Implementation and Enhancement of Endogenous Security Mechanisms in Photovoltaic Data Storage and Transmission

[1]Brahmadesam Viswanathan Krishna, [2]Sathvik Bagam, [3]Jayasri R, [4]Krishnaveni N, [5]Prasath R and [6]Tanweer Alam

[1]Department of Computer Science and Engineering, KCG College of Technology, Chennai, Tamil Nadu, India.
[2]Software Development Team Lead at Paycom, Masters in Computer Science, Oklahoma Christian University, Edmond, Oklahoma, United States, 73013.
[3]Department of Artificial Intelligence and Data Science, St. Joseph's College of Engineering, OMR, Chennai, Tamil Nadu, India.
[4]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai, Tamil Nadu, India.
[5]Department of Computer Science and Engineering, RMK College of Engineering and Technology, Puduvoyal, Tiruvallur, Tamil Nadu, India.
[6]Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah, Saudi Arabia.
[1]krishna.cse@kcgcollege.com, [2]sathvikbagam7@gmail.com, [3]jayasrir@stjosephs.ac.in, [4]drnkrishnaveni@veltech.edu.in, [5]prasathr05@gmail.com, [5]tanweer03@iu.edu.sa

Correspondence should be addressed to Brahmadesam Viswanathan Krishna : krishna.cse@kcgcollege.com

Article Info
Journal of Machine and Computing (https://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202505077
Received 12 September 2024; Revised from 19 December 2024; Accepted 27 February 2025.
Available online 05 April 2025.

Abstract – The global energy market is migrating toward sustainable renewable energy sources (RES), with solar energy (SE) being the most significant due to its abundance and reliability. Photovoltaic (PV) converts SE into electricity, relying on data integrity and security. However, digitized data has cybersecurity vulnerabilities, including data breaches and attacks. Traditional security systems can provide essential protection but fail to address PV's dynamic and distributed nature, leading to gaps in defense against evolving cyber threats. The study proposes an endogenous security model for improving data transmission and storage within PV. It uses a Verification Feedback Mechanism (VFM) to integrate routing methods, compute efficient data paths, schedule them periodically, and verify their integrity. The model also incorporates cryptographic key infrastructure and key management protocols to ensure secure data transmission and management. This approach addresses challenges in data forging and ensures the integrity of the network's components. The study compared two methods and found one model superior in communication integrity and system adaptability. It achieved latency statistics below 20 *ms* and maintained Network Throughput (NT) at 9.2 *Gbps* even when attacked, demonstrating its effectiveness in securing PV from multiple cyberattacks.

Keywords – Photovoltaic Systems, Renewable Energy Sources, Solar Energy, Cyber-Physical Security, Data Transmission Integrity Rates**.**

I. INTRODUCTION

Since the global community is migrating towards Renewable Energy Sources (RES), photovoltaics (PV) has evolved into a significant component of the Renewable Energy (RE) geographical region. By reducing the requirement to rely on natural resources, PV automatically transfers energy from ultraviolet radiation into electrical power, resulting in a more sustainable RE. Solar PV [1] differs among RE uses in that devices can be built up or down to connect to the Smart Grid (SG) at multiple levels, from individual residences to enormous solar power plants. Among other key measurements, data on PV radiation, electricity generation, and EC behaviours is necessary for PV functioning, management, and efficiency improvement [2]. The data improves the reliability and sustainability of PV installations, enabling proactive maintenance of these systems via proper data processing. In order to successfully monitor and maintain PV, it is essential to maintain the reliability and safety of these vital data.

However, digitizing and transmitting such data through the existing network connectivity infrastructure to perform all the PV-related operations has exposed the system to a complex environment of cybersecurity attacks [3]. The threats to PV can range from data breaches and unauthorized access to more complex attacks compromising the integrity and availability of critical PV energy data [4]. Further, as these PV are ultimately integrated into the national Smart Grid (SG), the complexity of handling and the probable impact of these attacks have become more imminent and have extended beyond the individual installations, which posed risks to the broader energy infrastructure's stability and reliability. In response to these challenges, various security mechanisms have been developed based on techniques such as encryption, authentication, and secure communication protocols. Such security mechanisms only frequently addressed the specific aspects of cybersecurity that had provided room to be exploited by knowledgeable attackers [5].

This is where endogenous security systems come into the field, which is unlike other security measures that are applied as external layers of security; the endogenous security systems, on the other hand, are models that are integrated into the core operational model of PV [6] It is implemented as core components of the system's design thereby ensuring the security and data integrity in the model during the data transmission and storage processes. However, few models have been developed using endogenous principles for PV data security, which provides plenty of room for proposing models in this domain. The motivation for this work is grounded in the limitations of existing security systems and the limited work on endogenous-based cyber-physical systems (CPS) for PV environments.

The proposed work addresses the above limitations and gaps by introducing an endogenous security model for the security and integrity of data transmission and storage in PV. The model employs an integrated routing model that employs three routing strategies: Data Integrity Forwarding (DIF), Load Balanced Forwarding (LBF), and Path Diversity Forwarding (PDF), together with Verification Feedback Mechanism (VFM) for ensuring reliability. This integrated model enables the simulation to take advantage of the best possible use of the resources on the network while reducing the probability of data manipulation [7]. In addition to evaluating the validity by finding paths with minimal collision and actively updating these paths according to network situations, the design includes techniques for computing the most appropriate data paths for scheduling [8]. The design employs a secret cryptographic key system and protocols for key management to ensure security when transmitting information via multiple routes. Based on the findings of the test of the security system using different performance indicators, the recommended security model attained data transmission reliability scores ranging from 97% to 99%, reliability rates ranging from 94% to 96%, and Network Throughput (NT) maintaining near 9.2 GBPS under attack issues with latency less than 20 *ms*.

The article is organized as follows: Section 2 provides the existing literature review, Section 3 shows the historical context of the work, Section 4 presents the recommended security model, Section 5 analyses the model, and Section 6 presents the conclusion. The paper is structured in the following order.

II. LITERATURE REVIEW

Based on the challenges experienced by CPS-based PV, their [7] paper provided an in-depth review. The study has emphasized the diversity of cyberattacks that have been the target of PV, which have ranged from data integrity to software-based attacks, and the work has also introduced a success rate metric to assess the impact of these attacks. It also explored model-based and data-driven approaches for threat detection and mitigation, highlighting the effective role of blockchain technology in securing software and CPS.

Authors [9] focused on the communication security of PV by a voltage regulation scheme. This scheme has been built to operate on two levels to reduce voltage deviation and voltage difference, and it incorporates a power compensation system for primary regulation and a consensus protocol for secondary regulation. The approach was built to handle the communication topology changes and delays by proposing predictive compensation for packet loss and significant delays. Their method's effectiveness is being validated using MATLAB simulations, and the results have shown the models' better performance.

Authors [10] have addressed the broader implications of integrating RE sources into SG, particularly the security problems associated with wireless data transmission and centralized power trading. They indicate a secure energy market approach appropriate for Smart Grid (SG), which employs Wireless Sensor Networks (WSN) for communication and is endorsed by blockchain.

In order to enhance the success of energy making and RE use, the authors propose a dual-chain design that saves electrical data about the blockchain while employing Smart Contracts (SC) for business decisions. In Germany's "Digitalization of the Energiewende" governance [11] has dealt with the evolution of the distribution systems, including SG. The study seeks to show how Controllable Local Systems (CLS) and Smart Metres collaborate to build a secure energy data network, enabling secure communication with distributed RES like PV cells and batteries. The reality that this reciprocal achievement meets data security regulations in Germany highlights the chance for enhanced management and integration of decentralized RES into SG. Their study shows the development and execution of a node controller that deals with the secure communication requirements associated with cloud-based RES [12].

