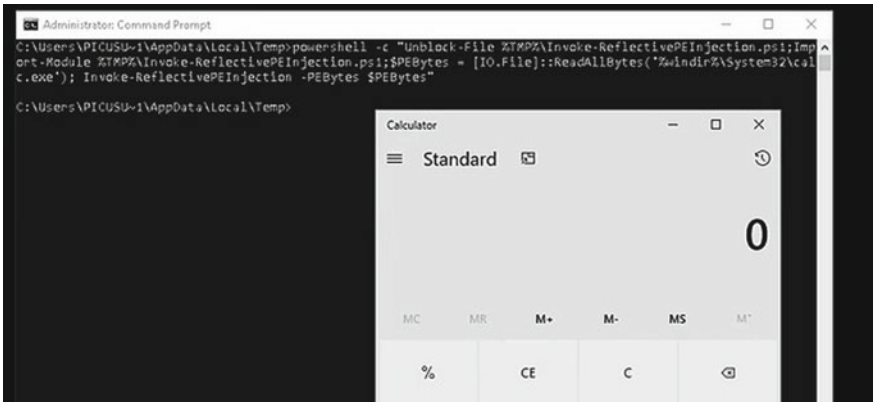


Fig. 4 Command used to invoke the reflective injection attack via an AI/DL

calling windows API's LoadLibrary that is likely to be detected by security tools, like SIEM.

For the purpose of this research, a commercially available software was used to demonstrate the attack by invoking the Reflective injection attack which can be used to simulate the reflective DLL injection technique attack using AI engine. In addition to loading a DLL or EXE into Windows PowerShell, the AI can reflectively load a DLL into a remote process (Figs. 3 and 4).

5.5 How to Auto Detect Autonomous Reflective Injection Attack (Critical Mitre Attack)

To auto detect the reflective DLL injection technique, configured systems like a SIEM tool embedded with autonomous cyber defence agent (ACyD) need logs (in this demonstration) that include PowerShell activities. Event log entries in Microsoft-Windows-PowerShell/Operational log includes such activities. The Event ID 4104 (sample script block logging) records is a legitimate blocks of code as they are genuinely executed by the PowerShell engine. Script block logging captures the

deciphered full contents of the code as it is executed. This includes the scripts and commands, as demonstrated in Fig. 5.

When the DLL is autonomously injected into the target process, the malware has to map the DLL's raw binary into virtual memory. It uses kernel32.dll and VirtualAlloc, GetProcAddress, and LoadLibraryA functions to fetch the right address of the injected export function. The researchers of this hack have developed rules that took advantage of this by utilizing the Microsoft-Windows-PowerShell/Operational log with the Event ID 4104.

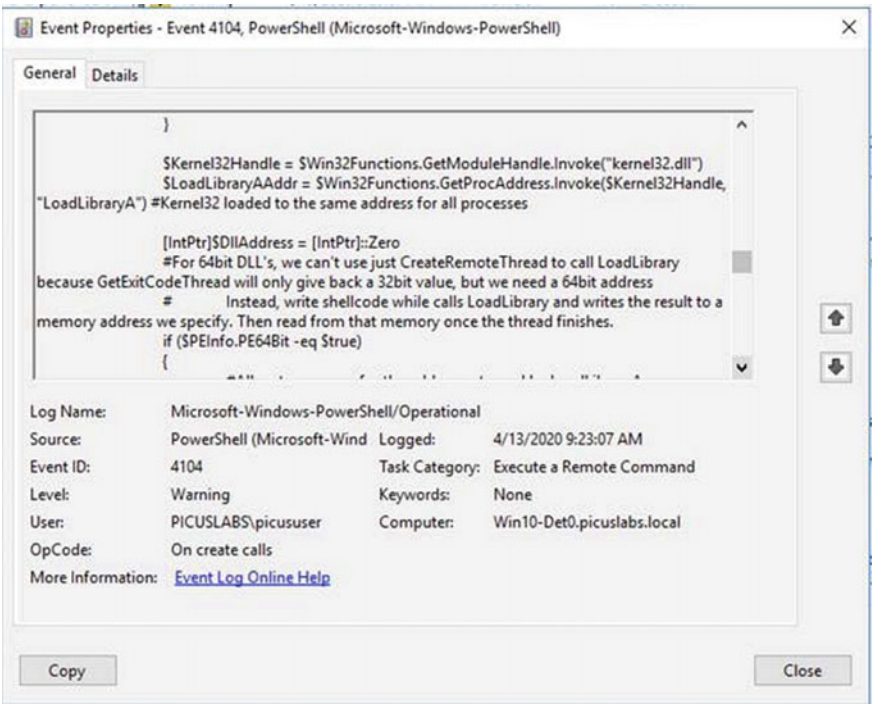


Fig. 5 Screen shot of activities in autonomous reflective DLL injection attack

```
Title: Reflective Portable Executable Injection via PowerShell

status: stable


description: Detects the attempt of reflective portable executable (DLL/EXE)
injection by PowerShell that uses API calls. This method is used by adversaries
to evade detection from security products since the execution is masked under a
legitimate process.
author: XXX(Redacted)


references:
   - https://attack.mitre.org/techniques/XXX/(Redacted)
  - https://attack.mitre.org/tactics/XXX/(Redacted)
   - https://attack.mitre.org/tactics/XXX/ (Redacted)


logsource:
   product: windows
   service: powershell/operational
   definition1: 'Requirements: Group Policy : Computer
Configuration\Administrative Templates\Windows Components\Windows
PowerShell\Turn On Module Logging'
   definition2: 'Requirements: Group Policy : Computer
Configuration\Administrative Templates\Windows Components\Windows
PowerShell\Turn On PowerShell Script Block Logging'


detection:
   selection:
      EventID: 4104
   keyword1:
      - '*kernel32.dll*'
   keyword2:
      - '*LoadLibraryA*'
   keyword3:
      - '*GetProcAddress*'
   keyword4:
      - '*VirtualAlloc*'
```

Fig. 6 Raw data screen dump of how the attack was classified after autonomous detection

5.6 *Analysis of the Detection and Corrective Action Taken*

As demonstrated in the flowchart above, the malware autonomous cyber defence module is made up of four functions: (1) System backup, (2) Systems monitoring, (3) System recovery, and (4) Passing intelligent data to DL/ML (Figs. 6 and 7).

Role of Each Function in the Defence and Remedial Cycle
The system backup ensures that key system data is backed up both on client and hosts. The system monitoring's role is to restore the operating system if it was corrupted by malware attack. The systems recovery's role is to recover the entire system using reverse engineering process after analysing the malware's signature and behaviour. This includes checking files for changes, privilege and security settings changes, hash and configuration values changes and file attribute and size changes.

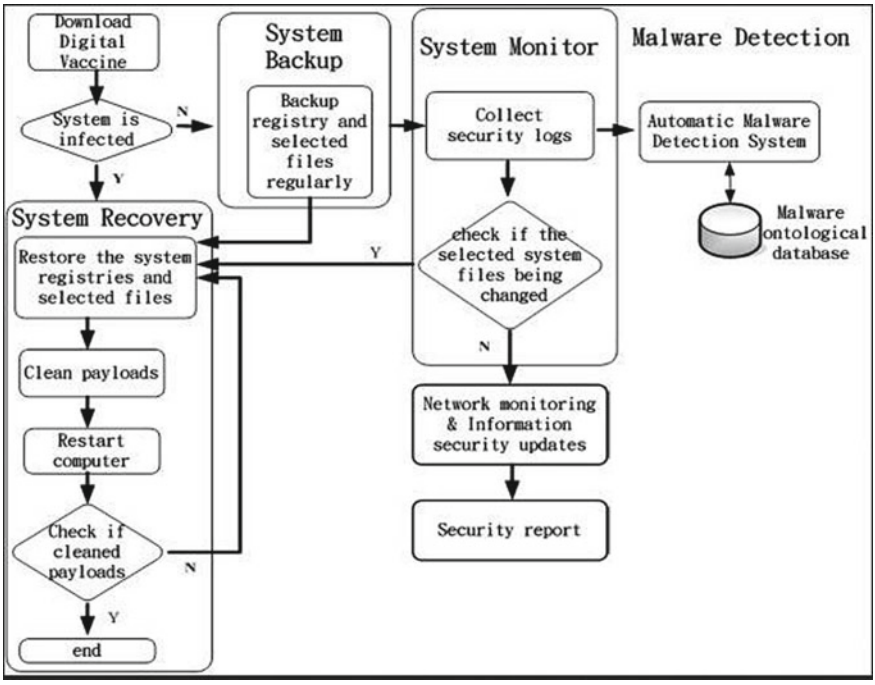


Fig. 7 Flowchart of autonomous corrective action (defence and remedial cycle). *Source* https://www.sciencedirect.com/science/article/pii/S0022000014001780#fg0040

5.7 User Behaviour Anomaly

The SIEM tool uses trend analysis to study user behaviour and pick out any abnormal deviations. Once this has been established, the behaviour is flagged as suspicious and alert is raised.

- Compliance violations
  SIEM tool was programmed to study incoming logs to detect data accuracy to distinguish illegally altered data not conforming to set compliance standard.
- Disruption of ICT services
  Set rules were set to detect malware activities to detect identity spoofing and Ddos attacks which could lead to the attacker taking over the ICT services, and demand ransom. Additionally, the SIEM tool, Network monitoring, (NM) was set to the entire network for traffic and packet analysis of known protocols and flag any anomaly.
- Controlling impacts of Botnets Control
  Smart AI/DL use botnet to infiltrate weak networks and communicate through command and control centres. The SIEM tool was programmed to detect illegal use of legitimate ports to divert traffic undetected. For example, the tool was programmed to persistently monitor traffic on port 80 for malicious http packet headers.

The SIEM tool Network monitoring tool gives detailed analytics on the entire network it is programmed to monitor. The NM 4.0 produces a simplified dashboard showing the status of collected data and monitored traffic which can then be further analysed for suspicious activity (Fig. 8).

The SIEM tool used for this research utilises a special function called SmartResponse. In this research, it is proposed that the SmartResponse be enhanced with

Fig. 8 Network traffic monitoring

ACyD (please refer to flowchart). The combined use of ACyD and SmartResponse will carry out auto response without the need for any manual intervention once threat is detected.

The combined force will automatically terminate unauthorised activities on detection. It will scan disable unused remote access/Remote Desktop Protocol ports:

Only permit systems to action legitimate and authorized programs as set out clearly in an established security policy.

Persistently audit administrative user accounts on a regular basis and adjust access controls to comply with least privilege access principle, to ensure that only legitimate accounts will access the systems resources that they are entitled to.

Persistently scan the network for listening and open ports and close all ports not required.

5.8 Education and Training

Several researches have shown that best method used by organisations to educate staff about security is to use fact-and-advice session, which are either instructor-led or computer-based training (CBT). For example, the use of video game which is fun, enjoyable and educational has been known to be very successful because participants get to learn to attack and exploit vulnerabilities in a dynamic setting and how to react to attacks and deploy countermeasures. Furthermore, participants get to master the concepts of cyber security from both adversarial and defence viewpoint.

It is well documented that cyber security risk is paramount for every organisation. Hence, implementing frameworks like Cyber essential skills from National Cyber Security Centre UK (NCSC), National Institute of Standards and Technologies (NIST), the IEC 62443 International Electrotechnical Commission for industrial control system and the ISO 27001 International Organization for Standardization for information security, will assist organisations in raising cyber security awareness within their staff. These standards assist organisations to identify and assess risks from potential attackers and help them choose the right procedure to be used for countermeasures.

For example, one of the most notorious hacker ever known, Kevin Mitnick is quoted to have said, “A company can spend hundreds of thousands of dollars on firewalls, intrusion detection systems and encryption and other security technologies, but if an attacker can call one trusted person within the company, and that person complies, and if the attacker gets in, then all that money spent in technology is essentially wasted”.

As the barest minimum, all staff that use computer systems must undergo internet security training and be made aware of all company policies on the use of computer systems. Being aware of policies like computer usage policy and others like information security policy, records management policy, retention and disposal schedules, archiving policy, data privacy policy, ICT policy, information sharing policy and

remote working policy, are essential, and everyone must undergo that training to secure organisation information security asset.

5.9 Legislative and Regulatory Controls

Many research outcomes have shown that having strong information governance that complements applicable government set of laws is good practice. As a minimum, data protection and information security policies that safeguards data to ensure that all information is legally and properly processed must be present.

Without a doubt, complying with General Data Protection Regulation (GDPR 2018) is compulsory to ensure that all personal data processed is legal. The consequence for non-compliance is serious for UK organisations. For example, the information commission office (ICO 2021) is legally bound to enforce breaches and impose fines for companies that do not comply with information security regulations. It is also important that Computer Misuse Act 1998 and the Information Technology Infrastructure Library (ITIL) guidelines are followed as they protect organisations information asset.

While the regulations are clear on consequences of non-compliance with data security, legislators have a difficulty in legislating against violation caused by AI/DL/ML, which is the focus of this research.

At present, there are no specific legislation to govern the action of AI. This is partly because laws have to be technology independent to make sure that future technology will still be subject to an overarching legal framework.

Listed below are recommended list of what organisations must comply with to ensure that data is secure:

6 Conclusion and Future Works

The main purpose of this research is to evaluate the challenges and opportunities in using AI and ML as decision tools against cyber attacks, and propose the future use of Autonomous Cyber Defence (ACyD) in tackling Autonomous intelligent Malware (AIM) as a tool to be embedded into Security Information and Events Management systems (SIEM).

However, based on numerous conclusions from peer reviewed literatures on the use of AI/ML/DL to auto detect cyber attacks, there wasn't a concrete detection mechanism currently adopted as a standard framework to stop auto generated cyber threats. Hence, there is a big gap in this area. Therefore, there is future work to be done to establish proper framework for combating auto generated cyber threats via the use of Intelligent malware generated by AI/DL/Ml.

On one hand, the technical research carried out using the ISE tools successfully demonstrated how Intelligent malware can be auto generated via an AI/Dl/ML

and how this can be stopped via auto generated defence using AI/DL/ML tools, conversely, this proposed framework is not widely used due to the fact that the ISE used to demonstrate this gap is a commercial product. Hence, the key barrier to entry to use this method is the high cost of implementation.

While some specific commercial tools and other simulation products were used in this research, they could be improved via the use of top Gartner quadrant SIEMs, File Integrity Monitoring, Endpoint detection and recovery tools and Network Monitoring commercial tools to provide agile setups/configurations against auto generated intelligent malware attacks. The detection and protection of data must be the ultimate goal to maintain confidentiality, Integrity and Availability (CIA) of data in every organisation.

Ongoing user awareness training should be made mandatory to ensure that everyone working in any organisation is equipped with the relevant knowledge to ensure that data security is safeguarded. Organizational security awareness specifically deals with staff attitude to data security and protecting organisation's information asset. This training framework must incorporate the training of auxiliary workers that come in contact with organisation information asset, including the use of non-disclosure clauses in contract to ensure information is asset is protected.

Training should also cover topics such as; proper methods for protecting sensitive information on computer systems, such as password policies and the use of two- factor authentication as well as and how to deal with Personal Identifiable Information (PII). Other computer security concerns, such as malware, phishing, and social engineering scams should be prominent in such programmes.

According to the European Network and Information Security Agency, "Awareness of the risks and available safeguards is the first line of defence for the security of information systems and networks."

Consequently, strong security governance framework must be at the core of organisation's plan.

In the words of a well known hacker, "A company can spend hundreds of thousands of dollars on firewalls, intrusion detection systems and encryption and other security technologies, but if an attacker can call one trusted person within the company, and that person complies, and if the attacker gets in, then all that money spent in technology is essentially wasted". Hence, the importance of training to raise information security awareness cannot be over emphasised.

While the use of AI/ML/DL is advocated in this research, there are key concern areas that more research is required. One key area that requires further research is the key concerns around the societal/ethical concerns surrounding the use of AI to auto eliminate cyber threats. While it is accepted that AI/ML/Dl have made great advancement in technological research, the fact that AI/ML/DL has the capability to regenerate and re-invent itself, is cause for concern. One of such key concern areas is Algorithmic bias which occurs when AI/ML "algorithm produces results that are systemically prejudiced due to erroneous assumptions in the machine learning process". The dangers of this has already been shown when Microsoft AI (Tay.Ai) in 2016, was able to learn to become racist in sixteen hours (The Guardian 2016) and was eventually taken offline due to the number of complaints that came in.

In a sci-fi depiction of what is possible with AI autonomous decision, BBC Click in its programme, "Is autonomous weapons a threat to humanity?" [40], the presenter, Spencer Kelly, was able to demonstrate the power of autonomous decision from AI/ML. It was obvious that the idea need to be carefully thought out because at some point, the AI through ML became independent of the human control.

Ethical concerns have been raised by the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems and a proposal to look into IEEE Standard on Algorithmic Bias Considerations has been lodged.

While these concerns are live and valid, the use of autonomous decision for cyber threat hunting is a game changer because it will invert the use of autonomous cyber defence for reactive mitigation and become pro-active offence/defence tool; an area still open for future research.

The researcher strongly believes that there is a big market for auto generated cyber defence, however, the human intervention is required for some specialist decisions. Hence, a combination of human and AI interaction cannot be ignored.

