Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review

Abhishek Pokhrel
Faculty of Computer Science
Østfold University College
1783 Halden, Norway
abhishep@hiof.no

Vikash Katta
Dept. Risk, Safety and Security
Institute of Energy Technology
1777 Halden, Norway
vikash.katta@ife.no

Ricardo Colomo-Palacios
Faculty of Computer Science
Østfold University College
1783 Halden, Norway
ricardo.colomo-palacios@hiof.no

ABSTRACT

The advancements in the field of internet of things, artificial intelligence, machine learning, and data analytics has laid the path to the evolution of digital twin technology. The digital twin is a high-fidelity digital model of a physical system or asset that can be used e.g. to optimize operations and predict faults of the physical system. To understand different use cases of digital twin and its potential for cybersecurity incident prediction, we have performed a Systematic Literature Review (SLR). In this paper, we summarize the definition of digital twin and state-of-the-art on the development of digital twin including reported work on the usability of a digital twin for cybersecurity. Existing tools and technologies for developing digital twin is discussed.

CCS CONCEPTS

• Security and privacy~Intrusion/anomaly detection and malware mitigation~Intrusion detection systems • Security and privacy~Intrusion/anomaly detection and malware mitigation • Computing methodologies~Modeling and simulation~Simulation types and techniques~Real-time simulation

KEYWORDS

Multivocal literature review, IoT, Digital Twin, Cybersecurity, Fault detection, Incident prediction

ACM Reference format:
Abhishek Pokhrel, Vikash Katta and Ricardo Colomo-Palacios. 2020. Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review. In *Proceedings of IEEE/ACM 42nd International Conference on Software Engineering Workshops (ICSEW'20). ACM, Seoul, Republic of Korea, 8 pages.* https://doi.org/10.1145/3387940.3392199

ICSEW'20, May 23–29, 2020, Seoul, Republic of Korea

ACM ISBN 978-1-4503-7963-2/20/05…$15.00
https://doi.org/10.1145/3387940.3392199

1 Introduction

The shift towards the 4th industrial revolution and Industry 4.0 has increased the level of automation of the industrial systems and processes and connectivity between them. In this revolutionary trend, one of the focus areas is creating transparency between the physical world and the digital world (cyberspace) and this led to a new concept, called digital twin [1]. Some related or partially overlapping conceptual terms are the digital counterparts, virtual twin, virtual object, product agent, and an avatar [2]. A digital twin is a replica model residing in a virtual environment that includes everything that is known about an object. It is a replica of physical counterparts that change with the current environment in real-time to assist in monitor, test, treat and maintain a physical system [16].

IoT is an extension and expansion of an Internet-based network that expands the communication among the physical world, the virtual world, the digital world, and the society by taking machine-to-machine communication (M2M) to the next level [5]. With the increasing number of heterogeneous devices connected to IoT and generating data, cybersecurity is an inevitable problem that must be solved for the further development of IoT [4]. IoT devices are becoming the new weakest link in the security chain in the modern world making themselves a powerful amplifying platform for cyberattacks [10]. Thus, it must be ensured that the IoT cybersecurity is not an afterthought when developing and using IoT devices [12]. While digital twin is regarded as one of the top technological trends, and more and more industries are looking into the usability of this technology to optimize the development and maintenance of their systems and processes, very less is known or published on the topic of usability of a digital twin for cybersecurity. This paper presents our work on an exploratory study, using a literature review, to review the published work on digital twin and their usability for cybersecurity. The concept of the digital twin was first coined by Michael Grieves back in 2002 at the University of Michigan [19]. Furthermore, this concept was greatly expanded in his book called, Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management [18]. Probably, the first definition of the digital twin was given by NASA as "an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin. The digital twin is ultra-realistic and may consider one or more important and interdependent vehicle systems, including propulsion/energy storage, avionics, life support, vehicle structure, thermal management/TPS, etc." [41]. Most literature and recent works show that this concept is mostly linked to maintenance, optimization, and simulation but this concept can also be used to enhance the safety and security of cyber-physical systems.

The rest of the paper is organized as follows. Section 2 describes the multivocal literature review methodology followed, and the research questions addressed. Section 3 presents the results of the literature review. Section 4 narrates the findings regarding the literature review and research objectives. Finally, Section 5 concludes the paper.

This paper is part of a Master-level Business Project done by a student (first author) at the Institute for Energy Technology (IFE). This paper presents a shortened version of the work -Literature review – reported in the project technical report [29].

2 Research Methodology

In this section, an outline of the research methodology based on a Multivocal Literature Review (MLR), research questions and data collection procedure, is presented.

2.1 Multivocal Literature Review (MLR)

A type of SLR study where the data from both scientific and grey literature (GL) are accepted is referred to as Multivocal Literature Review (MLR). This study covers the gap between academic literature and industrial/practical knowledge. MLR stands where purely research-based surveys or studies are unable to give significant decisions about complex interventions and a national or regional context for differences in implementation. Not all SLRs need to include GL and become MLRs but if extending the scope and including GL will bring value and benefits to study, researchers should consider evaluating MLR [17].