In order to verify the system's security and reliable performance, the study emphasized the importance of security measures adapted to the demands of the network of communications. In an example experiment in India, their model exhibited more significant enhancements in peak-clipping, valley-filling effects, and overall load features, proving the value of the method. More significantly, for the combination of distributed energy and improving the effectiveness of RES,

[13] have been studying CPS difficulties associated with multi-station systems. The study includes several RE sources to supply a supplemental design and techniques for securing the natural environment of Smart Energy Stations (SES). Security zone and congestion isolation systems are two techniques they suggest for improving SES-CPS in the context of potential attacks.

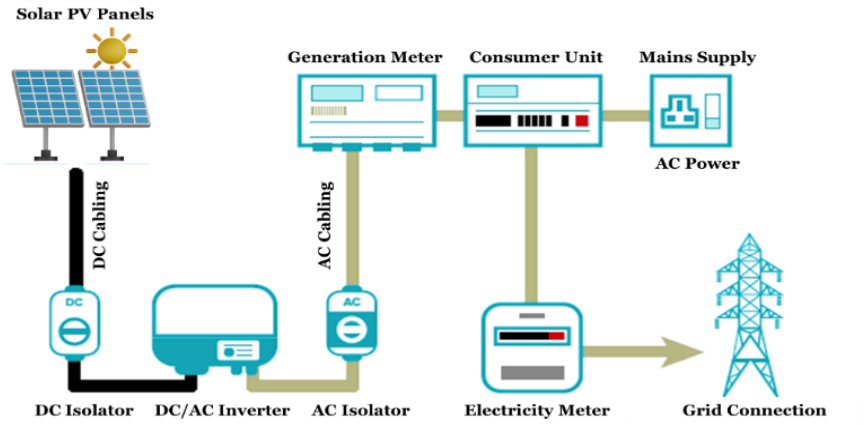

Fig 1. PV System.

Previous investigations have concentrated on enhancing the quality and performance of Photovoltaic Storage Systems (PVSS) in an Energy Blockchain (EB) context [14]. They achieve this by developing a task-matching model employing the Genetic Algorithm – CLOUD-Gale-Shapley (GA-CLOUD-GS). The work developed a method for integrating subjective and objective weights for matching tasks. Several computer simulations, sensitivity studies, and comparison analyses have demonstrated the model's performance and provided helpful information for Publicly Verifiable Secret Sharing (PVSS) task balancing in an EB environment [15]. Applying SG to solar PV, their research focuses on threat identification and risk estimation. Their approach involved identifying, assessing, and mitigating CPS risks specific to SG with those of solar PV integration. Utilizing the proposed Scheme for Trans-disciplinary Research for India's Developing Economy (STRIDE) for threat classification and the following proposed model, the DREAD stands for (Damage potential, Reproducibility, Exploitability, affected users, and Discoverability) threat-risk ranking model for prioritization, and the proposed study has been practical in the process of identifying the high-risk threats that have included information disclosure and elevation of privilege.

III. BACKGROUND

Structure of Solar PV

A typical solar PV (**Fig 1**) comprises several key components that combine to convert sunlight into electrical energy that can be used in homes/industries in the power grid [16-20].

The components of the PV include:

- *Solar PV Panels:* The solar PV panels capture sunlight and convert it into Direct Current (DC) electricity. These panels contain PV cells made from semiconductor materials exhibiting PV effects.
- *DC Cabling and DC Isolator:* The DC electricity generated by the panels is transmitted via DC cabling. This wiring is connected to a DC isolator, a safety device that disconnects the PV from the electrical circuit for maintenance.
- *DC/AC Inverter:* This device converts the DC electricity from the solar panels into Alternating Current (AC) electricity.
- *AC Isolator:* The AC isolator provides a point of disconnection for the AC converted from the solar panels.
- *Generation Meter:* The generation meter is connected to the inverter and measures the amount of AC electricity the solar PV produces.
- *Consumer Unit:* Also known as the fuse box, the consumer unit is where the electricity is distributed to different circuits within the home.
- *Electricity Meter:* The electricity meter records the amount of electricity the SG consumes.
- Connection to the SG: The system is connected to the main supply SG.
- *Mains Supply:* The main supply represents the household's connection to the public electricity SG.

Data Generation in Solar PV

Data generation in solar PV is a continuous process collected from the operators, owners, and utility companies. Making informed decisions about maintenance energy usage and enabling performance optimization using the data is possible.

Data Collected
Solar PV Panels: P_{pv}
Power Output Data (P_{out}) *:* Each panel generates data on the amount of electrical power (P_{out}) it produces.
Environmental Data (E_{data}) *:* Solar panels are often equipped with sensors that collect environmental data (E_{data}), such as irradiance and temperature.

DC/AC Inverter: $I_{dc/ac}$
Voltage and Current Data ($V_{dc}, I_{dc}, V_{ac}, I_{ac}$ *):* The inverter captures data on the DC voltage (V_{dc}) and current (I_{dc}) from the PV panels and the AC voltage (V_{ac}) and current (I_{ac}) it outputs to the grid.
Efficiency Data (η_{inv} *):* It also records its operational efficiency (η_{inv}), measuring how well it converts DC to AC power.

Generation Meter: M_{gen}
Energy Generated Data (E_{gen}*) :* The generation meter logs the total energy produced (E_{gen}), usually in kilowatt-hours (kWh).

Consumer Unit: CU
Load Distribution Data (LD_{data} *):* The consumer unit provides data on load distribution (LD_{data}).

Electricity Meter: M_{elec}
Consumption Data (C_{data} *):* Records how much energy is consumed by the SG (C_{data}) and track energy exported back to the grid.
Net Usage Data (N_{usage} *):* Provides net usage data (N_{usage}), which is the difference between energy produced and consumed.

Monitoring and Control Systems: MC_{sys}
Performance Data (P_{data}) *:* These systems aggregate all the data (P_{data}) it provides a comprehensive overview of the system's performance from various components.
Alerts and Fault Data (A_{data}) *:* They also generate alerts and log fault data (A_{data}).

Data Communication: D_{com}
Transmission Data (TD_{data} *):* The system includes data transmission components that relay all collected data (TD_{data}) to a central monitoring point or off-site data centre for further analysis.

Data Communication in Solar PV Power Plants
In solar PV power plants, the usage of Supervisory Control and Data Acquisition (SCADA) systems is dynamic in the task of enabling remote management of several field devices, including sensors, smart meters, Remote Terminal Units (RTU), and Intelligent Electronic Devices (IED). These systems comprise a network of components that coordinate to ensure operational efficiency and reliability.

The Components of a SCADA include

- *Data Acquisition Units:* These are responsible for measuring and gathering key monitoring parameters like voltage, current, temperature, and irradiance.
- *RTU:* These units serve as the intermediary, collecting data from the acquisition units, processing it, and relaying it to the primary control system.
- *Communication Networks*: The system's backbone, facilitating data transmission from the acquisition units to the control centre.
- *System Servers:* At the heart of the SCADA, many servers analyze and display the collected data for further action.

This includes

- *Front-End Servers:* Tasked with aggregating data from the field devices.
- *Historian Servers:* These servers act as data repositories, archiving data input for future reference and analysis.
- *Web Servers and Human-Machine Interfaces (HMI):* They provide a visual representation of the data and the status of the power plant, allowing for real-time monitoring and control.

Central Control Center
The control center contains SCADA, in which the application servers dissect the incoming information, making it user-friendly for operators to assess and act upon. This system performs data collection and visual analysis functions to ensure that each PV power plant operates at its peak.

