Cyber Physical Systems and Game Theory Integration

Haritha Deepthi and Siddiq Moideen

Department of Information Technology & Department of Computer Engineering, PSG Polytechnic College, Coimbatore, Tamil Nadu, India
harithadeepthibtech@gmail.com; siddiqmoideen07@gmail.com

Abstract: Cyber Physical Systems are Multidisciplinary Engineered Embedded Systems that cling to Computation, Communication & Control (3C) Technologies to integrate Full-Fledged Physical Systems and Control & Computational Resources/Elements through Cyberspace that Control & Alter Physical Environments. Thereby, Exposing the ecosystem to Security Threats. An Attacker, who roots through the control system can remain undetected via masking certain measurement signals & renders a portion of the system unobservable, which is an Observability Attack. Such Observability attacks can be analyzed through the Game-Theoretical approach. Where, Attacker's Strategy Set bags up all viable Masked Measurements. Defender's Strategy Set includes all viable columns to defend the system, which are further Quantified & Analysed. Eventually, Multiple Nash Equilibria are identified; Thereupon, an Optimal Strategy is used to retain the ecosystem from a Cyber-Attack.

Keywords: Cyber Physical Systems, Hybrid Systems, Open System Interconnection, Real-time Systems, Industry 4.0, Cyber Twin, Heterogeneity, Parallelism, Nash Equilibrium, Prisoner's Dilemma, Strategic form, Stochastic Game, Value of state, Quality of state, Minimax Q-Learning, Learning rate, Exploration rate, Naïve Q-Learning, Optimal policy

INTRODUCTION TO CPS

Cyber Physical System, is a multidisciplinary system that conducts feedback control over widely distributed embedded computing systems in a mixture of Control, Computation and Communication technologies, an integral mixture of existing traditional embedded and network systems. Physical system data modules collect data through distributed field devices in the CPS system, then pass data to the information processing layer as per the assigned tasks and demands of services such as statistical signal processing, feedback control, data security processing and data uncertainty management. The modern CPS are highly collaborative, dynamic and solid with the physical ecosystem. The potential benefits of the convergence of 3C technologies for developing next-generation engineered systems A.K.A. Cyber Physical Systems, are wide-ranging and highly transformative via efficient computation, distributed sensing, high-level decision-making algorithms, control over wireless/wired communication networks and multi-object optimization; engineered cyber physical systems are in many societal critical domains. Physical systems & Technical systems are developed and designed to be more & more reliable, efficient, smart, robust and secure.

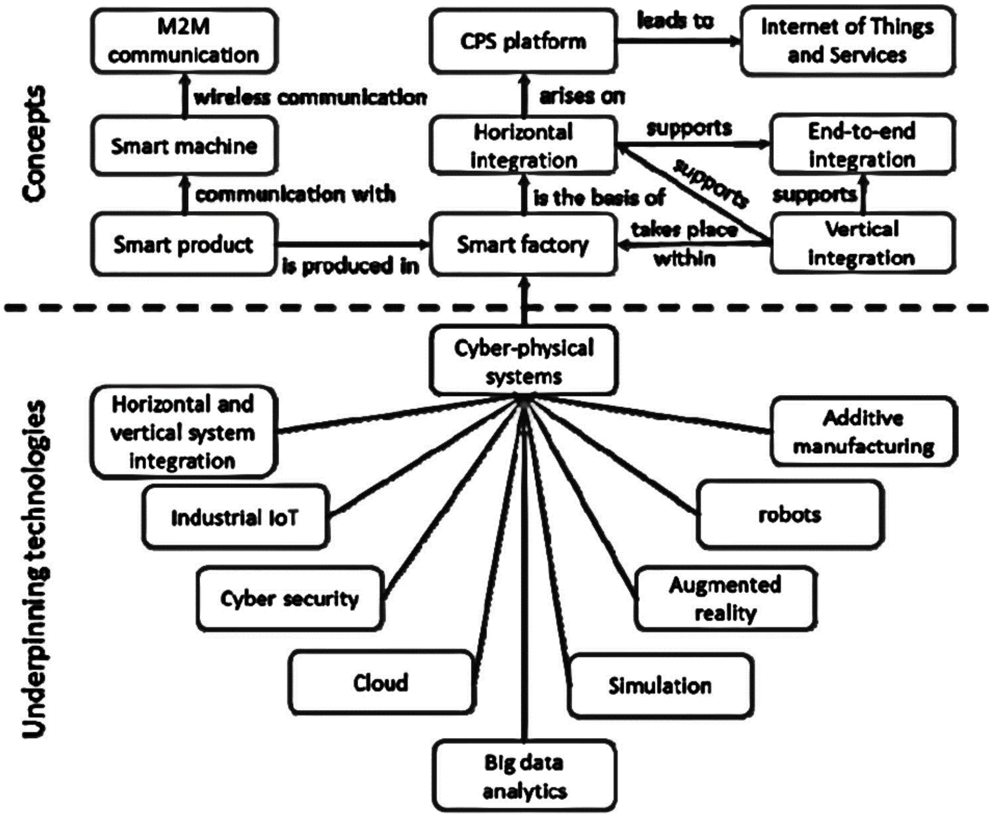


Figure 1: CPS Framework

CPSs cover a vast spectrum of applied areas, that enable more & more economic and technological drivers through shared design tools & abstract principles. Creating an atmosphere that opens a wide range of functional opportunities and possibilities. Advancements in CPSs bores through the leading edge engineered sectors, such as:

- Avionics
- Automotives

- Critical infrastructure control
- Distributed robotics
- Defence systems
- Environmental control
- Energy conservation
- Process control
- Manufacturing
- Med Tech
- Traffic safety & control
- Smart structures

With such high notions, the scope of CPS and integration of Cloud computing is about to bring the next big Industrial Revolution, Industry 4.0. The complete factory can be made into a digital twin in cyberspace. Any changes in the cyber twin will reflect in the physical world which could boost production & efficiency via varied implications.

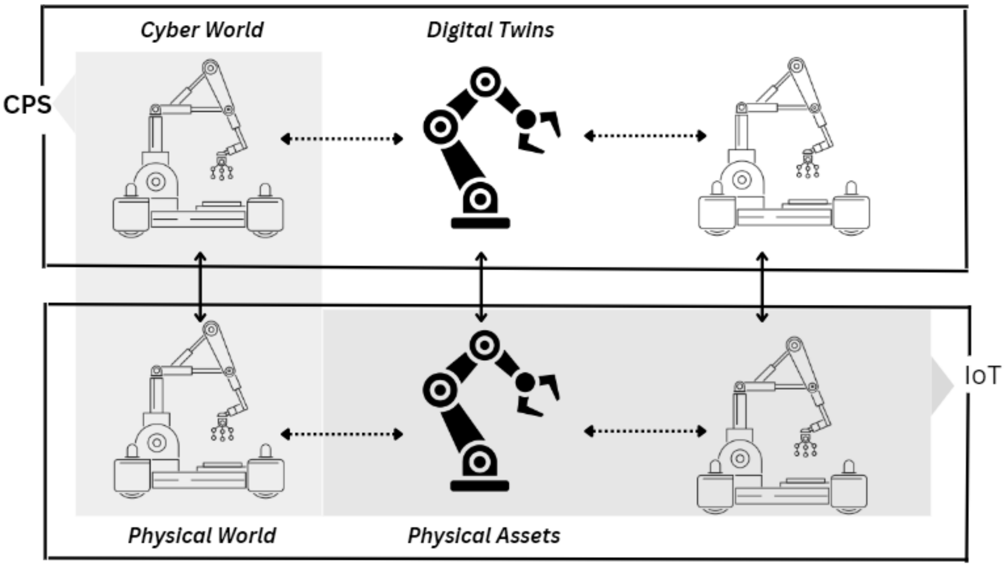


Figure 2: Digital Cloning via CPS implications

Wireline and wireless data networks were non-existent decades ago. Therefore, advancements in CPS's network technologies are destined for the third generation of control systems. There has also been enormous growth in the complexity of programming, hardware and software abstractions and there has also been an enormous evolution of distributed systems in the field of control systems.