The academic literature available on a digital twin is in a significantly limited amount. Our findings indicated that including GL to this study can give significant advantages as the scope of the digital twin is very broad and ambiguous. We found examples of numerous practitioners, companies and their sources being ignored in previous studies. Those sources can not only be academic studies and books but also a wide variety of unpublished non-research literature available online such as blog posts, white papers, presentation videos, and tools. Furthermore, the digital twin is a field of both scientific and practical interest where the research for this topic is growing enormously since the last decade. Besides, to our knowledge, we also found there is no MLR study conducted on this topic until now. So, we chose to conduct our study utilizing MLR according to the guidelines described by Garousi [17]. The MLR process conducted followed six steps: a) Define research questions, search strategy and search string, b) Define inclusion/exclusion criteria, c) Perform a search on the selected databases, d) Filter/Selection procedures among the founded papers, e) Review of the papers that passed filtering process and f) Summarise results and findings.

2.2 Research Questions

The main purpose of the MLR is to observe, classify, review, document, analyze and summarize the state of the art and practice related to digital twin and its use on cybersecurity. Based on the purpose and objective of the study, we have the following two research questions.

RQ1: What are the reported uses of the digital twin?
RQ2: What are the reported uses of the digital twin in cybersecurity for fault detection and incident prediction?

2.3 Search Strategy

The scope of this step is to characterize the search and evaluation strategy which allows thorough search of the available literature needed to answer the proposed research questions. This helps in identifying the primary sources from the scientific and grey literature [6].

2.3.1 Search data sources. Once the search string has been defined, the next step is to determine where the searches will be performed. We performed the searches on following digital libraries to identify formally published literature: IEEE Xplore; Science Direct; Google Scholar; ACM Digital Library; Springer Link; Web of Science. We have used a conventional web search engine, Google to find out the grey literature including white papers, blogs, articles, videos, etc. For searching, the rules of each search engine on an individual basis had been considered since the search engines have different search categories, constraints and various degrees of support for logical search keyword compositions.

2.3.2 Search term. The following search string has been used to perform searches in the digital libraries:

"Digital Twin" AND (("cyber security") OR ("cyber-security") OR ("cybersecurity")) AND (("incident prediction") OR ("fault detection") OR ("prediction") OR ("detection"))

2.3.3 Search process. This process allows the selection of primary studies from the scientific literature. This process is composed of four phases that follow the test-retest approach to find the scientific primary studies [7]. The same steps were also used to identify grey literature while searching on Google. The four phases are:
Phase 1. Initial search: We will apply the search string to each digital resource to identify the studies/literature related to the topic under review. Only sections of title, abstract and keywords are considered in this phase to find the relevant studies.
Phase 2. Remove duplicates: Papers found during the first phase of the selection process will be checked for duplicates. There may be the same paper retrieved from different digital sources. If any paper is found to be duplicated, only the paper which provides more

data, complete references of the publication, etc are selected for further process.

Phase 3. First selection process: In this phase, all the selected papers after Phase 2 will be evaluated with inclusion/exclusion criteria. During this phase, the title and abstract of each paper will be reviewed. If the paper is out of the inclusion criteria it will not be selected for the further review process. But if the paper fulfils the criterion for inclusion then it will be included for the next phase of the selection process.

Phase 4. Second selection process: The papers selected after Phase 3 will be thoroughly reviewed. It is a rigorous study to evaluate whether a paper contains relevant information for the topic under review. If the paper contains non-relevant information, it will not be selected. The quality of the work presented in the paper is also performed in this phase.

2.3.4 Search selection. The inclusion/exclusion criteria are a method for source selection if the papers retrieved from the searches can be included or excluded in the study.

Inclusion Criteria

The criteria for the inclusion are:

- The literature should explicitly discuss the state of art concepts of the digital twin.
- The literature must be related to the research questions.
- The literature should discuss various implementations of the digital twin.
- The literature should discuss the application of digital twins in cybersecurity.
- The literature should have its implementations discussed in detail.

Exclusion Criteria

The criteria for the exclusion are:

- The literature that is written in any language other than English.
- The literature where the abstract and full text is not available.
- The literature that is not relevant to our research.
- The literature that is duplicated/repeated.

2.3.5 Search execution. From the search strings, a total of 751 articles were retrieved, 660 articles were excluded at the first phase after the Title, Keywords and Abstract were reviewed. And this resulted in 91 articles being selected from Phase 1. In Phase 2, all the duplicate articles were removed, which left 77 articles. And then, those articles were checked with inclusion/exclusion criteria which reduced the total articles to 48. After the final phase – Phase 4- of quality assessment, the total number of final selected papers was 17. The detail categorization and analysis of the search results as per the individual database in each phase are shown below.

