LEMD-GARF: Lightweight Edge-based Malware Detection Using Genetic Algorithm Feature Selection and Random Forest Classification

Paper 565

Ali Rezazadeh, Majid Nezarat, Ali Kalate,
Hadi Shahriar Shahhosseini, and Amin Rezaei

Speaker Bio

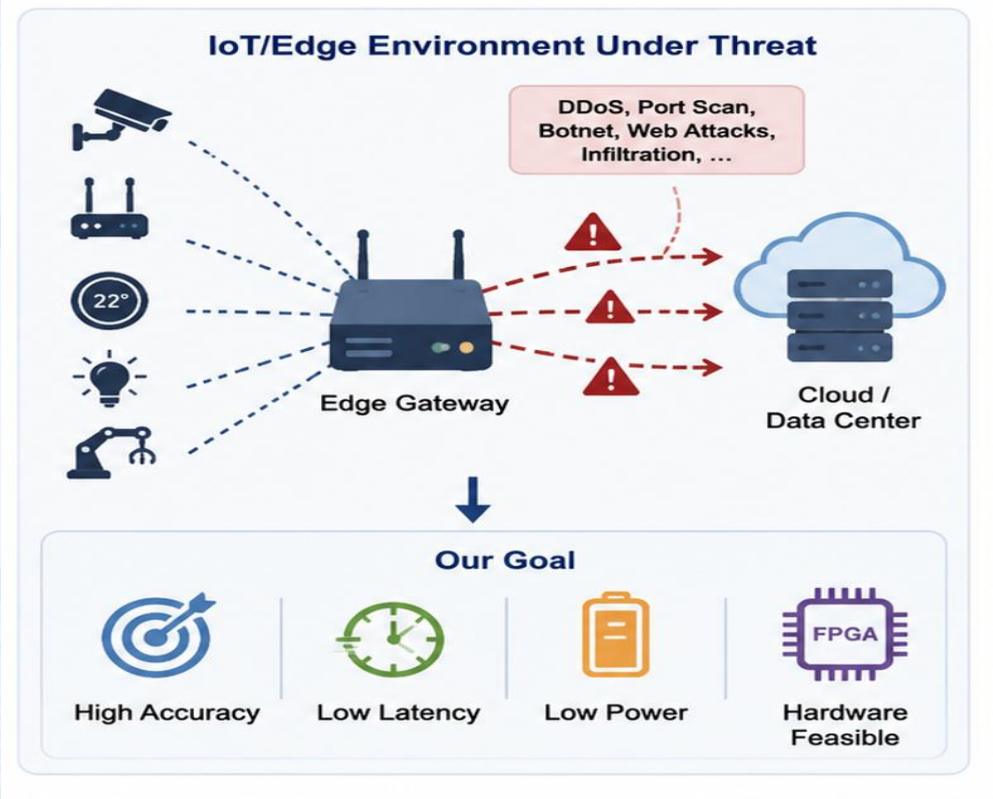
Dr. Amin Rezaei is an Associate Professor in the Department of Computer Engineering and Computer Science at California State University, Long Beach. He obtained his Ph.D. in Computer Engineering from Northwestern University. He has a decade of experience in hardware security, computer architecture, and machine learning, with more than 50 peer-reviewed scientific articles at flagship venues such as DAC, ICCAD, DATE, and ASP-DAC. He is a senior member of IEEE and a lifetime member of ACM and AAAI and has served on the technical program committees of many major conferences in his area.

