Iot Profiling With Transductive Transfer Learning

Ms M Vineela, Madishetty Nandini, Ranga Akshitha
[1]Associate Professor, Department Of Cse, Bhoj Reddy Engineering College For Women, India.
[2,3]B. Tech Students, Department Of Cse, Bhoj Reddy Engineering College For Women, India.

ABSTRACT

The rapid proliferation of Internet of Things (IoT) devices across various domains has introduced significant challenges in device management and security profiling. Traditional machine learning approaches to IoT profiling often rely on large amounts of labeled data, which are difficult to obtain in dynamic and heterogeneous IoT environments. This paper proposes a novel framework for IoT device profiling using ***Transductive Transfer Learning (TTL)****, a technique that enables knowledge transfer from a labeled source domain to an unlabeled target domain. The proposed system effectively classifies IoT devices and assesses their vulnerabilities by leveraging behavioral features extracted from network traffic data.*

The methodology incorporates statistical feature selection techniques and evaluates the performance of multiple machine learning models, including Random Forest, Gradient Boosting, and Support Vector Machines. To validate the transferability and robustness of the approach, extensive experiments were conducted using diverse datasets such as CIC IoT 2022, IMC 2019, and IoT Sentinel. The results demonstrate high classification accuracy and reliable vulnerability assessment across varying environments. This work contributes to advancing secure and scalable IoT network management by reducing the dependency on labeled data and enabling real-time device identification and risk evaluation.

Keywords*: Internet of Things (IoT), Device Profiling, Transductive Transfer Learning, Domain Adaptation, Feature Extraction, Vulnerability Assessment, Machine Learning, CIC IoT Dataset, IoT Security, Random Forest Classifier*

1-INTRODUCTION

The rapid evolution of the Internet of Things (IoT) has transformed numerous industries, enabling seamless connectivity and communication across a wide array of devices. With the proliferation of IoT devices, effective profiling has become crucial for managing networks and ensuring robust security. IoT profiling involves identifying and categorizing devices based on their behavior and characteristics, which aids in monitoring network activity and optimizing resource allocation.

However, profiling IoT devices poses unique challenges due to the diversity of device types, the dynamic nature of IoT environments, and the limited availability of labeled data for training machine learning models. In this context, transductive transfer learning emerges as a promising solution, leveraging knowledge from related domains to improve the profiling process. Unlike traditional methods that rely on extensive labeled data from the target domain, transductive transfer learning adapts learned knowledge to the target domain, thereby enhancing model performance even when labeled data is limited.

The motivation for this paper is driven by the need for more effective and efficient IoT profiling techniques capable of handling the increasing complexity and scale of IoT networks. Current methods often face challenges with accuracy and resource consumption, underscoring the demand for

innovative approaches. This paper aims to develop a transductive transfer learning framework specifically designed for IoT device profiling, enhancing both identification accuracy and profiling efficiency.

This paper focuses on profiling a diverse range of IoT devices and evaluating the proposed framework's performance across multiple datasets. It is essential to recognize the limitations of this study, particularly concerning the generalizability of the findings to all IoT environments.

2.LITERATURE SURVEY

The identification and profiling of IoT devices has become a critical area of research due to the exponential growth in connected devices and the corresponding security threats. Several studies have explored machine learning and transfer learning techniques to address the challenges of device heterogeneity and lack of labeled data in IoT environments.

Danso et al. (2024) proposed a framework leveraging the transferability of machine learning algorithms to profile IoT devices. Their study highlighted the effectiveness of using domain adaptation techniques to maintain classification accuracy across varied datasets such as CIC IoT, IMC, and IoT Sentinel. The approach emphasized reducing the dependency on labeled data by training models on one domain and applying them to another using Transductive Transfer Learning (TTL), resulting in reliable profiling across networks.

Fan et al. (2020) introduced an IoT device identification method based on semi-supervised learning. Their research focused on situations where labeled data is sparse, which is common in IoT scenarios. By leveraging behavioral features and a limited set of labeled instances, the system demonstrated the potential of machine learning in real-world deployment, though challenges remained in cross-domain adaptability.

Almomani and Rahman (2022) conducted a comprehensive literature review on IoT adoption and highlighted the importance of lightweight, adaptive profiling solutions. They emphasized the need for secure, scalable frameworks that can operate in resource-constrained environments and adapt to emerging threats, which aligns with the rationale behind applying TTL to IoT security.

Existing studies also reveal that traditional supervised learning models often struggle with domain shifts, where training and test environments differ significantly. Researchers have therefore turned to transfer learning, particularly TTL, to improve model robustness. However, many earlier efforts lacked extensive empirical validation across multiple datasets and did not address vulnerability assessment as an integral part of device profiling.

The current work builds on these foundations by integrating statistical feature selection, cross-domain evaluation, and security vulnerability analysis using real-time datasets. The proposed methodology enhances profiling accuracy and device-level risk assessment without the need for extensive labeled data in the target domain.