Digital Source	No. of Initial Result	Phase1	Phase2	Phase3	Phase4
IEEE Xplore	1	1	1	1	1
Science Direct	47	14	14	10	1
Scopus	50	8	6	5	3
Google Scholar	587	56	45	28	11
SpringerLink	65	11	10	3	1
Web of Science	1	1	1	1	0
Total	751	91	77	48	17

Table 1 Overall search results per digital database

For grey literature, firstly, we performed a google search using the general terms and keyword based as defined in Section 2.3.2. The search string resulted in 9 articles along with over 3800 results in the News section and around 3000 results in the Videos section. We have observed that pertinent results usually appear only in the first few pages. We have checked the first 10 pages (i.e., somewhat a search “saturation” effect) and only continued further if needed, e.g. when the 10th page results still looked relevant and not mentioned in the previous pages. From these grey sources, we have selected approximately 5 sources including a YouTube video, blogs, and news articles. Google chrome browser was used for the search where the respective google account was logged out and all the sessions and browsing history were cleared to ensure that the search results will not be directed towards one’s preferences. The selected grey articles are [39], [40], [42], [43], [44].

To ensure the inclusion of all relevant articles, we have used forward and backward snowballing as recommended by MLR guidelines, on the set of sources already in the pool. Forward snowballing is identifying articles that have cited the articles found in the search and backward snowballing is identifying articles from the reference lists. The articles found after snowballing are [2], [9], [22], [30], [31], [32], [35], [36], [37], [38].

All the finalised articles were used to answer two research questions listed in Section 2.2. The sources and their relevance to the appropriate research questions are shown below:

Research Questions	References
RQ1	[2], [9], [15], [22], [24], [27], [30], [31], [32], [34], [35], [36], [37], [38], [39], [40], [42]
RQ2	[3], [10], [11], [13], [14], [15], [16], [21], [24], [27], [31], [33], [35], [43], [44]

Table 2 Sources and their relevance to research questions

3 Results

In this section, the authors present the results of the MLR according to the research questions.

3.1 RQ1: What are the reported uses of the digital twin?

To answer this question, we have collected information related to digital twin concepts, implementation, case studies in industrial sectors.

The digital twin concept has become one of the most exciting R&D areas in autonomous and intelligent manufacturing with the rapid development of advanced technologies in mechanical, electrical and information technologies, as well as virtual simulation and data acquisition technologies [38]. Smart manufacturing is one of the strategic priorities shared by all the major manufacturing initiatives such as Industry 4.0 and the Industrial Internet [36]. Manufacturing is shifting from traditional knowledge-based intelligent manufacturing to data-driven and knowledge-enabled smart manufacturing which is supported by smart technologies such as IoT, digital twin, big data and cyber-physical systems (CPS) [36].

As Chris O'Connor stated, digital twin, can be implemented for whole product life cycle management (design, production, operation, testing to maintenance) [42]. In the product design phase, the digital twin can be used to design new products in a more responsive, efficient, and informed manner. The authors of [36] discussed how System of Systems (SoS) level digital twin can shorten the design cycle with great improvement in the process and also reducing the cost in terms of time and money. Similarly, authors of [31] proposed a digital twin model of a tool system and discussed how a feedback mechanism of a digital twin can effectively help to correct errors in the tool system design stage. Digital twin replaces the traditional design processes and practices, that used a "build it and tweak it" approach, instead, this concept allows manufacturers to validate the design against the requirements in the early stages of the process [34]. And several companies (e.g. NASA and US Air Force) had applied the digital twin technology in the product design phase to predict the future performance and status of the system built [38].

In the product manufacture phase, with the use of a digital twin, the manufacturing system can be highly enhanced [38]. A digital twin of a system can enclose Industry 4.0 along with Information and Communications Technology (ICT) advances such as CPS, IoT and cloud computing which are not only helpful in the development and design stages but also in the production stages to check that production is running smoothly, detect wear and tear without needing to halt production or predict component failure and other disruptions [2] [35]. Digital twin plays a pivot role as the impetus in the vision of the "smart manufacturing" and can influence future manufacturing from the following aspects [27]:

- Digital twin for manufacturing assets: All the manufacturing assets can be connected and replicated in cyberspace via their digital twin.
- Digital twin for people: Digital twin can also represent people/workers including their data (weight, health data, activity data, and emotional status) which can aid in establishing models to better understand the wellbeing and working conditions of humans in a factory.
- Digital twin for factories: A live factory environment can also be abstracted with digital twins with its complete operational visibility and flexibility.
- Digital twin for production networks: By connecting manufacturing assets, people and service via digital twin, every aspect of business can be virtually represented.

Furthermore, this is often a case that the global trend tends to focus on fostering a pure machine to machine manufacturing environment in the future where it overlooks the impact of the human component on production and business performance. In [24], a human-centric paradigm within Industry 4.0 towards Smart Factory harnessing the technology of digital twin has been presented. Their main aim is to deliver ubiquitous valuable knowledge within and among all the elements of the smart factory (including employees) to achieve full information symmetry. For this, an architecture of the Service-oriented digital twin is defined where apps and services are used which can be accessed at any place and at any time via the cloud enabling both on-site and remote monitoring.