References

1. Chen H, Wang FY (2005) Guest editors' introduction: artificial intelligence for homeland security. IEEE Intell Syst 20(5):12–16
2. Dasgupta D (2006) Advances in artificial immune systems. IEEE Comput Intell Mag 1(4):40–49
3. Huang K, Siegel M, Madnick S (2018) Systematically understanding the cyber attack business: a survey. ACM Comput Surv (CSUR) 51(4):1–36
4. Bonab AB, Rudko I, Bellini F (2021) A review and a proposal about socio-economic impacts of artificial intelligence. In: Business revolution in a digital era, pp 251–270
5. Théron P, Kott A (2019) When autonomous intelligent goodware will fight autonomous intelligent malware: a possible future of cyber defense. In: MILCOM 2019–2019 IEEE military communications conference (MILCOM), pp 1–7. IEEE
6. Lewis JA (2018) Rethinking cybersecurity: strategy, mass effect, and states. Rowman & Littlefield
7. PWC P (2017) Embed ethics within business practices. Strat Fin
8. Hatcher WG, Yu W (2018) A survey of deep learning: platforms, applications and emerging research trends. IEEE Access 6:24411–24432
9. Golovko VA (2017) Deep learning: an overview and main paradigms. Opt Mem Neural Netw 26(1):1–17
10. Harnad S (1992) The turing test is not a trick: turing indistinguishability is a scientific criterion. ACM SIGART Bull 3(4):9–10
11. Akhtar N, Mian A (2018) Threat of adversarial attacks on deep learning in computer vision: a survey. IEEE Access 6:14410–14430
12. Liu Z, Li J, Li J, Jia C, Yang J, Yuan K (2014) SQL-based fuzzy query medianism over encrypted database. Int J Data Warehouse Min (1JDWM) 70(4):71–87
13. Kuzlu M, Fair C, Guler O (2021) Role of artificial intelligence in the Internet of Things (IoT) cybersecurity. Discover Internet of Things 1(1):1–14
14. Cyber Threat Intelligence (2014) https://www.bibliotecadeseguranca.com.br/wp-content/uploads/2015/09/EY-cyber-threat-intelligence-how-to-get-ahead-of-cybercrime.pdf
15. Goldman ZK, McCoy D (2015) Deterring financially motivated cybercrime. J Nat'l Sec L Pol'y 8:595

16. Rasthofer S, Arzt S, Miltenberger M, Bodden E (2016). Harvesting runtime values in android applications that feature anti-analysis techniques. In: NDSS, Feb 2016
17. Wong MY, Lie D (2016) IntelliDroid: a targeted input generator for the dynamic analysis of android malware. In: NDSS, vol 16, pp 21–24, Feb 2016
18. Avdiienko V, Kuznetsov K, Gorla A, Zeller A, Arzt S, Rasthofer S, Bodden E (2015) Mining apps for abnormal usage of sensitive data. In: 2015 IEEE/ACM 37th IEEE international conference on software engineering, vol 1. IEEE, May 2015, pp 426–436
19. Russell, Norvig S (2010) Artificial intelligence. In: A modern approach. Prentice Hall, Englewood Cliffs, NJ
20. Kholidy HA (2021) Autonomous mitigation of cyber risks in the cyber-physical systems. Futur Gener Comput Syst 115:171–187
21. Kaplan A, Haenlein M (2020) Rulers of the world, unite! The challenges and opportunities of artificial intelligence. Bus Horiz 63(1):37–50
22. Xu Z, Ray S, Subramanyan P, Malik S (2017) Malware detection using machine learning based analysis of virtual memory access patterns. In: Design, Automation & Test in Europe Conference & Exhibition (DATE 2017), pp 169–174. https://doi.org/10.23919/DATE.2017.7926977
23. Hashemi H, Azmoodeh A, Hamzeh A, Hashemi S (2017) Graph embedding as a new approach for unknown malware detection. J Comput Virol Hacking Tech 13. https://doi.org/10.1007/s11416-016-0278-y
24. Ye Y, Chen L, Hou S et al. (2018) DeepAM: a heterogeneous deep learning framework for intelligent malware detection. Knowl Inf Syst 54, 265–285. https://doi.org/10.1007/s10115-017-1058-9
25. Olalere M, Abdullah MT, Mahmod R, Abdullah A (2016) Identification and evaluation of discriminative lexical features of malware URL for Real-Time classification. http://repository.futminna.edu.ng:8080/jspui/handle/123456789/10598
26. Karbab EB, Debbabi M, Derhab A, Mouheb D (2020) Scalable and robust unsupervised android malware fingerprinting using community-based network partitioning. Comput Secur 97:101965
27. Wang HH, Yu L, Tian SW, Luo SQ, Pei XJ (2020) Malicious webpages analysis and detection algorithm based on BiLSTM. Int J Electron Bus 15(4):351–367
28. Li Jh (2018) Cyber security meets artificial intelligence: a survey. Frontiers Inf Technol Electronic Eng 19:1462–1474. https://doi.org/10.1631/FITEE.1800573
29. Jiang J, Han F, Ling Q, Wang J, Li T, Han H (2020) Efficient network architecture search via multiobjective particle swarm optimization based on decomposition. Neural Netw 123:305–316
30. Dwivedi S, Vardhan M, Tripathi S (2020) An effect of chaos grasshopper optimization algorithm for protection of network infrastructure. Comput Netw 176:107251
31. Yusof AR, Udzir NI (2019) Systematic literature review and taxonomy for DDoS attack detection and prediction. Int J Digit Enterp Technol 1(3)
32. Najada HA, Mahgoub I, Mohammed I (2018) Cyber intrusion prediction and taxonomy system using deep learning and distributed big data processing. In: 2018 IEEE Symposium Series on Computational Intelligence (SSCI), pp 631–638. https://doi.org/10.1109/SSCI.2018.8628685
33. Mane N, Verma A, Arya A (2020) A pragmatic optimal approach for detection of cyber attacks using genetic programming. In: 2020 IEEE 20th international symposium on computational intelligence and informatics (CINTI). IEEE, Nov 2020, pp 71–76
34. Geetha R, Thilagam T (2021) A review on the effectiveness of machine learning and deep learning algorithms for cyber security. Arch Comput Meth Eng 28(4):2861–2879
35. Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf AJ (2020) A systematic review on supervised and unsupervised machine learning algorithms for data science. In: Supervised and unsupervised learning for data science, pp 3–21
36. Firlej M, Taeihagh A (2021) Regulating human control over autonomous systems. Regul Gov 15(4):1071–1091
37. Taeihagh A, Ramesh M, Howlett M (2021) Assessing the regulatory challenges of emerging disruptive technologies. Regul Gov

38. Sun X, Xie Y, Luo P, Wang L (2017) A dataset for benchmarking image-based localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 7436–7444
39. Gu J, Sun B, Du X , Wang J, Zhuang Y, Wang Z (2018) Consortium blockchain-based malware detection in mobile devices. In: IEEE Access 6, pp 12118–12128. https://doi.org/10.1109/ACCESS.2018.2805783
40. Köchling A, Wehner MC (2020) Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. Bus Res 1–54

Secure Deployment of IOT Devices

Setareh Jalali Ghazaani, Michael Faulks, and Sina Pournouri

Abstract The ubiquitous internet of things significantly improves every aspect of our daily lives. IoT devices and their use remain a big area of opportunity, but they are complicated by a lack of regulation as well as numerous security and privacy issues caused by design and setup flaws. Many current attacks against SMEs demonstrate that IoT devices make the networks vulnerable and expand the attack surface. Considering the widespread use of IoT devices and the security flaws they have, various parties have tried to provide security frameworks to teach users how to securely deploy these devices. They aimed to advocate that IoT devices should be subjected to strict security and privacy rules in isolated subnetworks, which has been proven to be a promising technique for securing networks, devices, and data. However, these frameworks are aimed at IT professionals rather than average users. In this study, we tried to educate normal users to securely deploy IoT devices. To achieve this goal, we have provided a set of best practices collected from existing standard frameworks. We have demonstrated the implementation of these security measures in two different scenarios using various network devices and with consideration of SME limitations. Some of the security measures are directly related to the device, and there is not much the consumer can do. However, if the technology is supported by the device, the users should be educated accordingly. To successfully achieve the aim of the study, we will investigate the existing vulnerabilities of smart devices and evaluate the existing guidelines for secure deployment of IoT devices. Then we will implement the current best practices for safeguarding computer networks, with a focus on IoT challenges and finally, we will pave the way to propose a practical framework for safely deploying IoT devices in small and medium enterprises.

Keywords IoT devices · OWASP · DDoS · Cyberattack · Security policy · Privacy · Ethics

S. J. Ghazaani · M. Faulks · S. Pournouri (✉)
Sheffield Hallam University, Sheffield, UK
e-mail: S.Pournouri@shu.ac.uk

H. Jahankhani et al. (eds.), *Blockchain and Other Emerging Technologies for Digital Business Strategies*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-98225-6_10

1 Literature Review

The Internet of Things, or IoT, was first created by Kevin Ashton in 1999 and refers to a network of interconnected objects that exchange data and communicate with one another [1]. This technology allows two or more devices to send and receive data over the internet by connecting to each other in a more efficient way than human interconnections [2]. Although these devices led to an easier life for their users, little consideration was paid to the security aspect. Many search engines and websites, such as Shodan and Insecam, release a list of accessible vulnerable devices across the internet, leaving the IoT environment at the risk of exploiting the vulnerabilities and eventually data breaches. For example, the DDoS attack against Dyn, a major DNS service provider, was conducted by compromising thousands of IP-Cameras in its network [3]. IoT devices expose UDP or TCP ports to both local and global entities across the network, so both insiders and outside attackers can exploit the vulnerabilities in both the IoT device and the network. In this chapter, we will discuss the structure of IoT devices, their security challenges, and the attack scenarios at each layer.

1.1 IoT Architecture

Each layer of the Internet of Things is described by the roles and instruments that are utilized in that layer. There are a few different points of view on the number of IoT layers. The number of the layers depends on the IoT application [4]. Based on the first and basic architecture, IoT devices are made up of three layers: the perception layer, the network layer, and the application layer, each of which has its associated security concerns. There is no unique approach to architecture. Because of the diverse forms of attacks, the offered security solutions for one architecture may not be suitable for another [5]. To safeguard these devices from data loss and tampering, these levels must be protected and secure [6]. The application layer provides a specific application for users to connect to the IoT devices. The network layer is responsible for processing and transmitting data, and the perception layer includes the hardware part of the device.

Because IoT is related to real-world phenomena like healthcare, agriculture, grids, and weather, and choices are made based on sensing and monitoring, security is a top priority [7]. Both data and device security should be considered to provide security. The integrity and confidentiality of data are the fundamental concerns of data security, whilst systems must be safeguarded from stealthy attacks [8].

1.2 IoT Security Challenges

The explosive growth of the Internet of Things necessitates the implementation of appropriate security and privacy regulations to avoid any system vulnerabilities or threats. Reliability, scalability, and power consumption are all major problems in the IoT devices. All of the aforementioned layers are critical when it comes to security because the ultimate goal of an attacker is to shut down a service or gain unauthorized access to a specific piece of information by assaulting a single layer [9]. According to the Open Web Application Security Project (OWASP), IoT security mitigations should be applied to all of the IoT tiers [10]. In this section the security challenges for IoT devices have been discussed.

1.2.1 Perception Layer

As vital as safeguarding data transferred to or from the device is securing the physical device itself. The most common risks in the physical layer are due to the poor hardware and application protection of IoT equipment. To protect the obtained data and the privacy of the owners, the devices at this level must be secured [11]. Hardware selection is also important for software selection and protecting IoT devices. Due to their low radio frequency operation mode and computational capabilities, IoT devices have limited capacity for transmitting big messages. In terms of hardware, IoT devices rely on microcontrollers, which have a wide range of performance characteristics. 8-bit, 16-bit, and 32-bit microcontrollers are the most popular microcontrollers available. As per [12], For certain ultra-low-cost applications, 4-bit microcontrollers are popular. Only a few simple commands are normally included in these ultra-low-cost microcontrollers. As a result, typical cryptographic algorithms will require a large number of cycles to perform, making them inefficient in terms of both time and energy for applications employing these devices. Furthermore, some microcontrollers have limited amount of RAM and read-only memory (ROM) [5].

1.2.2 Network Layer

Because of the variety of devices and communication protocols used in IoT systems, as well as the many interfaces and services available, conventional IT network solutions are not adequate for implementing security mitigation. In fact, the conventional security measures used in a traditional network might not be adequate. Lack of transport encryption, for example, refers to an insecure communication link between a device and the Cloud, a device and a gateway, a device and mobile applications, a device and another device, and communication between a gateway and the Cloud. However, to meet the needs of IoT operations, communication protocols such as Bluetooth, ZigBee, PLC, WiFi, 4G, and 5G may be used. Some benefits of this protocol are scalability and the fact that it may be self-maintained, requires little power, and

has a low operational cost [13]. Available security protocols for IoT devices were discussed and presented in [14]. It is observed that no protection against fragmentation attack have been provided in physical layer, network layer and application layer. Also, replay attack protection is also not supported in the physical layer, network layer and 6LoWPAN layer.

Encryption mechanism refers to the operation by which the confidentiality of the data would be guaranteed during the communication. The generic encryption methods cannot be applied to the IoT devices due to the resource limitation. RSA, AES and ECC are the common crypto systems that are widely used in IoT devices, nevertheless, the fast and simple nature of such lightweight cryptographic encoding methods made them vulnerable to the side channel analysis attacks [15]. The main obstacle to encrypting the devices is the device's simplicity, such as sensors. However, to protect the privacy and security of users, it is essential to include lightweight encryption in devices.

Authentication capabilities, end-to-end traffic encryption, a secure boot-loading procedure, the enforcement of digital signatures during firmware updates, and transparent transactions are all challenges with this layer [13]. Inadequate authentication and authorization protocols are a common vector for gaining access to IoT devices. MQTT, DDS, Zigbee, and Zwave are the protocols that now offer authentication in IoT systems. Unfortunately, most of these protocols are implemented insecurely, resulting in the loss of sensitive data such as device information, credentials, and network configuration data [16]. Reported that that defective deployment of MQTT, which functions as one of the main links of IoT communications, disclose sensitive information of both devices and servers to the adversaries using raw commands [17]. Furthermore, as MQTT may be used to update IoT devices' software and firmware, IoT devices are more vulnerable to cyberattacks [18]. Even if the developer has given the authentication mechanisms required for IoT communications, pairing, and messaging, the communication can still be intercepted. Furthermore, insecure network services may allow a bad actor or threat to explore and disseminate within the network. Authentication is currently the most prevalent security approach for achieving secure network connection. However, due to the constraints of the IoT devices, implementing IPsec in the context is not practical [16]. Also hardcoded credentials, which are commonly utilized in IoT devices, result in insufficient security configurability. Because multiple devices use the same password, and these passcodes are publicly available, hardcoded credentials are easy to corrupt.

1.2.3 Application Layer

In terms of IoT software, there are a variety of operating systems, for instance ARM Mbed, Brillo (Google Android Things), Ubuntu core, RIOT OS and Contiki OS, that are designed to work within the memory, size, and power constraints imposed by IoT devices. Security and privacy should be supported by IoT OS. Most of these standard operating systems, on the other hand, are incapable of meeting the security needs for IoT infrastructures [19]. Also, to increase the security of the IoT devices, they

must be patched and updated on a regular basis [12]. Nonetheless, one of the issues that will expose IoT devices to a variety of security vulnerabilities is the irregular security patching [20].

Due to the exposure and communication of IoT devices to the internet, new security requirements have been generated which application must adhere to [21]. Attackers can load malicious modules on nodes. Also it is possible to exploit the available vulnerabilities in the operating systems. Insecure web and cloud interfaces are weaknesses that could be exploited. As a result, security measures on cloud gateways are required to prevent malicious actors from altering configurations. At the application layer, biometrics and multi-level authentication could be a useful option for access control. Unreliable connectivity, hostile environments, and poor data and privilege protection are just a few of the issues that come with securing IoT devices [22].

There are several security challenges in IoT devices. Due to their limited resources, lack of security software, diversity of network interfaces (Ethernet, Bluetooth, Wi-Fi, ZigBee, to name but a few), and human factors like not changing default passwords, IoT devices are considerably easier to attack than traditional devices. Hence it can be concluded that IoT devices and smart networks are the main attack surface for adversaries [23]. The goal of this study is to provide a less vulnerable, safe network for restricted devices that can deliver all security and privacy functions.

1.3 IoT Attacks Based on Three-Layered IoT Architecture

By manipulating with physical weaknesses, abusing network or routing protocols, and breaking into devices using an encryption attack, an attacker can cause damage to IoT systems. Physical layer attacks target hardware devices, network layer attacks attack the IoT system's network, and application layer attacks use malware, trojans, and viruses to exploit the vulnerabilities in the application layer. In this section we will briefly discuss the possible attacks in each layer.

1.3.1 Physical Layer Attacks

The hardware devices of an IoT system are the target of physical attacks. The following are examples of physical attack

- Social Engineering: Social engineering is the process of deceiving an IoT system's end users in order to gain sensitive data [24].
- Radio Frequency Interference: An attacker using a device to disrupt the connectivity of IoT devices is known as radio frequency interference attack. When the attacker is near the device, jamming and RF interference can occur.
- Reverse Engineering: In order to find the vulnerabilities, the attacker breaks the targeted device down and step by step. Following the discovery of a list of known

and unknown vulnerabilities in the device, the attacker can use them to attack other devices in the same network.
- Tampering: the physical alteration of the device is known as tampering attack. This will result in access to login passwords, encryption keys, and other sensitive data [25].