Illustrated in **Fig 2**, the communication network of a PV showcases the interconnectivity between the local control centers, each dedicated to the management of a singular PV power plant. These centers are linked through a vast area network, bridging the communication gap via routers, ensuring seamless data flow and centralized control.

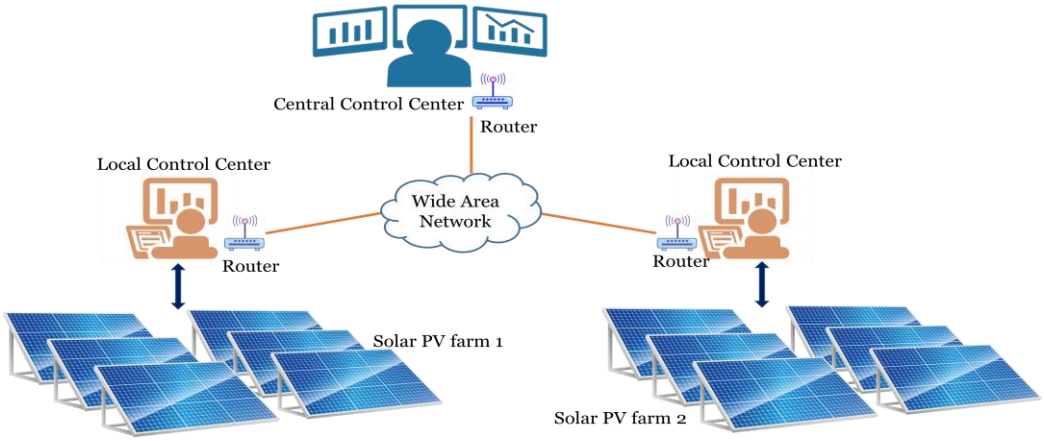

Fig 2. Communication Network For A PV Monitoring System.

Data and Power Integration in Communication Lauer

The data and power network is structured into three integral layers, as shown in **Fig 3** the physical infrastructure of the PV, the data transmission backbone of the communication network, and the user-interfacing application layer.

The PV Power System Layer

The PV power system layer is the foundation of a large-scale solar plant. The physical layer is where sunlight is captured and converted into usable electricity. This layer comprises all the essential equipment, including:

- *PV Modules:* The solar cells that capture sunlight and convert it into electrical energy.
- *Junction Boxes and Circuit Breakers:* Safety devices that protect the system from electrical malfunctions.
- Protection Devices: Equipment designed to shield the system from overloads and short circuits.
- *PV Inverters:* Devices that convert the DC generated by the PV modules into AC suitable for the power grid.
- *Power Cables:* Conductive wires that transport electricity throughout the plant.
- *Grid Connection Points:* Interfaces where the PV plant connects to the external power grid.
- *Transformers and Substations:* Apparatus that adjust voltage levels for efficient transmission and distribution.

Solar panels are connected in series to form a module string, which increases the voltage output. Multiple strings are then grouped and connected in a string combiner box. The string combiner boxes route the electricity to the Power Condition Unit (PCU), the system's main component. It converts the DC from the panels into AC. The specific type of transformer used depends on the overall plant design.

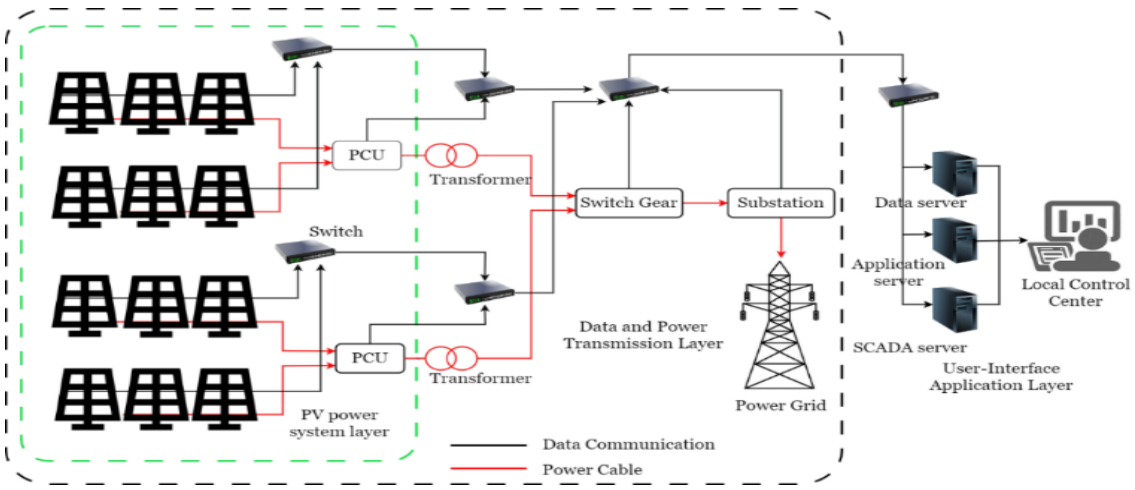

Fig 3. Layout of SG Integration of A Large-Scale PV Power Plant.

Communication Network Layer

The communication network layer acts as the central nervous system of a large-scale PV power plant. It facilitates the critical two-way flow of information between many subsystems and the local control centre.

This Layer Comprises Several Key Elements

Communication Devices

A network of interconnected devices like cables, routers, and switches ensures reliable data transmission throughout the plant.

Sensor Nodes and Measurement Devices
These intelligent devices gather data from the PV power system layer. They might monitor voltage, current, power output, ambient temperature, or wind speed.

Data Transmission
The communication network layer transmits two main categories of data:

- *Monitoring Data from the PV subsystem:* This includes real-time information on the performance of various components, allowing for early detection of potential issues.
- *Meteorological Parameters:* Data on weather conditions, such as solar irradiance and wind speed, is crucial for optimizing energy production and maintenance schedules.

Application Layer
The application layer acts as its central brain. Here is how it transforms collected data into intelligent control:

Control Center Command
It receives data from the communication network layer, encompassing information on:

- PV Panel Performance
- Inverter Status
- Transformer Efficiency
- Grid Connection Health

Table 1. Attacks/Threats and Its Implications

Attack Vector	Cyber Attacks	Potential Impact
Data Transmission	Data Interception and Theft	Compromise of data confidentiality; unauthorized access to sensitive information
Communication Networks	Denial of Service (DoS) Attacks	Disruption in monitoring and control capabilities; operational downtime
	Man-in-the-Middle (MitM) Attacks	Data tampering; incorrect operational commands leading to system inefficiency
SCADA	Malware and Ransomware	Manipulation of operational data; ransom demands for system control
	Injection Attacks	System malfunction; unauthorized control
Smart Meters and Sensors	Firmware Tampering	Inaccurate data reporting; energy theft
	Protocol Exploits	Unauthorized access; data manipulation
Internal Network	Insider Threats	Unintentional data breaches; intentional sabotage
Software and Digital Interfaces	Software Vulnerabilities	Exploitation leading to system compromise; data corruption
	Phishing and Social Engineering	Credential theft; unauthorized system access

Data Analysis and Storage
The control centre houses server systems that collect, process, and store this data. And employ applications to analyze these data to identify trends and potential issues and optimize performance.

Decision-Making and Control Actions
The control center can make informed decisions and take appropriate actions based on the analyzed data.