In the product operation and maintenance phase, the digital twin can be implemented to improve the safety and reliability of any system. Fault prediction based digital twin if used in the life cycle of the system can be of great significance to improve the reliability, availability, and safety of the equipment or machine. With a digital twin, it is possible to process information from complex or different sources and monitor the real working conditions. The prediction results can be placed in the digital model to show the accuracy of the cutting amount or the better feed of the tool system and system behaviour can be automatically optimised [31].

In summary, as stated in [32], the various applications of the digital twin in industrial production can be summarised: Increased productivity; Reduced complexity; Time and cost savings; Identification of new business fields; Increased quality; Increased cyber-security; Reduction or risks; Increased efficiency; Improved safety and reliability; Real-time remote monitoring and control; Predictive maintenance and scheduling; Scenario and risk assessment; Efficient and informed decision support system; Better documentation and communication.

Some common examples are: Tesla is working on developing a digital twin for every car it produces, allowing for synchronous data transmission between cars and its factory [38] to improve the safety and reliability of the vehicles. DNV GL explains the major incident that happened in the North Sea where unmanned barge collided with offshore oil platforms [39] could have been predicted if implemented digital twin [40]. Similarly, [37] presents a proposal for reengineering the structural life prediction and management process with a real-life scenario for an aircraft where each type of physics has its separate model. Laaki et al. [22] demonstrated a digital twin framework to perform remote surgery, investigating mobile networks.

Current commercial and open-source digital twin solutions

We explored currently available tools and technologies which are either fully commercial or open-source solutions. Some of the commercial software solutions implementing industrial digital twin technology are [29] [9]: General Electric (GE) digital twin, Windchill developed by PTC, Dassault Systèmes (DS) software called "Built to Operate", DXC's digital twin, "Simcenter 3D" developed by Siemens, Bosch IoT Suite, Microsoft's IoT service, Seebo digital twin, and ANSYS Twin Builder. On the other hand, the opensource community is also increasingly developing software and hardware solutions for the experiments in Industry 4.0 and Smart Manufacturing. Nonetheless, there are only a few opensource tools for developing digital twins: Eclipse Ditto and CPS Twinning [15].

3.2 RQ2: What are the reported uses of the digital twin in fault and incident prediction?

Through this research question, our goal was to explore digital twin implementation in cybersecurity for fault and incident prediction through the retrieved primary studies. We describe reported use cases of digital twins to improve the security of a system and to support security analysis.

Intrusion Detection

In a study conducted by Eckhart and Ekelhart [13], the authors presented a framework called CPS Twinning where the input to any system is implicitly defined so as it is relevant to security and safety analysis to detect intrusions. In another study [14], the same authors extended their framework, CPS Twinning so that digital twins can be generated based on the specification of CPSs. Intrusions can be easily detected with a real-time comparison of specification-based digital twin signals and real device signals. Additional security rules can be defined in the specifications for continuous safety and security. The same authors gave a new framework based on digital twin "A Cyber Situational Awareness Framework" where they combine their previous works to add more features such as visual feedback to users for detected intrusions and 'record and replay' feature that assists in recovering past states [15].

Anomaly Detection

In a study conducted by Gaikwad et al. [16], the authors concluded that the digital twin approach in additive manufacturing (AM) can detect potential anomalies introduced through malicious cyber-physical intrusions while utilizing in-situ sensor data and machine learning. A similar approach is presented in [3], where a digital twin architecture for AM is introduced that uses a formal specification language, signal temporal logic (STL). This digital twin architecture will detect anomalies on each process signals if any STL specification is violated.

A digital twin, specifically called dynamic digital model (DDM), proposed for transmission systems is proven to detect failures/anomalies before they occur by simulating the current signals with historical data or data trend and comparing with the measured signals [21].

Another interesting idea is presented by Dr. Datta on the digital twin in the form of digital swarm or flock [11], where hundreds of instances of the physical event are created as a digital duplication which may provide a wealth of information/data to understand patterns, predict faults, detect anomalies and use true "big" data to feed other functions such as security threats policy, intruder detection.

Monitoring (Remote and On-site)

A service-oriented digital twin becomes a virtual representation of the CPS enabling ubiquitous knowledge accessible at fingertips through apps and services. This model can provide remote monitoring through different services such as fault diagnosis and state monitoring service, prognostics and scenario optimisation service, manufacturing scenario execution service and the notification service. The same model also enables on-site monitoring with all those remote services and in addition "Augmented Assistance & Tutoring Service" [24]. A similar type of model is developed by STEP Tools Inc. [27], called Digital twin machining application that allows real-time remote monitoring through Web-based applications to access all the data in the Digital Twin. Another tool MTConnect-based Cyber-Physical Machine Tool allows near real-time remote machine monitoring [27].