There’s another aspect; the Internet of Things, where physical objects are assigned unique addresses and are interconnected with each other, targeted at connecting the entities over the Internet. All of it combined leads to another platform revolution. In such high time, it is necessary to re-examine both the policies as well as standard mechanisms. By mechanisms, regulating & standardizing the implementation of a system, while by policies, setting up a regulatory protocol. There are many challenges ahead that are to be addressed in the mere future; As technology evolves, Cyber threats evolve too.

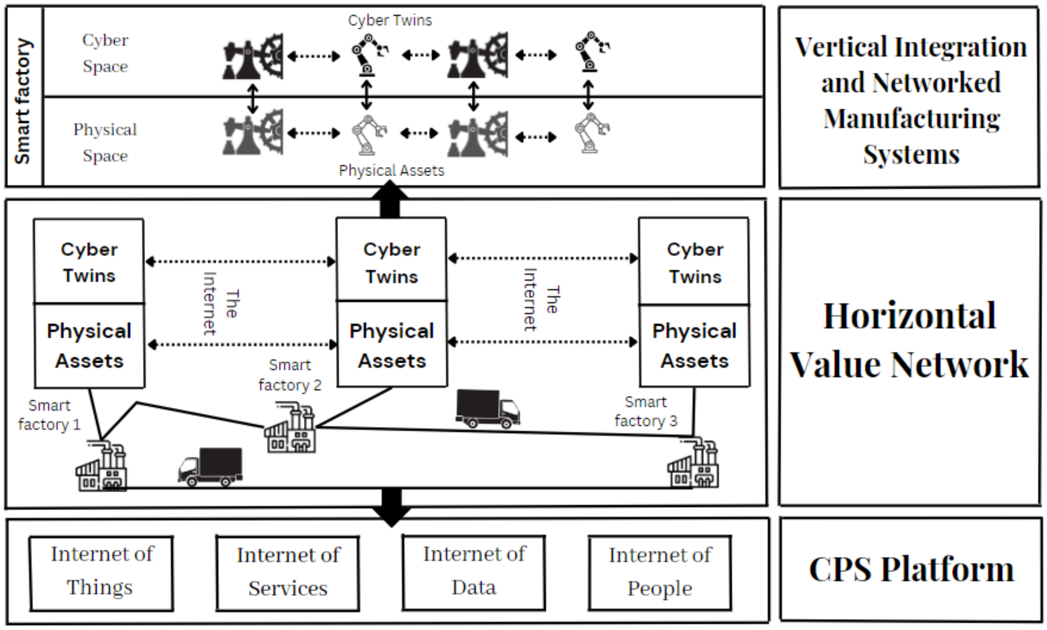


Figure 3: Industry 4.0 Principles

There can be a massive economic benefit that will bring forth integral changes to the existing engineered physical systems. The scientific research community defines Cyber-Physical System technology from *different perspectives*. Some describe CPS as the engineered physical systems that are designed to perform monitoring, controlling, which are patched up via Networking technologies & computational resources. The Cyber Physical System is deeply intertwined with the real-world elements that are to be connected across technologies at different paradigms, aimed to attain responses in real time.

Such descriptions at sight, the definition of CPS boils down to; "The Cyber-Physical Systems are next-generation engineered systems that are multi formational, structured through mathematical & computational abstractions that reflect physical actuation in a simulated environment"

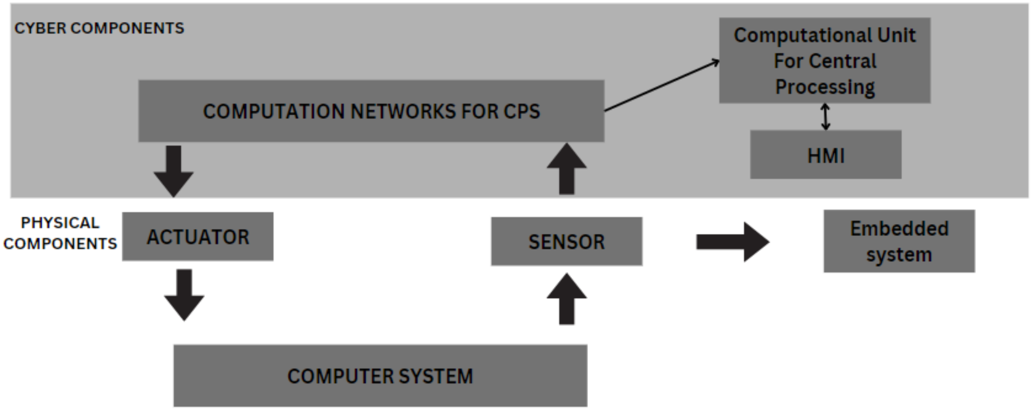


Figure 5: CPS Block Diagram

CPS ARCHITECTURE

Cyber Physical Systems compile information from real mechanical ecosystems & relay it through the telecommunication systems to the centralized decision-making module, which are processed & sent back to physical systems. The CPS architecture encapsulates the two major components: the Physical Layer & the Cyber Layer. A hybrid approach is a credible option for CPS, that takes into consideration of both discrete & continuous variable dynamics in synchronization.

In CPS, sensors actively monitor the differential changes respective to the standard set points which are known to be process variables, with a complex set of mathematical equations, control variables are obtained by the system controllers. In combination with a local or cloud decision-making module the right actuation is derived and new control variables are developed to perform the designated process. CPS's Human Machine Interface either through CLI or GUI provides an interactive platform for the system for the user end.

The general working mechanism of CPS bags: Monitoring, Networking, Computation and Actuation. The physical layer and cyber layer are channeled through industry protocols. A Cyber Physical System consists of multiple actuators and dynamic/static sensor networks integrated with the intelligent decision system.

Different types of Cyber-Physical Systems component integration are stationed based on their decisive connectivity & communication.