This Might Involve

- Adjusting inverter settings to optimize power output
- Activating maintenance protocols for faulty equipment
- Regulating power flow to meet grid requirements

Attack Vectors and Cyber Threats in PV Power Systems
The threats on PV can be considered attack vectors that target PV's physical components and cyberinfrastructure. The **Table 1** outlines the attack vectors and cyber threats specifically targeting PV power systems:

IV. PROPOSED ENDOGENOUS SECURITY IN PV SYSTEMS

Assumptions

To validate the effectiveness of the proposed endogenous CPS for PV data transmission and storage, we proceed under the following assumptions:

- *Resource Sufficiency:* The central controller (managing data flows and security protocols) and networked PV devices possess ample processing capabilities and resources, ensuring minimal processing delays.
- *Bounded Propagation Delay:* The propagation delay within the PV's communication network is limited, allowing the central controller to monitor data packet movements accurately within acceptable time frames.
- *Adequate Bandwidth:* The communication links within the PV have sufficient bandwidth to minimize data packet loss due to network congestion, ensuring robust data transmission.
- *Secure Controller and Communication:* The central controller and the control channel employed for data communication are secured using standard protocols such as Transport Layer Security (TLS), safeguarding against unauthorized access and data breaches.
- *Robust Encryption and Authentication:* The encryption and authentication mechanisms in place are considered secure against breaches (*e.g.*, encryption that cannot be easily broken and digital signatures that cannot be forged).

Threat Model

The endogenous security system is designed with the following threat models in mind for PV data transmission and storage systems:

- *Focus on Core System Security:* The primary concern is securing the core components of the PV's communication network. Threats related to peripheral or "edge" components, such as individual PV modules or local inverters, fall outside the immediate scope of this research.
- *Malicious Component Manipulation:* There is a risk of malicious components within the network due to compromised hardware or software vulnerabilities that result in the manipulation of data packets or flow rules.
- *Exclusion of Certain Network Attacks:* Attacks related to protocols, such as TCP/IP or OSPF, are considered beyond the scope of this model.

Endogenous Security Model for PV Data Transmission

The endogenous security system for data transmission within PV uses enhanced routing strategies and a robust VFM to secure the data communication network. The model employs three routing approaches: (i) Data Integrity Forwarding (DIF), which is for establishing accuracy and data consistency through data comparison from multiple paths and allows only the verified data; (ii) Load-Balanced Forwarding (LBF), handles the congestion by distributing data along different paths that were selected using predefined load capacity and (iii) Path Diversity Forwarding (PDF) is employed to select data transmission path from route pools to ensure randomness. The VFM employs the system to identify security attacks and anomalies by recognizing communication errors.

For processing, time management, and data path security and integrity validation, the structure includes the following segments:

Enhanced Path Computing (EPC)

By examining the current network state and security hazards, EPC uses one of three algorithmic routing methods to determine the most efficient and secure routes for data transmission.

Dynamic Path Scheduling (DPS)

Implementing real-time variations in security and network circumstances, this DPS unit objectives data transfers on demand.

Path Authentication and Feedback Verification (PAFV)

This PAFV unit focuses on the data transfer route to identify malicious use behavior and information errors.

Data Path Validation Checking (DPVC): By verifying what is received to the implied structure, this DPVC unit validates the packets of data integrity.

The operational flow of the processed model is presented below:

- Upon initiating data transmission, the proposed system evaluates which routing strategy is most appropriate based on current network conditions and security protocols.
- The controller then calculates the optimal paths for routing by employing the EPC module to respond to the network's needs adaptively.

Depending on The Routing Strategy Selected

- DIF routes duplicate data packets through different paths for cross-verification at the destination.
- LBF distributes data across available paths in alignment with their capacity to ensure a balanced load.
- PDF routes data packets via randomly selected paths to obfuscate the transmission pattern and enhance security.

Data packets undergo a final verification process at their destination to confirm their integrity. Any detected irregularities are marked as anomalies.

The Packet Verification Process Is Illustrated As Follows
Initiation of Probe Packet
The central control system initiates the process by dispatching a probe packet.

Verification Mark Generation
As the probe packet is transmitted, the first switch in its path generates two types of marks based on the flow's characteristics and the packet's unique hash value:

- *Verification Flow Rule Mark (e):* This mark verifies flow against the predefined security rules based on flow information, flow entry point, and packet hash value.
- *Verification Data Content Mark (t):* This mark validates the data content's integrity.

Intermediate Switch Processing
As the packet moves through subsequent switches in the network:

- The verification information from the preceding switch (S_{i-1}) is encrypted with a key (K_i) and incorporated into the current switch's (S_i) Verification data. This layered approach ensures that each switch adds its unique verification mark to the packet.
- Similarly, data verification information is cumulatively embedded and updated at each switch, with the current switch's data verification mark appended to the packet.

Final Verification and Routing Decision
Upon reaching the destination, the switch collects multipath information from various ports and forwards it to the control system. The control system then:

- A consistent comparing and ruling algorithm is used to evaluate the data from different paths. Key comparison metrics include the *Flow_ID, Datapath_ID, Buffer_ID,* and the data message's hash value.
- The VFM scrutinizes the flow rules and data if discrepancies are detected. This facilitates precise identification and swift rectification of any anomalies in the paths.
- In cases where data paths need to be quickly changed, the system is designed to efficiently reroute specified data over alternate switch ports.

Enhanced Path Computing (EPC) and Dynamic Path Scheduling (DPS)
The control system identifies the optimal data transmission paths across the PV by deploying discovery packets, which map the entire network layout. Upon receipt of the initial data packet at a network node N_0, this node signals the control system, initiating the path optimization process. The objective of this path determination and optimization process is to select paths that minimize the number of intersecting nodes while satisfying specific transmission standards, EQU (1)
Minimize the intersection set

$$\sum_{(p_i,p_j)\in P,i<j} \left|I(p_i,p_j)\right|, \tag{1}$$

where for every pair of paths p_i and p_j in the intersection set $I(p_i,p_j)$ from source N_{src} to destination N_{dst}, each is part of the flow paths subject to constraints:

Constraints
Node Diversity Requirement
For any selected group of paths within the possible flow sets, no single node should be standard across all paths within the group, ensuring that compromised nodes cannot affect the entirety of the data transmission, EQU (2)

$$\forall p \in P, \forall n \in p, "n_{\text{end}} " \notin p', \exists p' \in P \setminus p \tag{2}$$

Bandwidth Limitation
The collective bandwidth of the selected multipath set must not exceed a predefined bandwidth threshold, ensuring adequate data flow without congestion, EQU (3)

$$\sum_{(p\in P, l\in p)} bw_p^l \leq B_{\text{total}} \tag{3}$$

Link Integrity Assurance
The operational status of each selected link is marked as 1, while unselected links are marked as 0. This does not account for potential link failures, EQU (4)

$$\forall l_{\text{active}} \in L, l_{\text{active}} = \{1 \text{ if } l \in p, 0 \text{ otherwise } \} \tag{4}$$

Node Count Restriction
Within any selected path set, the count of nodes should not surpass the number of secure, operational nodes capable of proper data forwarding, EQU (5)

$$\forall g \in P, \left|N_g\right| \leq |N_{\text{secure}}|, \text{where } P \in \text{ Paths} \tag{5}$$

Transmission Latency Bound
The cumulative transmission delay across any selected path must not exceed a maximum delay, ensuring timely data delivery, EQU (6)

$$\sum_{l \in p}\ t_l \leq T_{\max}, \forall p \in P \tag{6}$$

Data Path Validation and Selection
Upon data initiation at a network node N_0, this node communicates the data flow details to the central control system. The control system logs the request details and computes multiple data paths from N_0 to the destination node N_d, and selects paths that minimize common points to enhance security and reliability. The transmission time can define the validation period for critical data flows where security outweighs latency concerns. Data packets from N_0 are duplicated and dispatched across selected paths based on predefined rules. For example, replication forwarding might occur from N_0 to N_1, N_2, and N_3 by specified actions. The decision on which paths to use considers the delay requirements, with the residual packets directed consequently to ensure timely delivery to N_d.