Virtual Commissioning

A six-layer digital twin architecture presented in [33] consists of a practical layer, capable of the role of virtual commissioning. The digital model can control the sequence of process events without data flow between the physical and digital objects. This is essential when rapid prototyping and testing are needed. For example, detecting misconfiguration while testing on the virtual environment before integrating with the real environment, detecting manipulation by an attacker when the physical device is not consistent, perform system tests on virtual model and simulation [13]. This approach can lead to better product quality and increased safety.

Autonomy

The digital twin can be used for forward simulations to anticipate the consequences of actions in each situation. This information can be used to create autonomous systems that are able to execute high-level task specifications without explicitly being programmed. These systems can take autonomous decisions over action alternatives to react to variation in products, production volume and the context of intrusions, cyber-attacks or in case of fault detection [35]. An example of an autonomous system is given in [10], where the digital twin supports the simulation of Smart Car's operational lifecycle to perform security, safety and operational analyses, predict insights and create analytics-based decisions that are automatically sent to respective stakeholders, either suggesting operational changes or performing operational decisions with or without human interventions.

Predictive Analytics

The digital twin based predictive maintenance software takes in the real-time data and analyses it against historical data about the physical models' failure modes, their criticality, their weak parts and scenarios where the physical model can be possibly infiltrated [44]. The future condition of a physical twin can be predicted before any malfunction occurs by forecasting errors and problems in facilities [33]. Similarly, ABB [43] presented their work on how digital twins can be integrated into the marine industry. The digital twin is implemented onboard and onshore and the data collected in their model in onboard twin goes through edge analytics to predict the condition of the actual physical asset and the information is sent onshore for cybersecurity analysis.

Similarly, [31] proposes a new model called the five-dimensional digital twin model of tool system over the three-dimensional model presented in most studies. This model performs data analytics (descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics) for identifying fault patterns and making decisions.

Documentation and communication

The digital twin provides a mechanism for identifying and describing behaviors and can be used as a mechanism for communication and documentation to illustrate equipment behaviors [33].

4 Discussions

The literature review conducted presents a detailed analysis of the concept of digital twin, its definition, its history, and use cases. The review started with the justification of the need for the MLR and grey literature and followed by the MLR settings to extract necessary literature relevant to the study under review. It involved an examination of the literature found in the database using a search string. The main aim of the review was to answer the research questions mentioned in Section 2.2. The literature review focused on those studies that addressed proper concepts of digital twin and its application on cybersecurity as reported by scientific and grey literature. The papers preferred after source selections (i.e. mentioned in Section 2.3.5) were published in the last 10 years with an increase in studies since 2015. Most research is done by the organizations in Europe, North America. Some of the companies are such as DNV GL – Digital Solutions, IFE, General Electric, Hirotec Corporation, Rockwell Automation, ABB and Siemens [4]. The results of the different literature evaluations, the conclusions of the analyzed studies and the perceived benefits of their authors allow us to conclude that digital twin has potential in the industrial sector and can be extended to many other fields for improving safety and security.

Our paper contributes a new procedure for performing MLR that establishes all the steps involved in the search and evaluation strategies to get the quality primary studies from both the scientific and grey literature sources. The procedure is based on the practices and guidelines proposed in [17] that take into consideration grey literature in systematic reviews. By conducting this MLR, we identified multiple primary studies that were further evaluated using the quality assessment. After this, both research questions were discussed using the quality papers found in the MLR we performed.

We found that the digital twin can be applied through the whole product lifecycle from concept to product implementation. One can simulate and evaluate a product at a conceptual stage before manufacturing the product itself. This usage of a digital twin has been proven to be applicable in autonomous industries, spacecraft industries, agriculture industries, and several other industries. Our study shows that digital twin can be used to improve the security and safety of the system. Digital twin, if implemented can detect intruders, anomalies and we can set up a notification or alarm service or maybe we can set up our system to automatically change the configuration to be safe from cyberattacks. We can utilise data analytics to predict future behavior. We can use the technologies of augmentation reality (AR) to detect the malfunction in real-time. There are lots of applications of digital twin through which we can predict the threats or faults in the system to be safe from any hazards. All to all, there are few grey literatures where it is imagined, we (human beings) will have our own digital twin to assist in the medical sector [45]. There is an exponential increase in the research on this field since the last decade.

We also found different names (types) of digital twins with their multiple implementations. In case of any organization using high fidelity simulation models – Digital twin of a process or part of the process, it has many implementations [34] [1] such as

- Stand Alone Digital Twin – Disconnected
- In-line Digital Twin
- Digital twin with optimal and model-predictive control techniques combined with advanced machine learning capabilities
- Embedded Digital Twin.

We also found new approaches using digital twin concept namely, Digital swarm or flock [11], Cyber Twin [2], Machine Tool Digital Twin (MTDT) [23], Digital Twin-based Cyber-Physical Production System(DT-CPPS) [12], Edge-deployed digital twin [46], Digital Avatar [46].