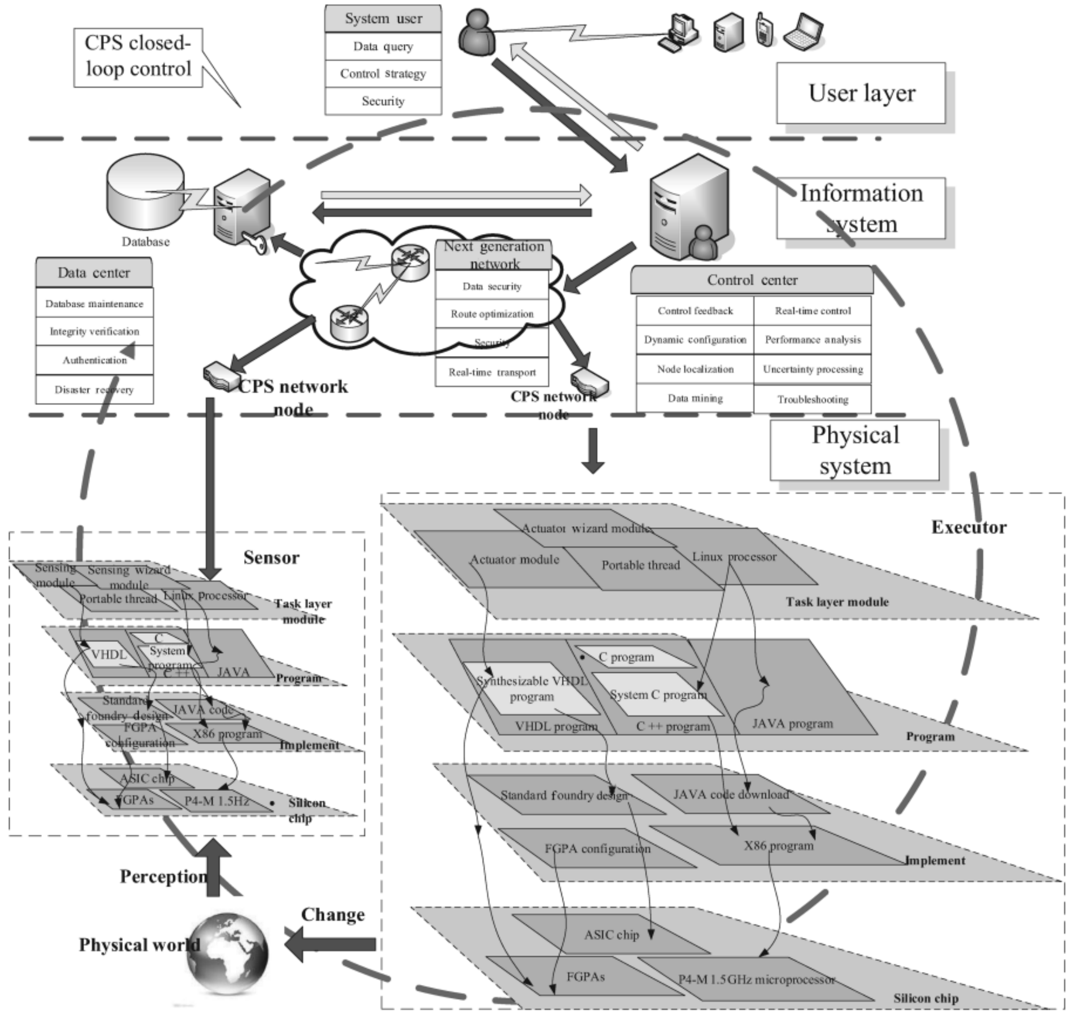


Figure 6: Typical CPS Architecture

Cyber physical systems are marked through their heterogeneous information flow, distinct elements in feedback systems and decision-making modules. The computational layer uses information & knowledge from physical environmental input that comprises a vast combination of key functions; in subject to the class of application, the right actuation is developed and are arrayed through process execution.

The key underlying architecture of Cyber Physical systems are partitioned into 7 intrinsic levels of the Open System Interconnection (OSI) model from the application layer to the physical layer.

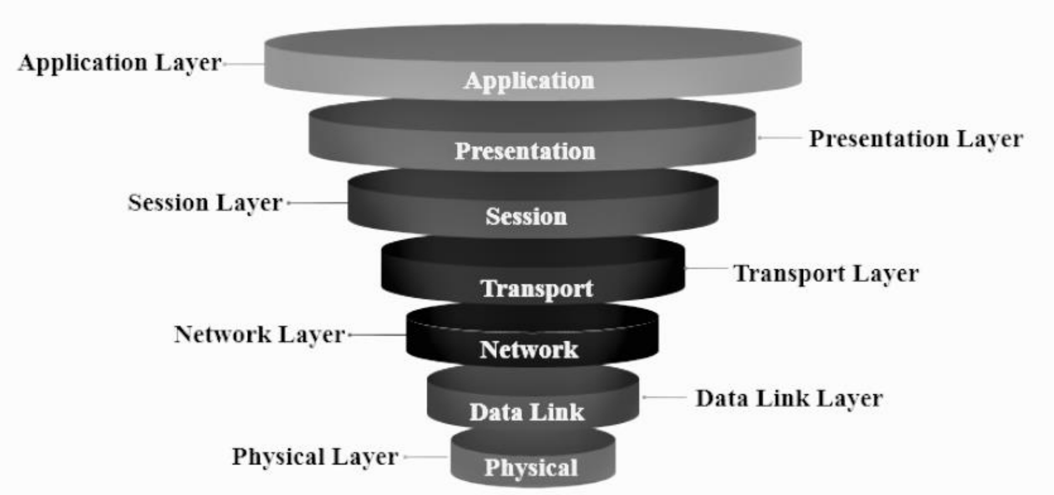


Figure 7: OSI Layers

Fundamental Architecture Levels

Physical layer

The physical layer of the OSI model lays the fundamental groundwork for the CPS architecture that comprises monitoring elementary components, sensors & actuators which receive the analog signals from the working physical environment and are transformed into digital signals for transmission. Attacks at this level are often via external physical forces.

Data link layer

The data-link layer achieves transmission of data through the physical layer onsets. That performs data framing & physical addressing, assigning MAC addresses to the nodes in the ecosystem. Regulated through Logical Channel Management & Media Access Control sublayers. Any attacks to this layer will rattle the node identification mechanism, leading to the failure of network bridging.

Network layer

The network layer transforms the data frames into data packets and are routed through IPv4/IPv6 and ICMP protocols that take up the role of path determination, converting MAC addresses to networking IP addresses i.e., logical addressing. Any attacks at this level, disrupt the data transmission path, leading to the elementary sensory failure of the system.

Transport layer

The transport layer uses User Datagram & Transmission Control protocols to break down the packets into smaller segments that ensure end-to-end connectivity and data reliability. Any attacks at this level directly affects the latency of the data transmission, thereby the real-time mark cannot be achieved.

Session layer

The session layer takes up the role of interhost communication, which plays an integral role in unifying the stacked layers, Ensuring the orderliness of the data transmitted; Handles exceptions by inserting labels into long messages until the session closes, such that data transmission need not start from the beginning. Any attacks at this level will affect the message’s integrity.

Presentation layer

The presentation level takes up the role of presenting the ascertain manner i.e., respective data formatting and data encryption in order to preserve information security through various protocols such as SSL and symmetric & asymmetric encryption algorithms. Any attacks at this layer will expose the plain text of the transferred message.

Application layer

The application is construed as the computational core space at the user end compiling the traversed data into a desired process or actuation, presented in a human-friendly context. Analysis, control & regulation can be made via the decision-making unit or the respective user in the space. Any attacks at this layer will obstruct the intrinsic working of the system, leading to colossal damage.

CPS CHARACTERISTICS & COMPLEXITIES

Physical System

Physical system, an important half of the Cyber Physical system, actuates and makes changes in the real world. A system that compiles the network, computation

& machinery hardware constituents that effectively meet the demands of the user with relevant spatial scales.

Uncoordinated Change

A technology that is in perennial development & research, the technology will not be stable enough to tackle modern problems and cyber threats over time. Advancements to the systems as in the novel standards are a must.

Pattern Abstraction

Existing The concept of concurrency & synchronization was not a problem back when writing instructions to a basic primal microprocessor architecture, as the complexity of the architectures evolved, factors such as computation, concurrency & scheduling posed a major problem when working with embedded systems which cannot be solved via the existing programming language until a certain set point.

Size & Computability

An industry-grade Cyber Physical systems are robust enough to tackle the major downsides of the ecosystem, bagging up a significant scale of computational resources, elementary sensory units, a decentralized decision-making system and its immersive longevity & connectives throughout the network and ecosystem.

Security by Design

Development of such newer technology, security was not always the prior concern, the attainment of desired task or target was a major concern. But the integration of both physical & cyber technologies tripled the chances of system failure through the security threats; tampering, counterfeiting, malicious injections & intrusions in various modes. Right from the branches of sensing, feedback, communication, computation to the resultant actuation at every possible level.

The existing cyber physical systems are reliable enough to solve modern industrial problems but not resilient enough to secure their resources & control via their built design. the security mechanisms such as assert-centric, data-centric, network-centric and user-centric security mechanisms post an underlying basis for CPS security.

There is a bright side in integrating the cyber aspect into a physical system, with the right level of blending and utilization, we can produce a superior defense system for the ecosystem. A retrospective approach would be the best for tackling the problems, recording & accounting for the threats and rolling out the testbed environments in the virtual space for trial & error implementation.