Outline

- ✓ Introduction
- ✓ Related Work & Motivation
- ✓ Problem Statement & Contributions
- ✓ Proposed Framework
- ✓ Experimental Setup
- ✓ Results
- ✓ Conclusion & Future Work
- ✓ References

Introduction

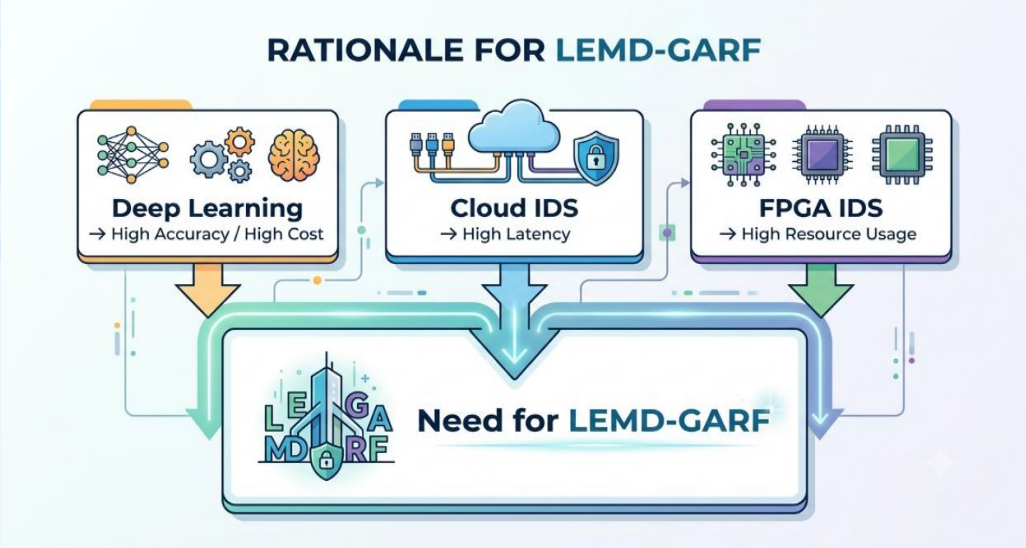
- The rapid growth of IoT and edge computing increases exposure to cyber threats.
- Cloud-based security solutions introduce latency, bandwidth overhead, and privacy concerns.
- Edge devices require lightweight and real-time malware detection mechanisms.



Related Work & Motivation

❑ **Existing Approaches**

- DL methods achieve high accuracy but require high computational resources.
- Cloud-based IDS solutions introduce communication delays.
- FPGA-based methods consume high hardware resources and power.



❑ **Research Gap:** Current solutions fail to simultaneously provide high accuracy, low latency, low power consumption, and efficient hardware implementation.

Problem Statement & Contributions

- **Problem Statement:** Detect malicious network traffic on resource-constrained edge devices while maintaining high detection performance and low hardware overhead.

- **Main Contributions**

- GA-based feature selection
- Random Forest malware classifier
- FPGA implementation on Xilinx XCU200
- Real-time detection with low power consumption

Proposed Framework

❑ **LEMD-GARF Workflow:**

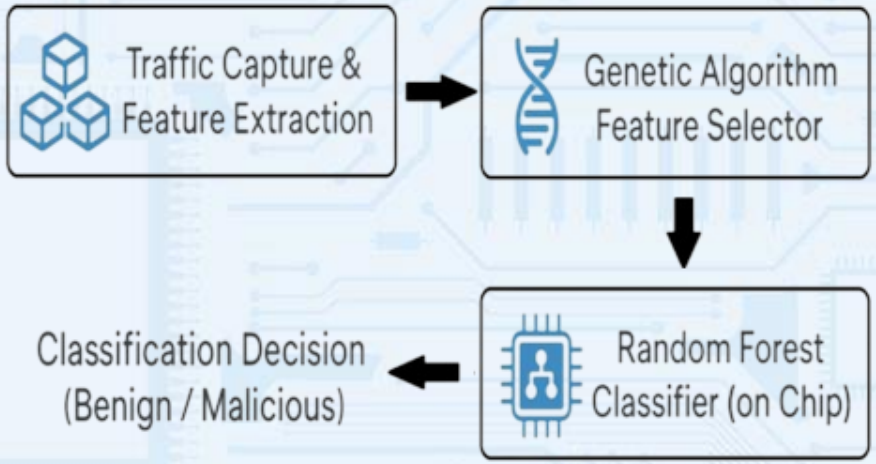
Raw Traffic →

Pre-processing →

GA Feature Selection →

RF Classification →

Malware / Benign



❑ **Key Idea:** Reduce redundant features before classification to improve efficiency and hardware feasibility.