And, we found many different definitions of the digital twin. Some of them are,

- According to Dr. Datta [11], the concept of digital twin posits that the flow of data, process, and decision is captured in a software avatar that mimics the operation.
- The digital twin is a technology to model assets with all their geometrical data, kinetic functionality and logical behavior using digital tools [25].
- A digital twin is virtual (i.e. digital), it includes both static(i.e. design documents, process specifications, and so forth) and dynamic(i.e. data acquisition and simulation) parts, and it addresses every instance of its twin product or process for its total life cycle [20].
- The digital twin is a virtual counterpart to actual physical devices(entities) that combines many Artificial Intelligence(AI) -based technologies and methods, real-time predictive analyses, and forecasting algorithms performing on top of Big Data derived from the Internet of Things(IoT) sensors and acquired historical data [8].
- The digital twin is defined as a virtual model in the cloud of a product, process, physical asset, that is persistent even if its physical counterpart is not always online/connected [5].
- The digital twin is a simulation process using physical models and sensors to acquire data and complete mapping in virtual space to reflect the corresponding entity's lifecycle process [26].
- The digital twin is a simulation environment that can accurately represent the temporal dynamics of the real-world systems [28].
- A virtual image of the physical device or system is created that simplifies the accessibility of data and verification of properties. This structured collection of data and algorithms form the device's "digital twin" [46].

Hence, more research and experimental outcomes are needed to observe the full potential of a digital twin as a learning scope.

5 Conclusions

This paper presents a multivocal literature review (MLR) on the state of the art and practice related to digital twin and its use in cybersecurity. MLR includes both scientific and grey literature. In this paper, we have reported the published work on digital twins and their usage in several industrial sectors. We have also reported published work on the usability of a digital twin for cybersecurity, and the list of tools and technologies used for the development of digital twin. Our future work will focus on the development of digital twin approaches for cybersecurity incident prediction. We will build our approach based on the existing methods and tools identified through the MLR.

ACKNOWLEDGMENTS

This work is done as a part of the CybWin project funded by the Norwegian Research Council.

REFERENCES

[1] Edmary Altamiranda and Eliezer Colina. 2019. A System of Systems Digital Twin to Support Life Time Management and Life Extension of Subsea Production Systems. In *OCEANS 2019 - Marseille*, 1–9. DOI:https://doi.org/10.1109/OCEANSE.2019.8867187

[2] Mikel Armendia, Aitor Alzaga, Flavien Peysson, Tobias Fuertjes, Frédéric Cugnon, Erdem Ozturk, and Dominik Flum. 2019. Machine Tool: From the Digital Twin to the Cyber-Physical Systems. In *Twin-Control: A Digital Twin Approach to Improve Machine Tools Lifecycle*, Mikel Armendia, Mani Ghassempouri, Erdem Ozturk and Flavien Peysson (eds.). Springer International Publishing, Cham, 3–21. DOI:https://doi.org/10.1007/978-3-030-02203-7_1

[3] Efe C. Balta, Dawn M. Tilbury, and Kira Barton. 2019. A Digital Twin Framework for Performance Monitoring and Anomaly Detection in Fused Deposition Modeling. In *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, 823–829. DOI:https://doi.org/10.1109/COASE.2019.8843166

[4] Florian Biesinger and Michael Weyrich. 2019. The Facets of Digital Twins in Production and the Automotive Industry. In *2019 23rd International Conference on Mechatronics Technology (ICMT)*, 1–6. DOI:https://doi.org/10.1109/ICMECT.2019.8932101

[5] Theodor Borangiu, Damien Trentesaux, André Thomas, Paulo Leitão, and Jose Barata. 2019. Digital transformation of manufacturing through cloud services and resource virtualization. *Computers in Industry* 108, (June 2019), 150–162. DOI:https://doi.org/10.1016/j.compind.2019.01.006

[6] Alejandro Calderón, Mercedes Ruiz, and Rory V. O'Connor. 2018. A multivocal literature review on serious games for software process standards education. *Computer Standards & Interfaces* 57, (March 2018), 36–48. DOI:https://doi.org/10.1016/j.csi.2017.11.003

[7] Alejandro Calderón, Mercedes Ruiz, and Rory V. O'Connor. 2018. A multivocal literature review on serious games for software process standards education. *Computer Standards & Interfaces* 57, (March 2018), 36–48. DOI:https://doi.org/10.1016/j.csi.2017.11.003

[8] Violeta Damjanovic-Behrendt. 2018. A Digital Twin-based Privacy Enhancement Mechanism for the Automotive Industry. In *2018 International Conference on Intelligent Systems (IS)*, 272–279. DOI:https://doi.org/10.1109/IS.2018.8710526

[9] Violeta Damjanovic-Behrendt and Wernher Behrendt. 2019. An open source approach to the design and implementation of Digital Twins for Smart Manufacturing. *International Journal of Computer Integrated Manufacturing* 32, 4–5 (May 2019), 366–384. DOI:https://doi.org/10.1080/0951192X.2019.1599436

[10] Violeta Damjanovic-Behrendt, Michaela Mühlberger, Cristina de Luca, Thomos Christos, Heinz Weiskirchner, and Edin Arnautovic. Violeta Damjanovic-Behrendt. 45.