1.3.2 Network Layer Attacks

As discussed earlier, the network layer is used by the IoT devices to send information from the physical layer to a server or another device for processing. The followings are some of the attacks and security vulnerabilities that have been discovered.

- DDOS Attack: A distributed denial of service attack occurs when an attacker floods a significant amount of traffic to a certain server. This will result in unavailability of the server and disturbing the customer traffic. It should be considered that DDoS attacks are not limited to IoT devices and applications. However due to the poor configuration of IoT devices, they are easy targets for attackers seeking to expand their botnet armies. As part of its attack [26].
- Spoofing: In IoT networks, data is encrypted and sent over the network using IP address-based routing protocols. To interrupt network traffic, an attacker can utilize transport protocols to copy, change, or resend IP addresses. An attacker can establish fake routing nodes, transmission pathways, and error messages to launch spoofing attacks in IoT networks [27].
- Data Transfer Attack: IoT applications generate a large amount of data that is in circulation from one site to another. This signal flows from sensors to servers to the cloud or to the application and deploy a variety of technologies. As a result, IoT applications are more vulnerable to attacks.
- Man in the Middle Attack: A man-in-the-middle attack collects and alters data sent between two nodes in an IoT network [28]. All nodes in the network can capture data packets that don't interact across devices. If desired, network devices can intercept and read the data. As a result, the purpose of this attack is to sabotage traffic by altering data sent over the IoT network. Many open and uncontrolled devices can be found in IoT networks. In an MITM attack, these insecure devices have the potential to be the source of traffic [24].
- Access Attack: Gaining an unauthorized access to the smart network is known as an access attack. The attacker can remain undetectable for a long period of time. The purpose of this attack is to steal sensitive and vulnerable data that could harm the user or the IoT network. IoT devices send and receive sensitive data, making them particularly vulnerable to this sort of attack [22].

1.3.3 Application Layer Attacks

To get access to the application layer and obtain sensitive data, software attacks are carried out. Some of the attacks used to target data in the application layer are listed below.

- Sniffing Attacks: Malicious software is used by attackers to monitor IoT network traffic. It gives an attacker the ability to intercept and read susceptible data over an IoT network. The vulnerable data transfer protocols in ioT devices and ecosystems leaves the network at the risk of such attacks [29].
- Code and Database Injection Attacks: Malicious code injection attacks can make IoT systems vulnerable. An attacker injects harmful code into susceptible entry points. Scripting is used to infiltrate malicious scripts or code into trustworthy websites and databases. A successful attack could result in the IoT account being compromised and the entire IoT network being damaged [27].
- Phishing Attacks: When a user is deceived into clicking messages, accessing online pages, or opening communication messages that appear to be from reputable sources, this attack happens. The user is usually duped into clicking on links that contain hazardous material, such as malware, or into entering sensitive information into input boxes. Both of these are then taken by the adversary [24].
- Theft of Data: IoT devices handle sensitive information, and much of it is transferred. This data is more vulnerable to cyber-attacks to data that is stored in a secure location. As users are aware that their IoT devices are vulnerable to attacks, they are also more cautious to enter private information on them [30].

1.4 Smart Devices in SMEs

Because any cyber incident can have a significant impact on a company's profitability, our hyper-connected world has become a dangerous place for businesses. According to a long-term survey conducted by the Business Continuity Institute, cyber-attacks and data breaches have been identified as the top two hazards for the past three years. Large organizations are more aware of those concerns, according to the report, with around 57% of respondents expressing concern. Only around 30 % of small and medium-sized businesses (SMEs) are concerned about being threatened [31]. According to several statistics, nearly half of all cyber-attacks are directed against small and medium-sized businesses (symantec, 2017). Small businesses can profit from the Internet of Things by automating simple processes like inventory management and procurement [32]. IoT is being used by over 60% of businesses, with only 9% having no IoT deployments at all (eclips, 2019).

Given the enormous potential for profit generation through new business models, SMEs are generally the most willing to adopt new technologies [33]. It is crucial to safeguard our SMEs, as they are considered to play an important part in the global economy. They employ two-thirds of the workforce in European countries and create

roughly 60% of total added value. Furthermore, SMEs are far more vulnerable than larger corporations. Regarding the financial loss point of view in data breaches, although the impact is less in absolute terms, it is significant when compared to their revenue, and it will also have a longer discovery period, implying a greater impact. Most SMEs will not get a second chance in the event of an incident: around 60% of businesses will close within six months of an attack [34].

Due to the increasing acceptance of digital technologies such as Cloud Computing and the Internet of Things (IoT), as well as the creation of a wide range of devices (e.g., PCs, servers, mobile devices, etc.) and business practices, cybersecurity in SMEs is becoming a significant concern (e.g., Bring Your Own Device, remote access, use of cloud-based apps and services, etc.)

Manufacturers of IoT devices are expected to ensure firmware integrity, traffic encryption, and adhere to strict software development rules; yet there are numerous unprotected devices on the market [35]. The users who directly utilize and/or manage these devices are household customers, estate managers, and network managers. These users should ideally follow a rigorous procurement process when acquiring and only install secure products that have strong encryption and are properly maintained. Users confront severe obstacles such as a lack of awareness, insufficient operational testing, a lack of automated asset management, and limited network monitoring abilities, all of which lead to the installation of vulnerable devices in their networks [36] Also, the average household consumer does not have the knowledge or resources to assess a device's security posture prior to purchase, and user manuals for such consumer gadgets do not include information regarding security features and/or threats [37].

Standards and guidelines for protecting IoT devices are being developed by government policymakers, regulatory authorities, and business partnerships. From prior offline models, guidelines for companies entering the IoT area aim to define meaningful security advice. These guidelines were aimed at IT experts and are hard to be implemented by normal users. There are numerous advocates of IoT security, and current methodologies are diverse; nonetheless, there is no obvious path to attaining proper quality. It is critical that security be considered throughout the device's life cycle, from creation to installation. Because the network is such an important element of the IoT ecosystem, we have focused on IoT network security in this paper.

1.5 Existing Guidelines for Secure Integration of IoT Devices

The aim of this research work is providing a guideline for average users to adopt the security practices for preventing cyber-attacks with consideration of IoT devices in the network. In order to compile a list of best practices we have examined the recommended security measures by the existing frameworks. We performed a study of the literature in both academic and government contexts.

The risks associated with IoT devices, systems, and services are numerous and constantly changing. As a result, it is critical to comprehend what must be safeguarded

and to devise appropriate security procedures to guard against cyber-attacks. While a formal set of standards on IoT security has yet to be developed, the industry initially recognized the need for guidelines in the early 2010s, and the debate has persisted since then and the debate has continued ever since. There are some IoT security and privacy standards of practice that provide recommendations and practices for both manufacturers and users to improve and safeguard security and privacy. DDoS attacks are mitigated by these suggestions. Only a few of standards are the focus of our investigation. The year 2017 has been chosen as the cut-off date for two reasons: Rapid IoT industry development may have rendered some notions obsolete, although recent papers frequently reference and align with previous ones in critical areas. Furthermore, rather than mere brochures or articles, industry standards established by reputable organizations or manufacturer groups were evaluated.

The Internet of Things Security Foundation's research, as well as the UK's Department for Digital, Culture, Media, and Sport's (DCMS) report titled "**Code of Practice for Consumer IoT Security**," published in October 2018, have classified critical details and best practices that must be followed and reviewed by the manufacturer, service provider, retailer, and consumer [38]. Not only does it help numerous IoT applications, such as wearable health trackers, smart homes, home automation, smart cameras, and alarm systems, comply with the General Data Protection Regulation (GDPR) but also it guides consumers to protect their privacy.

In 2020, the **National Institute of Standards and Technology** (NIST), which is part of the US Department of Commerce, released a paper titled "IoT Device Cybersecurity Capability Core Baseline" (NISTIR 8259A). The authors define an IoT device cybersecurity capability core baseline as a set of device capabilities that are commonly required to support common cybersecurity features that safeguard data, systems, and ecosystems. The proposed baseline is the result of a collaborative effort to create a list of common capabilities that is not detailed [39]. This document describes initiatives designed to improve the cybersecurity of manufactured products, reducing the number of attacked IoT devices as a result. However, the document does not offer guidance on how to implement these security guidelines.

ENISA has published several documents on secure IoT development. The document "Baseline Security Recommendations for IoT" was released in 2017 [40]. The goal of this project was to gain insight into the security needs of the Internet of Things, with a focus on Critical Information Infrastructure and services. The paper provides a comprehensive examination of current cybersecurity threats, as well as a thorough set of safeguards for IoT systems. Based on the findings of their research, the authors prepared a set of suggestions, expert opinions, best practices, and industry security measures. This publication also includes a comprehensive list of other IoT security standards, which can serve as a useful starting point for further analysis [41]. Standards from many sources were mapped by ENISA, for example, includes ISO standards, whereas the DCMS does not.

Another main source for securing IoT devices is the **European Telecommunications Standards Institute's** (ETSI) EN 303 645 standard on Cyber Security for Consumer Internet of Things. It includes citations to a number of major works, including the DCMS and ENISA documents mentioned above [42]. The document

mainly focuses on the european regulations and standards and attempts to provide a baseline in production level of IoT devices. For instance, in one of the projects, it is attempting to create a framework for managing cryptographic identities in a complicated environment of multi-function gadgets. Hence this is aimed at manufacturers and managers and professionals rather than consumers.

Because there are so many diverse parties involved in this security gap, including as manufacturers, suppliers, deployers, and network operators, the **US Department of Homeland Security** has developed a strategic approach for protecting IoT [43]. Integrating security into the design phase, advancing security updates and vulnerability management, building on proven security practices, prioritizing security measures based on potential impact, promoting transparency across IoT, and connecting properly and thoroughly are the core values to address IoT security challenges. For IoT developers, service providers, IoT manufacturers, and industrial and business level users, these proposed methods will help manage for security.

Specific industries appear to be developing their own standards in order to improve the safety and cyber security of their systems. For example, the latest TISAX: **Trusted Information Security Assessment exchange** standard for the automobile industry includes essential cyber security best practices standards (supply chain security [44]). This is especially critical for automobile IoT devices, since a hacked cyber security system could jeopardize their safety.

Following best practice recommendations also aids firms in demonstrating regulatory and legal compliance (e.g., **ISO27001** standard facilitates **GDPR** compliance). The EU Network and Information Systems Regulations (**NIS** Regulations), for example, provide legal steps to improve the overall degree of security (both cyber and physical stability) of network and information systems that are vital for the implementation of digital services (online markets, online search engines, and cloud computing services) and public infrastructure. This was implemented because the internet and private networks and information systems are becoming increasingly important enablers of our society and economies. As a result, it's critical to maintain a high level of network and information security across the board (NIS). It is worth mentioning that ISO is currently developing its own set of guidelines for secure deployment of IoT devices. **ISO/IEC CD 27400** "Cybersecurity—IoT security and privacy—Guidelines". As of August 2021, it is still under development [45].

However, the suggestions in aforementioned standards are too high-level and do not provide any system-specific information, obtaining in-depth system knowledge and interpreting those recommendations in the context of that knowledge is necessary for deriving effective security rules. Also, to verify those recommendations using formal tools, a significant amount of effort is required, including translating high-level recommendations into low-level security rules and preparing these rules for security verification (e.g., identifying data sources and converting them into formal languages). Eventually, user's lack of thorough understanding of IoT device technology, privacy and security, there are few effective mitigating activities made to preserve the security of the consumers. Also, regarding the Gap analysis of threats against standards, it is reported by Piasecki et al. [46] that there is no standard which addresses the issue of internal threats in the smart ecosystems.

The chosen security practices were driven from this document. However, many of the given measures are out of user's control and should have been implemented by the manufacturers. Also, since the focus of this paper is on prevention, the detection were not included. Also, the privacy risks associated with IoT device data sharing are outside the scope of this study.

1.6 Related Academic Works

Although all IoT platforms are concerned about securing IoT systems and users' privacy, there are certain distinctions in order to satisfy customer needs between consumer and enterprise IoT. Most academics and experts agree that in the case of consumer IoT, vulnerabilities affecting privacy pose a greater threat to the average user [47]. Because of the structure of personalized commodity systems, attackers may use such flaws to track user behavior or steal personal information. In order to decrease the possibility of an attack, preventative steps for minimizing the vulnerabilities should be taken.

According to Muhammad A. Iqbal et al., traditional security policies cannot be applied to IoT system architectures due to the various communication protocols and stacks. This is due to the IoT's heterogeneous structure, devices, and a scarcity of resources. To protect data security concerns, he recommends a solid network security infrastructure [48].

Shin and Kwon [49] have examined the existing authentication methods for Internet of Things devices and discovered that they are susceptible to attacks such as sensor node/end user impersonation, brute force assaults, Dos, and node capture attacks, replay attacks, and a slew of other threats. A lightweight key agreement protocol, based on the Internet Exchange Key server, has been suggested by Lavanya and Natarajan [50]. Despite the fact that this protocol provides end-to-end security between IPv6 and 6LoWPAN, it is only suitable for IP-based devices if the proper authentication mechanism is given to them. In 2017, Wu et al. [51] developed a mutual authentication protocol that is both privacy-preserving and lightweight; however, it does not handle the dynamic nature of IoT devices, and devices must be kept synced with the cloud server. An authentication technique that is both safe and efficient has been suggested by Srinivasa et al. [52] for multi-gateway wireless sensor networks. On the bright side, the proposed approach has no negative impact on the scalability and functioning of a wireless sensor network, nor on the functionality of the registration process for both end users and sensor nodes, as previously stated. On the disadvantage, nevertheless, it requires a lengthy and expensive computing procedure.

One of the most valuable work is a secure IoT architecture for smart cities based on Chakrabarty and Engels' black SDN proposal [53]. However, due to the restricted nature of IoT nodes, the proposed architecture does not provide a full SDN implementation, making IoT nodes subject to new sorts of risks and cyberattacks, such as node capturing, eavesdropping, and manipulation. The architecture also reduces

network efficiency and makes routing more difficult [54]. Also, due to the cost of implementing SDN/NFV based networks, it is not a clever idea to migrate SME network to SDN-based network.

In regard to edge computing, the most important consideration is the security of consumer's data. Their data are at risk of being compromised as soon as the user leaves the house. An adversary can check the electricity or water consumption and detects the user presence in the house. The user must monitor and protect network connections to ensure that such sensitive data is not accessed by outsiders [55].

One significant disadvantage of machine learning as a solution for IoT security is the difficulty in choosing the optimal machine learning algorithm for the purpose. Additionally, generating and locating the appropriate data collection is challenging. To ensure the success of machine learning algorithms, the appropriate data collection and algorithm must be utilised to build a functional IoT security solution. IoT devices may generate a massive quantity of data, the majority of which is redundant or worthless. This may cause complications when integrating machine learning, since researchers must manage with outliers and ambiguous data. Due to the limitations inherent with IoT devices, such as limited memory and processing resources, established methods such as unsupervised learning (recorded access logs) are irrelevant, since the logs cannot be stored locally on the IoT device [56].

Finally, the primary disadvantage of blockchain is the installation of the blockchain network. Due to the fact that all transactions are public, there is a possibility that private information will be exposed. Another issue with blockchain is the hardware required to run such systems. As the blockchain network grows in size and additional miners are added, storage space, prices, and the rate at which data is disseminated across the network may all rise. Thus, scalability is a critical issue to address while developing a blockchain system [57].

It is critical to take preventative measures in order to reduce susceptibility and the likelihood of an attack. Security devices must be built in a realistic, planned way in order to guarantee the safe and uninterrupted functioning of many protocols, communications, and services. As a result, the security of the entire system must be carefully considered and designed [58].

1.7 Research Approach

The goal of reviewing the entire ecosystem is to verify that all aspects of technology are safe (or, more realistically, secure within the constraints of a IoT device). Because of the linked nature of the Internet of Things, the security of any component in this ecosystem can and will have an impact on the security of all other components. For example, a flaw in access management can easily result in unauthorized access and control of embedded hardware, allowing a malicious insider to carry out attacks that could jeopardize the safety and security of the product user, bystanders, and the physical environment, or launch a distributed denial-of-service (DDoS) attack using compromised devices.

When detecting threats with network layer data in IoT contexts, mobility characteristics should be taken into account. However, for this scenario, as of small business it is based on a fixed network architecture. The fixed network architecture could handle the connectivity of computers to the network more efficiently than the IoT network through operational or technical controls.

We have built a small network to emulate an SME network including IoT devices. 4 windows 10 systems as well as 2 IoT devices (an IP Camera and a Raspberry Pi) have been implemented in the network. It has been assumed that IoT devices are fixed in the network and are connected to both local network and the internet. We have provided 2 scenarios by which we have demonstrated the implementation of security measures using a router and firewall.

1.7.1 Security Measures for IoT Implementation

In this section, we will go over a simplified model of an IoT device's lifecycle. We go over the primary and secondary stages that a gadget goes through during the course of its life. This enables comprehension of the link between an IoT device manufacturer and the end-users who will be influenced by the manufacturer's security-related design decisions. The lifecycle of an IoT device encompasses all of the major stages that a device can go through from invention through disposal—disposed-of, or otherwise never used again [59].