Proposed Framework (Cont.)

❑ Mutual Information Filtering

Threshold = 0.01

❑ Genetic Algorithm

Population = 50

Generations = 30

Mutation = 0.2

Crossover = 0.5

❑ Result

80+ Features → 25 Selected

❑ Classification: (RF)

100 Trees

Gini Impurity

Majority Voting

❑ Takeaway

≈70% reduction in feature dimensionality

Proposed Framework (Cont.)

❑ Hardware Components

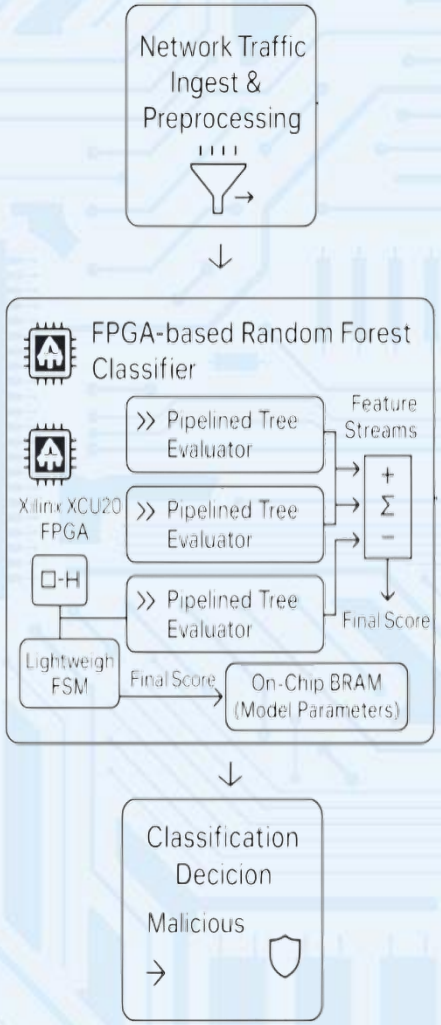
- Tree Evaluator
- BRAM Storage
- FSM Controller
- Voting Aggregator

❑ FPGA Target

- Xilinx XCU200

❑ Benefits

- Parallel execution
- Low latency
- Energy efficiency



Experimental Setup

- **Dataset**

  CIC-IDS2017

- **Attack Types**

  DDoS, PortScan, Botnet, FTP Patator, SSH Patator, Web Attacks

- **Evaluation Metrics**

  Accuracy, Precision, Recall, F1-score

- **Data Split**

  80% Training, 20% Testing

Results (Classification)