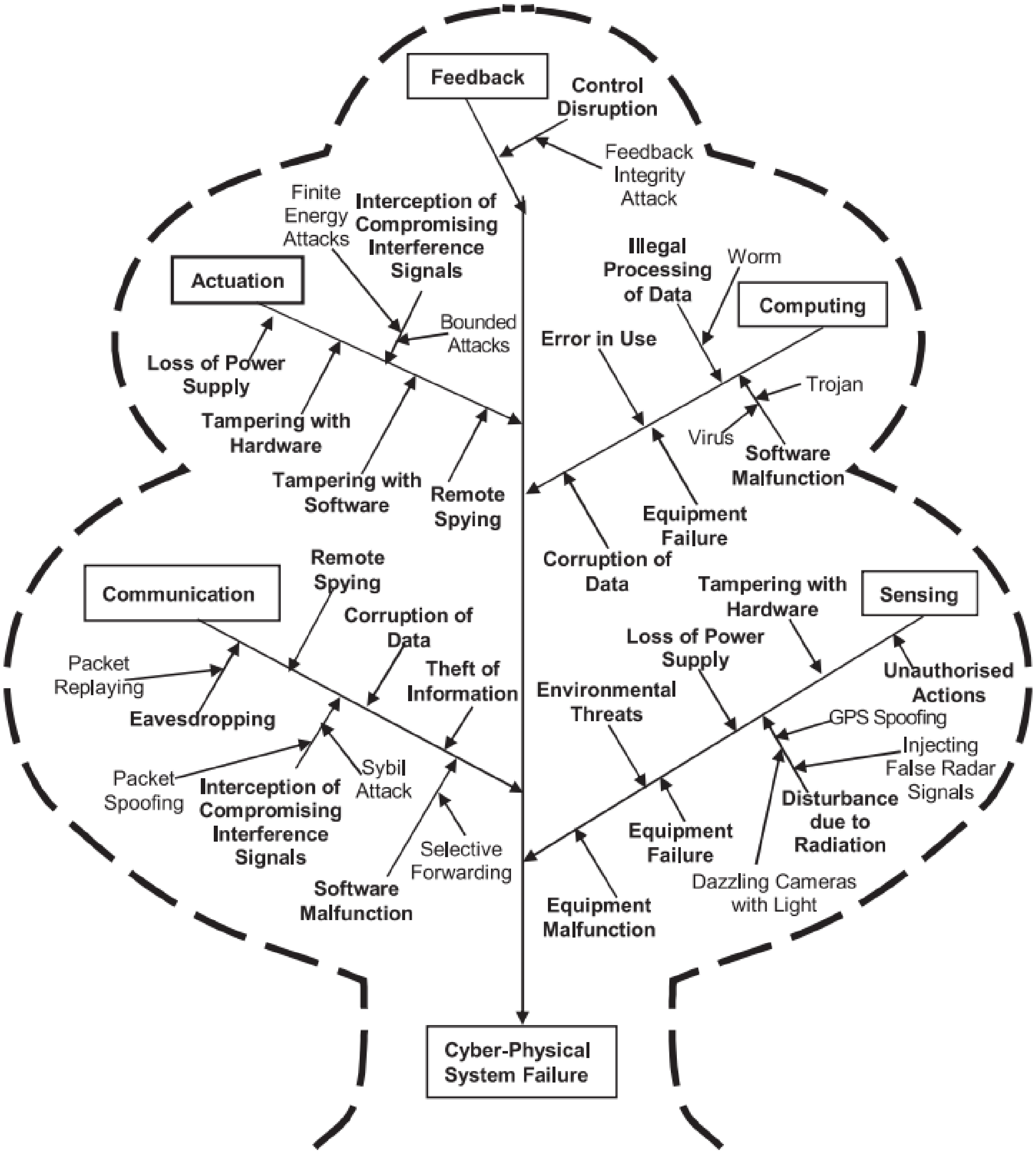


Figure 8: Cyber Physical Attacks

Notions of large-scale industrial implementation focus on data integrity, system reliability, data confidentiality, efficiency & high degree of redundancy in production systems, a packed unit that deals with all-sorts security threats is a must for an industry-grade production system.

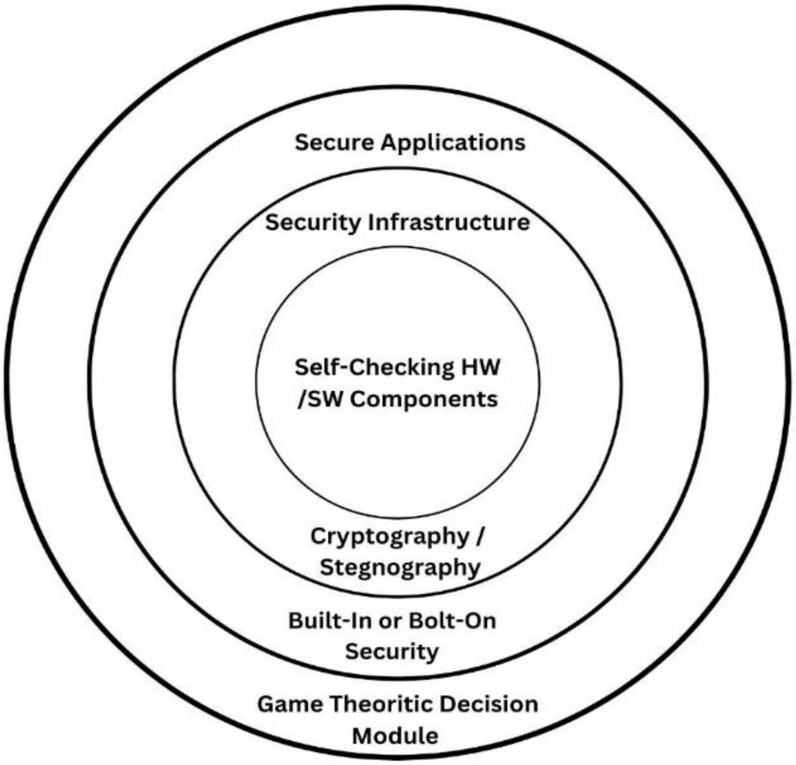

Secure Applications
Security Infrastructure
Self-Checking HW /SW Components
Cryptography / Stegnography
Built-In or Bolt-On Security
Game Theoritic Decision Module

Figure 9: Holistic System Design Approach

In order to attain such a structured system; the incorporation of elf-monitoring hardware & software components, the inclusion of cryptography & steganography for data integrity and confidentiality, and bolt-on security with the security system infrastructure layered with game theoretical decision-making module would be the best approach to build a CPS model.

Information System

The Information system plays a crucial role in accounting & recording the intrinsic facets of the system; Control variables, sensory outputs, response time, regulatory readings, feedback transmission, production system loggings and mechanical parameters. That underlies the basis for the digital twinning concept, the collected data are warehoused & mined for knowledge discovery and are tuned in such a way that the physical production environments are mimicked in the cyberspace. Can be controlled & actuated through the virtual end and reflected in the physical world.

Heterogeneity

Cyber Physical system represents highly distributed, complex, hybrid, real-time and closed-loop systems. Multiple dimensions of the system are to be accounted for as far as its characterization is concerned. CPS is a complete mixture of different technologies, different approaches & different behaviorisms. Each component supports the other to maintain its integral stability. The theory of networking, theory of computation, theory of control and theory of complexity make up a tough course to attain synchronization & parallelism in real-time. Thus, arising a major issue in the heterogeneous approach.

Real-Timelessness Nature

The one & only prior concerns in the Cyber Physical systems are real-timelessness nature, the inability of the system to sync with the physical world timeline but on its exponential virtual time-space. Most of the simulated systems often do not realize the parallelism considering its asynchronous framework. In the unity of the physical & cyber segments, it is hard to realize the environmental inputs and actuate at the desired pace through the existing programming languages with such number of abstractions.

Dynamics

Cyber Physical systems are deeply intertwined & tightly coupled system, each & every parameter are so dependent such that, any change to a single parameter measurement will affect the entire dynamics of the ecosystem triggering a chain reaction in elementary variable measurements in the system. That could even lead to a colossal damage to the ecosystem.

REAL-WORLD CYBER PHYSICAL THREATS

Cyber, Cyber-Physical & Physical attacks are enforced on the Cyber Physical systems; that influence the system at different levels.

Attacks are categorized respectively based on where the injuries are. Attacks that target & hit the system’s software & internal level and not the sensors & actuators are categorized into “Cyber Attacks”.

Attacks that physically target & hit the system’s components in the real world are categorized into “Physical Attacks”.

Attacks that target & hit the system’s physical layer which affects and alters the existing real world by means of cyberspace & masking techniques are categorized into “Cyber Physical Attacks”.

Real-World Instances

Smart Monitoring System Attack

User safety would be at risk if a wearable device, home-assist devices or any smart device’s security is breached. The attacker can actively monitor and record the moves of the victim via the breached system for their profit.

Industrial Control System Attacks

CPSs are applied in a wide arrange of fields, such as aviation, water treatment systems, nuclear plants, construction, sewage system, industries, etc. any considerable vulnerability in the system could lead to catastrophic failure.

