



# Towards Efficient Machine Learning Method for IoT DDoS Attack Detection

Pavitra Modi

Contact Email: pmodi@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)

Acknowledgement: I would like to thank my undergraduate research project's supervisor Dr. Rongxing Lu, UNB and CIC.



## ABSTRACT

With the rise in the number of IoT devices and its users, security in IoT has become a big concern to ensure the protection from harmful security attacks. In the recent years, different variants of DDoS attacks have been on the rise in IoT devices. Failure to detect DDoS attacks at the right time can result in financial and reputational loss for victim organizations. These attacks conducted with IoT devices can cause a significant downtime of applications running on the Internet. Although researchers have developed and utilized specialized models using artificial intelligence techniques, these models do not provide the best accuracy as there is always a scope of improvement until 100% accuracy is attained. We propose a hybrid feature selection algorithm that selects only the most useful features and passes those features into an XGBoost model, the results of which are explained using feature importances. Our model attains an accuracy of 99.993% on the CIC IDS 2017 dataset and a recall of 97.64 % on the CIC IoT 2023 dataset. Overall, this research would help researchers and implementers in the field of detecting IoT DDoS attacks by providing a more accurate and comparable model.

## Proposed Feature Selection Algorithm

- **Pearson Correlation [1] Selection:** The feature selection algorithm finds out the Pearson correlation coefficient between all the features and the output label. The features with positive Pearson values greater than positive mean of Pearson values and the features with negative Pearson values less than negative mean of Pearson values are selected in this step.
- **Kendall [2] and Spearman [3] correlation selection:** The feature selection algorithm finds out the Pearson correlation coefficient between all the features and the output label. The algorithm finds out the Spearman and Kendall correlation coefficients for each feature, finds out the mean of both the correlation coefficients keeping the positive and negative means separate. By performing this, we now have two means: positive mean of Spearman and Kendall correlation values and negative mean of Spearman and Kendall correlation values. Features with Spearman and Kendall means greater than the above values are selected in this step.
- **Mutual Information Gain [10]:** Mutual Information gain values for each feature is found and mean of all values is also found. The features with mutual information gain values greater than average are selected.
- **Result:** The Set union of the features selected in Step 1 and 2 is found into set *final*. The result set of selected features is found by the set intersection of features in the set *final* and the features in Step 3.

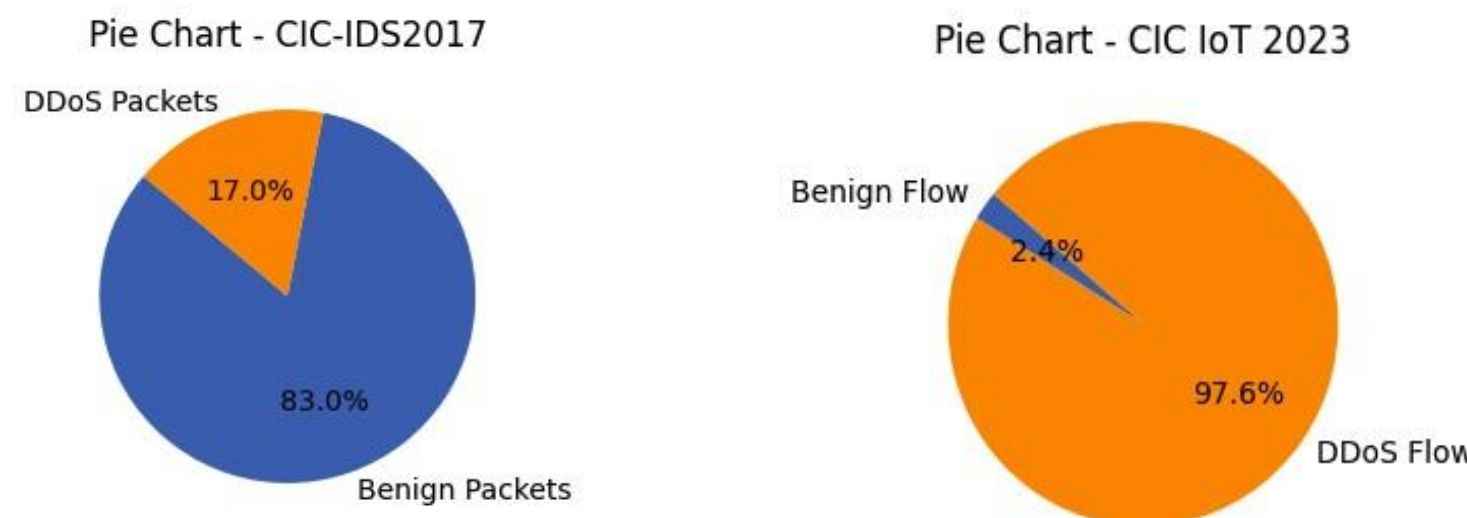
## Discussion

- The features selected are passed into different models as seen in the Experimental results section. The XGBoost [8] Model resulted in giving superior accuracy in CIC IDS 2017 dataset of 99.993 % compared to [4,5,6] and it resulted in a superior recall value of 97.64 % compared to [7] for the CIC IoT 2023 dataset.
- To understand the features contribute the most to the success of the model, feature importances are shown below.

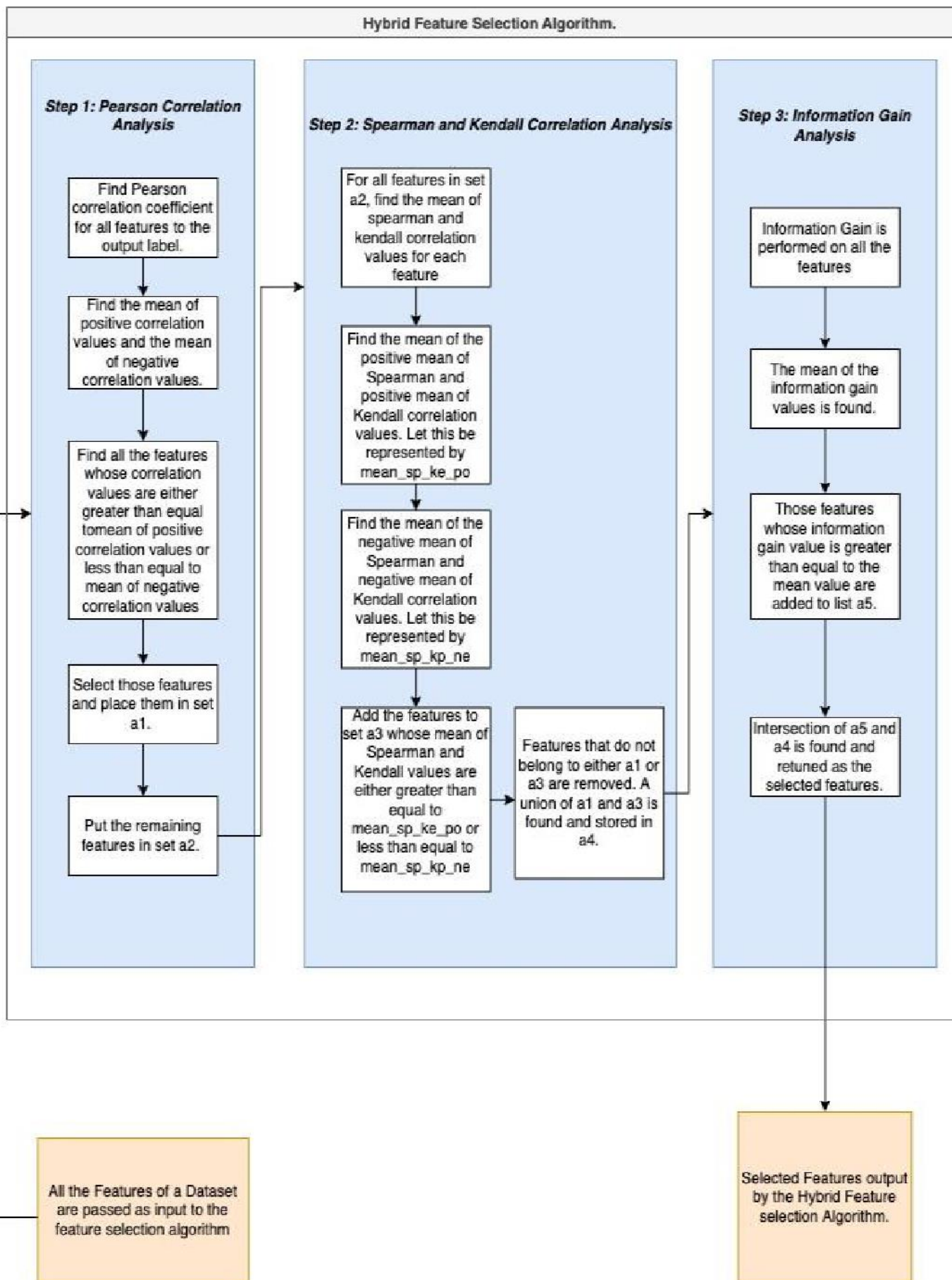
## References

- [1] K. Pearson, "On the theory of contingency and its relation to association and normal correlation." (1904).
- [2] M. G. Kendall, "A NEW MEASURE OF RANK CORRELATION", *Biometrika*, Volume 30, Issue 1-2, June 1938, Pages 81–93, <https://doi.org/10.1093/biomet/30.1-2.81>
- [3] C. Spearman, "The Proof and Measurement of Association between Two Things." *The American Journal of Psychology*, vol. 100, no. 3/4, 1987, pp. 441–71. *JSTOR*, <https://doi.org/10.2307/1422689>. Accessed 18 July 2024.
- [4] U. S. Chanu, K. J. Singh, and Y. J. Chanu, "A dynamic feature selection technique to detect ddos attack," *Journal of Information Security and Applications*, <https://www.sciencedirect.com/science/article/abs/pii/S2214212623000303> (accessed Jul. 21, 2024).
- [5] D. Kshirsagar & S. Kumar, "Towards an intrusion detection system for detecting web attacks based on an ensemble of filter selection techniques", *CyberPhysical Systems*, 9(3), 244–259, <https://doi.org/10.1080/23335777.2021.2023651>
- [6] R. Doshi, N. Aphorpe and N. Feamster, "Machine Learning DDoS Detection for Consumer Internet of Things Devices," *2018 IEEE Security and Privacy Workshops (SPW)*, San Francisco, CA, USA, 2018, pp. 29–35, doi: 10.1109/SPW.2018.00013.
- [7] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensor* (2023)
- [8] T. Chen and C. Guestrin. "Xgboost: A scalable tree boosting system." *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016.
- [9] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", *4th International Conference on Information Systems Security and Privacy (ICISSP)*, Portugal, January 2018.
- [10]J. R. Quinlan, "Induction of decision trees." *Machine learning* 1 (1986): 81–106.

## Dataset Distribution



## Proposed Method

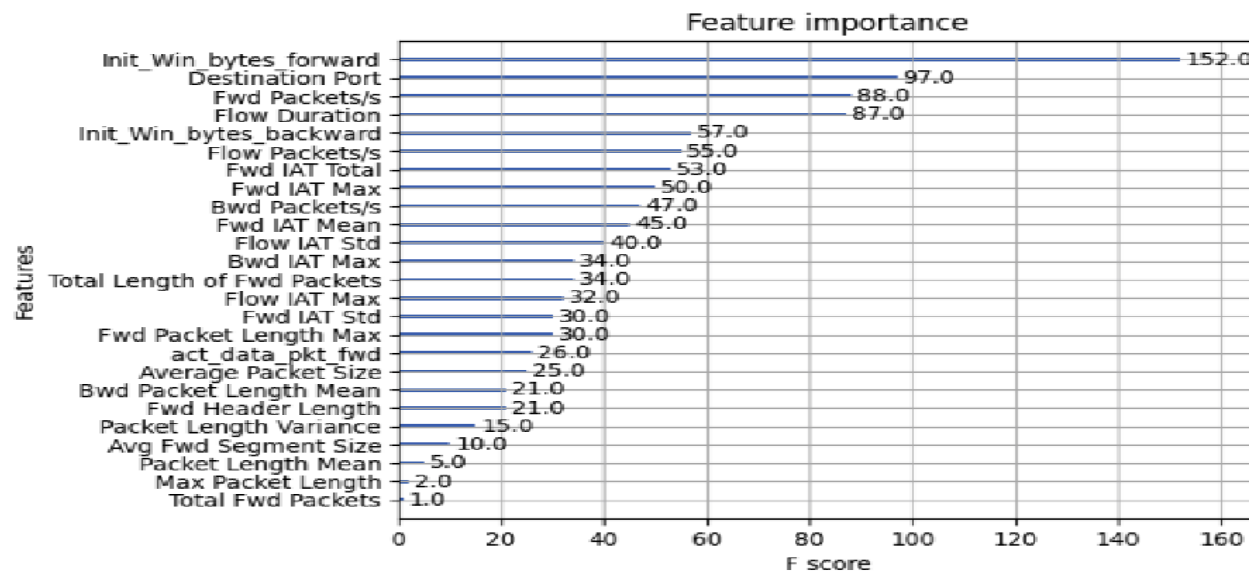


## Feature Importance and Experimental Results

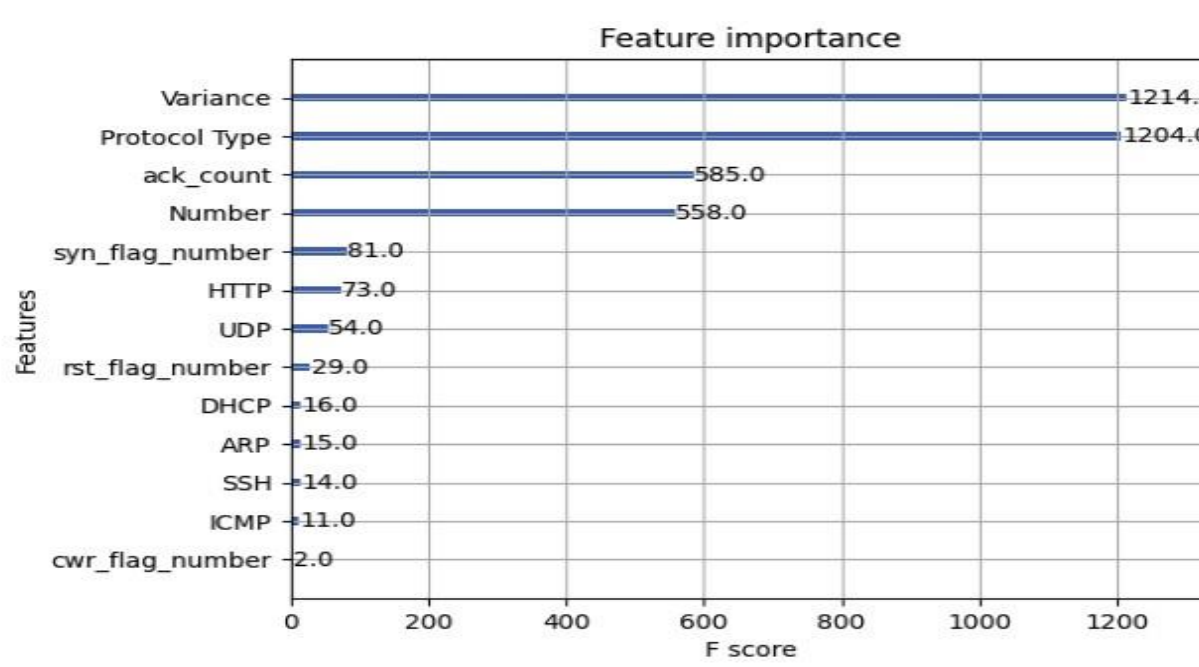
Metrics	Random Forest	Decision Trees	XGBoost	Linear SVM	KNN
Accuracy	99.989 %	99.986 %	99.993 %	88.516 %	99.867 %
Precision	99.989 %	99.986 %	99.993 %	92.962 %	99.867 %
Recall	99.989 %	99.986 %	99.993 %	88.516 %	99.867 %
F1-Score	99.989 %	99.986 %	99.993 %	83.530 %	99.867 %
Training Time	52.94 second s	5.13 second s	1.95 second s	139.37 second s	78.82 second s

Part A: Experimental Results – CICIDS 2017 and CIC IoT 2023

Metrics	Random Forest	Decision Trees	XGBoost
Accuracy	97.631%	96.528%	97.642%
Precision	95.45%	95.39%	95.33%
Recall	95.12%	96.52%	97.64%
F1-Score	95.28%	95.95%	96.47%
Training Time	3018.23 seconds	1185.16 seconds	34.196 seconds
Mean Squared Error	0.033	0.034	0.023



Part B: Feature Importance – CIC IDS 2017 and CIC IoT 2023







# IoT-LiteLine: Lightweight Pipeline for IoT Device Identification, Profiling and Monitoring

Alireza Zohourian, Dr. Sajjad Dadkhah, Dr. Ali Ghorbani

Contact Email: [alireza.Zohourian@unb.ca](mailto:alireza.Zohourian@unb.ca)

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)

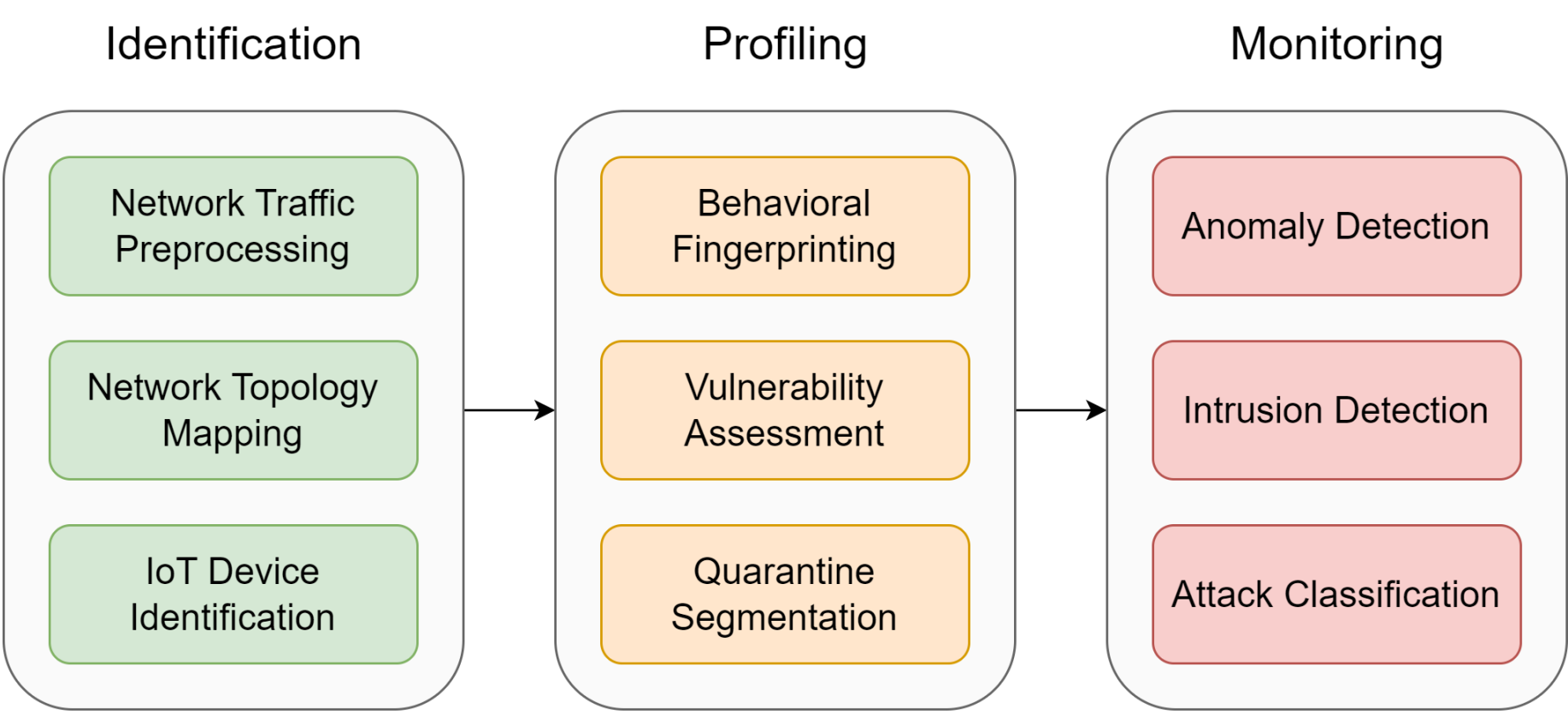


## ABSTRACT

Internet of Things (IoT) has become an inseparable part of human life. It is constantly being adopted in different use cases such as healthcare, transportation and critical infrastructure for real-time monitoring and automation. Although IoT helps organizations and manufacturers to increase efficiency and productivity, it comes with several drawbacks such as cybersecurity. The extensive adoption of IoT and the huge deployment of IoT devices lead to a complex network of low-power devices that mostly do not have enough computation capabilities to support sufficient security measures. Consequently, there is a need for an extra layer of protection for these vulnerable devices. Therefore, we propose a Lightweight Pipeline for IoT device identification, profiling and monitoring (IoT-LiteLine) that identifies IoT devices in a network, builds baseline profiles for each device and constantly monitors each individual device for behavior that is different from the baseline profile.

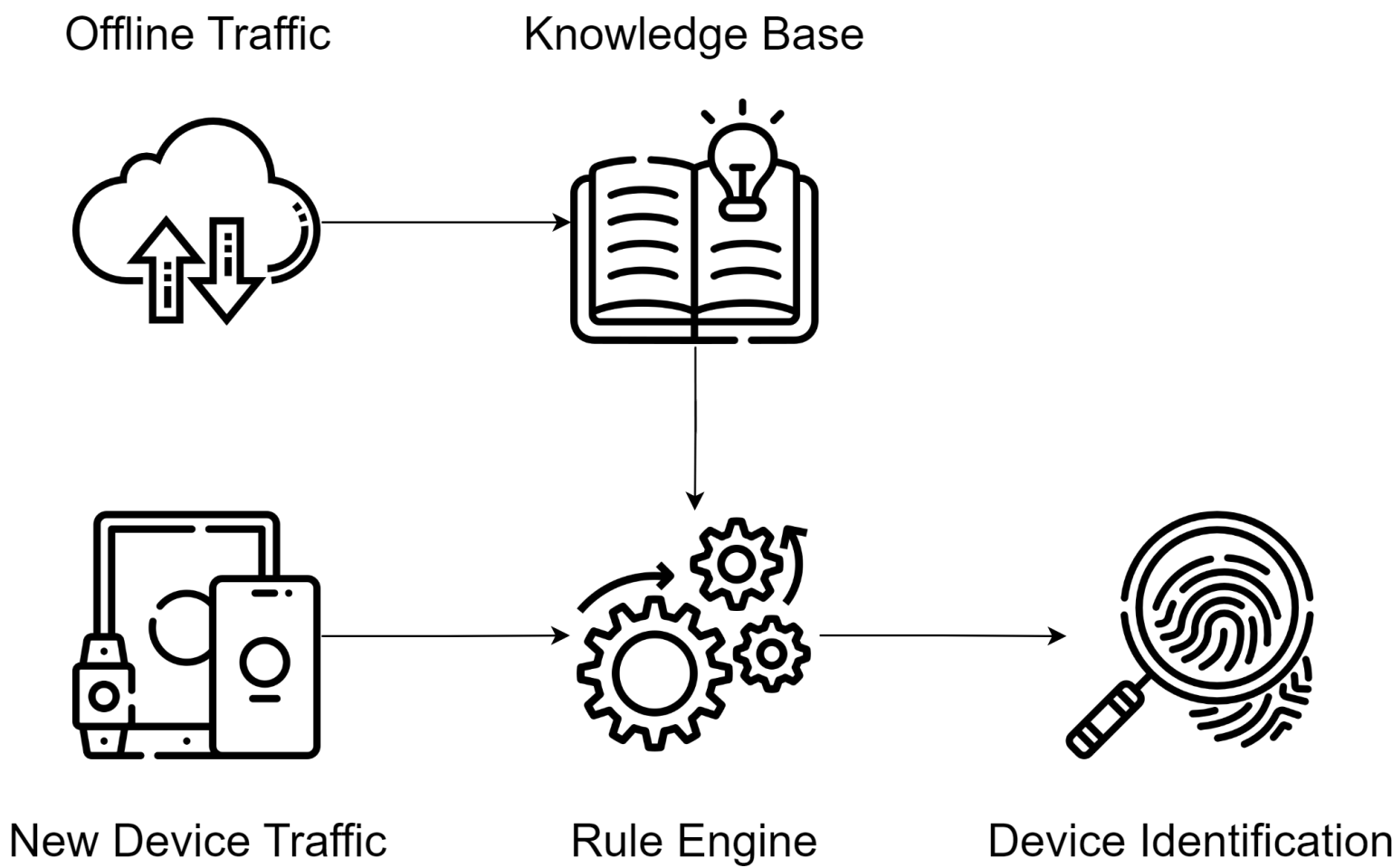
## Model Design

The proposed framework is a pipeline that identifies IoT devices and tracks their behaviour in real-time and includes three steps: Identification, Profiling and Monitoring. During identification, the network traffic of the devices are preprocessed for network topology mapping and IoT device Identification. This stage identifies all IoT devices on the network and gives a visual network topology, separating IoT devices from Non-IoT ones. Next, it extracts behavioral fingerprints to create a baseline profile for each individual device and performs automated vulnerability assessment to further segment/quarantine the IoT device. Lastly, during the monitoring stage, the framework constantly monitors the devices and uses the baseline profile to look for any change in the devices' behavior for further anomaly detection, intrusion detection or attack classification.



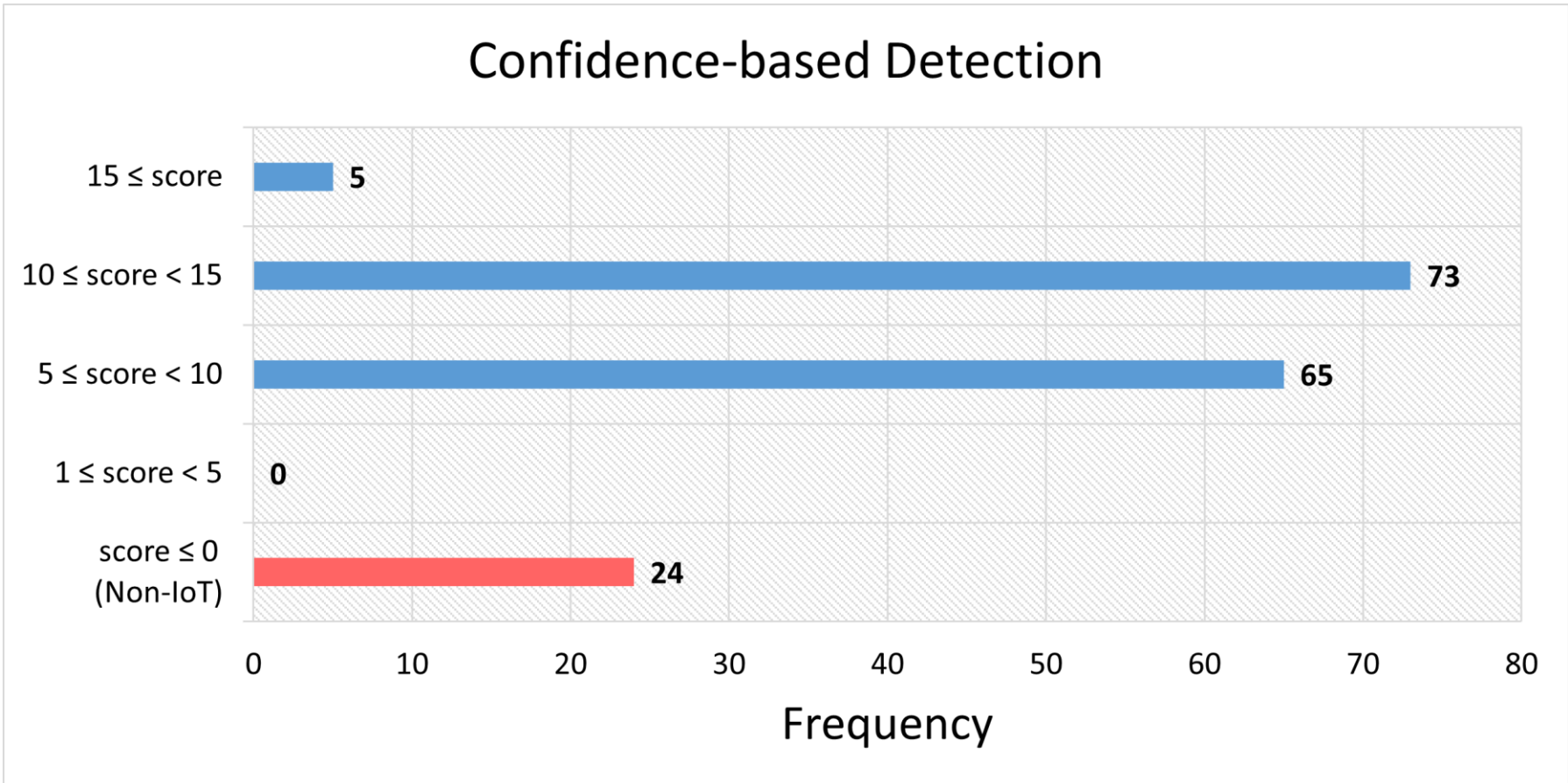
## Identification

The IoT device identification module (RIoT-DevIde) is a rule-based model that uses specific information from packet headers to identify IoT devices. The knowledge base uses specific packet header information from offline IoT and non-IoT traffic to find signatures for IoT and non-IoT devices and extracts rules for the rule engine. For a new device, a short initial traffic is passed on to the rule engine and identification is accomplished.



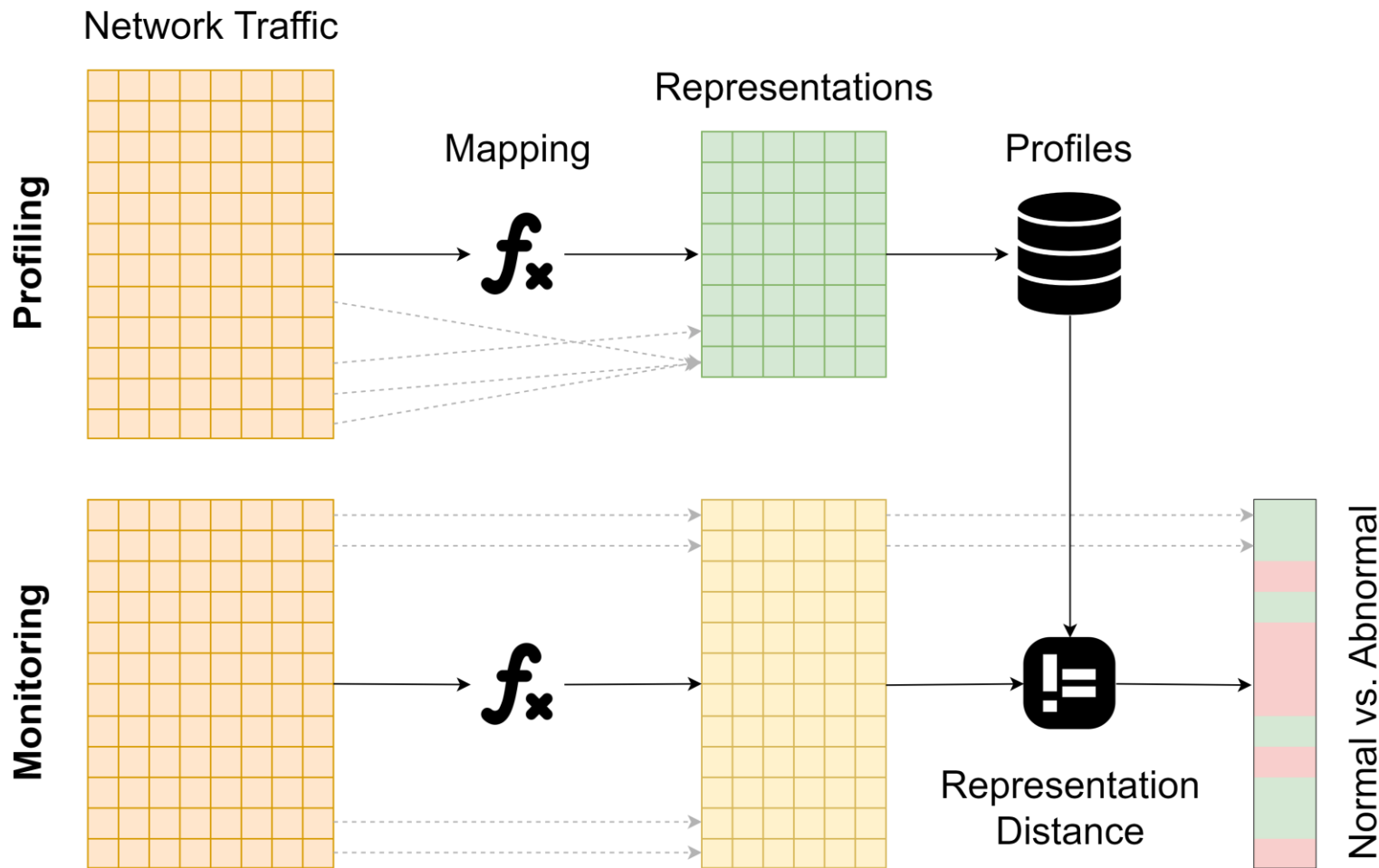
## Experimental Results

After extracting a total of 10 rules, we employed them in a confidence-based approach in which each triggered rule contributes to the identification to some extent, resulting in an identification score. All IoT and non-IoT devices were correctly identified in an average of less than 5 seconds of network traffic. Furthermore, the majority of IoT devices were identified with a high confidence score.



## Profiling and Monitoring

The IoT profiling and monitoring leverage a lightweight intrusion detection system (IoT-PRIDS) that maps the benign packets to “representations” and stores them as device profiles. Later, every incoming packet is mapped to its corresponding representation, and its distance from the representation set (device profile) is computed. Should the distance exceed a predefined threshold, the packet is identified as abnormal.



## Experimental Results

We tested our model on about 30 attacks from the CICIoT2023 dataset which specifically detected low-rate attacks such as web attacks with a very high accuracy and minimal runtime overhead.

Attack	Metrics	Packets		Flows	
		2022	2023	2022	2023
Backdoor Malware	Accuracy	0.9819	1.0000	1.0000	1.0000
	Precision	0.7674	1.0000	1.0000	1.0000
	Recall	0.9429	1.0000	1.0000	1.0000
	F1-Score	0.8462	1.0000	1.0000	1.0000
Browser Hijacking	Accuracy	0.9841	0.9815	0.9977	0.9983
	Precision	0.9055	0.4737	0.9851	0.9890
	Recall	1.0000	1.0000	1.0000	1.0000
	F1-Score	0.9504	0.6429	0.9925	0.9944
Command Injection	Accuracy	0.9852	0.9900	0.9998	1.0000
	Precision	0.5764	0.7391	0.9899	1.0000
	Recall	0.9235	1.0000	1.0000	1.0000
	F1-Score	0.7098	0.8500	0.9949	1.0000
SQL Injection	Accuracy	0.9127	0.9907	1.0000	1.0000
	Precision	0.1251	0.6087	1.0000	1.0000
	Recall	0.9044	1.0000	1.0000	1.0000
	F1-Score	0.2198	0.7568	1.0000	1.0000
Uploading Attack	Accuracy	0.9847	0.9917	0.9997	1.0000
	Precision	0.6501	0.7857	0.9889	1.0000
	Recall	0.9291	1.0000	1.0000	1.0000
	F1-Score	0.7650	0.8800	0.9944	1.0000
Cross Site Scripting	Accuracy	0.9881	0.9925	0.9998	1.0000
	Precision	0.5238	0.7333	0.9848	1.0000
	Recall	0.9308	1.0000	1.0000	1.0000
	F1-Score	0.6704	0.8462	0.9924	1.0000

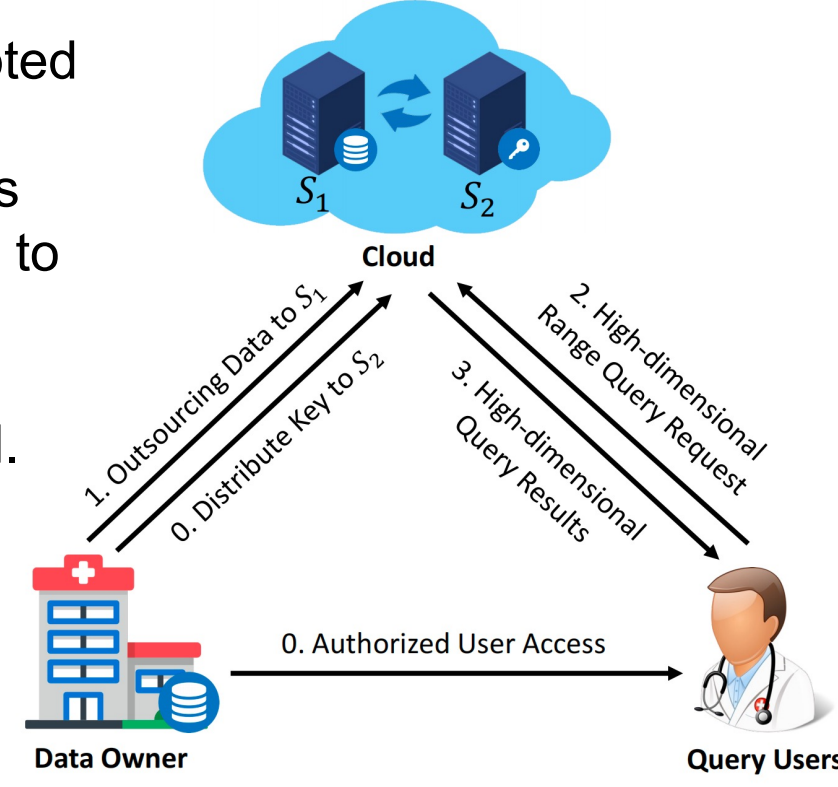


## ABSTRACT

The Internet of Things (IoT) boom has enabled Internet Service Providers (ISPs) to collect an enormous amount of high-dimensional data. Performing range queries on such data can effectively reuse them to help ISPs offer better services. Owing to the low cost and high resource utilization of cloud computing, an increasing number of ISPs are inclined to outsource data and services to it. However, as the cloud is not fully trusted, data need to be encrypted before being outsourced, which inevitably hinders many query services, e.g., range queries. Existing privacy-preserving range query schemes struggled to extend to high-dimensional scenarios and did not support dimension selection. Aiming at this challenge, we propose an efficient and privacy-preserving high-dimensional range query scheme (PHRQ) based on an iMinMax tree and symmetric homomorphic encryption (SHE) technique while supporting dimension selection. Security analysis and performance evaluation show that our scheme is privacy-preserving and efficient in high-dimensional range query processing.

## System Model

- Data Owner (DO): Outsourcing encrypted dataset.
- Two Cloud Servers: Storing ciphertexts and offering recommendation services to users.
- Users: The users can initiate POI recommendation requests to the cloud.



## Security Model

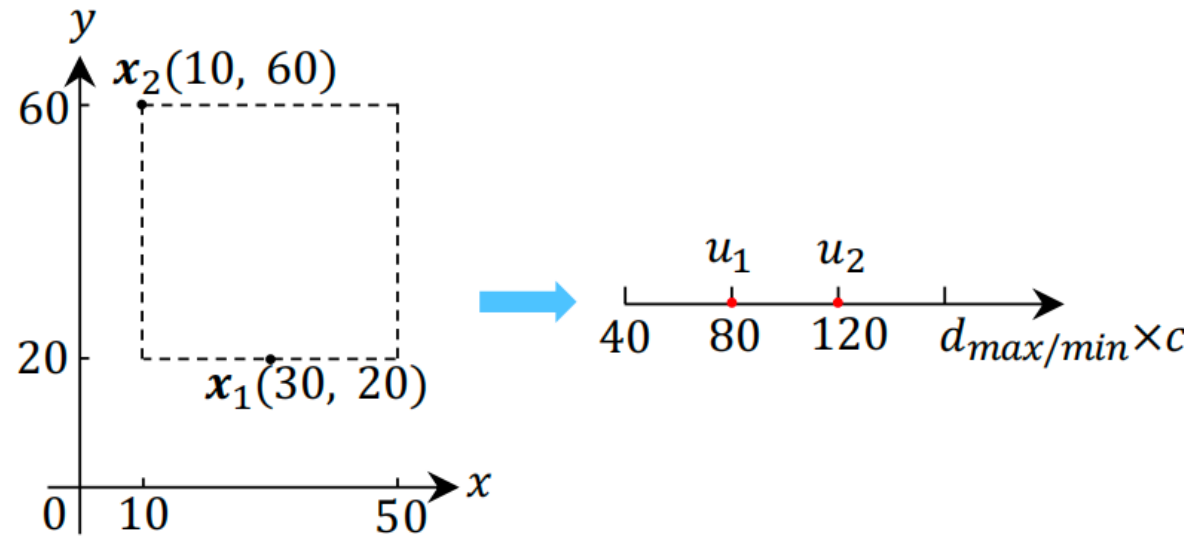
- DO: trusted;
- Users: honest;
- Cloud Servers: honest-but-curious.

## Our Proposed Scheme

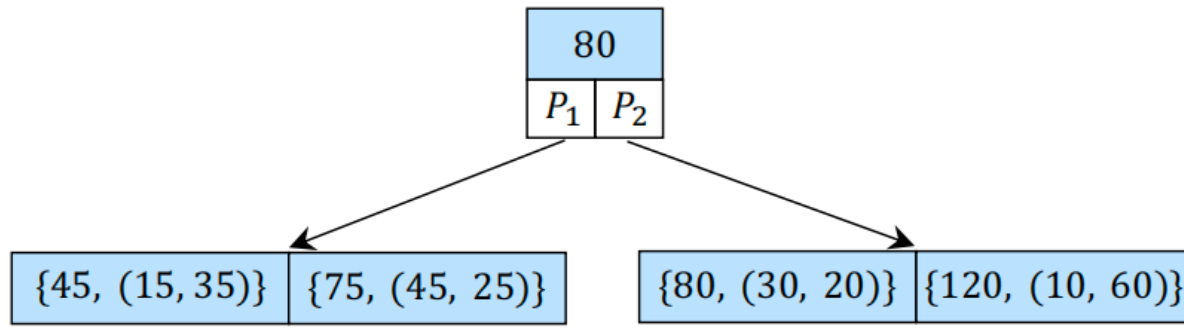
### 1. Data Outsourcing.

Let  $\mathcal{X} = \{\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{id}) | i = 1, 2, \dots, n\}$  be a high-dimensional healthcare dataset. The data owner outsources  $\mathcal{X}$  according to the following steps.

- Map high-dimensional records to one-dimensional index values.



- Index the index value dataset with a B+ tree to form an iMinMax tree.



- Encrypt the iMinMax tree with SHE.

The data owner encrypts each  $u_i$  in internal nodes to  $E(u_i)$ . For each  $\{u_i, \mathbf{x}_i\}$  in leaf nodes, he/she encrypts it to  $\{E(u_i), E(\mathbf{x}_i)\}$ . Then, the data owner sends the encrypted iMinMax tree  $E(T)$  to  $S_1$ .

### 2. Privacy-preserving Range Query.

Definition (Privacy-preserving High-dimensional Range Queries): Given an encrypted dataset  $E(\mathcal{X})$  and an encrypted query token  $E(Q)$ , a privacy-preserving high-dimensional range query is to find a set  $Result(Q)$  where each record satisfies  $E(\mathbf{x}_i) \in E(Q)$ .

- Map the high-dimensional range query  $(Q, \alpha)$  to one-dimensional intervals.

$$\begin{aligned} E(q'_j) &= [E(j * c + t_1 * (c - \hat{q}_{min1} - (q_{j1} - h_j)) + q_{j1} - h_j), \\ &E(j * c + t_2 * (q_{j2} - h_j - \hat{q}_{min2}) + q_{j2} - h_j)]. \end{aligned} \quad (1)$$

- Compute the encrypted mapped sub-dimensional range.

$$\begin{cases} [E(\phi_{q_{j1}}), E(\phi_{q_{j2}})] = E(q'_j); \\ [E(\psi_{q_{j1}}), E(\psi_{q_{j2}})] = [E(j * c), E((j + 1) * c)]. \end{cases} \quad (2)$$

$$E(\tilde{q}_j) = \begin{cases} [E(\phi_{q_{j1}}), E(\phi_{q_{j2}})] & \text{If } E(\alpha_j) = E(1); \\ [E(\psi_{q_{j1}}), E(\psi_{q_{j2}})] & \text{otherwise.} \end{cases} \quad (3)$$

- Range query algorithm based on the encrypted iMinMax tree.

**Algorithm 3:** High-Dimensional Query Over Ciphertexts.

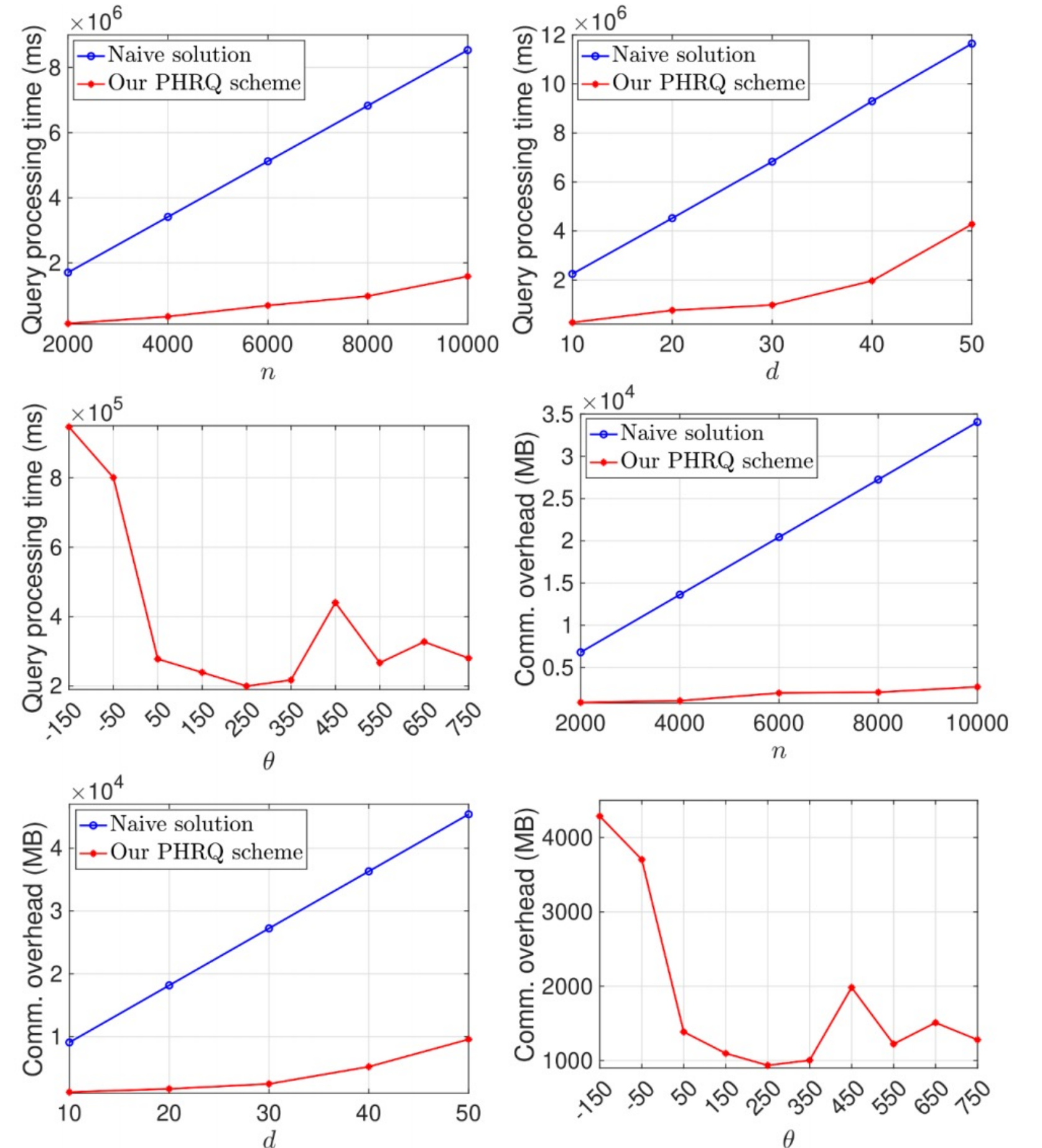
**Input:** The iMinMax tree  $E(T)$ ;  
The query token  $E(Q)$ ;  
The unpruned query  $E(W)$ ;  
**Output:** Result set  $Result(E(Q))$ ;  
// Filter  
1:  $C = \emptyset$ ; // Initialize the candidate set  
2: **for each**  $E(q'_j) \in E(W)$  **do**  
3:  $Result(E(q'_j)) = Filter(E(T).root, E(W))$ ;  
4:  $C = C \cup Result(E(q'_j))$ ;  
// Verification  
5: **for each**  $\{E(u_i), E(\mathbf{x}_i)\} \in C$  **do**  
6: Determine  $E(\mathbf{x}_i) \in E(Q)$  by MRD protocol;  
7: **if**  $E(\mathbf{x}_i) \in E(Q)$  **then**  
8:  $S_1$ : computes  $E(\mathbf{v}_i) = E(\mathbf{x}_i) + \mathbf{r}_i$ ;  
9:  $S_1 \rightarrow S_2$ :  $S_1$  sends  $\{E(\mathbf{v}_i), AES_{ack}(\mathbf{r}_i)\}$  to  $S_2$ ;  
10:  $S_2$ :  $\mathbf{v}_i = Dec(sk, E(\mathbf{v}_i))$ ;  
11:  $S_2$ :  $Result(E(Q)) = Result(E(Q)) \cup \{\mathbf{v}_i, AES_{ack}(\mathbf{r}_i)\}$ ;  
12: **return**  $Result(E(Q))$ ;

**Algorithm 4:** Filter(Node  $E(node)$ , Query  $E(W)$ ).

1: **if**  $E(node)$  is a leaf node **then**  
2: **for each** index value  $E(u_i)$  in  $E(node)$  **do**  
3: Determine  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  by ORD protocol;  
4: **if**  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  **then**  
5:  $Result(E(q'_j)) = Result(E(q'_j)) \cup \{E(u_i), E(\mathbf{x}_i)\}$ ;  
6: **else**  
7: **for each** index value  $E(u_i)$  in  $E(node)$  **do**  
8: // **Case 1:**  $E(u_i)$  is the last index in  $E(node)$   
9: Determine  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  by ORD protocol;  
10: **if**  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  **then**  
11:  $Filter(E(node.P_i), E(W))$ ;  
12:  $Filter(E(node.P_{i+1}), E(W))$ ;  
13: **else**  
14: Determine  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  by data comparison protocol;  
15: **if**  $E(u_i) < E(q'_{j1})$  **then**  
16:  $Filter(E(node.P_{i+1}), E(W))$ ;  
17: **else**  
18:  $Filter(E(node.P_i), E(W))$ ;  
19: // **Case 2:**  $E(u_i)$  is not the last index value  
20: Determine  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  by ORD protocol;  
21: Determine  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  by data comparison protocol;  
22: **if**  $E(u_i) \in [E(q'_{j1}), E(q'_{j2})]$  or  $E(u_i) > E(q'_{j2})$  **then**  
23:  $Filter(E(node.P_i), E(W))$ ;

## Performance Evaluation

- Computational Cost and Communication Overhead of the Query Processing Algorithm.



## Conclusions

We proposed an efficient and privacy-preserving high-dimensional range query scheme based on an iMinMax tree while supporting dimension selection. First, we constructed an iMinMax tree for high-dimensional data. Then, based on an SHE technique, we designed a suite of privacy-preserving protocols to guarantee the privacy of high-dimensional range queries. To support dimension selection, we converted sub-dimensional range queries to high-dimensional queries and designed a sub-dimensional range determination protocol to protect the privacy of sub-dimensional queries. Further, we proposed our PHRQ scheme. Security analysis and performance evaluation proved that our scheme is secure and efficient.