This is important to include in discussions about best practices, especially for IoT devices, which are designed to be long-lasting, mostly idle, and frequently connect our digital and physical worlds. Decisions made early in the lifespan might have an impact on later stages. While the stakeholders are mostly responsible for implementing the security measurements during the design and development of the IoT devices, there are some security mechanisms which the users should deploy in order to secure the device while installing, configuring, using and finally disposing or transferring the ownership of the device. Although the security challenges in the creation phase have direct impact to the installation, usage, and decommissioning, the focus of this paper is on the security practices that end-users can deploy to secure their systems, and data. Hence, the security practices in the creation phase are out of scope of this work.

In order to secure the IoT environment, all the security measurements should be applied to various levels to guarantee the safety of the data and resources [60]. However, it depends on the IoT device if it supports these security technologies. For instance, some of the devices have hardcoded credentials and it is not possible to change the default credentials. Hence, most of these security measurements are in control of manufacturers in development stage rather than users. The focus of this study is the security measurements which can be applied to the network and systems by the users to prevent the cyber-attacks.

To a large extent, best practices should address the requirements essential to guarantee security and privacy in the IoT network. Some of these behaviors are

entirely technical, as they are part of the device's build and design. Some requirements, such as vulnerability disclosure, are essentially organizational. We discovered a number of vulnerabilities with a simple evaluation utilizing standard, readily available techniques that might have been simply resolved by known best practices.

For service provision to be successful, it is critical to verify that only authorized data from genuine IoT devices is being utilized and processed. It has been suggested that a network-level security design be implemented since it is not feasible to apply standard protection techniques to smart devices in order to address the difficulties of regulating and managing in conventional networks and platforms. The issue with conventional networks is that they are incapable of dealing with the large and heterogeneous number of connected devices, as well as the massive amount of data that must be processed [61]. There has been a great deal of study done utilizing SDN in attempt to identify rogue IoT devices and attacks against them. This innovative network technology reduces the complexity of conventional networking from end users by separating the data plane from the control plane. It also enables heterogeneous networks with rapid development via the use of rules, which is a first in the networking industry. The combination of IoT with SDN has the potential to meet the expectations of those concerned with control and management [62]. With software-defined networking (SDN), the control plane from physical devices, as in conventional network design, is transferred to programmable controllers, and network intelligence is centralized. As a consequence, the network is presented to the application and policy engines as if it were a single switch, thanks to the global perspective of the network that it maintains Because SDN allows instructions from the SDN controllers rather than processing a large number of protocol standards, networking devices are made simpler as a result of SDN [63]. In centralized SDN-based networking the system faces the danger of single point of failure because of the centralized controller, however, it is simpler to maintain the network with a centralized controller. The absence of consistency in the decentralized system, on the other hand, is a major drawback. In this research we have tried to manage the network with a centralized SDN device to improve the security of network.

Authentication

Authentication is the first line of defense for majority of systems and networks. Before entering the network, devices should be authenticated in order to keep hostile devices out of the IoT ecosystem and prevent falsified data from spreading throughout the network [64]. There are many attacks such as MIDM which can be mitigated by applying authentication. In terms of IoT devices, as a first step the default credentials should be modified. Also, users should prevent reusing passwords across diverse types of devices. As suggested by NCSC, the password should be changed to a three- random word which the user can remember [65]. After resetting the password, double-check that the user's manual's default password no longer works. If it does, the credentials may have been hard-coded, in which case you should return the device or choose an alternative product. Also, make sure any hardware reset mechanisms

on the device are password-protected and not easily accessible, as these may restore the device's default manufacturer's passwords.

Blocking Unnecessary Ports and Services

Another matter to be considered in the context of network-based weaknesses is related to port blocking policies [66]. To this end, [67] investigated IoT connection via IPv4 and IPv6 and made many eye-opening discoveries. The authors observed that a substantial number of IoT sites are exclusively available through IPv6 and that certain IoT protocols are more accessible via IPv6 than via IPv4. The researchers discovered that in 46% of instances, the Telnet service's exposure was higher over IPv6 than over IPv4. Additionally, the authors contacted IoT network operators to validate their results and discovered that many default port openings are accidental, raising concerns about IoT security. Also, although network isolation can help prevent IoT threats, many IoT devices can connect and communicate with one another via alternative means, such as UPnP (Universal plug and play), Bluetooth, Zigbee, and other wireless and wired peer-to-peer (P2P) networks that are enabled by default [68]. Some IoT devices support remote management via a web interface as well as command-line methods such as Telnet or SSH. If possible, disable these services on the device (see the user's handbook), or use a router or firewall to block ports 22, 23, 2222, and 2323. If a device requires only internal access within the building, completely block it at the firewall.

Network segmentation

Network segmentation, in which you partition the network into subnets to restrict the flow of traffic across different zones in your business to prevent malware from travelling sideways, is one technique to proactively avoid network attacks [69]. IoT devices should be separated from the rest of your network. Separate network cabling and switches for vital equipment, or at the very least separate VLANs to contain threats to logical networks of comparable devices, may be required in enterprises. This may take the shape of a set of reserved IP or NAT addresses that you may limit and monitor from the management interface of your broadband router at home. The scalability of IoT networks is almost endless now that IPv6 has been implemented (2^{128} addresses possible). It's critical to implement network security through network segmentation and segregation in order to prevent direct Internet access to IoT devices [70].

Network Monitoring

The organization should keep track of the connected devices. This can be achieved by running networks scans to detect the available devices in the network. For the times

in which the smart device is not in use, consider whether equipment, such as desktop PCs, Wi-Fi routers, DVRs, and others, could be turned off during office hours, during vacations, and so on [71]. Connecting certain devices like smart TVs, smart speakers, and DVRs to a power strip with an on/off switch or connecting your primary router into an outlet with an automatic timer, may suffice. Remove the obsolete devices. However, the security measurements before disposal/transfer of ownership should be considered. All of the stored information on the device should be wiped as requested by legal obligations [37].

Packet Filtering

Several strategies for preventing DoS attacks that work by monitoring incoming traffic exist. These include involve filtering traffic from a given IP Address, restricting the number of packets that can be transmitted from a single IP Address, and forwarding and dumping any packets from specific IP Addresses without ever allowing them to reach their intended destination [72]. To make a filtering assessment, packet inspection can look at any or all of the elements such as source IP address, source port, destination IP address, destination port, IP protocol, and packet header information. In this scenario, a firewall examines all packets of a message that attempt to flow over the network and denies those that do not comply with the security policy [73].

Securing Layer 3 Devices

A router and firewall are important components of networks and the Internet because they determine the optimum way through the network to reach a destination and regulates data packet flow. Furthermore, when the administrator implements a good security policy, the router delivers good security solutions. Routers perform a variety of tasks, including traffic forwarding between two or more local networks, filtering, encrypting, relaying, and analyzing data streams. In fact, these functions have an impact on the confidentiality, integrity, and availability of information links in crucial network security mechanisms. It is recommended to disable all superfluous protocols and services in the router configuration in order to prevent an attacker from using it to harm the network and network devices configuration or steal sensitive information. Because routers are commonly used to handle network traffic and connect at least two separate networks, security is critical to prevent network intrusion and unauthorized router access [73]. To allow devices to locate other devices on your network, many routers use UPnP and port forwarding technologies. Unfortunately, cyber criminals can use these technologies to gain access to your network's devices, such as smart cameras. To avoid this risk, you should disable UPnP and port forwarding on your router; instructions can be found in the router's manual or on the manufacturer's website [65].

Device Management

IoT devices should be configured correctly and be updated as soon as the patch released from the company. Based on the worth of other assets in your network, you should check for updates more frequently. Look for updates on the manufacturer's website, and if patches are no longer available, consider replacing the device. Some gadgets download and install updates automatically. For those who don't, check with the manufacturer on a regular basis and apply updates as soon as they become available [71]. Prior to installation, the file should be updated and downloaded from a secure server, and the files should be signed and properly verified before being used [10]. When IoT devices are in use, make sure to protect every device on your network, particularly "legacy" laptops. In addition to the IoT-specific protections described above, install firewalls and antivirus/antimalware software on all servers, desktops, and laptops, and configure tablets and smartphones with appropriate security configuration (strong passcodes, two-step authentication, disable automatic Wi-Fi connections, be selective when installing apps).

It's possible that viruses are present if the IoT product starts to become slow or malfunctioning. Some virus is kept in memory and can be easily eliminated by restarting the device. Try a factory reset if the device is still slow or unresponsive after a reboot. Increased internet consumption or billing charges could signal that your device has been hijacked. this issue can be addressed by a factory reset and a password change [71].

Encryption

Data encryption, as the name implies, encrypts data so that it is secured even if the encrypted data packets are intercepted across a communication channel by an unauthorized user. This is a very effective method of protecting data and ensuring its security in the case of an attack [24].

If the device does not support encrypted channel for communication over wireless connection, it should be connected through wire. Consider purchasing a different device or disconnecting the device from the rest of your network if it does not enable encryption. The worth of all other assets on the same network as a possibly compromised device is the question, not the value of the item itself.

End-Point Detection and Response

Although anti-virus (AV) software is useful for protecting against malware, it is not the most efficient solution since signatures must be updated on a regular basis and anti-virus does not defend against zero-day attacks. The access management technique is another option to anti-virus software. Systems and apps that are allowed to access devices are permitted to do so; however, any system or application that is not authorized to access devices will be immediately banned [74].

Information Security Policies and Procedures

Choosing IoT vendors who follow best practices is perhaps just as crucial as applying the aforementioned best practices. The Manufacturer IoT Security Guidance project of the OWASP project provides security advice and considerations for IoT goods [10]. As part of a well-rounded security awareness program at work, frequent security training (every six months, if practicable) should include phishing awareness, installing software and OS updates, avoiding unauthorized devices including IoT, and using strong passwords or passphrases. Ensure that all employees are aware of the kind of devices in use and how they could be exploited by an attacker over the Internet [70].

Disposal/transfer of Ownership

During the disposal or transferring the ownership of the device all the personal information stored on it should be deleted. In most of the devices performing a factory reset erased the stored data on the device. Also, it should be disconnected from the associated accounts or paired devices. The users should be aware that the removeable media should be disposed separately from the device [6].

Back up

Data backups in the traditional sense prevent data loss. Offsite backups, often known as air-gapped backups, are stored offsite. This can be critical in preventing a loss of operational capacity after a disaster. Having an updated back up is the only solution the organization can have against ransomware attacks.

2 Experimental Work

In this chapter, we will illustrate how the security of Internet of Things devices may be tested using the testbed architecture and components. Our studies are carried out in two phases: demonstrating security flaws in smart devices using a home router and protecting the smart ecosystem using a firewall.

In the first scenario, we have utilized an SDN-based router to centrally manage the network and implement the security measurements. In the second scenario we have provided an access point as well as a firewall to segregate the network and provide a fine-grained security for the smart eco-system. In the second scenario we have deployed a firewall to secure the network. The chosen firewall—Cisco ASA5516x—has a built-in router, hence there is no need to provide a router. However, the firewall has only 8 ports and if the organization has extra devices, they might require utilizing a switch. The management system for configuration of the device

has been connected to the management port. The first phase is conducted to show how the security requirements can be met using basic network devices however in the second phase we provided a higher level of security using a firewall. We have provided 3 workstations, 1 back-up system which would connected to the IP camera to store the recordings, a Raspberry Pi 4 and smart phone. The main reason that we have chosen an IP camera is that it is reported that they have been the main attack surface for 47% of IoT target attacks.

2.1 Phase 1: Securing Network Using a Router

In this phase we have assumed that the organization has the new generation of routers which have built-in security controls. We attempted to demonstrate how users can take advantage of these security measurements to acquire security for installing IoT devices in the system. Figure 1 depicts the network map for the first scenario.

The default IP address for accessing the router is 192.168.1.1 on port GigaLan1. The management machine has been assigned IP address 192.168.1.11 with default gateway as the router. This can be done using the Ethernet network interface settings from control panel of the workstation (Fig. 2).

To access the GUI of the router, open any browser and enter https://192.168.1.1 and default credentials are admin/admin (Fig. 3).

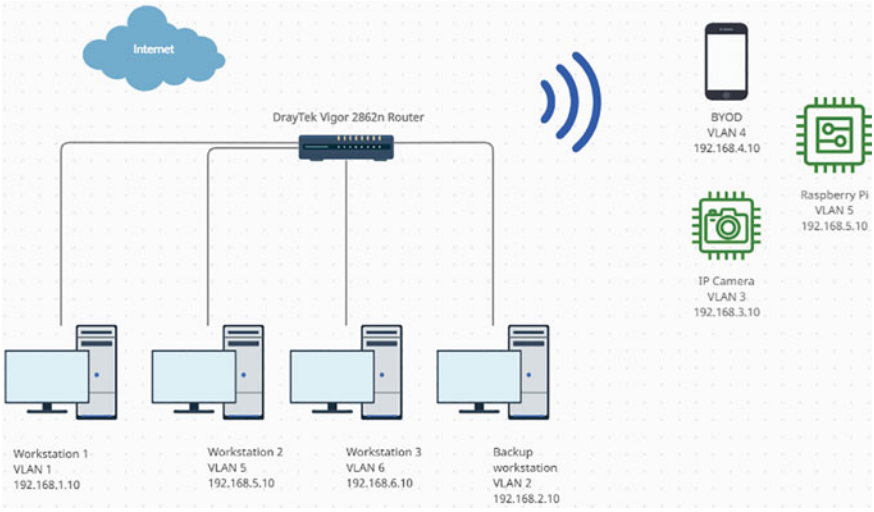


Fig. 1 First scenario network

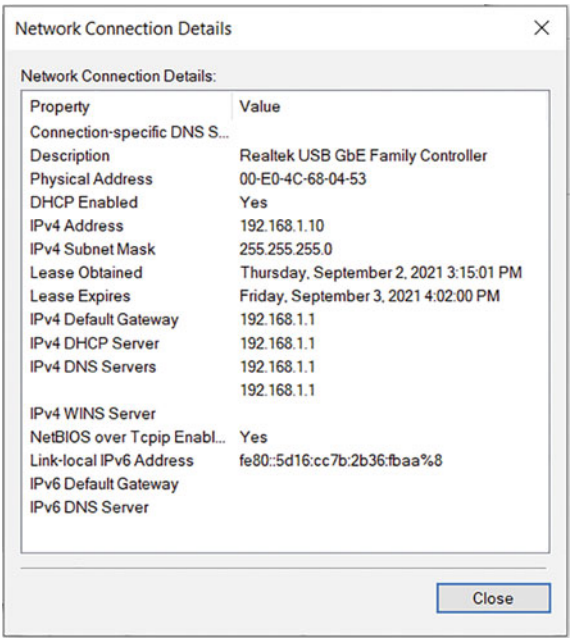


Fig. 2 Management workstation IP address

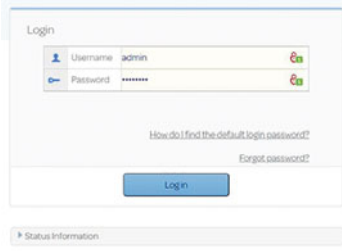


Fig. 3 Router default page

2.1.1 Hardening the Router/Layer 3 Device

1. The default credentials for the routers are either publicly available or are written on the device. They should be changed to secure the router device (Fig. 4).

2. The Router name and SSID can expose information about the device and internet provider and router; hence, it should be changed as well. By clicking on the system management tab, we have changed the router name and start hardening the device. All the unnecessary ports and services should be closed. We have disable ping from the internet to prevent the network scanning from the adversaries and minimize exposure of information regarding the network. For secure connection, we have only allowed the HTTPS and SSH connection and the

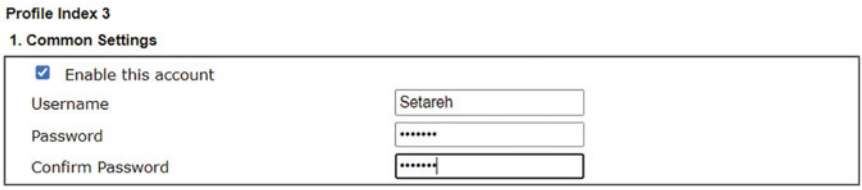


Fig. 4 Changing router default credentials

encryption has been set to TLS 1.1 and TLS 1.2. Also, the encryption for wireless devices have set to WPA2 (Fig. 5).

