Augmenting Cyber Physical Systems through Data Collection & Machine Learning: A Perspective

Haritha Deepthi[1*], Siddiq Moideen[2*], Jacquline Dorothy[3†]

[1*]Department of Information Technology, PSG Polytechnic College, Coimbatore, 641004, Tamil Nadu, India
harithadeepthibtech@gmail.com

[2*]Department of Computer Engineer, PSG Polytechnic College, Coimbatore, 641004, Tamil Nadu, India
siddiqmoideen07@gmail.com

[3†]Department of Basic Science, PSG Polytechnic College, Coimbatore, 641004, Tamil Nadu, India
jacqulinedorothy@gmail.com

Abstract:
Industry 4.0, destined to an astounding breakthrough in the field of Production & Manufacturing through leading-edge Cyber Physical Systems, Monitoring systems & Automation. The next big Industrial Revolution focuses on digitalizing the industries which is popularly known as Digital Twinning, any changes to the Digital Twin reflects the Real Physical World. Cyber Physical Systems are the key players in the 4th industrial Revolution, Cyber Physical Systems are amalgamation of Theory of Cybernetics & Mechatronics where Physical Plant i.e., Full-Fledged Hardware component Controlled/Monitored by Computational platform i.e., Computer-Based Algorithms moderated by Network Fabrics & Sensors. CPS integrates the dynamics of physical processes, software & networking. Where components of Physical and Computational elements are deeply intertwined. The Backdoors of the system aren't robust enough to tackle modern-day Cyber Threats. Digital Twinning gives us an upper-hand in both Security and Production perspectives. Our paper focus on enhancing the production & security of the CPS through Machine learning approach, analysing the digital asserts statistically to set a favourable pay.

Keywords: Industry 4.0, Cyber Physical Systems, Digital Twin, Game Theory, Machine Learning, Simulated Systems, Security Mechanisms, Graph Database

I. INTRODUCTION TO CPS

Cyber Physical Systems are multidisciplinary systems that conduct feedback control on widely distributed embedded computing systems through combination of communication, computation and control technologies. Modern CPS are able to realize the real-time, dynamic, safe and reliable collaboration with physical systems represented via embedded system. They are Integral mixture of existing network systems and traditional tightly coupled embedded computing systems comprising a large number of physical systems composed of intelligent wired/wireless actuators & sensors. Where, Physical system data modules collect data by distributed field devices in CPS system, then pass data to the information processing layer as per the complete given tasks and demands of services by information processing technologies. The potential benefits of the convergence of 3C technologies for developing next-generation engineered systems are wide ranging and highly transformative via efficient computation, distributed sensing, high-level decision-making algorithms, control over wireless/wired communication networks formal verification technologies and multi-objective optimization; engineered cyber physical systems are implied in branches of mechatronics, biology, computer science and chemistry which are integrated into many societal critical domains such as construction, energy, transportation, and medical systems. Physical & Technical systems are developed and designed to be more & more reliable, efficient, smart, robust and secure.

With such high notions, the scope of CPS and integration of Cloud computing are about to bring the next big Industrial Revolution, Industry 4.0. The complete factory can be made into digital twin in the cyber space, any changes in the cyber twin will reflect in the physical world which could boost the production & efficiency in the field of avionics, distributed robotics, energy conservation, process control, smart structures, defence systems, critical infrastructure control, assisted living, environmental control, medical systems, manufacturing and traffic safety & control. This is the core design of the CPS, making affects in both physical and cyber worlds in synchronisation.

II. INDUSTRY 4.0

Industry 4.0, the next stage in the organization and control of the entire industrial value stream along the lifecycle of a product based on dynamic, self-organized, cross-organizational, real-time optimized value networks, which can be optimized according to criteria such as availability, costs and consumption of resources. Thus, solving the existing problems in social infrastructure, safety, security, resource efficiency, standardization, work organization, work design, training and regulatory framework. Industry 4.0 is a collective term for technologies and concepts of value chain organization.