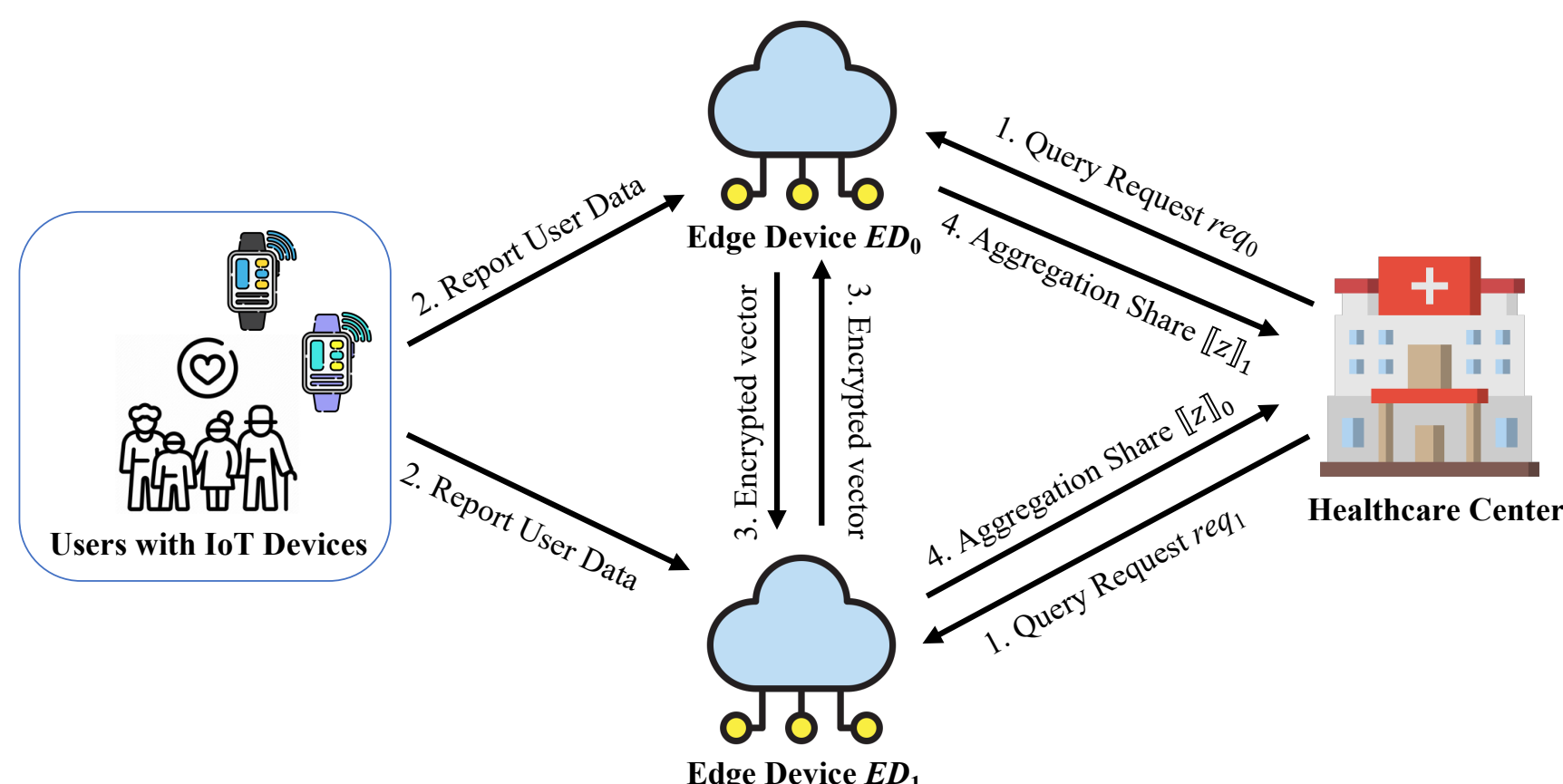


## ABSTRACT

Aging in place (AiP) has been adopted internationally as a response to population aging. Range aggregation is a fundamental method in AiP to enable the healthcare center to have a comprehensive view of health trends and better monitor the overall health in a given area. However, this aggregation process inevitably introduces security and privacy risks, drawing significant research attention. Existing privacy-preserving schemes supporting aggregation often fail to meet the specific needs of range aggregation in AiP or incur high computational costs. To address these challenges, we propose an efficient edge-based privacy-preserving range aggregation scheme for the AiP system. Our scheme employs the superincreasing sequence to ensure that users can obtain multiple types of aggregation results in a query and utilizes the one-time matrix encryption and the additive secret-sharing technique to safeguard sensitive information. Security analysis demonstrates that our proposed scheme preserves privacy during range aggregation. In addition, extensive experiments indicate its high efficiency.

## System Model

- Users with IoT Devices Server:** Each user periodically reports their health data to edge devices by using their IoT devices (e.g., wearable devices or other resource-constrained sensors).
- Edge Devices ( $\mathcal{ED}$ ):**  $\mathcal{ED}$  is situated on the edge of the network and can be regarded as a link between users and healthcare center.  $\mathcal{ED}$  is responsible for processing users' data sent by their IoT devices and offers the range aggregation query results to the healthcare center.
- Healthcare Center ( $\mathcal{HC}$ ):** To observe users' health status,  $\mathcal{HC}$  initiates a query with a range  $(\alpha, \beta)$  and sends it to  $\mathcal{ED}$ . In response,  $\mathcal{HC}$  can get three types (count, sum, average) of aggregation results from  $\mathcal{ED}$ .



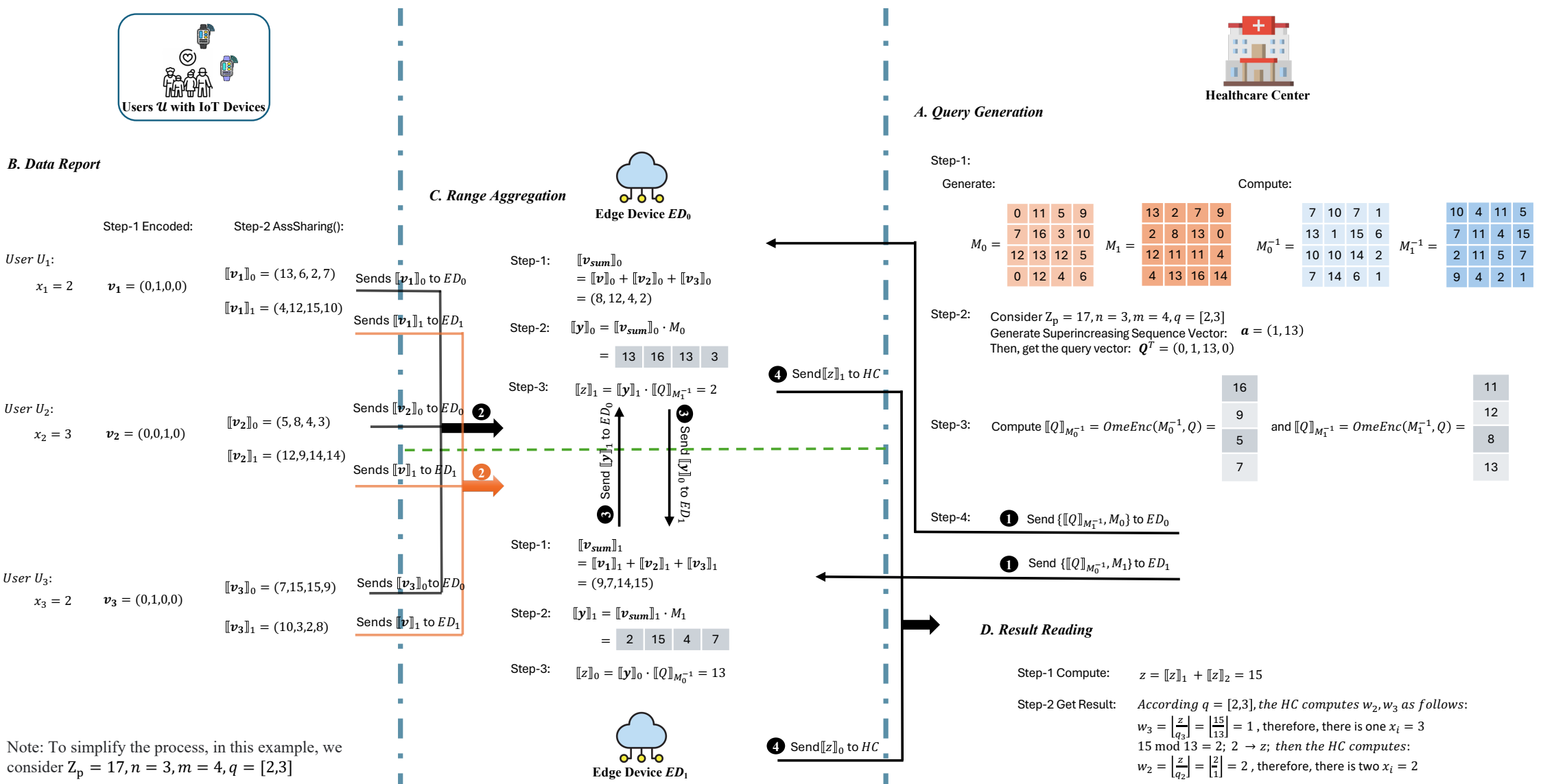
## Security Model

- All users to be honest, i.e., they faithfully follow the protocol and send data to edge devices.
- Healthcare center ( $\mathcal{HC}$ ) is viewed as honest, meaning that it precisely generates query requests.
- Edge devices ( $\mathcal{ED}$ ) are considered to be honest-but-curious, they are interested in both users' private data and  $\mathcal{HC}$ 's queries.

## Design Goals

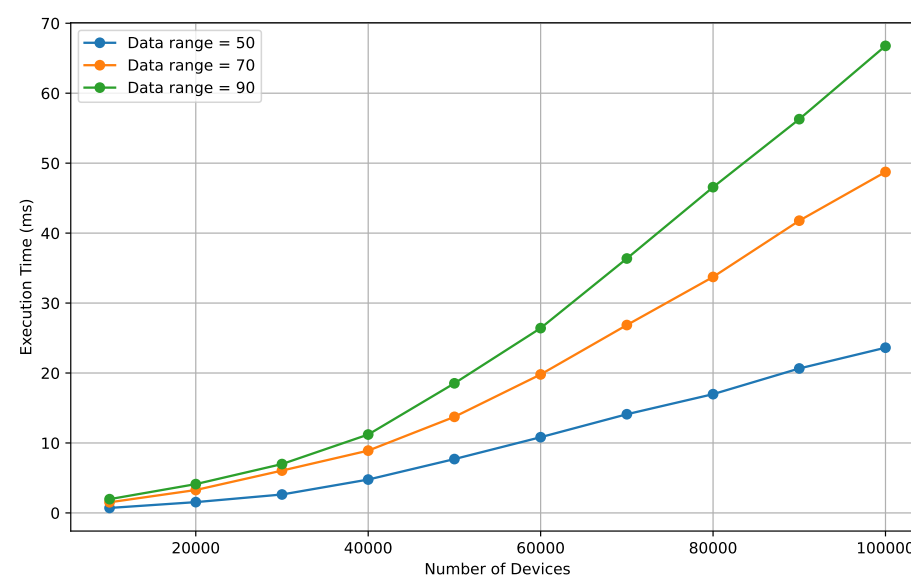
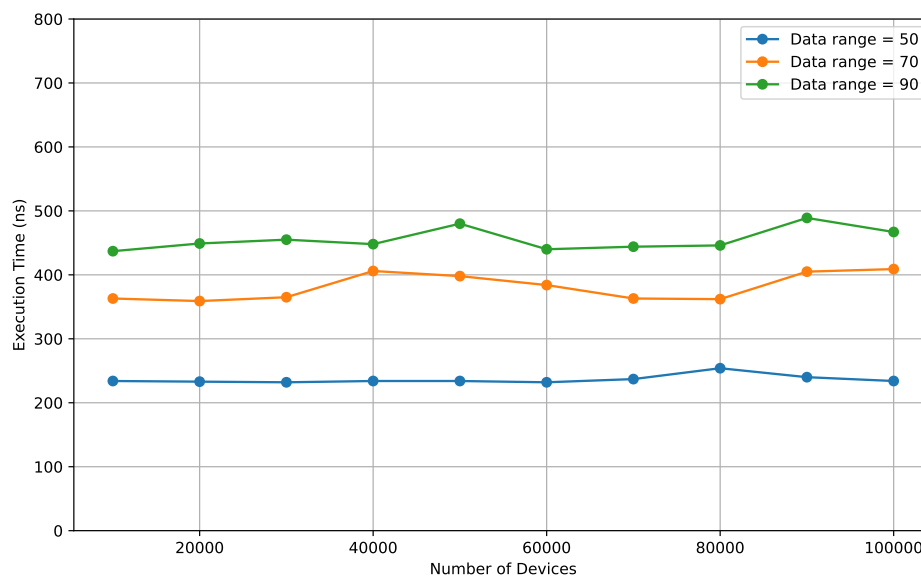
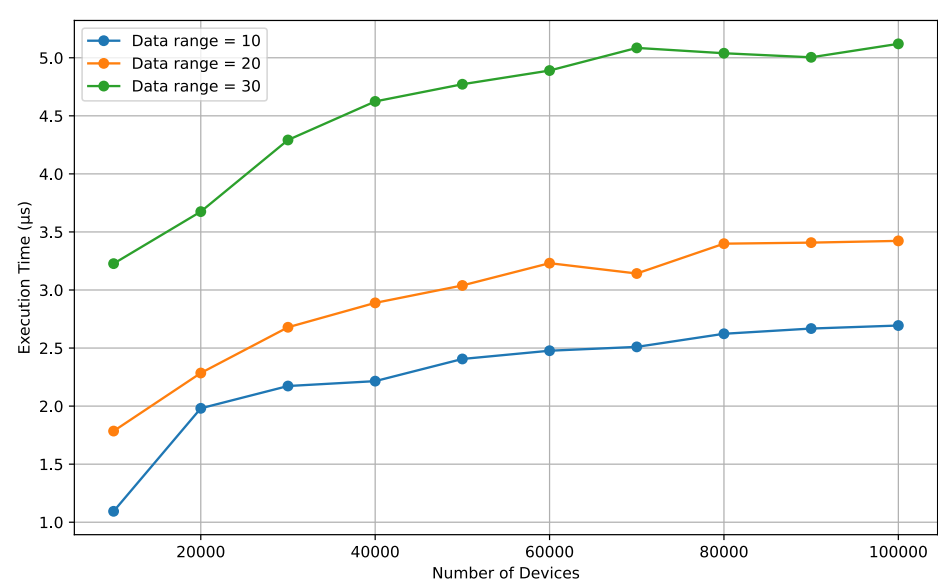
- Preserve each user's private information against  $\mathcal{ED}$  and  $\mathcal{HC}$ .
- Preserve  $\mathcal{HC}$ 's aggregation types and the query range  $(\alpha, \beta)$  against  $\mathcal{ED}$  and users.
- Efficient in terms of both computational costs and communication overhead.

## Our Proposed Scheme

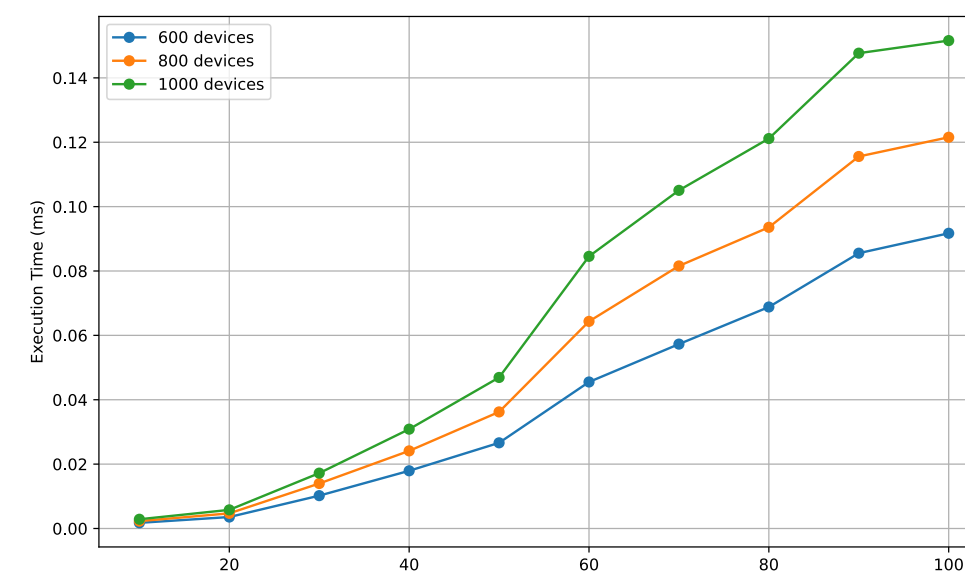
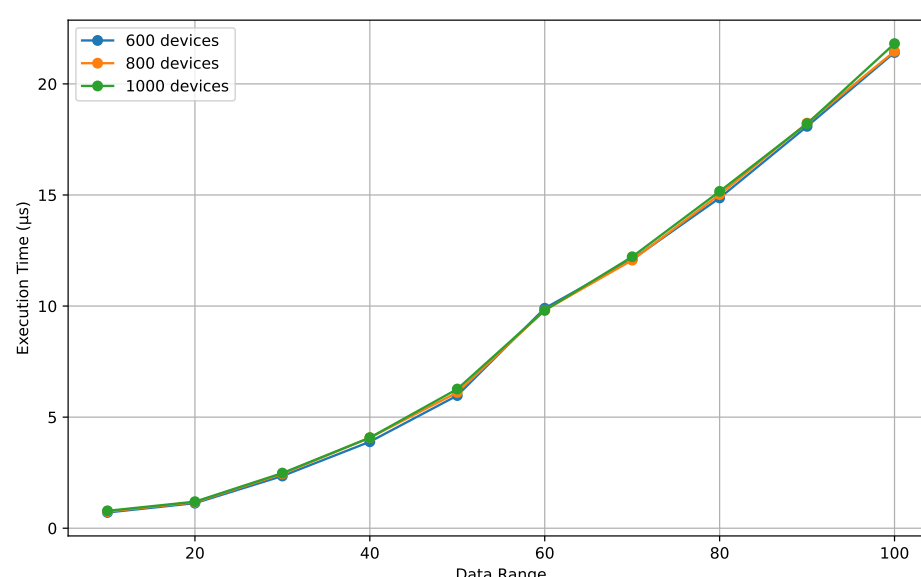
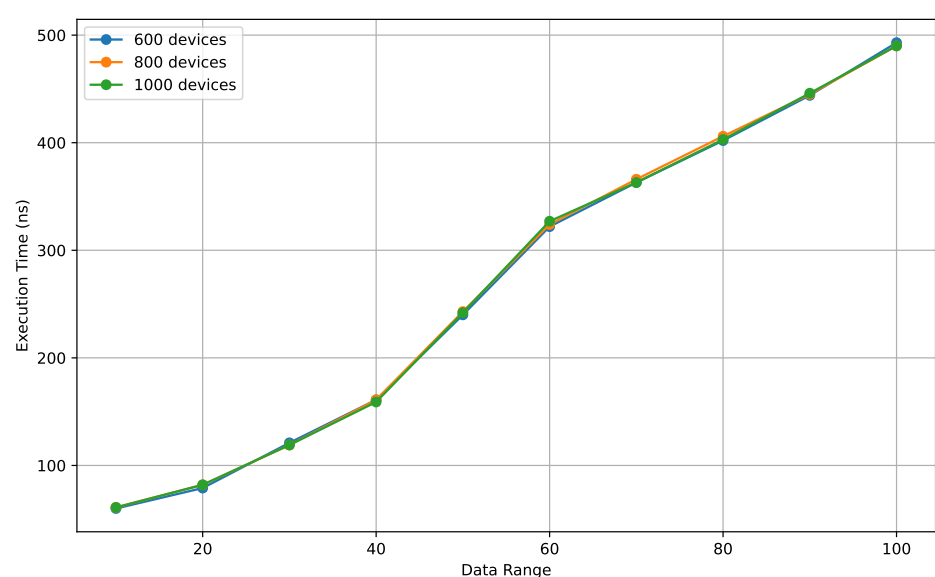


## Experimental Analysis

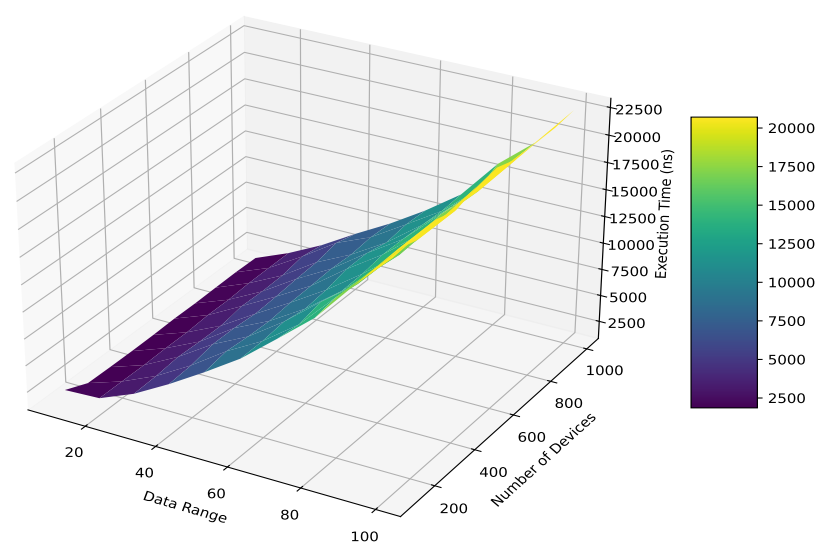
Part A. The execution time varies with the number of devices



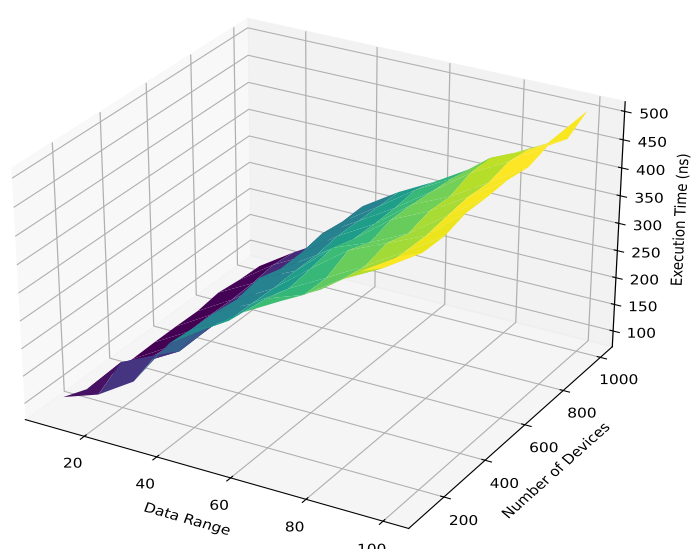
Part B. The execution time varies with the data range



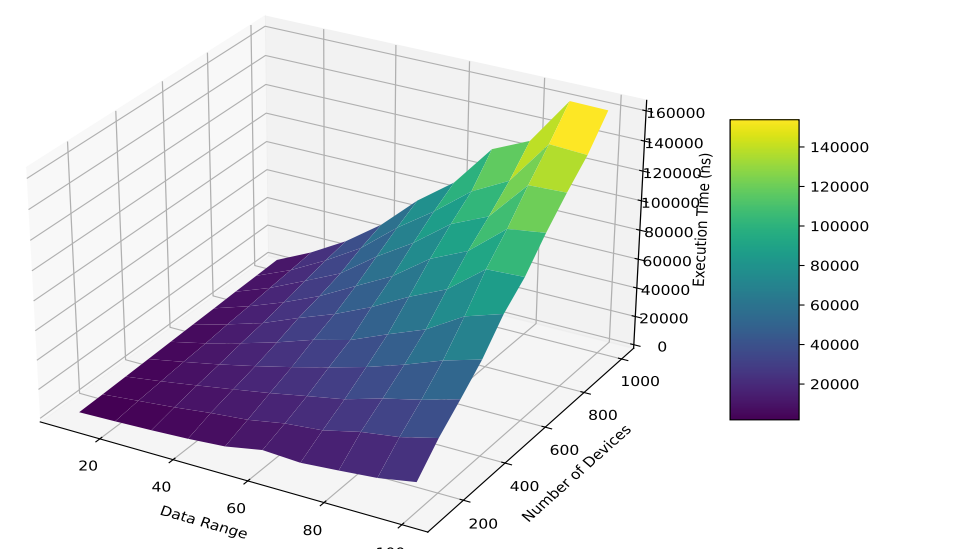
Part C. The execution time varies with the number of devices and the data range



(a) The execution time of the healthcare center



(b) The execution time of users



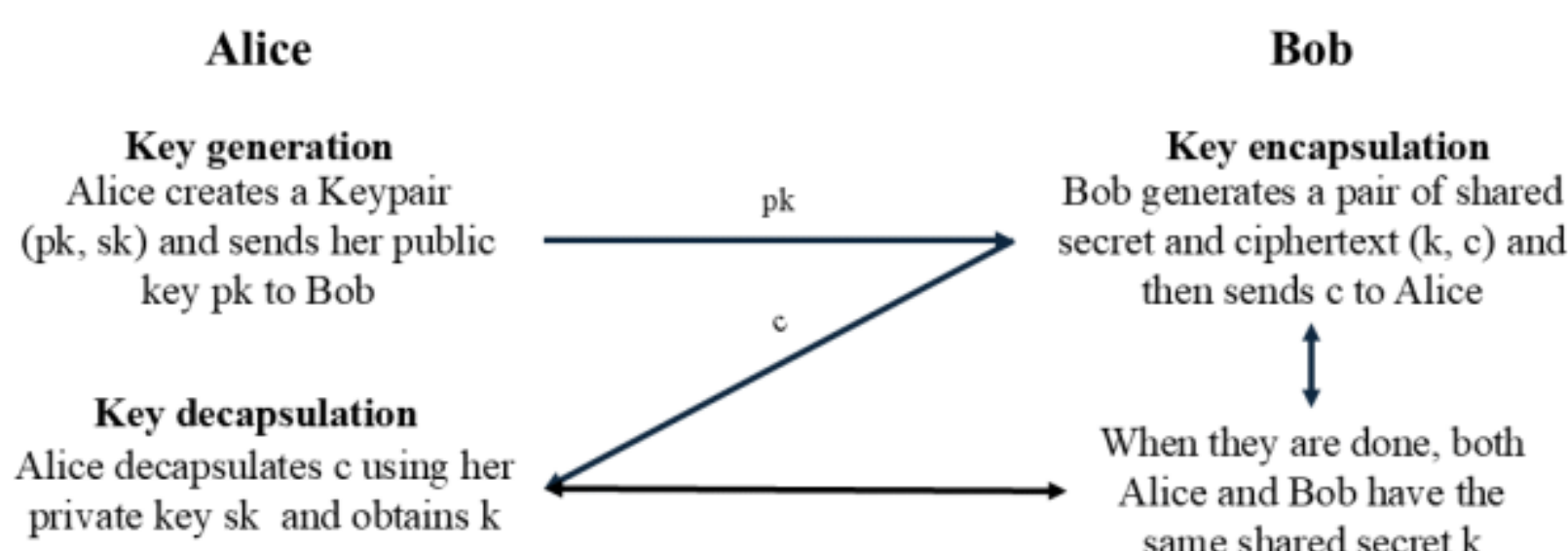
(c) The execution time of edge devices



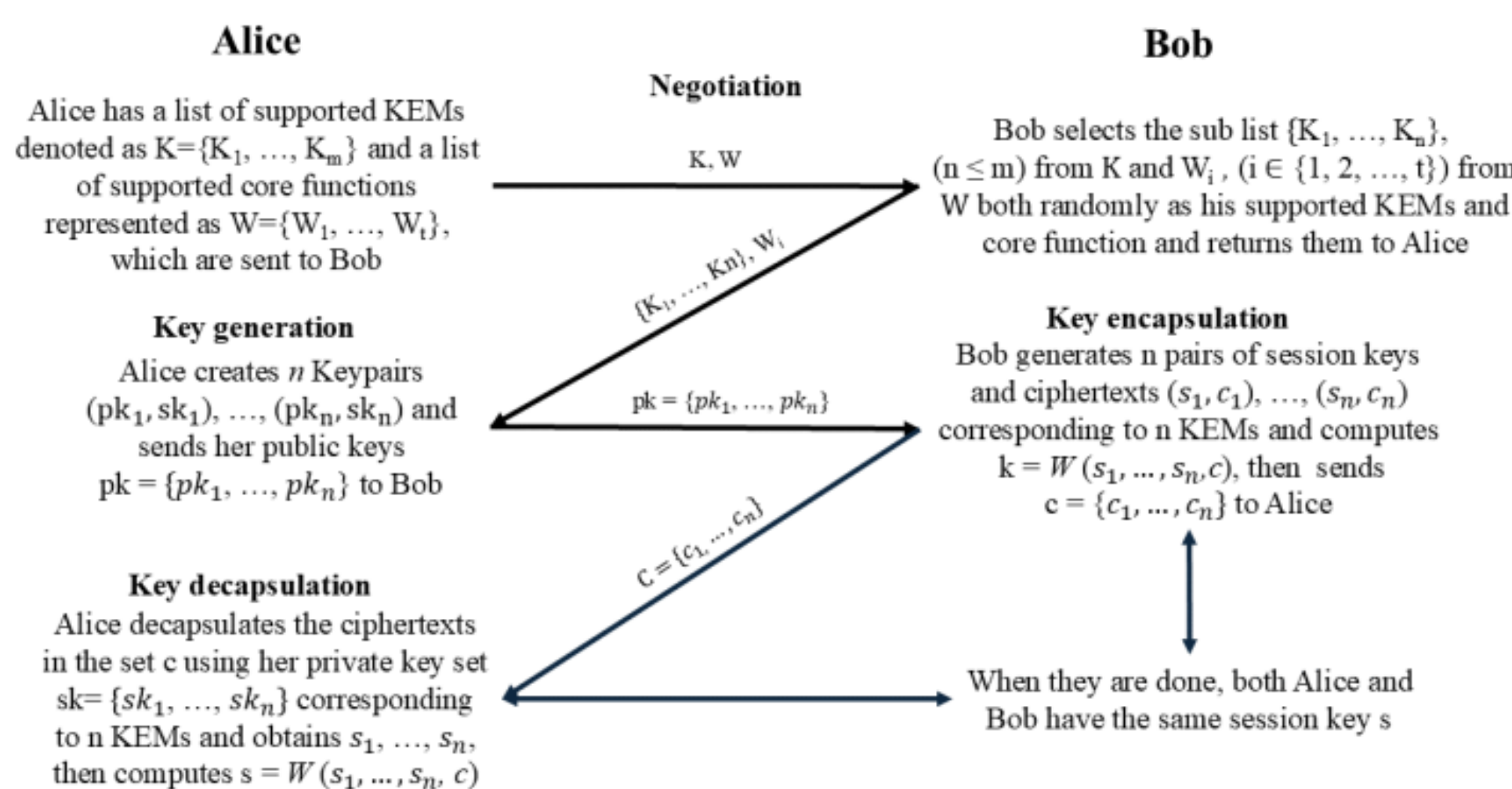
## ABSTRACT

Hybrid Key Encapsulation Mechanisms (KEMs) are emerging as a viable approach to counter the potential risks posed by quantum computing by integrating diverse cryptographic techniques and primitives. The significance of Hybrid KEMs in the quantum landscape is driven by two main factors: facilitating a seamless transition from classical cryptographic algorithms to post-quantum algorithms, thereby supporting cryptographic agility, and underscoring the need for diverse cryptographic strategies, as relying solely on post-quantum algorithms may not be sufficient. This work provides a comprehensive review of current approaches to combining KEMs, with a particular focus on their efficiency. We investigate various KEM combinations, evaluating their cryptographic characteristics and computational efficiency. Our analysis reveals that the time and memory overhead associated with combiners is minimal, indicating that performance discrepancies among different combiners are negligible. Consequently, security becomes the primary factor in choosing a combiner. These insights are crucial for identifying optimal KEM combinations that meet specific cryptographic needs.

## Key Encapsulation Mechanism (KEM)



## Hybrid Key Encapsulation Mechanism



## Overview of KEM Combination Types

### XOR

$$W(k_1, \dots, k_n, c_1 \dots c_n) = k_1 \oplus \dots \oplus k_n$$

### XOR-then-PRF

$$(k_1, \dots, k_n, c_1 \dots c_n) \mapsto F(k_1 \oplus \dots \oplus k_n, c_1 \dots c_n)$$

### PRF-then-XOR

$$W(k_1, \dots, k_n, c_1 \dots c_n) = \bigoplus_{i=1}^n F_i(k_i, c_1 \dots c_n)$$

### Chain of blockcipher-then-PRF

$$\text{Let: } \pi(k_1, 0) = E_1, \pi(k_2, E_1) = E_2, \dots, \pi(k_n, E_{n-1}) = E_n$$

$$K = F(E_n, c_1 || \dots || c_n)$$

## Methodology

We utilized the Liboqs library from the Open Quantum Safe (OQS) project to implement four combiners: XOR, XOR then PRF, PRF then XOR, and chain of block ciphers then PRF, Using one of the ingredients as salt and Improved XOR. In our methodology, we combined the RSA KEM as a pre-quantum KEM with four post-quantum KEMs: BIKE-L5, Classic-McEliece-8192128, HQC-256, and Kyber1024.

## Key Findings

### ▪ Negligible Impact of Combiners on Time and Memory Usage

- ✓ Combiners have minimal effect on performance in hybrid KEM

### ▪ Security as the Primary Consideration

- ✓ Given minimal performance differences, security should be prioritized when selecting combiners

### ▪ Similar Memory Usage Across Post-Quantum KEMs

- ✓ All post-quantum KEM candidates in the NIST Standardization Process, except ClassicMcEliece-8192128, show similar memory usage

### ▪ Kyber-1024 Identified as the Most Time-Efficient

- ✓ Among the evaluated candidates, Kyber-1024 offers the best time efficiency

### ▪ Recommendation: 'PRF-then-XOR' Combiner

- ✓ The 'PRF-then-XOR' combiner is recommended for securely combining post-quantum and pre-quantum KEMs, such as Kyber

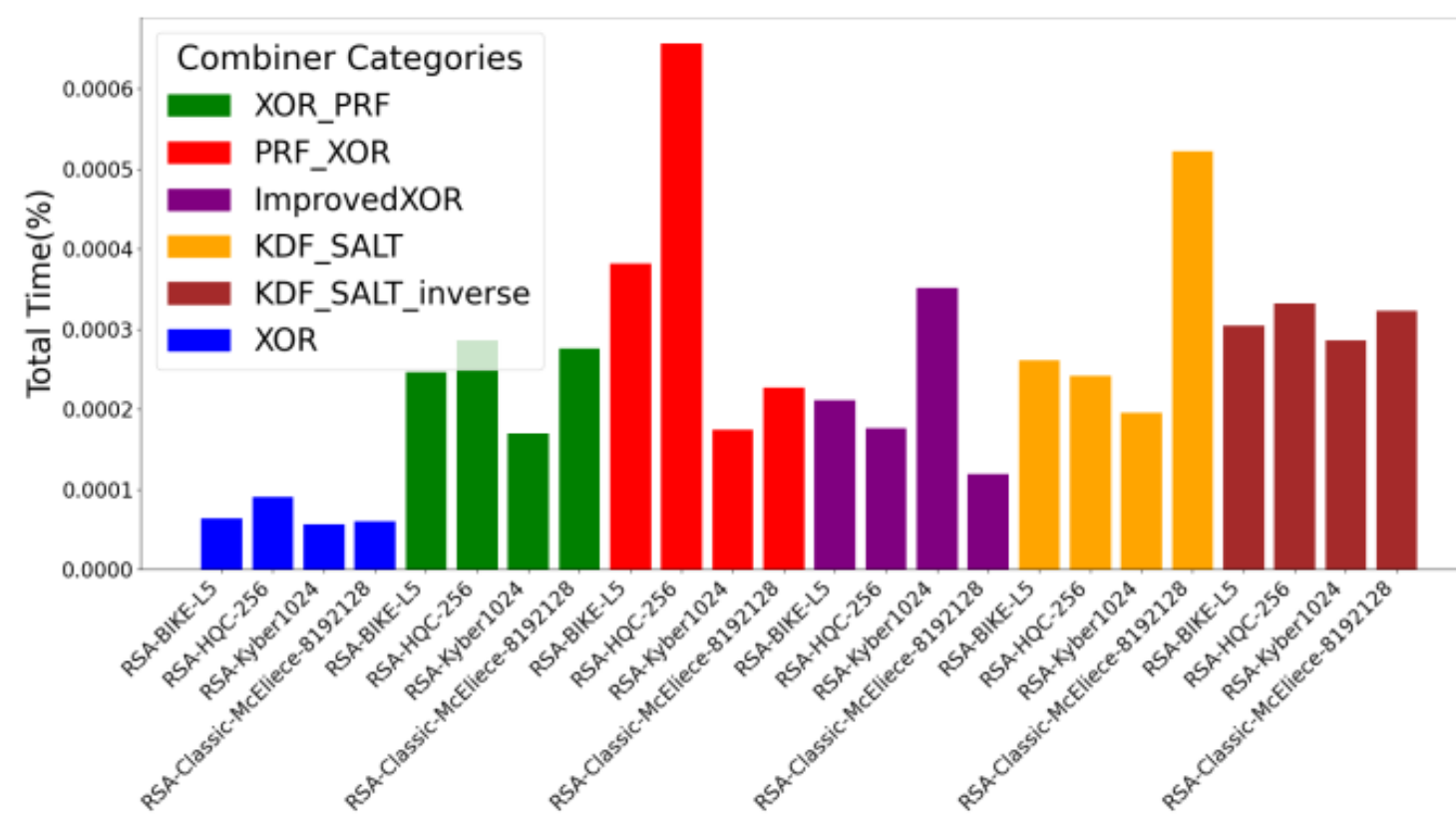
### ▪ Future Work: McEliece Key Size Optimization

- ✓ Focus should be on improving the efficiency of post-quantum KEMs, such as optimizing the key size of McEliece

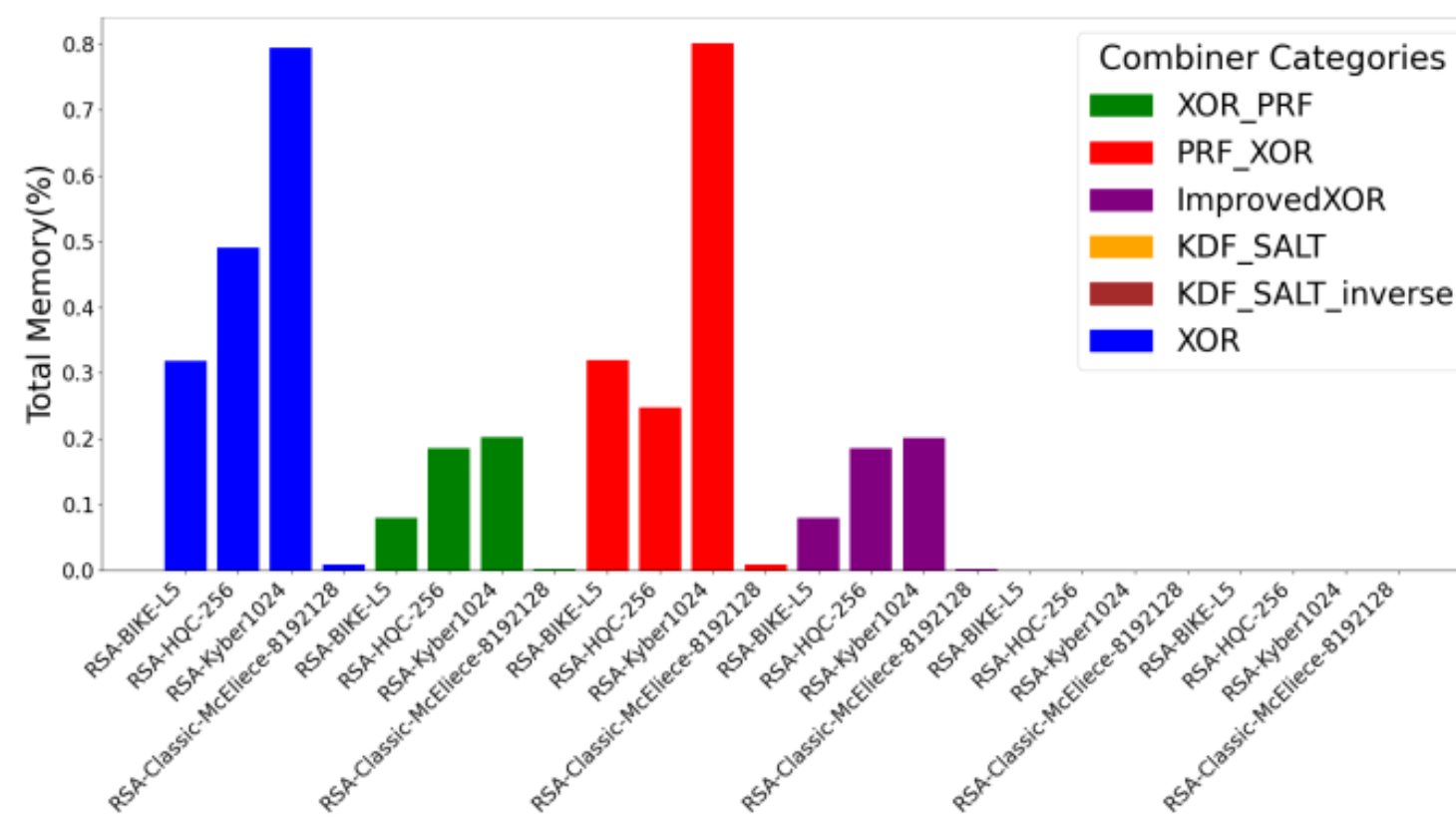
### ▪ Need for More Robust Combiners

- ✓ Development of stronger combiners is essential for ensuring maximum security in post-quantum cryptography

## Performance Evaluation of Hybrid KEM Combiners



- ✓ The percentage of time attributable to combiners in all combinations is very small (less than 0.001%) and considered negligible.
- ✓ The choice of combiner does not have a significant impact on time performance.



- ✓ The percentage of memory used by combiners in all combinations is also very small (less than 1%).
- ✓ The choice of combiner does not have a significant impact on memory consumption.

## References

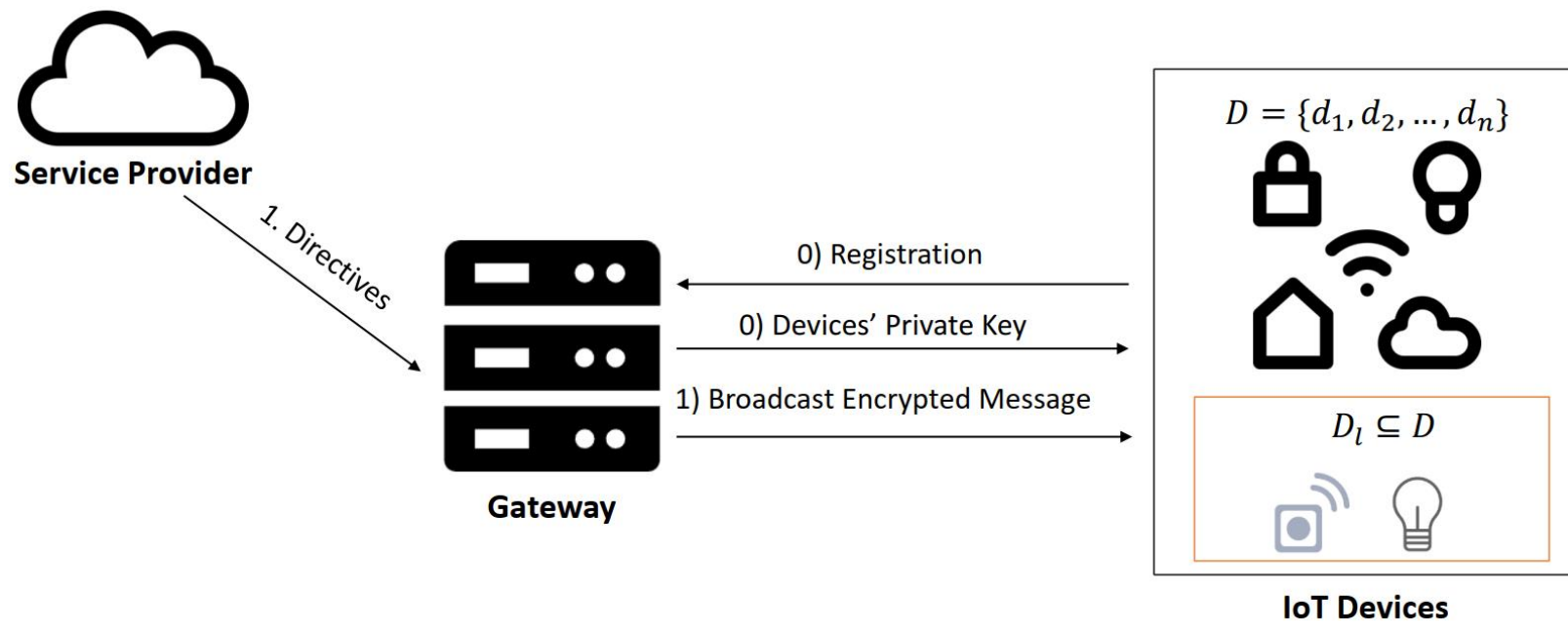
- [1] Federico Giacon, Felix Heuer, and Bertram Poettering. Kem combiners. In Michel Abdalla and Ricardo Dahab, editors, Public-Key Cryptography – PKC 2018,
- [2] NIST. Pqc standardization process: Announcing four candidates to be standardized, plus fourth round candidates. <https://csrc.nist.gov/News/2022/pqc-candidates-to-be-standardized-and-round-4>, 2022.
- [3] Daniel J Bernstein and Tanja Lange. Post-quantum cryptography. Nature, 549(7671):188–194, 2017.
- [4] Nina Bindel, Jacqueline Brendel, Marc Fischlin, Brian Gonçalves, and Douglas Stebila. Hybrid key encapsulation mechanisms and authenticated key exchange.
- [5] Takahiro Matsuda and Jacob C. N. Schuldt. A new key encapsulation combiner. In 2018 International Symposium on Information Theory and Its Applications (ISITA), pages 698–702, 2018.
- [6] Eric Crockett, Christian Paquin, and Douglas Stebila. Prototyping postquantum and hybrid key exchange and authentication in tls and ssh. Cryptology ePrint Archive, Paper 2019/858, 2019.
- [7] Open Quantum Safe. C library for prototyping and experimenting with quantum-resistant cryptography.
- [8] Sara Ricci, Patrik Dobias, Lukas Malina, Jan Hajny, and Petr Jedlicka. Hybrid keys in practice: Combining classical, quantum and post-quantum cryptography. IEEE Access, 12:23206–23219, 2024.



## ABSTRACT

The global trend toward aging populations shows the significance of Aging in Place (AiP), which necessitates advanced technologies that enhance both the safety and autonomy of the elderly within their familiar environments. In this context, secure and efficient communication within Internet of Things (IoT) networks for AiP systems becomes crucial. In this paper, we present a novel threshold authenticated encryption scheme designed specifically for AiP contexts. Our proposed scheme integrates the ElGamal threshold decryption with a binary fuse filter, effectively minimizing the frequency of communication group key updates thereby reducing communication overhead. Furthermore, our scheme applies the ASCON encryption algorithm to secure messages' contents, ensuring the transmitted data's security. Security analysis confirms that our proposed scheme satisfies the security requirements, which ensures confidentiality and integrity. Performance evaluations also validate its efficiency, highlighting its advantages in terms of communication, storage, and computation overheads.

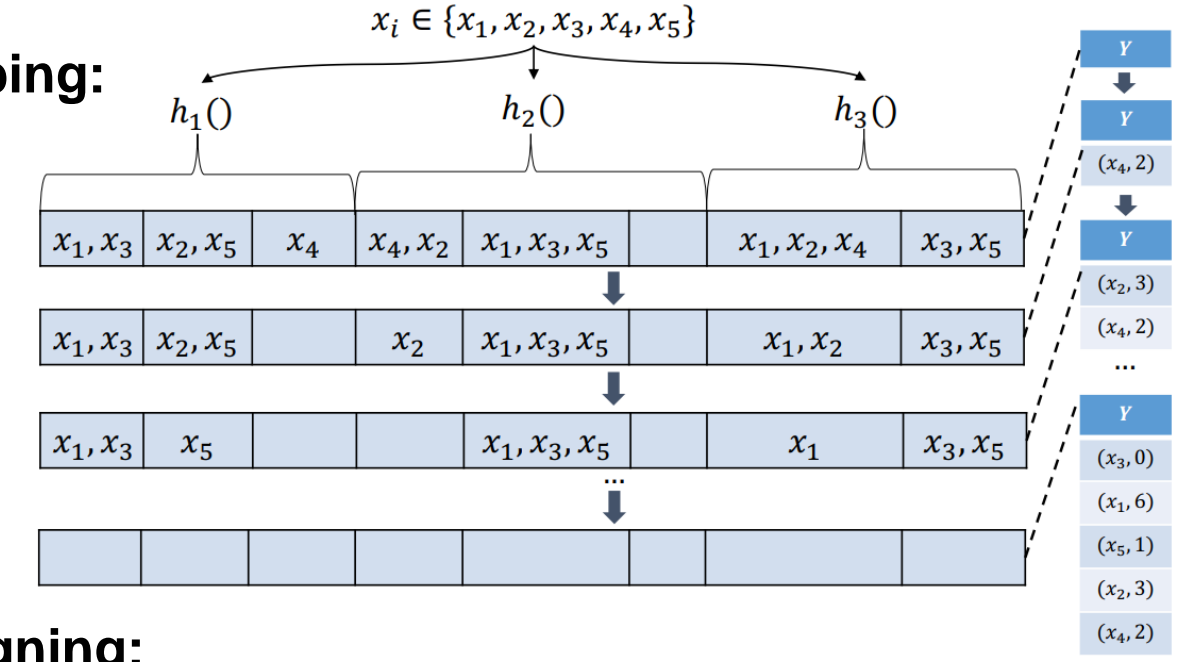
## System Model



- **IoT Devices ( $D$ ):**  $D$  is a group of  $n$  IoT devices.  $D_l \subseteq D$  are devices that temporarily leave the network while retaining the group key.
- **Gateway ( $GW$ ):** The  $GW$  acts as the central node within our system with superior computational and storage capabilities compared to the IoT devices.

## Binary Fuse Filter

### Mapping:



### Assigning:

$$A[index_i] = h_f(z_i) \oplus A[h_1(z_i)] \oplus A[h_2(z_i)] \oplus A[h_3(z_i)]$$

### Checking:

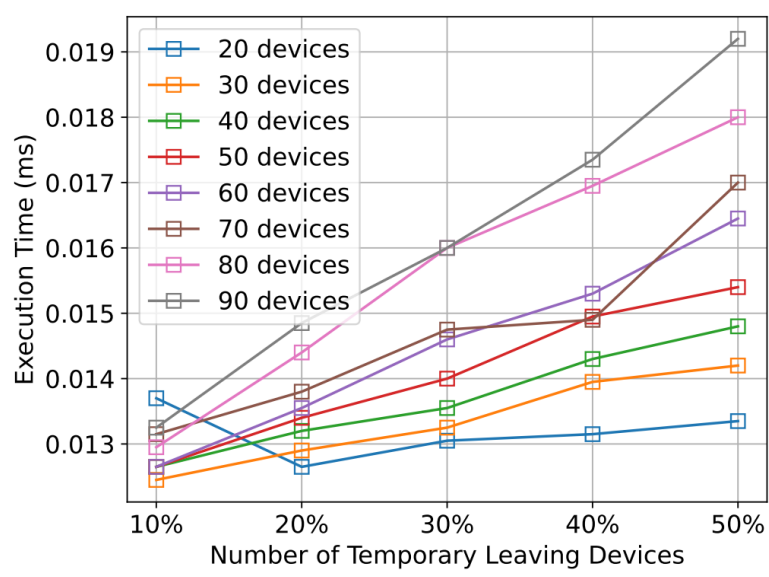
$$h_f(z') \stackrel{?}{=} A[h_1(z')] \oplus A[h_2(z')] \oplus A[h_3(z')]$$

The **Binary Fuse Filter** is an optimized data structure designed for efficiently verifying the existence of keys within a target set while maintaining a low space overhead.

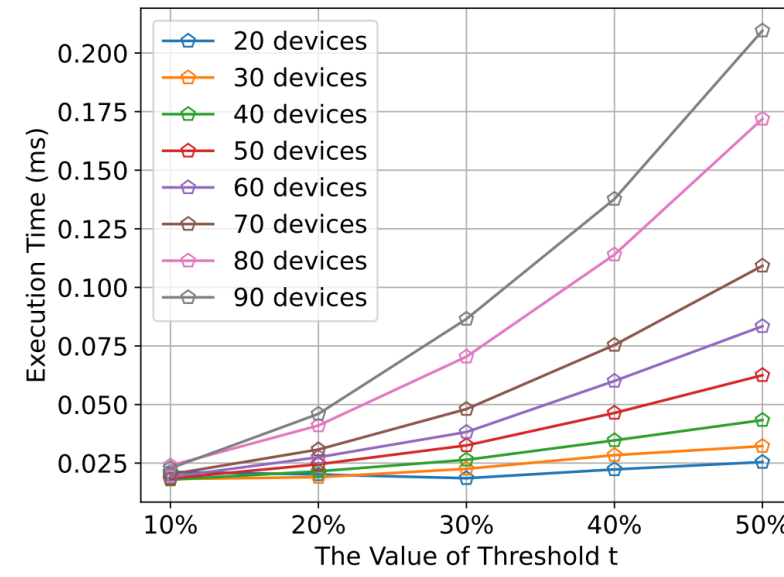
## Security Model

- All parties in this model are semi-honest. The model is based on the assumption that  $SS$ ,  $GO$  and  $Q$  do not collude.
- For  $GO$ , the topological structure of a graph, distance between vertices, and existence of paths between vertices cannot be disclosed to any party. However, both  $GO$  and  $Q$  know the vertices set  $v_1, \dots, v_m$ .
- For  $Q$ , the vertices  $v_s$  and  $v_t$ , specific hop number, and query distance between  $v_s$  and  $v_t$  should not be leaked.

## Performance Evaluation



The execution time of  $GW$  to update  $D_l$  and encrypt  $k_s$



The execution time of  $d_l$  to recover  $k_s$

## ElGamal Threshold Decryption

### Initialization

1. Selecting a large prime number  $p$  and forming a group  $G = \langle g \rangle$  with prime order  $q$ , where  $g \in Z_p^*$  as a generator.
2. Randomly chosen A private key  $s \in Z_q^*$  and compute public key  $y = g^s \mod p$ .
3. Constructed a polynomial  $f(x)$  of degree  $t - 1$  as follows:
$$f(x) = a_{t-1}x^{t-1} + \dots + a_2x^2 + a_1x + s$$
4. Each participant  $P_i$  with a pseudo identity  $x_i \in Z_p^*$ , distributor computes a fragment key  $s_i = f(x_i)$  and distributes to  $P_i$ .