Stuxnet-2010, attack on Iran's nuclear power plants caused damage to several nuclear plants which could have also been a Threat to the whole world as result, due to a Security breakthrough.

Smart Grid CPS Attack

Attackers can do cyberattacks on Smart grids and Blackout an entire state or a country. This make destabilizes the complete system & economy. The social life of the affected areas would crumble.

MedTech System attack

Centralized distributed systems are susceptive to insider attacks & spying. An intruder can pose threat to a victim by jamming the RF signals of the implant device that use to keep the victim in health and the attacker can damage or tamper with the medical device and make machines do what the attacker wants to.

Smart Vehicles Attacks

Modern car manufacturers are coming up with new technologies that revolve around self-driving capabilities, efficient resource utilization & remote access/control functionalities. CPS renders a large portion of current industry standard cars. If an attacker could break into the system, Vitim's life could be in danger.

GAME THEORY

Game Theory, primarily a mathematical framework that analyses the decision-making skill set 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 that can be applied to wide-ranging domains such as economics, politics, computer science, biology, philosophy & so on. Game theory is a field that examines the dynamics between players in a game and the strategies they use.

It looks at situations where multiple parties engage with one another following a specific set of regulations. These players could be individuals, groups, artificial entities, organizations, or businesses. The outcome of game theory is shaped by the choices made by each player and the rewards they anticipate, which reflects their level of contentment when making decisions. With this in mind, Players then make choices and take actions that result in the highest possible payoffs for themselves.

Game

A game is described as a trio of elements ($N, (A_i)_{i\in N}, (u_i)_{i\in N}$), where:

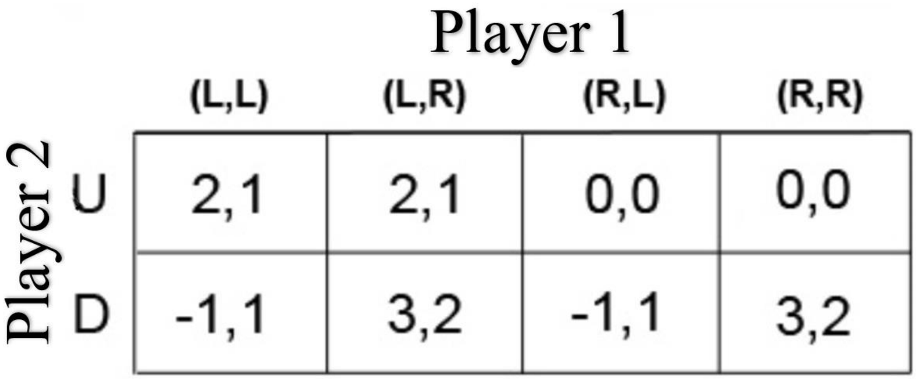
- N is a set of players represented by the numbers 1, 2, ..., n.
- A_i is the set of actions available to player i.
- A is the set of all possible action combinations, represented as $(a_i)_{i\in N}$ where each a_i belongs to A_i.
- u_i is the payoff function for player i, linking each action profile (a1, . . . , an) to a value in the set R. This value can either be a profit to maximize or a cost to minimize.

The action profiles can also be expressed as $(a_j\ )_{j\in N}$, where (a1, ..., an) equals (a_i, a_{-i}). Here, a_{-i} represents the action profile of all players except player i.

Normal and Extensive form Representation

The strategic\normal form is a model of game representation that presents players, strategies & playoffs in the game structured in a dynamic matrix. The players are assigned with respective matrix sides. Each player's strategies are fashioned through N cross N rows & columns.

Then the playoffs are plotted within the block spaces presenting every possible combination of actions. The games that are modeled in normal form are often played simultaneously or the players involved must have no idea about the strategic playoffs of the other player involved.



		Player 1			
		(L,L)	(L,R)	(R,L)	(R,R)
Player 2	U	2,1	2,1	0,0	0,0
	D	-1,1	3,2	-1,1	3,2

Figure 10: Normal Form

Where in the other case, if the player involved has some idea about the strategic playoffs of the player, the game is usually represented in the form of an extensive model. Presenting such a scenario in terms of normal form could lead to more & more complexities.

The extensive form is presented like a tree structure concerning time sequencing moves either simultaneously or in the game the of past, the players are classified as per the first parent-to-child vortex points and the bottom of the tree represents the chosen playoffs of the player.

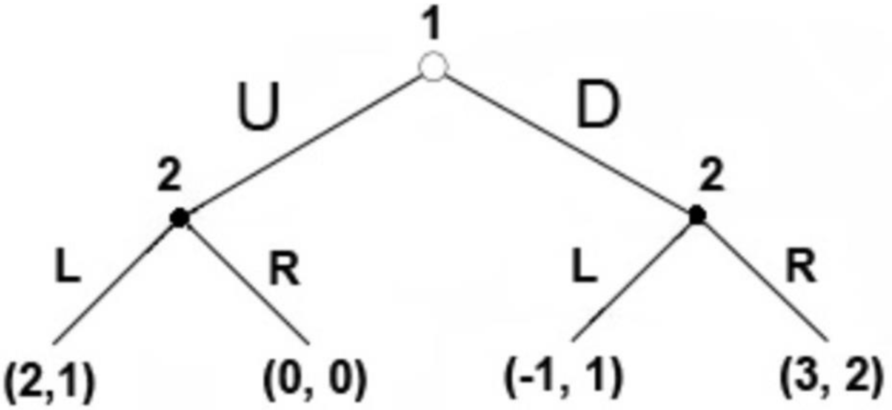


Figure 11: Extensive Form

In order to analyze the strategy of the player, bottom-up approach is the best way to determine the decisive mode of a rational player in an extensive game, traversing from the last vortex to the connected nodes in the decision tree to the parent vortex, which is also stated as backward induction in a multi-player extensive game scenario.

Types of games

In game theory, different types of games help us analyze different problems. They are categorized as per the supposing game symmetry, several players involved, the game model & cooperation among players.

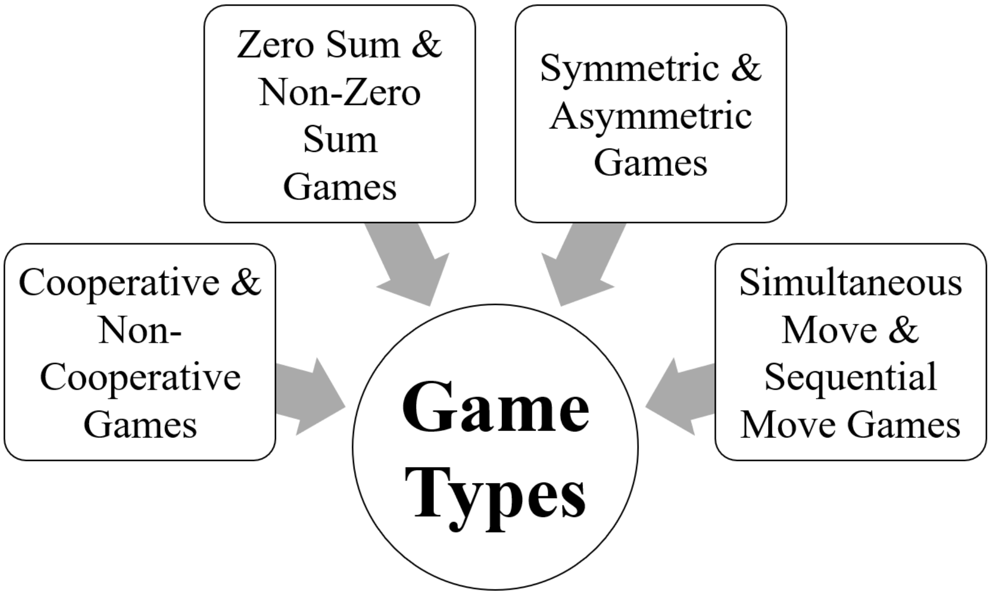

Zero Sum & Non-Zero Sum Games
Symmetric & Asymmetric Games
Cooperative & Non-Cooperative Games
Game Types
Simultaneous Move & Sequential Move Games

Figure 12: Game Types

Cooperative & Non-cooperative Games

A game is considered cooperative when players can establish agreements that are enforceable by external means, such as legal contracts. Conversely, a game is considered non-cooperative when players are unable to form alliances or when all agreements must rely on internal enforcement mechanisms, such as through the use of credible threats.

