CLUSTER-AWARE ANOMALY DETECTION FOR HETEROGENEOUS IOMT TRAFFIC

Visvakaanth KT
Department Of Computer Science
R.M.K Engineering College
Chennai, India
Visvakamal225566@gmail.com

Abstract— Hospitals contain numerous heterogeneous IoMT devices, each with distinct traffic patterns and usage behaviors. Most hospital networks rely on a single, global intrusion detection system (IDS) to monitor all devices, which may fail to detect malicious traffic masked by legitimate device signals. In this study, we demonstrate that clustering IoMT device traffic and training separate Isolation Forest (IF) models for each cluster improves anomaly detection performance. Using the CCIoMT2024 dataset, we simulated malicious traffic resembling an infusion pump and compared detection performance between a global IF and cluster-specific IF models. Results indicate that cluster-specific IFs detect anomalies faster and more reliably, identifying threats that the global model missed entirely in several cases. These findings highlight the importance of device-aware IDS deployment in heterogeneous IoMT environments for improved Cybersecurity.

Keywords— IoMT, intrusion detection, Isolation Forest, KMeans clustering, anomaly detection, device-specific IDS

I. INTRODUCTION

Hospitals increasingly rely on Internet of Medical Things (IoMT) devices to monitor patients, administer treatments, and manage hospital operations. These devices exhibit highly heterogeneous behavior—some transmit data continuously, while others are idle for periods before generating bursts of traffic. Despite this diversity, most hospital networks deploy a single, global intrusion detection system (IDS) to monitor all device traffic, primarily for efficiency and resource constraints.

However, a global IDS may fail to detect malicious traffic from specific devices because anomalous patterns can be masked by normal traffic from other heterogeneous devices. For instance, a malicious data flow from an infusion pump may be interpreted as normal by a global IDS if other devices' traffic exhibits similar patterns or variability. This limitation poses significant Cybersecurity risks in hospital environments, where timely detection of threats is critical.

In this study, we investigate whether clustering IoMT devices with similar traffic patterns and training cluster-specific Isolation Forest (IF) models can improve anomaly detection performance. Using the CCIoMT2024 dataset, we simulate malicious traffic from infusion-pump-like devices and compare detection performance between a global IF trained on all devices and cluster-specific IFs trained on traffic from similar devices.

Our results demonstrate that cluster-specific IF models detect anomalies faster and more reliably than a global model, highlighting the importance of device-aware IDS deployment. The contributions of this work are summarized as follows:

- This work demonstrates the limitations of a global IDS in heterogeneous IoMT environments.
- A cluster-based IDS approach using KMeans clustering and Isolation Forests is proposed.
- An empirical evaluation shows that cluster-specific models detect malicious traffic faster and more reliably than a global IDS.

II. RELATED WORK

Traditional hospital IDS systems often monitor all devices with a single global model. While computationally efficient, such models may fail to detect device-specific anomalies. Clustering techniques can help group similar device traffic, allowing for more targeted anomaly detection and faster identification of malicious flows. However, few studies provide empirical evidence comparing global versus cluster-specific IDS performance on real-world IoMT datasets. This work addresses this gap by applying clustering and Isolation Forests to the CCIoMT2024 dataset, demonstrating that cluster-specific models can detect malicious traffic more reliably than a global IDS.

III. METHODOLOGY

A.Dataset Description

The experiments in this work utilize the publicly available CCIoMT2024 dataset [1], which contains traffic data from multiple Internet of Medical Things (IoMT) devices commonly found in hospital environments. Each network flow in the dataset corresponds to a specific device and captures features such as packet timing, size, and protocol type. The dataset is heterogeneous, with devices exhibiting diverse traffic patterns—some continuously active, while others remain idle for extended periods before transmitting bursts of data. This diversity reflects real-world hospital networks, where a single global intrusion detection system (IDS) must monitor all device traffic simultaneously.

B. Malicious Traffic Simulation

To evaluate the effectiveness of global versus cluster-specific IDS models, selected rows from the dataset were modified to represent malicious traffic flows. The altered traffic was designed to closely resemble that of an infusion pump device, ensuring that detection would be challenging in the presence of heterogeneous device traffic. These modified flows serve as controlled anomalies for testing detection performance.

C. *Global IDS Using Isolation Forest*

A single Isolation Forest (IF) model was trained on the entire dataset, representing a conventional global IDS approach. This model monitors all device traffic collectively, without distinguishing between device types. Isolation Forest is an unsupervised anomaly detection algorithm that identifies anomalies by isolating data points through random partitioning of the feature space [2]. The simulated malicious flows were input into the global IF, and predictions and anomaly scores were recorded over multiple windows to assess detection effectiveness.

D. Clustering of Device Traffic

To account for the heterogeneous nature of device traffic, KMeans clustering was applied to the dataset to group similar flows. Each cluster represents a subset of devices with comparable traffic characteristics. Separate CSV files were created for each cluster, facilitating the training of device-specific IDS models. The cluster most closely resembling infusion pump traffic was selected for further analysis.