[11] Shoumen Palit Austin Datta. 2017. Emergence of Digital Twins - Is this the march of reason? *jim* 5, 3 (November 2017), 14–33. DOI:https://doi.org/10.24840/2183-0606_005.003_0003

[12] Kai Ding, Felix T. S. Chan, Xudong Zhang, Guanghui Zhou, and Fuqiang Zhang. 2019. Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors. *International Journal of Production Research* 57, 20 (October 2019), 6315–6334. DOI:https://doi.org/10.1080/00207543.2019.1566661

[13] Matthias Eckhart and Andreas Ekelhart. 2018. Towards Security-Aware Virtual Environments for Digital Twins. In *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security - CPSS '18*, ACM Press, Incheon, Republic of Korea, 61–72. DOI:https://doi.org/10.1145/3198458.3198464

[14] Matthias Eckhart and Andreas Ekelhart. 2018. A Specification-based State Replication Approach for Digital Twins. In *Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy - CPS-SPC '18*, ACM Press, Toronto, Canada, 36–47. DOI:https://doi.org/10.1145/3264888.3264892

[15] Matthias Eckhart, Andreas Ekelhart, and Edgar Weippl. 2019. Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 1222–1225. DOI:https://doi.org/10.1109/ETFA.2019.8869197

[16] Aniruddha Gaikwad, Reza Yavari, Mohammad Montazeri, Kevin Cole, Linkan Bian, and Prahalada Rao. 2019. Toward the Digital Twin of Additive Manufacturing – Integrating Thermal Simulations, Sensing, and Analytics to Detect Process Faults. *IISE Transactions* 0, ja (December 2019), 1–22. DOI:https://doi.org/10.1080/24725854.2019.1701753

[17] Vahid Garousi, Michael Felderer, and Mika V. Mäntylä. 2019. Guidelines for including grey literature and conducting multivocal literature reviews in

software engineering. *Information and Software Technology* 106, (February 2019), 101–121. DOI:https://doi.org/10.1016/j.infsof.2018.09.006

[18] Michael Grieves. 2011. *Virtually perfect: Driving innovative and lean products through product lifecycle management*. Space Coast Press.

[19] Michael Grieves and John Vickers. 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. . 85–113. DOI:https://doi.org/10.1007/978-3-319-38756-7_4

[20] Christos Koulamas and Athanasios Kalogeras. 2018. Cyber-Physical Systems and Digital Twins in the Industrial Internet of Things [Cyber-Physical Systems]. *Computer* 51, 11 (November 2018), 95–98. DOI:https://doi.org/10.1109/MC.2018.2876181

[21] André Kummerow, Dennis Rösch, Cristian Monsalve, Steffen Nicolai, Peter Bretschneider, Christoph Brosinsky, and Dirk Westermann. 2019. Challenges and opportunities for phasor data based event detection in transmission control centers under cyber security constraints. In *2019 IEEE Milan PowerTech*, 1–6. DOI:https://doi.org/10.1109/PTC.2019.8810711

[22] Heikki Laaki, Yoan Miche, and Kari Tammi. 2019. Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery. *IEEE Access* 7, (2019), 20325–20336. DOI:https://doi.org/10.1109/ACCESS.2019.2897018

[23] Chao Liu, Hrishikesh Vengayil, Yuqian Lu, and Xun Xu. 2019. A Cyber-Physical Machine Tools Platform using OPC UA and MTConnect. *Journal of Manufacturing Systems* 51, (April 2019), 61–74. DOI:https://doi.org/10.1016/j.jmsy.2019.04.006

[24] Francesco Longo, Letizia Nicoletti, and Antonio Padovano. 2019. Ubiquitous knowledge empowers the Smart Factory: The impacts of a Service-oriented Digital Twin on enterprises' performance. *Annual Reviews in Control* 47, (January 2019), 221–236. DOI:https://doi.org/10.1016/j.arcontrol.2019.01.001

[25] Xinxin Lou, Yun Guo, Yuan Gao, Karl Waedt, and Mithil Parekh. An idea of using Digital Twin to perform the functional safety and cybersecurity analysis. 12.