The study of cooperative games often employs the lens of cooperative game theory, which offers a holistic perspective by examining the structure and strategies involved in predicting the formation of coalitions, the actions taken by these groups, and the payoffs that result.

On the other hand, traditional non-cooperative game theory takes a broader approach, predicting the actions and payoffs of individual players and analyzing Nash equilibria.

This emphasis on individual payoffs can lead to the "Tragedy of the Commons," where resources are overutilized inefficiently. Additionally, a non-cooperative game theory also explores the impact of bargaining processes on the distribution of payoffs among coalitions.

In many scenarios, the ideal situation would be to rely on a single theory. However, when there is a lack of sufficient information to the model, the formal bargaining process or when the model becomes too intricate to be useful in real-life situations, cooperative game theory offers a simplified solution. This approach enables the examination of the game as a whole without having to make assumptions about the bargaining powers involved.

Nash Equilibrium

The Nash Equilibrium is a principle in game theory that deals with non-cooperative games. It states that each player can maximize their outcome based on the strategies of the other players in the game. This is reached when no player has any reason to change their strategy, even if they are aware of the other players' strategies. It is a decision-making theorem that suggests a player can achieve their desired outcome by sticking to their initial strategy, taking into account the actions of other players. Hence, it is utilized to determine the optimal response in different scenarios.

In the Nash Equilibrium, each player's approach is the best possible considering the moves made by the other players. As a result, every player benefits and obtains the outcome they had aimed for.

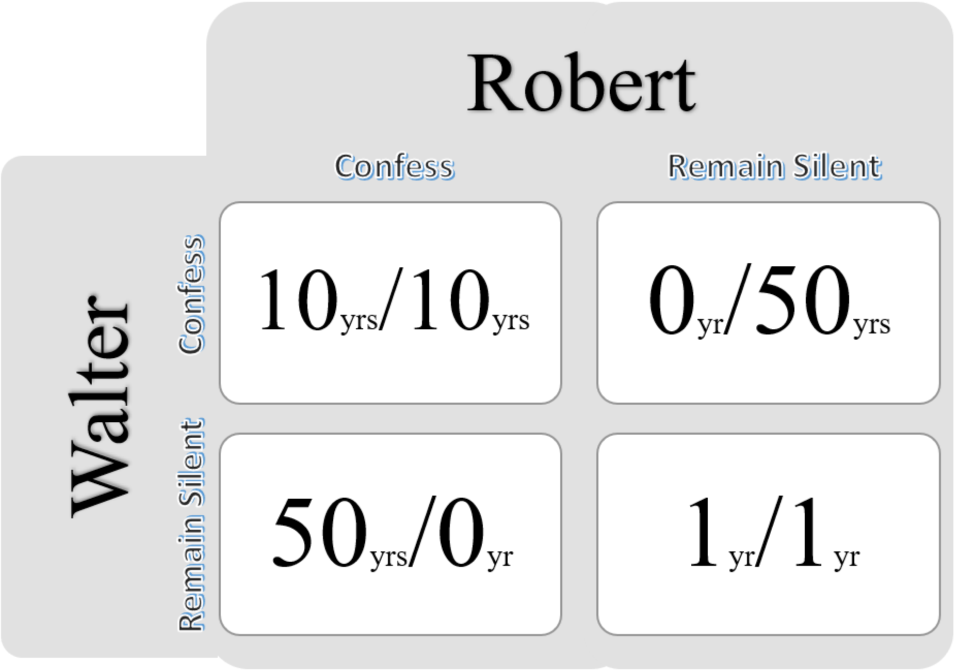
Prisoner’s Dilemma

The prisoner's dilemma is a well-known scenario in game theory that illustrates why rational individuals may not choose to cooperate, even if it seems like it would benefit them both.

The prisoner's dilemma represents a situation in which each person is motivated to act in a manner that results in a suboptimal outcome for the group as a whole.

In the classic example of the prisoner's dilemma, two members of a cartel named Robert and Walter are arrested and placed in solitary confinement with no means of communicating. The prosecutors do not have enough evidence to prosecute the two individuals on the primary charges.

Prosecutors hope to get both of them sentenced on a lesser charge. Simultaneously, the prosecutors offer each prisoner a bargain.



		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 13: Decision Matrix

Each prisoner is allowed either to betray the other by testifying that the other committed the crime or to cooperate with the other by remaining silent. The options offered are:

- If both confess, they each serve 10 years in prison.

- If Robert confesses and Walter stays silent, Robert is set free and Walter serves 50 years in prison.

- If Walter confesses and Robert stays silent, Walter is set free and Robert serves 50 years in prison.

- If both stay silent, they each serve only 1 year in prison.

Note, it is assumed that the prisoners have no means of affecting each other's sentence after the fact and that their decision will not impact their reputation in the future.

We’d feel both remaining silents would be the best option. But they will not opt for it. Both of them will betray each other i.e., they would confess. Because that is human psychology, this feels like the best option for both 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.

The incentive to betray a partner provides a greater benefit than cooperation, leading to the conclusion that if both individuals are purely rational and self-interested, they will each choose to betray the other. This results in the only possible outcome being mutual betrayal, which is referred to as the Nash Equilibrium.

The prisoner's dilemma serves as a model for understanding cooperative behavior in various real-life scenarios. Although the situation may not always fit the exact criteria of the classic or repeated version of the game, the term "prisoner's dilemma" can still be applied to situations where two parties could greatly benefit from cooperation or face negative consequences for not doing so. However, these parties may struggle to coordinate their actions, whether due to difficulty, cost, or other factors.

Zero sum & Non-zero sum Games

In Zero-sum or constant-sum games, players' choices do not affect the available resources. The total benefit in zero-sum games, for every combination of strategies, always equals zero, meaning one player's winnings are equal to the losses of the others. Poker is a prime example of a zero-sum game where a player's gain is directly linked to their opponents' losses. Such games are often associated with activities like gambling and theft, but not with a trade that can yield mutual gains.

A constant-sum game can be turned into a zero-sum game by adding a dummy player referred to as "the board," whose losses balance out the players' net winnings. Other zero-sum games include coin flipping and popular board games like chess.

In contrast, non-zero-sum games, such as the prisoner's dilemma, studied by game theorists, can have outcomes that result in a net gain or loss for the players. In non-zero-sum games, one player's gain doesn't necessarily result in another player's loss.

Simultaneous & Sequential Games

Simultaneous games refer to games in which both players make their moves at the same time, or in which later players are not aware of the previous player's actions. In such cases, normal form notation is utilized to depict the game.

On the other hand, dynamic games involve a situation where later players possess some understanding of prior actions. This information may not be complete, but rather limited. For instance, a player may be aware that an earlier player didn't take a specific action, yet not know which another move was taken. Extensive form notation is employed to represent such sequential games in this scenario.

It's important to note that one extensive-form game can be equivalent to multiple normal-form games. As a result, equilibrium concepts used in simultaneous games do not apply to analyzing sequential games due to their complexity.

Symmetric & Asymmetric Games

In a symmetric game, the rewards for a particular strategy are not dependent on the players themselves, but rather on the strategies employed. If switching the players does not alter the payoffs for the strategies, the game is considered symmetric. A number of popular 2x2 games, such as the prisoner's dilemma and stag hunt, fall into this category.

Asymmetric games, on the other hand, are characterized by a lack of identical strategy sets for both players. Games such as the dictator game and ultimatum game are examples of asymmetric games with different strategies for each player. Despite this, a game can have the same strategies for both players and still be considered asymmetric if other factors, such as the payoffs, are not equal.

DEFENSE-ATTACK MODEL

When creating a defense model in the realm of cyber-physical systems, we focus on parties who have conflicting interests: the attacker and the defender. The defender, typically the system administrator, is responsible for managing the system and their primary objective is to protect the cyber and physical infrastructure from any harmful actions.