Upon arrival at N_d, the data packets are aggregated, and a validation message is relayed to the control system. A consensus mechanism validates the data, where packets sharing identical flow identifiers are considered. The network will consider these authenticated channels as secured routes for subsequent communications if the data is verified precisely. The system uses corrective methods like switching and route reconfiguration if differences evolve. The system will start a security alert once all the required measures have been taken to deal with the errors. This warning will be used to clean up and repair affected routes.

Path Authentication and VFM
The following methods are used to develop this system to identify and mitigate assaults.

Key Distribution and Flow Identification

- Utilizing asymmetric cryptography, the controller maintains a key pair ($K_{\text{public}}, K_{\text{private}}$) and distributes individual session keys (K_{session}) to network nodes (N_i).
- A unique flow identifier (FID) is generated for each data flow based on its characteristics:

$FID = HashFn(PortEntry \parallel FlowHeader)$

Here, $HashFn$ denotes a secure hash function, PortEntry is the input path for the data flow, and Flow Header includes dangerous header details like source and destination identifiers.

Probe Packet Dispatch and Response

- When the starting node N_0 meets data d missing a match, it computes $Hash(d)$, forwarding this along with d to the controller. The controller, in response, techniques a payload encompassing FID, a flow header, a protection timestamp TS, and a signature $\text{Sig}_{K_{\text{private}}}$ $(FID \parallel TS)$ to ensure authenticity and temporal integrity.
- The first node N_0 then propagates d along with its Hash $(d), FID, TS$, and $S_{ig_{K_{\text{prinate}}}}$ to the next node in the sequence, N_1.

Validation and Encryption at Intermediate Nodes

- At each node N_i, based on a predefined Validation Flag (VF), the data may undergo authentication to verify the integrity of the transmitted signature and the associated time stamp, ensuring that the VF has not been altered.
- An encryption function at the node N_1 appends an authentication tag Tag_{N_1} to d, computed as: $\text{Tag}_{N_1} = \text{MAC}$ $C_{K_{\text{session}}}(\text{PortEntry}\ (N_1) \parallel \text{AuthPayload})$ where $\text{AuthPayload} = FID \parallel TS$, and MAC is the message authentication code generated using the session key.

Data Forwarding and Comprehensive Verification

- Successive nodes N_i determine the necessity for authentication via VF, continuously appending verification tags to ensure end-to-end integrity.
- Upon arrival at the terminal node N_n, a consolidated packet containing d, Tag_{N_n}, and Hash (d) is dispatched to the controller for final validation, employing a comprehensive check against the original flow details.

Controller's Final Authentication and Feedback

- The controller executes a thorough verification for packets sharing an FID, leveraging the session keys to authenticate the flow's transmission path and data content.
- This step facilitates precisely identifying anomalous or compromised nodes, enabling swift corrective actions to reestablish secure data pathways.

The following algorithm presents the steps in detail about the proposed security architecture.

Algorithm 1: Endogenous Security for PV Data Transmission
Input: Network topology, Data requests
Output: Secure and optimized data transmission paths
1 Initialize keyPair $(K_{\text{public}}, K_{\text{private}})$ for the controller.
2 Distribute K_{session} to all network nodes N_i.
3 Initialize FlowTable as an empty dictionary.
Enhanced Path Computing (EPC)
1. **For Each** data request DR in Data requests, **Do**
   1.1. $FID = HashFn(DR.PortEntry \parallel DR.FlowHeader)$
   1.2. $Paths = DiscoverPaths(DR.Source, DR.Destination)$
   1.3. $OptimalPaths = ComputeOptimalPaths(Paths)$
   1.4. $FlowTable[FID] = OptimalPaths$
Dynamic Path Scheduling (DPS)
2. **For Each** FID in FlowTable.keys() **Do**
   2.1. $SelectedPath = SelectPathBasedOnStrategy(FlowTable[FID])$
   2.2. If $strategy == DIF$ then
      2.2.1. $PerformDIF(SelectedPath)$
   2.3. Else, if $strategy == LBF$, then
      2.3.1. $PerformLBF(SelectedPath)$
   2.4. Else
      2.4.1. $PerformPDF(SelectedPath)$
Path Authentication and Feedback Verification (PAFV)
3. **For Each** path in SelectedPath, **Do**
   3.1. $ProbePacket = GenerateProbePacket(K_{public}, K_{private}, FID)$
   3.2. Send ProbePacket through the path.
   3.3. **For Each** node N_i in path **Do**
      3.3.1. $Tag_{N_i} = AuthenticateNode(N_i, ProbePacket, K_{session})$
      3.3.2. Append Tag_{N_i} to ProbePacket.
   3.4. $Feedback = CollectFeedback(ProbePacket)$
   3.5. If $VerifyFeedback(Feedback)$ is False then
      3.5.1. Alert and recompute OptimalPaths excluding compromised path.
4. **For Each** path in SelectedPath, **Do**
   4.1. $DataTransmission(path)$
(i) Function $PerformDIF(path)$

- Copy and forward packets through different paths
- Cross-verify at destination
- Forward only if data is consistent

(ii) Function $PerformLBF(path)$

- Distribute data across paths based on capacity
- Ensure balanced load

(iii) Function $PerformPDF(path)$

- Select paths randomly
- Increase unpredictability for attackers

(iv) Function $DiscoverPaths(source, destination)$

- Use LLDP or a similar protocol to discover all possible paths
- Return list of paths

(v) Function $ComputeOptimalPaths(paths)$

- Apply heuristic algorithms to find optimal paths minimizing node intersection
- Return optimal paths

(vi) Function $SelectPathBasedOnStrategy(paths)$

- Determine strategy based on current network conditions and security needs
- Return the selected path for data transmission

Key Management and Distribution

Let $RSA(K_{pub}, K_{priv})$ represent the RSA algorithm generating public key (PuK) and private key (PrK) pairs for asymmetric encryption, where K_{pub} denotes the PuK and K_{priv} signifies the corresponding PrK. Similarly, $AES(K_{\text{sym}})$ symbolizes the generation of symmetric keys (SyK) utilizing the AES-256 standard, with K_{sym} indicating the symmetric key.

Key distribution is facilitated through a secure channel, denoted as SC, which employs Transport Layer Security (TLS) protocols to ensure the confidential transfer of K_{sym} to network components. This process is as $SC_{TLS}(D_i, K_{sym})$, where D_i is the i^{th} device in the network receiving its SyK, K_{sym}.

The key management is integrated into CSP as follows:

DIF: For each data packet P_{data} encryption is applied using the symmetric key $Enc_{AES}(P_{data}, K_{\text{sym}})$ before transmission, this encrypted data ensures that integrity checks during DIF are performed on secure content, thereby preserving data confidentiality and integrity across the transmission paths.

LBF and PDF: The selection of paths for LBF and PDF is predicated on the availability of secure communication channels. Let P_{opt} represent the optimal path selected through the heuristic evaluation of path security and network conditions, formulated as SelectPath $(D_i, K_{\text{sym}}, \text{NetState}) \rightarrow P_{\text{opt}}$, where NetState encapsulates the current network state, including load and security posture.

VFM: The integrity of feedback and probe packets, P_{feedback}, is secured through digital signatures $\text{Sig}_{RSA}(P_{\text{feedback}}, K_{\text{priv}})$, ensuring the authenticity and non-repudiation of the feedback sent to the control system for anomaly detection and system adjustments.