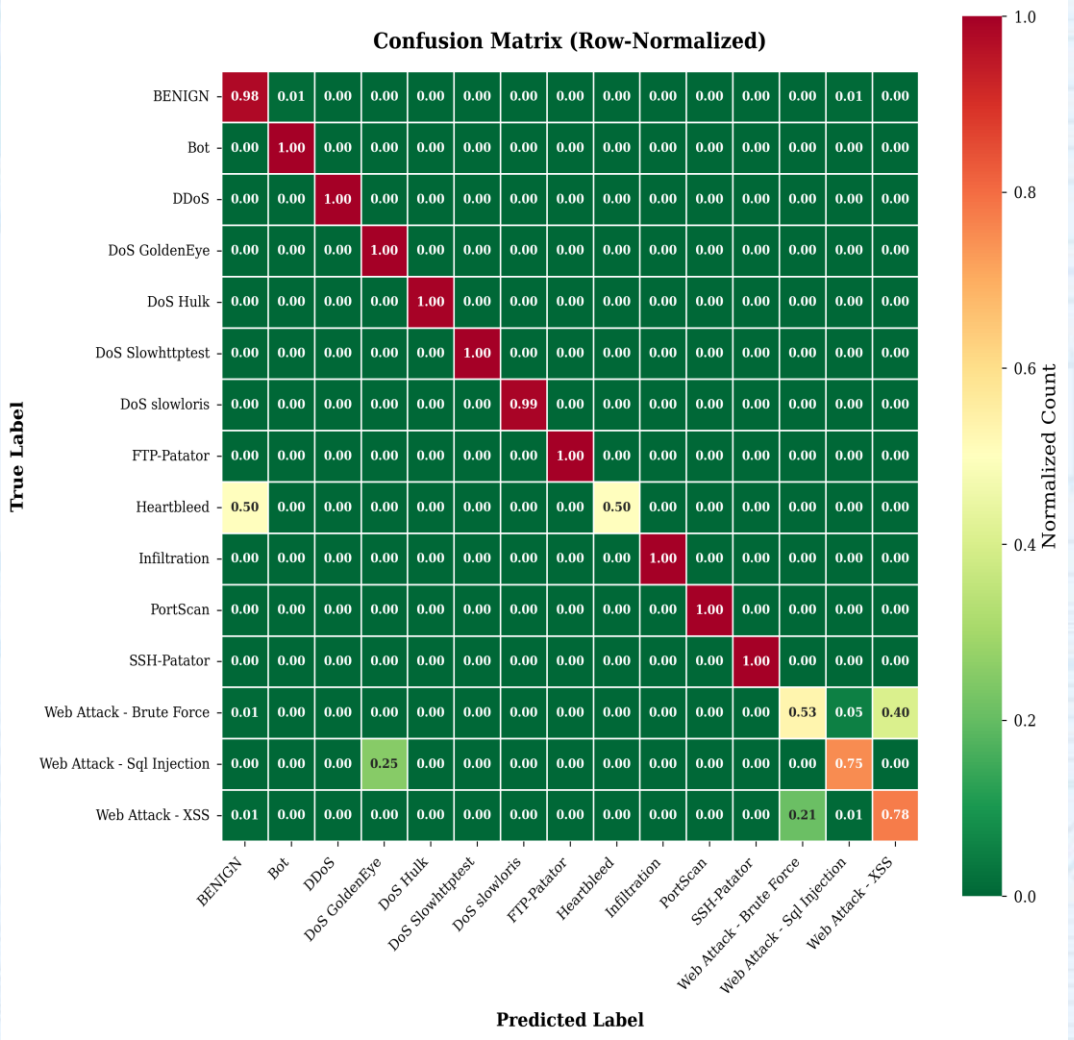
Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ZETROS [1]	99.83	99.70	99.75	99.72
PLLM-CS [2]	95-100	94-99	94-99	94-99
Hybrid ML for ICS [3]	99.10	98.80	98.70	98.75
FPGA-based Frequency Transformation [4]	97.72	94.43	99.18	96.75
LEMD-GARF (Proposed)	98.19	99.56	98.19	98.83

Note: High precision indicates low false alarm rates.

Results (Confusion Matrix)