On the other hand, the attacker is the adversary who tries to breach the end target system. The interaction between the attacker and the defender is modeled using actual data from security incidents to provide a realistic representation.

Defender

The defender is the responsible party for ensuring the security of the CPS against malicious attacks. To do so, the defender has access to a set of monitors that act as a safeguard for the system. The primary goal of this player is to make appropriate responses in an anticipatory manner, utilizing their limited view of the system's status and relying on the monitors designed into the CPS.

Defence State

'DS_x', the perceived state of an attack as seen by the defender. The information that the defender utilizes is based on the monitoring systems, however, it is limited in its level of detail due to the lack of granularity in the data available in the database. This means that the observations made by the defender do not provide a complete picture of the user's actions.

Defender Action

'D', Set of actions available to the defender in a specific state. In the realm of security incident response & detection, a admonitor is tasked with identifying changes in the system status. However, these detections may not accurately reflect the actions of an attacker. The admonitor may overlook an action or wrongly categorize a harmless action on the CPS as a harmful attack, resulting in false negatives or false positives.

Given this uncertainty, the defender must make informed decisions based on imperfect information. To simplify the process, the defender's actions are regulated and abstracted into two options: Respond or Not Respond, assuming that there is an appropriate response for each action.

Attacker

The attacker is the adversary who gains access to the system intending to compromise its security. Attacks can range from a solitary action to a series of actions. For this analysis, our focus is limited to attacks that involve multiple steps with the ultimate objective of infiltrating the CPS.

Attack State

'AS_x', state of an attack which denotes the extent of the intrusion. Every attack state is given a numerical value, referred to as a reward, which measures the harm inflicted on the target CPS. The greater the damage, the more severe the impact on the system and/or the higher the level of unauthorized control. The progression from one state to another is contingent on the outcome of the action taken.

Activity

The attacker has a range of moves, referred to as a_i at their disposal. These actions could result in a takeover of the Cyber Physical system or, if the attacker chooses to stay in the same state, it will result in repetition. The specific set of activities that can be taken within the state AS_x is represented as A_x, which is a portion of the larger group of action set '*A*'. The relationship between the actions and the attack states is depicted often through a state transition illustration.

Transition Matrix

The probability of moving from one state *(s)* to another *(s')* as a result of a specific action is represented by $P_a\,(s, s')$. In the context of an attack model, a transition matrix is used to show the likelihood of a successful attack on a cyber-physical system. The effectiveness of the defender's monitoring system determines if the attack will be detected or missed, and the transition matrix accounts for this uncertainty. The reward earned by the attacker as a result of transitioning from state s to s' through a specific action is referred to as the Immediate Reward $R_a\,(s, s')$. This reward is a numerical value that depicts the benefit the attacker gains from a successful attack.

Attacker-Defender Interaction

The decisions made by the attackers in a system are not isolated, but rather, they are linked to the decisions made by their opponents. Thus, the interaction between the two players is modelled. After an attacker takes an action, the defender chooses their response based on information obtained from their monitoring system.

A successful transition to the intended state for the attacker can only occur if the defender fails to respond appropriately. On the other hand, if the defender detects the attack and takes action, the attacker is forced to transition to the default state.

In a zero-sum game scenario, a successful attack results in an immediate reward for the attacker and a symmetrical loss for the defender. The attack state will change as a result of the attack execution.

However, if the defender successfully detects the attack and responds accordingly, the attack state will reset to the default for the identified attacker and the defender will receive a reward, with a corresponding loss for the attacker.

GAME MODEL IMPLEMENTATION

A game with two or more rational players in conflict is made up of individuals trying to maximize their rewards through the identification of the best possible strategy for each state.

The concepts of state quality and state value are used to indicate the expected outcome of a player's decision within the framework of the game model.

Stochastic Game

A Stochastic game / Markov game is a type of repeated game with uncertain transitions, played by one or more players. The game progresses through a series of rounds. At the start of each round, the game is in a particular state, and the players choose actions.

Each player then receives a reward that is based on the current state and their chosen actions. The game then moves to a new state, the distribution of which is determined randomly based on the previous state and the actions taken by the players. This process is repeated at the new state, and the game continues for either a limited or unlimited number of rounds.

The final payoff for a player is typically calculated as the discounted sum of the rewards received in each round or as the limit inferior of the average rewards in each round. The collection of all available moves for a player is referred to as the set of actions, denoted as A.

The payoff or immediate benefit from a player's choice of action and the opponent's move in a particular state, is represented by the reward function $R(s, a, o)$, where s represents the attack state in the t-th iteration and o represents the opponent's action. To gauge the anticipated reward of a player's decision, the concepts of state quality and state value are employed.

A Stochastic game or Markov game is a type of repeated game in which one or more players take part and involves probabilistic transitions. The game consists of a series of rounds, each starting in a specific state.

In a gaming scenario, players select their moves at the start of each round. The resulting rewards for each player are based on the current state and the chosen actions. The game then shifts to a new state, which is generated randomly based on previous state and player moves. This cycle of selecting moves, receiving rewards and transitioning to new states continues for either a fixed number of rounds or an indefinite amount of time.

A player's total reward can be depicted as either the sum of all rewards from each round with discounting or as a lower bound estimate of the average reward from each round. The set of actions *A* encompasses all the options a player can pick from. The reward *R* for a particular round is dependent on the state of play *s*, the player's action *a*, and the opponent's action *o* at that stage.

The state value and state quality concepts are utilized to estimate the expected reward for a player's move.

Value of state

V(s), The expected reward for a player, is calculated based on the player starting from a specific state and utilizing the optimal strategy. The value represents the highest reward that the player can hope to receive, under the assumption that the opponent will select an action o that minimizes the expected reward.

To increase the state value, the player seeks to determine the ideal value policy, which is a distribution of the possible actions they can take in a given state.

$$V(s) = \max_{\pi} \min_{o\prime \in Os} \sum_{a\prime \in As} \pi(s, a')Q(s, a', o')$$

Discount factor

The weight placed on future versus current rewards is determined by the value assigned to γ. If a player prioritizes current rewards over future ones, a value of 0 is assigned. Conversely, a value of 1 is given to a player who aims for a high, long-term reward.

Quality of state

The expected reward a player can receive by taking actions a and o from state s and subsequently following the optimal policy is represented by *Q*(*s, a, o*) . The "Quality of state" is the sum of the immediate reward from the current iteration

(R_{t-1}) and the expected reward that results from transitioning to state *s' (V^t (s'))*, which was calculated based on previous iterations. It's important to note that the value of the state is influenced by a discount factor (γ).

$$Q^{t+1}(s,a,o) = R^{t+1}(s,a,o) + \gamma V^t(s')$$

Optimal policy

The set *π* signifies the distribution of potential actions *π(s, .))* at each state *s*. The goal is to choose *π* in such a way that it maximizes the expected reward for the player, which is indicated by the *V(s)* value of the state. *π(s, a)* specifies the probability of taking action *a* in a given state, with *π* being the overall distribution that optimizes the value of state *V(s)*.

$$\pi(s,.) = arg \max_{\pi\prime(s,.)} \min_{o\prime\in Os} \sum_{a\prime\in As} \pi(s,a')Q(s,a',o')$$

Minimax Q-Learning

Security games often face the challenge of an unrealistic assumption of complete information and rationality. Players in these games usually make decisions based on limited information, but try to overcome this challenge through learning.

To deal with this situation, Minimax Q-Learning is utilized as a decision-making algorithm. It doesn't require complete knowledge of the attack model, instead, it combines past results with the current earnings and future expected rewards to make informed decisions by assigning partial weight to historical data.

Quality of state

In Minimax Q-Learning, the assessment of a state is determined by its estimated value, which is captured by the state-value function *V(s)*. This function provides a projection of the reward a player could expect to receive if they follow the best possible strategy.

Minimax Q-Learning aims to identify the optimal policy that yields the highest expected reward, and this is achieved by progressively refining the state-value function estimations.

The calculation of state's quality in Minimax Q-Learning involves an examination of the estimated rewards that could be obtained from each possible action that can be taken from that state. The algorithm considers the reward that would result from each action, taking into account the expected reward of the state that would follow, as well as the expected reward of future states.