The Key Management System (KMS) automates the processes of key generation, distribution, rotation, and revocation within the model as $KMS(K_{pub}, K_{priv}, K_{sym}, SC_{TLS})$. This system ensures that cryptographic keys are dynamically managed in response to network events, security incidents, or predefined schedules, enhancing the resilience of the communication infrastructure.

V. IMPLEMENTATION AND EVALUATION

Implementation Details

Hardware Environment

The hardware setup consists of commercial-grade solar panels rated at 300 *W* peak power connected to a central inverter with a capacity of 10 *kW*. The network infrastructure is built using Cisco Catalyst 2960-X Series Switches. Each component within the PV is interfaced with Raspberry Pi-4 Model B devices for real-time data processing and encryption.

Software Environment

The control system software is developed on the Node-RED platform, and the security model is encoded using Python 3.8 and Scapy libraries for packet manipulation and PyCryptoDome for cryptographic functions. For simulation, the NS-3 network simulator is employed. **Table 2** presenting the configuration of the system setup:

Table 2. System Configuration

Component	Specification/Tool	Description/Function
PV Modules	300W peak power	High-efficiency solar panels capturing SE.
Inverter	10kW capacity	Converts DC to AC power, equipped with network interfaces for secure data communication.
Data Loggers	Raspberry Pi 4 Model B	Collects and records system performance data, equipped with secure transmission capabilities.
Network Devices	Cisco Catalyst 2960-X Switches	Facilitates encrypted data communication within the PV system.
Control System	Node-RED platform	Manages monitoring, data flow, and implementation of security algorithms.
Security Software	Python 3.8, Scapy, PyCryptoDome	Executes the endogenous security model, including cryptographic functions and packet manipulation.
Simulation Tools	NS-3	Models the PV and network infrastructure for testing under simulated conditions.

Performance Metrics

The effectiveness of the endogenous CPS is quantitatively assessed using the following metrics, each represented with corresponding formulas to ensure precise evaluation, EQU (7) to EQU (10).

Data Transmission Integrity (I)

$$I = \frac{N_{\text{correct}}}{N_{\text{total}}} \times 100\% \tag{7}$$

where N_{correct} is the number of data packets received accurately at the destination, and N_{total} is the total number of data packets sent.

System Resilience to Attacks (R)

$$R = 1 - \frac{T_{\text{attuct}} - T_{\text{normal}}}{T_{\text{normal}}} \tag{8}$$

here, T_{attack} represents NT under attack conditions and T_{normal} signifies system NT under normal operation.

Network Throughput (T)

$$T = \frac{D_{\text{total}}}{t} \tag{9}$$

where D_{total} is the total data transmitted over the network in a specific timeframe, and t is the time taken.

Latency (L)

$$L = t_{\text{destination}} - t_{\text{source}} \tag{10}$$

with $t_{\text{destination}}$ being the time a packet is received at its destination and t_{source} the time it was sent from its source. This measures the delay introduced by security protocols.

The above metrics were compared against the attacks such as (i) Data Flow Manipulation, (ii) DoS Attacks, (iii) Spoofing Attacks, and (iv) Path Compromise. The proposed models' performance was compared against the works, and the results analysis for the above metrics are discussed below:

The transmission integrity analysis of the compared models was done using different packet sizes and attacks. The **Fig 4** shows the transmission integrity of the compared models for different types of attacks. The proposed model shows better transmission integrity scores ranging from 97% to 99% compared to other models. The compared models show limited performance, particularly for DoS and Spoofing attacks, for which the proposed model displayed better performance.

The **Fig 5** shows the comparison of transmission integrity against different packet sizes. The proposed model performs better with 96% and 99% integrity rates for all packet sizes tested ranging from 500 to 2000 *bytes*. The performance decreased as the packet size increased, which was visible across all models, but even then, the proposed model showed better performance.

The analysis of "Resilience to Attacks" (R) across different security models is shown in **Fig 6**. For Data Flow Manipulation and Path Compromise attacks, the proposed model maintained a resilience rate of 95%, which is close to normal conditions. For DoS Attacks, the proposed model shows a higher resilience rate of 96%, which is a better rate considering it is a high-intensity threat. For Spoofing Attacks, the proposed model had shown a resilience rate of 94%. Among the compared models, Isozaki et al. showed resilience rates ranging from 88% to 91%, with the lowest resilience for Path Compromise attacks and the highest for Spoofing Attacks. Zhang et al.'s model showed resilience rates between 90% and 93%, which was balanced compared to the other models except the proposed model.

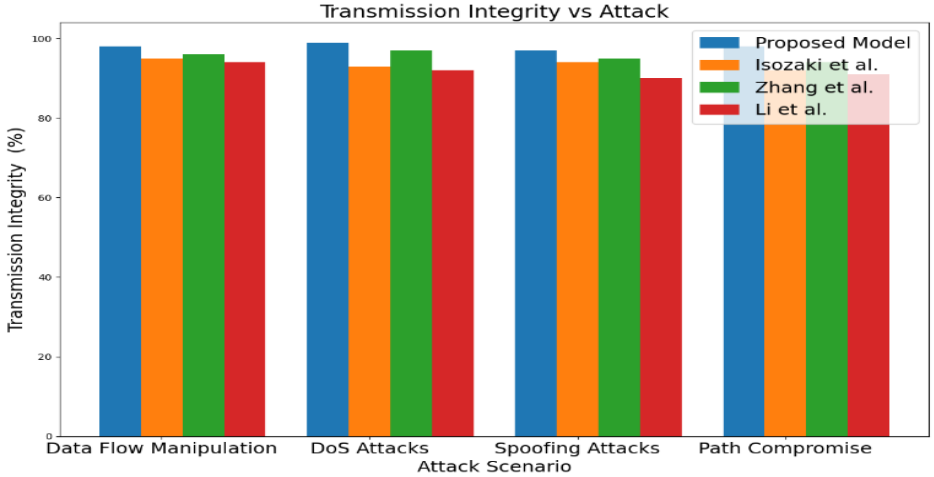

Fig 4. Transmission Integrity Vs Attack.

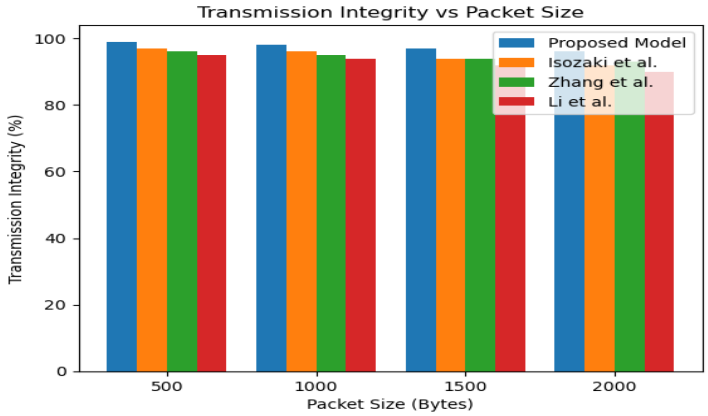

Fig 5. Transmission Integrity vs Packet Size.

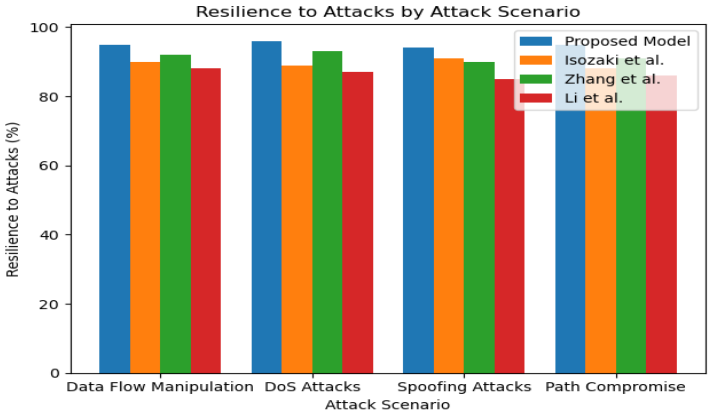

Fig 6. Resilience vs Attacks.