E. *Cluster-Specific IDS Using Isolation Forest*

A second Isolation Forest model was trained exclusively on the selected cluster, representing a cluster-specific IDS. The same malicious flows used in the global model experiments were tested on this cluster-specific IF. This approach isolates device-specific patterns and enables more sensitive detection of anomalies that may be obscured in global traffic

F. *Comparison of Global vs Cluster IDS*

Detection performance was evaluated by comparing the predictions and anomaly scores of the global and cluster-specific IF models. Key metrics include the window number at which a flow is first identified as anomalous and the magnitude of the anomaly scores. Results indicate that cluster-specific IDS models detect malicious flows more rapidly and reliably than the global model, particularly for traffic resembling infusion pump devices.

IV. RESULTS

A. Detection Performance of Global IDS

The global Isolation Forest (IF) model was trained on the entire CCIoMT2024 dataset. Selected malicious flows resembling infusion pump traffic were tested on this model. Table I summarizes the detection performance for representative flows. A value of "Detected" indicates that the anomaly score exceeded the detection threshold in that window, while "Missed" indicates the flow was not identified as malicious within the observation period.

Table I – Global IDS Detection Results

Flow (iloc)	Windows to Detection	Detected?
5	6	Yes
18	–	No
35	6	Yes
57	3	Yes
74	1	Yes
92	–	No
110	5	Yes
145	2	Yes
180	–	No
215	5	Yes

B. Detection Performance of Cluster-Specific IDS

The cluster-specific IF model was trained only on the subset of data representing infusion pump-like traffic. The same malicious flows were tested on this model. Table II summarizes the results

Table II – Cluster Specific IDS Detection

Flow (iloc)	Windows to Detection	Detected?
5	3	Yes
18	1	Yes
35	2	Yes
57	2	Yes
74	2	Yes
92	1	Yes
110	1	Yes
145	2	Yes
180	1	Yes
215	2	Yes

C. Comparative Analysis

The results indicate that the cluster-specific IDS consistently detects malicious flows faster than the global IDS. Notably, flows iloc 18, 92, and 180 were completely missed by the global IDS but were detected by the cluster-specific model. In cases such as iloc 74 and 145, detection by the cluster-specific model was only one window slower than the global model. In addition, cluster-specific models generally identified anomalies within one to two detection windows, whereas the global IDS often required multiple windows or failed to detect the malicious flows entirely.

Overall, the cluster-specific approach demonstrates enhanced sensitivity and reliability in identifying device-specific anomalies in heterogeneous IoMT traffic.

V. DISCUSSION

The results demonstrate that a single global intrusion detection system (IDS) trained on heterogeneous IoMT traffic can fail to reliably detect device-specific malicious behavior. Because hospital networks contain multiple IoMT devices with fundamentally different traffic patterns, anomalous behavior from one device type may be masked by the normal behavior of others when analyzed collectively.

In contrast, the cluster-specific IDS approach improves detection by isolating traffic with similar characteristics before applying anomaly detection. By training an Isolation Forest on a cluster representing infusion pump–like traffic, the model becomes more sensitive to deviations that would otherwise appear normal in a global context. This explains why the cluster-specific IDS detected malicious flows earlier and, in some cases, detected attacks that were entirely missed by the global IDS.

From a practical perspective, these findings suggest that resource-efficient hospital security architectures can still benefit from device-aware or cluster-aware IDS designs. Instead of deploying a separate IDS for every device, clustering similar IoMT devices provides a compromise between scalability and detection accuracy.

However, this study has limitations. The malicious traffic was synthetically generated by modifying existing flows, which may not fully capture the complexity of real-world attacks. Additionally, the number of clusters and the choice of clustering algorithm may influence detection performance. Future work could explore adaptive clustering, online learning, or supervised approaches when labeled attack data is available.

VI. CONCLUSION

This paper investigated the effectiveness of global versus cluster-specific intrusion detection systems for heterogeneous IoMT traffic in hospital environments. Using the CCIoMT2024 dataset, malicious traffic resembling infusion pump behavior was evaluated using Isolation Forest models trained on both global and clustered data.

Experimental results show that the cluster-specific IDS consistently detected malicious traffic faster and more reliably than the global IDS. In several cases, the global model failed to detect malicious flows entirely, while the cluster-based model identified them almost immediately. These findings indicate that homogeneous or cluster-aware IDS designs are better suited for securing complex IoMT environments. The comparison was conducted under identical experimental conditions to isolate the effect of clustering on detection performance.

Overall, this work highlights the importance of considering device heterogeneity in hospital network security and demonstrates that unsupervised clustering combined with anomaly detection can significantly improve threat detection without excessive resource overhead

REFERENCES

[1] C. C. IoMT Research Group, “CCIoMT2024: A Comprehensive Dataset for Internet of Medical Things Traffic Analysis,” 2024. [Online]. Available: https://www.unb.ca/cic/datasets/cc-iomt-2024.html

[2] F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation Forest,” in Proc. IEEE Int. Conf. on Data Mining (ICDM), 2008, pp. 413–422.