The combination of these estimations produces the state's quality evaluation. The algorithm continually updates its estimates of the state-value function until it reaches the optimal policy. Quality of a state in Minimax Q-Learning reflects the player's best approximation of the expected reward that would result from executing the optimal policy from that state.

The higher the state's quality, the more valuable it is considered to be for the player, increasing the probability that the player will choose the action corresponding to that state.

$$Q^{t+1}(s,a,o) = \alpha Q^{t}(s,a,o) + (1-\alpha)R^{t+1}(s,a,o) + \gamma V^{t}(s')$$

Learning rate

α modulates the player's learning capacity by assigning a real number between 0 and 1. If the learning rate is 0, it signifies that the player has complete learning ability. On the other hand, if the rate is 1, it means that the player only considers the most recent information.

When the player has full learning ability, it does not take into account the immediate reward *R(s, a, o)* or the expected future reward *V(ns),* but instead maintains the state quality constant.

At the start of the game, where prior results to learn from are missing, α is assigned a value of 1.0. As *Q(s, a, o)* accumulates information about the performance of previous iterations, α gradually decreases.

Exploration rate

exp is a unique factor that influences how much the player deviates from the ideal policy. Unlike Markov games, where the optimal solution is established from the outset, Q-Learning requires experimentation to determine the best course of action.

The exploration rate indicates the likelihood that the player will deviate from the optimal policy in order to assess the outcomes of different moves.

A low exp value will lead to the player following a similar approach as in a Markov game, while a high value will prompt the player to make more random choices.

Naive Q-Learning

In a security game scenario, both the attacker and defender often face the challenge of limited information about their opponent. While the attacker may gather some information about the target system from publicly available resources, the amount of knowledge is still restricted.

In a similar vein, the defender also plays against an unknown adversary. To address this situation, Naive Q-Learning is employed. This approach optimizes strategy without relying on information about the opponent's actions.

Utilizing only the limited information about the immediate rewards and the player's own information to determine the most optimal policy.

Quality of state

In Naive Q-Learning, the quality of state refers to the value that the algorithm assigns to a particular state. This value represents the level of desirability of the state in terms of achieving a goal or maximizing rewards.

The calculation of the quality of the state takes into account the rewards received for taking certain actions in the state, as well as the projected rewards from future states. As new information is encountered, the algorithm continually updates the quality of state, thus continually refining its approach.

$$Q^{t+1}(s,a) = \alpha Q^{t}(s,a) + (1-\alpha)R^{t+1}(s,a) + \gamma V^{t}(s')$$

Value of state

Value of a state represents the significance assigned to it by the algorithm. It symbolizes the predicted cumulative reward that would be attained in the long run if the agent is in that state and follows a specific course of action.

The value is continuously adjusted as the algorithm gains new insights from the rewards obtained from executing actions in the state, as well as the rewards estimated from subsequent states. By updating the value of a state, the algorithm is capable of molding its approach over time.

$$V(s) = \max_{\pi} \sum_{a\prime \in As\prime} \pi(s,a')Q(s,a')$$

Optimal policy

The main objective is to identify the sequence of actions that will yield the highest long-term reward for a given state. This is achieved by the algorithm using a trial-and-error approach, continually adjusting the value assigned to each state based on the observed rewards of specific actions and the estimated rewards from future states.

Over the course of iterations, the algorithm becomes more informed and develops an understanding of which actions lead to the best outcomes, leading to the creation of the optimal policy.

$$\pi(s,.) = arg \max_{\pi\prime(s,.)} \min_{o\prime \in Os} \sum_{a\prime \in As} \pi(s,a')Q(s,a')$$

CONCLUSION

Cyber Physical Systems are Multidisciplinary engineered systems that could transform the lens we see upon the world. Indeed, combining multiple theories into a single system, arises problems in several multitudes, but the right blend of these technologies fabricates a robust hybrid system that is resilient enough to tackle all modern-day problems.

The invergence of CPS are the underlying basis for Industry 4.0. The concept of digital twinning is about to make a pivotal mark in the field of production and manufacturing. With overwhelming levels of potence & complexity, the idea of industry-level deployment is at unease which is also open research.

The integration of both Physical & Cyber aspects poses newer threats to the environment, the game theoretical approach when implied on varied sections of the ecosystem, the CPS can construct the righty columns to defend the system, before the attack even arises based on the testbed environments, several players, playoffs and strategies of the respective game theoretical model. Opting for the proposed Stochastic, Minimax Q-Learning or Native Q-Learning game; we’ll be able to attain the ideal Quality of state such that we’d be able to preserve the integrity of the ecosystem.

ACKNOWLEDGEMENT

We would like to acknowledge and give my warmest thanks to Dr. B. Giriraj, Principal, PSG Polytechnic College, who made this work possible. We would also like to thank Mr. A. Kathiresan, Head-In-Charge, Department of Information Technology for allowing us to proceed with the stages of writing.

We would like to give our special thanks to Dr. K. Anitha Kumari, Associate Professor, Department of Information Technology, PSG College of Technology, the guiding force and for the outstanding support in proposing script writing and in completing this successfully.

Finally, I would like to thank in the name of creator and parents for the completion of the writing.

REFERENCES

Mahmoud Parto, Pedro Daniel Urbina Coronado, Christopher Saldana, and Thomas Kurfess, "Cyber-Physical System Implementation for Manufacturing with Analytics in the Cloud Layer", ASME, JCISE-20-1330, July 14, 2021.

Yongkui Liu, Xun Xu, "Industry 4.0 and Cloud Manufacturing: A Comparative Analysis", MANU-16-1445, October 6, 2016.

Siva CHAITANYA Chaduvula, Adam Dachowicz, Mikhail Atallah, Jitesh Panchal," Security in Cyber-Enabled Design and Manufacturing: A Survey", Research Gate, DOI:10.1115/1.4040341, July 2018.

Tobias Post, Rebecca Ilsen, Bernd Hamann, Hans Hagen, Jan C. Aurich," User-Guided Visual Analysis of Cyber-Physical Production Systems", JCISE-15-1310, February 16, 2017.

Hausi A. Müller, "The Rise of Intelligent Cyber-Physical Systems", IEEE, DOI: 10.1109/MC.2017.4451221, 18 December 2017.

Kyoung-Dae Kim, P. R. Kumar," Cyber–Physical Systems: A Perspective at the Centennial", IEEE, DOI: 10.1109/JPROC.2012.2189792, 03 April 2012.

Yabing Huang, Jun Zhao," Cyber-Physical Systems with Multiple Denial-of-Service Attackers: A Game-Theoretic Framework", IEEE, DOI: 10.1109/TCSI.2021.3098335, 27 July 2021.

Amit Tyagi, N. Sreenath, "Cyber Physical Systems: Analyses, challenges and possible solutions", ResearchGate, DOI: 10.1016/j.iotcps.2021.12.002, December 2021.

Rasim Alguliyev, Yadigar Imamverdiyev, Lyudmila Sukhostat, "Cyber-physical systems and their security issues", ScienceDirect, DOI: 10.1016/j.compind.2018.04.017, 14 May 2018.

Keywhan Chung, Charles A. Kamhoua, Kevin A. Kwiat, Zbigniew T. Kalbarczyk, Ravishankar K. Iyer, "Game Theory with Learning for Cyber Security Monitoring", IEEE, DOI: 10.1109/HASE.2016.48, 03 March 2016.

Ankica Barišićab, Ivan Ruchkinc, Dušan Savićd, Mustafa Abshir Mohamed, Rima Al- Ali, Letitia W. Lig, Hana Mkaouarh, Raheleh Eslampanahi, Moharram Challengeri, Dominique Blouinh, Oksana Nikiforovaj, Antonio Cicchettik, “Multi-paradigm modeling for cyber–physical systems: A systematic mapping review”, ScienceDirect, DOI: 10.1016/j.jss.2021.111081, 6 September 2021.

Khan Md Shafi Ahad Siddique B.S., “A Game Theoretic Framework to Secure Cyber Physical Systems (CPS) Against Cyber Attacks”, Master’s Thesis, Texas State University, December 2018.