The evaluation of "Network Throughput" (*T*) across various security models is shown in **Fig 7**. Under normal operation, the proposed model recorded an NT of 10 Gbps, which is better than other models with Zhang et al.'s model is the next model reaching 9.8 Gbps. For DoS Attacks, the NT for the proposed model decreases to 9.2 Gbps, which is followed by Zhang et al.'s model that achieved 8.5 Gbps; in the case of Spoofing attacks, the proposed model achieved NT of 9.6 Gbps, and for Path Compromise situations, the proposed model demonstrates an NT of 9.7 Gbps. Out of all the models, the one next to the proposed model was Zhang et al., and Li et al. was the least-performing model.

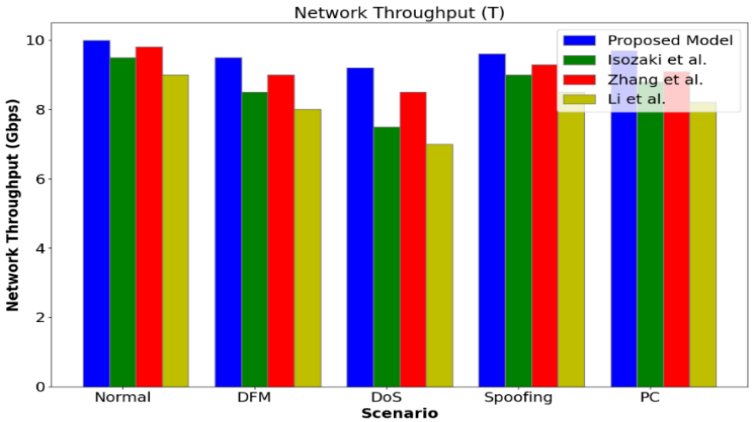

Fig 7. Throughput vs Attacks.

The analysis of "Latency" (*L*) is presented in **Fig 8** . In normal operations, the proposed model exhibits the lowest latency of 20 *ms*, Zhang et al. have shown a performance of 22 *ms*, and Isozaki et al. and Li et al. had shown higher latency of 25 and 28 *ms* respectively. For data flow manipulation, the proposed model shows 25, followed by Zhang et al. at 27 ms and Isozaki et al. at 30 ms, which comes at the last. A similar trend is understood across all attacks; the model had shown 30 ms latency for DoS attacks. For Spoofing Attacks, the proposed model shows lower latency at 23 *ms;* for Path

Compromise attacks, the proposed model exhibits a latency of 22 *ms*. Among all models, Zhang et al. come next to the proposed model, and Li et al. scored the lowest performance across all attacks.

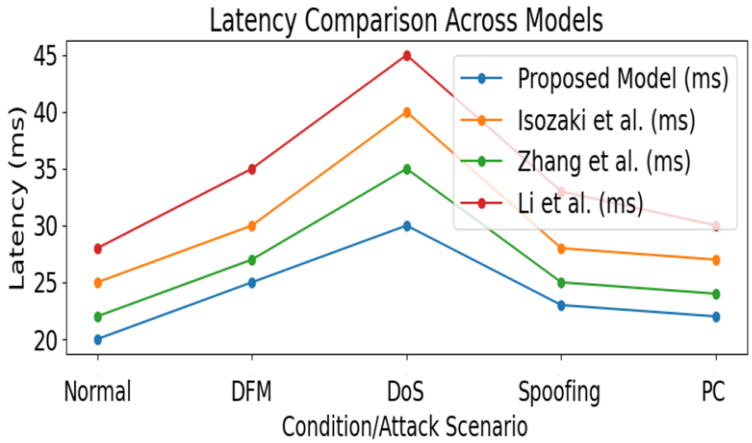

Fig 8.Latency vs Attacks.

Fig 9 shows the Packet Loss Rate (PLR) across different cyberattack scenarios. For the normal operation, the proposed model exhibits a packet loss rate of 0.2%, the lowest among the models, indicating high data transmission reliability. Isozaki et al. record 0.5%, Zhang et al. 0.4%, and Li et al. 0.6%. The proposed model maintains the PLR at 0.4% for the data flow manipulation. Isozaki et al. have a rate of 1.0%, Zhang et al. 0.8%, and Li et al. the highest at 1.2%. For the DoS attacks, the rates increase due to the attack's nature, with the proposed model at 1.0%, demonstrating resilience. Isozaki et al.'s PLR is 2.0%, Zhang et al.'s 1.8%, and Li et al.'s 2.5%. As for the spoofing attacks, the proposed model shows a 0.5% PLR, compared to Isozaki et al. at 0.9%, Zhang et al. at 0.7%, and Li et al. at 1.0%. The proposed model records a 0.3% PLR for the Path Compromise, indicating effective rerouting strategies. Isozaki et al. have 0.8%, Zhang et al. 0.6%, and Li et al. 0.9%.

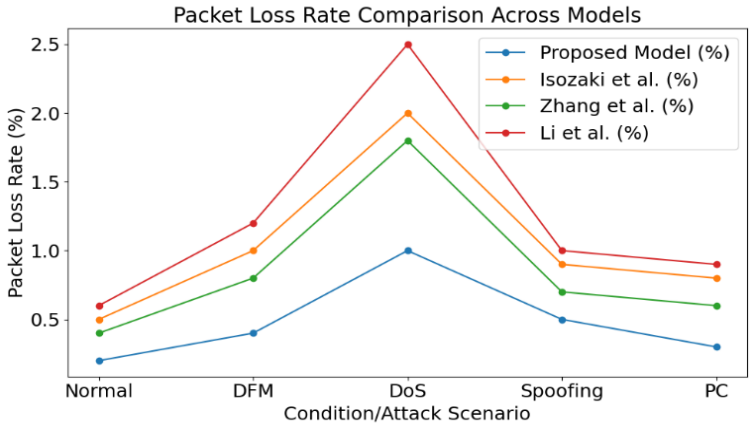

Fig 9. Packet loss vs Attacks.

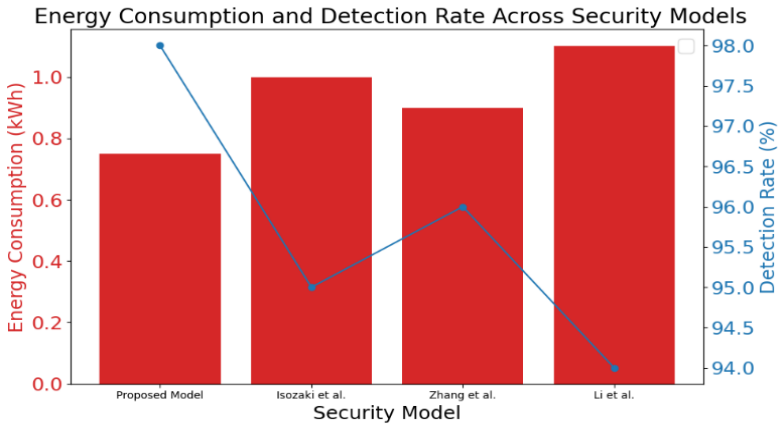

Fig 10. EC and Detection Rate (DR).