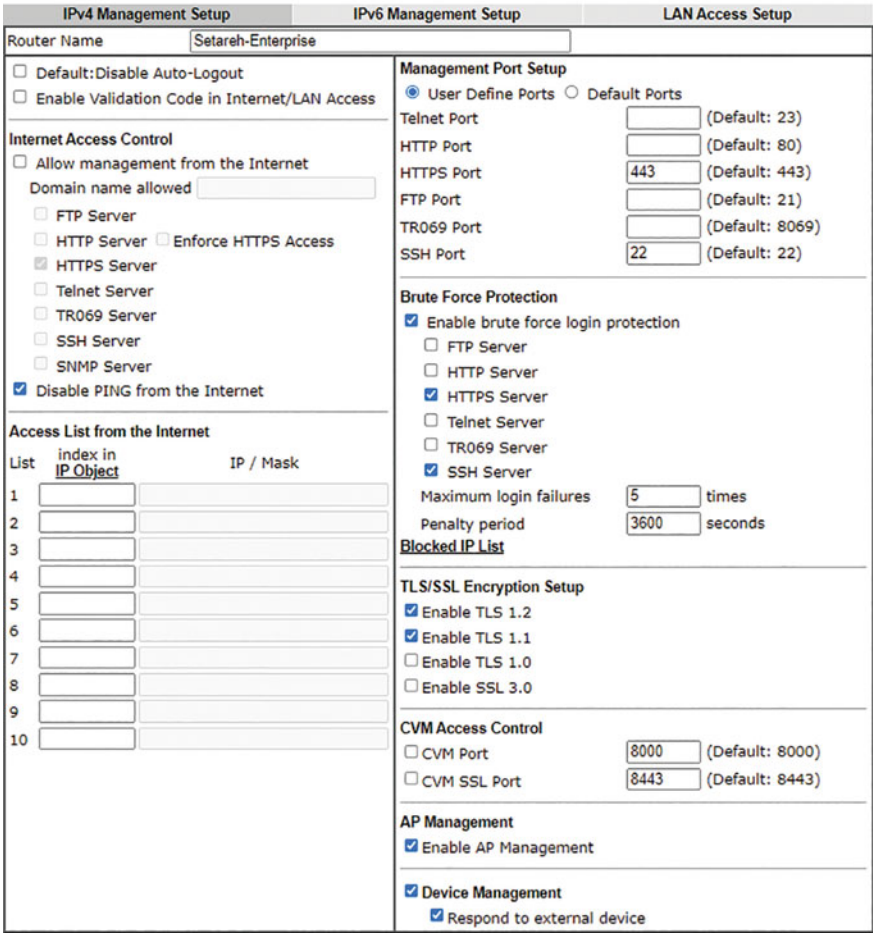


Fig. 5 Hardening router

2.1.2 Network Segmentation

The router can segregate the network to subnetworks. In order to isolate the devices, we have followed the following steps:

1. We have entered to the VLAN configuration and split them in to 5 different subnetworks (Main desktops, Backup desktop, Camera, Raspberry Pi, BYOD including mobile devices and guests) (Fig. 6).

2. For each subnetwork the IP address ranges have been established (Fig. 7).

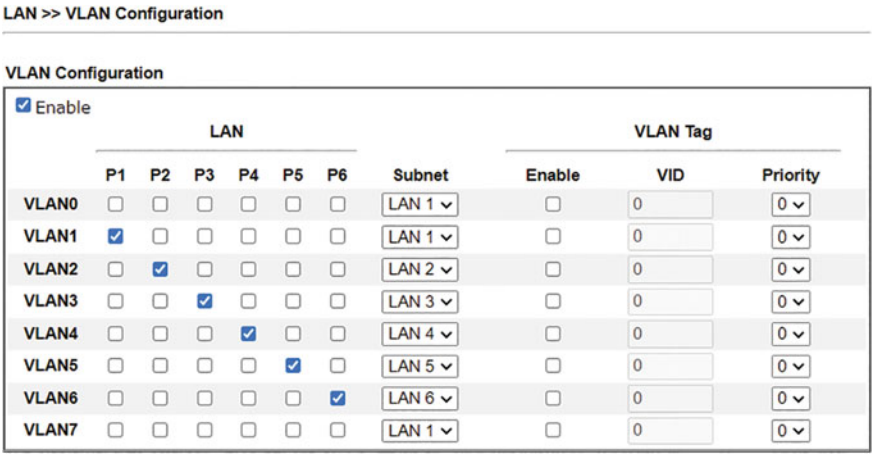


Fig. 6 Segregation of subnetworks

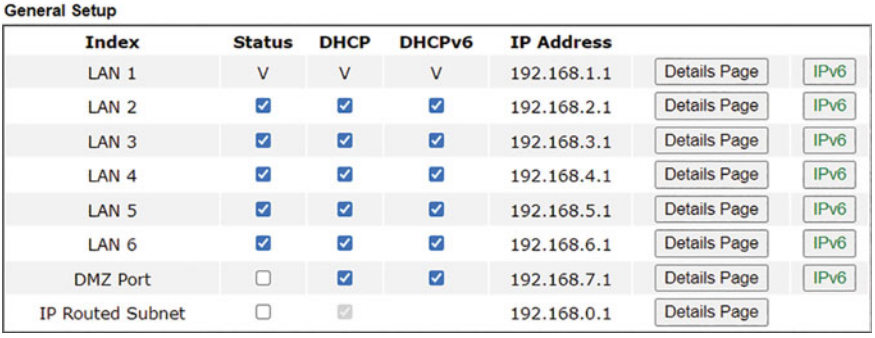


Fig. 7 Establishing IP ranges

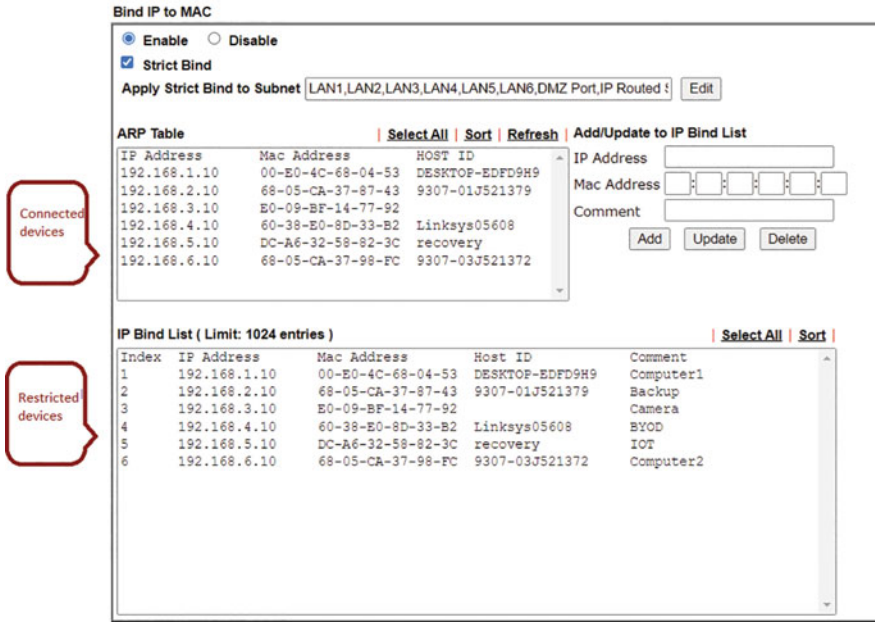


Fig. 8 Binding IP address to mac address

2.1.3 Monitoring the Devices

In the management tab, for allowed devices to connect we bind the IP addresses to the mac addresses and named each device in the comment field. By doing so, not only the organization can monitor the connected devices but also it allows the organization to create rules for the communication of each device (Fig. 8).

2.1.4 Access Management

As discussed in the previous section, the router is able to dedicate a rule to restricted devices for communication among them. Hence, only the specified LAN's are able to communicate to each other through a specific port. In order to store the recorded videos from the camera to back up desktop, we have allowed the communication from LAN2 to LAN3 (Fig. 9).

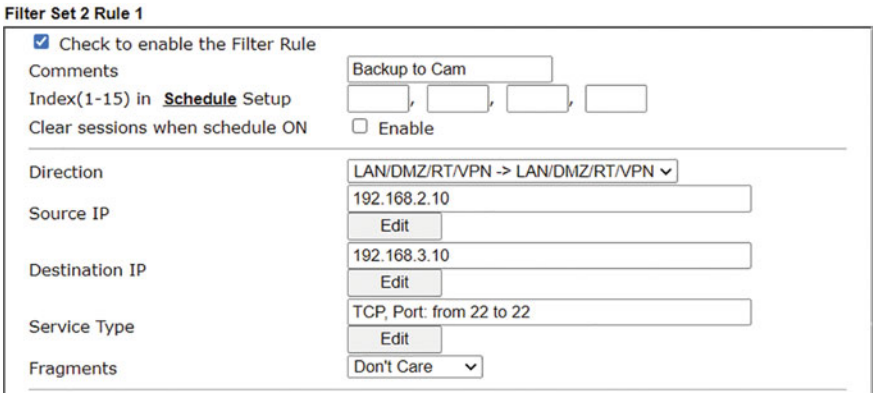


Fig. 9 Establishing a rule

2.2 *Phase 2: Using Extra Firewall*

3 Desktop machines and a raspberry Pie were connected to the firewall using wire. An access point was provided to connect mobile devices and all the other IoT devices which supports wireless communication (Fig. 10).

The second phase illustrates securing the IoT network using firewall in the following steps:

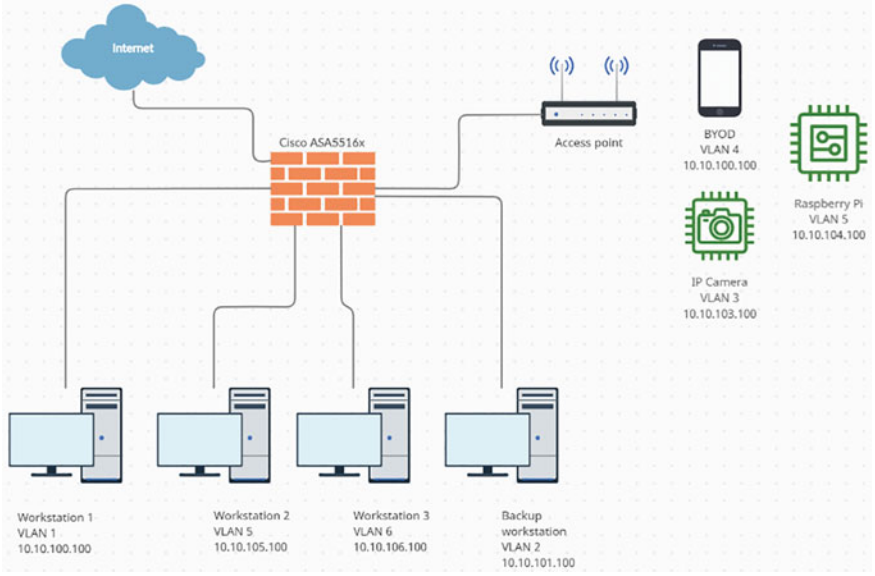


Fig. 10 Second scenario

```
ciscoasa> enable
Password:
ciscoasa#
```

Fig. 11 Enabling firewall

```
ciscoasa# configure terminal
ciscoasa(config)# hostname Setareh-Enterprise
Setareh-Enterprise(config)#
```

Fig. 12 Naming host

```
Setareh-Enterprise(config)# interface Management 1/1
Setareh-Enterprise(config-if)# ip address 192.168.45.45 255.255.255.0
Setareh-Enterprise(config-if)# nameif management
Setareh-Enterprise(config-if)# security-level 100
Setareh-Enterprise(config-if)# no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 13 Management interface

2.2.1 Hardening and Setting up the Firewall

In order to provide security for the desktop systems and IoT devices which can be connected using a router, we have connected them directly to a firewall. This is mainly because they can enjoy the security provided by the firewall to the maximum level. For configuring the router, we have used Putty software. The following steps illustrate the basic configuration of the firewall:

1. Initially to configure the firewall we enter the enable mode and then the configuration mode (Fig. 11).
2. Then we have renamed the firewall name (Fig. 12).
3. Then we have set up the management Interface as shown in Fig. 13:
   Table 1 describes each command:
4. We have set up a SSH channel for management Interface to connect to the host securely (Fig. 14):
5. Finally, in order to secure the SSH channel we have provided a secure password. It is worth mentioning that it is not possible to change the username for admin account on firewall (Fig. 15):

2.2.2 Segregation of Network

As shown in Fig. 2, we have provided 7 separate networks. Computer devices each have a separate zone to monitor the accessibility and packet filtering. For each of the

Table 1 Firewall command description

Command	Description
Interface	This command shows which socket is being used for the management host
Nameif	Name of Management security zone
Security-level	The Cisco ASA Firewall makes use of so-called "security levels," which show how trustworthy one interface is in comparison to another interface. It ranges from 100 to 0, the greater the degree of security, the more trustworthy the interface is considered to be. An interface with a high security level can communicate with an interface with a low security level, but the reverse is not feasible unless an access-list is configured to allow this kind of communication
No shutdown	To enable the interface

```
Setareh-Enterprise(config)# ssh 192.168.45.46 255.255.255.255 management
Setareh-Enterprise(config)#
```

Fig. 14 Setting up a SSH channel

```
Setareh-Enterprise(config)# username admin password 0$N+arbFPOhWn privilege 15
Setareh-Enterprise(config)#
```

Fig. 15 Setting up a username and password

IoT devices which can be connected through a wire, we have provided one separate network. The security level has been set to 0 to isolate the device. However, since the camera needs to communicate with one of the desktop systems to store its data and recordings, we have dedicated a separate desktop system and set the security level of the back-up system to 50 to permit the communication between the camera and the back-up system. This is mainly because, if the camera becomes compromised, we could contain the attack since the communication is one-way; hence, they are not fully connected. We placed the Raspberry Pi in the IoT subnetwork and connected an access point to the guest subnetwork for the mobile devices as well as all the wireless devices such as BYOD. The following shows the segregation of devices into seven separate LANs (Figs. 16, 17, 18, 19, 20, 21 and 22).

```
Setareh-Enterprise(config)# interface GigabitEthernet1/1
Setareh-Enterprise(config-if)#  nameif Computer1
Setareh-Enterprise(config-if)#  security-level 100
Setareh-Enterprise(config-if)#  ip address 10.10.100.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 16 Setting up LAN 1

```
Setareh-Enterprise(config)# interface GigabitEthernet1/2
Setareh-Enterprise(config-if)#  nameif Backup
Setareh-Enterprise(config-if)#  security-level 50
Setareh-Enterprise(config-if)#  ip address 10.10.101.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 17 Setting up LAN 2

```
Setareh-Enterprise(config)# interface GigabitEthernet1/3
Setareh-Enterprise(config-if)#  nameif Cameras
Setareh-Enterprise(config-if)#  security-level 0
Setareh-Enterprise(config-if)#  ip address 10.10.102.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 18 Setting up LAN 3

```
Setareh-Enterprise(config)# interface GigabitEthernet1/4
Setareh-Enterprise(config-if)#  nameif Guest
Setareh-Enterprise(config-if)#  security-level 0
Setareh-Enterprise(config-if)#  ip address 10.10.103.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 19 Setting up LAN 4

```
Setareh-Enterprise(config)# interface GigabitEthernet1/5
Setareh-Enterprise(config-if)#  nameif IOT
Setareh-Enterprise(config-if)#  security-level 0
Setareh-Enterprise(config-if)#  ip address 10.10.104.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 20 Setting up LAN 5

```
Setareh-Enterprise(config)# interface GigabitEthernet1/6
Setareh-Enterprise(config-if)#  nameif Computer2
Setareh-Enterprise(config-if)#  security-level 100
Setareh-Enterprise(config-if)#  ip address 10.10.105.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 21 Setting up LAN 6

```
Setareh-Enterprise(config)# interface GigabitEthernet1/7
Setareh-Enterprise(config-if)#  nameif Computer3
Setareh-Enterprise(config-if)#  security-level 100
Setareh-Enterprise(config-if)#  ip address 10.10.106.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
Setareh-Enterprise(config)#
```

Fig. 22 Setting up LAN 7

```
Setareh-Enterprise# show interface ip brief
Interface                  IP-Address      OK? Method Status                Protocol
GigabitEthernet1/1         10.10.100.1     YES manual up                    up
GigabitEthernet1/2         10.10.101.1     YES manual up                    up
GigabitEthernet1/3         10.10.102.1     YES manual up                    up
GigabitEthernet1/4         10.10.103.1     YES manual up                    up
GigabitEthernet1/5         10.10.104.1     YES manual up                    up
GigabitEthernet1/6         10.10.105.1     YES manual up                    up
GigabitEthernet1/7         10.10.106.1     YES manual up                    up
GigabitEthernet1/8         unassigned      YES unset  administratively down down
```

Fig. 23 Interface status

```
Setareh-Enterprise# show arp
        Computer1 10.10.100.100 00:E0:4C:68:04:53
        Backup 10.10.101.100 68:05:CA:37:87:43
        Cameras 10.10.102.100 E0:09:EF:14:77:92
        Guest 10.10.103.100 60:3B:E0:8D:33:B2
        IOT 10.10.104.100 DC:A6:32:58:82:3C
        Computer2 10.10.105.100 EC:8E:B5:3E:7E:93
        Computer3 10.10.106.100 68:05:CA:2F:EA:80
```

Fig. 24 Monitoring devices in Firewall

Figure 23 illustrates the available LANs in the firewall. As it can be seen, desktop systems have the ability to communicate with all of the systems, while the camera, Raspberry Pi, and BYOD are isolated to their own subnetwork.

2.2.3 Monitoring the Devices

Not only can the organization monitor the subnetworks within their network, but also it is feasible to monitor the available devices on the network using the "show arp" command (Fig. 24).

2.2.4 Access Management

As mentioned earlier, we scored the security level of the workstations to 100, the back-up system to 50 and the rest to 0. Based on the firewall default setting, devices/LANs

```
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit tcp interface Computer1 interface Computer2
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit tcp interface Computer2 interface Computer1
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit tcp interface Computer1 interface Computer3
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit tcp interface Computer3 interface Computer1
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit tcp interface Computer2 interface Computer3
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit tcp interface Computer3 interface Computer2
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit udp interface Computer1 interface Computer2
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit udp interface Computer2 interface Computer1
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit udp interface Computer1 interface Computer3
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit udp interface Computer3 interface Computer1
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit udp interface Computer2 interface Computer3
Setareh-Enterprise(config)# access-list COMPUTER_to_COMPUTER extended permit udp interface Computer3 interface Computer2
Setareh-Enterprise(config)# access-group COMPUTER_to_COMPUTER global
```

Fig. 25 Default access management list

with higher security scores are able to interact with the lower ones by default. However, the ones with the lowest or even the same level are isolated. Hence, we only needed to permit access among the workstations. Figure 25 shows the current level of access for each device.

2.2.5 Threat Detection

The ASA's threat detection system serves as a first line of protection against cyber-attacks. In order to establish a baseline for traffic on the device, threat detection operates at Layers 3 and 4 by monitoring packet drop data and compiling "top" reports based on traffic patterns. In contrast, a module that offers intrusion prevention or next generation intrusion prevention services detects and mitigates attack vectors up to Layer 7 on traffic that has been allowed by the ASA but does not view the traffic that has already been discarded by the ASA. As a result, threat detection and intrusion prevention systems (IPS) may collaborate to offer a more complete threat defense. Figure 26 indicates the basic threat detection in ASA firewalls.