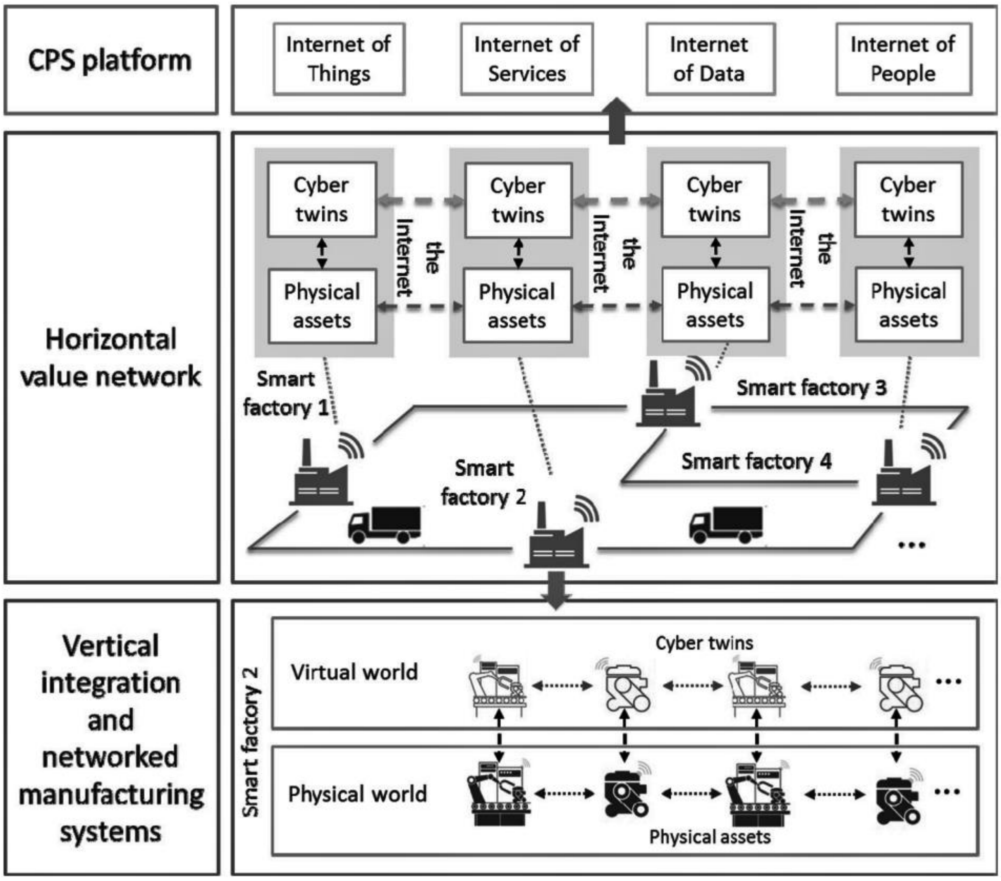


Figure 1: Industry 4.0 Principles

Within the modular structured smart factories of Industry 4.0, CPS monitor physical processes, create a virtual copy of the physical world and make decentralized decisions. Over the IoT, CPS communicate and cooperate with each other and humans in real-time. Via the IoS, both internal and cross organizational services are offered and utilized by participants of the value chain. In a smart factory, each and every entities in the system are equipped with sensors, controlled by software and connected to the Internet communication with each other therefore forming a seamless interconnected co-operative production system in real-time where all the physical production elements in the physical world have a cyber twin in the virtual world destined to achieve the global optimization of production. Industry 4.0 will give rise to novel CPS platforms geared toward supporting collaborative industrial business processes and the associated business networks.

CPS provide a critical support to the horizontal and vertical network manufacturing system integration. The widespread application of CPS gives rise to the generation of industrial big data, which requires big data analytics and cloud technology for analysis and storage.

III. CLOUD MANUFACTURING

Cloud Manufacturing is a new networked manufacturing paradigm that establishing the cyber end of the Industry 4.0 by organizes manufacturing resources over networks according to the client's requirements in order to provide on-demand manufacturing services via networks & cloud manufacturing service platforms thus, enabling ubiquitous, convenient, on-demand network access to a shared pool of virtual configurable manufacturing resources that can be easily upgraded in need. In the cloud manufacturing mode, providers can supply their requirements to the platform for requesting services to manufacturing end for all kinds product life-cycle stages such as product design, management, manufacturing and testing, which will be transformed into services and then pooled into the cloud manufacturing platform. This is an advanced manufacturing business approach that focuses directly on the core manufacturing issues and paying less attention to issues like demographic change and urban production.