The EC and DR accuracy are analyzed in **Fig 10**. The proposed model demonstrates the lowest EC at 0.75 kWh. Zhang et al.'s model follows at 0.90 kWh, Isozaki et al.'s model consumes 1.00 kWh, and Li et al.'s model has the highest at 1.10 kWh. The proposed model achieves a 98% DR for the analysis, the highest among the compared models. Zhang et al.'s model has a DR of 96%, Isozaki et al.'s model records a 95% DR, and Li et al.'s model has the lowest at 94%.

VI. CONCLUSION AND FUTURE WORK

Due to weaknesses in digitalized PV operations, the move toward renewable energy sources (RES) has made protecting critical systems from cyberattacks more important. Using routing methods like DIF, LBF, and PDF, this research has developed an autonomous security system that protects data privacy while it is being sent and stored. A cryptographic key system improves network communication privacy. This study recommends models for data integrity, system reliability, latency, and network traffic. The proposed model does better in all of these ranges than competing models.

Future research should explore machine learning (ML) for predictive threat detection and expand the model to include numerous RES.

CRediT Author Statement

The authors confirm contribution to the paper as follows:

Conceptualization: Brahmadesam Viswanathan Krishna, Sathvik Bagam, Jayasri R, Krishnaveni N, Prasath R and Tanweer Alam; **Methodology:** Brahmadesam Viswanathan Krishna and Sathvik Bagam; **Software:** Jayasri R, Krishnaveni N, Prasath R and Tanweer Alam; **Data Curation:** Brahmadesam Viswanathan Krishna and Sathvik Bagam; **Writing-Original Draft Preparation:** Brahmadesam Viswanathan Krishna, Sathvik Bagam, Jayasri R, Krishnaveni N, Prasath R and Tanweer Alam; **Visualization:** Jayasri R, Krishnaveni N, Prasath R and Tanweer Alam; **Investigation:** Brahmadesam Viswanathan Krishna and Sathvik Bagam; **Supervision:** Viswanathan Krishna, Sathvik Bagam; **Validation:** Brahmadesam Viswanathan Krishna, Sathvik Bagam; **Writing- Reviewing and Editing:** Brahmadesam Viswanathan Krishna, Sathvik Bagam, Jayasri R, Krishnaveni N, Prasath R and Tanweer Alam; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

[1]. C. Breyer et al., "On the History and Future of 100% Renewable Energy Systems Research," IEEE Access, vol. 10, pp. 78176–78218, 2022, doi: 10.1109/access.2022.3193402.
[2]. S. Panneerselvam, S. K. Thangavel, V. S. Ponnam, and S. Sengan, "Federated learning based fire detection method using local MobileNet," Scientific Reports, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-82001-w.
[3]. A. Alshahrani, S. Omer, Y. Su, E. Mohamed, and S. Alotaibi, "The Technical Challenges Facing the Integration of Small-Scale and Large-scale PV Systems into the Grid: A Critical Review," Electronics, vol. 8, no. 12, p. 1443, Dec. 2019, doi: 10.3390/electronics8121443.
[4]. D. P. F. Möller, "Cybersecurity in Digital Transformation," Guide to Cybersecurity in Digital Transformation, pp. 1–70, 2023, doi: 10.1007/978-3-031-26845-8_1.
[5]. Mahalakshmi, R. L. Kumar, K. S. Ranjini, S. Sindhu, and R. Udhayakumar, "Efficient authenticated key establishment protocol for telecare medicine information systems," Industrial, Mechanical And Electrical Engineering, vol. 2676, p. 020006, 2022, doi: 10.1063/5.0117522.
[6]. S. Saeed, S. A. Altamimi, N. A. Alkayyal, E. Alshehri, and D. A. Alabbad, "Digital Transformation and Cybersecurity Challenges for Businesses Resilience: Issues and Recommendations," Sensors, vol. 23, no. 15, p. 6666, Jul. 2023, doi: 10.3390/s23156666.
[7]. J. Ye et al., "A Review of Cyber–Physical Security for Photovoltaic Systems," IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 10, no. 4, pp. 4879–4901, Aug. 2022, doi: 10.1109/jestpe.2021.3111728.
[8]. "An E-Commerce Based Personalized Health Product Recommendation System Using CNN-Bi-LSTM Model," International Journal of Intelligent Engineering and Systems, vol. 16, no. 6, pp. 398–410, Dec. 2023, doi: 10.22266/ijies2023.1231.33.
[9]. A. C. S. Robert Vincent and S. Sengan, "Effective clinical decision support implementation using a multi filter and wrapper optimisation model for Internet of Things based healthcare data," Scientific Reports, vol. 14, no. 1, Sep. 2024, doi: 10.1038/s41598-024-71726-3.
[10]. R. Lokeshkumar, O. Mishra, and S. Kalra, "Social media data analysis to predict mental state of users using machine learning techniques," Journal of Education and Health Promotion, vol. 10, no. 1, p. 301, 2021, doi: 10.4103/jehp.jehp_446_20.
[11]. G. Heilscher et al., "Integration of Photovoltaic Systems into Smart Grids Demonstration of Solar-, Storage and E-Mobility Applications within a Secure Energy Information Network in Germany," 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), pp. 1541–1548, Jun. 2019, doi: 10.1109/pvsc40753.2019.8980532.
[12]. U. Chadha et al., "Powder Bed Fusion via Machine Learning-Enabled Approaches," Complexity, vol. 2023, pp. 1–25, Apr. 2023, doi: 10.1155/2023/9481790.
[13]. R. K. Poluru and R. Lokeshkumar, "Meta-Heuristic MOALO Algorithm for Energy-Aware Clustering in the Internet of Things," International Journal of Swarm Intelligence Research, vol. 12, no. 2, pp. 74–93, Apr. 2021, doi: 10.4018/ijsir.2021040105.

[14]. B. R. R. Reddy and R. L. Kumar, “A Fusion Model for Personalized Adaptive Multi-Product Recommendation System Using Transfer Learning and Bi-GRU,” Computers, Materials & Continua, vol. 81, no. 3, pp. 4081–4107, 2024, doi: 10.32604/cmc.2024.057071.
[15]. S. Kunjiappan, L. K. Ramasamy, S. Kannan, P. Pavadai, P. Theivendren, and P. Palanisamy, “Optimization of ultrasound-aided extraction of bioactive ingredients from Vitis vinifera seeds using RSM and ANFIS modeling with machine learning algorithm,” Scientific Reports, vol. 14, no. 1, Jan. 2024, doi: 10.1038/s41598-023-49839-y.
[16]. N. Krishnadoss and L. Kumar Ramasamy, “A study on high dimensional big data using predictive data analytics model,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 30, no. 1, p. 174, Apr. 2023, doi: 10.11591/ijeecs.v30.i1.pp174-182.
[17]. A. L. Karn et al., “Fuzzy and SVM Based Classification Model to Classify Spectral Objects in Sloan Digital Sky,” IEEE Access, vol. 10, pp. 101276–101291, 2022, doi: 10.1109/access.2022.3207480.
[18]. N. Krishnadoss and L. K. Ramasamy, “Crop yield prediction with environmental and chemical variables using optimized ensemble predictive model in machine learning,” Environmental Research Communications, vol. 6, no. 10, p. 101001, Oct. 2024, doi: 10.1088/2515-7620/ad7e81.
[19]. P. Krishnamoorthy et al., “Effective Scheduling of Multi-Load Automated Guided Vehicle in Spinning Mill: A Case Study,” IEEE Access, vol. 11, pp. 9389–9402, 2023, doi: 10.1109/access.2023.3236843.
[20]. P. Selvam et al., “A Transformer-Based Framework for Scene Text Recognition,” IEEE Access, vol. 10, pp. 100895–100910, 2022, doi: 10.1109/access.2022.3207469.