```
Setareh-Enterprise(config)# threat-detection rate dos-drop rate-interval 600 average-rate 100 burst-rate 400
Setareh-Enterprise(config)# threat-detection rate dos-drop rate-interval 3600 average-rate 80 burst-rate 320
Setareh-Enterprise(config)# threat-detection rate bad-packet-drop rate-interval 600 average-rate 100 burst-rate 400
Setareh-Enterprise(config)# threat-detection rate bad-packet-drop rate-interval 3600 average-rate 80 burst-rate 320
Setareh-Enterprise(config)# threat-detection rate acl-drop rate-interval 600 average-rate 400 burst-rate 800
Setareh-Enterprise(config)# threat-detection rate acl-drop rate-interval 3600 average-rate 320 burst-rate 640
Setareh-Enterprise(config)# threat-detection rate conn-limit-drop rate-interval 600 average-rate 100 burst-rate 400
Setareh-Enterprise(config)# threat-detection rate conn-limit-drop rate-interval 3600 average-rate 80 burst-rate 320
Setareh-Enterprise(config)# threat-detection rate icmp-drop rate-interval 600 average-rate 100 burst-rate 400
Setareh-Enterprise(config)# threat-detection rate icmp-drop rate-interval 3600 average-rate 80 burst-rate 320
Setareh-Enterprise(config)# threat-detection rate scanning-threat rate-interval 600 average-rate 5 burst-rate 10
Setareh-Enterprise(config)# threat-detection rate scanning-threat rate-interval 3600 average-rate 4 burst-rate 8
Setareh-Enterprise(config)# threat-detection rate syn-attack rate-interval 600 average-rate 100 burst-rate 200
Setareh-Enterprise(config)# threat-detection rate syn-attack rate-interval 3600 average-rate 80 burst-rate 160
Setareh-Enterprise(config)# threat-detection rate fw-drop rate-interval 600 average-rate 400 burst-rate 1600
Setareh-Enterprise(config)# threat-detection rate fw-drop rate-interval 3600 average-rate 320 burst-rate 1280
Setareh-Enterprise(config)# threat-detection rate inspect-drop rate-interval 600 average-rate 400 burst-rate 1600
Setareh-Enterprise(config)# threat-detection rate inspect-drop rate-interval 3600 average-rate 320 burst-rate 1280
Setareh-Enterprise(config)# threat-detection rate interface-drop rate-interval 600 average-rate 2000 burst-rate 8000
Setareh-Enterprise(config)# threat-detection rate interface-drop rate-interval 3600 average-rate 1600 burst-rate 6400
Setareh-Enterprise(config)# threat-detection basic-threat
Setareh-Enterprise(config)# threat-detection scanning-threat shun duration 3600
Setareh-Enterprise(config)# threat-detection statistics access-list
```

Fig. 26 Basic threat detection

```
Setareh-Enterprise(config)# class-map inspection_default
Setareh-Enterprise(config-cmap)#  match default-inspection-traffic
Setareh-Enterprise(config-cmap)# exit
Setareh-Enterprise(config)# policy-map type inspect dns preset_dns_map
Setareh-Enterprise(config-pmap)#  parameters
Setareh-Enterprise(config-pmap-p)#     message-length maximum client auto
Setareh-Enterprise(config-pmap-p)#     message-length maximum 512
Setareh-Enterprise(config-pmap-p)#     dns-guard
Setareh-Enterprise(config-pmap-p)#     protocol-enforcement
Setareh-Enterprise(config-pmap-p)#     nat-rewrite
Setareh-Enterprise(config-pmap-p)# exit
Setareh-Enterprise(config-pmap)# exit
Setareh-Enterprise(config)# policy-map global_policy
Setareh-Enterprise(config-pmap)#  class inspection_default
Setareh-Enterprise(config-pmap-c)#   inspect dns preset_dns_map
Setareh-Enterprise(config-pmap-c)# exit
Setareh-Enterprise(config-pmap)# exit
Setareh-Enterprise(config)# service-policy global_policy global
```

Fig. 27 Enabling default service policy

2.2.6 Service Policy

When configured using the Modular Policy Framework, service policies offer a consistent and flexible method to define ASA functionalities. A policy is included by default in the configuration, which matches all default application inspection traffic and applies specific inspections to all traffic on all interfaces, regardless of which interface is being used (a global policy). Figure 27 illustrates the configuration of the default policy:

2.2.7 VPN

Employees are increasingly requesting the ability to work from anywhere using corporate computers as well as personal mobile devices. The Cisco AnyConnect Secure Mobility Client allows you to enable your workers to accomplish this while also providing the essential security to assist in guaranteeing that your company is secure and protected. A unified security endpoint agent, such as Cisco AnyConnect, protects the business by providing various security services. It also offers the visibility and control you need to determine who is accessing the extended business and which devices they are using to do so. In addition to remote access and posture enforcement, Cisco AnyConnect's extensive security services portfolio includes capabilities such as online security features and roaming protection. It provides the security features required to offer a strong, user-friendly, and highly secure mobile experience [75]. We have provided a dedicated zone (outside) for remote access and enabled it for users to reach the systems remotely. We have assumed that the user needs to connect to the backup system, hence the VPN is created for remote access to that device. Table 2 describes each command that was used for creating the VPN channel:

1. Creating the "outside" zone for remote Access (Fig. 28)

Table 2 OWASP Top 10 security risk for IoT devices and mitigation strategies

OWASP Top 10 security risk for IoT devices	Mitigation
Weak or guessable password	Modification of all default credentials for each of the network component and IoT devices
Insecure network security	Blocking insecure ports and services as well as network segmentation
Insecure ecosystem interface	Blocking insecure ports for data communication among applications, changing default credentials for interfaces
Lack of secure update mechanism	Mostly out of user control, consumers can only check for the updates on regular basis or special circumstances
Use of insecure or outdated components	Conducting research prior to purchasing the device, if the device cannot be removed from the network it should be isolated in a separate subnetwork
Insufficient privacy protection	Prevent to provide sensitive information or store them on the device
Insecure data transfer and storage	Unnecessary ports should be blocked. Communication should be allowed through secure channels and prevent to store data on the device
Insecure default setting	customize the default setting for both device and all other network components to the best level suits the organization security requirements
Lack of physical hardening	In order to prevent the tampering attacks, IoT devices should be kept in secure places with physical security. Also

```
Setareh-Enterprise(config)# interface GigabitEthernet1/8
Setareh-Enterprise(config-if)#  nameif Outside
Setareh-Enterprise(config-if)#  security-level 0
Setareh-Enterprise(config-if)#  ip address 10.10.200.1 255.255.255.0
Setareh-Enterprise(config-if)#  no shutdown
Setareh-Enterprise(config-if)# exit
```

Fig. 28 Setting up an outside interface

2. AnyConnect access on the ASA's "Outside" interface should be enabled (Fig. 29).

```
Setareh-Enterprise(config)# webvpn
Setareh-Enterprise(config-webvpn)# enable Outside
INFO: WebVPN and DTLS are enabled on 'Outside'.
```

Fig. 29 Enabling the outside interface

```
Setareh-Enterprise(config)# ip local pool SSLClientPool 192.168.100.1-192.168.100.50 mask 255.255.255.0
```

Fig. 30 Creating local address IP pool

```
Setareh-Enterprise(config)# object network BACKUP-HOSTS
Setareh-Enterprise(config-network-object)# host 10.10.101.100
Setareh-Enterprise(config-network-object)# exit
Setareh-Enterprise(config)# object network VPN-HOSTS
Setareh-Enterprise(config-network-object)#  subnet 192.168.100.0 255.255.255.0
```

Fig. 31 Exempting NAT for 8.3 version and later

```
Setareh-Enterprise(config)# nat (Backup,Outside) source static BACKUP-HOSTS BACKUP-HOSTS destination static VPN-HOSTS VPN-HOSTS
```

Fig. 32 Specifying source of IP

```
Setareh-Enterprise(config)# username userA password N+arbFPOhWn
Setareh-Enterprise(config)# username userA attributes
Setareh-Enterprise(config-username)# service-type remote-access
```

Fig. 33 Creating an account for remote users

3. Creating a local IP address pool to be used for assigning IP addresses to distant users (Fig. 30).
4. Exemption from NAT should be configured for traffic between internal LAN and distant users.
   (a) If you are using ASA version 8.3 and later (Fig. 31)
   (b) Then the source of IP address for host and destination should be defined (Fig. 32)
5. Create usernames that will only be used for remote access via AnyConnect (Fig. 33).
6. Finally, we have created a tunnel group profile to specify the parameters of the connection (Fig. 34).