### Encryption

To encrypt message  $m \in Z_p^*$  the ciphertext  $C$  is computed as:

$$C = E(m) = (\alpha, \beta) = (g^k \mod p, m y^k \mod p)$$

### Threshold Decryption:

Each participant  $P_i$  with pseudo identity  $x_i$  and fragment key  $s_i$  computes a Lagrange coefficient

$$L_i(0) = \prod_{i \neq j}^{j \leq t} -\frac{x_j}{x_i - x_j} \mod q$$

The original message  $m$  is reconstructed by:

$$m = \beta \left( \prod_{i=1}^t a^{s_i L_i(0)} \right)^{-1}$$

## Proposed Scheme

### Initialization (Set up the system and distribute keys to IoT devices)

1. The gateway ( $GW$ ) selects a large prime number  $p$  and a generator  $g$ . It then creates a public-private key pair:
  1. Private key:  $s$  (master private key)
  2. Public key:  $y = g^s \mod p$
  3. These public system parameters ( $p$ ,  $g$ , and  $y$ ) are shared with all devices.
2.  $GW$  determines a threshold  $t$ , the number of devices needed to decrypt messages. A polynomial  $f(x)$  of degree  $t - 1$  is generated using random coefficients and the master private key  $s$ .
3.  $GW$  assigns each device a unique private key  $s_i = f(x_i)$  using their pseudo-identity  $x_i$ . These keys are securely distributed to each IoT device.

### Verification and Decryption (Devices collaborate to decrypt the message)

1. A subset of devices ( $D_g$ ), where at least  $t$  devices are active, collaborates. A leader device  $d_l$  is chosen. Each device computes an intermediate value  $\alpha_i = \alpha^{s_i} \mod p$  and sends it to the leader.
2. The leader computes Lagrange coefficients  $L_i(0)$  using the identities of the participating devices. This is essential for reconstructing the session key.

### Threshold Authenticated Encryption (Encrypt a message from $GW$ and broadcast it securely to IoT devices)

1.  $GW$  checks for devices that temporarily left the network and constructs a filter  $F_l$  listing them.  $GW$  generates a signature  $\sigma$  for  $F_l$  using the master key  $s$  and broadcasts the pair  $\{F_l, \sigma\}$ .
2.  $GW$  creates a session key  $k_s$ , which is used to encrypt the communication.  $GW$  encrypts the session key using the public key  $y$  and broadcasts the encrypted session key  $c_k$ .
3.  $GW$  encrypts the actual message  $m_s$  using the ASCON algorithm with the session key  $k_s$ , producing an encrypted message  $c_m$  and a tag  $t_g$ .  $GW$  broadcasts  $\{c_k, c_m, t_g, r_n, a_s\}$  where  $r_n$  is a random number and  $a_s$  contains additional info like the timestamp.
3. The leader uses the Lagrange coefficients and intermediate values to reconstruct the session key  $k_s$ . The leader decrypts the message using  $k_s$ . If successful, the message  $m_s$  is recovered.
4. The leader shares the session key  $k_s$  with other devices (not on the temporary leave list) to decrypt the message.



## Abstract

Data exfiltration, the unauthorized transfer of sensitive information, is a critical threat to organizations, leading to potential financial, legal, and reputational harm. This system presents a solution to detect and prevent data exfiltration in Amazon Web Services (AWS), addressing limitations of AWS GuardDuty in detecting and preventing DNS and ICMP tunneling. The system integrates key tools such as Suricata for traffic monitoring, AWS CloudWatch for real-time alerting, and Lambda functions for automated threat response. The evaluation shows 100% detection and prevention success, with minimal false positives.

## The Problem



### AWS GuardDuty (Threat Detection Service)

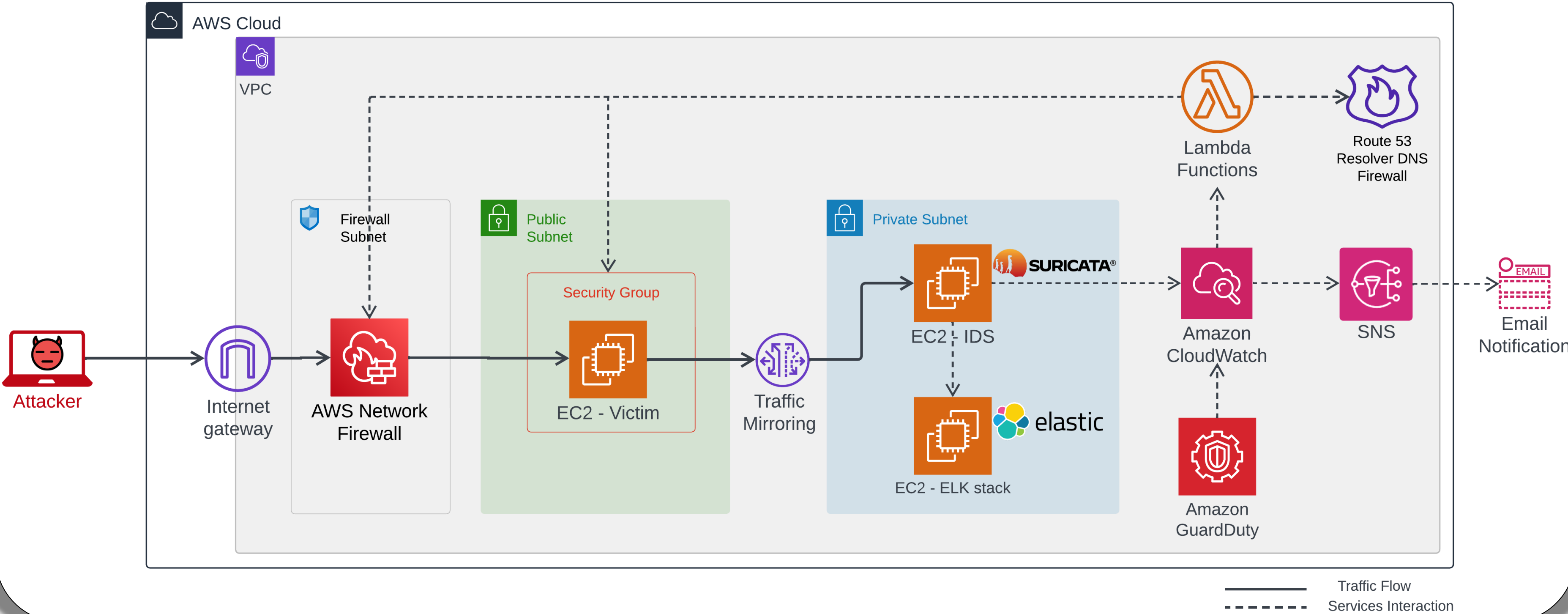
- Lacks the capability to detect DNS tunneling when using non-AWS DNS resolvers (e.g. Google DNS).
- Lacks the capability to detect ICMP tunneling.
- No built-in response mechanism in case a threat is detected.

## The Solution








### Intrusion Detection and Prevention System

- IDS: Based on Suricata to monitor and analyze network traffic.
- IPS: An automated event-driven response mechanism using AWS CloudWatch, which triggers Lambda functions to block threats with AWS Network Firewall and Route 53 DNS Resolver, performing real-time actions such as domain blocking and EC2 instance isolation.

## System Architecture



## System Components

IDS Components		
	Traffic Mirroring	mirrors traffic from the victim machine to the IDS (Suricata) machine for real-time analysis.
	EC2 - IDS	Monitors the mirrored traffic to detect threats based on predefined rules.
IPS Components		
	CloudWatch	Triggers other services based on Suricata alerts.
	SNS	Sends email notifications to system administrators.
	Lambda	Automatically executes response actions such as blocking malicious domains, IP addresses, or isolating compromise machines.
	Route 53 Firewall	Blocks DNS requests to malicious domains when the AWS-provided DNS resolver is used.
	Network Firewall	Blocks traffic to malicious IP and DNS traffic if instances are using external DNS resolvers.

## System Evaluation and Results

- A Domain Generation Algorithm (DGA) was used to generate 1,000 domains, which were used to test the system's detection and prevention capabilities.
- System performance was tested against DNS tunneling tools (Dnscat2, Iodine, Dns2tcp).
- Legitimate DNS traffic was generated (1,008,481 in total).

Tool	Detection Rate	Prevention Rate	Avg. Block Time (s)	False Positives	False Positives Rate
Dnscat2	100%	100%	4.22	19	0.0018%
Iodine	100%	100%	4.25	0	0
Dns2tcp	100%	100%	4.13	0	0

## Future Work

Future work includes extending the system to detect and prevent other data exfiltration methods, improving ICMP tunneling detection rules, incorporating machine learning algorithms, and utilizing the predefined ELK stack rules and integrate them with the system to enhance and expand the detection capabilities.





# Vulnerabilities in Autonomous Vehicle V2X Communication: Safety at Risk

Ishan Randeniya, Saqib Hakak

Contact Email: [ishan.r@unb.ca](mailto:ishan.r@unb.ca), [saqib.hakak@unb.ca](mailto:saqib.hakak@unb.ca)

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)



## ABSTRACT

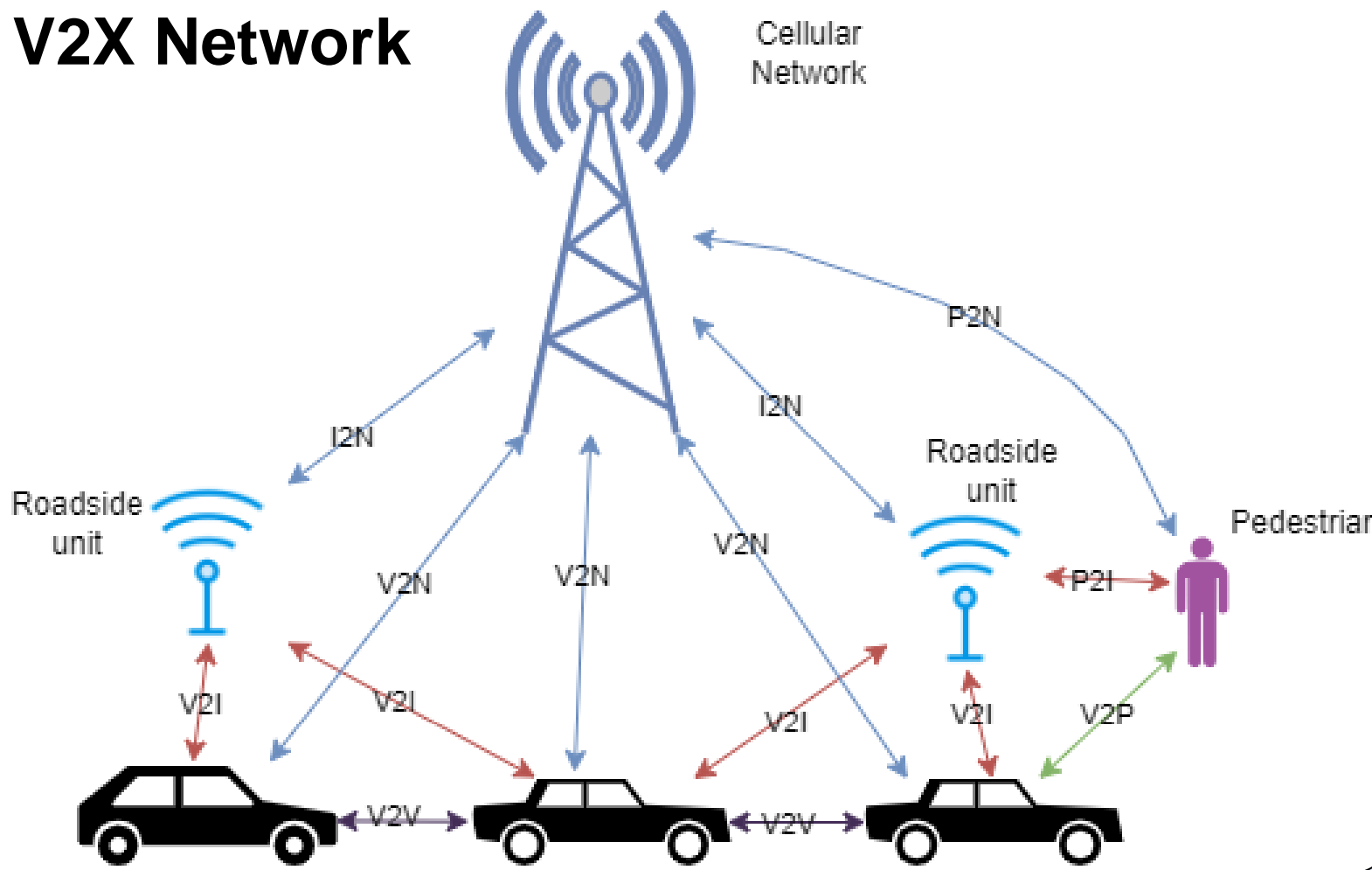
Vehicle-to-Everything (V2X) communication is poised to revolutionize transportation by enabling vehicles to communicate with each other and their surroundings. However, this connectivity also introduces vulnerabilities that malicious actors can exploit, jeopardizing safety and efficiency. This research explores the various attack vectors targeting V2X communication, including those affecting availability, authenticity, confidentiality, and integrity. We delve into the specific methods employed in these attacks, such as jamming, denial of service, spoofing, data manipulation, and replay attacks. Furthermore, we examine the potential consequences of successful attacks, emphasizing the critical need for robust security measures in V2X networks. By understanding these vulnerabilities and their implications, we can develop effective strategies to safeguard V2X communication and ensure the safe and efficient deployment of connected and autonomous vehicles.

## V2X Communication

V2X communication, or Vehicle-to-Everything communication, refers to the wireless exchange of information between a vehicle and any entity that may affect or be affected by it. It's a comprehensive system designed to improve road safety, traffic efficiency, and overall driving experience.

Type	Technology	Summary	Protocol
V2V	DSRC, C-V2X	Direct communication between vehicles for safety & efficiency	IEEE 802.11p (DSRC), 3GPP PC5 (C-V2X)
V2I	DSRC, C-V2X, RSUs	Communication between vehicles & roadside infrastructure for traffic & safety updates	IEEE 802.11p (DSRC), 3GPP PC5 (C-V2X), ETSI ITS-G5
V2P	BLE, Wi-Fi Direct	Communication between vehicles & pedestrians/cyclists for improved safety	Bluetooth LE, Wi-Fi Direct
V2N	Cellular networks (4G/5G)	Connects vehicles to cloud services for real-time data & enhanced navigation	TCP/IP, HTTP, MQTT

## V2X Network



## Vulnerabilities in V2X Communication

- Vehicle-to-Everything (V2X) communication, while promising significant advancements in road safety and traffic efficiency, also introduces several vulnerabilities that can be exploited by malicious actors.
- Attacks on Availability:** These attacks aim to disrupt or deny access to V2X services, impacting the real-time exchange of critical safety information.
  - Jamming:** Overloading the communication channel with noise, preventing legitimate messages from being transmitted or received.
  - Denial of Service (DoS):** Flooding the system with requests, overwhelming its resources and making it unavailable to legitimate users.
- Attacks on Authenticity:** These attacks focus on compromising the identity verification mechanisms in V2X, leading to the impersonation of legitimate entities.
  - Spoofing:** Masquerading as a trusted vehicle or infrastructure, sending false information to mislead other participants.
  - Sybil Attack:** Creating multiple fake identities to gain undue influence or disrupt the network's operation.
- Attacks on Confidentiality:** These attacks aim to gain unauthorized access to sensitive data transmitted over V2X networks.
  - Eavesdropping:** Intercepting and decoding messages to gain access to confidential information such as location data or personal details.
  - Man-in-the-Middle (MitM):** Intercepting communication between two parties, allowing the attacker to modify or eavesdrop on the exchange.
- Attacks on Integrity:** These attacks aim to modify or corrupt data transmitted over V2X, impacting the reliability of information.
  - Data Manipulation:** Altering or injecting false information into V2X messages, leading to incorrect decision-making or unsafe actions.
  - Replay Attacks:** Capturing and retransmitting legitimate messages at a later time, potentially causing disruptions or confusion.

## Securing V2X Communication

### Robust Authentication and Encryption:

It's imperative to implement strong cryptographic algorithms for data confidentiality and integrity, establish a robust Public Key Infrastructure (PKI) for secure certificate management, and utilize Hardware Security Modules (HSMs) for enhanced key protection.

### Secure Communication Protocols

It's crucial to employ Message Authentication Codes (MACs) for detecting message tampering, implement precise time synchronization to prevent replay attacks, and adhere to secure coding practices along with rigorous testing and vulnerability assessments of V2X software and firmware.

### Intrusion Detection and Prevention Systems (IDPS)

Strengthen V2X security through AI-powered anomaly detection, network segmentation, and firewalls to proactively identify and contain threats.

### Physical Layer Security

Robust sensor protection and anti-jamming measures are vital, along with redundant communication channels for enhanced resilience.

### Standardization and Interoperability

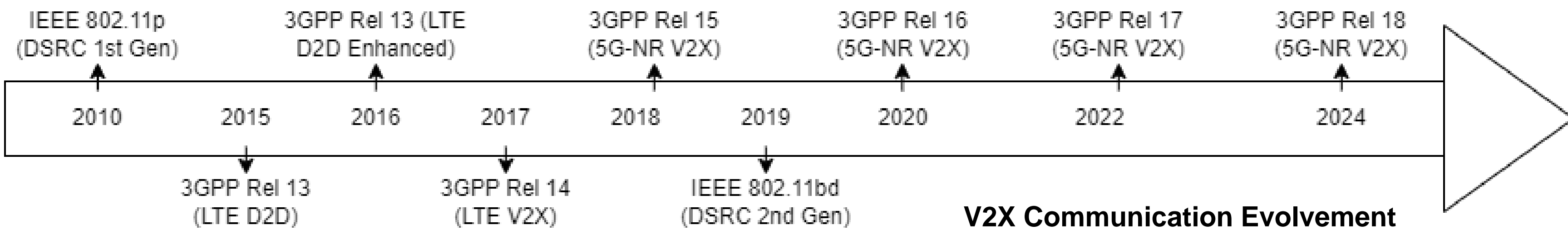
It's essential to foster collaboration among industry stakeholders. This collaboration aims to establish common standards and protocols, thereby ensuring seamless interoperability across different systems and minimizing vulnerabilities that arise from inconsistencies.

### User Education and Awareness

Fostering industry collaboration is key to establish shared standards, thus ensuring interoperability and reducing system vulnerabilities.

### Continuous Security Monitoring and Improvement

Maintain V2X security through regular vulnerability assessments and penetration testing, along with a robust process for timely updates and patches to address any identified weaknesses.



V2X Communication Evolvement





CIC

# Post-Attack Mitigation in Digital Substations: A Practical Approach

Mahdi Abrishami, Kwasi Boakye-Boateng, Hossein Shokouhinejad, Kishore Sreedharan, Shabnam Saderi Oskoue, Rongxing Lu, Ali Ghorbani

Contact Email: mahdi.abrishami@unb.ca, kwasi.boakye-boateng@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)



## Overview

### System Level

- Purpose:** Disrupting a service provided by the controller (e.g., SSH) to make it inaccessible from outside.
- Potential Attack Scenario:** An adversary aims to compromise the CPU usage observation conducted by the controller. They execute an SSH brute force attack to access the controller's terminal. Then, they stop the CPU observer script, and by having the sudo credentials, they change the existing sudo password so that no one can access the system from outside anymore.
- Defence Scenario:** A keylogger dumps the key strokes constantly in a file. A Flask server is listening to a non-conventional port. The operator authenticates to this server. After the successful authentication, the SSH dump files are zipped and sent back to the control center.

### Software Level

- Purpose:** Disrupting a running supervisory application on the controller.
- Potential Attack Scenario:** An adversary aims to compromise the CPU usage observation conducted by the controller. They execute an SSH brute force attack to access the controller's terminal.
- Defence Scenario:** In case of receiving the signal that is intended for showing the delay in writing, the operator SSH's to the controller to run the auxiliary code stored in a certain location.

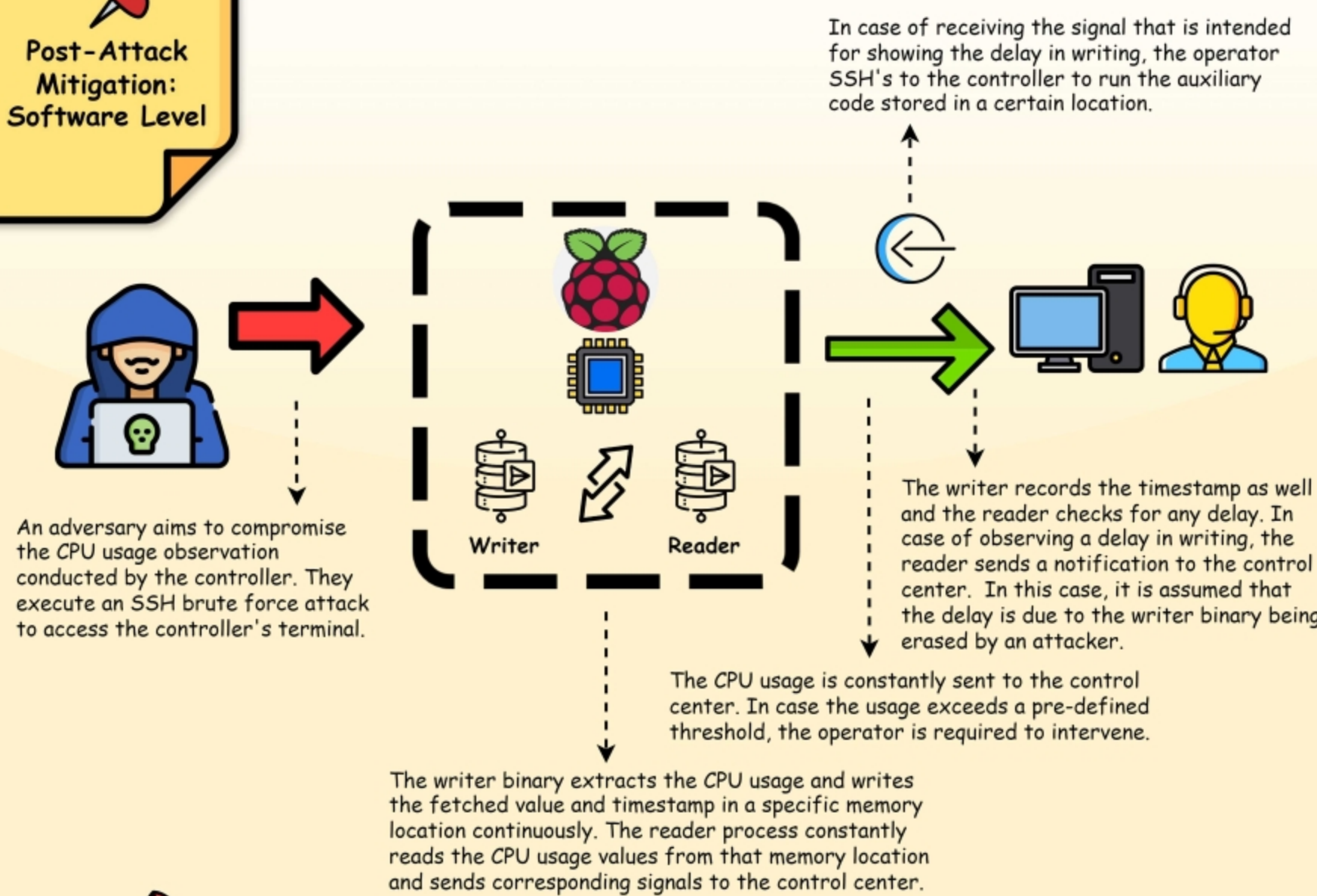
### Firmware Level

- Purpose:** Compromising the firmware to prevent the controller from booting an operating system.
- Potential Attack Scenario:** An adversary aims to wipe out the firmware. They gain an access through an SSH brute force attack. Having the root privileges, the adversary would be able to format the disk partitions containing the firmware.
- Defence Scenario:** A software-level program regularly checks for the integrity of the firmware. The status is continuously sent to the watchdog as heartbeats. If any changes in the integrity are observed, the watchdog could be signaled to reset the system (e.g., by not sending anymore heartbeats). A hardware-based watchdog timer ensures system reliability by monitoring for regular heartbeat signals from the controller. If these signals are not received within a set period, it triggers a system reset. Upon resetting the controller by the Watchdog, booting is started from an external memory containing the firmware and OS.

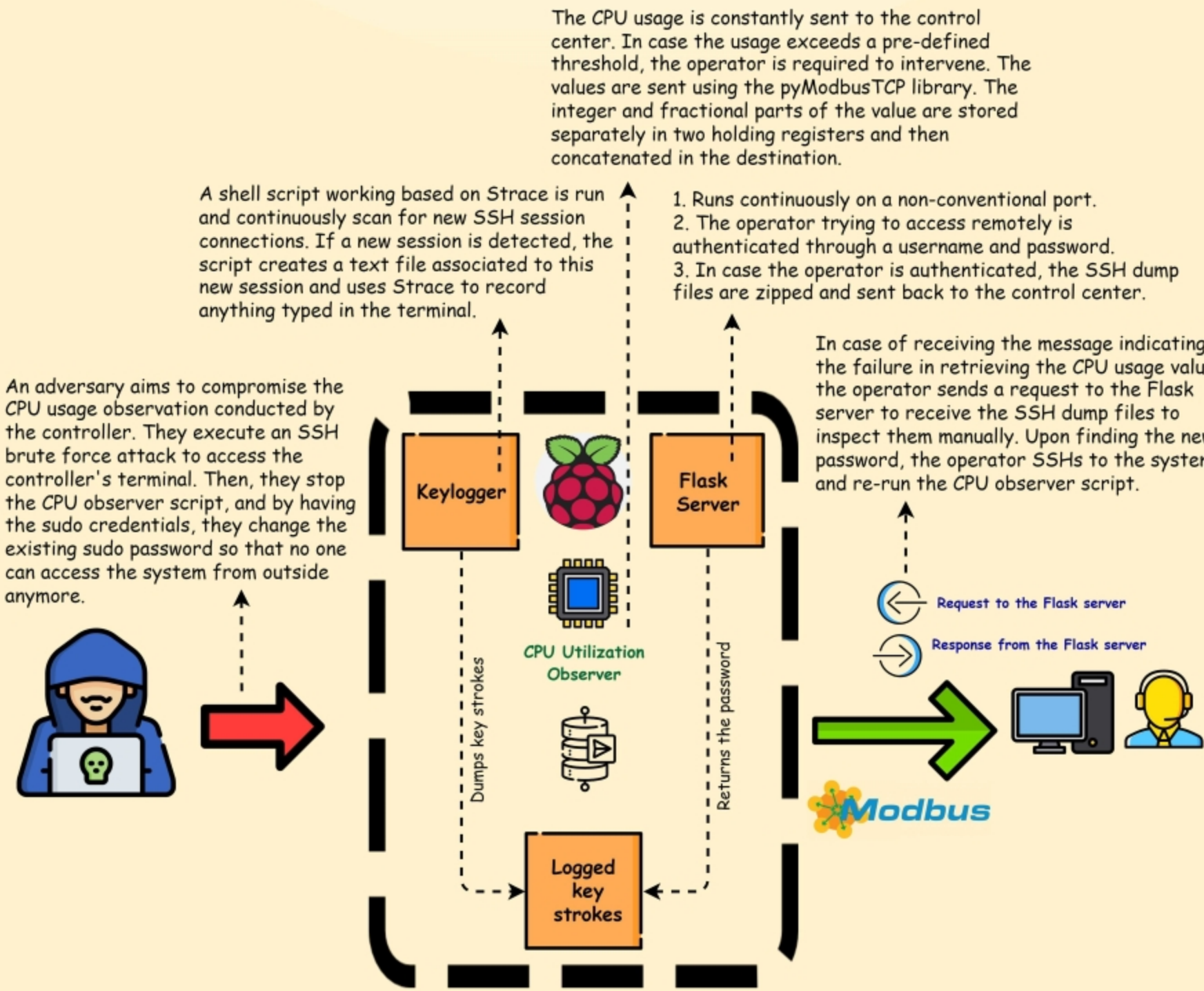
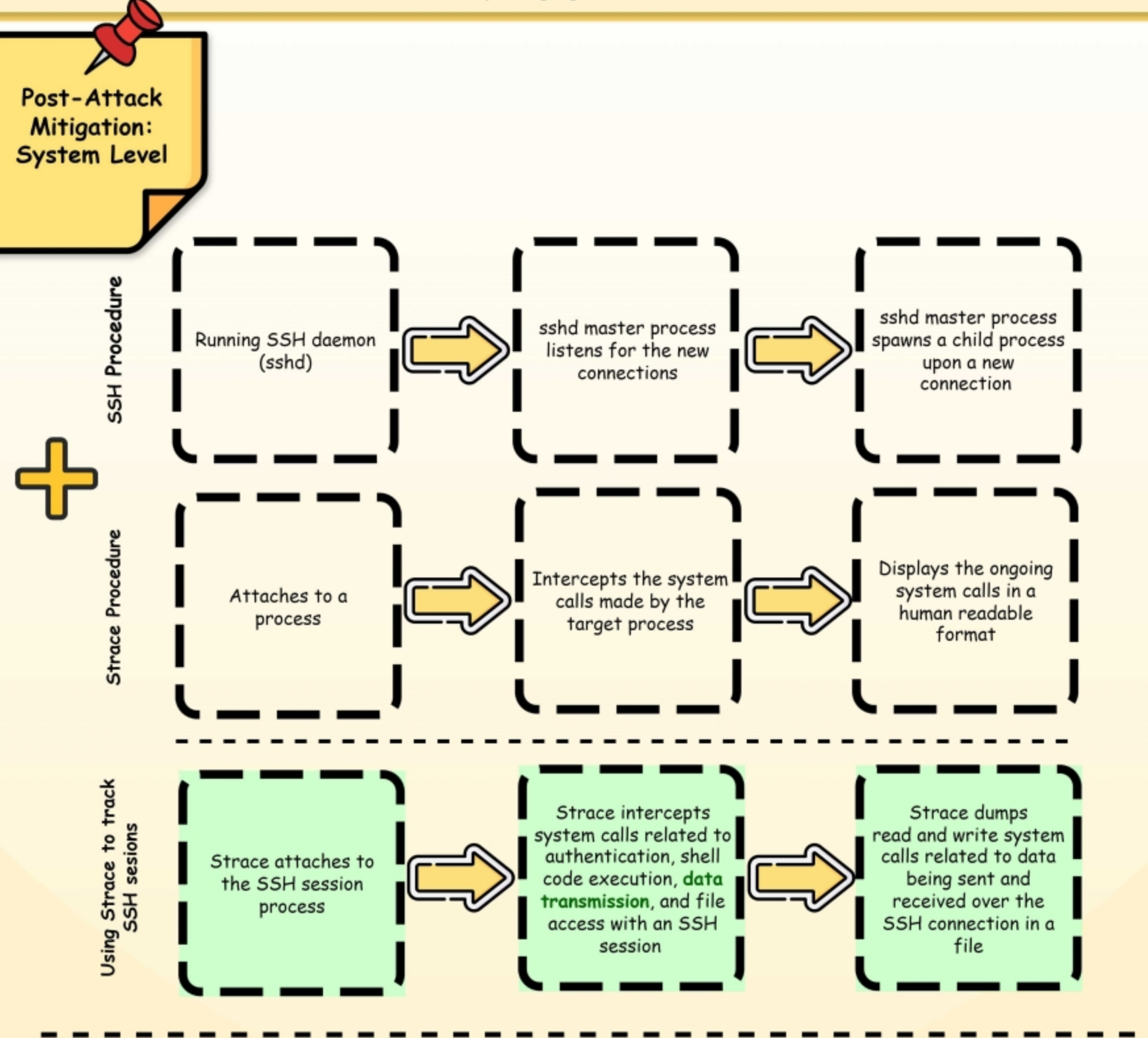
### Bootloader Level

- Purpose:** Disrupting the operating system through compromising the bootloader to prevent it from booting properly.
- Potential Attack Scenario:** An adversary aims to wipe out the bootloader. They gain an access through an SSH brute force attack. Having the root privileges, the adversary would be able to compromise bootloader-related files.
- Defence Scenario:** A software-level program regularly checks for the integrity of the bootloader. The status is continuously sent to a watchdog as heartbeats. If any changes in the integrity are observed, the watchdog could be signaled to reset the system (e.g., by not sending anymore heartbeats). A hardware-based watchdog timer ensures system reliability by monitoring for regular heartbeat signals from the controller. If these signals are not received within a set period, it triggers a system reset. Upon resetting the controller by the Watchdog, new OS is booted from the network.

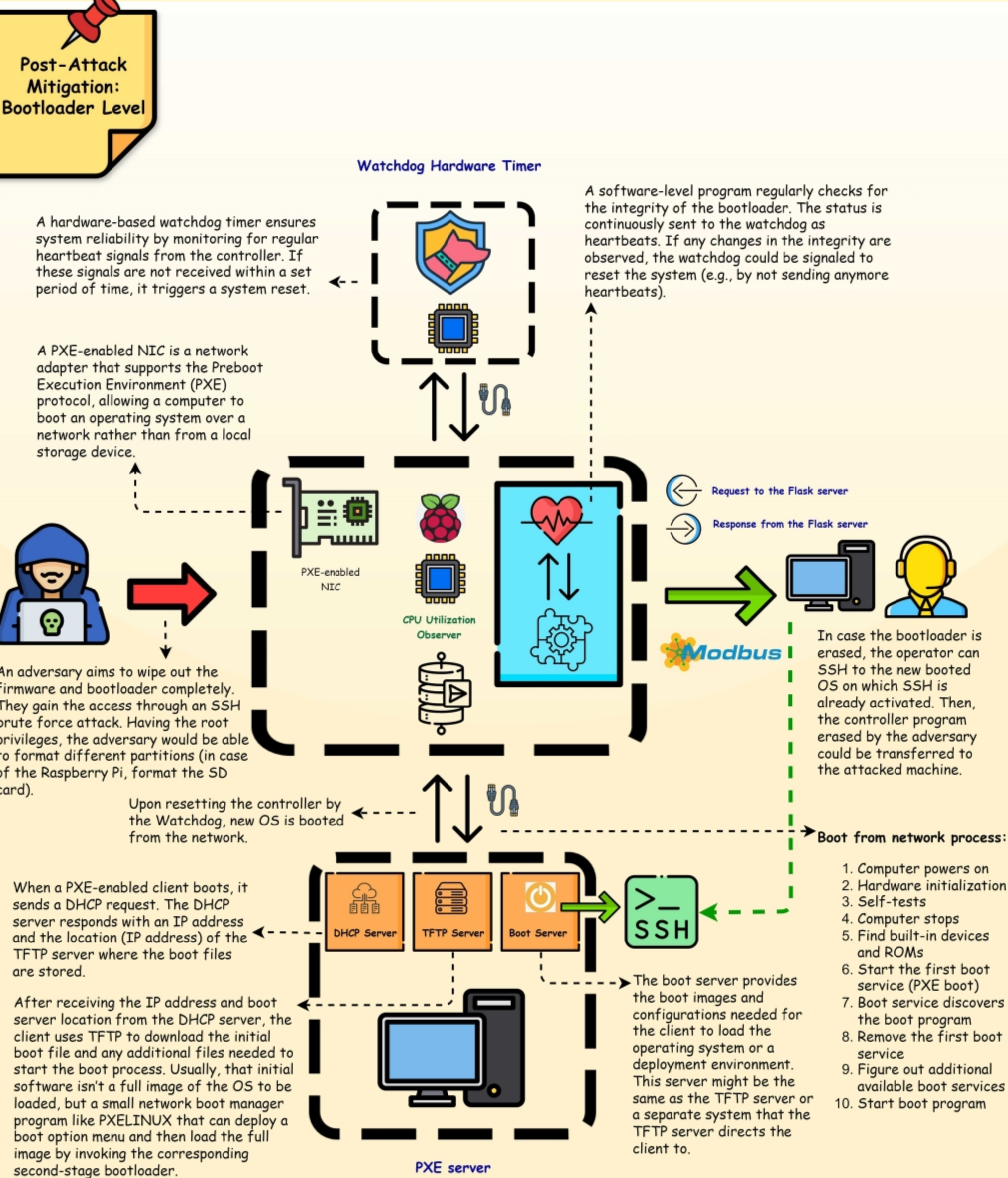
## Post-Attack Mitigation: Software Level



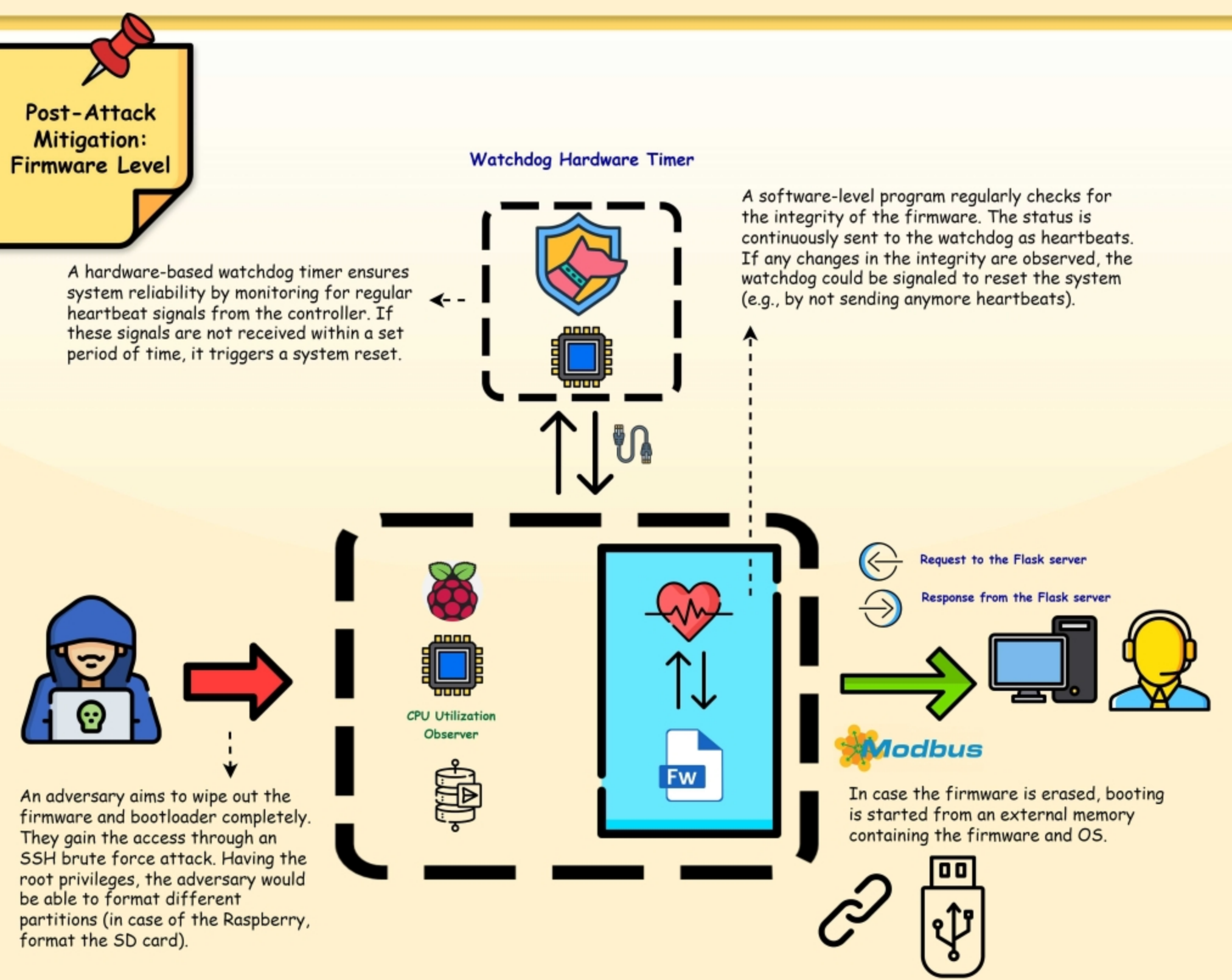
## Post-Attack Mitigation: System Level



## Post-Attack Mitigation: Bootloader Level



## Post-Attack Mitigation: Firmware Level

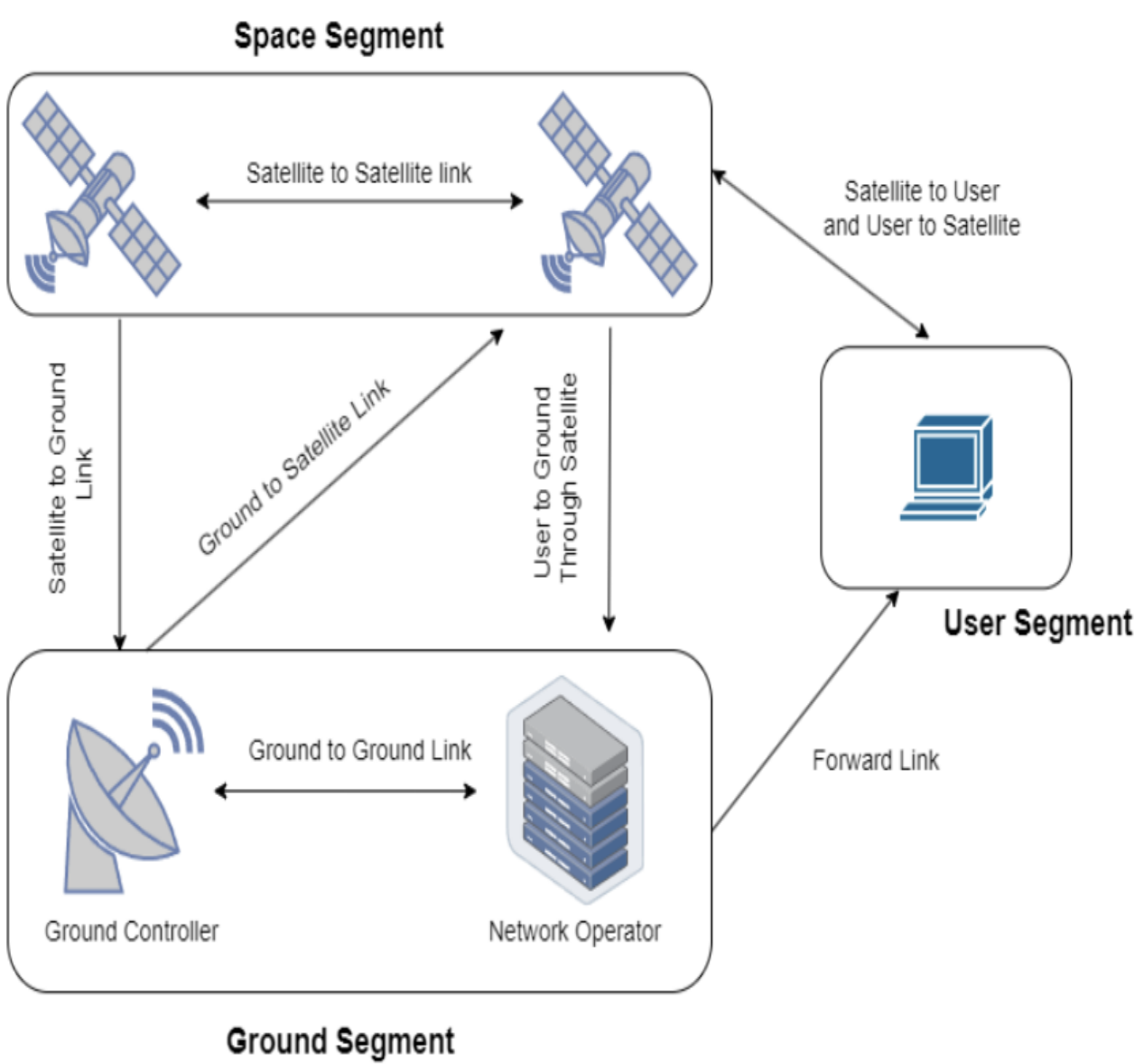




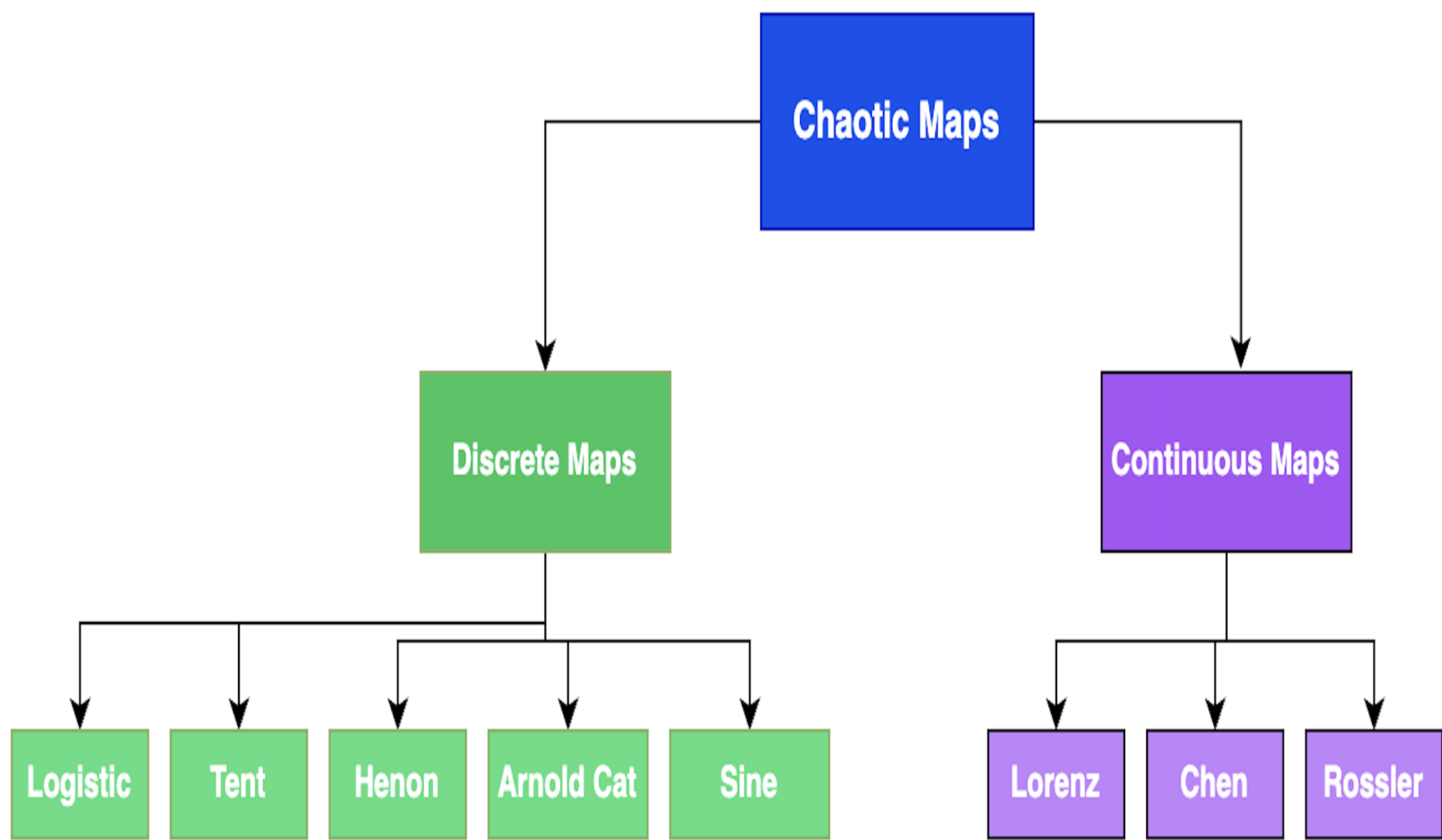
Introduction

- Chaotic systems, rooted in the study of the three-body problem in 1913 [1] and significantly advanced through the discovery of the Logistic map in 1976 [2], are known for their unpredictable yet deterministic behavior.
- Chaotic systems are characterized by their extreme sensitivity to initial conditions and parameters, making them highly effective for secure encryption.
- Chaotic systems are preferred in advanced encryption algorithms due to their properties such as determinacy, ergodicity, and sensitivity to initial conditions
- One notable application of chaotic maps is in the realm of satellite communications, where the secure transmission of sensitive image data is crucial [3].
- Traditional encryption methods often falter due to the high redundancy and strong pixel correlations in image data, and using chaotic maps can significantly enhance security.
- There are several existing works in the field of satellite chaotic image encryption which use Hyperchaotic, Multidimensional and Enhanced Chaotic systems to encrypt images in satellite communication.
- The effectiveness of these algorithms is evaluated based on criteria such as key space and resistance to differential attacks, ensuring robust protection against potential threats.

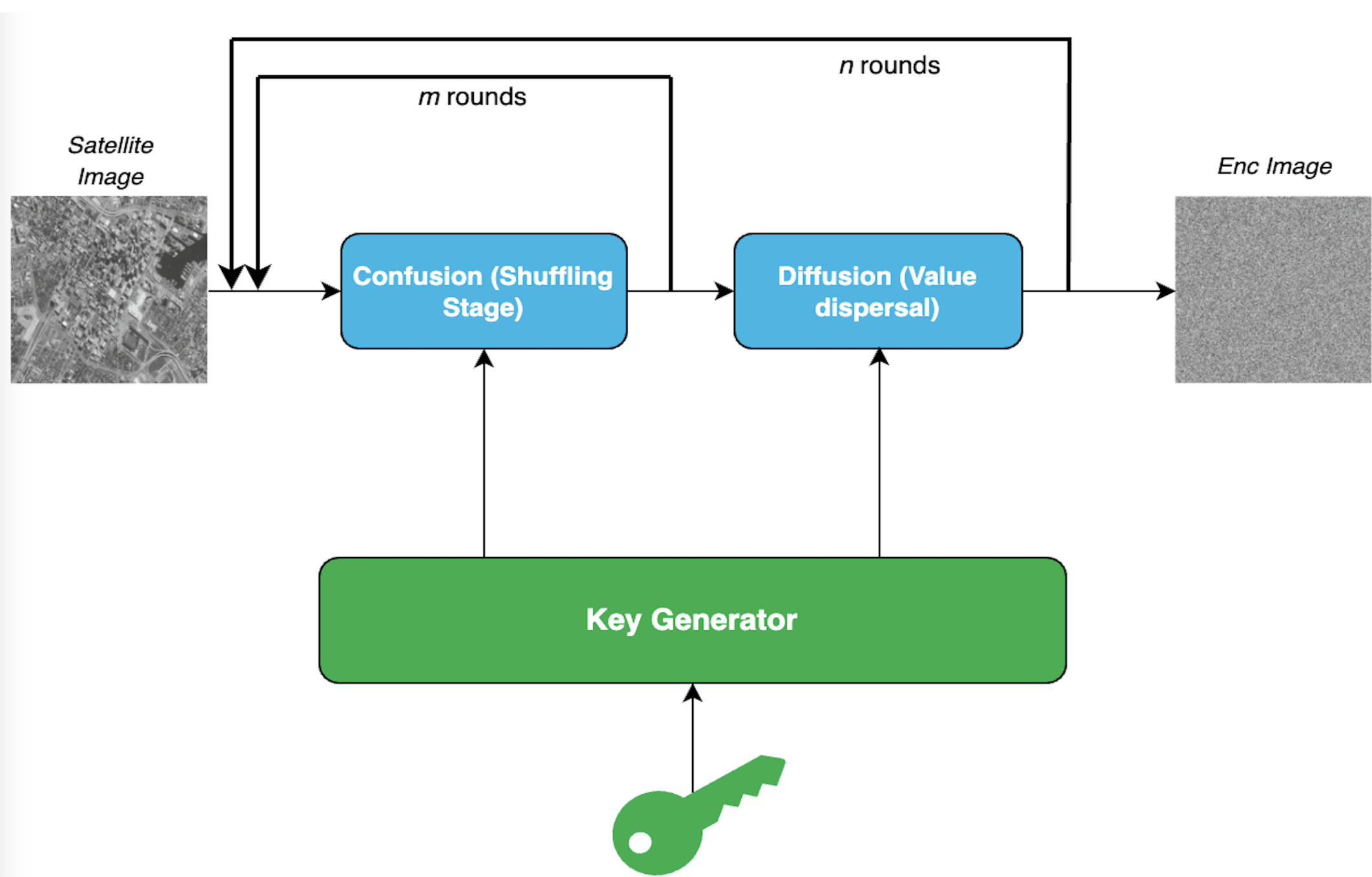
Satellite Communication Systems Architecture



Categorization of Chaotic Systems



General Chaos Based Image Encryption Process



Evaluation of Chaotic Encryption Algorithms

- Key sensitivity Analysis** - Evaluates encryption robustness by assessing how small key changes affect ciphertext
- Correlation Coefficient analysis** - Measures resistance to statistical attacks by evaluating pixel relationships in encrypted images. Lower correlation between adjacent pixels (vertically, horizontally, diagonally) indicates stronger encryption, with ideal values approaching zero.
- Histogram analysis** - Evaluates pixel intensity distribution in encrypted images. Effective chaos-based encryption produces uniform histograms, indicating high randomness and minimal exploitable patterns
- Differential attack** - Assesses encryption sensitivity to small input changes. Chaos-based schemes, highly sensitive to initial conditions, resist these attacks well.
- Information Entropy Analysis** - Entropy measures encryption randomness, with values near 8 indicating high unpredictability. Robustness against noise attacks (e.g., Gaussian, salt and pepper) is crucial for satellite transmissions

Future Research Directions

- Quantum Resistant Encryption** - Explore quantum-safe chaos-based encryption methods
- Efficient Hardware Implementation** - Design compact, efficient circuits for FPGAs and other platforms
- Hybrid Encryption Schemes** - Integrate chaos-based methods with traditional encryption (e.g., AES, RSA)
- Defending Zero-Day Attacks** - Develop robust methods against emerging threats

[1] Zhang, Bowen, and Lingfeng Liu. "Chaos-based image encryption: Review, application, and challenges." Mathematics 11.11 (2023): 2585.