3. BACKGROUND

IOT DEVICES AND THEIR CHARACTERISTICS

The **Internet of Things (IoT)** has become a key topic in technology, representing a growing ecosystem of connected devices. IoT can be defined as a network of interconnected physical devices—such as vehicles, home appliances, and industrial systems—that are embedded with electronics, sensors, and network connectivity. These devices can collect and exchange data, enabling them to communicate with each other and interact with the

environment, ultimately creating smart, responsive systems.

Key Characteristics of IoT Devices:

1. **Connectivity**: IoT devices are designed to be connected anytime and anywhere, allowing remote interaction and control, such as adjusting a smart thermostat from a smartphone.
2. **Intelligence and Unique Identification**: Each IoT device has a unique identity and can process data to make decisions autonomously. For instance, a smart sensor can adjust HVAC systems based on real-time temperature readings.
3. **Self-Configuration**: IoT devices often have self-configuring capabilities, enabling easy integration into existing networks with minimal user intervention during setup.
4. **Interoperability**: Standardized protocols allow different IoT devices to communicate seamlessly, enabling cross-device functionality and fostering innovation by reducing data silos.
5. **Scalability**: IoT systems are built to accommodate a growing number of devices and data, maintaining efficiency and performance as the network expands.
6. **Embedded Sensors and Actuators**: Sensors detect environmental changes, while actuators perform actions based on this data, enabling automation, such as lighting control based on occupancy.
7. **Data-Driven Operation**: IoT systems gather and analyze large amounts of data, which is used to enhance efficiency and inform decisions. This data-centric approach allows continuous improvement and adaptation.
8. **Security Concerns**: Due to inherent resource constraints (such as limited power and memory), IoT devices are often more vulnerable than traditional computing devices. These limitations make it challenging to implement robust security protocols directly on the devices, increasing their susceptibility to cyber threats.
9. **Context Awareness**: IoT systems are often context-aware, meaning they can adjust their actions based on the surrounding environment, providing tailored and efficient user experiences.

TRANSFER LEARNING IN MACHINE LEARNING

Transfer learning is a machine learning approach where a model developed for one task is applied to a related, yet distinct, task or dataset. This technique leverages the knowledge a model has gained in one domain (known as the *source domain*) to enhance its performance in another, often related domain (the *target domain*). Transfer learning is particularly beneficial in scenarios where labeled data is limited or costly to obtain, allowing models to generalize well in the target domain by building upon the insights acquired from a well-labeled source domain.

Traditionally, machine learning models are trained from scratch for each new task, assuming that the training and test data originate from the same feature space and data distribution. However, in real-world applications, data distributions often vary, and starting anew for every task can be inefficient. Transfer learning addresses this by reusing and adapting pre-trained models, which helps improve model generalizability and efficiency without needing large volumes of labeled data for each task.

4-METHODOLOGY

This section outlines the methodology used for IoT profiling using transductive transfer learning, detailing the data collection, feature extraction, and transductive learning framework.

DATA COLLECTION

Data collection is the foundation of effective IoT device profiling. In this methodology, data is collected from various IoT devices across multiple environments or networks (e.g., different labs). This

includes obtaining network traffic data and metadata from IoT devices, such as packet transmission logs, device identifiers, and communication patterns. Key datasets, like the CIC IoT 2022 dataset, are used as the primary source, while other datasets (e.g., IMC 2019 and IoT Sentinel) serve as target domains to evaluate the adaptability and robustness of the model across different distributions. The data collected must encompass a diverse set of device activities to accurately reflect real-world usage and vulnerabilities.

FEATURE EXTRACTION

Feature extraction focuses on selecting and transforming data variables that help distinguish one IoT device from another. After data collection, relevant features are derived from the raw network data. Common features include:

- **Packet counts, DNS requests, and IP addresses**: Indicate communication frequency and destinations.
- **Network protocol types and response times**: Highlight device communication behaviors.
- **Power usage and operational states**: Indicate device activity and potential vulnerabilities.

These features are carefully chosen to maximize the classifier's accuracy while minimizing computational costs. The ANOVA statistical technique is often applied to select the most informative features, focusing on those that best differentiate device types.

