Notification No.:591/2025

Date of Award:30/12/2025

Name of the Scholar: Qawsar Gulzar

Name of the Supervisor: Prof. Khurram Mustafa

Name of the Department: Computer Science, Faculty of Sciences.

Title of the Ph.D. Thesis: Intelligent Detection of Adversarial Cyber Attacks and Anomalies in Cyber-Physical Systems.

Findings

The integration of Cyber-Physical Systems (CPS) and the Internet of Things (IoT) has transformed modern infrastructures such as smart cities, healthcare, transportation, and industrial automation, but it has also significantly expanded the cyberattack surface. Real-time constraints, limited computational resources, and tight coupling between physical and cyber components make these systems vulnerable to complex, multi-vector attacks. Traditional security mechanisms are insufficient, while existing machine learning and deep learning–based intrusion detection systems (IDS) struggle with high-dimensional and imbalanced data, lack of interpretability, poor generalization to zero-day attacks, and deployment challenges on resource-constrained edge devices. These limitations motivate the need for adaptive, lightweight, interpretable, and intelligent IDS tailored to CPS and IoT environments.

This thesis develops a comprehensive suite of AI-driven IDS frameworks combining machine learning, deep learning, and reinforcement learning paradigms. The research begins with baseline ML models and introduces a hybrid IPCA–KNN/LR approach for efficient early-stage detection on CPS datasets such as SWaT. It then advances to attention-enhanced RNN, LSTM, and Bi-LSTM models capable of capturing temporal dependencies and improving interpretability, validated on SWaT, WADI, and GHL datasets. For IoT and edge environments, a lightweight RNN model with correlation-based feature selection is proposed. A robust hybrid architecture, DeepCLG, integrating CNNs, LSTM, GRU, Capsule Networks, and attention mechanisms, achieves superior multiclass detection on CICIoT2023 and UNSW-NB15 datasets. To further address class imbalance and intelligent decision-making, reinforcement learning–based frameworks using Q-learning and Double DQN optimize feature selection, minority class handling, and classification, with explainability supported through SHAP analysis.

Extensive evaluations across seven benchmark CPS and IoT datasets demonstrate that the proposed models achieve high accuracy, robustness, scalability, and interpretability while remaining suitable for real-time and edge deployments. By integrating dimensionality reduction, attention mechanisms, hybrid deep architectures, reinforcement learning, and explainable AI, this work addresses key challenges such as data imbalance, temporal dependency, resource constraints,

and transparency. The thesis concludes that a multi-paradigm AI approach significantly enhances IDS resilience and effectiveness and is extendable to decentralized and federated learning settings. Future research directions include adversarial training, continual and online learning, privacy-preserving federated IDS with differential privacy, blockchain-enabled trust mechanisms, and further optimization for edge deployment, paving the way for trustworthy, adaptive, and next-generation IDS for complex CPS and IoT ecosystems.

This thesis is structured into five chapters, each presenting a distinct yet interconnected contribution toward the development of intelligent, robust, and scalable cyberattack detection systems for CPS and IoT environments.

Chapter 1, entitled *"Hybrid Cyber-Attack Detection Model on Cyber-Physical Systems Using Machine Learning Techniques,"* introduces a hybrid IDS that integrates Incremental Principal Component Analysis (IPCA) with K-Nearest Neighbors (KNN) and Logistic Regression (LR). Evaluated on the SWaT dataset, this chapter demonstrates the effectiveness of dimensionality reduction for real-time CPS security and shows that the IPCA–KNN model outperforms IPCA–LR, achieving a precision of 0.997, recall of 0.996, and F1-score of 0.996.

Chapter 2, entitled *"Interdisciplinary Framework for Cyber-attacks and Anomaly Detection in Industrial Control Systems Using Deep Learning,"* proposes an attention-driven, lightweight deep learning framework using Deep RNN, LSTM, and Bi-LSTM models with feature selection. Extensive evaluations on SWaT, WADI, and GHL datasets shows that the attention-based Deep LSTM achieves superior detection performance and computational efficiency, highlighting its suitability for diverse industrial control system environments.

Chapter 3, entitled *"Enhancing Network Security in Industrial IoT Environments: A DeepCLG Hybrid Learning Model for Cyberattack Detection,"* presents the DeepCLG architecture, which integrates CNN, LSTM, GRU, and Capsule Networks with attention mechanisms to capture spatial, temporal, and hierarchical features. Tested on CICIoT 2023 and UNSW-NB15 datasets, DeepCLG consistently outperforms state-of-the-art methods, achieving high multiclass accuracy and low false alarm rates, thereby demonstrating strong adaptability in real-world IIoT scenarios.

Chapter 4, entitled *"Enhancing Industrial IoT Security Through Explainable Deep Reinforcement Learning Framework,"* introduces a deep reinforcement learning–based IDS that employs Double Deep Q-Networks (DDQNs) for synthetic minority sample generation and final classification, along with Q-learning for optimal feature selection. Evaluated on highly imbalanced datasets such as X-IIoTID and EdgeIIoTset, this framework achieves balanced and accurate intrusion detection, significantly improving performance on minority classes.

Chapter 5 concludes the thesis by synthesizing the findings across all proposed models, highlighting their effectiveness, robustness, and practical deployability. It also discusses the limitations of the current work and outlines future research directions, including scalability,

privacy-preserving learning, and advanced edge-based deployment strategies, thereby positioning the research within the broader context of next-generation CPS and IoT security.