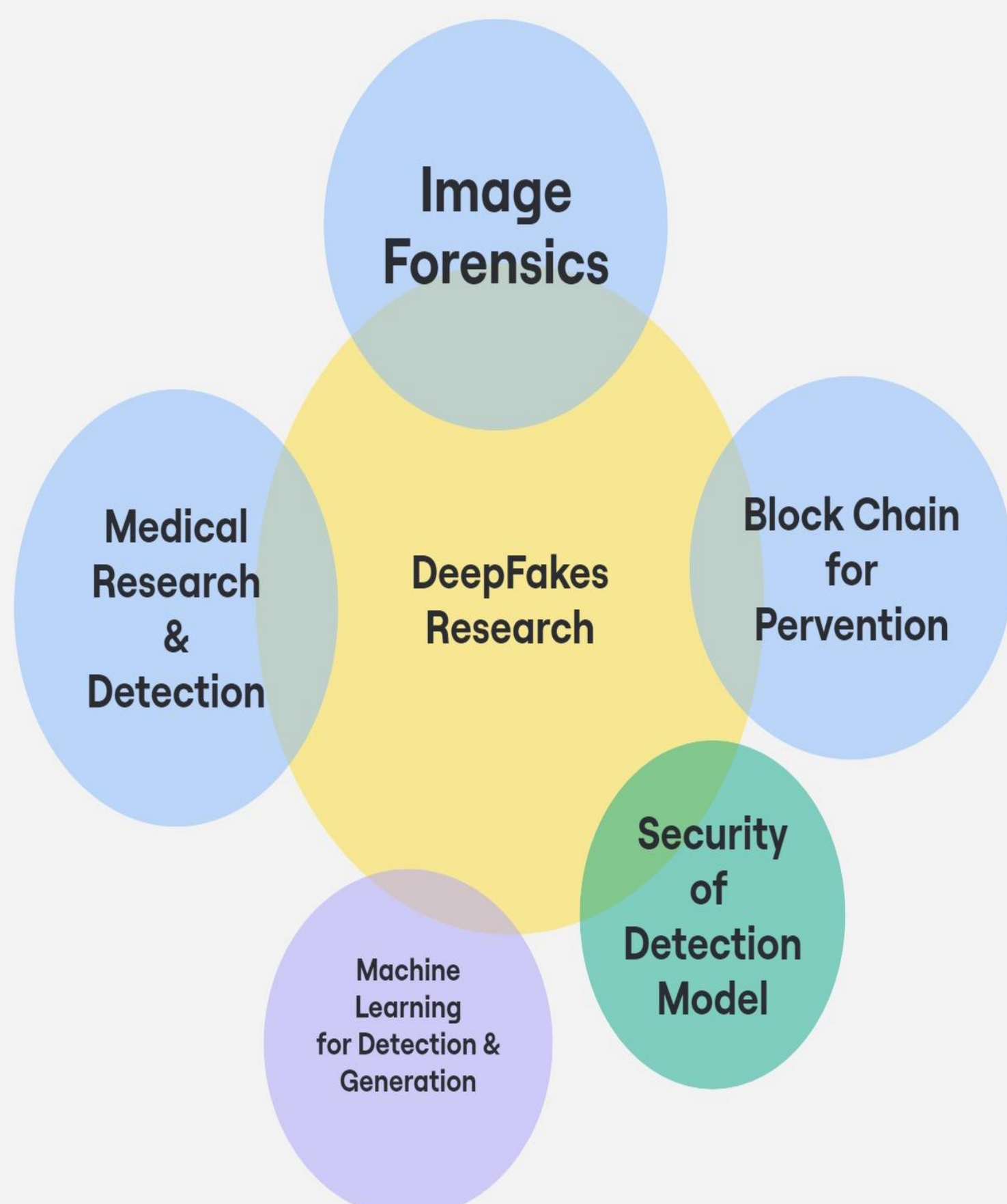
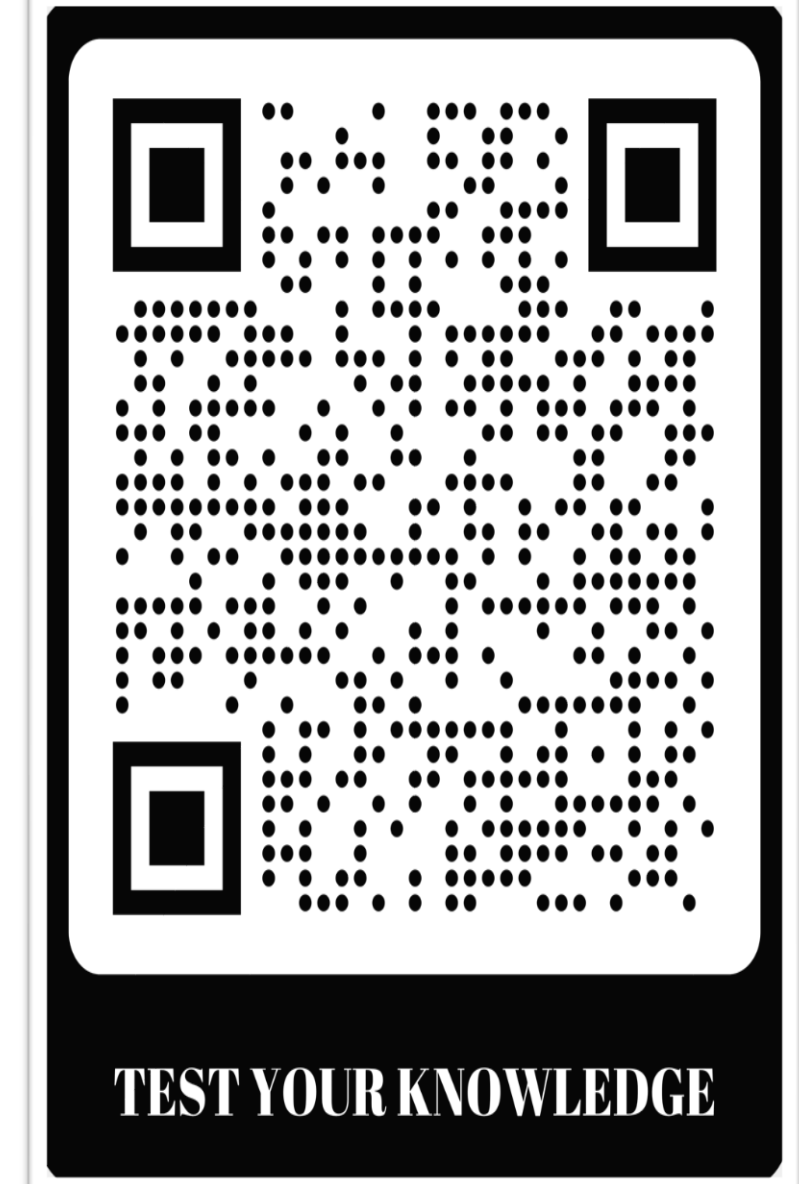
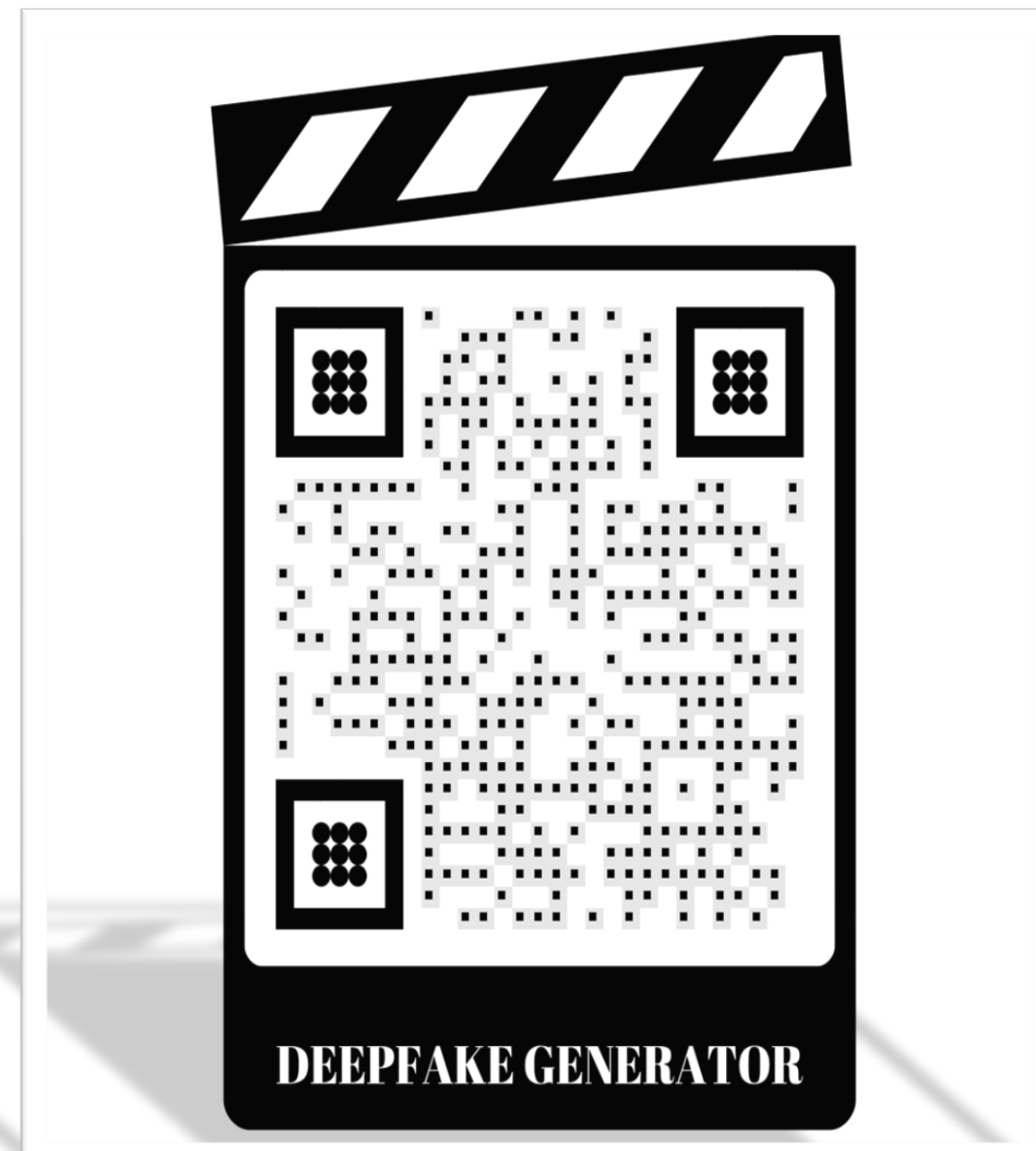
[2] May, Robert M. "Simple mathematical models with very complicated dynamics." Nature 261.5560 (1976): 459-467.

[3] Maral, Gerard, Michel Bousquet, and Zhili Sun. Satellite communications systems: systems, techniques and technology. John Wiley & Sons, 2020.

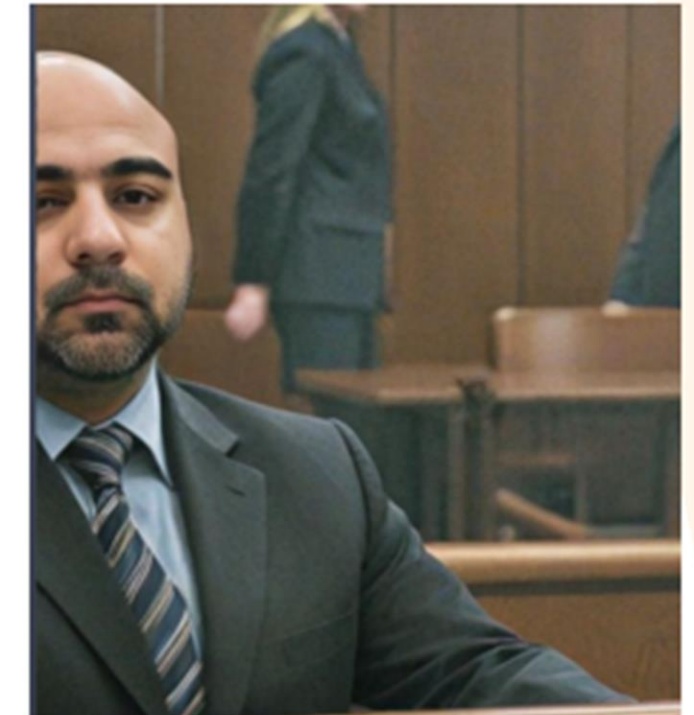


## Introduction

Deepfakes, which involve the creation of synthetic images, videos, and audio, are poised to significantly impact various sectors, including finance, entertainment, the judiciary, and politics. With the rise of advanced and easily accessible applications that enable the generation of user-created content, producing and sharing manipulated or fabricated media has become more straightforward. This underscores the critical need to educate the public about this technology and its potential consequences. There is also a need for detectors that can provide a comprehensive analysis of diverse types of Deepfakes in an understandable manner.



## My Deepfake



## Future Research

- ❖ Secure Multimedia
- ❖ Fast & Lightweight Detection
- ❖ User-oriented explainable system
- ❖ Traceable Multimedia
- ❖ Generalized Detection System

## Acknowledgment

This research is supported by the New Brunswick Innovation Fund (NBIF) under grant reference number RAI 2021-057 and the Harrison McCain Foundation Young Scholars Award under grant reference number HMF2023 YS-1.





CIC

# Assessing the impact of cyber threat in substations: A practical approach

Kishore Sreedharan, Kwasi Boakye-Boateng, Hossein Shokouhinejad, Mahdi Abrishami, Shabnam Saderi Oskoue, Rongxing Lu, Ali Ghorbani

Contact email: kishore.sreedharan@unb.ca, kwasi.boakye-boateng@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)

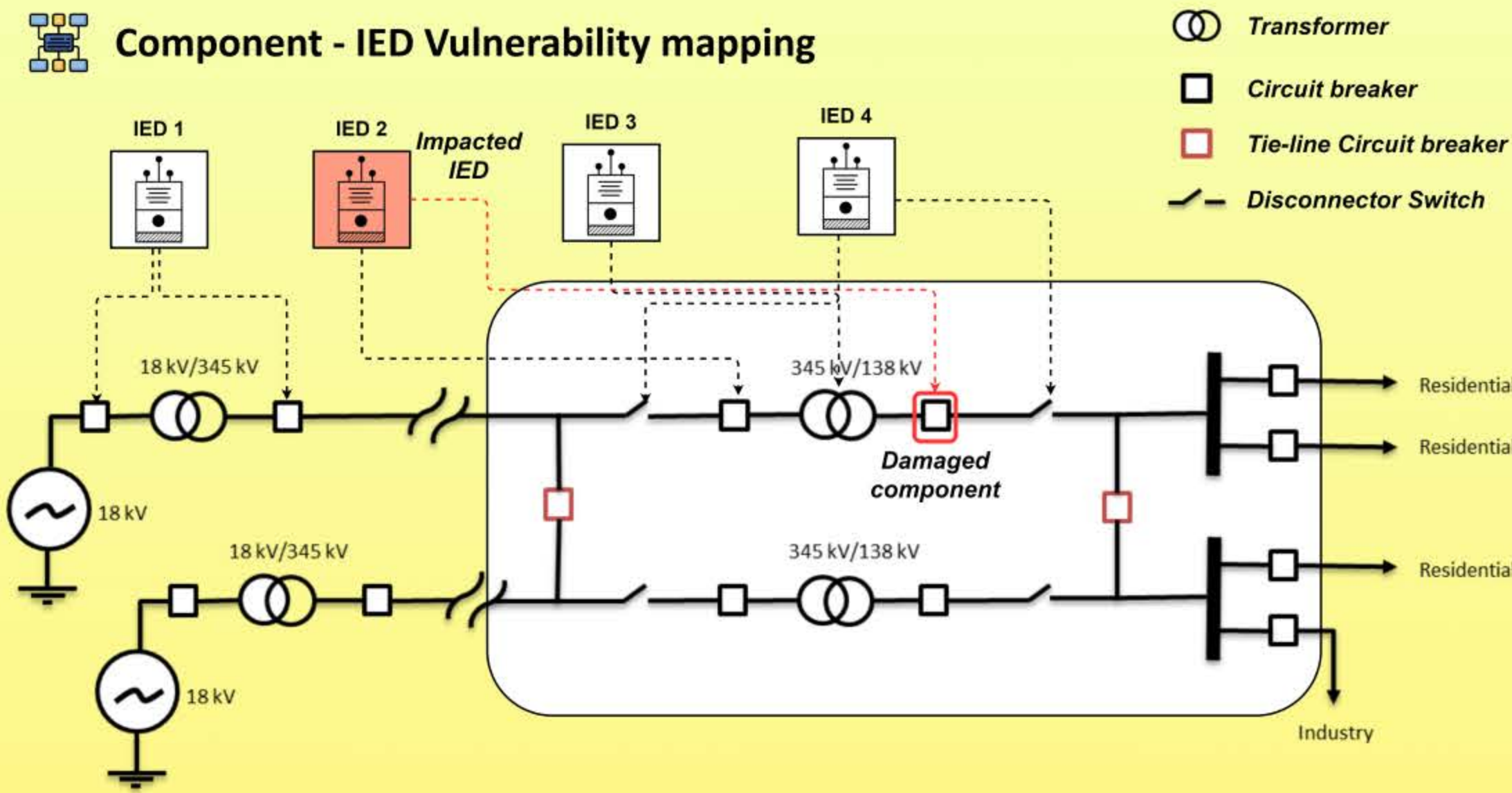


## Abstract

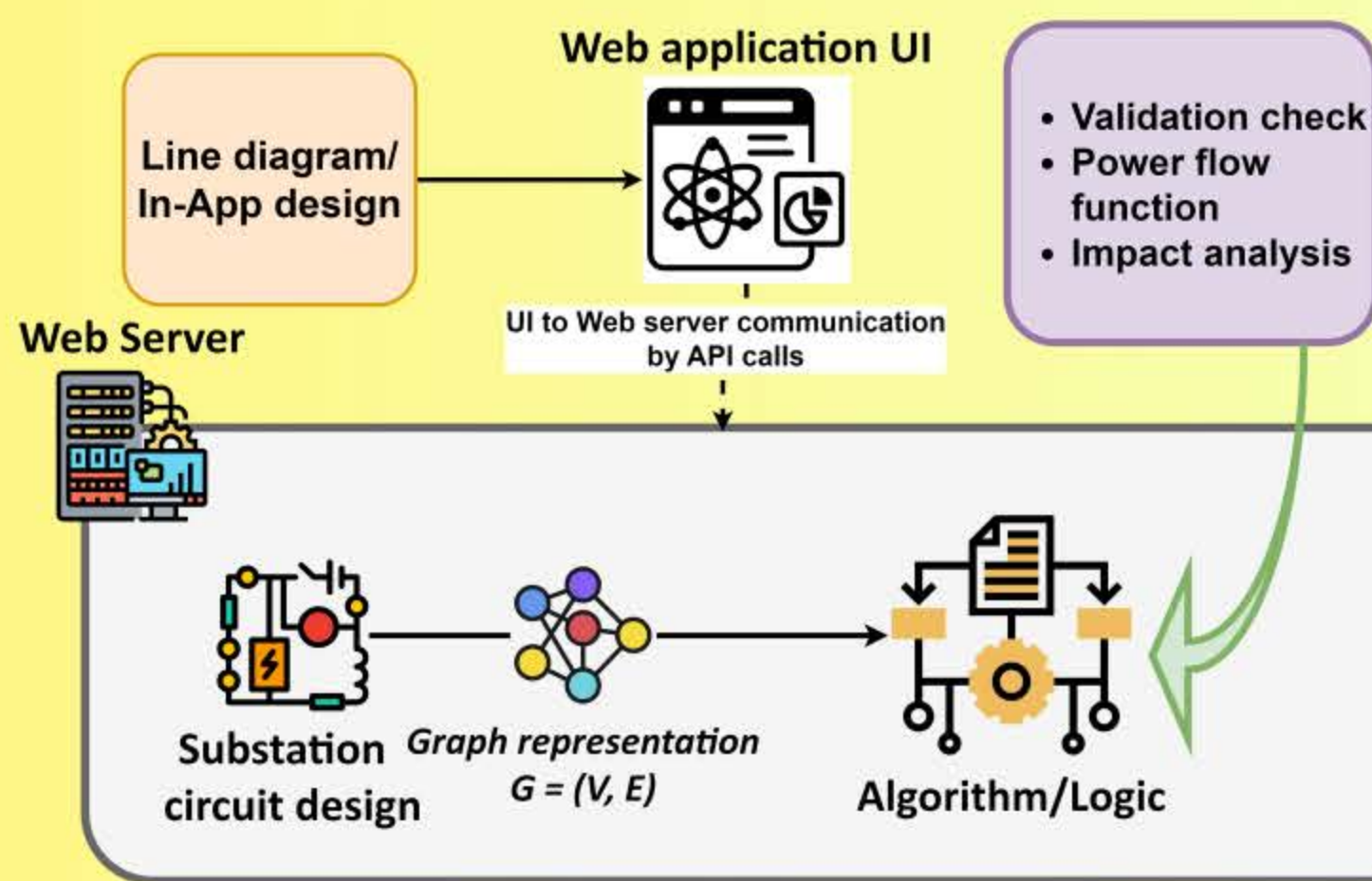
The advancement of smart grid technologies has significantly improved the automation and control of power distribution from generation to consumers, positioning substations at the core of these operations. However, the integration with cyber infrastructure has made substations increasingly vulnerable to cyber-attacks. While risk assessments are commonly employed to address these threats, impact assessments remain relatively rare. In particular, the investigation of how attacks targeting Intelligent Electronic Devices (IEDs)—often the primary focus of cyber-attacks—affect the physical components of substations has not been thoroughly studied, especially across various bus bar arrangement scenarios. This research aims to model the impact of cyber-attacks from a physical domain perspective to identify the most critical physical devices within substations. By mapping these devices to their controlling cyber components, we seek to enhance the understanding of potential vulnerabilities. To achieve this, we are developing a proof-of-concept tool that analyzes the Single Line (SL) diagram of a substation to generate comprehensive impact assessments. This tool will highlight the most critical IEDs, thereby aiding in the formulation of more effective defense strategies against cyber threats in smart grid environments.



## Component - IED Vulnerability mapping



## Proposed tool's System Design



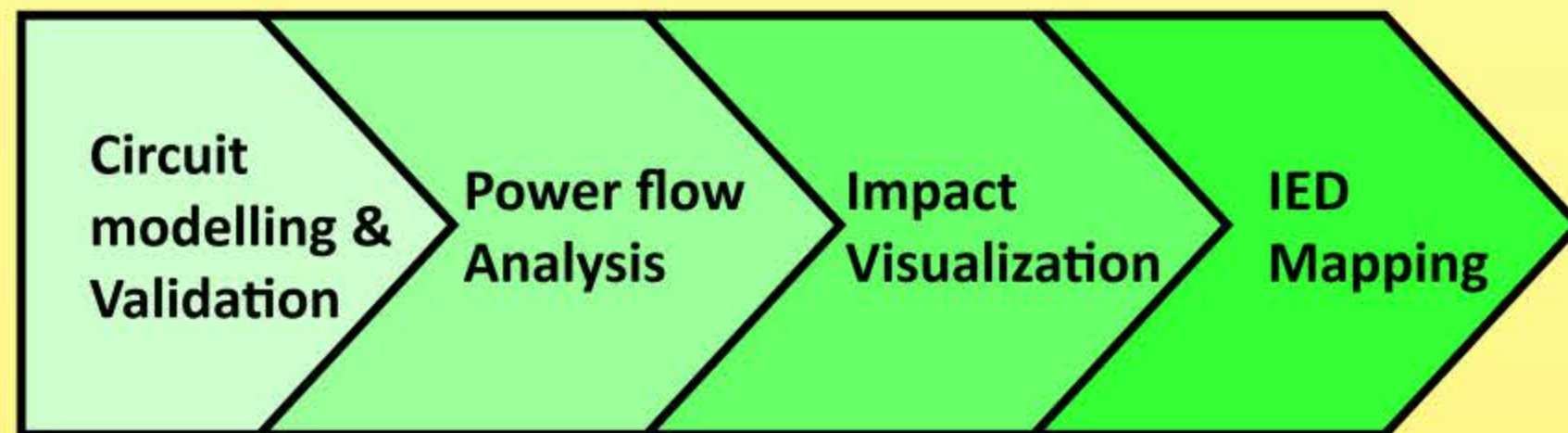
## The Maths behind the hood

Aspect	Notation in mathematical form
Substation	$\Gamma = (\Phi, \Lambda)$ $\Gamma$ denotes the substation as graphical representation
Electrical Components	$\Phi = \{\phi_1, \phi_2, \dots, \phi_n\}$ $\Phi$ denotes the set of electrical component and $\phi_i$ denotes the individual components
Line Connections	$\Lambda = \{(\phi_i, \phi_j), \dots\} \mid \phi_i, \phi_j \in \Phi$ $\Lambda$ denotes the set of electrical connections in the substation.
Power flow function	$P_{f \rightarrow \phi_i, \phi_j} = f_p(\phi_i, \phi_j) \begin{cases} 1 \\ 0 \end{cases}$ $P_{f \rightarrow \phi_i, \phi_j}$ is the power flow status between $\phi_i$ and $\phi_j$
Importance of component	$\mathbb{I}_{\phi_i} = \frac{w(\phi_i)}{W} \mid W = \sum_{j=1}^n w(\phi_j) \mid \forall \phi_j \in \Phi$ $\mathbb{I}_{\phi_i}$ is the importance of the electrical component; $W$ is the cummulative weight of each component in the substation
Impact Analysis	$Impact(\phi_i) = \frac{\sum_{\phi_j \in A(\phi_i)} Impact(\phi_j)}{\max_{\phi_j \in \Phi} Impact(\phi_j)} \mathbb{I}_{\phi_j} \cdot  A(\phi_i) $ represents the impact of component $\phi_i$ $A(\phi_i)$ is the set of components affected by $\phi_i$ ; $\mathbb{I}_{\phi_j}$ is the importance of component $\phi_j$ ; $ A(\phi_i) $ is the number of affected components; $\max_{\phi_j \in \Phi} Impact(\phi_j)$ is the maximum impact across all components in the set $\Phi$ .

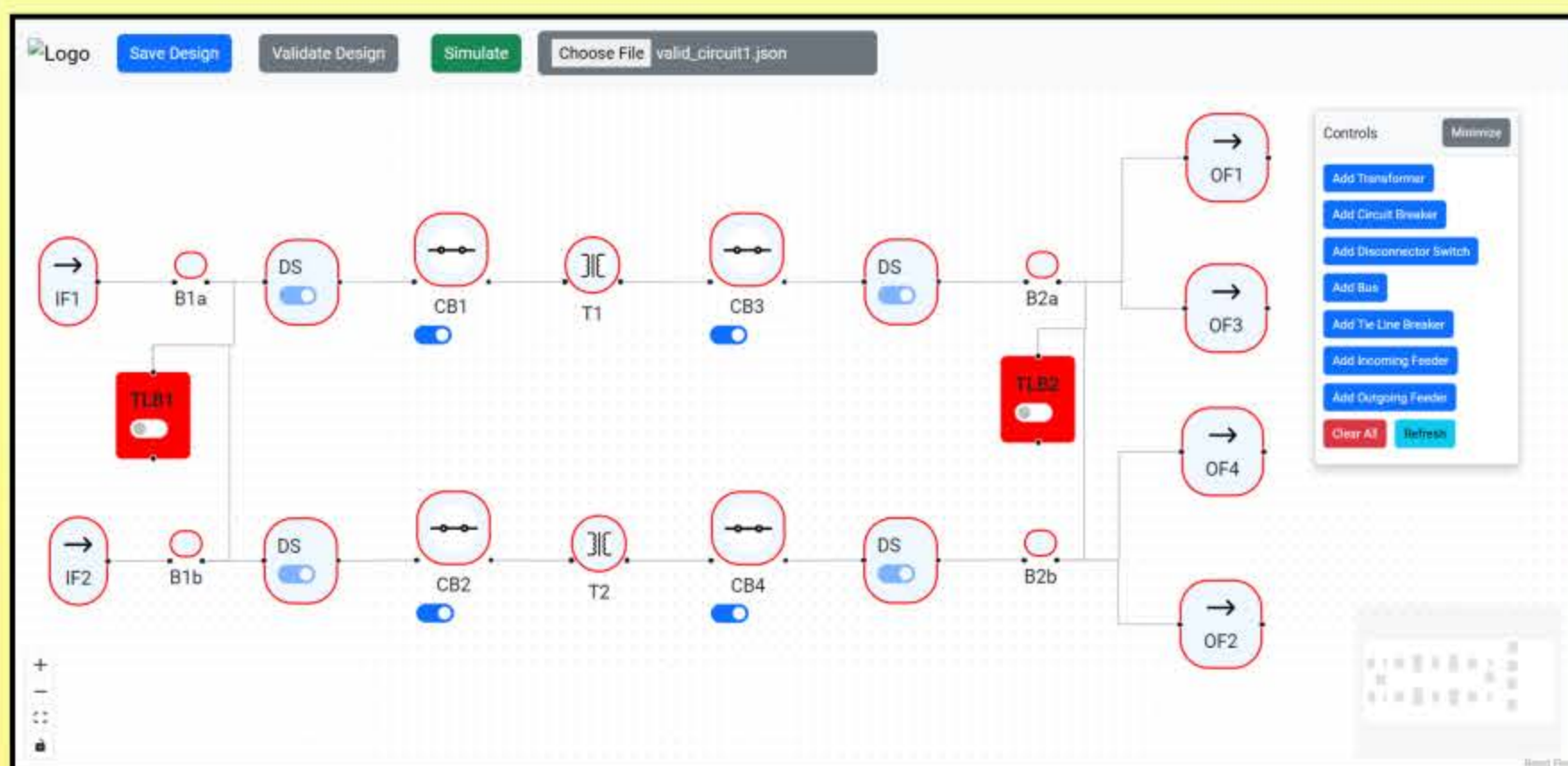


## Data Flow and System Integration

The proposed tool will model the substation by allowing users to upload a line diagram or use a drag-and-drop designer to position components. The core representation of the substation diagram is based on graph theory. The tool validates the circuit to ensure compliance with standards and simulates power flow. By applying network theory, it calculates the impact of potential component failures. Additionally, by backtracking each component's Intelligent Electronic Device (IED), the tool identifies vulnerable IEDs, providing operators with the most up-to-date information for informed decision-making.



## Prototype of the proposed tool







# RuleSense: Efficient and Lightweight Rapid Rule-Based Anomaly Detection for IIoT Streaming Data

Amir Firouzi, Sajjad Dadkhah, Heather Molyneaux, Ali A. Ghorbani\*

Contact Email: amir.firouzi@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)

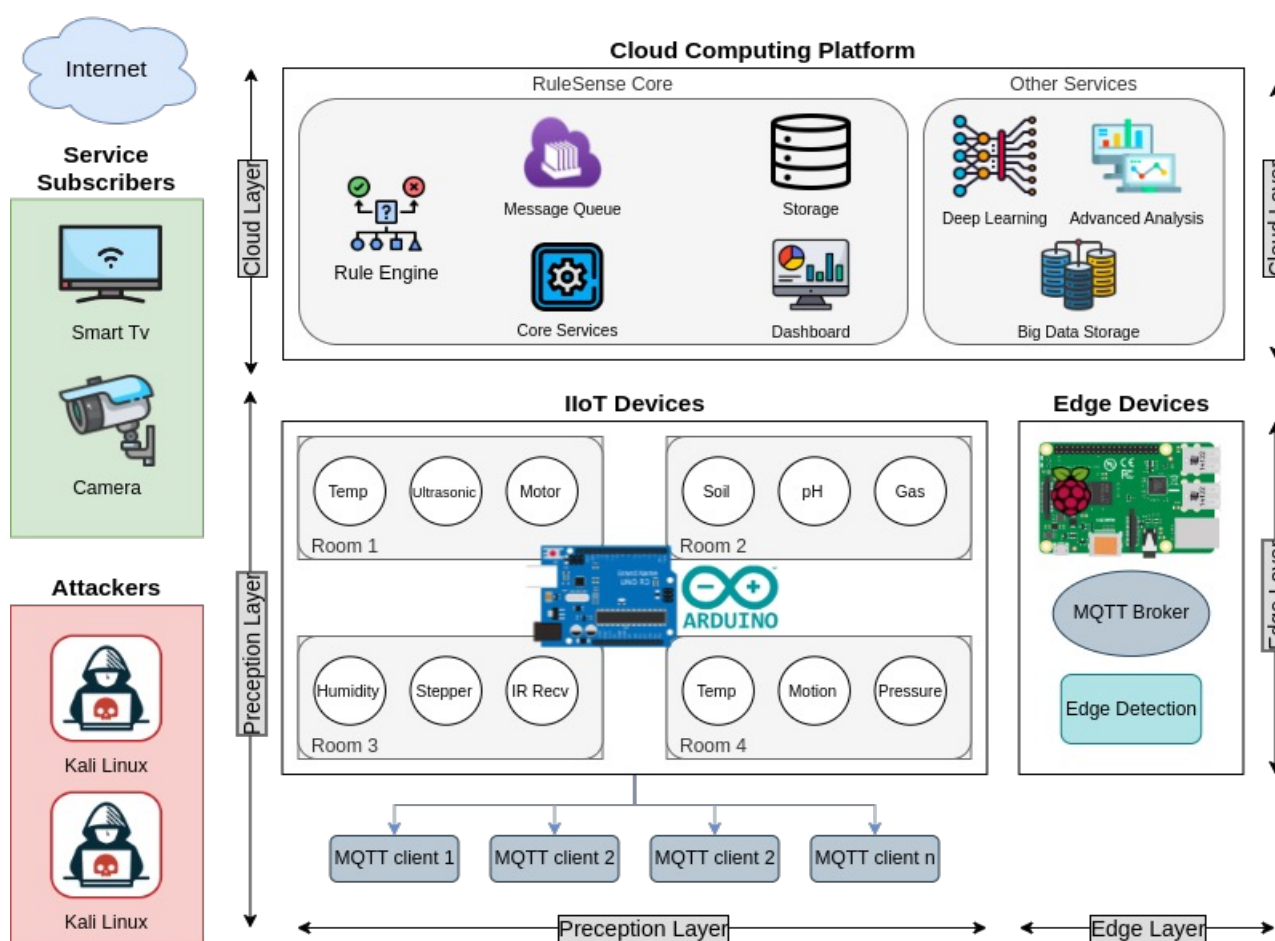


## ABSTRACT

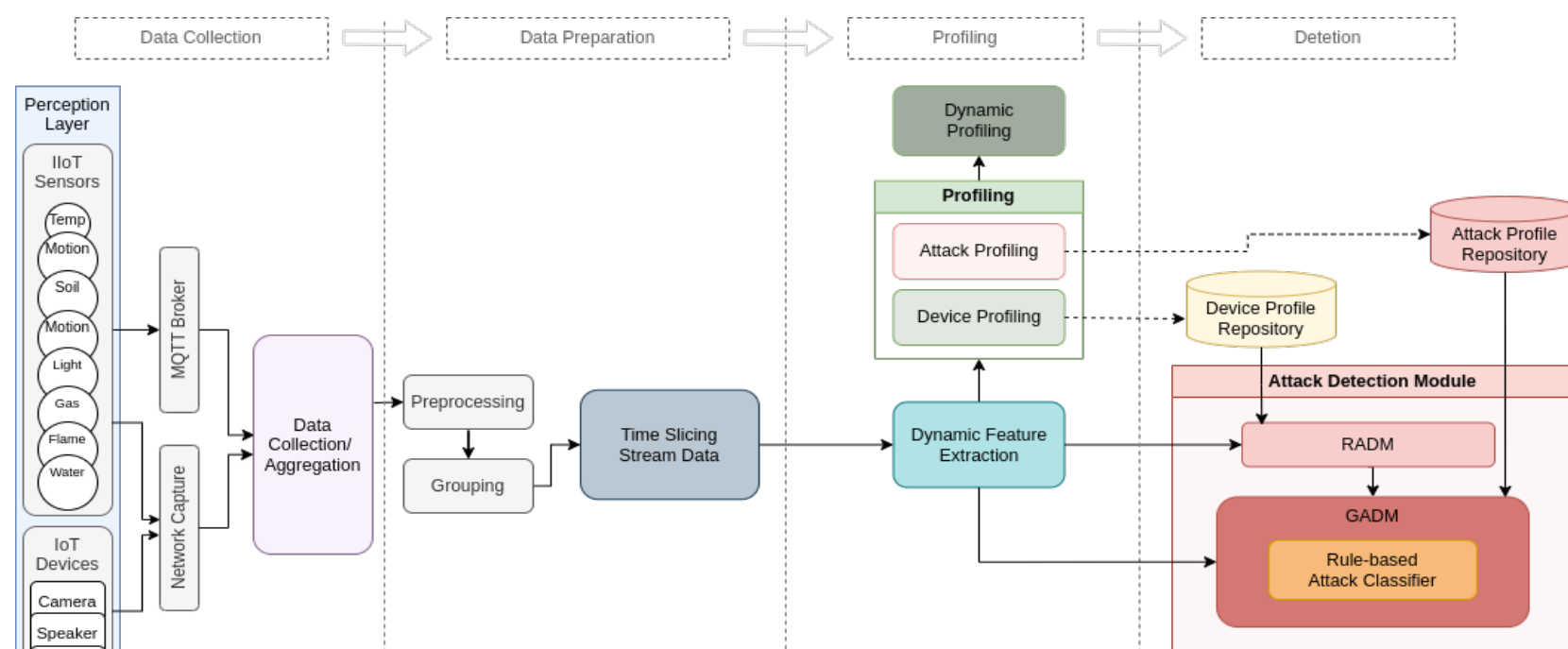
The rapid growth of Industrial Internet of Things (IIoT) technology generates vast amounts of sensor data, making anomaly detection essential for maintaining performance, security, and sustainability. This paper introduces RuleSense, a rule-based anomaly detection framework specifically designed for IIoT environments. RuleSense employs a three-layer architecture—perception, edge, and cloud—integrating both network and sensor data to enable early anomaly detection at the edge. This reduces latency, enhances system responsiveness, and minimizes resource consumption, particularly beneficial in environments with unreliable internet connections. By profiling normal device behavior, RuleSense effectively detects anomalies without training and handles unbalanced, real-time data streams. It achieves a detection accuracy of 99.35% in IIoT environments and 92% in IoT settings, along with high F1-scores. To the best of our knowledge, RuleSense is the first rule-based anomaly detection framework tailored for IIoT, offering a robust solution for real-time anomaly detection.

## Framework Architecture

- Perception Layer:**
  - Data Sensing
  - Data Capture
  - Mqtt Broker
  - Transmission to Edge
- Edge Layer:**
  - Local Data Processing
  - Rule-based Anomaly Detection
- Cloud Layer:**
  - Data Storage
  - Advanced Rule Extraction
  - In-depth analysis



## Framework Workflow



The RuleSense framework consists of four phases: **Data Collection** gathers sensor and network data; **Data Preparation** cleans, groups, and time-slices data; **Profiling** creates device and attack profiles through feature extraction and dynamic updates; and **Detection** uses RADM for edge anomaly detection and GADM in the cloud for attack classification.

## Profiling Algorithm

The profiling algorithm includes the **Device Profiling Algorithm**, which captures normal device behavior using feature vectors, distance calculations (Manhattan for numerical, Jaccard for string features), and weighted profile vectors. The **Attack Profiling Algorithm** characterizes attacks by comparing attack data against device profiles, using distance metrics to measure deviations.

These profiles are generated through a series of mathematical steps, including mean distance calculations and weighted vector products, enabling effective anomaly detection and attack classification.

$$\mathbf{V}_{A_i f_j} = \{\mathbf{v}_{A_i f_j S_1}, \mathbf{v}_{A_i f_j S_2}, \mathbf{v}_{A_i f_j S_3}, \dots\} \quad \text{for all } A_i \in \mathcal{A}, f_j \in \mathcal{F} \quad (7)$$

$$\mathbf{V}_{A_i f_j} = \{\mathbf{v}_{S_1}, \mathbf{v}_{S_2}, \mathbf{v}_{S_3}, \dots\} \quad \text{where } \mathbf{v}_{S_i} = A_i f_j S_i \quad \text{for all } A_i \in \mathcal{A}, f_j \in \mathcal{F} \quad (8)$$

$$\text{dist}_{A_i f_j, P_{D_i f_j}} = \begin{cases} \text{distance}(A_i f_j S_1, P_{D_i f_j}), \\ \text{distance}(A_i f_j S_2, P_{D_i f_j}), \\ \dots, \\ \text{distance}(A_i f_j S_n, P_{D_i f_j}) \end{cases} \quad (9)$$

$$\mathbf{V}_{D_i f_j} = \{\mathbf{v}_{D_i f_j S_1}, \mathbf{v}_{D_i f_j S_2}, \mathbf{v}_{D_i f_j S_3}, \dots\} \quad \text{for all } D_i \in \mathcal{D}, f_j \in \mathcal{F}$$

$$\mathbf{V}_{D_i f_j} = \{\mathbf{v}_{S_1}, \mathbf{v}_{S_2}, \mathbf{v}_{S_3}, \dots\} \quad \text{where } \mathbf{v}_{S_i} = D_i f_j S_i \quad \text{for all } D_i \in \mathcal{D}, f_j \in \mathcal{F}$$

$$\text{dist}_{D_i f_j} = \frac{1}{n} \sum_{k=1}^n \text{distance}(D_i f_j S_k)$$

where  $n$  is the number of slices, and

$$\text{distance}(D_i f_j S_k) = \begin{cases} \text{Manhattan\_dist}(D_i f_j S_k) & \text{if } f_j \text{ is a numerical feature} \\ \text{Jaccard\_dist}(D_i f_j S_k) & \text{if } f_j \text{ is a string feature} \end{cases} \quad (10)$$

$$P_{D_i f_j} = \text{dist}_{D_i f_j} \cdot \mu \quad (11)$$

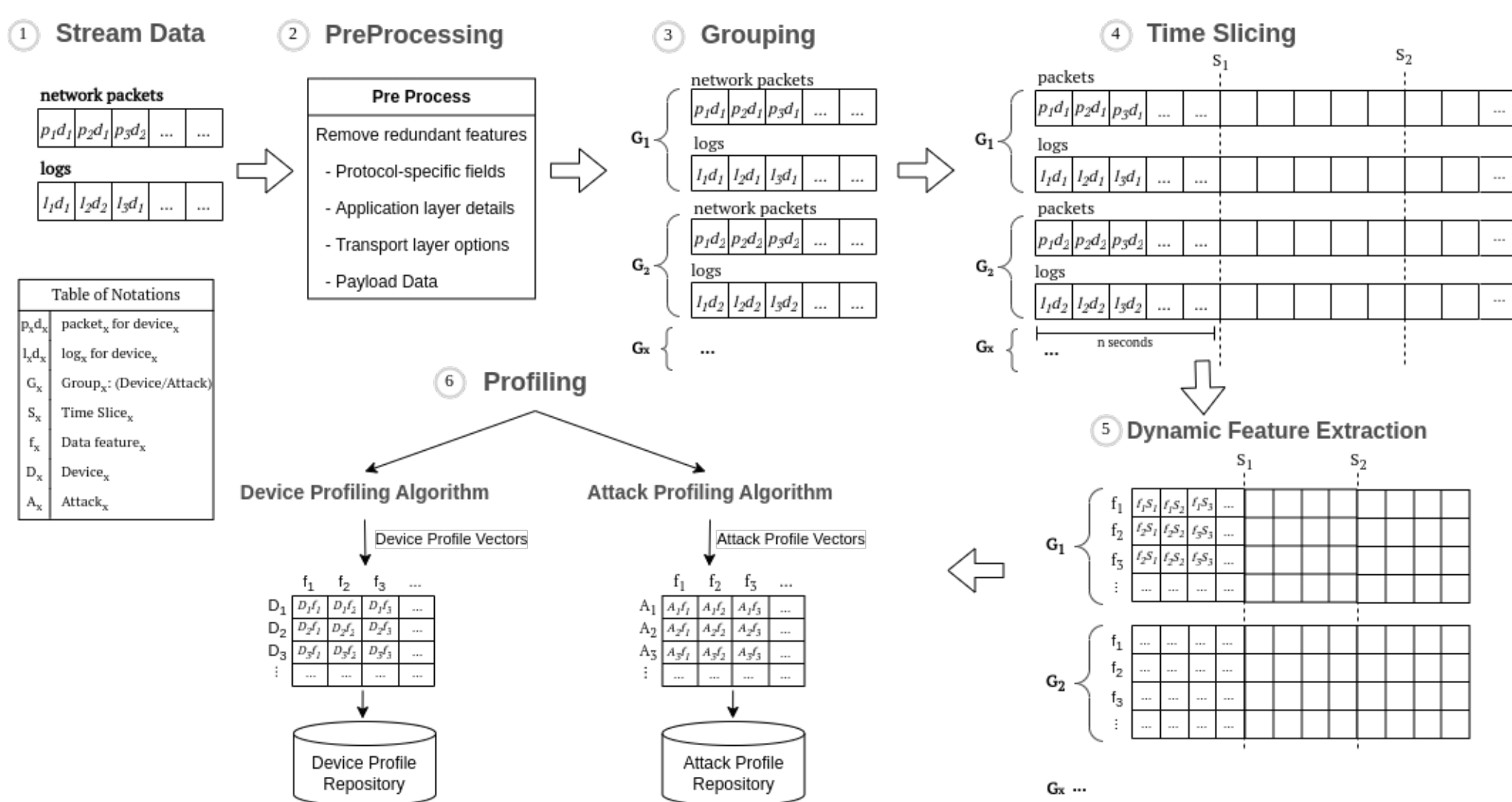
$$P_{D_i} = \{P_{D_i f_1}, P_{D_i f_2}, P_{D_i f_3}, \dots\} \quad \text{for all } D_i \in \mathcal{D} \quad (12)$$

$$\mathbf{P}_{A_i} = \{P_{A_i f_1}, P_{A_i f_2}, P_{A_i f_3}, \dots\} \quad \text{for all } A_i \in \mathcal{A} \quad (13)$$

### Device Profiling

### Attack Profiling

## Profiling Mechanism



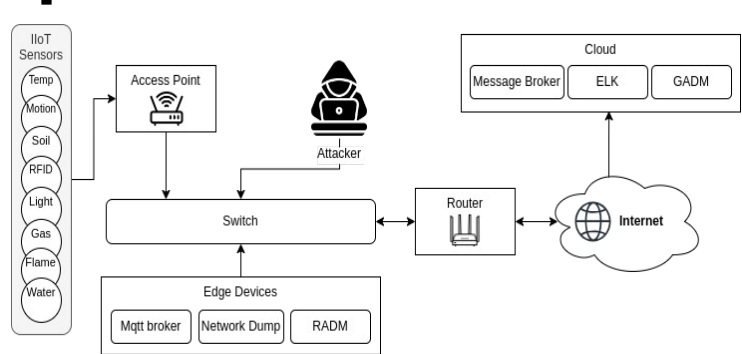
The profiling mechanism includes six steps: **Stream Data** handles unordered stream data; **Preprocessing** removes redundant features; **Grouping** organizes data by devices or attacks; **Time Slicing** segments data into fixed intervals; **Dynamic Feature Extraction** extracts features into vectors; and **Profiling** generates device and attack profiles using dedicated algorithms.

## Testbed and Implementation

A real IIoT testbed generates authentic data for RuleSense, featuring sensors, an edge layer with an MQTT broker, RADM for threat detection, and cloud integration for advanced analysis.



We built our IIoT devices with 15 Arduino boards equipped with industrial sensors. Data is transmitted via WiFi using MQTT for reliable, fast transfer, with each board sending data to a distinct broker topic

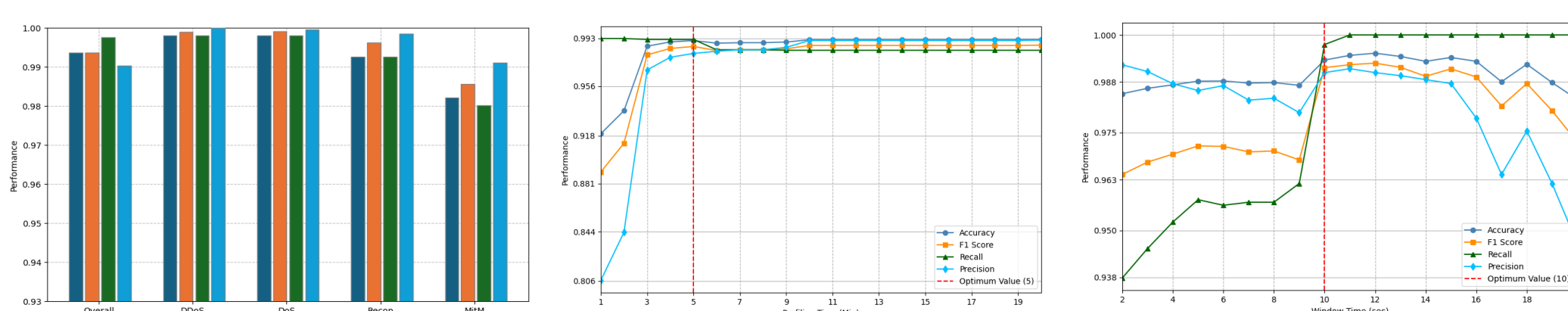


The architecture enables data collection with IIoT sensors, an MQTT broker, network dump tool, RADM for threat detection, and cloud integration, ensuring scalable and robust data handling for RuleSense.

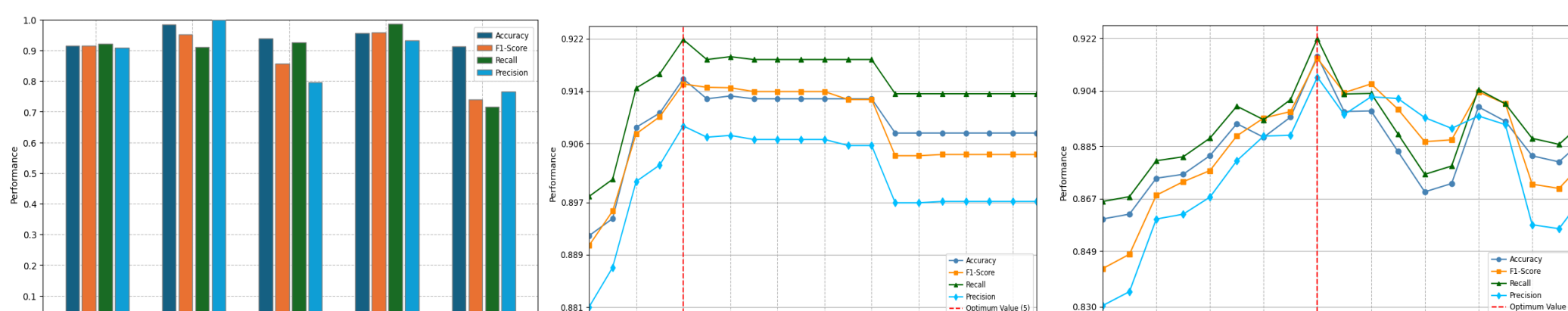
No.	Application	IIoT Sensors	Sensor Types
1	Weather	Temperature & Humidity	DHT11
		Linear temperature	LM35
		Analog temperature	KY-013
2	Soil Moisture	Digital temperature	DS18B20
		Atmospheric pressure	BMP-180
		Soil Moisture	YL-69
3	Sound	Big Sound Detector	KY-037
4	Motion	Small Sound Detector	KY-038
		PIR Motion Sensor	HC-SR501
		Ceramic Vibration Sensor	SW-420
5	Vibration	Water Level Sensor	YL-83
6	Water	Analog gas detector	MQ-2
7	Gas	Analog Alcohol detector	MQ-3
8	Steam	Steam Sensor	KSO203
9	RFID	RFID Sensor	RFID-RC522
10	Accelerometer & Gyroscope	Inertial Digital Acceleration	ADXL345
11	Proximity	ALS Infrared LED Optical	APDS-9930
12	Collision	Collision (Crash Sensor)	KY-031
13	Ultrasonic	Ultrasonic Sensor	HC-SR04
14	Flame	Flame Detector	KY-026
15	Light & Gesture	Light & Gesture Detection Sensor	APDS-9960

## Experiments and Evaluations

**Efficiency of RuleSense Attack Detection:** We evaluated the attack detection algorithm on our dataset generated from the testbed.



**Efficiency of RuleSense Attack Classification:** We evaluated the attack classification algorithm on our dataset generated from the testbed.