Cloud manufacturing relies on Process and Digital Twin Modelling Solutions with modern developments of communication protocols and connection architectures The implementation of different layers in cloud manufacturing requires different technologies such as IoT, H2M interface, virtualization, cloud computing, semantic web and servitization technologies are needed to analyse the manufacturing resources and transform the physical resources into virtual resources.

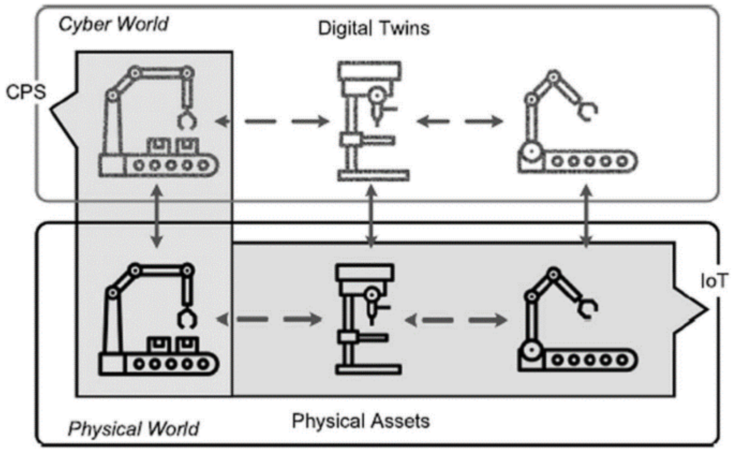


Figure 2: Digital Cloning via CPS implications

These Digital Twins of the physical world are developed and modelled to attain different extents that mimics the physical systems in real-time, they tackle the problems via an open-loop monitoring systems in-times even powered with AI-feedback systems for complex operation in the manufacturing. Digital twins adapt to wide variety of environmental scenarios & predicts different areas of production process. The concept of self-control in a decentralized production system is based on the ability of several elements of the system to act and decide autonomously.

A. Simulated Production System

A simulation is used to acquire production data of a virtual factory, all of the factory's components such as machine tools, resources and the product workpieces are replicated into a cyberworld. Each product has a product type and each machine has a machine type, so there are different types of virtual products and machines. The work plan, i.e., mimicking the order of operations required to produce a final product in the physical ecosystem. Thus, the product type defines which operations need to be processed sequentially in order to finish the product, while the machine type defines the operation the machine is capable to perform an operation with specific requirements is encoded into the machine type.

Many Digital Twin methods integrate models of physical systems with information and result from both computer simulations and CPS connected data. The concept by summarizing the technology needed for a complete Digital Twin to predict system's production and behaviour such as manufacturing time, device microstructure & architecture, part defects and integration of temperature and velocity fields from both experimental measurement and numerical simulation.

In simulation production systems, we have an upper hand of fast forward the time as per the recorded sequential growth of the ecosystem in time in integration with ML prediction models. Thereby, giving out an exceptional futuristic statistical plot points for the ecosystem, paving the way for improvements and augmentation.

B. Security Mechanisms

The major security objectives are Confidentiality, Integrity, Availability and Accountability of the ecosystem; Smarter the ecosystem is, times the vulnerable it is. Security mechanisms are classified into the following four categories: asset-centric security, network centric security, data-centric security, and user-centric security. The process of combining security mechanisms poses a complex challenge as it requires a proper combination of security mechanisms at every level as the system will further integrates into multiple independent layers while achieving multi-objective security might which might also leave chance for undesired security gaps or redundancy that make the system vulnerable to attacks.