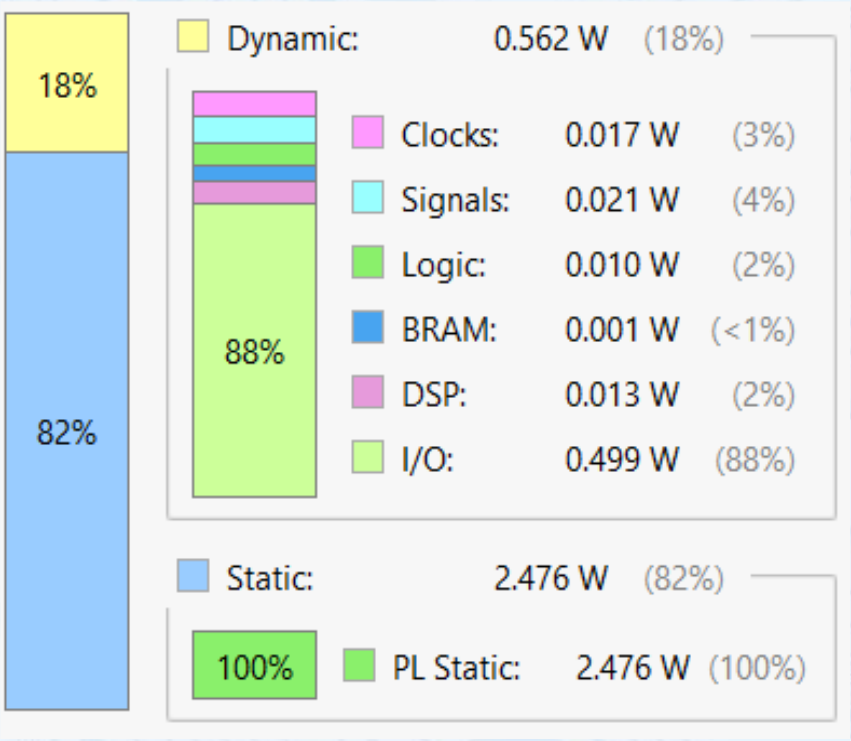
Strong Detection:
Bot, DoS, PortScan, FTP Patator, and SSH Patator

Challenging Detection:
Web Attack-Brute Force, Web Attack-XSS, SQL Injection, and Heartbleed



Results (FPGA Implementation)

Metric	Value
Frequency	143.22 MHz
Latency	0.189 µs
Throughput	5.29 MSamples/s
Power	2.476 W



Method	LUT	LUTRAM	FF	BRAM	DSP	IO	BUFG	Power(W)
LEMD-GARF (Proposed)	516	150	1483	0.5	19	81	1	2.476
FPGA-based Frequency Transformation [4]	88K	NR	NR	103.5	12	NR	NR	10.827

Conclusion & Future Work

❑ Conclusion

- GA reduced features from 80+ to 25
- RF achieved 98.19% accuracy
- FPGA implementation achieved 0.189 μs latency
- Power consumption limited to 2.476 W
- Suitable for real-time edge security

❑ Future Work

- Online learning capability
- Adaptive feature selection
- Deployment on heterogeneous FPGA platforms

References

[1] C. A. Baykara, I. Safak, and K. Kalkan, “ZETROS: A zero-trust IoT network security framework using distributed blacklisting, trust scoring and smart contracts,” Computer Networks, vol. 271, article 111601, 2025.

[2] M. Hassanin, M. Keshk, S. Salim, M. Alsubaie, and D. Sharma, “PLLM-CS: : Pre-trained Large Language Model (LLM) for cyber threat detection in satellite networks,” Ad Hoc Networks, vol. 166, issue C, 2025.

[3] M. M. Aslam, A. Tufail, L. C. De Silva, and R. Apong, “Multi-feature hybrid anomaly detection in ICS: An integration of ML, DL, and statistical techniques,” in ACM Workshop on Secure and Trustworthy Deep Learning Systems (SecTL), pp. 43-51, 2025.

[4] Z. Hu, H. I. Hasegawa, Y. Yamaguchi, and H. Shimada, “Enhancing detection of malicious traffic through FPGA-based frequency transformation and machine learning,” IEEE Access, vol. 12, pp. 2648–2659, 2024.

[5] N. Tekin, A. Acar, A. Aris, A. S. Uluagac, and V. C. Gungor, “Energy consumption of on-device machine learning models for IoT intrusion detection,” Internet of Things, vol. 21, article 100670, 2023.

[6] C. Sun, B. Chen, Y. Bu, S. Zhang, D. Zhang, and B. Jiang, “Lightweight traffic classification model based on deep learning,” Wireless Communications and Mobile Computing, vol. 1, article 3539919, 2022.

[7] A. Mudgerikar, P. Sharma, and E. Bertino, “Edge-based intrusion detection for IoT devices,” ACM Transactions on Management Information Systems, vol. 11, no. 4, 2020.

THANK YOU