# FIGS: A Lightweight Intrusion Detection Framework For Highly Imbalanced IoT Environments

Zeynab Anbiaee, saba.anbiaee@unb.ca

Sajjad Dadkhah, sdadkhah@unb.ca

Ali A. Ghorbani, ghorbani@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)



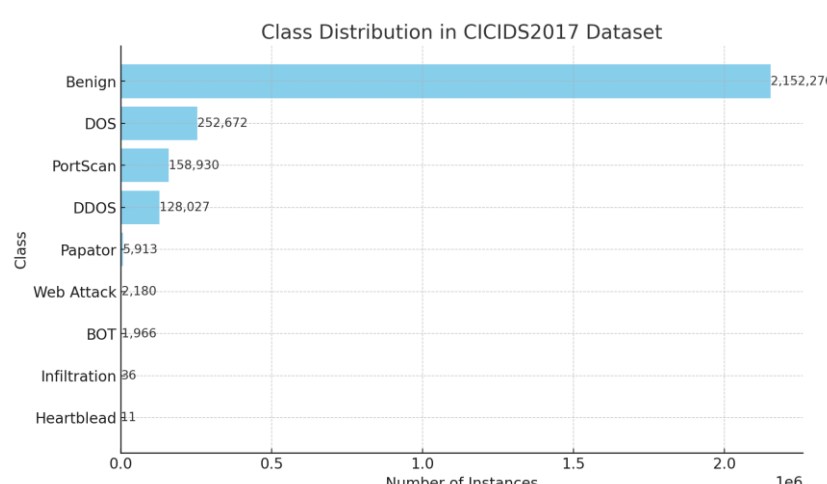
## ABSTRACT

The development of IoT environments has increased the challenges of security against cyber threats and attacks, especially because of the imbalanced nature of attack traffic where underrepresented but critical attacks were ignored. Traditional IDS often fail to provide a balance between the detection rate of majority and minority classes. We propose FIGS (Feature Importance GAN SMOTE), an innovative Lightweight Intrusion Detection Framework designed to address the challenge of class imbalance in IoT environments. FIGS integrates sensitivity-based feature importance analysis, Generative Adversarial Networks(GAN), and Synthetic Minority Over-sampling Technique(SMOTE) to generate high-quality synthetic data for minority attack classes. FIGS enhanced minority class detection while lessening the computational overhead and effectively reducing noise during data generation. FIGS substantially improves the detection rate and decreases the false positive rate for minority categories, particularly the Bot attacks that state-of-the-art algorithms struggled with.

## MOTIVATION

- The problem of class imbalance can result in biased models favoring majority classes, making the environment more vulnerable to intrusions.
- Our model uses GAN not only to generate synthetic data but also to identify and utilize important features.
- The proposed Generalized Imbalance Ratio (GIR) addresses the limitation of traditional imbalance metrics by incorporating not only the sample size but also weighting factors that reflect the importance of each class.

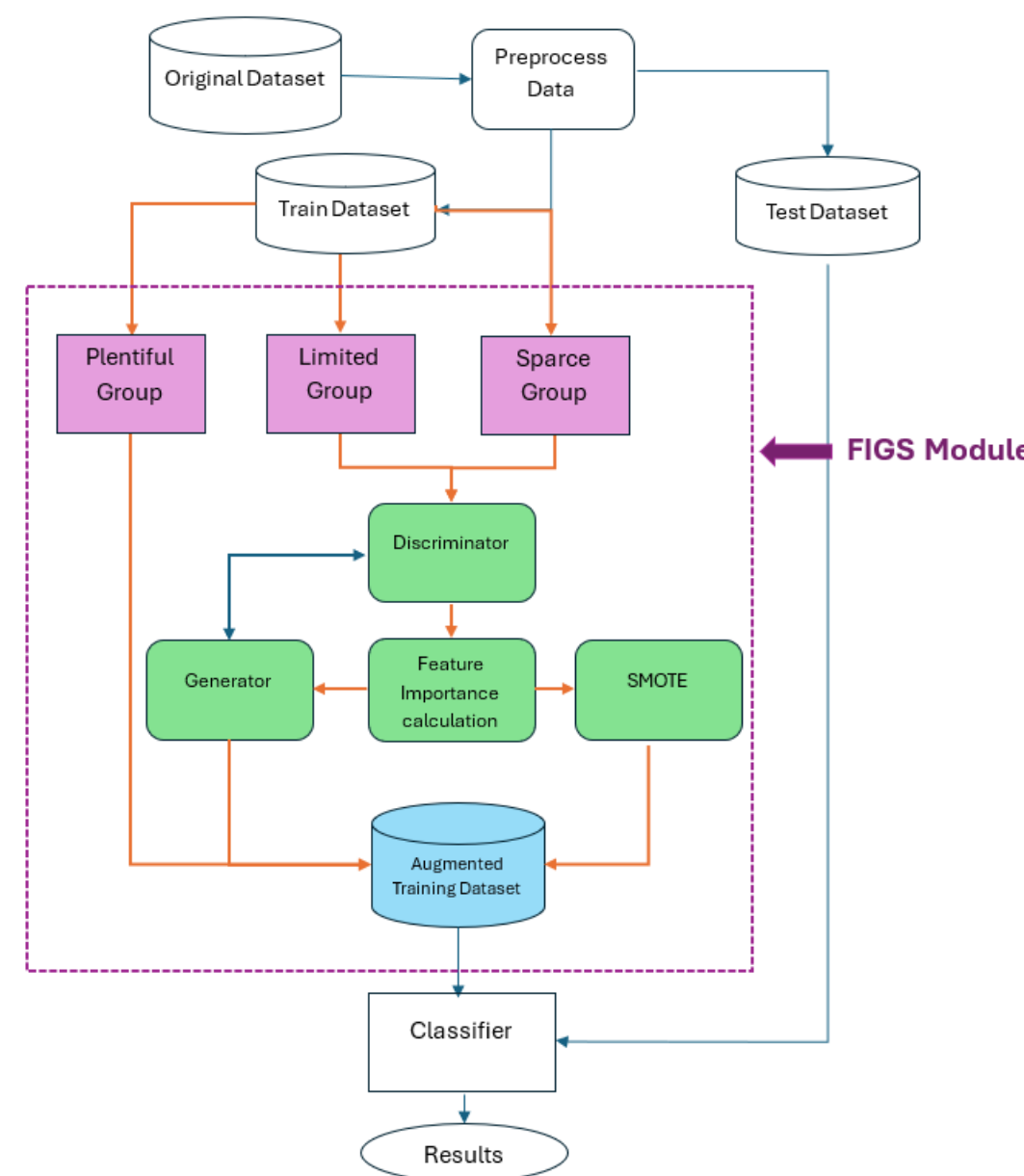
$$GIR = \frac{w_{maj} \times n_{maj}}{w_{min} \times n_{min}}$$



GIR VALUES AND CATEGORIES FOR EACH CLASS

Class	GIR Value	Category
Benign	0.5	Plentiful
DOS	4.26	Plentiful
PortScan	6.77	Plentiful
DDOS	8.41	Plentiful
Patator	182.02	Limited
Web Attack	493.84	Limited
BOT	547.71	Limited
Infiltration	29,900.99	Sparse
Heartbleed	97,830.73	Sparse

## Proposed FIGS Model Framework



## FIGAN

FIGAN operates by integrating CGAN with a feature selection method that dynamically identifies the most important features using sensitivity analysis, where small perturbations in input features are used to measure their impact on the discriminator's output. This process helps FIGAN generate targeted synthetic data by focusing on the most critical features, enhancing the classification ability while reducing unnecessary data complexity. Unimportant features are set to zero, ensuring that FIGAN remains lightweight and suitable for resource-constrained IoT environments.

$$I_i = |D(\mathbf{x}) - D(\mathbf{x} + \epsilon \cdot \mathbf{e}_i)|$$

## FSMOTE

FISMOTE refines the traditional SMOTE approach by applying oversampling only to the most important features identified through sensitivity analysis, which measures the impact of small changes in each feature on the discriminator's output. This targeted method generates higher-quality synthetic data that better represents minority classes, reducing noise and enhancing the overall effectiveness of the augmentation process.

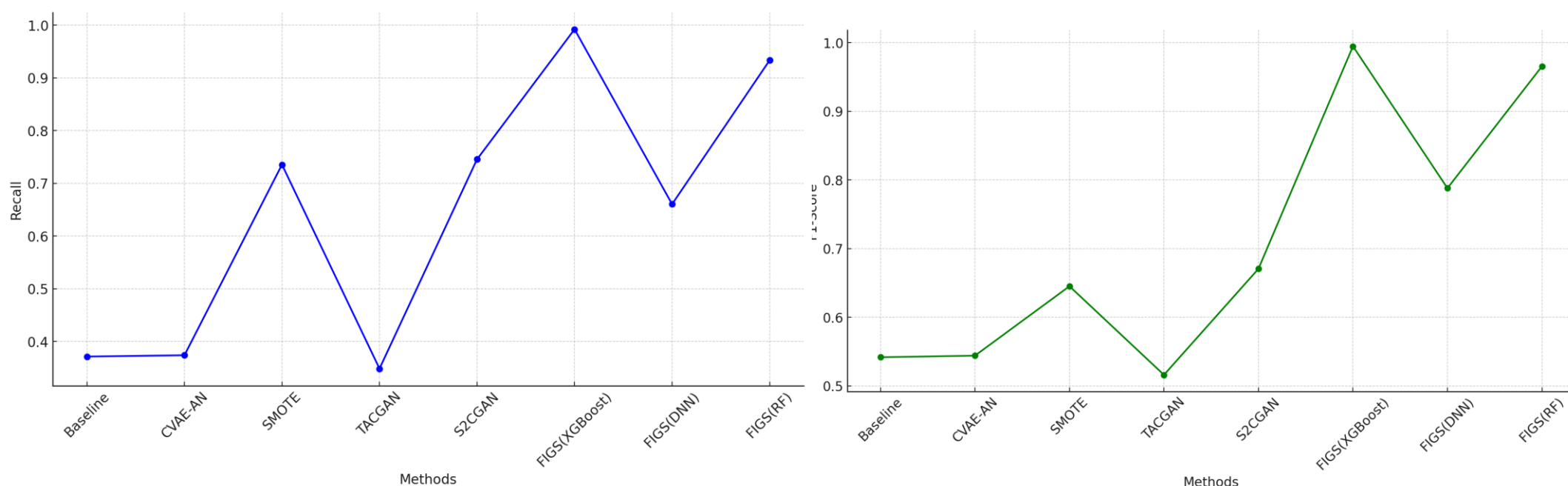
$$I_i = |D(\mathbf{x}) - D(\mathbf{x} + \epsilon \cdot \mathbf{e}_i)|$$

## Experimental Analysis

Experimental Results Of Multiclassification Include The Original And Four State-Of-The-Art-Model And FIGS with Three Different Classifiers

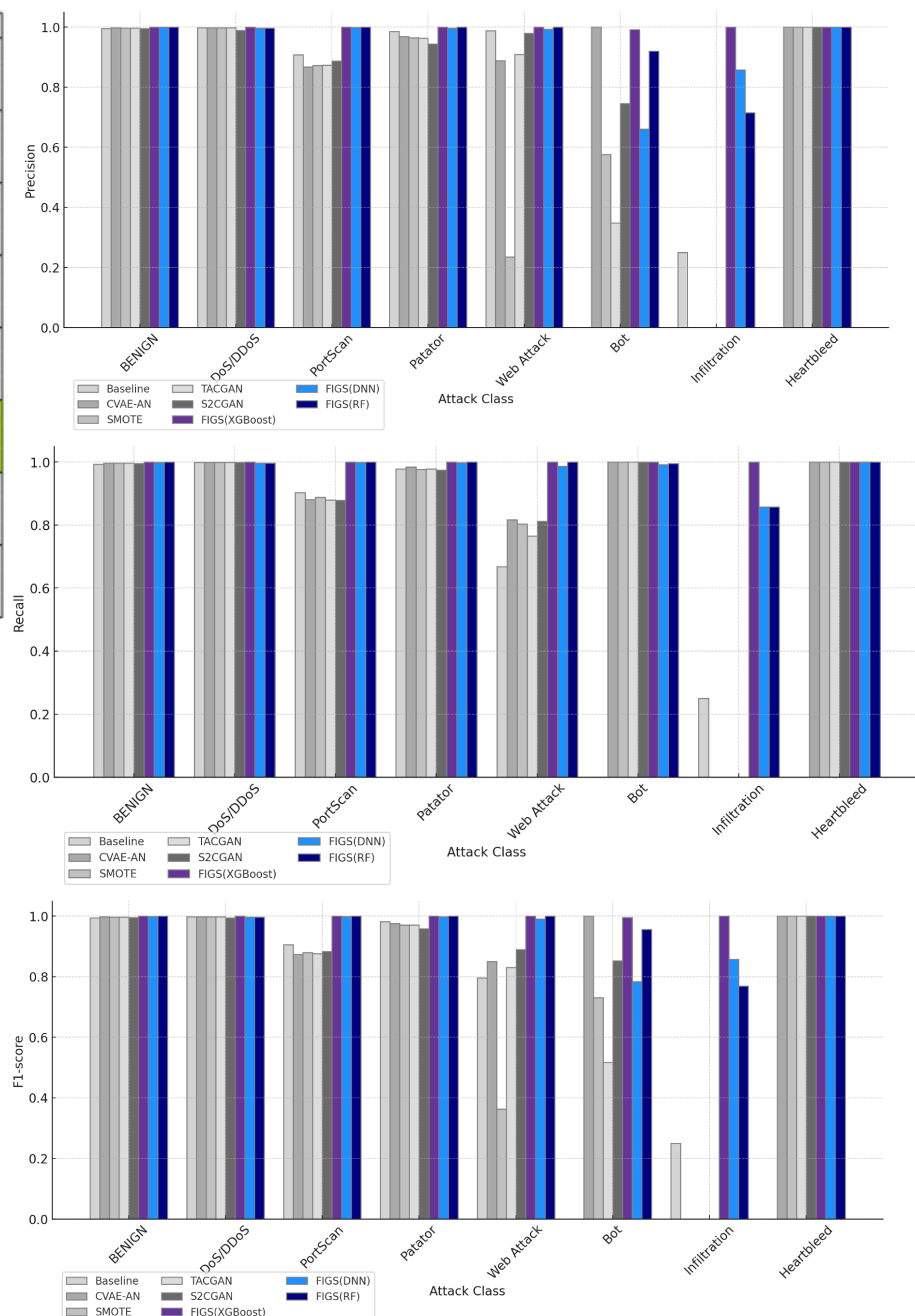
Methods	Metric	Baseline	CVAE-AN	SMOTE	TACGAN	S2CGAN	FIGS(XGBoost)	FIGS(DNN)	FIGS(RF)
BENIGN	Precision	0.9949	0.9979	0.9971	0.9965	0.9955	1.000	0.9998	0.9999
	Recall	0.9916	0.9877	0.9847	0.9898	0.9903	1.000	1.000	0.9999
	F1-Score	0.9932	0.9928	0.9908	0.9931	0.9929	1.000	0.9999	0.9999
DoS/DDoS	Precision	0.9984	0.9904	0.9878	0.9904	0.9896	0.9981	0.9695	0.9996
	Recall	0.9994	0.9984	0.9993	0.9990	0.9993	0.9988	0.9993	0.9963
	F1-Score	0.9944	0.9949	0.9935	0.9947	0.9944	0.9984	0.9842	0.9965
PortScan	Precision	0.9082	0.8671	0.8722	0.8878	0.9016	1.000	1.000	0.9999
	Recall	0.9369	0.9818	0.9644	0.9607	0.9417	1.000	0.9994	0.9999
	F1-Score	0.9223	0.9209	0.9160	0.9228	0.9212	1.000	0.9997	0.9999
Patator	Precision	0.9856	0.9682	0.9845	0.9824	0.9433	0.9996	0.9971	1.0000
	Recall	0.9884	0.8922	0.9895	0.9895	0.9924	1.0000	0.9931	0.9996
	F1-Score	0.9870	0.9785	0.9870	0.9860	0.9672	0.9998	0.9951	0.9998
Web Attack	Precision	0.9874	0.8881	0.2350	0.9000	0.9794	1.0000	0.9934	1.0000
	Recall	0.9358	0.8922	0.9404	0.9083	0.9794	0.9892	0.9698	0.9720
	F1-Score	0.9566	0.8902	0.3760	0.9041	0.9794	0.9946	0.9815	0.9858
Bot	Precision	1.0000	1.0000	0.5746	0.9928	0.6091	0.9973	0.9763	1.0000
	Recall	0.3715	0.3740	0.7354	0.3486	0.7455	0.9920	0.6604	0.9332
	F1-Score	0.5417	0.5444	0.6451	0.5160	0.6705	0.9946	0.7879	0.9654
Infiltration	Precision	0.0000	0.2500	0.8333	0.0000	1.0000	1.0000	1.0000	1.0000
	Recall	0.0000	0.1429	0.7143	0.0000	0.7143	0.8571	0.7143	0.7143
	F1-Score	0.0000	0.1818	0.7692	0.0000	0.8333	0.9231	0.8333	0.8333
Heartbleed	Precision	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000
	Recall	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000
	F1-Score	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000

Metric Comparison Across Different Models For Bot Attack Class



FIGS outperforms state-of-the-art methods in detecting minority attack classes, such as Bot, Infiltration, and Heartbleed, and enhances recall and F1 scores without adding unnecessary complexity or computation. It consistently delivers superior or matching results in Plentiful categories while being more computationally efficient, making it an effective and reliable solution for real-world intrusion detection in highly imbalanced datasets.

Metric Comparison Across Different Models for All The Classes





ABSTRACT

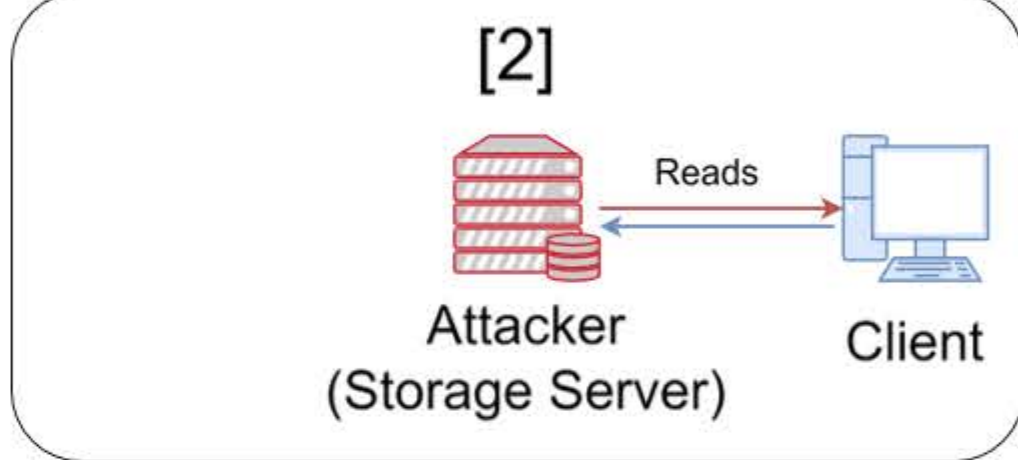
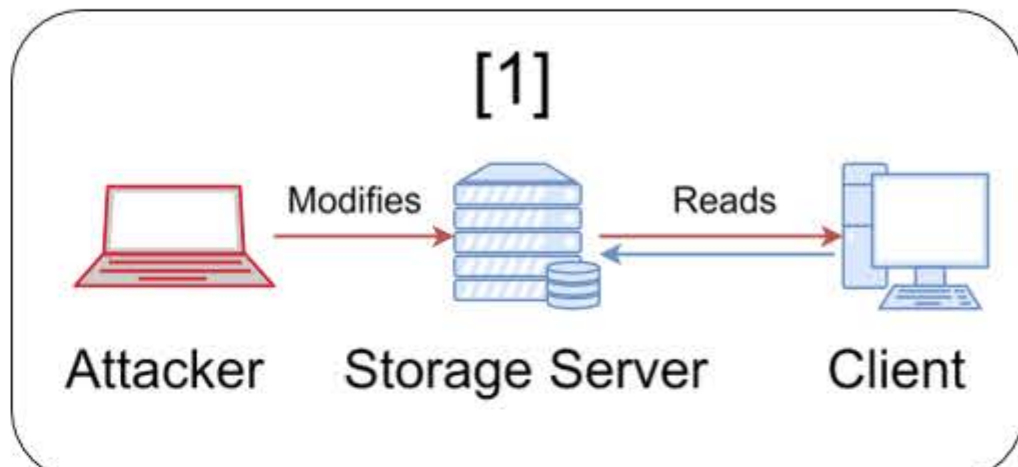
The use of software-defined storage (SDS) systems to store sensitive data is becoming increasingly prevalent. However, these systems primarily implement security measures to ensure the confidentiality and availability of stored data, with limited consideration for the protection of its integrity. This presents a significant risk; Therefore, we propose a methodology to incorporate integrity-protecting measures for ciphertext stored at rest. To demonstrate the practical challenges and opportunities of such measures we integrated “authenticated encryption with associated data” (AEAD) ciphers into the widely used SDS system Ceph, specifically, into its block storage interface, to secure the integrity of stored data and metadata. Ultimately, we identify the characteristics that an SDS system should possess to adopt our methodology.

Ceph Architecture

- Software Defined Storage System
  - Decouple Hardware
- Object Storage backend
  - Metadata
  - Object Data
- Exposes data in multiple formats
  - File, Block, Object Interface
- Integrated into serval Cloud Platforms
  - OpenStack, Kubernetes, etc.
- Create private cloud platform
  - Data Sovereignty
  - Store Sensitive Data

Problem Statement

- Distributed System
  - Client get data across network
  - Private or public service
- Client-side encryption
  - No cryptographic method to establish integrity
  - Block Storage often none
- Client should not have to trust storage server
- Focus on Ceph block storage interface
- How to minimize runtime overhead?

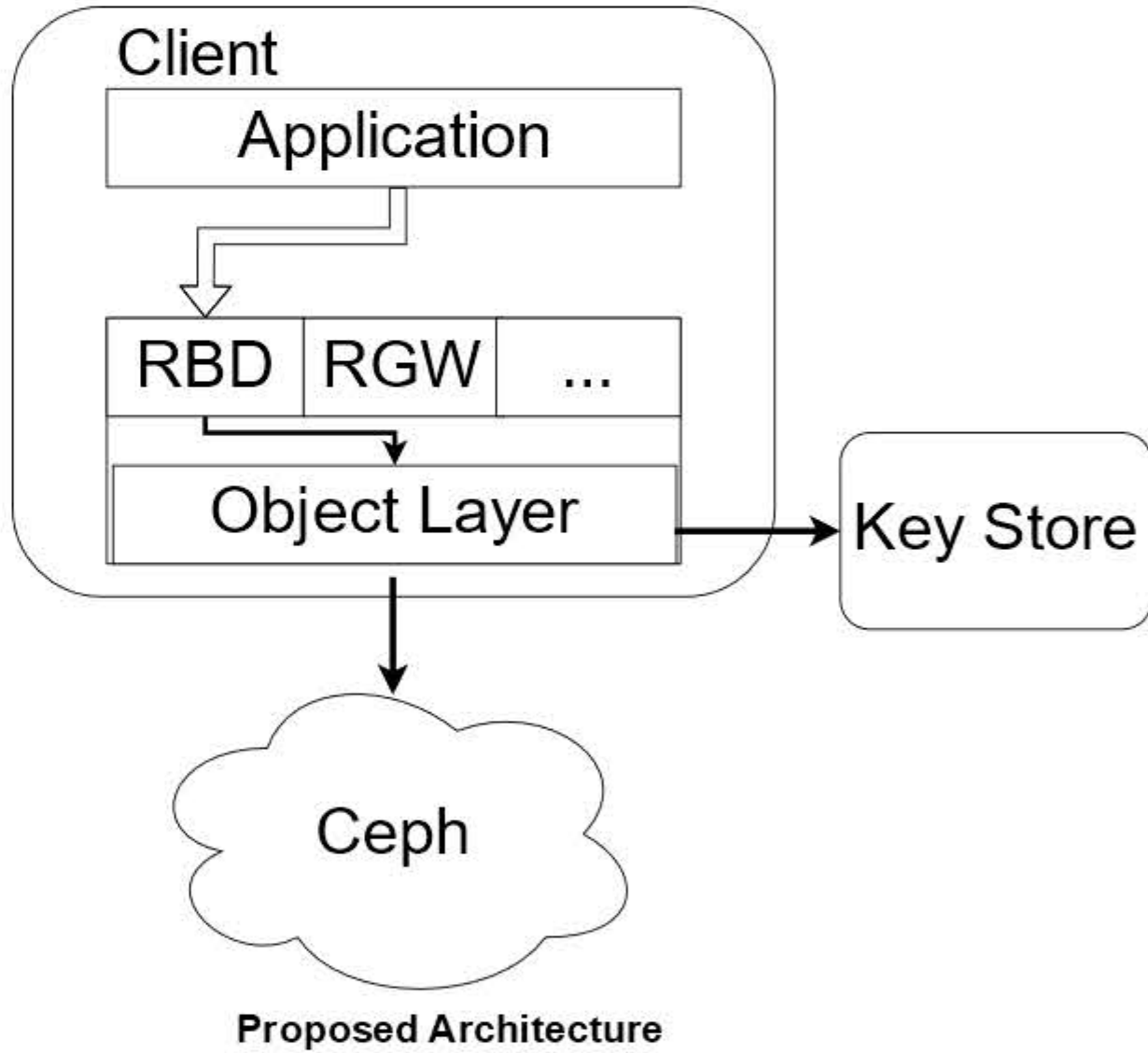


Attack Scenario

Storage System Client-Side Encryption Integrity Verification Integrity Verification Block Store

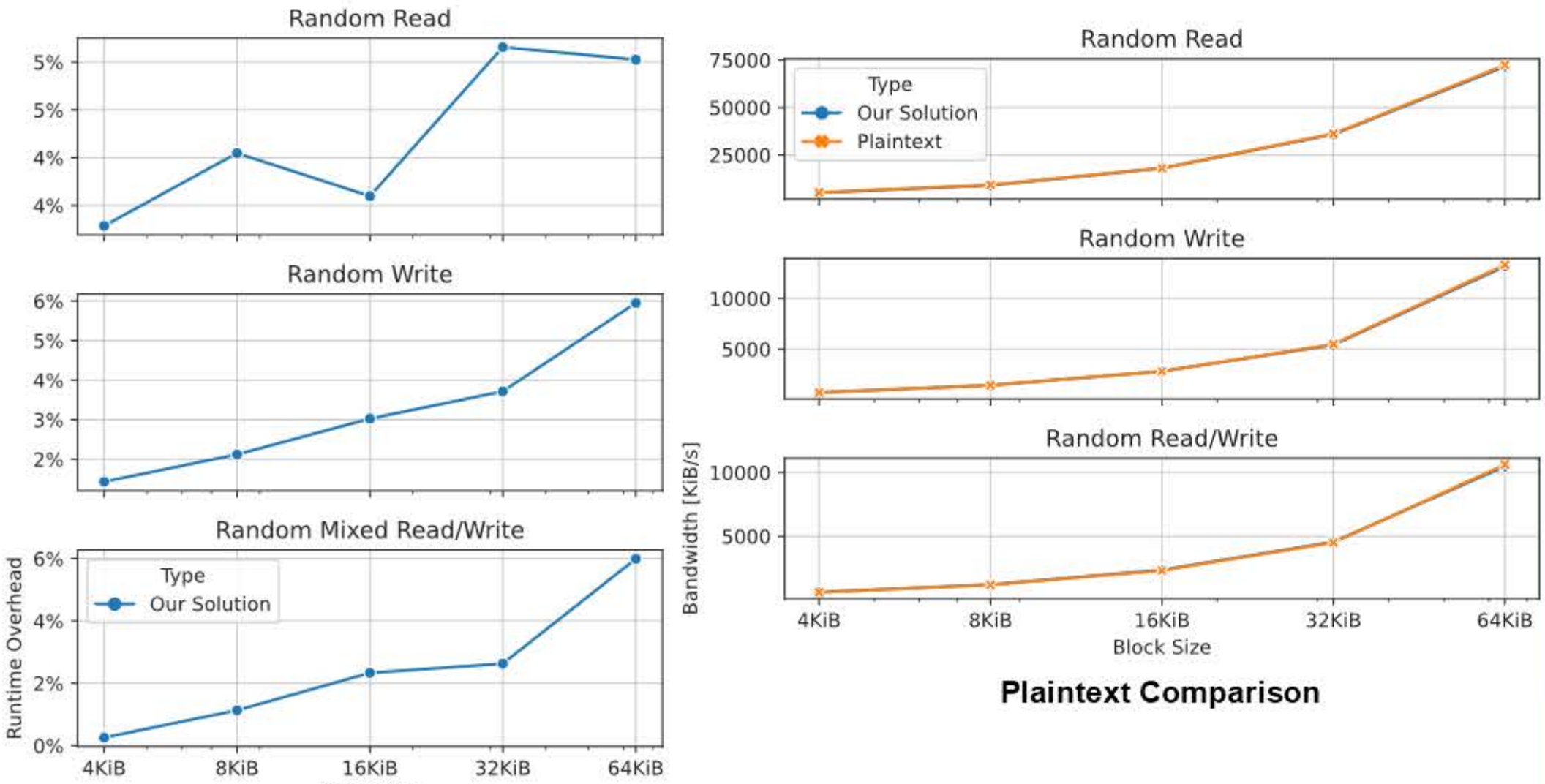
Ceph	Y	N	N
Hadoop	Y	N	N/A
OpenStack-Swift	Y	N	N/A
Lustre	Y	N	N/A
GPFS	Y	N	N/A
Azure	Y	Y	N
Google-Cloud	Y	Y	N
AWS	Y	Y	N

Proposed Method



- Client-Side encryption using AEAD ciphers
- Get key from external key store
- Store authentication tag (+nonce) in Ceph
- Applicable to other Ceph interfaces

Experimental Analysis



Results:

- Random IO lower overhead
  - Most common use case for BS
- Better for small block sizes (smaller 64KiB)
  - Common block size for Linux filesystems

Benchmarks Configuration:

- 1 GiB IO
- 16 Concurrent Requests
- Random & Sequential IO modes

Open Questions

- How applicable to other Software Defined Storage Systems?
- How to much data should be encrypted under one key?
- How to enable a distributed key management?
  - How to enable a proper key life cycle?
- Should 4 KiB block size be mandatory?
  - How to make it configurable?





# A Communication-efficient Conjunctive Query Scheme under Local Differential Privacy

Ellen Z. Zhang, Yunguo Guan, Rongxing Lu, Harry Zhang

Contact Email: rlu1@unb.ca

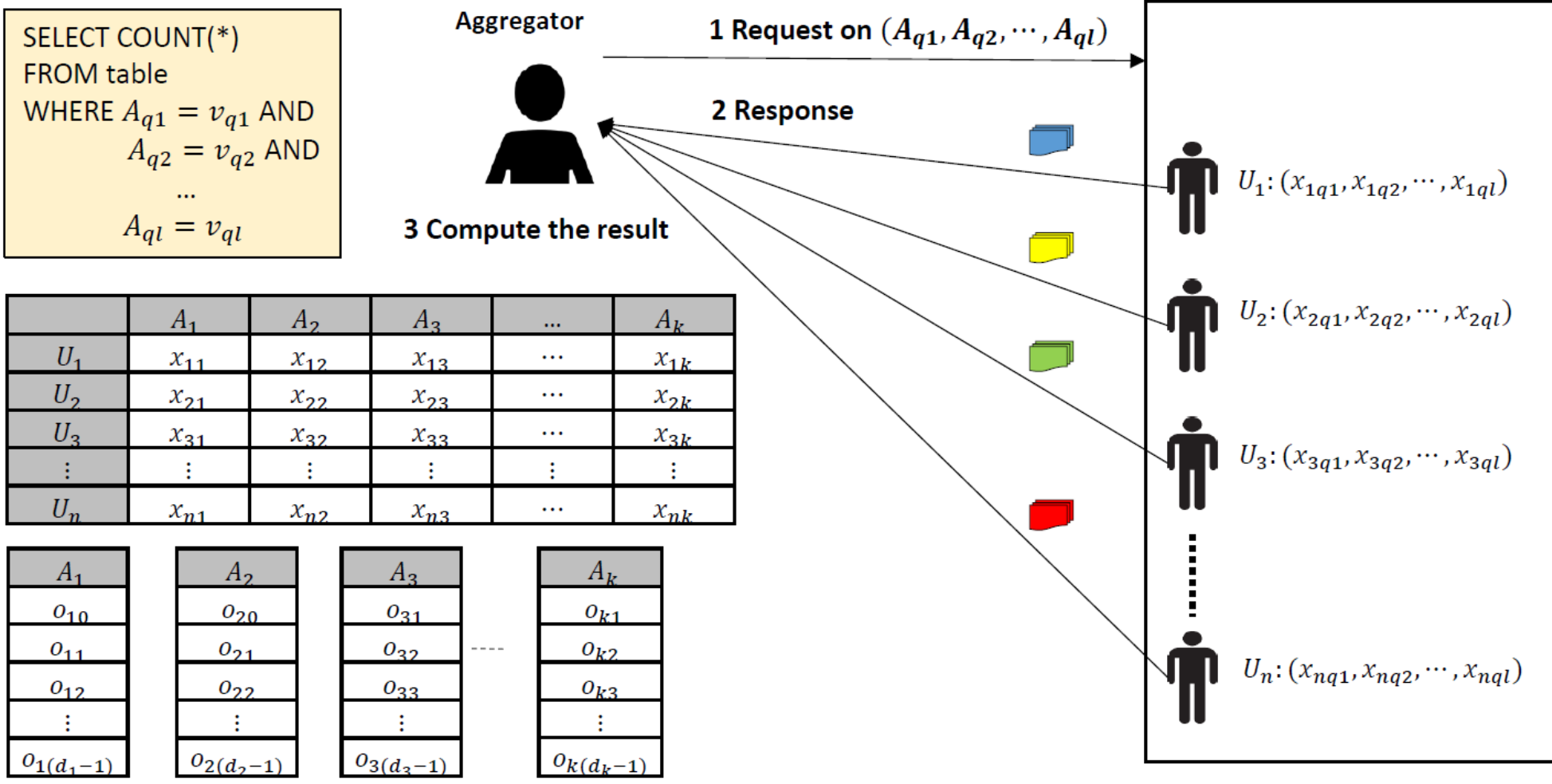
Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)



## ABSTRACT

Crowdsourcing has become a widely used method for data collection and analysis, yet its privacy remains a challenge. In this work, we present a new efficient and privacy-preserving conjunctive query scheme for crowdsourcing scenarios. The scheme employs the Local Differential Privacy (LDP) technique to ensure both query privacy and high communication efficiency. Specifically, when an aggregator launches a conjunctive query to a set of crowdsourcing users, the query condition will not be leaked. To respond the query, each user just needs to return one bit back to the aggregator. By integrating prefix encoding technique, our proposed scheme can also efficiently support conjunctive queries with one range query condition. Detailed security analysis shows our proposed scheme can achieve desirable security requirements. In addition, performance evaluations also indicate its efficiency. Furthermore, extensive experiments demonstrate our proposed scheme can achieve high accuracy while ensuring  $\epsilon$ -LDP.

## System Model



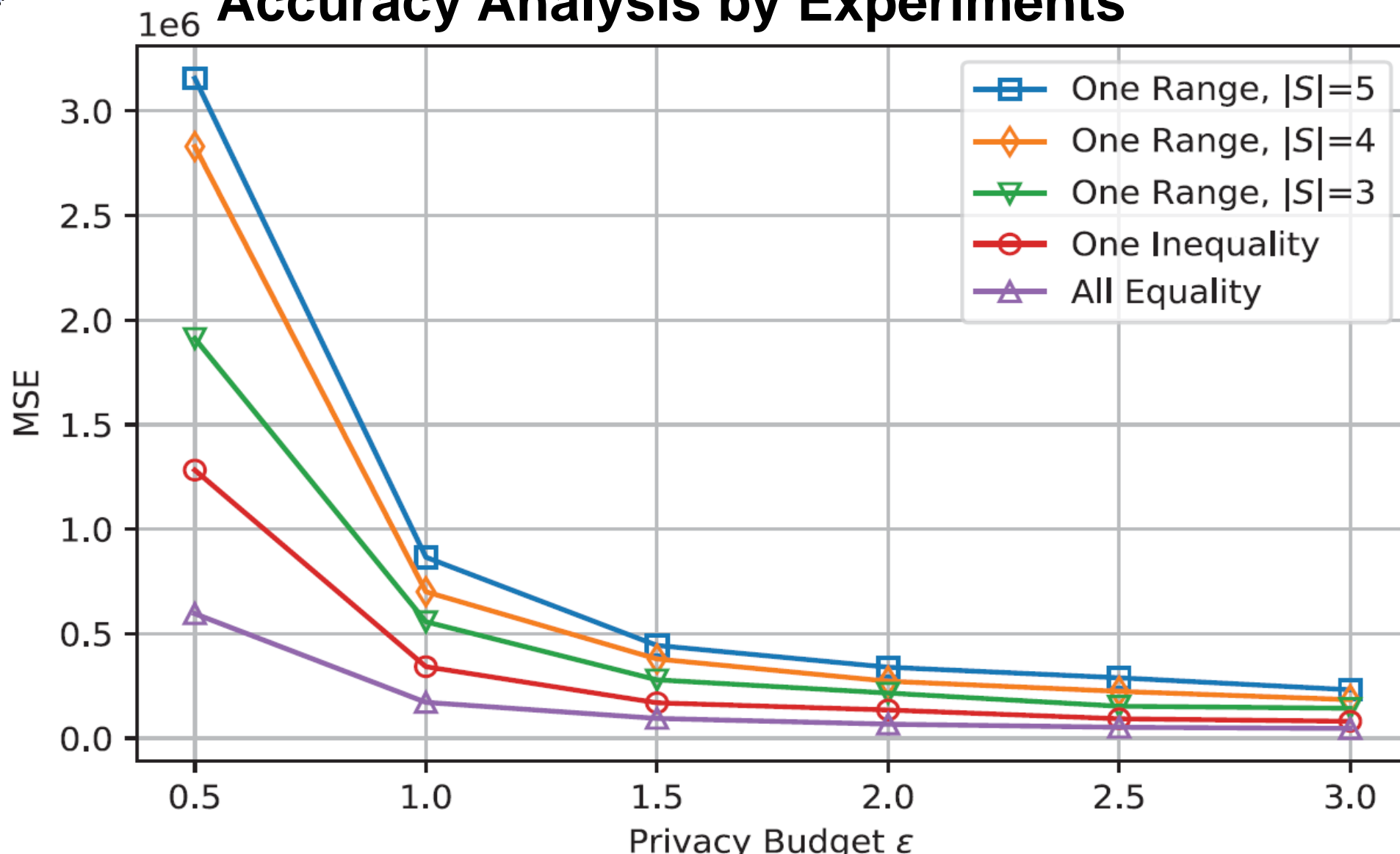
## System Model & Design Goal

- Security Model: All users are honest, Aggregator is honest-but-curious
- Design Goal: privacy and accuracy, communication efficiency

## Main Idea of Our Proposed Scheme

For each categorical attribute  $A_i \in \mathcal{A}$ , its categories  $\{o_{i0}, o_{i1}, \dots, o_{i(d_i-1)}\}$  are expressed as indices  $\{0, \dots, d_i - 1\}$ . Then, for a specific conjunctive query on  $\{A_{q1}, A_{q2}, \dots, A_{ql}\}$ , the query conditions  $(o_{q1}, o_{q2}, \dots, o_{ql})$  can be one-to-one mapped into a string  $v = o_{q1} || o_{q2} || \dots || o_{ql}$ . This string has a bit length  $\sum_{i=1}^l \lceil \log_2 d_{qi} \rceil$  and resides within a string space  $\mathcal{V}$  of size  $|\mathcal{V}| = d = \prod_{i=1}^l d_{qi}$ . For each user  $U_j \in \mathcal{U}$ , if  $U_j$ 's response value  $X_{jq} = \{x_{jq1}, x_{jq2}, \dots, x_{jqk}\}$  for the query  $\{A_{q1}, A_{q2}, \dots, A_{ql}\}$  is mapped into a string  $v_j \in \mathcal{V}$ , and  $v_j = v$ , then  $U_j$ 's response will be counted in the conjunctive query result  $c(v)$ , i.e.,  $c(v) = \sum_{j=1}^n \mathbf{1}_{\{v_j=v\}}$ . To hide the value  $v_j$  from the aggregator,  $U_j$  can employ the random response technique to perturb  $v_j$  before reporting it.

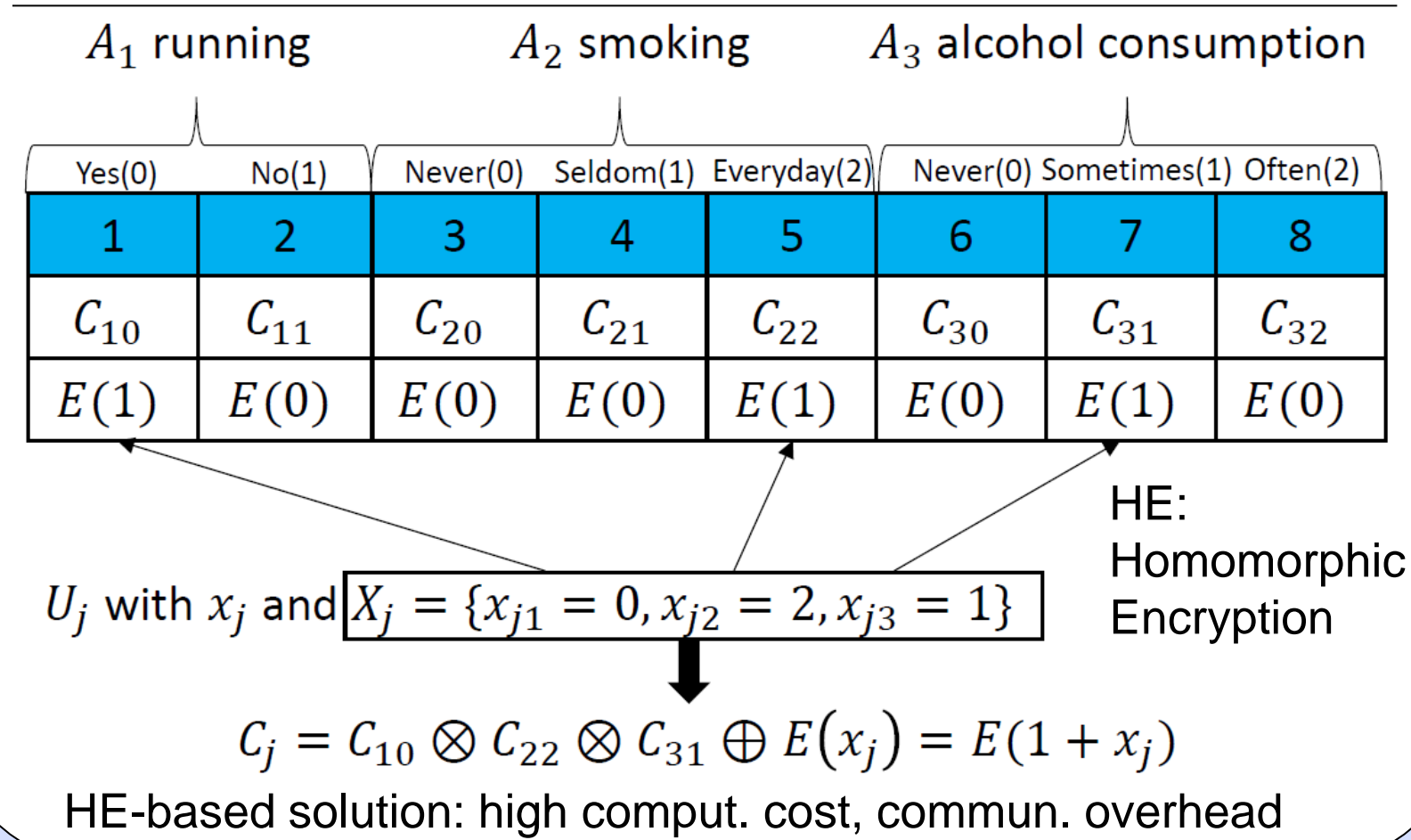
## Accuracy Analysis by Experiments



MSE versus varying privacy budget  $\epsilon$  with  $n=38,531$  users. In All Equality, the conjunctive query includes 25 equality conditions, one per attribute, and AG counts users that satisfy these conditions. In One Inequality, the query has 24 equality and one inequality conditions, requiring AG to process two queries and compute the difference between their results. In One Range, the query includes 24 equality conditions and a range condition. This range condition is evaluated in three cases, with  $|S|=3$ ,  $|S|=4$ , and  $|S|=5$ , respectively.

## A Simple Example & HE-based Solution

Consider three categorical attributes  $(A_1, A_2, A_3)$ :  $A_1$  (running) with values  $\{\text{yes}(0), \text{no}(1)\}$ ,  $A_2$  (smoking) with values  $\{\text{never}(0), \text{seldom}(1), \text{everyday}(2)\}$ , and  $A_3$  (alcohol consumption) with values  $\{\text{never}(0), \text{sometimes}(1), \text{often}(2)\}$ . Let  $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$  be a set of users, where each user  $U_j \in \mathcal{U}$  holds personal data  $X_j = \{x_{j1}, x_{j2}, x_{j3}\}$  regarding these three attributes. For example, if  $U_j$  likes running, smokes everyday, and drinks alcohol sometimes, the value of the personal data is  $X_j = \{x_{j1} = 0, x_{j2} = 2, x_{j3} = 1\}$ . Now, consider an aggregator who wants to determine how many users like running, smoke everyday and drink alcohol sometimes; the aggregator will send the query to all users.



## Local Differential Privacy

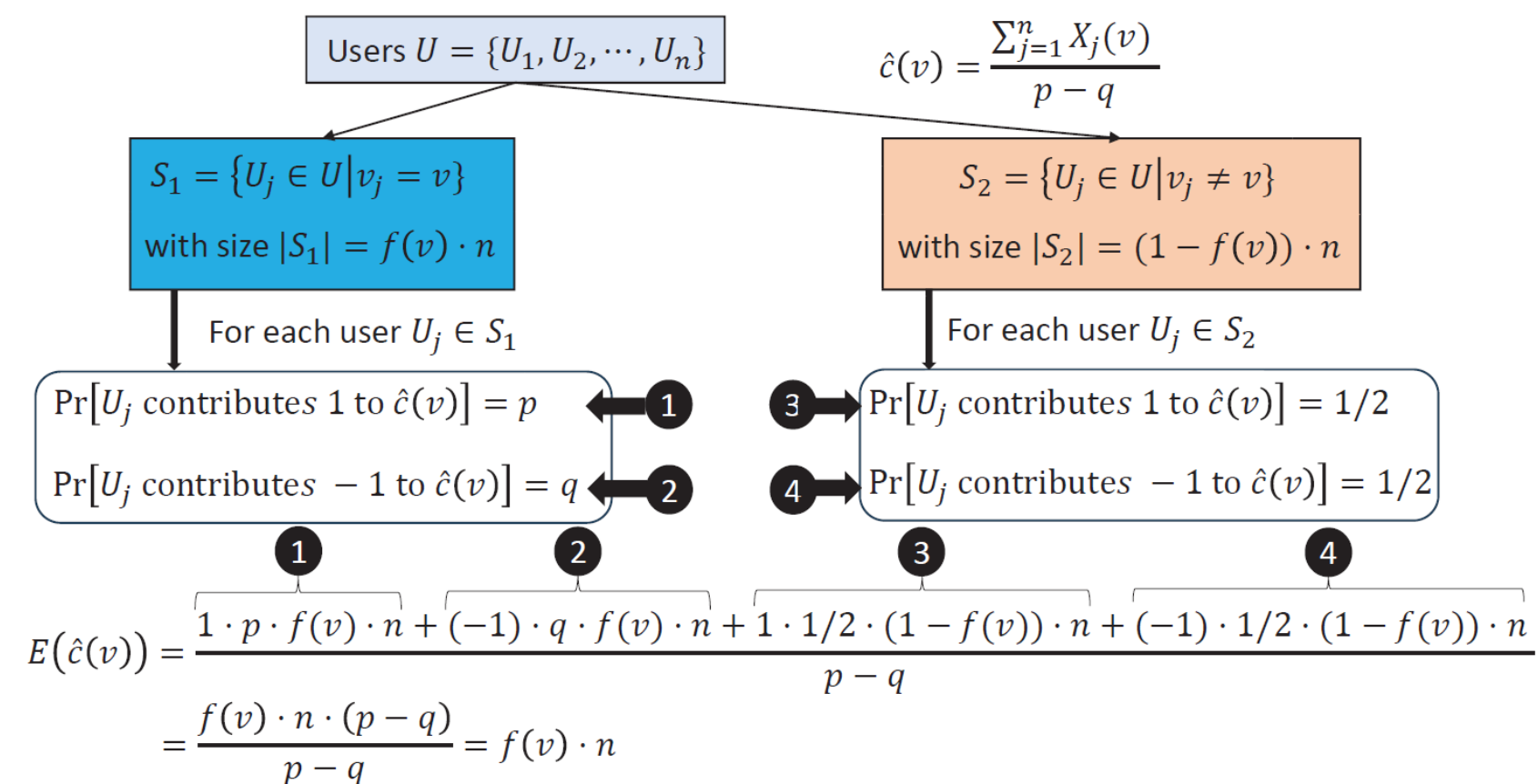
**Definition 1. (Local Differential Privacy)** A randomized algorithm  $\mathcal{RA} : \mathcal{I} \rightarrow \mathcal{O}$  satisfies  $\epsilon$ -Local Differential Privacy (or  $\epsilon$ -LDP), where  $\epsilon \geq 0$ , if and only if for any two inputs  $x, x' \in \mathcal{I}$  and for any output  $y \in \mathcal{O}$ ,

$$\Pr[\mathcal{RA}(x) = y] \leq e^\epsilon \cdot \Pr[\mathcal{RA}(x') = y].$$

## Correctness & Accuracy

**Theorem 1.** For any query string  $v$  in the string space  $\mathcal{V}$ , let  $f(v)$  be the fraction of occurrences when  $v_j = v$ , for  $j = 1, 2, \dots, n$ ,  $\hat{c}(v) = \frac{\sum_{j=1}^n X_j(v)}{p-q}$  is an unbiased estimator of  $c(v) = \sum_{j=1}^n \mathbf{1}_{\{v_j=v\}}$ , i.e.,  $E(\hat{c}(v)) = f(v) \cdot n$ .

Let  $f(v)$  be the frequency for one value  $v$  in the domain, divide users into two sets



**Theorem 2.** The variance of the estimator  $\hat{c}(v)$  is bounded by  $\text{Var}(\cdot) = n \cdot \left( \frac{e^\epsilon + 1}{e^\epsilon - 1} \right)^2$  in our proposed scheme.