a) Asset-centric security revolves around the device or asset, which are classified into hardware security and software security. *Hardware Security* focuses on establishing the validity of physical components. For traceability, liability reasons, counterfeit detection, typically through leveraging randomness that is extrinsic or intrinsic to the part which can be defended through Physical Marking/Watermarking and Crypto processors. *Software Security* solutions are aimed at protecting software against attacks such as tampering or causing run-time misbehaviour through invalid inputs that the software erroneously treats as legitimate, later exploits the existing software bugs and vulnerabilities; introduced during the software development cycle which can be defended through Antivirus Software, Tamper-resistant software and Digital watermarking.

b) Network-centric Security aims at providing secure communication among different assets in the network, a major facet for cloud-based design and manufacturing. The security mechanisms are hit through HTTPS, Secure Multicasting, Virtual private networks, Blockchains, Onion routing and Intrusion detection systems.

c) Data-centric security aims at protecting data throughout its lifetime starting from its creation, transmission, updating and storage which can be achieved through Steganography and Encryption model that bags up Symmetric Encryption, Asymmetric encryption, Secret sharing, Secure multiparty computation, Cryptographic hash functions and Digital certification techniques.

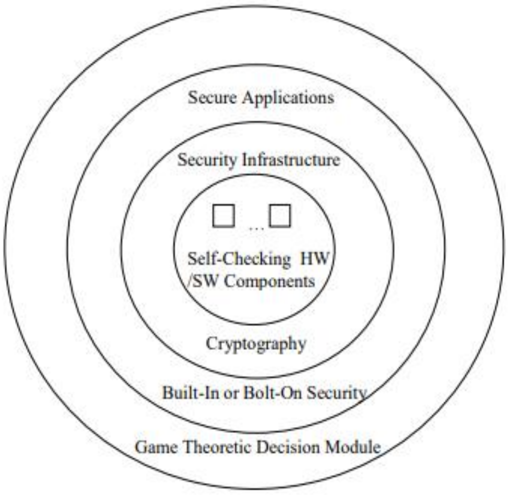

Secure Applications
Security Infrastructure
Self-Checking HW /SW Components
Cryptography
Built-In or Bolt-On Security
Game Theoretic Decision Module

Figure 3: Holistic System Design Approach

d) User-centric security revolves around advancements in sensor technologies and communication methods have made it possible to monitor product use in real time. Which can be achieved through Authorization, Authentication, Access control and Anonymity. such mechanisms help designers to come up with a security architecture that prevents an attack. A parallel branch, Forensic analysis includes damage estimation, identifying the method of attack and its source.

C. Data Collection

Virtual simulation approach is the most efficient way for data collection; the nodes, entities & products of the ecosystem records each & every measurement in every aspect. Since, the workflow of the system is pre-designed in order to mimic the physical world which adapts & synchronises itself with the help of actuators & sensors. This process bags up the M2M interaction, Resource consumption, M2H interaction, Product lifecycle, Machine temperature, Fault registry, Latency, Idle period, Machine health and System log data in the form of Comma Separated values. Later, mould into advanced structural models.

Adopting Graph Database(NoSQL) approach for Industrial Data Analytics in varied industry-related applications, network operations, object tracking & asset and data management due to its scalability, versatility, efficiency and flexibility. GDBs perfectly improve the data management, include path finding with weighted and time-related path properties, mapping dependencies of various system components to capture potential weak points, and communications between various networked elements. Conversely, query languages are used to extract data, including comparing node properties, traversing the database and subgraph matching.

The data model will consider two types of system entities: dynamic and static. The class of static entities are those entities that do not change in time which covers testbed setup profiles, network interfaces, testbed components and their settings. These entities are normally predetermined or collected in the initialization of each measurement. The class of dynamic entities captures various system events since they tend to change in time which covers network traffic, machine status reports and information flows in the testbed. These entities are dynamically added into the data set whose properties and quantities are determined by the measured data.

IV. INDUSTRIAL DATA ANALYTICS

Industrial data analytics plays an essential role in achieving the smart factory vision and improving decision-making in various industrial applications. The main facets of this domain are highly distributed data ingestion, management, repository, governance and analytics. Industrial analytical approach can avoid down-times, costly failures and tends to effective maintenance decisions, deployed for improving factory operations through improving machinery utilization and predicting production demands, improving product quality by analysing market demands and reducing defective products, and enhancing supply chain efficiency by analysing risk factors and making accurate logistic plans and schedules.