[26] Hongfang Lu, Lijun Guo, Mohammadamin Azimi, and Kun Huang. 2019. Oil and Gas 4.0 era: A systematic review and outlook. *Computers in Industry* 111, (October 2019), 68–90. DOI:https://doi.org/10.1016/j.compind.2019.06.007

[27] Yuqian Lu, Chao Liu, Kevin I-Kai Wang, Huiyue Huang, and Xun Xu. 2020. Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing* 61, (February 2020), 101837. DOI:https://doi.org/10.1016/j.rcim.2019.101837

[28] Edward O'Dwyer, Indranil Pan, Salvador Acha, and Nilay Shah. 2019. Smart energy systems for sustainable smart cities: Current developments, trends and future directions. *Applied Energy* 237, (March 2019), 581–597. DOI:https://doi.org/10.1016/j.apenergy.2019.01.024

[29] Abhishek Pokhrel. 2019. *BUSINESS PROJECT : DIGITAL TWIN APPROACH FOR CYBERSECURITY FOR FAULT PREDICTION*. Østfold University College, Halden, Norway. Retrieved from Available on Request.

[30] Qinglin Qi, Fei Tao, Tianliang Hu, Nabil Anwer, Ang Liu, Yongli Wei, Lihui Wang, and A. Y. C. Nee. 2019. Enabling technologies and tools for digital twin. *Journal of Manufacturing Systems* (October 2019). DOI:https://doi.org/10.1016/j.jmsy.2019.10.001

[31] Qianzhe Qiao, Jinjiang Wang, Lunkuan Ye, and Robert X. Gao. 2019. Digital Twin for Machining Tool Condition Prediction. *Procedia CIRP* 81, (January 2019), 1388–1393. DOI:https://doi.org/10.1016/j.procir.2019.04.049

[32] Adil Rasheed, Omer San, and Trond Kvamsdal. 2019. Digital Twin: Values, Challenges and Enablers. *arXiv:1910.01719 [eess]* (October 2019). Retrieved January 16, 2020 from http://arxiv.org/abs/1910.01719

[33] A. J. H. Redelinghuys, A. H. Basson, and K. Kruger. 2019. A six-layer architecture for the digital twin: a manufacturing case study implementation. *J Intell Manuf* (December 2019). DOI:https://doi.org/10.1007/s10845-019-01516-6

[34] Blaž Rodič. 2017. Industry 4.0 and the New Simulation Modelling Paradigm. *Organizacija* 50, 3 (August 2017), 193–207. DOI:https://doi.org/10.1515/orga-2017-0017

[35] Roland Rosen, Georg von Wichert, George Lo, and Kurt D. Bettenhausen. 2015. About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. *IFAC-PapersOnLine* 48, 3 (January 2015), 567–572. DOI:https://doi.org/10.1016/j.ifacol.2015.06.141

[36] Fei Tao, Qinglin Qi, Lihui Wang, and A. Y. C. Nee. 2019. Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering* (May 2019). DOI:https://doi.org/10.1016/j.eng.2019.01.014

[37] Eric J. Tuegel, Anthony R. Ingraffea, Thomas G. Eason, and S. Michael Spottswood. 2011. Reengineering Aircraft Structural Life Prediction Using a Digital Twin. *International Journal of Aerospace Engineering*. DOI:https://doi.org/10.1155/2011/154798

[38] Yu Zheng, Sen Yang, and Huanchong Cheng. 2019. An application framework of digital twin and its case study. *J Ambient Intell Human Comput* 10, 3 (March 2019), 1141–1153. DOI:https://doi.org/10.1007/s12652-018-0911-3

[39] 2015. Damaging North sea collision averted. *BBC News*. Retrieved August 27, 2019 from https://www.bbc.com/news/world-europe-35204070

[40] 2016. Digitalization, digital twins and cyber-physical systems...what's in it for us? *DNV GL - Software*. Retrieved July 9, 2019 from https://blogs.dnvgl.com/software/2016/04/digital-twins-structural-engineering/

[41] 501321main_TA11-MSITP-DRAFT-Nov2010-A1.pdf. Retrieved April 23, 2019 from https://www.nasa.gov/pdf/501321main_TA11-MSITP-DRAFT-Nov2010-A1.pdf

[42] *Genius of Things: An End to End Approach to IoT with Chris O'Connor*. Retrieved February 1, 2020 from https://www.youtube.com/watch?v=gUCCnVXgYvw

[43] Your systems may be optimized but digital twins could learn to do it better. Retrieved January 17, 2020 from https://new.abb.com/news/detail/24663/your-systems-may-be-optimized-but-digital-twins-could-learn-to-do-it-better

[44] A Digital Twin Approach to Predictive Maintenance. *InformationWeek*. Retrieved July 8, 2019 from https://www.informationweek.com/big-data/ai-machine-learning/a-digital-twin-approach-to-predictive-maintenance/a/d-id/1333331

[45] *Philips Digital Twin concept*. Retrieved March 11, 2020 from https://www.youtube.com/watch?v=H6JzPCbyVSM

[46] 01-80 ABB Review 2/2019. Retrieved January 17, 2020 from http://search.abb.com/library/Download.aspx?DocumentID=9AKK107492A3437&LanguageCode=en&DocumentPartId=&Action=Launch