```
Setareh-Enterprise(config)#  group-policy SSLClientPolicy internal
Setareh-Enterprise(config)# group-policy SSLClientPolicy attributes
Setareh-Enterprise(config-group-policy)# address-pools value SSLClientPool
Setareh-Enterprise(config-group-policy)# webvpn
Setareh-Enterprise(config-group-webvpn)# vpn-tunnel-protocol svc
Setareh-Enterprise(config-group-policy)# sysopt connection permit-vpn
Setareh-Enterprise(config)#
Setareh-Enterprise(config)# tunnel-group SSLClientProfile type remote-access
Setareh-Enterprise(config)# tunnel-group SSLClientProfile general-attributes
Setareh-Enterprise(config-tunnel-general)# default-group-policy SSLClientPolicy
Setareh-Enterprise(config-tunnel-general)# tunnel-group SSLClientProfile webvpn-attributes
Setareh-Enterprise(config-tunnel-webvpn)# group-alias SSLVPNClient enable
Setareh-Enterprise(config-tunnel-webvpn)# webvpn
Setareh-Enterprise(config-webvpn)# tunnel-group-list enable
```

Fig. 34 Creating a group policy

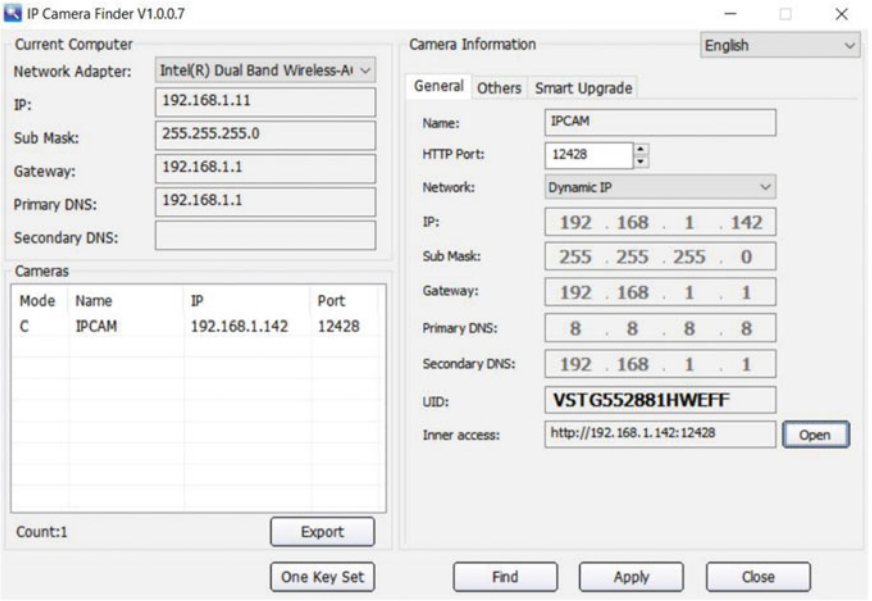


Figure 35- Finding camera's IP address and open port

Fig. 35 Finding camera's IP address and open port

2.3 Shared Phase: Securing IoT Device

Many of the security measurements suggested by current guidelines are out of users' control. They should have been implemented in the development stage. However, there are two primary actions which should be taken by consumers for every IoT device they own. The majority of IoT devices are accessible through web browsers by using open ports and default credentials. As a very first step, the default username and password for these devices should be modified. We showed in earlier sections how users may determine the IP address of an IoT device. We installed a tool called "IP camera finder" to determine the open port via which the smart device—the camera—communicates with the internet. Not only does it display the camera's IP address, but it also displays the available port for communication (Fig. 35).

We have accessed the camera using the web browser and default credentials (Figs. 36 and 37).

2.3.1 Changing Default Credentials

Since the default credentials are publicly available, it is essential to update them immediately after configuring the IoT device. This is accomplished by pressing the setup button and modifying the password (Fig. 38).

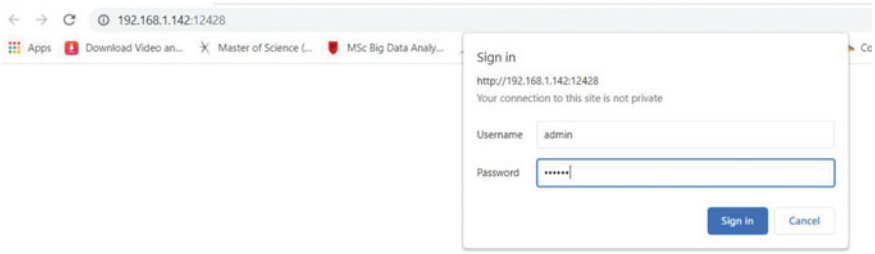


Fig. 36 Camera default page

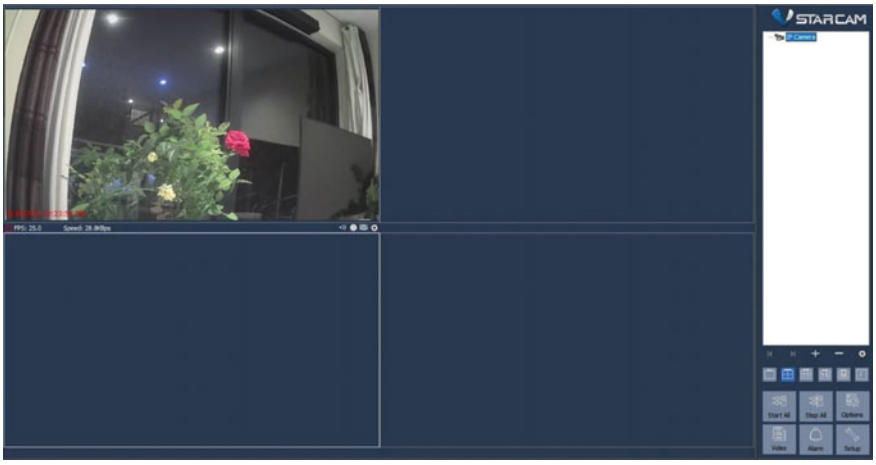


Fig. 37 Logging into the camera

Regarding the Raspberry Pi, after powering up the device, we were prompted to alter the default password (Fig. 39).

2.3.2 Updating Software

Deployed software for IoT devices is never bug-free. These bugs are being patched by the company, but it is the responsibility of the user to apply the latest patches to the device. We sought updated versions by clicking the setup and update buttons, respectively (Fig. 40).

Regarding Raspberry Pi we opened the terminal and updated the software using "Sudo apt update && upgrade" command (Fig. 41).

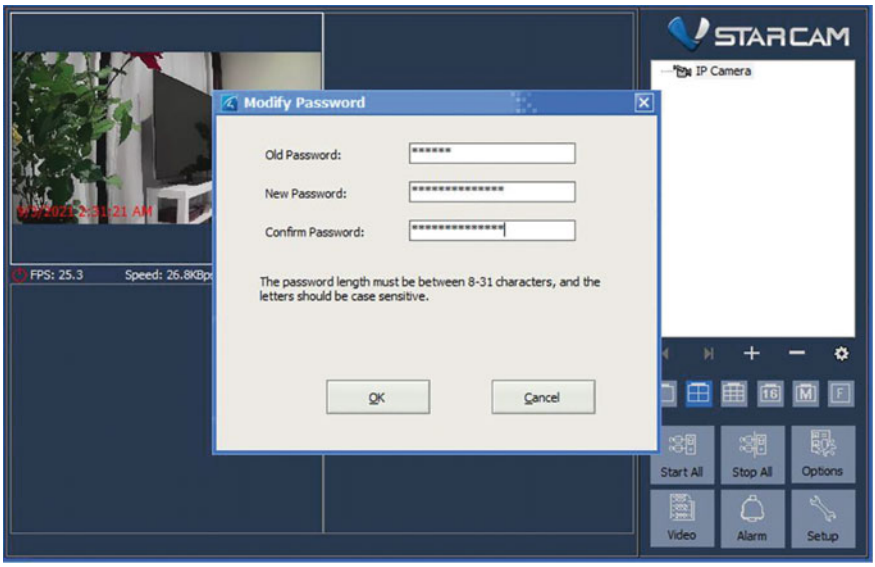


Fig. 38 Changing IP camera default password

Fig. 39 Changing Raspberry Pi default password

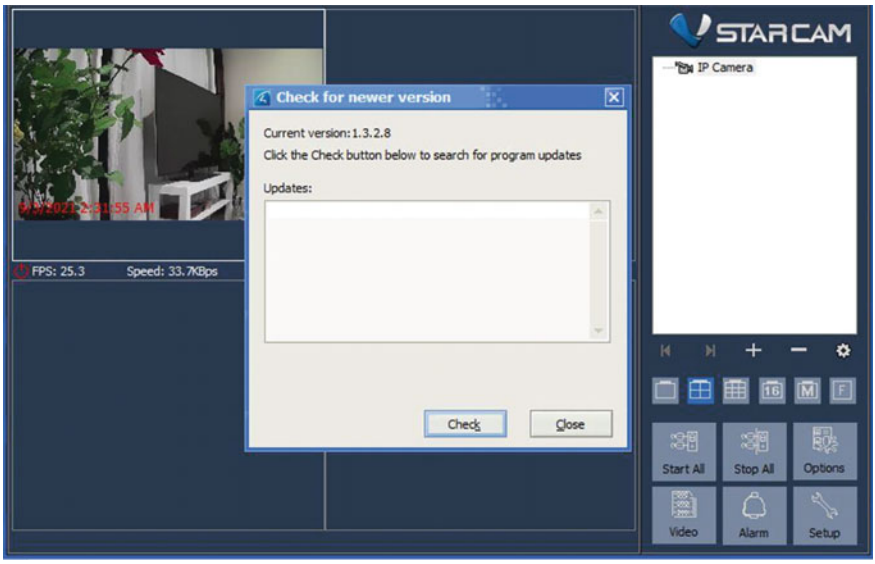


Fig. 40 Updating camera

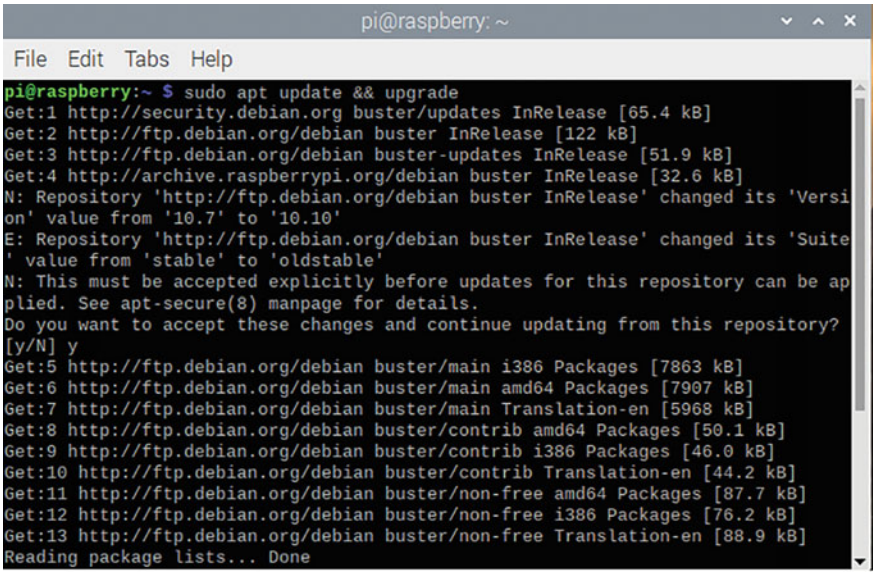


Fig. 41 Updating Raspberry Pi

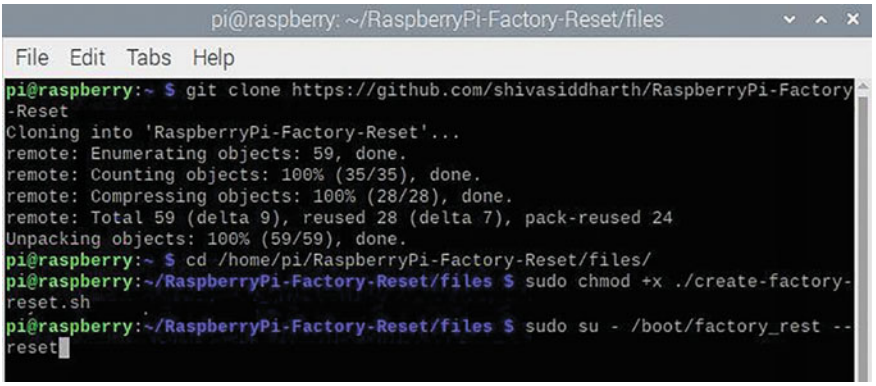


Fig. 42 Resetting raspberry Pi

2.3.3 Secure Disposal

Before discarding of any device, it is recommended that all stored data be deleted. To do so, the user needs to perform a factory reset. This may be accomplished with the IP camera by inserting a pin into the device until it restarts. To reset the Raspberry Pi, users must remove the SD Card, format it, rewrite the OS image, then re-plug it [76]. We have downloaded the image file from Github and re-write the OS image using a fresh OS. After entering the last command, the device will boot from the SD card with a new OS (Fig. 42).

3 Discussion

The goal of this study was to investigate how owners of smart devices can detect and protect IoT devices against cyber-attacks by securing their network. The majority of owners are completely unaware of the cyber dangers associated with IoT and have failed to educate themselves on the many security solutions available to avoid or minimize such risks. Although many recommendations exist for protecting the IoT environment, they lack a guideline for implementing such security measures.

To ensure the safety of sensitive device data, it is critical to address security measures in each domain in a proactive manner. Vulnerabilities in IoT devices are mitigated by implementing security measures properly to eliminate threats. The lack of security controls on devices is well-known. The numerous botnet attacks throughout the years have demonstrated this. For instance, the latest form, called "Hide and Seek," can survive a device reboot. We must keep ahead of the curve by adopting and implementing security policies to reduce possible damage as IoT attacks materialize and grow in complexity [77].

Performance Evaluation: To measure the effectiveness of the IoT-based security best practices demonstrated from the experiment, the "OWASP Top 10 security risk for IoT devices" is utilized. Each risk from this framework is evaluated to verify if the proposed mitigation techniques can reduce the risk level. Table 2 provides a summary of information on the mitigation techniques utilized. Many of these risks, however, may be addressed only by manufacturers, and the mitigation techniques cannot reduce the risk to an acceptable level. If the IoT device does not support security technologies, the customer is limited in their options. Hence, consumers must conduct research before purchasing IoT devices and be aware of the security limitations of the devices.

In order to reduce the risk, it is feasible to segment the network into different subnetworks and use various levels of protection, such as firewalls. For both scenarios, we have implemented network segmentation. However, since the IoT device must verify their identity before accessing resources and communicating, many IoT devices fail in this approach [78]. As soon as the connection is established by any of the other devices on the network, they are at risk. To mitigate this security risk, we have dedicated an isolated back-up system which is connected to IoT devices to store and process the smart device's data. The connection to the systems is only allowed through secure ports and an access management technique has been provided to restrict the access.

For devices that are not updated on a regular basis—such as medical devices and laboratory research equipment—the use of segregation strategies, such as VLANs, to isolate IoT devices has been recommended. However, for printers that are accessed by the general public, it is more appropriate to use firewalls as internal network boundaries and intrusion detection systems to control the data. Due to the router's restrictions, the first scenario was unable to identify and prevent threats; furthermore, the rules were limited to LANs rather than devices. However, in the second approach, after installing the firewall, it is feasible to granularly protect the computers and network. Not only is it possible to segment the network into subnetworks, but access control and data filtering rules may also be applied to individual devices rather than the whole LAN. Additionally, firewalls may be used to solve the issue of remote access via the usage of VPNs. The consistent implementation of this solution, however, may be difficult due to the possibility that users may mix IoT software on their work device with IoT software on their personal device, or that they will bring their own IoT device into the network. Insecure software application programming interfaces (APIs) and hardware interfaces, as well as a lack of integrity checks on the software images being loaded into the device, are some of the causes of malicious code injection attacks against IoT devices. Application developers and other users may employ insecure APIs to connect and communicate with IoT devices, raising the risk of harmful code injection attacks from unauthorized actors. To protect against such attacks, well-known best practices for securing API endpoints, such as input validation and IP address filtering, can be used. However, these security measures are in the first stage of the IoT device life cycle and can be implemented by manufacturers rather than users. Consumers should do research on security before purchasing

a device. Even ostensibly advanced devices sometimes fail to meet the lowest security standards specified by common checklists, indicating that more research and regulation on subtle issues in IoT security is needed.

On the other hand, manufacturers expose insufficient information about their devices' security characteristics to the public, which complicates market monitoring and gives customers little knowledge about the security of items prior to purchase. Consumers are seldom advised about cyber safety. Finally, we discovered a deficiency in the standardization of security-related information for IoT devices.

Authentication has traditionally been considered the first line of defense for the vast majority of IoT devices. It remains susceptible and cannot be regarded as a comprehensive solution for IoT security countermeasures. Almost all of the authentication techniques proposed up to this point make use of public keys, and since a public key may be stolen, they still expose the system to the danger of unauthorized access to information. Eavesdropping attacks are possible because some IoT devices lack an encrypted communication connection between the backend server and the device. As a result, the attacker can impersonate a genuine organization and gain access to unencrypted data delivered across its home network, potentially exposing personal network information. Traditional data encryption techniques meant for network security are typically too heavy to be implemented on IoT devices with limited memory. This involves the development of energy and computation constrained IoT devices with robust, lightweight encryption algorithms. Even for devices that don't require encryption, encrypting communication ensures the privacy of other entities in the IoT network. Low-cost demilitarized zones, which operate as buffer zones to prevent attackers from accessing data on personal home networks, can be set up at the consumer's end. Not only did we provide DMZ in both scenarios, but we also allowed communication inside the network only.

On the other hand, support personnel must be able to remotely diagnose operating system issues and apply critical upgrades, so many IoT devices include embedded credentials. Hackers may take use of this feature to penetrate a device's defenses. It is critical that users have the ability to alter their default credentials right from the start. However, even while employing complicated passwords would reduce the risk to a considerable degree, this form of authentication alone is susceptible to a variety of assaults, including dictionary attacks, which are particularly dangerous. Furthermore, due to the large number of users that use the same credentials for many accounts, as well as the simplicity with which this vulnerability in systems may be exploited, the danger of unauthorized access is very high in the event that one account is hacked.

In addition to the technical aspect of securing the device, organizations should consider personnel training as well. Proper training and awareness are the most crucial aspects of using such devices. All individuals who administer or utilize these devices should be adequately committed to this purpose. They must receive sufficient training for the code of conduct and usage from an experienced organizational professional.

Last but not least, in the case of being a victim of cyber-attacks, the most promising solution is to have a series of updated backups. Whether it's a malicious update from third parties or a ransomware attack, the only solution is an updated backup.

4 Conclusion

IoT has been in the spotlight for the past decade because it has simplified practically every aspect of life, including smart homes, manufacturing, medical, and transportation. These innovations have made life easier because the IoT can be accessible from anywhere in the world via smartphone or the internet. These advancements drew the attention of attackers, and several attacks are launched every day. Although there are numerous guidelines regarding securing IoT devices and networks, none of them has demonstrated how these security measurements can be applied. They were designed for IT admins with technical knowledge. However, many small and medium enterprises do not have this technical knowledge, and due to financial restrictions, they cannot afford to hire a security officer. In this study, we have discussed the three-layer architecture of smart devices and the various attack types at each layer. We have reviewed the existing guidelines and their security requirements. Then we selected only prevention techniques and applied them to a real-life network and devices to secure them from security terrorization. The implementation of the selected techniques has been demonstrated in two different networks to educate the normal users of IoT devices on how they can apply these security measurements without the help of IT experts in their network. Moreover, the security measures for IoT devices themselves have been discussed.

The IoT is a large integration of numerous layers, and the security concerns we found were not limited to a single layer, but rather affected the entire infrastructure and multiple layers. Instead of focusing primarily on individual scenarios and targeting solutions specifically for them, we should deploy numerous technologies in combination and study the specific conditions in order to find and solve these challenges. We have primarily focused on network segmentation to monitor and isolate smart devices. Using existing and available best practises, based on our findings, would be a huge improvement. However, implementing most of the best practices relies on the development stage and supported technologies by the device. Hence, consumers should conduct research on the security aspects of the devices prior to purchasing them. Unsupported security techniques, lack of regular patches, and lack of regulation in the development stage of the IoT life cycle are the biggest obstacles to securing IoT devices, which should be considered.

5 Limitations and Future Works

This study consisted of setting two different test beds in order to demonstrate the installation of recommended security measures. While in the first scenario, the SDN-based router is capable of providing only 8 VLANs, in the second scenario, there is no limit for providing various LANs as it is possible to add various switches to network maps and monitor all the devices. In the second scenario, we have three distinct zones of workstations since it was not feasible to connect them all to the

same LAN without using a switch. It is not feasible to connect the device directly to a firewall and share the same LAN. The company, on the other hand, might supply a switch, link the computers to it, and then connect the switch to the firewall. Additionally, it is feasible to utilize the firewall's graphical user interface. However, since the university was hesitant to reveal the credentials to students owing to their confidential rules and regulations, we used the ACL environment. It should be mentioned that internal threat analysis, detection methods, incoming traffic investigations, as well as auditing techniques, are not included in our analysis. Also, the development stage best practices for IoT devices were out of scope of this research work. Our goal was to help small and medium enterprises secure their networks and apply security practices considering the security challenges of smart devices as well as other resource limitations for deploying security practices in IoT environments.

This research has focused on the adoption of effective security practices for prevention of cyber-attacks in smart ecosystems. Various strategies, test beds and smart devices are yet to develop. For our future work, we will investigate the newest generation of routers, called Dream Machines, which are stated as an all-in-one security network device. Also, in future work, we will select best practices as a combination of detection and prevention techniques, considering the restricted knowledge and limited resources they have.

References

1. Kevin A (2009) That 'Internet of Things' thing. RFID J 22(7):97–114
2. Kenton W (2021) The Internet of Things, May 28. Retrieved from Investopedia.com. https://www.investopedia.com/terms/i/internet-things.asp
3. Hilton S (2016) Dyn analysis summary of friday Oct 21. Retrieved from Oracle+Dyn: https://perma.cc/YW5C-MDEV
4. Bansal S, Kumar D (2020) IoT ecosystem: a survey on devices, gateways, operating systems, middleware and communication. Int J Wireless Inf Networks 27:340–364. https://doi.org/10.1007/s10776-020-00483-7
5. Ahemd MM, Shah MA, Wahid A (2017) IoT security: a layered approach for attacks & defenses. In: International conference on communication technologies