## Flexibility on Conjunctive Queries

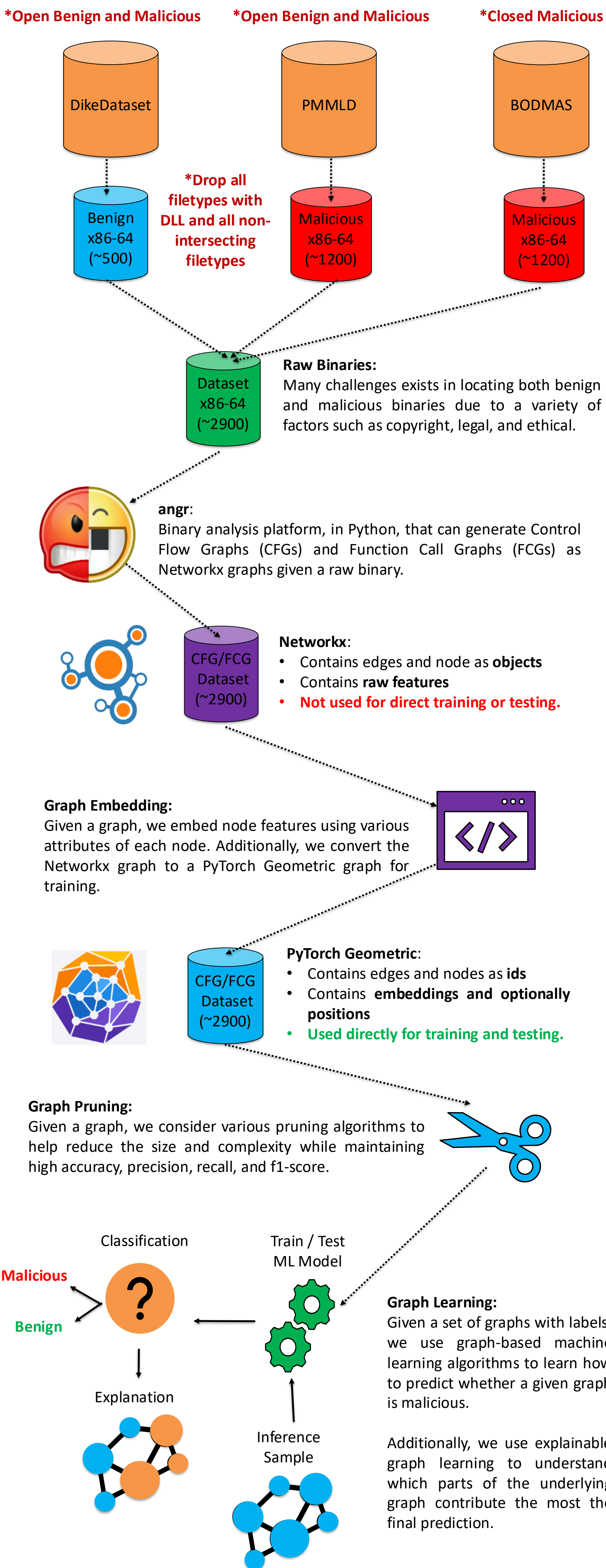
- All Equality - variance  $\text{Var}(\cdot)$ :  $U_j$  reports 1 bit
- One Inequality - variance  $2 \cdot \text{Var}(\cdot)$ :  $U_j$  reports 2 bits
- One Range - variance  $O(\log d_{ql}) \cdot \text{Var}(\cdot)$ :  $U_j$  reports  $O(\log d_{ql})$  bits



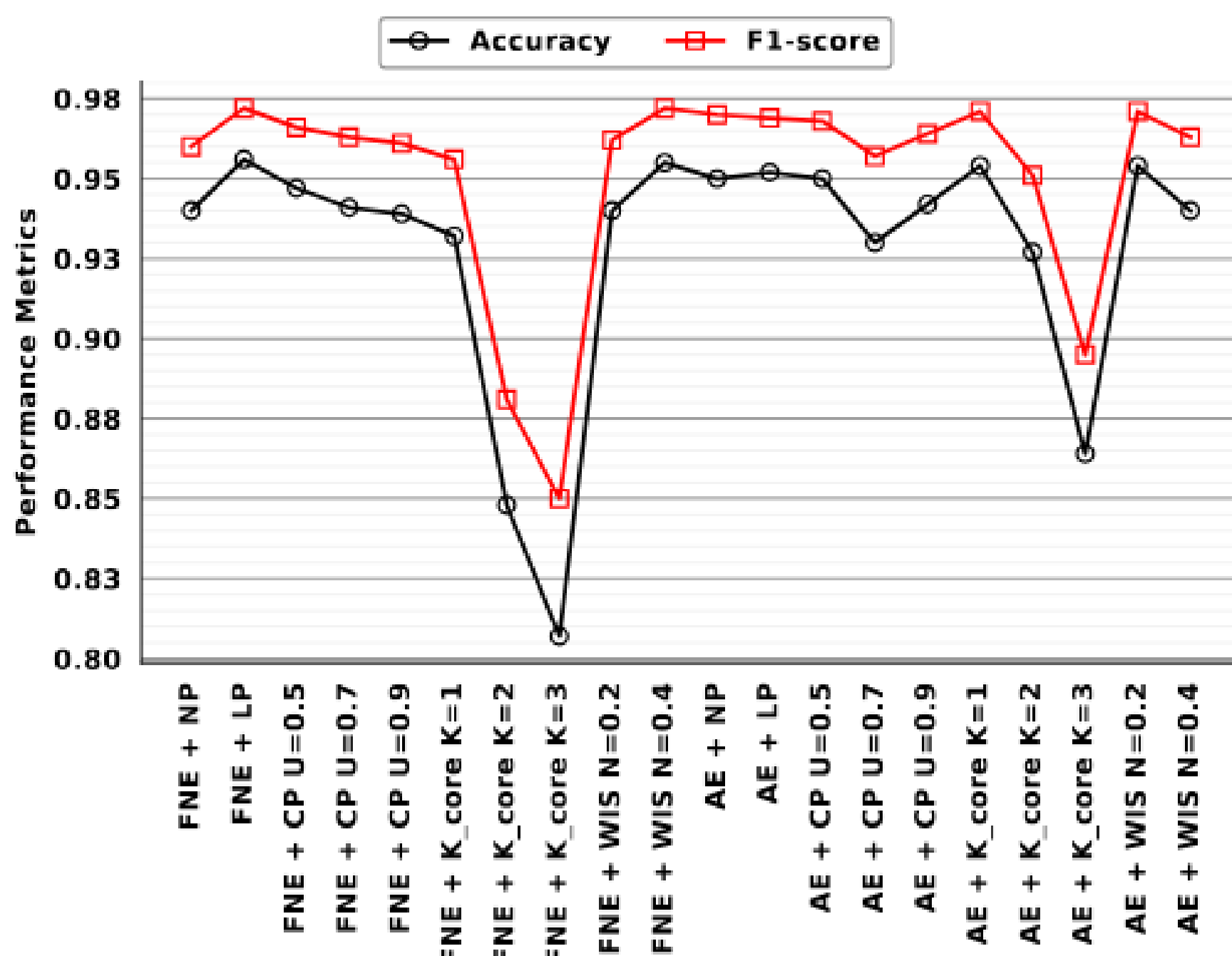
## ABSTRACT

Control Flow Graphs (CFGs) and Function Call Graphs (FCGs) have become pivotal in providing a detailed understanding of program execution and effectively characterizing the behavior of malware. These graph-based representations, when combined with Graph Neural Networks (GNNs), have shown promise in developing high-performance malware detectors. However, challenges remain due to the large size of these graphs and the inherent opacity in the decision-making process of GNNs. Our work addresses these issues by evaluating several graph pruning techniques to reduce graph size. Our analysis demonstrates that the Leaf Prune technique not only significantly reduces graph size but also maintains superior performance, offering a balanced approach to improving both efficiency and transparency in malware detection.

## Detection System Pipeline



## Model Metrics



In our work we use two embedding methods, Function Node Embedding (FNE), for Function Call Graphs (FCGs), as well as Assembly Embedding for Control Flow Graphs (CFGs). These methods help capture information about nodes and embed them so they can be use by graph-based machine learning algorithms. Additionally, we propose several pruning methods to prune graphs before they are used for training. In the above figure we can see some dramatic differences in the various pruning algorithms and embedding methods for the given model evaluation metrics.

## Pruning Metrics

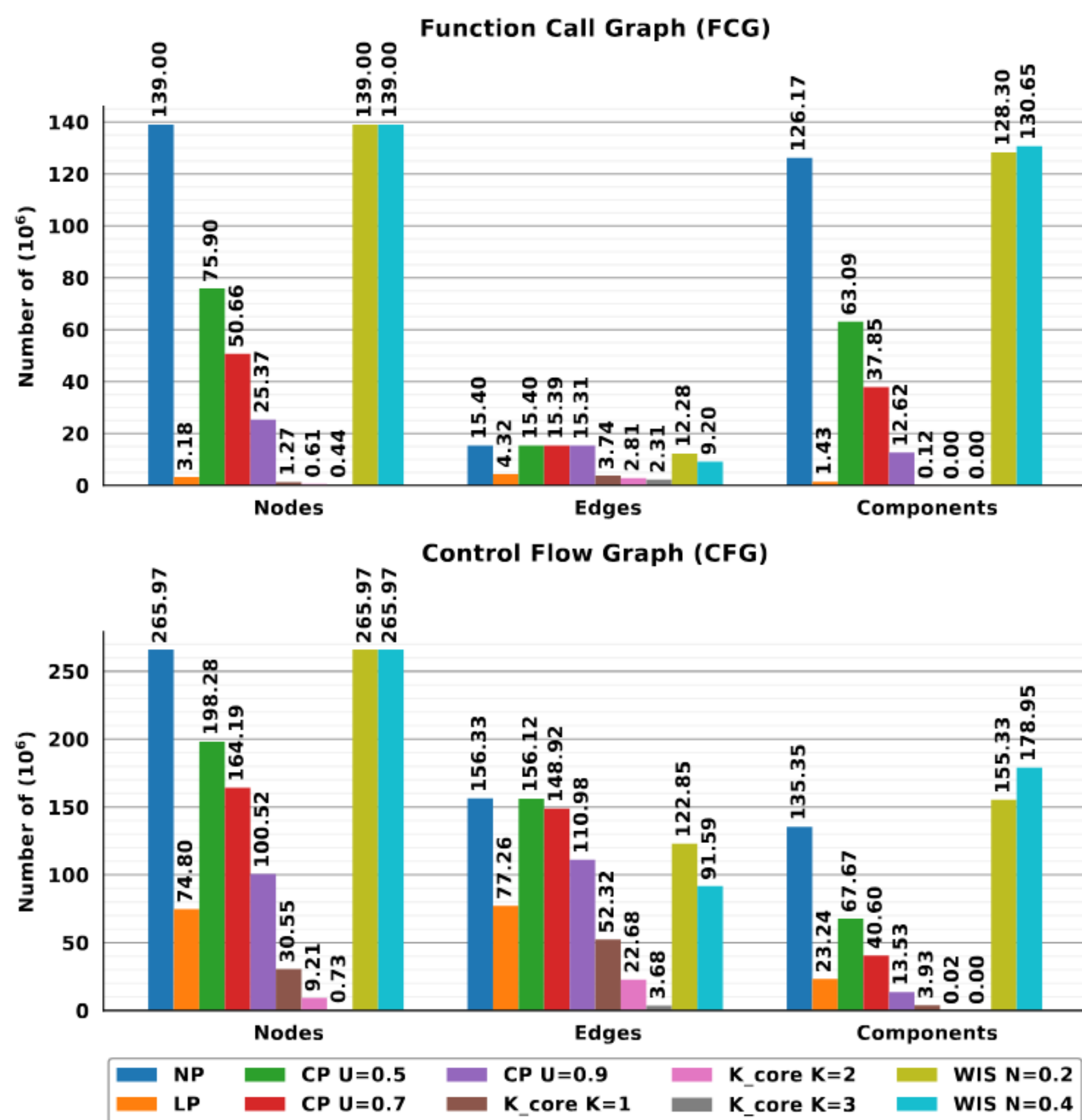


Fig. 4: Different graph pruning methods comparison

Since some graphs in our dataset are very large in terms of number of nodes, edges, and components it is important to understand how a given pruning algorithm alters the size and structure of the dataset as a whole. Here, we propose several pruning algorithms with both simple and complex methods to help reduce the size of the graphs in an intelligent way through a process known as graph sparsification.







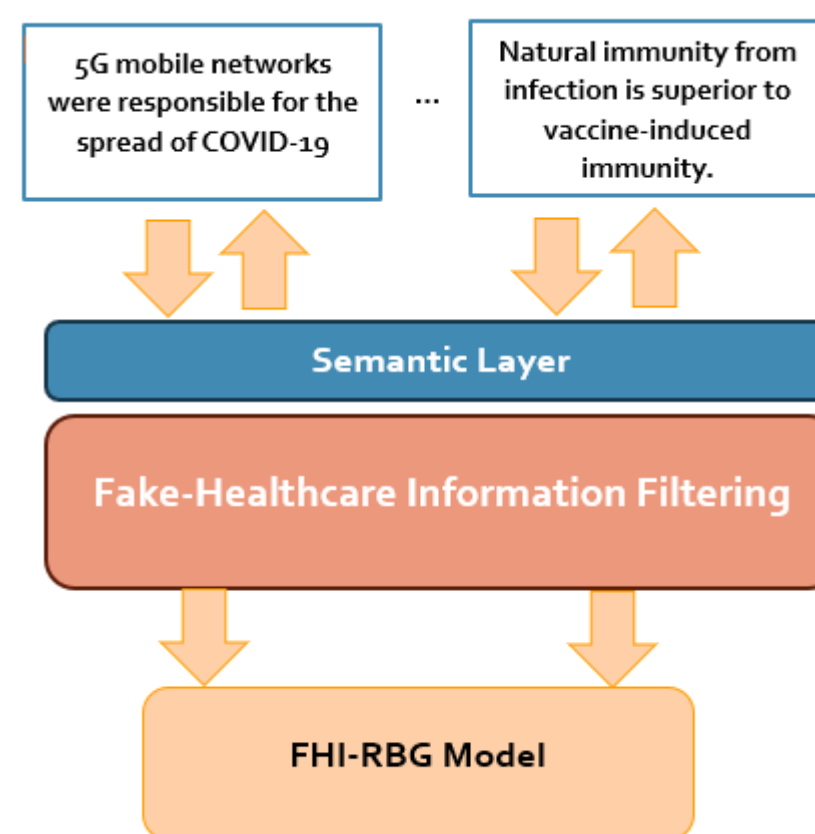
## ABSTRACT

In this study, we introduce a novel fine-tuning model combining RoBERTa and bi-GRU architectures for the detection of fake healthcare information, that call FHI-RBG. We also evaluated various pre-trained models, including BERT and DistilBERT, to compare their effectiveness. Our results demonstrate that the RoBERTa-based model offers superior accuracy in identifying misinformation. The proposed approach operates in two distinct phases: In Phase 1, we apply a hybrid method that integrates filtering and semantic similarity analysis to distinguish relevant inputs from noise. Following tokenization, the processed data is fed into our fine-tuned embedding model. In the subsequent phase, we assess the performance of BioBERT and ClinicalBERT for analyzing medical entities associated with false healthcare claims. Furthermore, we incorporate prompt engineering with few-shot prompts to enhance the re-evaluation and extraction of misleading information from the text.

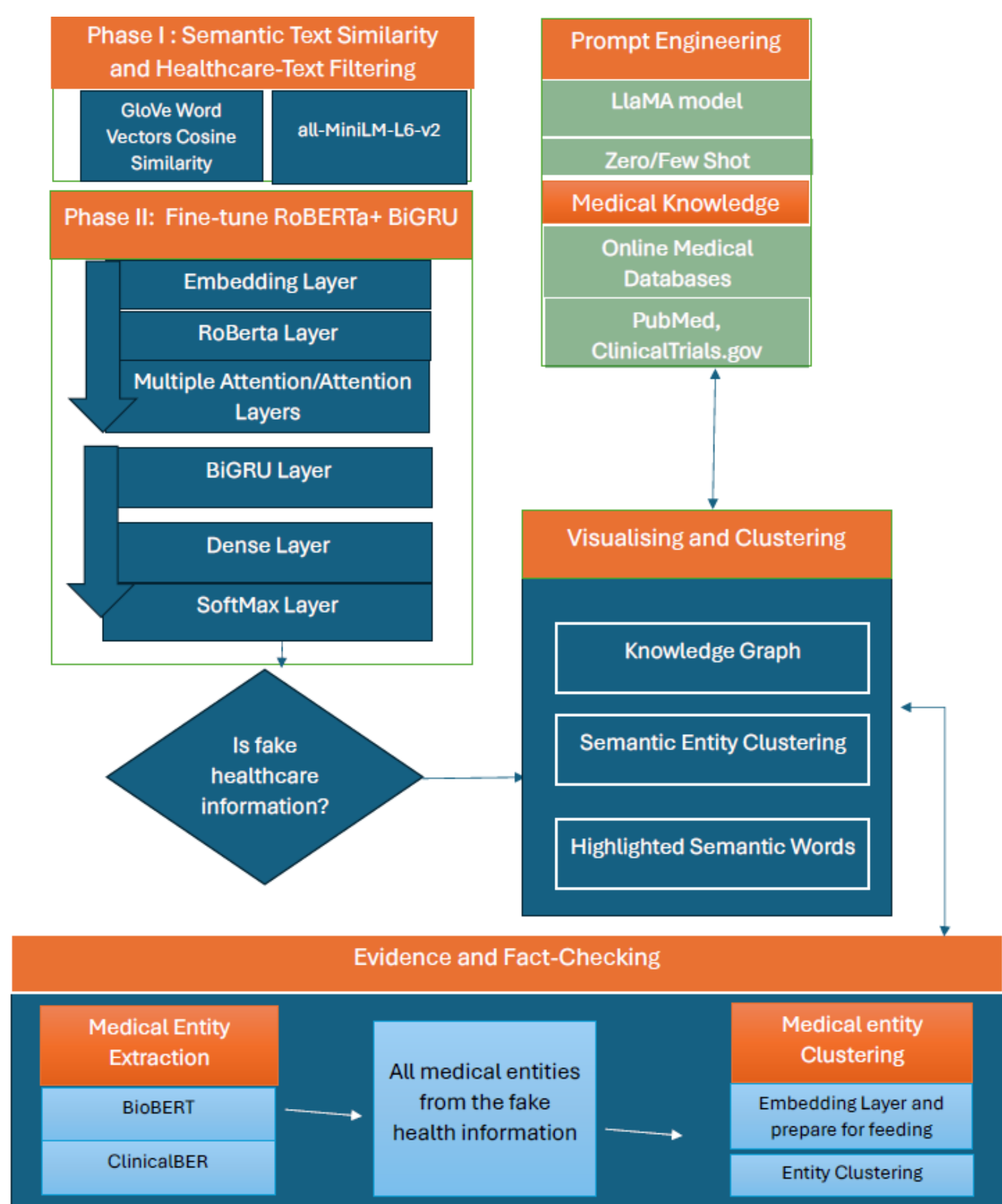
## The main contributions of the model

- Introduced an innovative model combining RoBERTa and bi-GRU to transform healthcare information into rich, dense embeddings, significantly enhancing the detection of fake content by capturing deeper semantic nuances.
- Utilized a structured approach involving initial semantic similarity analysis to filter relevant information and subsequent detailed entity analysis using BioBERT and ClinicalBERT, improving accuracy in identifying misleading healthcare claims
- Proposed a systematic fact-checking process incorporating advanced prompt engineering techniques with few-shot prompting, refining the model's ability to detect and extract false information.

## Semantic Filtering Mechanism



## Proposed Model



## Highlighted Model Phases

- ✓ **Phase I** : This phase employs a hybrid approach combining semantic similarity analysis and advanced filtering techniques. By leveraging the contextual embeddings from RoBERTa and bi-GRU, it effectively identifies relevant information and minimizes noise. This approach ensures that the data is filtered and processed to highlight meaningful content while discarding irrelevant or misleading information.
- **Phase II** : In the second phase, the model uses BioBERT and ClinicalBERT for in-depth analysis of medical entities, focusing on identifying and validating healthcare-related information. Additionally, this phase incorporates a systematic fact-checking procedure using advanced prompt engineering techniques with few-shot prompting.

## Contributions of Fact-Checking

- Investigated and extracted medical entities to determine which entities are most prominently featured in detected fake healthcare information, refining the model's focus on key misleading elements.
- Applied semantic clustering using Agglomerative Clustering to group similar healthcare information based on semantic content, which aids in identifying patterns and relationships within the data and improves the model's overall accuracy.

## Experimental Analysis

The below three figures show the performance of the fine-tune model based on different metrics for Bert, RoBERTa, DistBert models

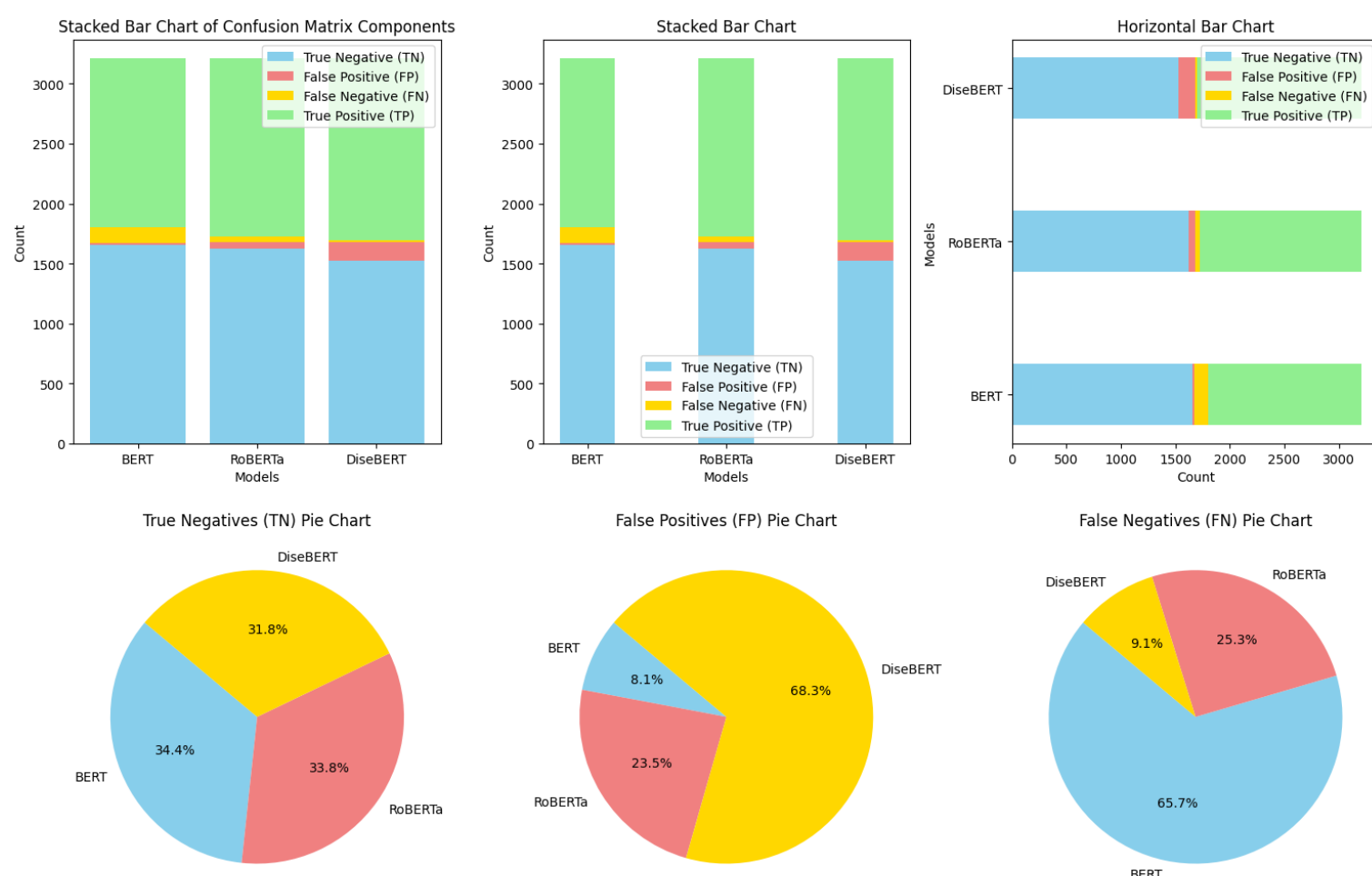


Fig (A)

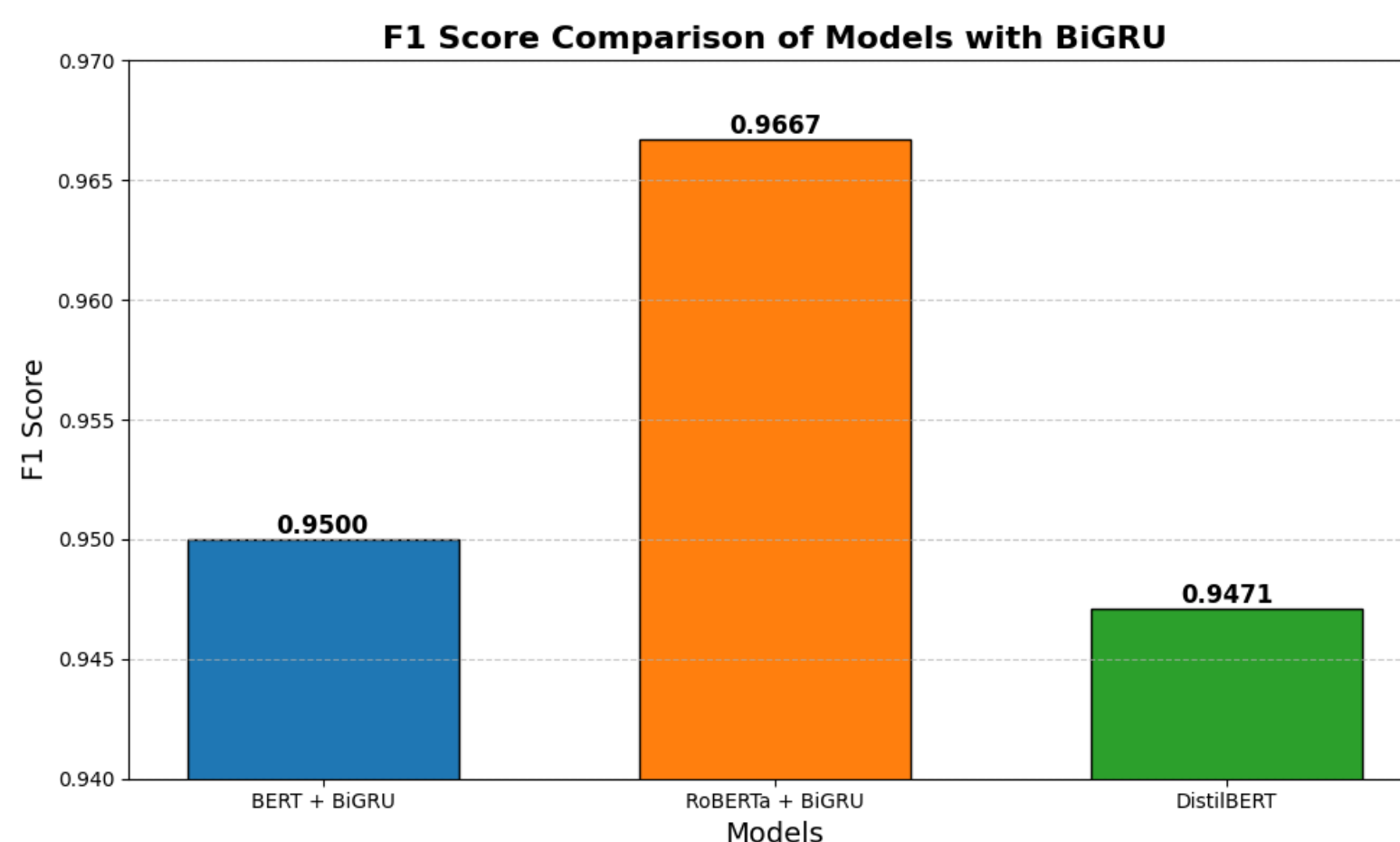


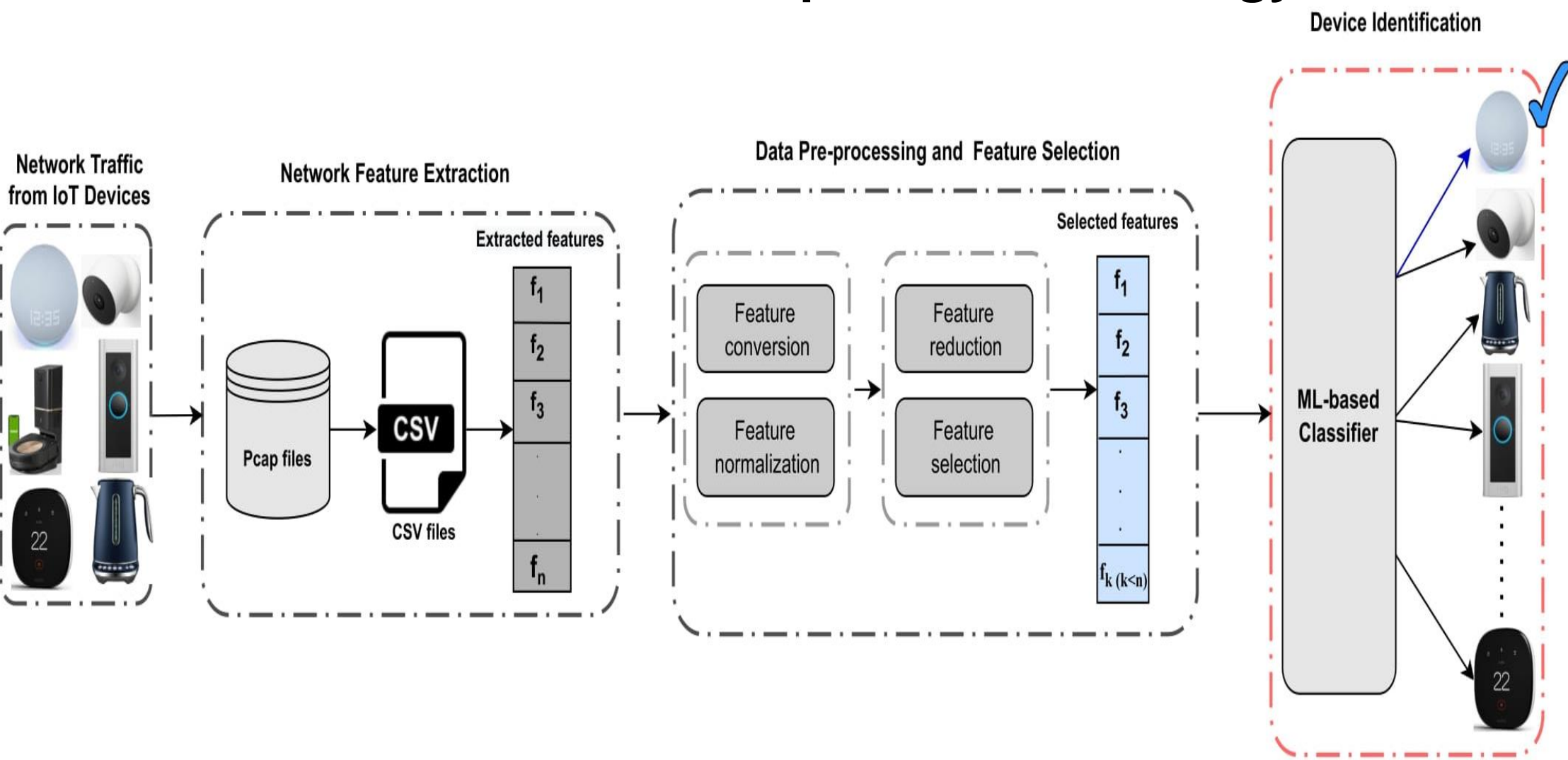
Fig (B)



## ABSTRACT

As the Internet of Things (IoT) landscape expands, new devices with various functionalities are continuously being integrated into the IoT ecosystem. When traditional systems, which involve human interaction, are replaced by devices, it becomes crucial to upgrade the conventional authorization and authentication systems. Traditional device identification approaches often struggle to accommodate the dynamic behaviors exhibited by IoT devices. In response, this work introduces an innovative approach that leverages enhanced behavioral features to generate a representation of device behavior. This representation is then employed to train machine learning models for classifying devices based on their behaviors. Furthermore, this work also considers special scenarios where the access management system lacks access to full network traffic data. In such cases, device identification is achieved based on HTTPS features and user agent information. We conducted experimental analyses using real data from state-of-the-art IoT device profiling datasets. The performance results indicate that our behavioral-based features have the capability to identify multiple IoT devices with various functionalities and vendors.

## Workflow of the Proposed Methodology



## Main Contributions

- ❖ We applied both packet-per-packet and flow-based analysis to capture a wide range of features from different perspectives, improving the accuracy and depth of analysis.
- ❖ Experiments were conducted on the CICIOT2022 (39 devices) and Aalto (31 devices) datasets, combining them to create a larger dataset of 70 devices. Feature selection and machine learning techniques were used to optimize training time and accuracy for IoT device identification.
- ❖ The device identification model was adapted for scenarios with limited payload access, such as encrypted traffic (HTTPS). Features like packet headers, JA3, handshake data, and User-Agent strings were extracted to maintain identification accuracy in these contexts.

## New Extracted Behavioral-Based Features

No	Feature Name	No	Feature Name	No	Feature Name	No	Feature Name
1	device mac (not included in training)	15	src port	29	http content len	43	sum p
2	timestamp (not included in training)	16	dst port	30	http response code	44	min p
3	epoch timestamp (not included in training)	17	port class dst	31	icmp data size	45	max p
4	src mac (not included in training)	18	tcp window size	32	icmp type	46	med p
5	dst mac (not included in training)	19	highest layer	33	icmp checksum status	47	average p
6	src ip	20	dns server	34	payload entropy	48	var p
7	dst ip	21	tls server	35	dns interval	49	q3 p
8	eth src oui	22	dns len ans	36	ntp interval	50	q1 p
9	eth dst oui	23	dns query type	37	most freq spot	51	iqr p
10	L4 tcp	24	dns len qry	38	min et	52	13 ip dst count
11	ttl	25	http content type	39	q1		
12	eth size	26	http request method	40	min e		
13	ip size	27	http host	41	var e		
14	payload length	28	http uri	58	q1 e		

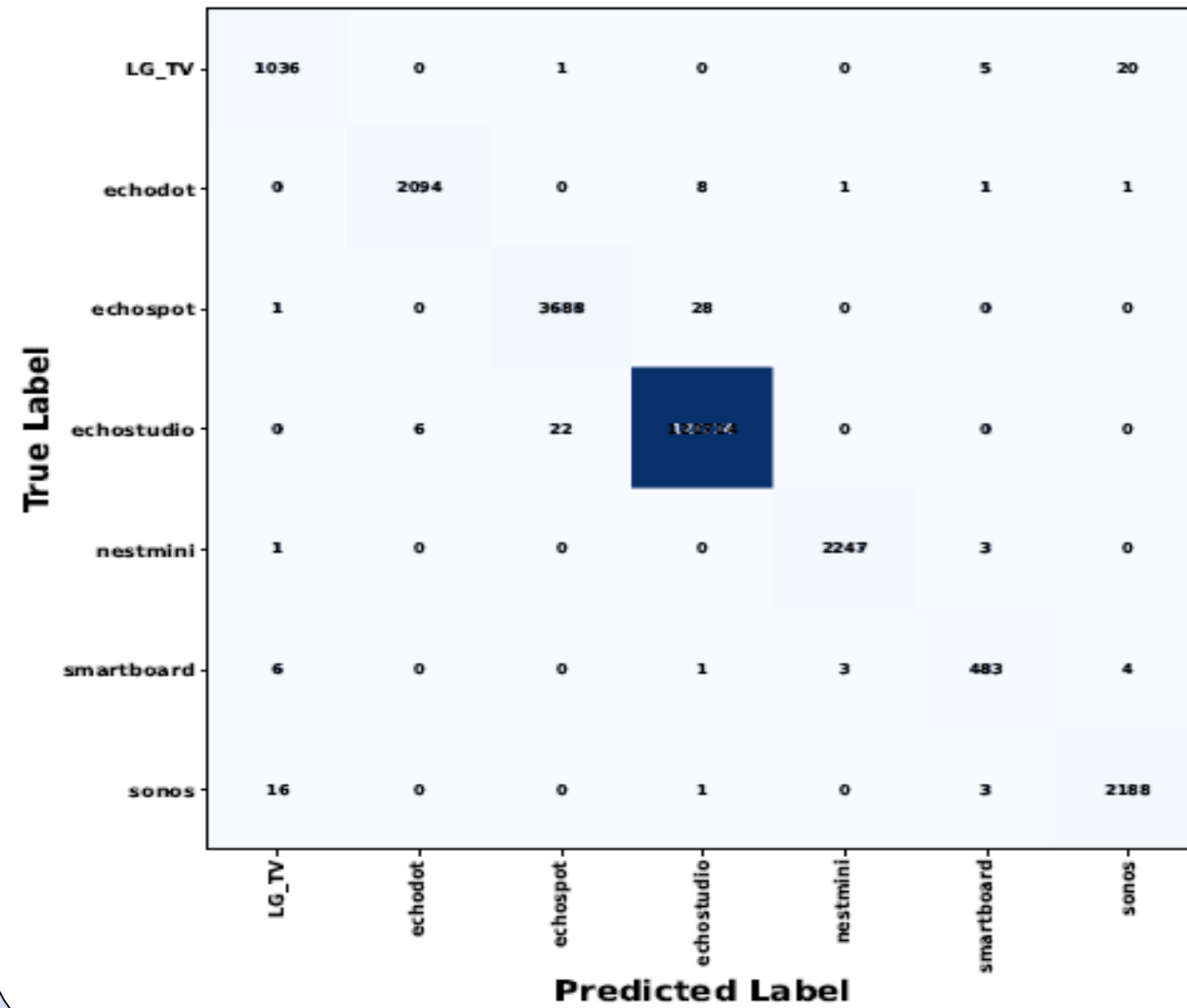
## New Extracted Features Exclusively from HTTPS Traffic and User-Agent Strings

No	Feature Name	No	Feature Name	No	Feature Name	No	Feature Name
1	JA3	14	handshake_version	27	q1	40	max_p
2	stream	15	handshake_cipher_suites_length	28	iqr	41	med_p
3	inter_arrival_time	16	handshake_ciphersuites	29	sum_e	42	average_p
4	time_since_previously_displayed_frame	17	handshake_extensions_length	30	min_e	43	var_p
5	14_tcp	18	handshake_sig_hash_alg_len	31	max_e	44	q3_p
6	14_udp	19	payload_entropy	32	med_e	45	q1_p
7	17_http	20	sum_et	33	average_e	46	iqr_p
8	17_https	21	min_et	34	var_e	47	user_agent_Browser
9	ttl	22	max_et	35	q3_e	48	user_agent_OS
10	eth_size	23	med_et	36	q1_e	49	user_agent_Device
11	ip_size	24	average_et	37	iqr_e		
12	payload_length	25	var_et	38	sum_p		
13	tcp_window_size	26	q3	39	min_p		

## Experimental Analysis

Classifier Technique	No of devices	Train Time (s)	Test Time (s)	No of samples	No of features	FS technique	Accuracy	Precision (MA)	Recall (MA)	F1 Score (MA)
Decision Tree (DT)	70	17.9	0.16	797,648	47 (all)	-	0.9971	0.9702	0.9701	0.9701
	70	7.4	0.1	797,648	35	PCC	0.9970	0.9673	0.9672	0.9670
	70	6.3	0.1	797,648	18	ANOVA	0.9625	0.8520	0.8518	0.8503
Selected Classifier	70	3.8	0.08	797,648	18	RFE	0.9973	0.9711	0.9712	0.9706
	70	5.9	0.1	797,648	21	GA	0.9824	0.9079	0.9072	0.9064
	70	1.7	0.10	797,648	7	PCC+ANOVA+RFE+GA	0.9490	0.8288	0.8233	0.8215

## Confusion Matrix of the Selected Classifier for Devices with Only HTTPS Traffic Data



### User-Agent String for an Amazon Echo Dot

Mozilla/5.0 (X11; Linux x86\_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/52.0.2743.82 Safari/537.36

### User-Agent String for a Google Nest Mini

Mozilla/5.0 (X11; Linux aarch64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/68.0.424.188 Safari/537.36 CrKey/1.54.250118

## Conclusion

This work introduces an innovative solution to the challenges arising from the expanding Internet of Things (IoT) landscape. Our novel access control system, powered by machine learning techniques, utilizes enhanced behavioral features for device classification, addressing the need for upgraded authorization and authentication mechanisms. We demonstrate the effectiveness of our approach, even in scenarios where network traffic data is limited, by identifying devices only through HTTPS features and user agent information. Through experimental analyses using real-world IoT device data in Pcap level, we validate the effectiveness of our method in accurately identifying devices across diverse functionalities and vendors.



ABSTRACT

This study examines the spread of true and false news on X, using a large dataset of ~126,000 stories tweeted by ~3 million people. The analysis reveals that false news spreads significantly faster, farther, deeper, and more broadly than true news. Surprisingly, this difference cannot be explained by user characteristics like followers, verified status, or activity level. Instead, the emotional responses it evokes, such as surprise and disgust, seem to play a significant role in its propagation. These findings challenge common assumptions and highlight the need for further research into the human behavior driving the spread of misinformation.

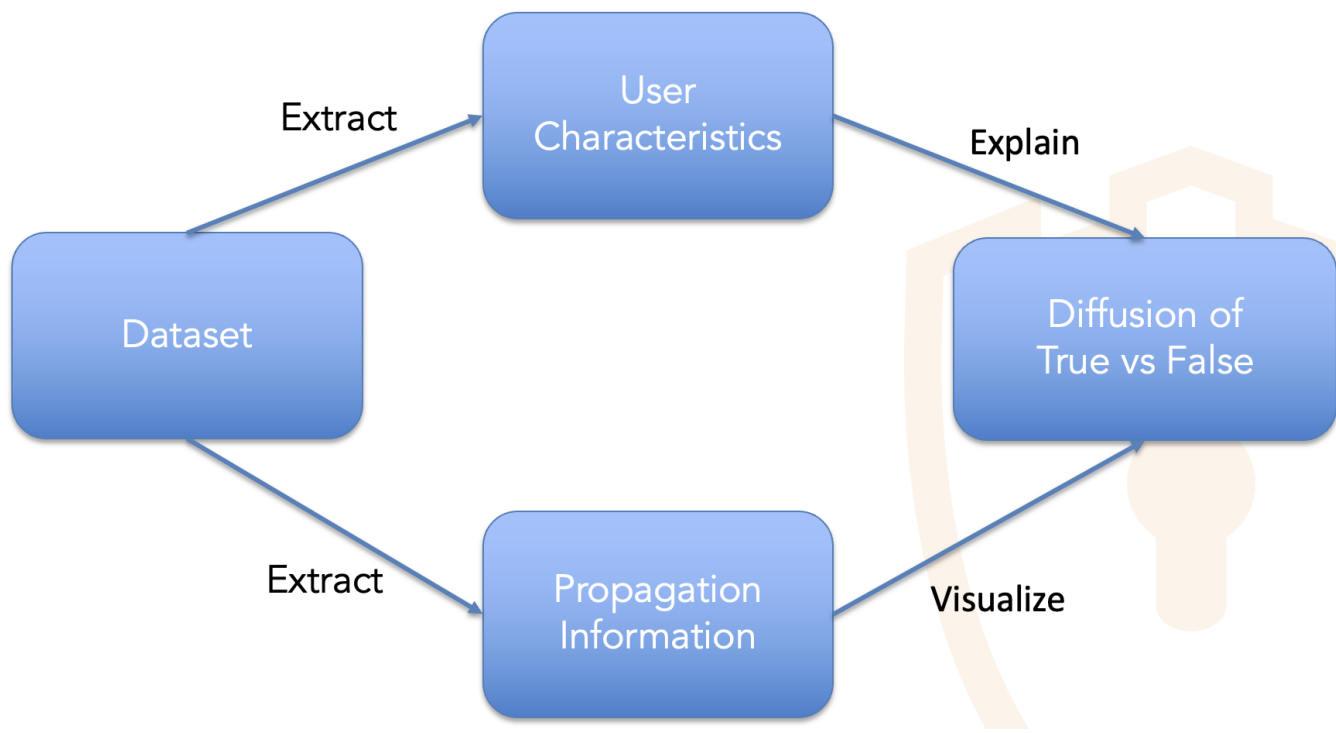
Research Methodology

RQ1: How do truth and falsity diffuse differently?

- Size: Total number of users exposed
- Depth: Levels of retweets from the original tweet
- Breadth: Maximum users exposed at any retweet level
- Structure Virality: Spread pattern capturing dispersion and depth.

RQ2: What Network/User characteristics can explain this differential diffusion?

- User Account Characteristics
- Emotion Response



Visualizing Rumor Propagation Graphs on X

- Delayed explosion
- Echo-chamber effect

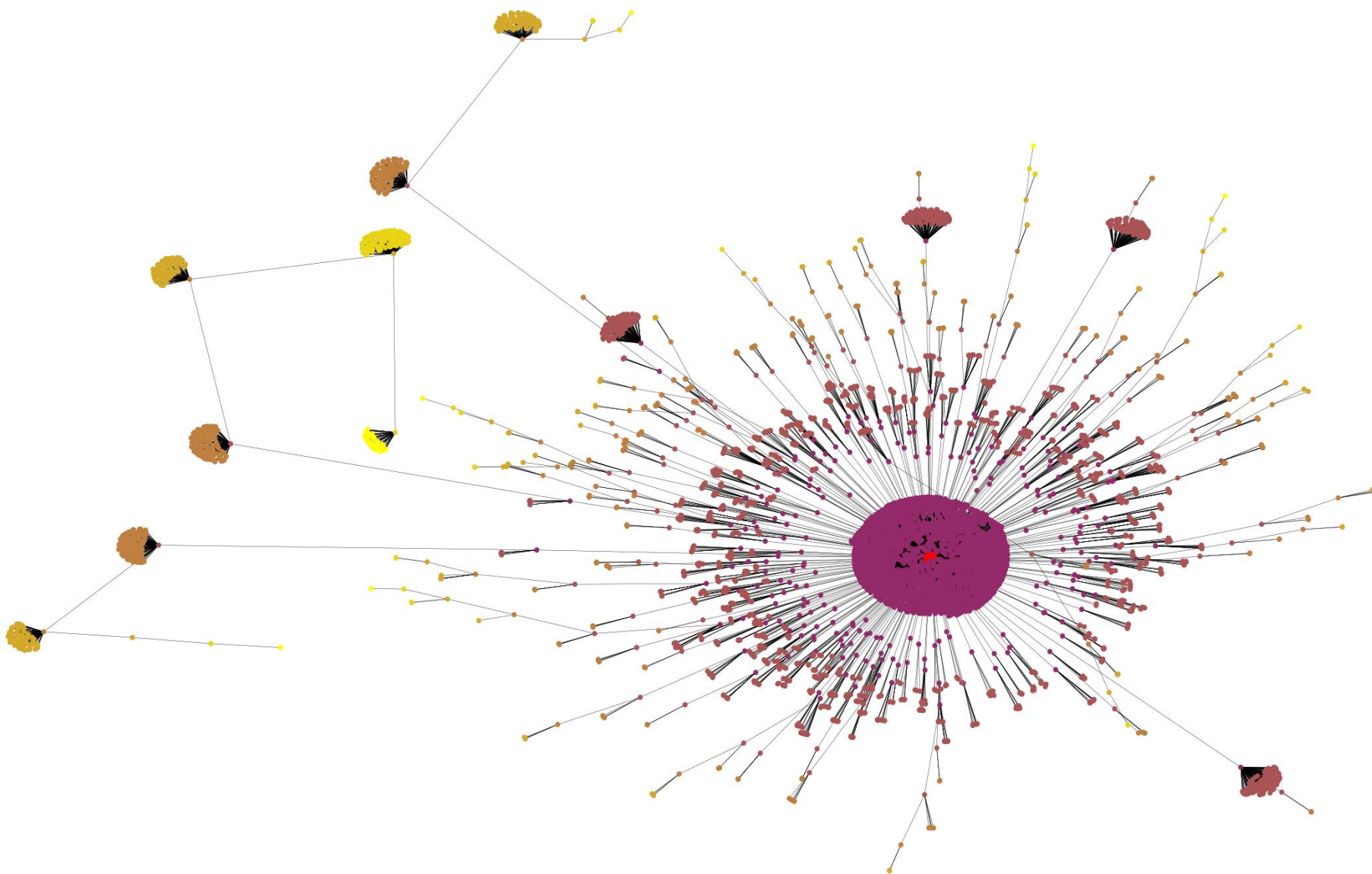


Fig A: A Viral Rumor Cascade

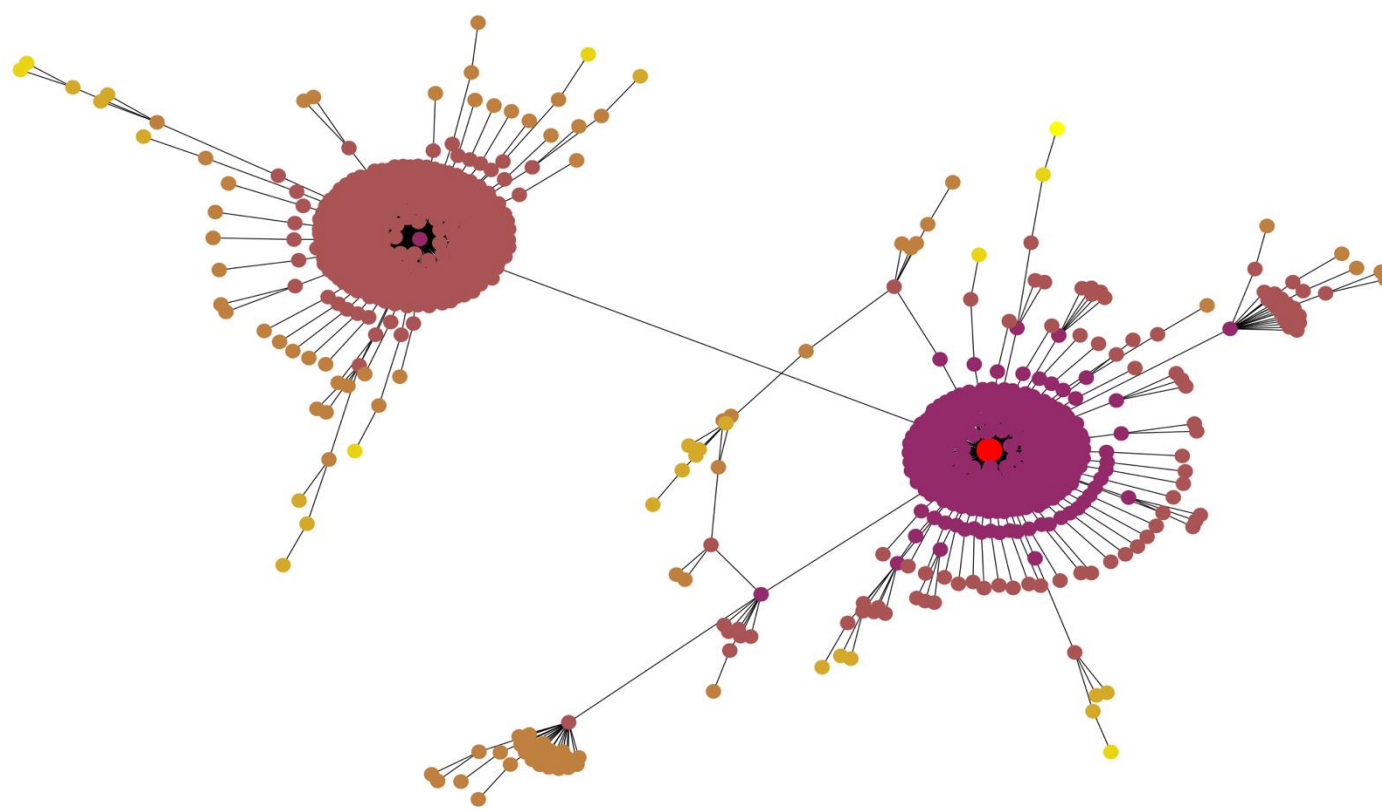
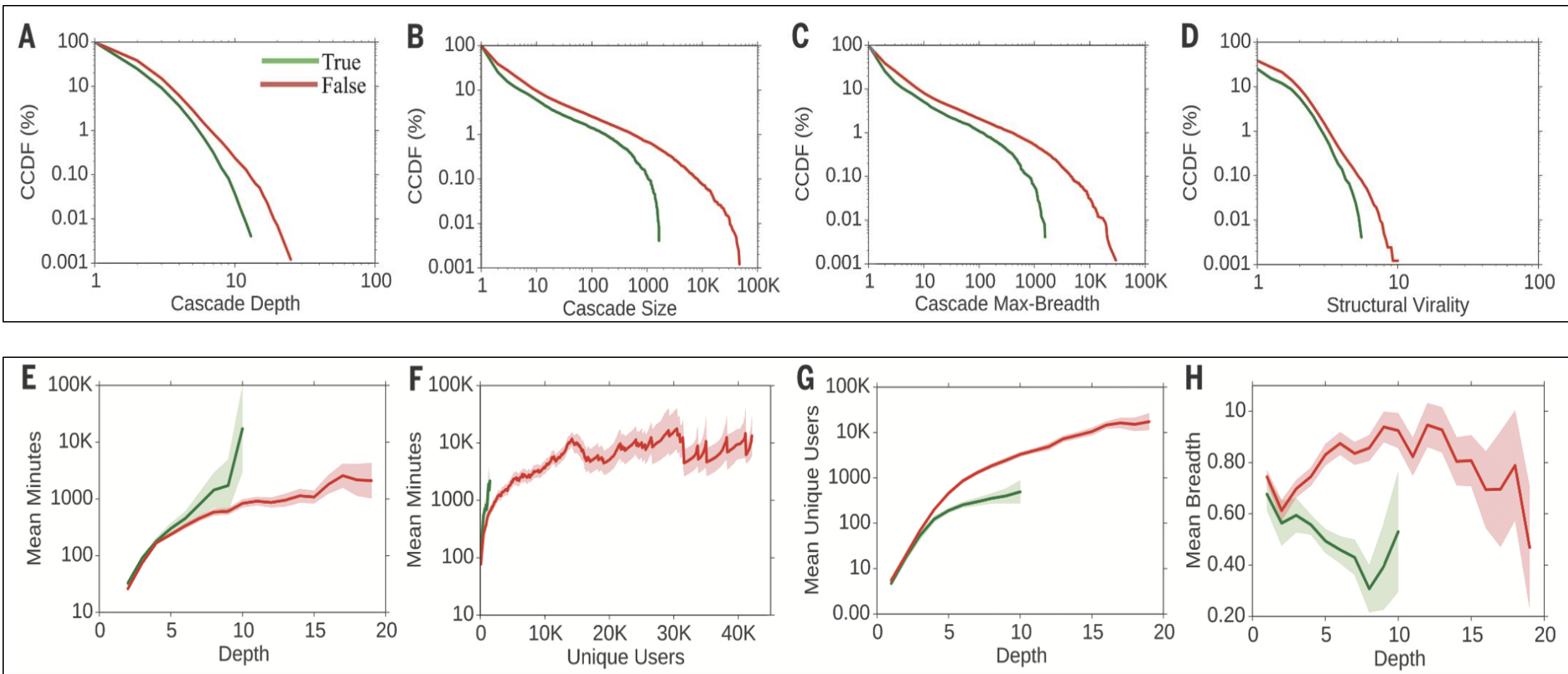


Fig B: A Typical Rumor Cascade

Analyzing Diffusion of True vs False Rumors



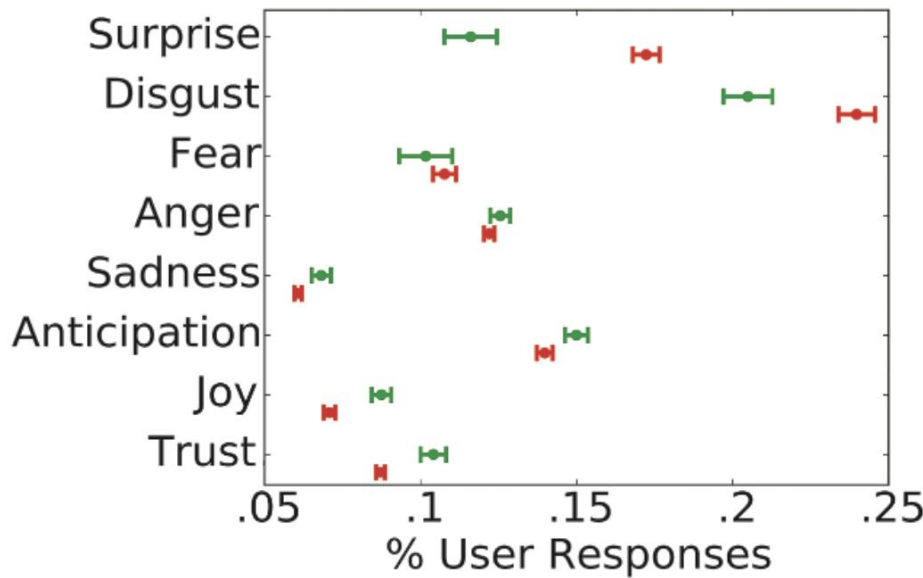
**Differential Diffusion:** False rumors reach greater depths, spread faster, and affect a larger audience compared to true rumors.