A. Visualization

Visualization Helps the us in building a mental map of the production systems considered an art of presenting data. A cognitive transition from one view to the others is a must, visualisation tends to mitigate a different perspective over the systems, latter influencing the decision-making process of the ecosystem through statistical comparison, interactions & workflows by means of data collected through the simulated systems in the cyberspace.

The overall goal is the optimization of the production process with respect to a diversity of parameters though the optimal solution is unknown to the user-end, visualization helps the clients to improve the factory ecosystem in wide range aspects. By that, users are enabled to approximate an optimal solution, thereby finding a sufficient solution to gain confidence in their production process.

B. Game Theory Implementation

Game Theory is primarily a mathematical framework analyses the decision-making of a player based on how they expect other players to make a decision i.e., determining optimal rational choices given a set of circumstances which can be applied in many fields such as economics, politics, computer science, biology, philosophy & so on where a game is considered be a set comprising of all possible input moves made by the defence and attacker end. Game theory depicts the game played between different players and the strategies of each player.

A game can be defined as interaction of different players according to a set of rules. Players may consist of individuals, machines, parties, companies or associations. The results of game theory depend upon behaviour of every other player present in the game and not only the current player.

Due to this reason, this approach is extremely scalable and versatile. The outcome of game theory also depends on the estimated payoff by each player before making decisions, which is a measure of the satisfaction obtained by each player by making that decision. Therefore, the players will perform actions and take decisions that would provide them the maximum payoff.

In game theory, there are different types of games that help us analyse different problems. They are categorised in the basis of number of players involved, cooperation among players & symmetry of the game.

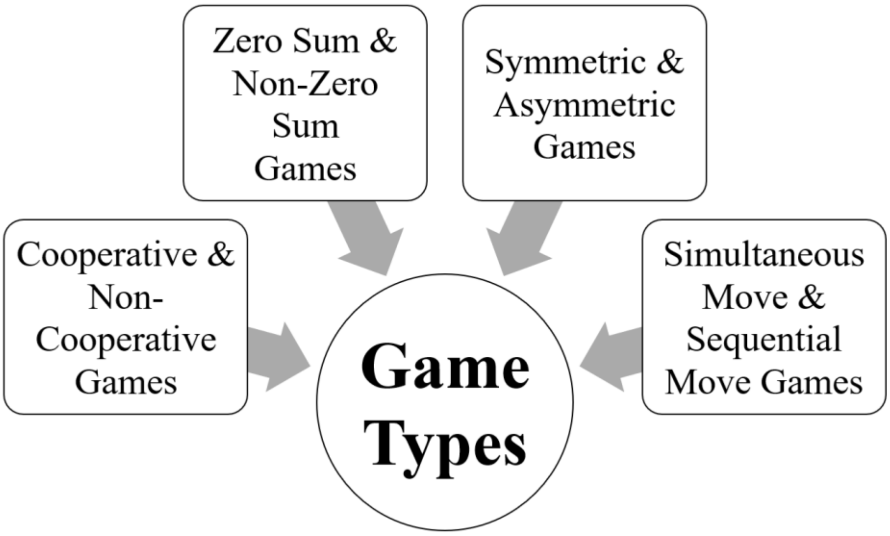
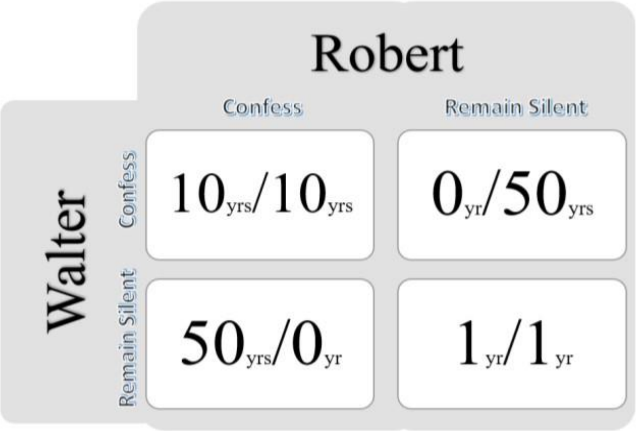