6. S. A. Kumar, Vealey T (2016) Security in Internet of Things: challenges, solutions and future directions. In: 49th Hawaii international conference on system sciences, Koloa, HI, pp 5772–5781
7. Gupta R, Tanwar S, Tyagi S, Kumar N, Obaidat MS, Sadoun B (2019) HaBiTs: Blockchain-based telesurgery framework for healthcare 4.0. In: Proceedings of International Conference on Computer Information and Telecommunicaton Systems (CITS), pp 1–5
8. Anand P, Singh Y, Selwal A, Alazab M, Tanwar S, Kumar N (2020) IoT vulnerability assessment for sustainable computing: threats, current solutions, and open challenges. IEEE Access 8:168825–168853. https://doi.org/10.1109/ACCESS.2020.3022842
9. Gurkan T, Dimitrios GK, Gungor VC, Cengiz G, Erhan T, Erman A (2017) A survey on information security threats and solutions for machine to machine (M2M) communications. J Parallel Distrib Comput 142–154. https://doi.org/10.1016/j.jpdc.2017.05.021
10. Owasp (2016) Project, manufacturer IoT security guidance. Open web application security. Retrieved from OWASP.ORG. https://www.owasp.org/index.php/IoT_Security_Guidance

11. Hamad SA, Sheng QZ, Zhang WE, Nepal S (2020) Realizing an Internet of secure things: a survey on issues and enabling technologies. IEEE Commun Surv Tutor 22(2):1372–1391. https://doi.org/10.1109/COMST.2020.2976075
12. Bertino E, Choo K-KR, Georgakopolous D, Nepal S (2016) Internet of Things (IoT): smart and secure service delivery. ACM Trans Internet Technol 16:1–7
13. Noor M, Hassan WH (2019) Current research on Internet of Things (IoT) security: a survey. Comput Netw 283–294. https://doi.org/10.1016/j.comnet.2018.11.025
14. Deshmukh S, Sonavane SS (2017) Security protocols for Internet of Things: a survey. Proceedings of International Conference on Nextgen Electronic Technologies (ICNETS2), pp 71–74
15. Hatzivasilis G, Fysarakis K, Papaefstathiou I, Manifavas H (2018) A review of lightweight block ciphers. J Cryptograph Eng 8(2)
16. Maggi DQ (2018) When machines cannot talk: security and privacy issues of machine-to-machine data protocols. Retrieved from https://www.blackhat.com/us-17/briefings.html
17. Samaila MG, Neto M, Fernandes DA, Freire MM, Inácio PR (2018) Challenges of securing Internet of Things devices: a survey. Secur Priv
18. Lundgren L (2017) Taking over the world through MQTT-AfterMath. Retrieved from https://www.blackhat.com/us-17/briefings.html
19. McBride J, Arief B, Hernandez-Castro J (2018) Security analysis of Contiki IoT operating system. In: International conference on embedded wireless systems, pp 278–283
20. Mckay KA, Meltem LB, Turan S, Mouha N (2017) Report on lightweight cryptography. https://doi.org/10.6028/NIST.IR.8114
21. Tuna G, Kogias DG, Gungor VC, Gezer C (2017) A survey on information security threats and solutions for machine to machine (M2M) communications. J Parallel Distrib Comput 109: 142–154 (2017)
22. Chen B, Wan J, Celesti A, Li D, Abbas H, Zhang Q (2018) Edge computing in IoT-based manufacturing. IEEE Commun Mag 56(9):103–109
23. Liu X, Qian C, Hatcher WG, Xu H, Liao W, Yu W (2019) Secure Internet of Things (IoT)-based smart-world critical infrastructures: survey, case study and research opportunities, 79523–79544. https://doi.org/10.1109/ACCESS.2019.2920763
24. Rizvi S, Orra R, Coxa A, Ashokkumar P, Rizvi MR (2020) Identifying the attack surface for IoT network. Internet of Things. https://doi.org/10.1016/j.iot.2020.100162
25. Jurcut AD, Ranaweera PS, Xu L (2020) Introduction to IoT security. In: Liyanage M, Braeken A, Kumar P, Ylianttila M (eds) IoT security: advances in authentication, pp 27–64
26. Kolias et al (2017) DDoS in the IoT: Mirai and other botnets. Computer 50(7):80–84
27. Sharma PK, Chen M-Y, Park JH (2018) A software defined fog node based distributed blockchain cloud architecture for IoT. IEEE Access 6:115–124 (2018)
28. Conoscenti M, Vetrò A, Martin JC (2016) Blockchain for the Internet of Things: a systematic literature review. In: IEEE/ACS 13th international conference of computer systems and applications (AICCSA), pp 1–6
29. Fan K, Wang S, Ren Y, Yang K, Yan Z, Li H, Yang Y (2019) Blockchain-based secure time protection scheme in IoT. IEEE Internet Things J 4671–4679. https://doi.org/10.1109/JIOT.2018.2874222
30. Jin Y (2014) Embedded system security in smart consumer electronics. In: Proceedings of the 4th international workshop on trustworthy embedded devices
31. BCI Horizon Scan Report (2018) Retrieved from BSI: https://www.bsigroup.com/LocalFiles/en-GB/iso-22301/case-studies/BCI-Horizon-Scan-Report-2018-FINAL.pdf
32. An M (2018) A practical approach to emerging tech for SMBs: AI, blockchain, cryptocurrencies, IoT, and AR/VR. Retrieved from https://blog.hubspot.com/news-trends/emerging-tech-forsmbs
33. COOK S (2021) 60+ IoT statistics and facts. Retrieved from campritech: https://www.compritech.com/internet-providers/iot-statistics/
34. Leclair J (2016, April 22) Testimony of Dr. Jane Leclair before the U.S. house of representatives committee on small business. Retrieved from http://bit.do/sme-leclair

35. Loi F, Sivanathan A, Gharakheili HH, Radford A, Sivaraman V (2017) Systematically evaluating security and privacy for consumer IoT devices. In: Proceedings of ACM IoT S&P
36. Hamza A, Gharakheili HH, Sivaraman V (2020) IoT network security: requirements, threats, and countermeasures. Comput Sci > Crypt Secur
37. Blythe JM, Sombatruang N, Johnson SD (2019) What security features and crime prevention advice is communicated in consumer IoT device manuals and support pages? J Cybersecur
38. Code of Practice for Consumer IoT Security (2018, October). Retrieved from Department for Digital, Culture, Media & Sport. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/971440/Code_of_Practice_for_Consumer_IoT_Security_October_2018_V2.pdf
39. Fagan M, Megas K, Scarfone K, Smith M (2020) IoT device cybersecurity capability core baseline. Technical report. National Institute of Standards and Technology
40. ENISA (2017) Baseline security recommendations for IoT. European Union Agancy for cyber Security
41. Baseline Security Recommendations for IoT in the Context of Critical Information Infrastructures (2017, November). Retrieved from European Union Agency for Network and Information Security. https://op.europa.eu/en/publication-detail/-/publication/c37f8196-d96f-11e7-a506-01aa75ed71a1/language-en
42. ETSI. (2020). EN 303 645 cyber security for consumer internet of things: baseline requirements, June 2020. Retrieved from https://www.etsi.org/deliver/etsi_en/303600_303699/303645/02.01.01_60/en_303645v020101p.pdf
43. Geiger H, Kleiner A, Woods B (2017) Communicating IoT device security update capability to improve transparency for consumers, 14 July 2017. Retrieved from National Telecommunications and Information Administration. https://www.ntia.doc.gov/files/ntia/publications/draft_communicating_iot_security_update_capability_-_jul_14_2017_-_ntia_multistakeholder_process.pdf
44. Vidich S (2021) Trusted information security assessment exchange (TISAX), Mar 8. Retrieved from Microsoft.com. https://docs.microsoft.com/en-us/azure/compliance/offerings/offering-tisax
45. ISO/IEC DIS 27400 (2021) ISO
46. Piasecki S, Urquhart L, McAuley PD (2021) defence against the dark artefacts: smart home cybercrimes and cybersecurity standards. Comput Law Secur Rev. https://doi.org/10.1016/j.clsr.2021.105542
47. Babun L, Sikder A, Acar A, Uluagac A (2018) IoTdots: a digital forensics framework for smart environments. Arxiv
48. Iqbal M, Oladiran G, Magdy A, Bayoumi A (2017) A review on internet of things (IoT): security and privacy requirements and the solution approaches. Global J Comput Sci Technol
49. Shin S, Kwon S (2018) Two-factor authenticated key agreement supporting unlinkability in 5G-integrated wireless sensor networks. IEEE Access 11229–11241. https://doi.org/10.1109/ACCESS.2018.2796539
50. Lavanya M, Natarajan V (2017) Lightweight key agreement protocol for IoT based on IKEv2. Comput Electr Eng 580–594. Retrieved from https://doi.org/10.1016/j.compeleceng.2017.06.032
51. Wu F, Xu L, Kumari S, Li XK, Kumar D (2017) An enhanced mutual authentication and key agreement scheme for mobile user roaming service in global mobility networks. Ann Telecommun 72:131–144
52. Srinivasa J, Mukhopadhyaya S, Mishrab D (2017) Secure and efficient user authentication scheme for multi-gateway wireless sensor networks. Ad Hoc Netw 147–169. Retrieved from https://doi.org/10.1016/j.adhoc.2016.11.002
53. Chakrabarty S, Engels DW, Member S (2016) Secure IoT architecture for smart cities. In: 13th IEEE annual consumer communications & networking conference (CCNC), pp 812–813. https://doi.org/10.1109/CCNC.2016.7444889
54. Alaba FA, Othman M, Hashem IA, Alotaibi F (2017) Internet of Things security: a survey. J Netw Appl. https://doi.org/10.1016/j.jnca.2017.04.002

55. Babun L, Celik Z, McDaniel P, Uluagac A (2021) Real-time analysis of privacy-(un) aware IOT applications. Proc Privacy Enhanc Technol 2021(1)
56. Khan AY, Latif R, Latif S, Tahir S, Batool G, Saba T (2020) Malicious insider attack detection in IoTs using data analytics. IEEE Access 8:11743–11753. https://doi.org/10.1109/ACCESS.2019.2959047
57. Liang X, Kim Y (2021) A survey on security attacks and solutions in the IoT network. In: IEEE 11th Annual computing and communication workshop and conference (CCWC), 0853–0859. https://doi.org/10.1109/CCWC51732.2021.9376174
58. Peters R (2018) Securing the industrial internet of things in OT networks. Retrieved from Fortinet. https://www.fortinet.com/blog/industry-trends/securing-the-industrial-internet-of-things-in-ot-networks
59. Garcia-Morchon O, Kuma SS, Sethi M (2019) RFC8576: Internet of Things (IoT) security: state of the art and challenges
60. Toy N, Senthilnathan T (2019) Light weight authentication protocol for WSN using ECC and hexagonal numbers. Indonesian J Electr Eng Comput Sci (IJEECS) 443–450
61. Iqbal W, Abbas H, Daneshmand M, Rauf B, Bangash YA (2020) An in-depth analysis of IoT security requirements, challenges, and their countermeasures via software-defined security. IEEE IoT J 7(10):10250–10276
62. Tayyaba SK, Shah MA, Khan OA, Ahmed AW (2017) Software defined network SDN based internet of things IoT a road ahead. In: Proceedings of ACM international conference on future networks and distributed systems p 15
63. Miettinen M, Marchal S, Hafeez I, Asokan N, Sadeghi A-R, Tarkoma S (2017) IoT sentinel: automated device-type identification for security enforcement in IoT. In: Proceedings of IEEE 37th international conference on distributed computing systems (ICDCS), pp 2177–2184
64. Rao TA, Ehsan-ul-Hagh (2018) Security challenges facing IoT layers and its protective. Int J Comput Appl
65. NCSC (2020) Smart security cameras using them safely in your home, March 3. Retrieved from NCSC.GOV.UK. https://www.ncsc.gov.uk/guidance/smart-security-cameras-using-them-safely-in-your-home
66. Neshenko N, Bou-Harb E, Crichigno J, Kaddoum G, Ghani N (2019) Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations. IEEE Commun Surv Tutor 21(3):2702–2733. https://doi.org/10.1109/COMST.2019.2910750
67. Czyz J, Luchie M, Allman M, Bailey M (2016) Don't forget to lock the back door! a characterization of ipv6 network security policy. Netw Distrib Syst Secur (NDSS)
68. Pauli D (2016) IoT worm can hack Philips Hue lightbulbs, spread across cities, Nov 10. Retrieved from Theregister.com. https://www.theregister.com/2016/11/10/iot_worm_can_hack_philips_hue_lightbulbs_spread_across_cities/
69. Ferencz K, Domokos J, Kovács L (2021) Review of Industry 4.0 security challenges. In: 2021 IEEE 15th international symposium on applied computational intelligence and informatics (SACI), pp 245–248. https://doi.org/10.1109/SACI51354.2021.9465613
70. Payne BR, Abegaz TT (2017) Securing the Internet of Things: best practices for deploying IoT devices. Comput Netw Secur Essentials
71. (2020) Tips to secure your internet of things advice. Australian cyber security. Retrieved from https://www.cyber.gov.au/sites/default/files/2020-08/Tips%20to%20secure%20your%20Internet%20of%20Things%20device%20%28AUG%202020%29.pdf
72. Mallikarjunan KN, Muthupriya K, Shalinie SM (2016) A survey of distributed denial of service attack. In: 10th International conference on intelligent systems and control (ISCO). https://doi.org/10.1109/ISCO.2016.7727096
73. Alabady SA, Al-Turjman F, Din S (2020) A novel security model for cooperative virtual networks in the IoT era. Int J Parallel Program 48(2):280–295
74. Gopal M, Meerolla G, Jyostna P (2018) Mitigating mirai malware spreading in IoT environment. In: Reddy Lakshmi Eswari, Magesh E (eds) In: 2018 International conference on advances in computing, communications and informatics (ICACCI), pp 2226–2230. https://doi.org/10.1109/ICACCI.2018.8554643

75. Cisco (2021) CLI Book 3: Cisco ASA series VPN CLI configuration guide, 9.7. Cisco. https://www.cisco.com/c/en/us/td/docs/security/asa/asa97/configuration/vpn/asa-97-vpn-config.pdf
76. Siddharth S (2020) Factory reset your Raspbian OS. Retrieved from Github.com. https://github.com/shivasiddharth/RaspberryPi-Factory-Reset
77. Rizvi S, Pipetti R, McIntyre N, Todd J, Williams I (2020) Threat model for securing internet of things (IoT) network at device-level. Internet of Things. Retrieved from https://doi.org/10.1016/j.iot.2020.100240
78. Gurunath R, Agarwal M, Nandi A, Samanta D (2018) An overview: security issue in IoT network. In: 2018 2nd international conference on I-SMAC (IoT in social, mobile, analytics and cloud)
79. Ali B, Ismail A (2018) Cyber and physical security vulnerabilityassessment for IoT based smart homes. Sensors 2–17
80. Ali M, Vecchio M, Pincheira M, Dolui K, Antonelli F, Rehmani M (2019) Applications of blockchains in the internet of things: a comprehensive survey. IEEE Commun Surv Tutor
81. Alladi T, Chamola V, Sikdar B, Choo KR (2020) Consumer IoT: security vulnerability case studies and solutions. IEEE Consum Electron Mag. https://doi.org/10.1109/MCE.2019.2953740
82. Burhan M, Rehman RA, Khan B, Kim B-S (2018) IoT elements, layered architecture. Sensors 1–38
83. Cappelli DM, Moore AP, Trzeciak RF (2012) The CERT guide to insider threats: how to prevent detect and respond to information technology crimes (theft Sabotage Fraud)
84. Chen K, Zhang S, Li Z, Zhang Y, Deng Q, Ray S, Jin Y (2018) Internet-of-Things security and vulnerabilities: taxonomy, challenges, and practice. J Hardw Syst Secur 97–110
85. Demiris G, Hensel BK (2018) Technologies for an aging society: a systematic review of "smart home applications." IMIA Yearbook Med Inf 47:33–40
86. Hair JF, Samouel, Page M (2015) The essentials of business research methods
87. Hill K (2015) This guy's light bulb performed a DoS attack on his entire smart house. Retrieved from Splinter. https://splinternews.com/this-guys-light-bulb-performed-ados-attack-on-his-enti-1793846000
88. Holst A (2021) statisa.com, Jan 20. Retrieved from Statisa Research Department. https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/
89. Ii N, Tech YM, Pai V (2018) Survey on IoT security issues and security protocols. Int J Comput Appl 180:975–987
90. Institute BC (2021) Supply chain resilience report 2021. Institute, Business Continuity
91. Kim A, Oh J, Ryu J, Lee K (2020) A review of insider threat detection approaches with IoT perspective. IEEE Access 8:78847–78867. https://doi.org/10.1109/ACCESS.2020.2990195
92. Labs M (2017) McAfee Labs threat report. Mcafee.com
93. Lim H-K, Kim J-B, Heo J-S, Han Y-H (2020) Federated reinforcement learning for training control policies on multiple IoT devices. Sensors. https://doi.org/10.3390/s20051359
94. Lin J, Yu W, Zhang N, Yang X, Zhang H, Zhao W (2017) A survey on Internet of Things: architecture enabling technologies security and privacy and applications. IEEE Internet Things J 4:1125–1142
95. LLC, P. I. (2018). 2018 Cost of insider threats: global. ObserveIT. Retrieved from https://153j3ttjub71nfe89mc7r5gb-wpengine.netdna-ssl.com/wp-content/uploads/2018/04/ObserveIT-Insider-Threat-Global-Report-FINAL.pdf
96. Novo O (2018) Blockchain meets IoT: an architecture for scalable. IEEE Internet Things J 5(2):1184–1195
97. Scrutton R, Beames S (2013) Measuring the unmeasurable: upholding rigor in quantitative studies of personal and social development in outdoor adventure education. https://doi.org/10.1177/1053825913514730
98. Sharma PK, Chen M-Y, Park JH (2018) A software defined fog node based distributed blockchain cloud architecture for IoT. IEEE Access 6:115–124
99. Sonicwall. (2021). *Sonicwall Cyber threat report.* Sonicwall. Retrieved from https://www.sonicwall.com/medialibrary/en/white-paper/mid-year-2021-cyber-threat-report.pdf

100. Theis M, Trzeciak RF, Costa DL, Moore AP, Miller S, Cassidy T, Claycomb WR (2020) Common sense guide to mitigating insider threats. https://doi.org/10.1184/R1/12363665.v1
101. Thomson I (2017). Firmware update blunder bricks hundreds of home 'smart' locks, Aug 2017. Retrieved from the register: https://www.theregister.co.uk/2017/08/11/lockstate_bricks_smart_locks_with_dumb_firmware_upgrade
102. Xu L, Guan Y, Singhal V (2021) Network attack trends: Internet of threats (Nov 2020–Jan 2021), Apr 12. Retrieved from https://unit42.paloaltonetworks.com/. https://unit42.paloaltonetworks.com/network-attack-trends-winter-2020/