User Characteristics & Emotional Response Analysis

	median		mean		mean (log)		stdv (log)		ks-test
	false	true	false	true	false	true	false	true	
followers	410	466	2234	5240	2.62	2.68	0.69	0.88	D=0.104, p~0.0
followees	383	509	1002	1707	2.59	2.72	0.85	0.96	D=0.136, p~0.0
verified	0	0	0.002	0.006	nd	nd	nd	nd	D=0.005, p<0.001
engagement	9.52	9.54	19.70	24.65	0.91	0.90	0.65	0.76	D=0.054, p~0.0
account age	982	1214	1072	1269	2.90	2.97	0.39	0.42	D=0.125, p~0.0

Part A: User Characteristics

Contrary to expectations, this experiment did not reveal any significant differences between users who spread true vs false news. Therefore, this could not explain the differential diffusion.



Part B: Emotional Response Analysis

False rumors evoke stronger emotions, such as surprise and disgust, which likely drive their rapid spread, while true news evoked trust, joy and anticipation.



## Abstract

In recent years, smart grid-based Electric Vehicle (EV) charging systems have increasingly faced vulnerabilities to Distributed Denial of Service (DDoS) attacks, especially through malicious authentication failures. These attacks typically involve monopolizing the Grid Server (GS), thereby hindering the authentication process for legitimate EVs. Despite the severity of this issue, no research has focused on detecting DDoS attacks exploiting weaknesses in EV authentication. This study introduces a DDoS attack detection model specifically designed for EV authentication. The approach involves developing a machine learning model involving unique feature selection and combination. The proposed approach has been evaluated using a new DDOS attack dataset. The model is engineered to optimize feature combination, aiming for high sampling resolution, minimal information loss, and robust performance under 16 distinct attack scenarios. The feature combination used in this study shows improved accuracy over traditional DDoS detection methods based on access time variation while minimizing information loss.

## Research Problem

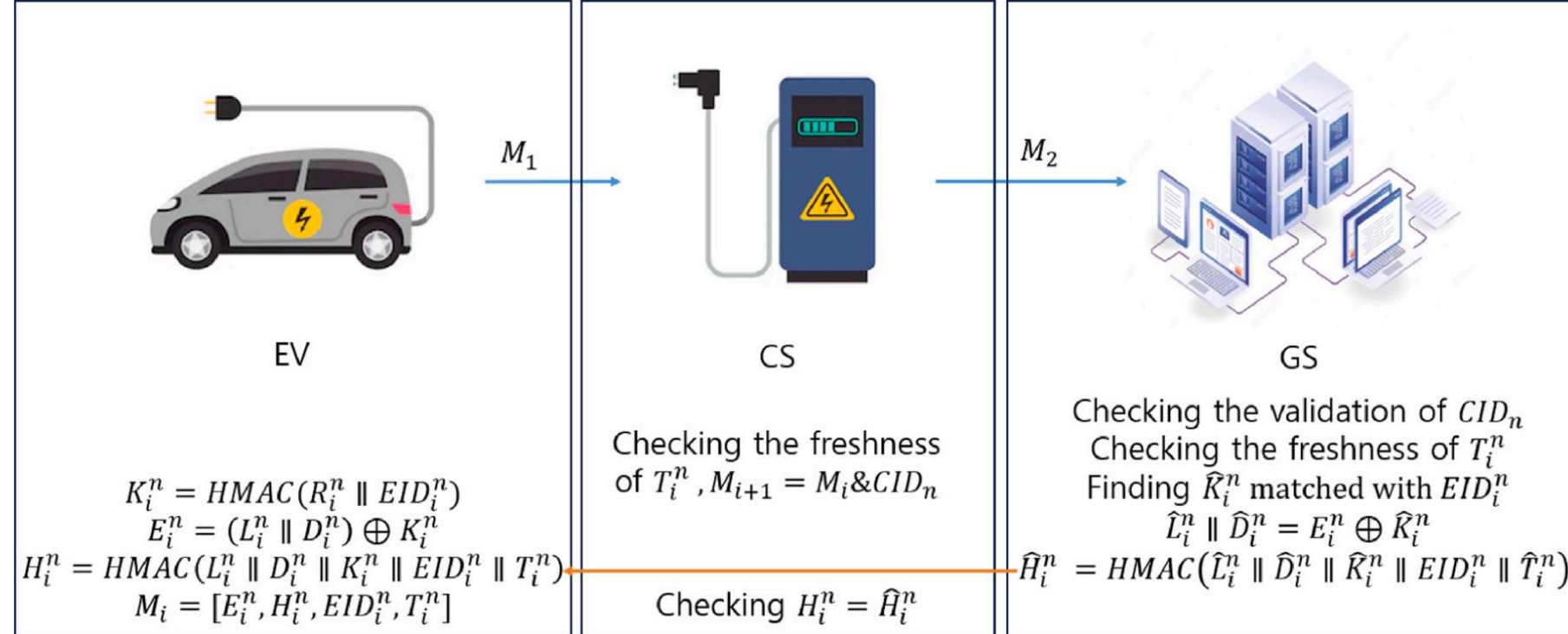


Fig. 1. EV Authentication protocol on the simulator (Kim et al., 2023a).

This work defines the problem as follows: there are "n" EVs, "m" Charging Stations (CS), and one GS involved across 16 scenarios, as depicted in Fig. 1. The focus is on various authentication attempts  $\omega$  made by the EVs to access the GS via the CSs. We aim to determine whether each authentication attempt  $\omega$  is legitimate or a fraudulent effort designed to

disrupt the authentication service. Let a set of EVs be  $EV_{\omega}^n = \begin{bmatrix} ev_1^1 & \dots & ev_i^1 \\ \vdots & \ddots & \vdots \\ ev_1^j & \dots & ev_j^j \end{bmatrix}$ ,  $1 \leq i \leq \omega$ ,  $1 \leq j \leq n$ , and a set of the CSs be  $CS_{\omega}^m = \begin{bmatrix} cs_1^1 & \dots & cs_k^1 \\ \vdots & \ddots & \vdots \\ cs_1^l & \dots & cs_k^l \end{bmatrix}$ ,  $k \leq m$ , where  $\forall i, \forall j \in N$ . In Fig. 1,

the parameters, such as  $H_i^n$ ,  $T_i^n$ ,  $\hat{H}_i^n$ , and  $\hat{T}_i^n$ , must be generated and communicated on the protocol in the same sequence for  $H_i^n = \hat{H}_i^n$  and  $T_i^n = \hat{T}_i^n$ . As a result, we get the following rule:  $0 < \dots < i - 1 < i < i + 1 < \dots < \infty$ . If an attacker cuts into the middle of the authentication processes of another EV, the parameters will be modified to  $H_{i+t}^{j-1}$  or  $T_{i+t}^{j-1}$ , where  $t$  is a count of continuous DDoS attacks. The parameters will become invalid between the EV, CS, and GS since  $H_{i+t}^{j-1} \neq \hat{H}_i^n$  or  $T_{i+t}^{j-1} \neq \hat{T}_i^n$ . Consequently, the legitimate EVs must wait for charging until the attacker finishes the number of  $t$  false authentication trials. Therefore, we need a novel attack detection model to accurately sense the DDoS attacks based on the EV false authentications.

## Methodology

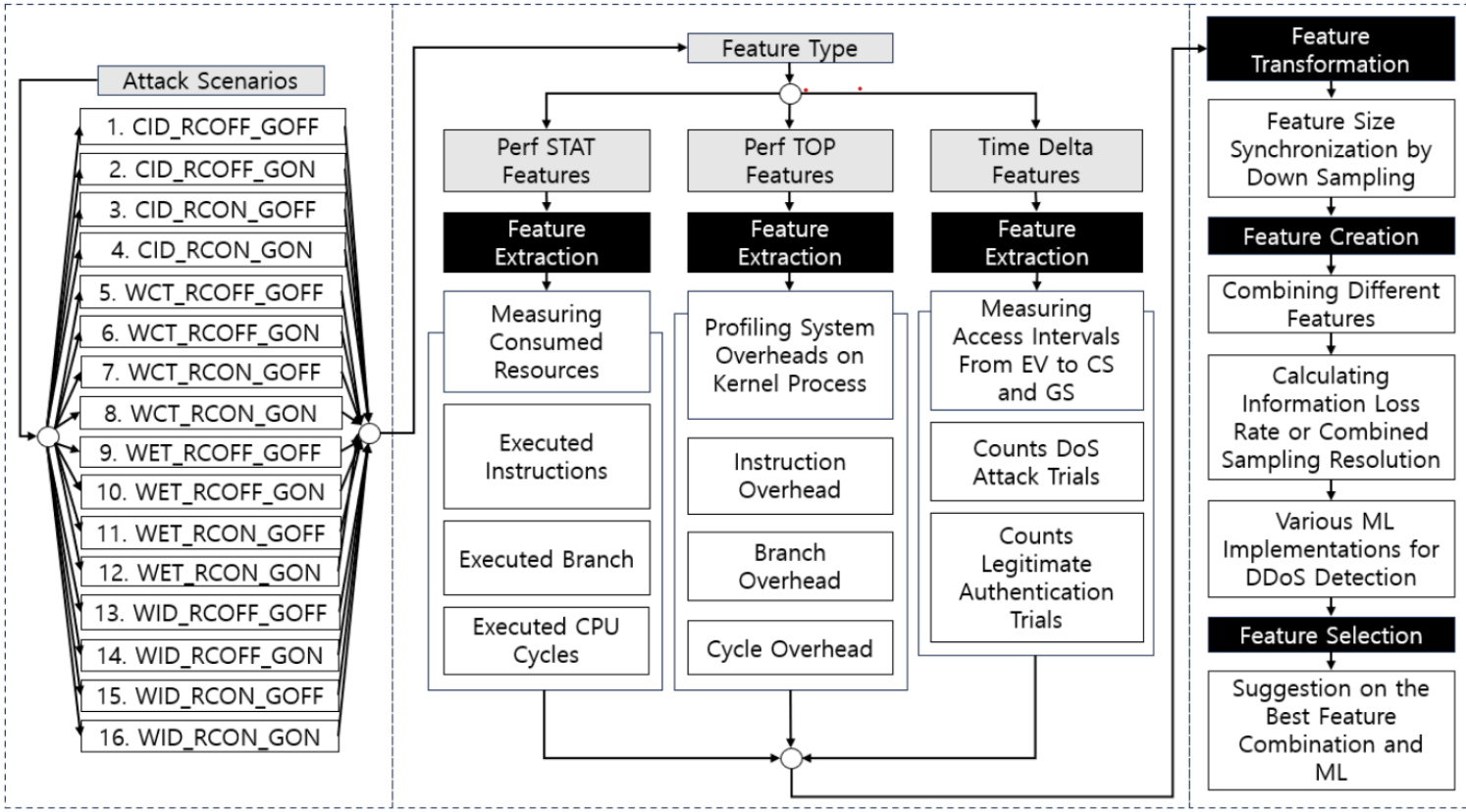


Figure 2: Schematic view of our attack detection model

- Fig. 3 illustrates the feature downsizing methods proposed in this work, following steps 1 through 7. In steps 1 and 2, the smaller feature is divided by an integer multiple of the size "m" of the larger feature. However, the orange portion cannot be evenly divided by this integer because the size "l" is smaller than "m," leading to internal fragmentation. In steps 3 and 4, indices pointing to the data points of the larger feature are extracted at intervals of "n." Additional indices may remain from the division between the small and large features, and in this case, "κ" extra indices are selected beyond the "m" indices. Steps 5 and 6 involve creating a new list by randomly selecting "m" indices from the "m + κ" indices extracted in step 4. Random selection ensures that the reduced list matches the size "m." Finally, the selected indices are sorted in ascending order to extract data points from the feature in step 2.

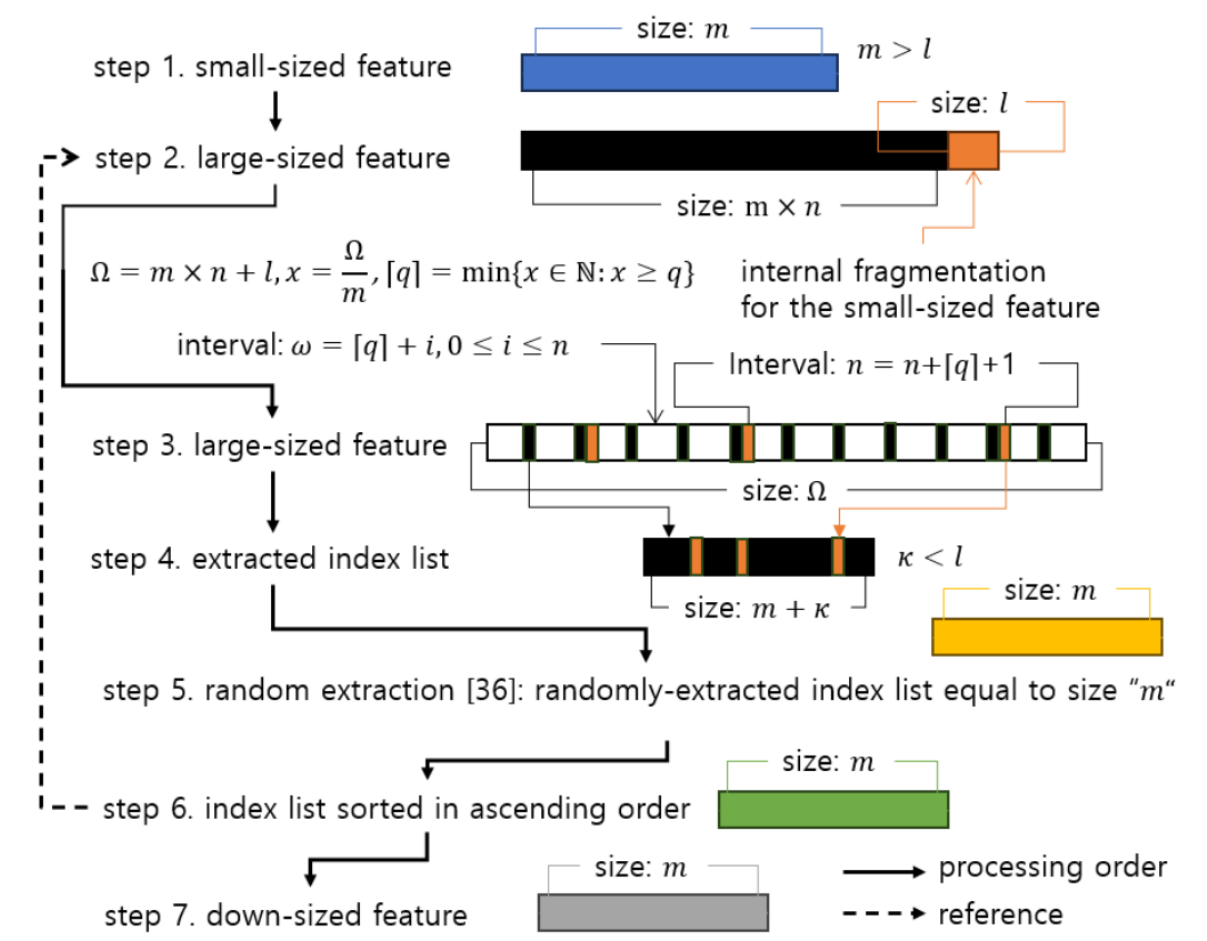


Figure 3: Feature size reduction process

## Result

- Detection Accuracy:** The proposed model achieved an average detection accuracy of over 95% across 16 different DDoS attack scenarios.
- Precision and Recall:** Both precision and recall exceeded 93% on average, with certain scenarios reaching up to 97%.
- Minimization of Information Loss:** The CLR approach applied during feature combination effectively suppressed the information loss rate to below 5%, ensuring that most of the original data's information was retained even after combination.
- Reduction of False Positive Rate:** The proposed model reduced the false positive rate to below 2%, ensuring high reliability in real-world operational environments.

## Conclusion

This study optimizes feature combinations for detecting DDoS attacks in EV charging infrastructures using ML, exploring 16 attack scenarios and evaluating ML classifications. Key features include Time Delta, Perf STAT, and Perf TOP, offering a comprehensive approach to DDoS detection. The study introduces a feature downsizing method to minimize information loss and identifies the best feature combination based on criteria like low information loss, high sampling resolution, F1 score, and feature size. Employing Perf TOP and Perf STAT significantly enhances detection accuracy compared to relying solely on Time Delta. The proposed sampling method preserves more information than traditional methods when downsizing features. The approach's effectiveness is supported by multiple regression and Spearman's correlation analysis. Future research will focus on applying grid search techniques to refine feature combinations and optimize ML model hyperparameters.

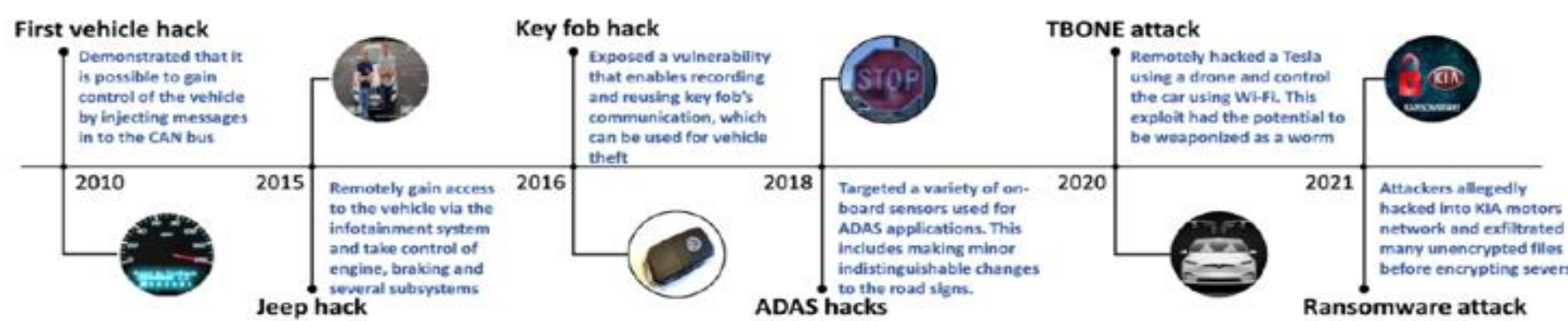


## ABSTRACT

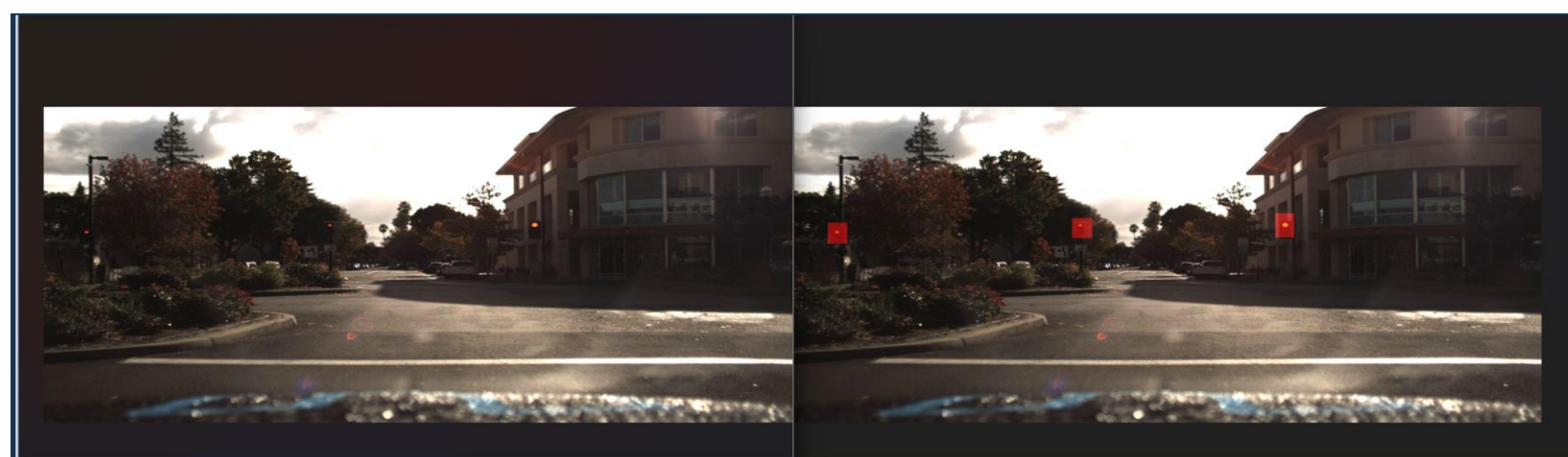
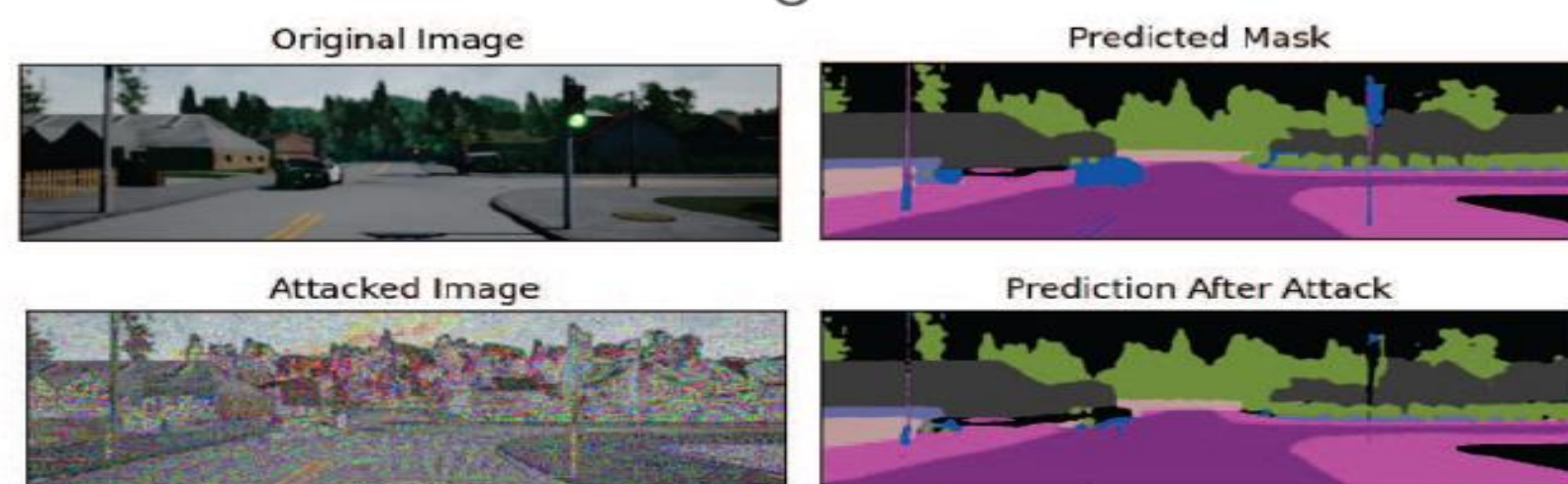
In the evolving landscape of autonomous vehicle technology, the threat of cyber attacks poses a significant challenge, particularly through the manipulation of traffic signal recognition systems. Such vulnerabilities can compromise the safety and efficiency of autonomous navigation, necessitating the development of more sophisticated anomaly detection methods. This research addresses the critical need for enhanced cybersecurity measures by investigating the resilience of neural network models, specifically Convolutional Neural Networks (CNNs) and autoencoders, against visual anomalies caused by altered traffic signals. Despite the increasing reliance on deep learning for autonomous vehicle perception, there is a lack of comprehensive strategies that effectively mitigate the risks associated with these cyber threats. Therefore, the study aims to evaluate and improve the accuracy and reliability of anomaly detection systems under various noise conditions, contributing to the safeguarding of autonomous vehicles against potential cyber attacks.

## Motivation

This research advances cybersecurity in AI-driven systems, particularly in autonomous vehicles, by improving anomaly detection and robustness in object detection systems. It lays the groundwork for safer, more efficient transportation, contributing to the resilience of AI technologies against cyber threats.



## Data Augmentation



## Data Augmentation

- The Self Driving *Car.v2-fixed-large.tensorflow* Is used in this experiment. To simulate traffic attack, we manipulated the images with color masking and complex random noise manipulations. **Blended Image =  $\alpha * (\text{Original ROI}) + \beta * (\text{Mask}) + \gamma$ .**
- A noise distortion was also added to the **image by flattening the ROI into a 2D array** and introducing noise, the noisy ROI is then reshaped **back to its original 3D form** which preserves the spatial structure of the image

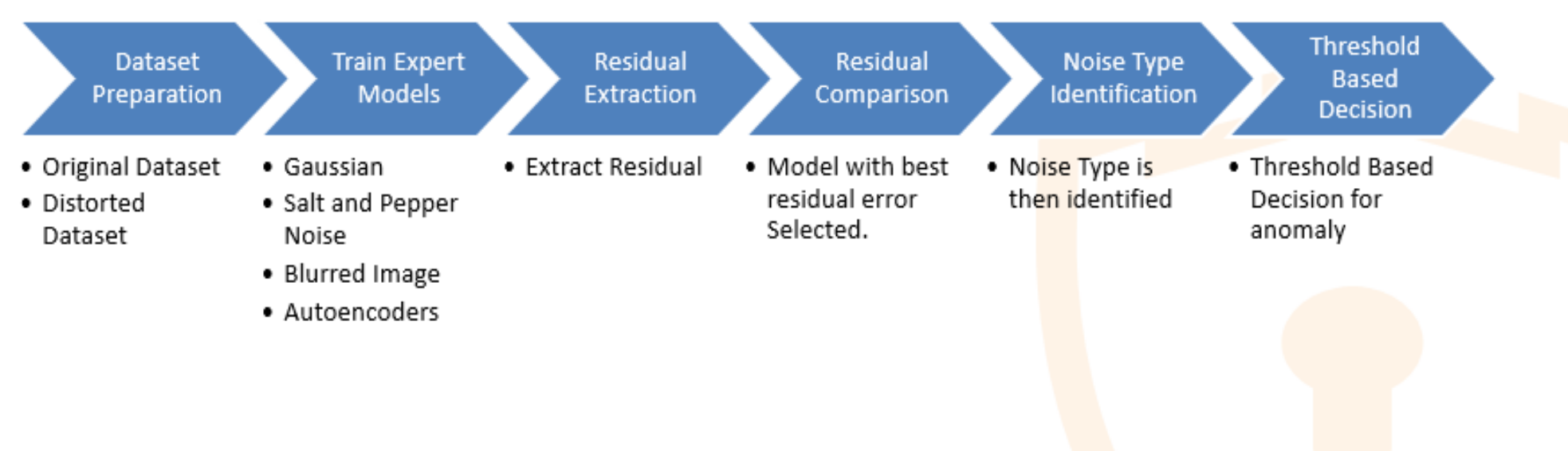
## Robustness Test

- In this experiment, we aimed to assess **the robustness of our neural network model to Gaussian noise**, a common type of statistical noise that simulates real-world environmental and sensor inaccuracies
- The aim is to check if our model reacts to just any noise to decide or its effectively **making decisions concerning traffic signals**.



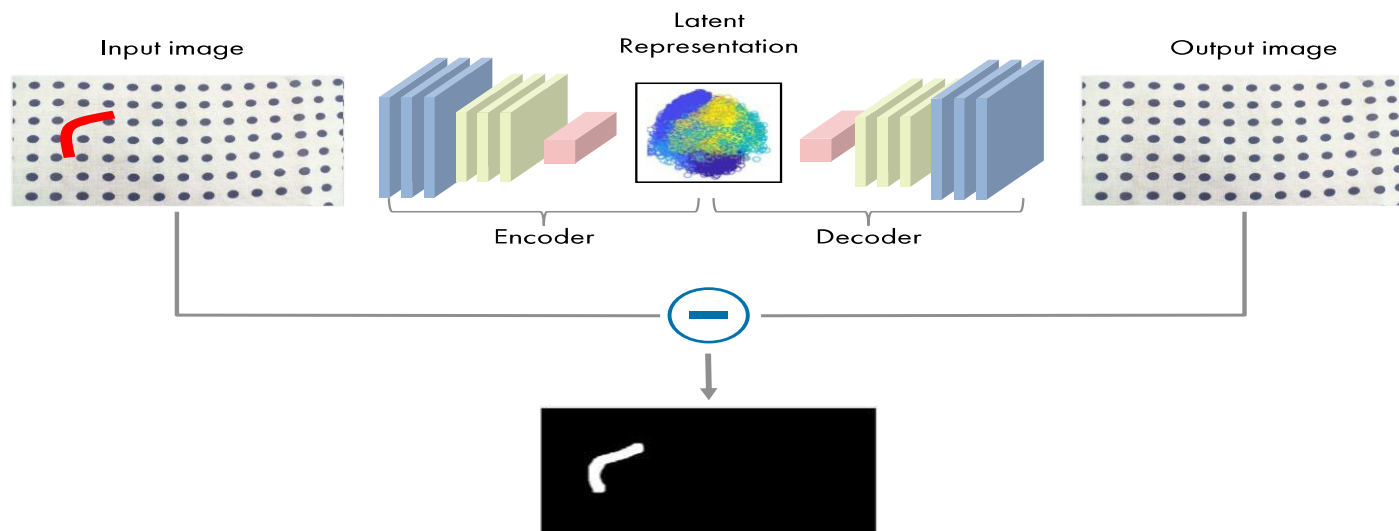
Fig. 3 Feature Visualization to see what our model is focused on

## Our Approach



## Detail Methodology

### Generative Approach



The MSE loss function  $L$  for a given input  $x$  and its reconstruction  $\hat{x}$  is given by:

$$L(x, \hat{x}) = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2$$

where  $n$  is the number of pixels in each image.

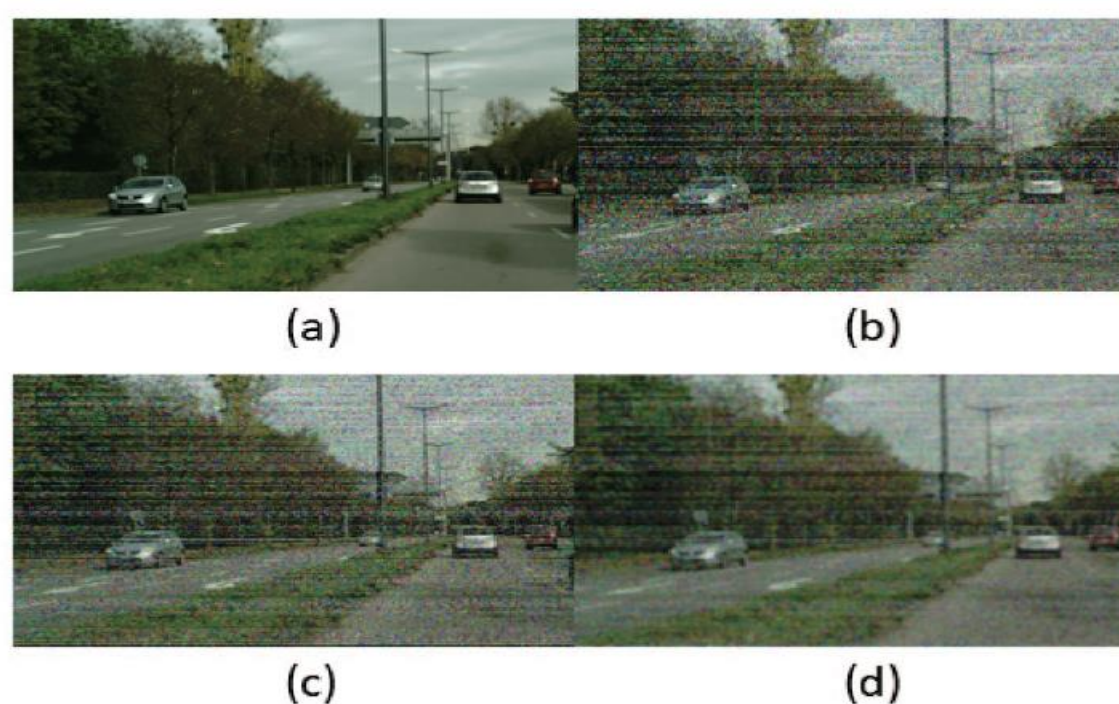
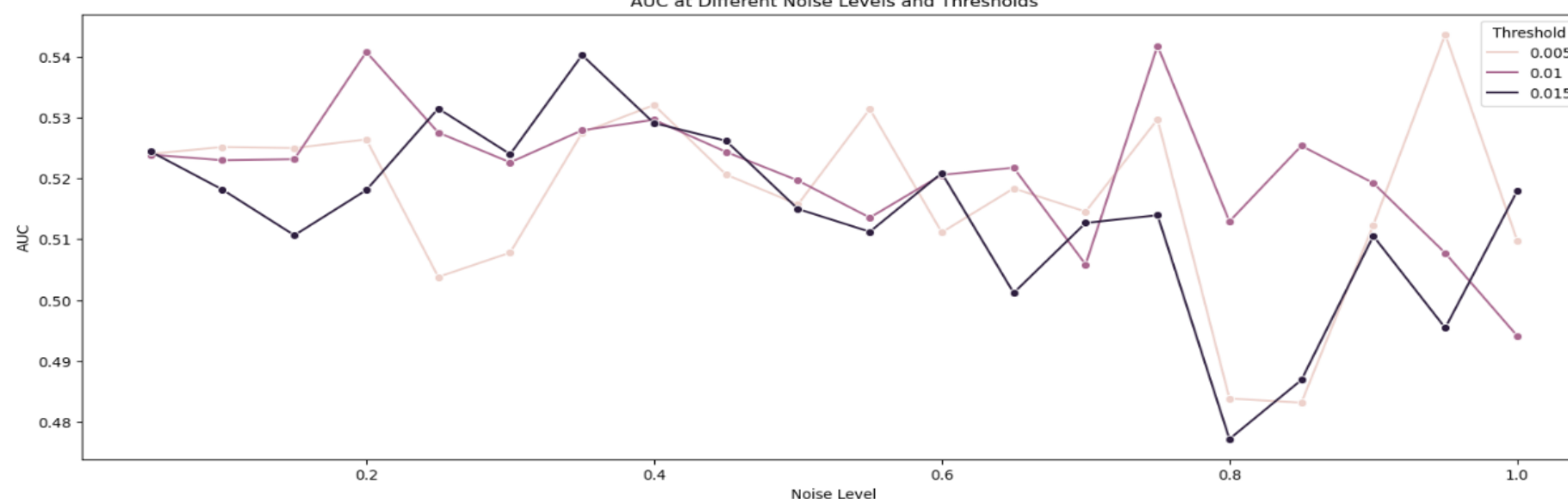


Fig. 6. Existing methods for removing noise: (a) Input Image from Cityscapes dataset. (b) Noisy image. (c) Result of applying bilateral filter. (d) Result of applying Gaussian smoothing.

AUC at Different Noise Levels and Thresholds



### Performance Evaluation

Metric	Value (%)
Accuracy	93
Reconstruction Error Threshold	85

Table 2: Autoencoder performance in anomaly detection

- RI-N Formula: Overlap of  $n$ -sized pixel patches:  
$$RI - N = \frac{\sum \min(\text{original patch}, \text{reconstructed patch})}{\sum \text{original patch}}$$
- RI-W Formula: Incorporates weight  $w_i$  for important regions:  
$$RI - W = \frac{\sum w_i \cdot \text{length of common subregion}}{\sum w_i \cdot \text{total length of region}}$$
- Weighting based on saliency or domain knowledge to emphasize critical regions.

### Novel Proposed Evaluation Method



## ABSTRACT

The curse of dimensionality is a well-recognized challenge in the machine learning field. As the internet grows and technological advancements progress, the data generated daily tends to have increasingly high dimensions, many of which are irrelevant, sparse, noisy, or redundant. This can negatively impact the efficiency and effectiveness of many algorithms. Dimensionality reduction techniques, particularly feature selection methods, are commonly used to address this issue. Specifically, this review will focus on recent developments in unsupervised feature selection through graph learning. These methods typically involve learning the manifold structure of the original data, projecting it to a lower-dimensional space, and selecting features using sparse learning, which retains only the most important and discriminative features. This review also covers the algorithms used to learn data structures, how sparse learning is applied, and the challenges of learning in different spaces.

## Feature Selection

- Get a meaningful feature subset of the original feature space by removing redundant and irrelevant features
- Maintaining the physical structure of the original features
- Preferring unsupervised methods due to labeling cost

## Graph Learning

- Constructing a graph affinity matrix to describe the local geometric structure of data
- Achieving feature selection by a sparse learning model, i.e., maintaining the local structure of the data with a subset of features



How to learn the affinity graph?



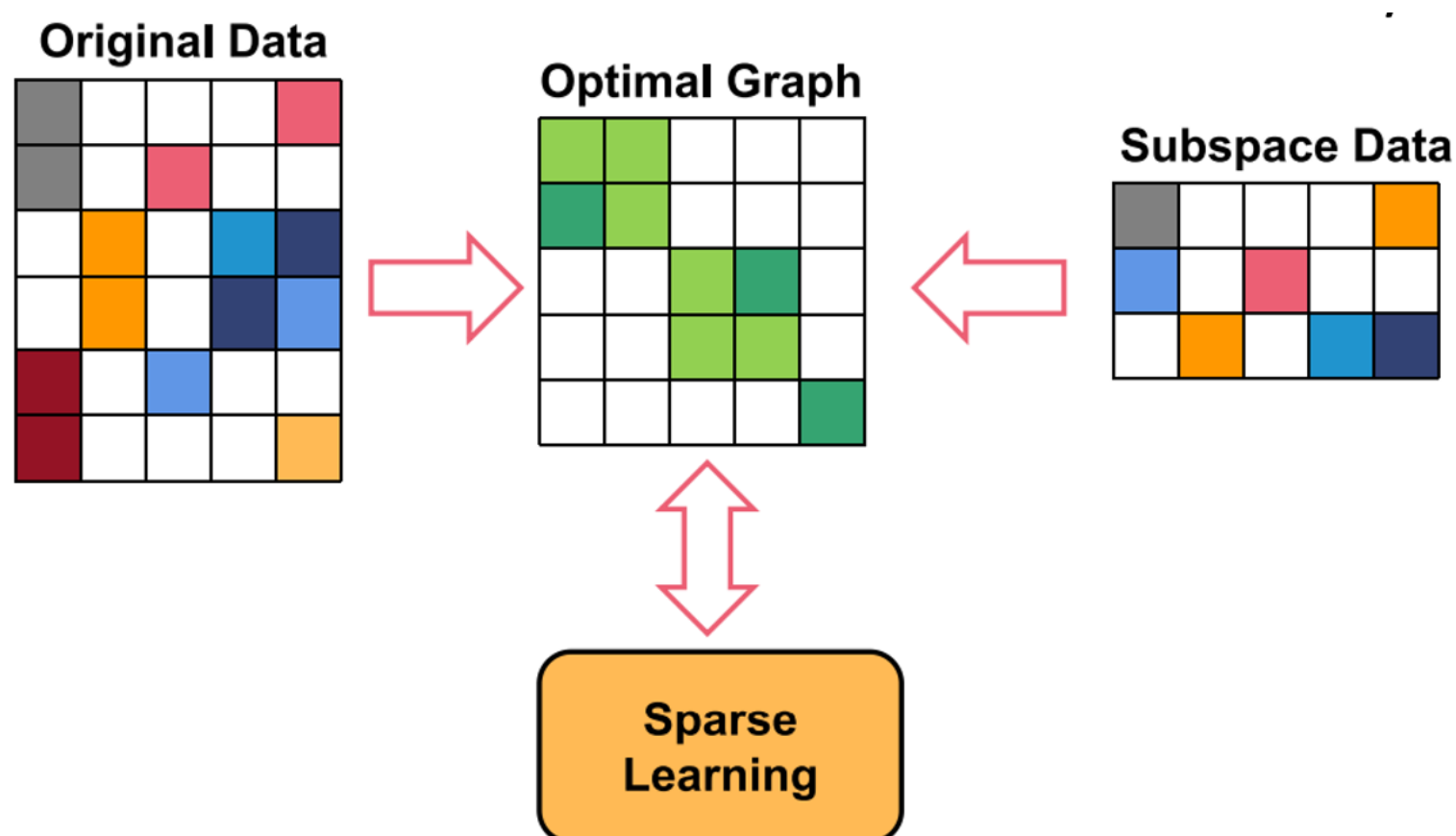
In what space the graph should be learned: Original space or subspace?



How to apply sparse learning?

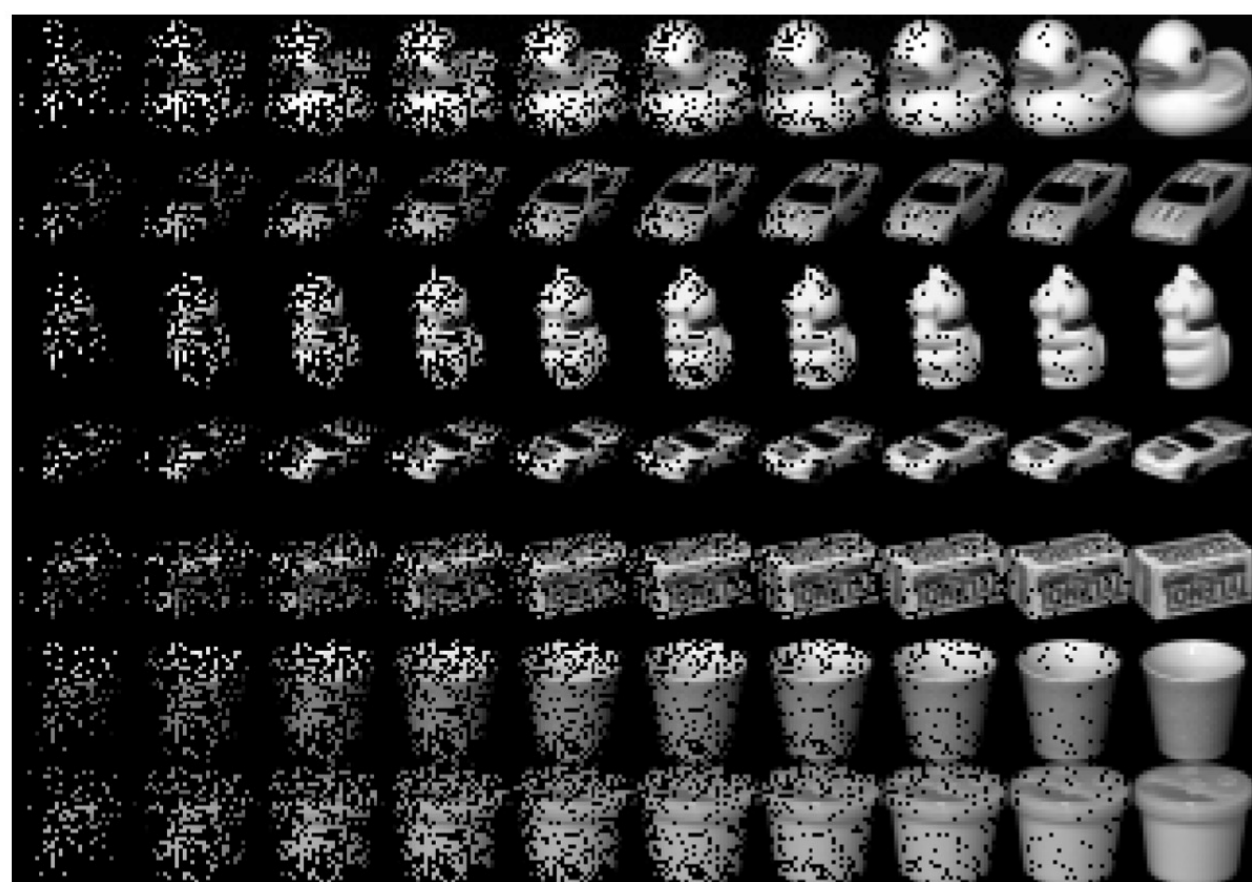
## Choosing the Space

- Original Space: Affected by a lot of noise and redundant information
- Subspace: Maintaining the similar structure to the original space is difficult
- Joint space: Original space Helps maintaining the structure of original data and subspace weakens the influence of original spatial noise and redundancy



## Visualization of the selected features

- Visualization of selected features in the Coil20 dataset. Even with very low selected features the manifold structure of the original data is maintained to a good extent.

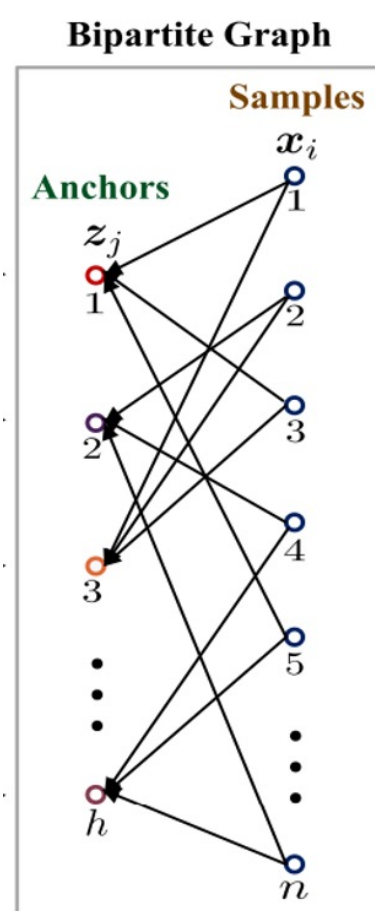
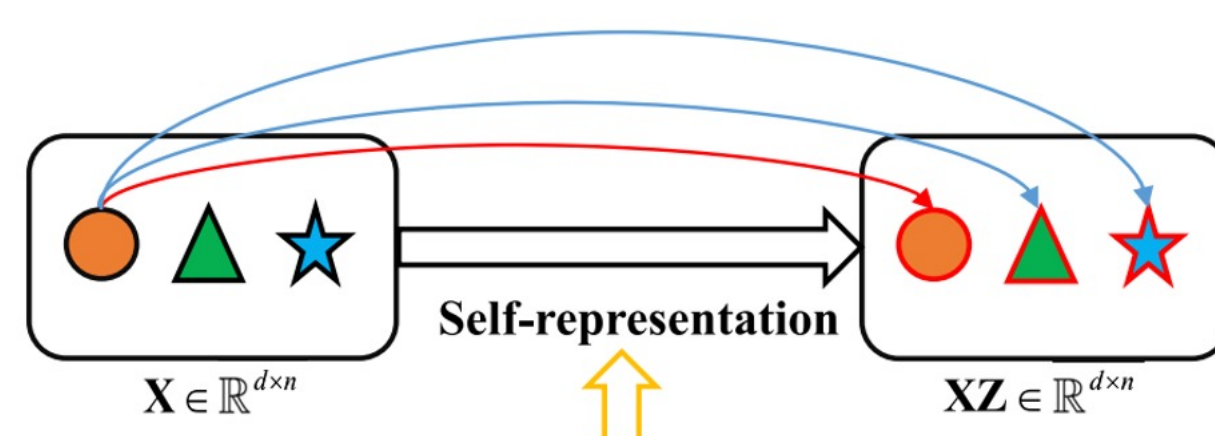


## Learning Affinity Matrix

- 1) Self-representation learning: Expressing each sample as a linear composition of all samples

$$\mathbf{x}_i = Z_{1i}\mathbf{x}_1 + Z_{2i}\mathbf{x}_2 + \dots + Z_{ni}\mathbf{x}_n$$

Utilizing contrastive learning to achieve this by using the sample and its self-representation as the positive pair and every other sample and their corresponding self-representation as negative pairs



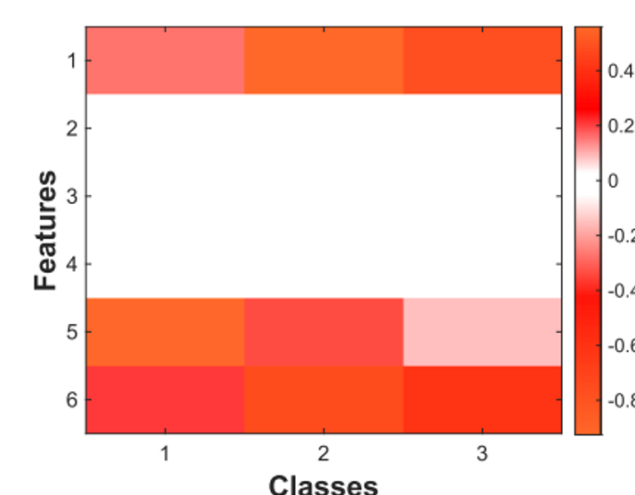
- 2) Anchor-based: Generate some anchors and calculate the distance between the anchors and the original samples

Using Balanced K-means based Hierarchical K-means (BKHK) algorithm to generate better anchors

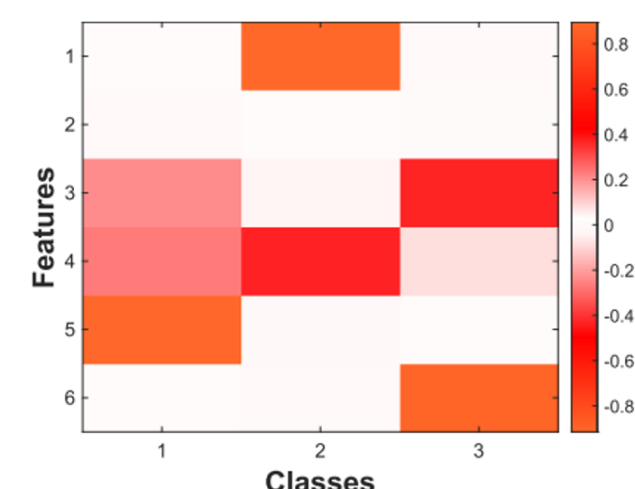
## Sparse Learning

Projecting the data to a subspace and enforcing sparsity via regularization

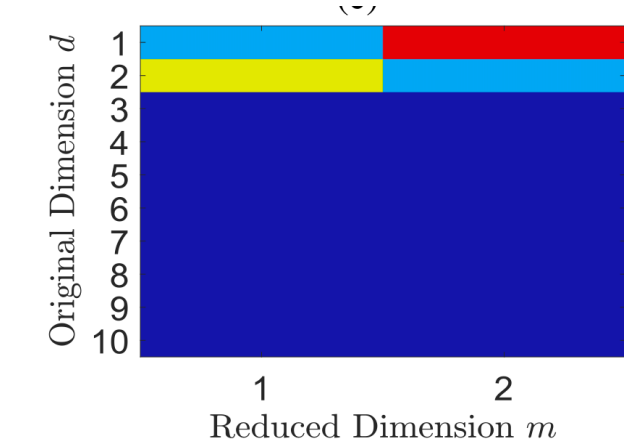
L 2, 1-norm: Row sparsity - Removes redundant features shared by all data



L 1, 2-norm: Column sparsity - Removes specific redundant features per class



L 2, 0-norm: Exact row sparsity – ensures that there is exactly k non-zero rows



## References

- [1] X. Dong, F. Nie, D. Wu, R. Wang, and X. Li, "Joint structured bipartite graph and row-sparse projection for large-scale feature selection," IEEE Transactions on Neural Networks and Learning Systems, 2024.
- [2] F. Nie, W. Zhu, and X. Li, "Structured graph optimization for unsupervised feature selection," IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 3, pp. 1210–1222, 2019.
- [3] Q. Zhou, Q. Wang, Q. Gao, M. Yang, and X. Gao, "Unsupervised discriminative feature selection via contrastive graph learning," IEEE Transactions on Image Processing, 2024.





CIC

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)

# Privacy-preserving Machine Learning in IoT Utilizing TEE and Lightweight Ciphers

Arash Kariznovi, Kalikinkar Mandal

Contact Email: arash.k@unb.ca, kmandal@unb.ca



## ABSTRACT

The rapid growth of IoT and resource-constrained devices has increased the demand for lightweight cryptography. In response, NIST has standardized the ASCON lightweight AEAD and hash algorithm for this purpose. Beyond secure IoT communications, IoT data analytics are essential for efficiency, innovation, decision-making, and predictive maintenance. In this paper, we propose a privacy-preserving machine learning (PPML) system for securely transporting IoT data to the cloud and enabling secure machine learning. Our protocol, based on a lightweight AEAD scheme and TLS, resists various attacks, and we use Intel-SGX for secure analytics. We prototype and evaluate the system on real-world datasets.