Zero Sum & Non-Zero Sum Games
Symmetric & Asymmetric Games
Cooperative & Non-Cooperative Games
Simultaneous Move & Sequential Move Games
Game Types

Figure 4: Game Theory Types

The prisoner's dilemma is one perfect example; how game is analysed in game theory which shows why two completely rational individuals might not cooperate, even if it appears that it is in their best interests to do so. Prisoner's dilemma is a situation where individual decision-makers always have an incentive to choose in a way that creates a less than optimal outcome for the individuals as a group.

Two members of a cartel named Robert & Walter were arrested and imprisoned. Each prisoner is in solitary confinement and they have no means of communicating with each other. The prosecutors lack sufficient evidence to convict the pair on the principal charge.



		Robert	
		Confess	Remain Silent
Walter	Confess	$10_{yrs}/10_{yrs}$	$0_{yr}/50_{yrs}$
	Remain Silent	$50_{yrs}/0_{yr}$	$1_{yr}/1_{yr}$

Figure 5: Decision Matrix

Prosecutors hope to get both of them sentenced on a lesser charge. Simultaneously, the prosecutors offer each prisoner a bargain. Each prisoner is given the opportunity either to betray the other by testifying the crimes committed by the other one or to cooperate with the other by remaining silent. The options offered are:

- If Robert & Walter, both betray each other i.e., if they both confess, each of them serves 10 years in prison.
- If Robert betrays Walter but Walter remains silent, Robert will be set free & Walter must serve 50 years in prison.
- If Walter betrays Robert but Walter remains silent, Walter will be set free & Robert must serve 50 years in prison.
- If Robert & Walter both remain silent, both of them will only serve 1 year in prison.

Note that it is implied that the prisoners will have no opportunity to reward or punish their partner other than the prison sentences they get and that their decision will not affect their reputation in the future.

We'd feel both remaining silent would be the best option. But the individuals won't opt for it; Both of them will betray each other i.e., they would confess. Because that's the human psychology, confession seems to the best option for both the parties. Individually the prisoner clings on luck that he would be set free, if the other prisoner does not confess & fears that what if I remain silent and the other confesses. Because betraying a partner offers a greater reward than cooperating with them, all purely rational self-interested prisoners will betray the other, meaning the only possible outcome for two purely rational prisoners is for them to betray each other and that is Nash Equilibrium A.K.A. Optimal state for all the participants. The prisoner's dilemma game can be used as a model for many real-world situations involving the cooperative behaviour.

With respective to the branch chosen in game theory, the Nash equilibrium (i.e., the optimal vulnerable point where the attacker aims at) in the CPS is found, Later the system is defended via appropriate measures, by integrating the ML techniques; the backdoors of the system will be unbreakable.

C. Machine Learning Approach

Machine learning is a branch of Artificial Intelligence. A system, rather than explicitly following the instructions fed to it; The system moulds itself to accomplish the set target through Supervised, Unsupervised and Reinforcement training that sets path to computer vision, .advanced security, prediction & assisted decision-making models which can be engineered into the systems.

Supervised Learning leans on the regression & classification models trained through labelled data inputs which bags up Gradient Descent, Linear Regression, Naive Bayes, Support Vector Machine, Decision Tree and Random Forest techniques.

Unsupervised Learning leans on the association & clustering models trained through unlabelled data inputs which bags up K-means clustering, k-nearest neighbours, Hierarchal clustering, Artificial Neural Networks, Principal Component Analysis and Singular value decomposition techniques.

Reinforcement learning leans on exploitation & exploration models which adapts & interacts with the environment to take decisions in order to maximize reward in a given situation through undefined data inputs.

Integrating AI into ecosystem is a pivoting factor in Industry 4.0, each and every path laid were to attain this particular stage of indulging intelligence to the ecosystem to make self-controlled/decentralised decisions. The path laid, journey of amalgamating 3C technologies, integrating embedded computing and cloud manufacturing into order to populate the invergence of Cyber-Physical systems & digital twinning; thereby, making it easier for data collection via simulation models in real-time which can be used to train the model through several machine learning algorithms implied in wide range of possibilities.

AI system can be enforced for decision-making, vulnerability check, industrial upgradation, optimal workflow, scope for improvements and optimal power/resource consumption by the analytical & predictive report generated. Combination of Cyber-Physical Manufacturing environment, Simulation systems, Game Theory and Machine Learning models; turns out to have an omnipotent impact over the upcoming years.

V. CONCLUSIONS

With high notions of Smart physical & simulated factories, the heights the production industries are about to reach are ‘out of the world’. Virtual factory in the cyberspace gives us an upper-hand over the data collection process, which can be further beautifully laid for analysis. By deploying machine learning algorithms, trained via the collected dataset to give out an accurate predictive model to optimize the production & development of the system. Mixture of Game Theory & ML algorithms boils a robust defence system, securing the backdoors & vulnerable weak points of CPS. Such approach towards Industrialisation breaks the fabric of present production ecosystem to a more resilient & mature factories.

COMPLIANCE WITH ETHICAL STANDARDS

Funding: This study does not involve any funding.

Conflict of Interest: There is no conflict of interest.

Ethical approval: This article does not contain any studies with animals or the human participants performed by any of the authors.

REFERENCES

Hausi A. Müller, “The Rise of Intelligent Cyber-Physical Systems”, *IEEE*, https://doi.org/10.1109/MC.2017.4451221, 18 December 2017.

Mahmoud Parto, Pedro Daniel Urbina Coronado, Christopher Saldana and Thomas Kurfess, “Cyber-Physical System Implementation for Manufacturing with Analytics in the Cloud Layer”, *ASME*, https://doi.org/10.1115/1.4051663, 14 July 2021.

Zhijia You and Lingjun Feng, "Integration of Industry 4.0 Related Technologies in Construction Industry: A Framework of Cyber-Physical System", *IEEE*, https://doi.org/10.1109/ACCESS.2020.3007206, 6 July 2020.

Yongkui Liu and Xun Xu, “Industry 4.0 and Cloud Manufacturing: A Comparative Analysis”, *ASME*, MANU-16-1445, https://doi.org/10.1115/1.4034667, 6 October 2016.

Kyoung-Dae Kim and P. R. Kumar,” Cyber–Physical Systems: A Perspective at the Centennial”, *IEEE*, https://doi.org/10.1109/JPROC.2012.2189792, 3 April 2012

Fernando Matsunaga, Vitor Zytkowski, Pablo Valle and Fernando Deschamps “Optimization of Energy Efficiency in Smart Manufacturing Through the Application of Cyber–Physical Systems and Industry 4.0 Technologies”, *ASME*, https://doi.org/10.1115/1.4053868, October 2022.

Siva Chaitanya Chaduvula, Adam Dachowicz, Mikhail J. Atallah and Jitesh H. Panchal, "Security in Cyber-Enabled Design and Manufacturing: A Survey", *ASME*, https://doi.org/10.1115/1.4040341, December 2018.

Tobias Post, Rebecca Ilsen, Bernd Hamann, Hans Hagen and Jan C. Aurich, "User-Guided Visual Analysis of Cyber-Physical Production Systems", *ASME*, https://doi.org/10.1115/1.4034872, June 2017.

Mikhail V. Chester and Braden R. Allenby, "Perspective: The Cyber Frontier and Infrastructure", IEEE, https://doi.org/10.1109/ACCESS.2020.2971960, 05 February 2020.

Vishruti Kakkad, Hitarth Shah, Reema Patel and NishantDoshi, "A Comparative study of applications of Game Theory in Cyber Security and Cloud Computing", ScienceDirect, https://doi.org/10.1016/j.procs.2019.08.097, August 2019.

Manu Suvarna, Ken Shaun Yap, Wentao Yang, Jun Li,Yen Ting Ng and Xiaonan Wang, "Cyber–Physical Production Systems for Data-Driven, Decentralized, and Secure Manufacturing - A Perspective", ScienceDirect, https://doi.org/10.1016/j.eng.2021.04.021, September 2021.