## System Model

**System Model:** We consider a real-world Cloud-IoT system consisting of three key entities, namely a set of IoT devices, a gateway, and a cloud server. The IoT devices are connected to the cloud via a gateway and periodically transmit IoT data to the cloud. For simplicity, we assume that all the devices belong to a single owner who wishes to perform various tasks on IoT data such as data collection, processing.

**Adversarial Model:** We consider semi-honest adversaries where an adversary may compromise some IoT devices or the cloud applications, and observes the execution of the protocol. The goal of the adversary is to learn any unintended information about other honest IoT devices' data or the trained model. We assume that the adversary can intercept the IoT data communications including record, replay, and modify network data and can compromise the cloud software applications.

## Problem Statement

Consider an Internet of Things (IoT) system where multiple IoT devices continuously stream high-dimensional data to a cloud service. The dataset generated by these IoT devices is denoted as  $D = \{ (x, y) \}$ , where each  $(x, y)$  represents a high-dimensional data point. The goal is for the data owner to train a regression model  $\theta$  using the dataset  $D$ , such that:

$$\theta \leftarrow \text{Training}(\theta, D, f)$$

where  $f$  is either a linear or logistic regression algorithm.

The primary challenge is securely transmitting the IoT data from the devices to the cloud, and ensuring that the training of the regression models (linear or logistic) on the dataset  $D$  in the cloud is performed privately and securely. The system needs to protect the privacy and integrity of the IoT data at multiple stages: In-transit, at-rest, In-use.

## Security Model

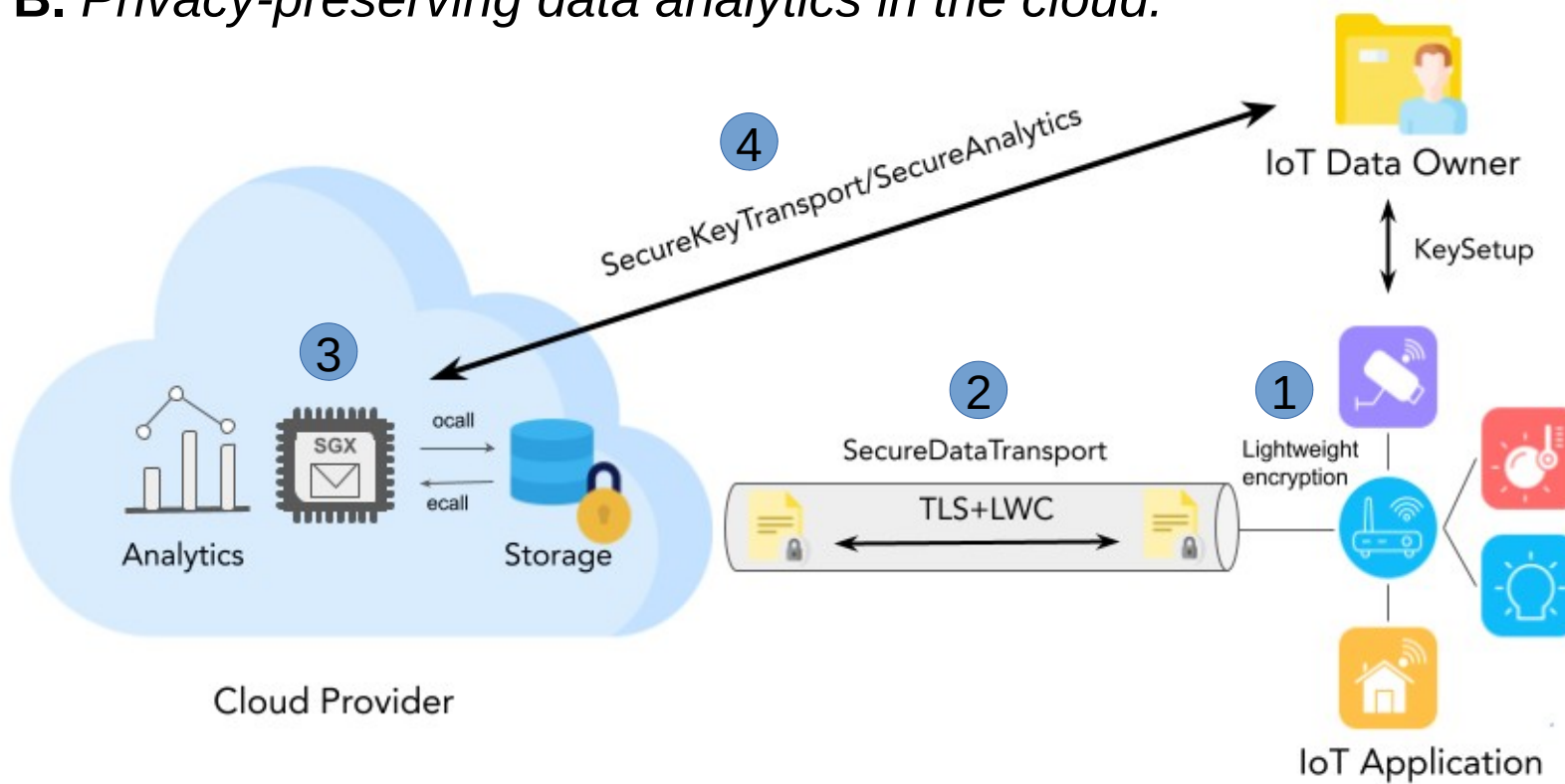
Our system ensures IoT data confidentiality, integrity, and privacy with multiple layers of protection: AEAD, TLS, and SGX attestation safeguard data from man-in-the-middle attacks, insider threats, and external cyber threats during data in-transit, at-rest, and in-use.

Assume the lightweight AEAD scheme is secure under indistinguishability under chosen plaintext attack (IND-CPA), TLS and EKEP protocols are secure, and the TEE (SGX) is trusted. Our system is secure against semi-honest adversaries.

## Our Proposed Scheme

We design a secure Cloud-IoT analytics system using practical cryptographic tools including SGX, TLS, and ASCON. The system handles two main tasks:

- Secure data transmission from IoT devices to the cloud.
- Privacy-preserving data analytics in the cloud.



Using the bring-your-own-encryption (BYOE) model, the data owner manages encryption keys. A double-encryption mechanism combines AEAD and a TLS-like protocol for protection against external and internal threats. Analytics are executed in a trusted environment to ensure data privacy.

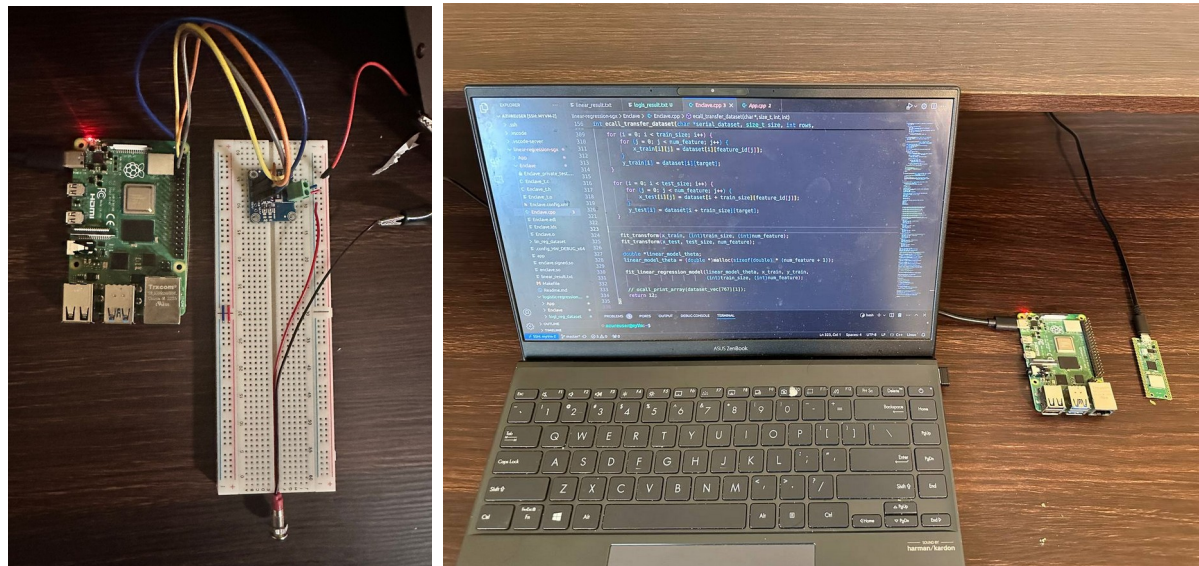
## Conclusion and Future Work

This framework presents a privacy-preserving ML system for Cloud-IoT that securely collects and trains models on fine-grained IoT data. It uses a lightweight AEAD scheme for efficient encryption and a TEE to protect data privacy during training. The system's practicality was demonstrated by encrypting IoT data, transmitting it securely to the cloud, and training models on real-world datasets.

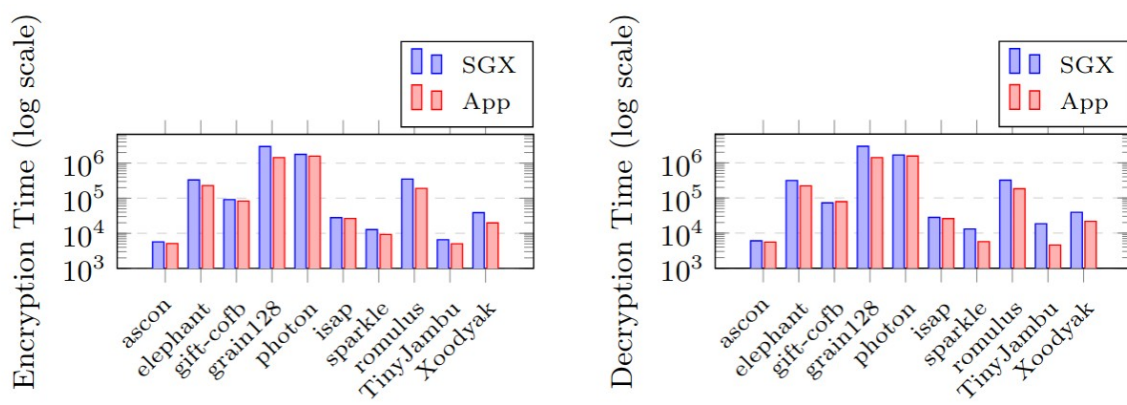
Future work will focus on implementing application-specific deep learning algorithms, exploring advanced encryption techniques, and enhancing the scalability of the system for larger datasets.

## Experimental Analysis and Results

Part A. Our Experimental setup



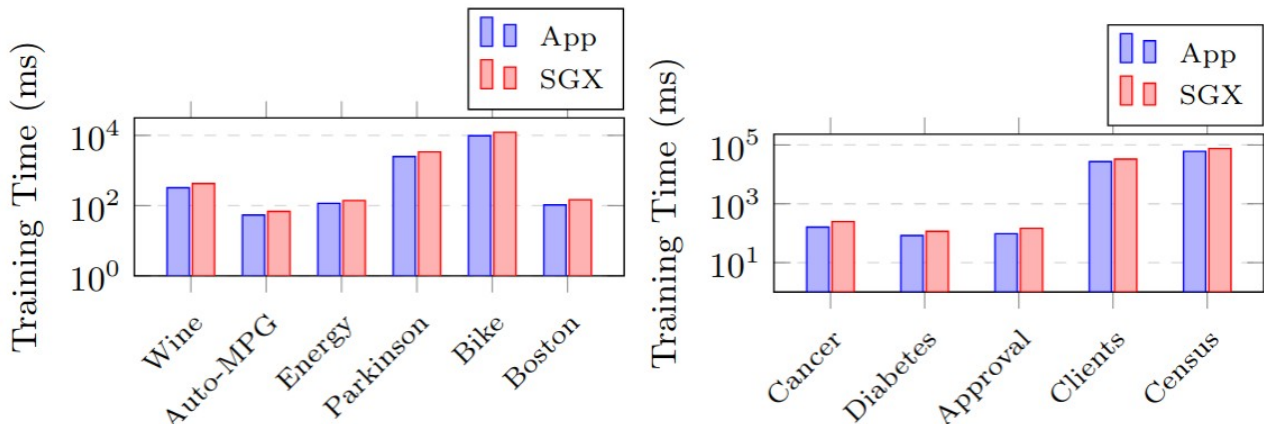
Part B. Experimental Result of NIST LWC in Azure cloud



Part C. Encryption result in/out SGX

AEAD Cipher	SGX Encryption (clock cycle)	App Encryption (clock cycle)	Overhead (x)
Ascon	5,708	5,122	1.11
Elephant	331,888	212,682	1.56
GIFT-COFB	90,864	72,918	1.25
Grain128	2,972,530	1,373,994	2.16
ISAP	28,020	26,520	1.06
Photon-Beetle	1,764,714	1,554,082	1.14
Romulus	350,980	157,296	2.23
TinyJambu	6,486	5,544	1.17
Sparkle	12,748	5,656	2.25
Xoodooak	38,764	19,818	1.95

Part D. Experimental Result of regression models



Part E. Model training results in/out SGX

Model Type	Name	n	d	Untrusted (ms)	SGX (ms)	Overhead (x)
Logistic Regression	Diabetes	9	768	83	115	1.387
	Credit Approval	14	652	96	145	1.515
	Breast Cancer	32	454	162	247	1.523
	Credit Card Clients	24	30,000	26,972	33,245	1.231
	US Census Income	20	48,842	60,276	75,595	1.255
Linear Regression	Auto MPG	8	392	53	68	1.283
	Boston Housing	14	506	104	147	1.415
	Energy Efficiency	9	768	115	138	1.217
	Wine Quality	12	1,599	323	425	1.316
	Parkinson	22	5,875	2,489	3,392	1.362
	Bike Sharing	12	17,379	9,777	12,248	1.251

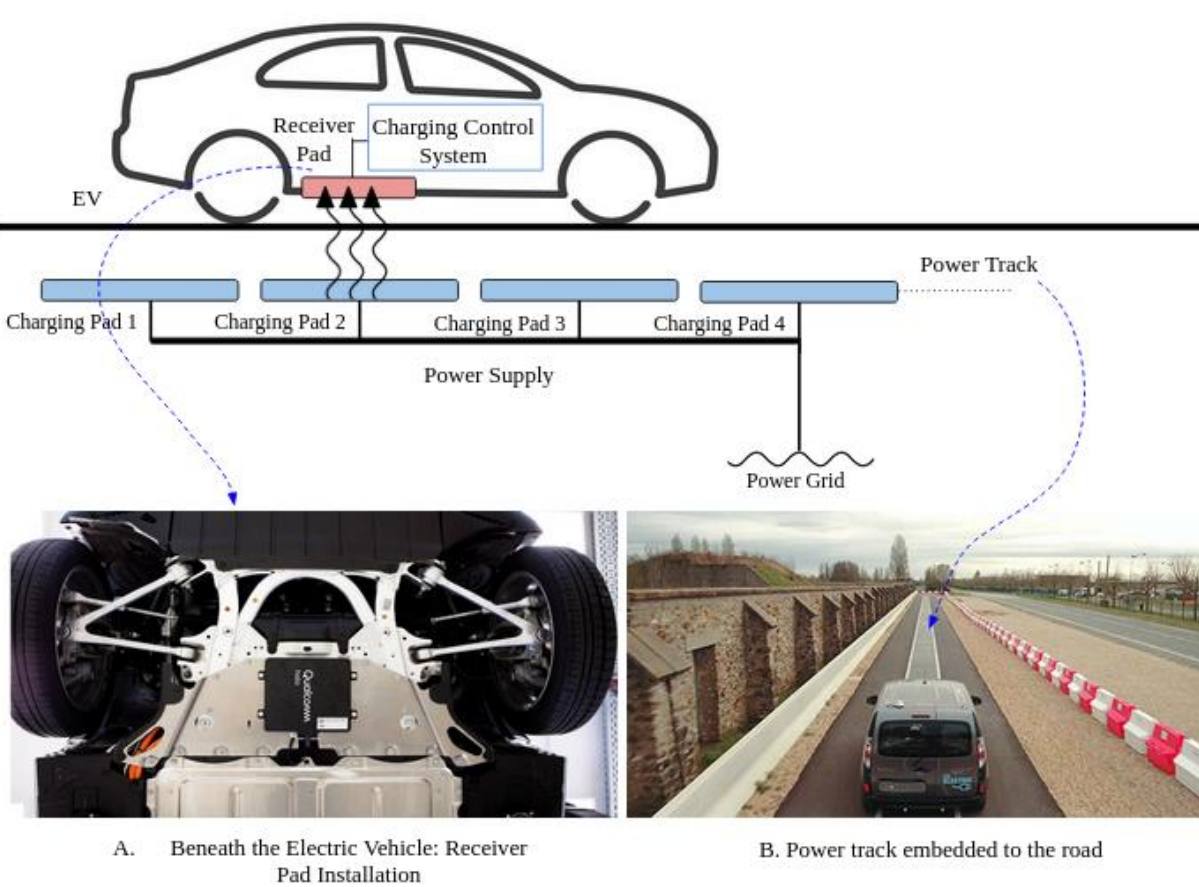
**Acknowledgment :** This work was supported by the NBIF RAI 2024 and NSERC Discovery grants.



## ABSTRACT

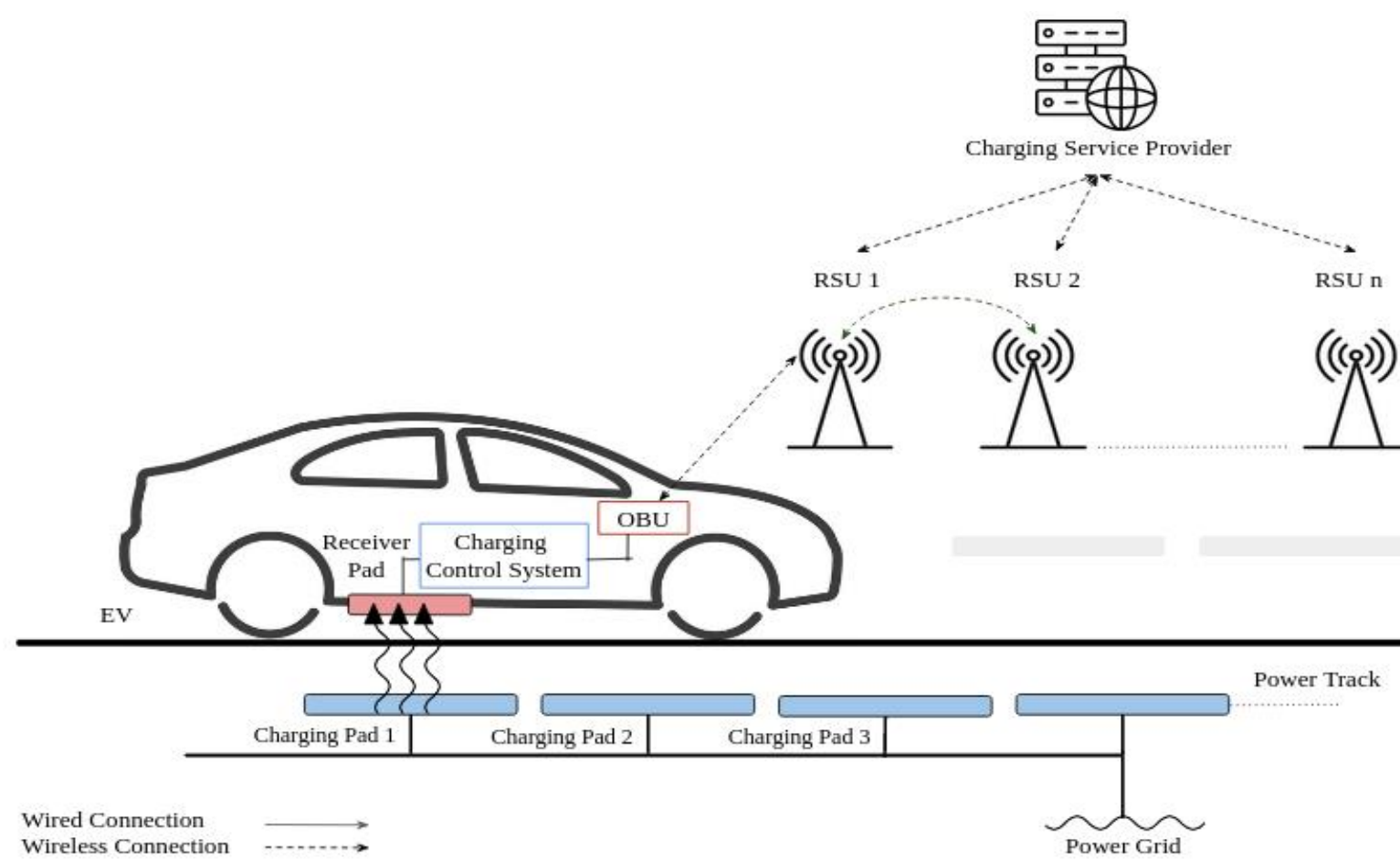
Electric Vehicles (EVs) are considered the predominant method of decreasing fossil fuels as well as greenhouse gas emissions. With the drastic growth of EVs, the future smart grid is expected to extensively incorporate dynamic wireless charging (DWC) systems, a significant advancement over traditional charging methods. DWC, offering the unique ability to charge vehicles in motion, introduces new infrastructures, complex network models and consequently, a massive attack surface. To accomplish the goal of such an enormous smart grid accompanying DWCs, the security of EV charging infrastructures has become a deciding factor. EV charging is vulnerable to cyberattacks as it has many attack vectors and many challenges to combat. Unlike the traditional charging services provided in a typical static charging station, the DWC has a complex network architecture which makes it vulnerable to many forms of cyberattacks. Authentication plays a crucial role in safeguarding the frontline security of this ecosystem. However, within the domain of DWC, the current academic landscape has seen limited attention dedicated to authentication protocols. This background signifies the necessity of a robust revocable anonymous authentication framework for dynamic EV charging.

## Components of Dynamic Wireless EV Charging System

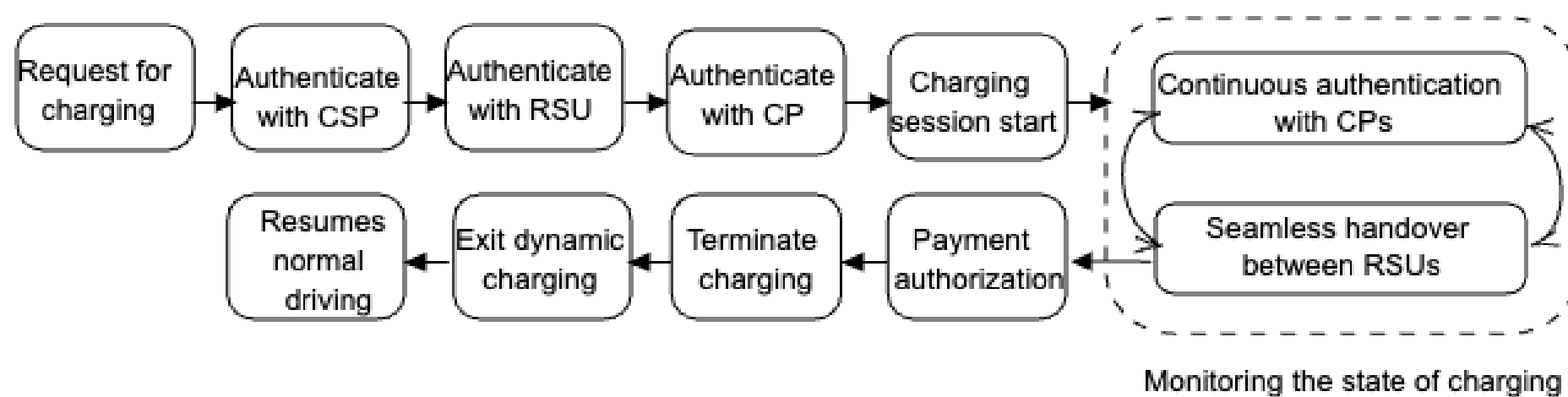


- 1. Power Supply:** It connects a dynamic charging system to the smart grid to receive power.
- 2. Charging Infrastructure/ Power Track:** It is embedded into the road surface or the side of the pavement, facilitating the wireless transfer of electric energy when vehicles drive over the track
- 3. Charging Pad (CP)/ transmitter (primary) pads:** The power track consists of a chain of CPs that use an electromagnetic field to transfer energy wirelessly to the EVs.
- 4. Receiver Pad/ secondary pad:** It is embedded at the bottom of the EV to receive electricity.

## Communication Model



## Theory of Operation



## Security & Privacy Features

- **Anonymity:** Service providers cannot identify users within the registered user set; authentication remains unlinkable
- **Mis-authentication Resistance:** Ensures that only registered users can authenticate
- **Unlinkability of Tickets:** Individual tickets cannot be linked to a specific set of user transactions. Prevents tracking of user activities, thereby preserving privacy, anonymity
- **Revocability:** Users listed in the revocation window cannot be authenticated
- **Coalition Resistance:** Both revoked and unregistered users, individually or in groups, cannot authenticate
- **Identity-Escrow Freeness:** No TTP can infer a user's identity or pseudonym
- **Backward Unlinkability:** Past authentications remain anonymous and unlinkable even after user revoked from the system
- **Revocation Auditability:** Users can verify their revocation status, preventing malicious SPs from falsely recognizing users as revoked

## Proposed Scheme

A novel authentication protocol is proposed for dynamic EV charging using dynamic accumulators and zero knowledge proofs.

The Charging Service Provider (CSP) employs a dynamic accumulator to manage a blacklist of users involved in malicious activities.

**Registration phase:** Users receive essential system parameters.

Based on these parameters, users generate pseudonyms or tickets, which are unlinkable to their real identities.

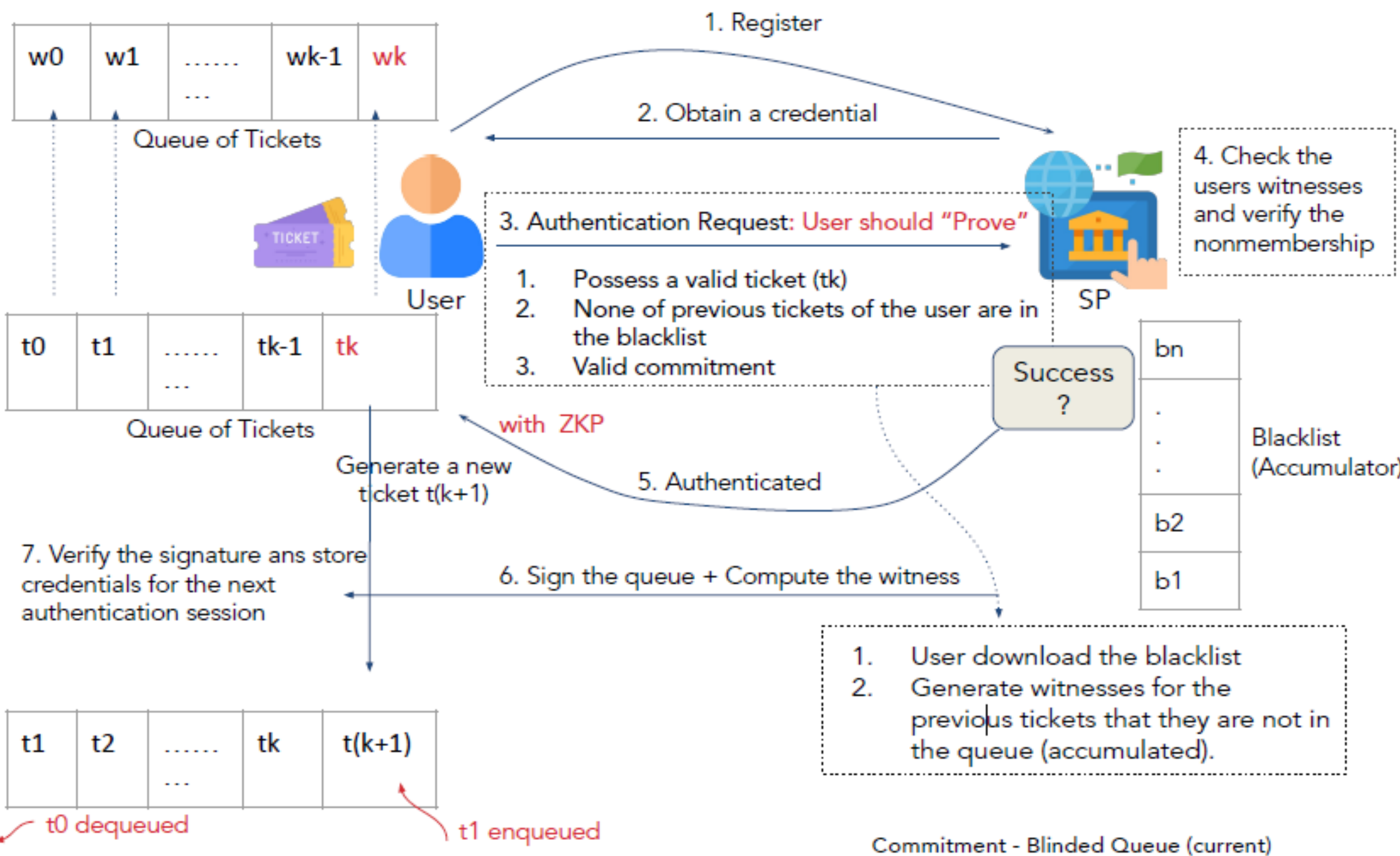
**Authentication phases:** Users submit a pseudonym. Additionally, they must prove, via zero-knowledge proofs, that their pseudonym is not included in the CSP's blacklist. Only users who can successfully provide this proof are authorized to access the charging service.

**Revocation of misbehaving users:** In the event of detected misbehavior or malicious activity, the CSP updates the blacklist by adding the pseudonym associated with the offending EV. This results in the user being denied future access to the system, as they will no longer be able to authenticate successfully.

**Pseudonym-based Authentication:** Pseudonyms ensure the user's real identity is not revealed. This prevents tracking or linking the user to their charging activities.

**Zero-Knowledge Proofs for Blacklist Verification:** Zero-knowledge proofs allow users to prove they are not blacklisted without sharing any personal information. This ensures both privacy and security while maintaining anonymity.

## System Model



## Cryptographic Primitive: Dynamic Accumulator

- A dynamic accumulator, or simply an accumulator, is a constant-size cryptographic construct that represents set membership.
- Elements may be added to (i.e., "accumulated"), or removed from, the accumulator.
- Anyone can prove in zero knowledge that certain element is "in" the accumulator if and only if the element has indeed been accumulated.

**In our work:**

users can authenticate by proving in zero knowledge that their pseudonym is in the accumulator, where the accumulator represents a "blacklist" of pseudonyms belongs to malicious users.

## Conclusion

This poster presents a framework that EVs to anonymously authenticate themselves to the untrustworthy charging service providers as well as the RSU and exchange information while protecting against malicious behavior by EVs. If a legitimate EV who gained access to the system successfully, converted to behaves maliciously, the framework can revoke the anonymity of the EV and reveal the real identity to cease future attacks. The proposed framework is designed to be robust against both malicious insider attacks and system-level threats.





# Fortifying Publish-Subscribe Communication: Advanced Security Solutions

Shabnam Saderi, Kalikinkar Mandal, Ali A. Ghorbani

{sh.saderi,kmandal,ghorbani}@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)



## Motivation

Publish-subscribe (pub-sub) systems, crucial for modern digital infrastructures, enable efficient, real-time communication across various sectors. Utilizing a decoupled communication model, these systems allow publishers to broadcast messages without specific recipients, supporting scalable and flexible data distribution. This is vital in sectors like smart grids and healthcare, where rapid dissemination of information is essential. Despite their benefits, pub-sub systems also pose significant security challenges.

- Security Vulnerabilities:** As pub-sub systems are increasingly integrated into critical operations, they become potential targets for sophisticated cyber threats. Vulnerabilities such as unauthorized access and message spoofing can disrupt operations and lead to substantial data breaches.
- Impact of Security Breaches:** The repercussions of security breaches extend beyond data loss, causing severe disruptions in services vital to public safety and economic stability. For example, breaches in financial or healthcare pub-sub systems can cause irreversible harm and erode public trust.

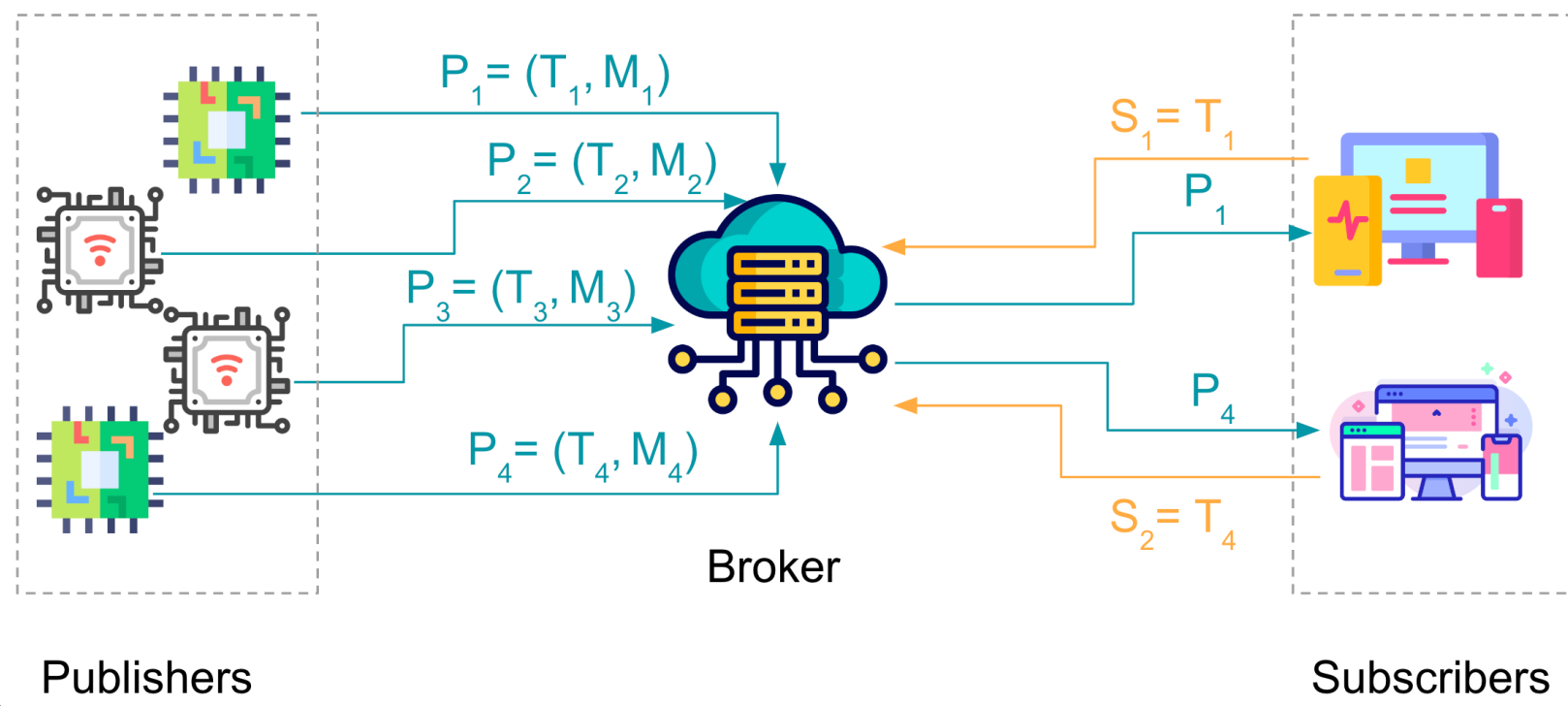
To ensure continuous service reliability and resilience against cyber-attacks, proactive security measures are essential:

- Proactive Measures:** Implementing early detection systems, conducting regular security audits, and investing in advanced security technologies are crucial steps toward mitigating risks.
- Security Training and Frameworks:** Engaging in continuous security training and adopting robust security frameworks can further strengthen the defenses of pub-sub systems, ensuring they can withstand evolving cyber threats.

## System Components

- Publishers:** Publishers send messages labeled with specific topics, initiating communication without knowing the subscribers. This allows for timely content delivery and topic adaptation based on current events.
- Subscribers:** Subscribers choose topics of interest and receive relevant messages, preventing information overload. They can either actively pull messages or passively receive them from the broker, enhancing system efficiency.
- Message Broker:** The message broker serves as the central node, managing the flow of messages between publishers and subscribers. It filters and routes messages by topic while overseeing queuing and delivery, ensuring reliable performance under varying loads.

## Communication Structure of Pub-Sub



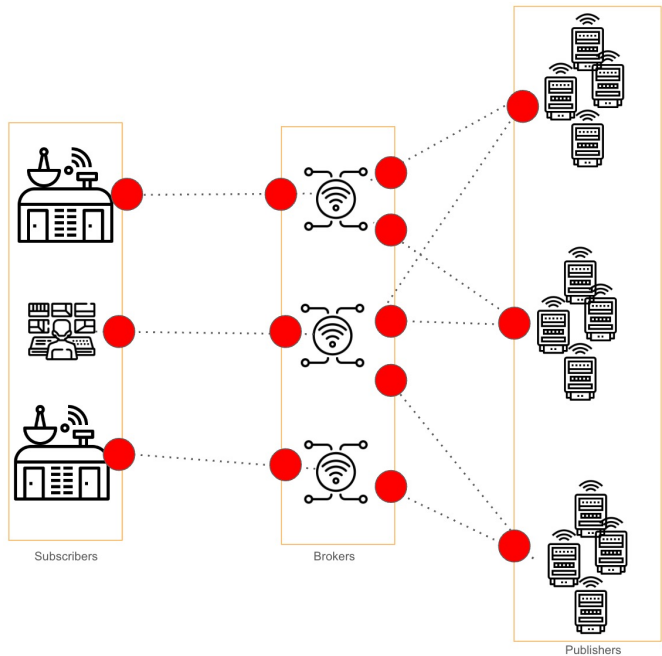
## Vulnerability Points in Pub-Sub

The diagram highlights key vulnerability points in pub-sub systems, showcasing security risks across publishers, brokers, and subscribers, each presenting unique threats to system:

- Publishers** are susceptible to data leaks if not properly secured.
- Brokers** face risks from MITM attacks, data tampering, and being a single point of failure.
- Subscribers** can introduce unauthorized access points and data integrity risks due to insecure endpoints.

### Importance of Addressing These Vulnerabilities:

- System-wide Impact:** A failure in any of these points can compromise the entire communication framework, leading to data breaches, misinformation, or operational failures.
- Mitigation Strategies:** Implementing end-to-end encryption, robust key management, decentralized broker models, and secure access control mechanisms can help reduce these risks.



## Common Vulnerability Exposures (CVEs) in Pub-Sub

This table illustrates examples of CVEs on various pub-sub protocols, highlighting typical vulnerabilities that affect system components:

CVE	Pub- Sub Protocol	Category	Affected Components
CVE-2024-27309	Kafka	Incorrect Authorization	Broker
CVE-2023-32315	XMPP	Unauthorized Access	Broker
CVE-2024-29195	AMQP	Data Corruption	Publishers/Subscribers
CVE-2018-1257	STOMP	DoS	Broker
CVE-2024-31486	MQTT	Confidentiality Loss	Publishers/Subscribers

## State-of-the-Art Approaches to Secure Publish-Subscribe Communication

- End-to-End Encryption** Ensures message integrity, confidentiality, and authenticity using protocols like TLS.  
**Challenges:** High resource demands and compatibility issues with legacy systems.
- Third-Party Involvement** Uses third-party services for key management, simplifying encryption processes.  
**Challenges:** Risk of third-party compromise and single point of failure.
- Attribute-Based Encryption (ABE):** Encrypts data based on subscriber-defined attributes, allowing fine-grained access control.  
**Challenges:** Complex key management and high resource demands.
- Homomorphic Encryption:** Allows computations on encrypted data, preserving privacy without decryption.  
**Challenges:** High computational overhead, limiting real-time use.
- Secure Multi-Party Computation (MPC):** Enables collaborative computation over private inputs, maintaining data privacy.  
**Challenges:** Communication overhead and complex implementation.
- Secure Enclaves:** Provides secure code execution within isolated hardware memory areas.  
**Challenges:** Limited compatibility and susceptibility to side-channel attacks.

## Future Directions

- Lightweight Cryptography:** Choosing efficient encryption for resource-constrained devices.
- Decentralized Brokers:** Use decentralized brokers instead of one to reduce single points of failure.
- Quantum-Resistant Encryption:** Prepare for future quantum threats with advanced cryptography.
- Privacy-Preserving Computation:** Enable secure data sharing with techniques like MPC.
- Scalable Key Management:** Improve key distribution methods for growing pub-sub.





## Introduction

The digital smart grid represents the convergence of information and operational technologies (IT/OT) with traditional electrical systems to enhance grid intelligence, resilience, and efficiency. It leverages IoT, big data analytics, and cybersecurity technologies to modernize electricity distribution.

### Core Components:

- Smart Meters: Utilize advanced metering infrastructure (AMI) for real-time monitoring and management of energy usage.
- SCADA Systems: Supervisory control and data acquisition (SCADA) systems for remote control and automation of electrical substations.
- Advanced Communication Networks: Deploy secure, high-bandwidth communication protocols (e.g., DNP3, IEC 61850) for grid data exchange.
- DERs and Microgrids: Integrate distributed energy resources (DERs) and microgrids using smart inverters and grid-edge technologies.
- EMS and DMS: Leverage energy management systems (EMS) and distribution management systems (DMS) for optimized grid operations and reliability.
- Cyber-Physical Security: Incorporate layered security architectures, combining physical security measures and cybersecurity technologies.

### Strategic Objectives:

- Operational Efficiency: Implement advanced analytics and automation to minimize losses and optimize energy flow.
- System Reliability: Enhance reliability through predictive maintenance, fault detection, and self-healing grid technologies.
- Sustainable Energy Integration: Seamlessly integrate renewable energy sources, supporting dynamic load balancing and energy storage solutions.

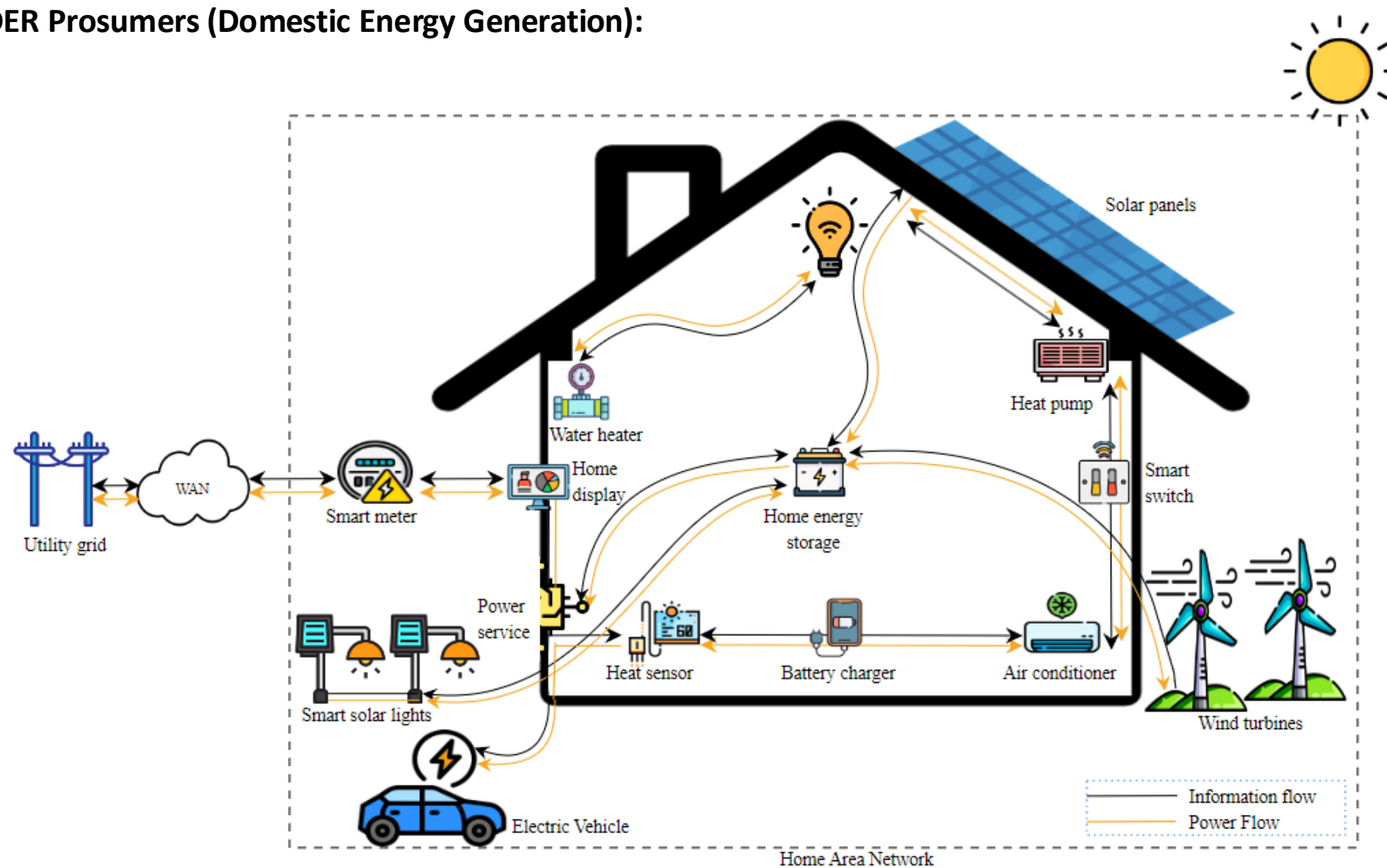
### Technical Challenges:

- Cybersecurity and Compliance: Address evolving cyber threats and regulatory compliance requirements for IT/OT convergence.
- Legacy System Integration: Overcome interoperability and integration challenges with aging grid infrastructure.
- Investment and ROI : Balance the capital investment in advanced grid technologies against expected operational efficiencies and return on investment (ROI).
- Data Privacy and Security: Ensure the integrity and confidentiality of consumer data and operational intelligence.

## Distributed Energy Resources and ICS Protocols

- DER in a smart grid range from domestic energy sources such as rooftop solar panels and individual wind turbines to large-scale systems like solar farms, offshore windmills, and expansive wind farms.
- These resources enable more efficient energy management, decentralized power generation, and enhanced grid resilience by integrating renewable energy into the grid.
- Industrial control systems (ICS) protocols are used for controlling and monitoring a range of industrial processes and systems such as distributed control systems (DCS), SCADA, and industrial automation systems (IAS).
- Widely used in industrial systems and critical infrastructures such as nuclear and thermal plants, water treatment facilities, oil extraction facilities and modern smart grids.
- Originally deployed physically isolated from external networks, with the main focus on real-time responses with extremely high availability and reliability. Thus, they lack inherent cybersecurity.

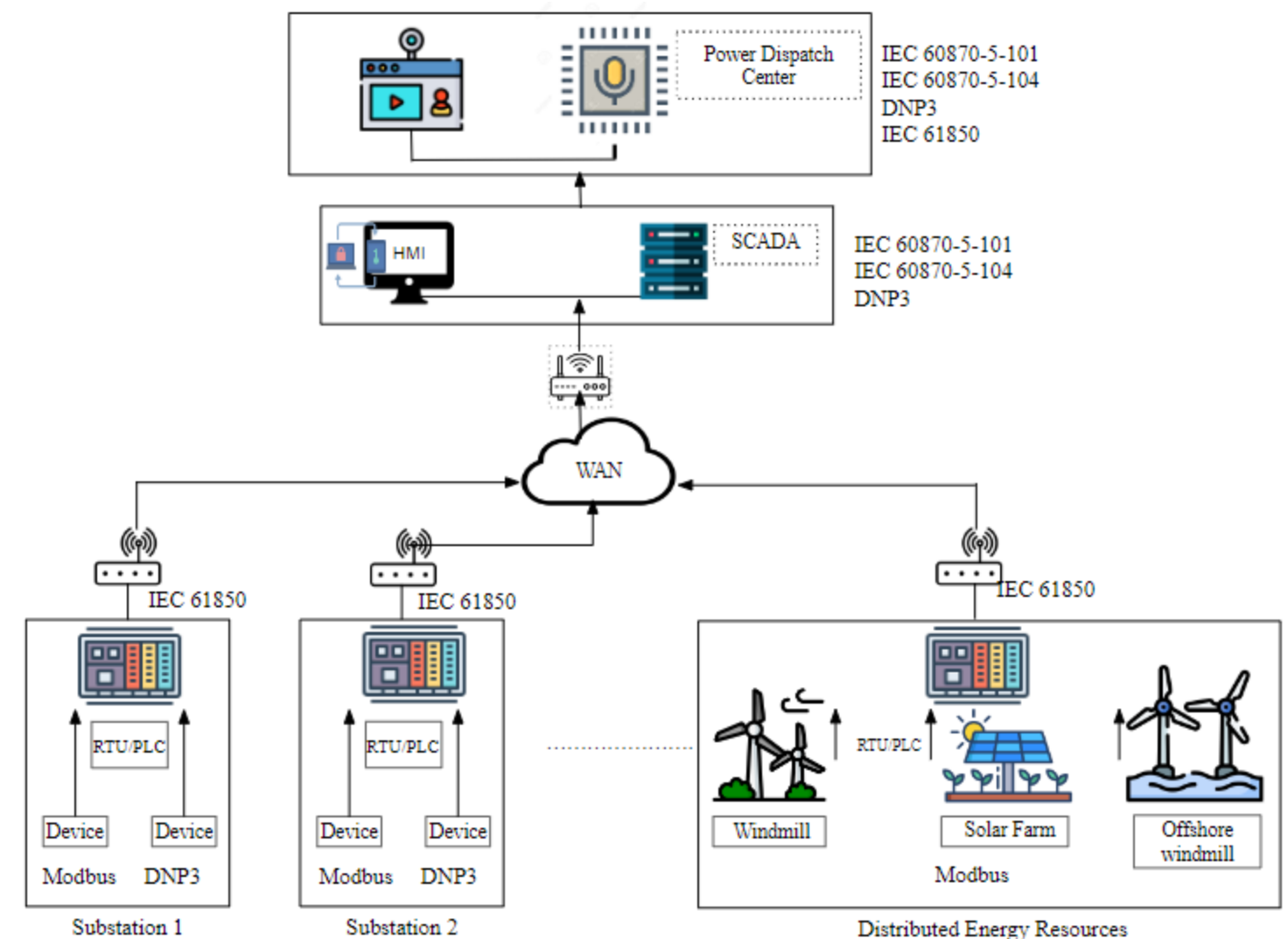
### DER Prosumers (Domestic Energy Generation):



## Threat Landscape for the Smart Grid

- Cyber Attacks:** Sophisticated attacks like malware and phishing threaten IT/OT systems. Incidents include Ukraine (2015/16) and WannaCry (2017).
- Physical Sabotage:** Infrastructure, such as substations, is vulnerable to direct attacks, exemplified by the Metcalf attack (2013).
- DoS/DDoS Attacks:** Target network and control systems to disrupt operations, with utility communications often attacked.
- Supply Chain Compromises:** Vulnerabilities in the supply chain can lead to system infiltration, as seen in the SolarWinds attack (2020).
- Quantum Computing Threats:** Emerging concerns over cryptographic standard vulnerabilities in the face of quantum decryption advancements

## DER and ICS Protocols in the Smart Grid



## ICS Protocol Level Vulnerabilities

The integration of Distributed Energy Resources (DERs) into the smart grid involves the use of several ICS communication protocols, each with unique cybersecurity considerations:

- IEEE 1815-DNP3 and Modbus are extensively utilized for process automation, yet they inherently lack encryption and authentication mechanisms, exposing them to data breaches.
- OpenADR and IEEE 2030.5 support more advanced security features but still require additional measures for comprehensive protection against sophisticated cyber threats.

### Detailed vulnerabilities in DNP3 and Modbus include:

- Data Interception and Modification: Alters commands.
- Fabrication Attacks: Generates counterfeit commands.
- Remote Exploitation: Allows unauthorized access.

## Proposed Scheme: Securing ICS Protocols using VPN

Our research focuses on enhancing the security of ICS communication by implementing VPN technology. We benchmarked Modbus, DNP3, and SNMP with and without WireGuard VPN. The below table represents the latency measurements in milliseconds:

Protocol	Time (ms)		Overhead
	no security	with security	
Modbus	0.971	2.1016	2.1x
DNP3	1.002	2.1711	2.1x
SNMP	0.964	2.1304	2.2x

Results signify that WireGuard VPN adds latency but tolerable overhead, indicating VPN's security benefits in ICS may outweigh performance costs.

## Challenges and Future Work

- Integrating DER systems via ICS protocols exposes critical infrastructures to the internet, increasing their susceptibility to cyber threats. As DER systems become more interconnected, the risk of unauthorized access or manipulation by malicious actors grows, creating potential threats to the reliability and safety of the entire grid.
- To safeguard against such vulnerabilities, robust security measures should be implemented, including IDS/IPS, encryption for secure communication, and strong access controls. These measures not only detect and prevent intrusions but also ensure that sensitive data is protected and only authorized personnel can access critical systems.
- Focus on developing and integrating advanced cybersecurity techniques to enhance smart grid security. Continuous innovation in cybersecurity strategies is essential to stay ahead of evolving threats and ensure that the smart grid remains resilient against potential cyberattacks.
- As preventive measures, developing real-time monitoring systems and robust incident response frameworks is crucial for identifying and mitigating cyber threats in DER systems. This includes employing machine learning algorithms to detect anomalous behavior and automating response strategies to contain potential breaches before they cause widespread disruption.

## Acknowledgement

This research work is supported by the NB Power Cybersecurity Research Chair Grant