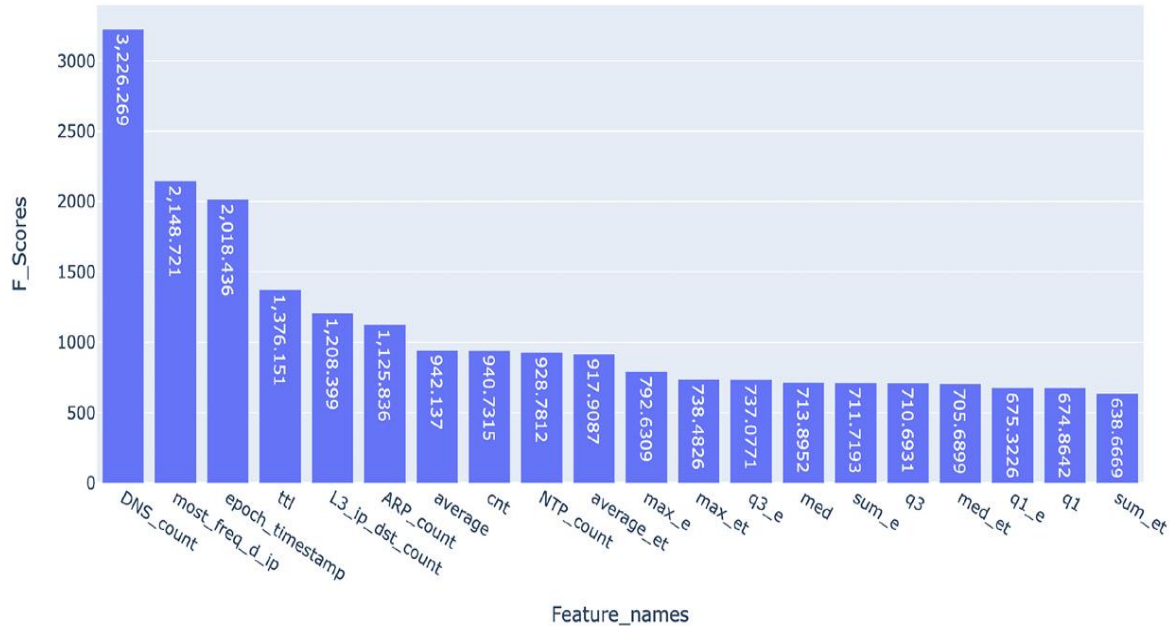

Fig. 4.1 Top 20 features extracted using ANOVA technique and their respective score on the CIC data set

TRANSDUCTIVE LEARNING FRAMEWORK

The transductive learning framework utilizes labeled source domain data to perform predictions in an unlabeled target domain, bridging domain differences without retraining the model for each new dataset.

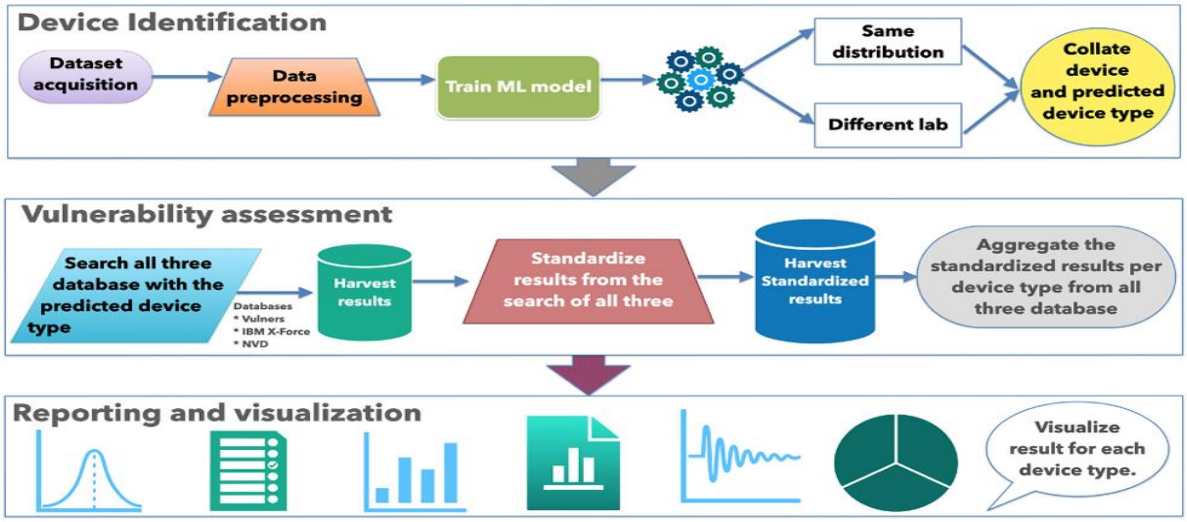

Fig. 4.2 Overall system framework of the proposed system

5-

ALGORITHM

The algorithm follows a structured approach to identify IoT device types across different domains. It leverages transductive transfer learning to apply a model trained on one dataset (source domain) to a different, but related, dataset (target domain) without requiring additional labeled data in the target domain. The key objectives of the algorithm are:

- To accurately classify device types in the target domain based on the source domain model.
- To assess vulnerabilities of identified devices by consulting security databases.
- To visualize the device types and associated security risks.

The algorithm can be divided into three main phases:

1. **Feature Extraction and Model Training in the Source Domain**: Prepares the data, selects relevant features, and trains the model on labeled source data.
2. **Transductive Transfer and Testing in the Target Domain**: Applies the trained model to make predictions in the target domain, where data lacks labels.
3. **Vulnerability Assessment and Visualization**: Maps identified device types to potential vulnerabilities and visualizes results on a dashboard for further analysis.

DETAILED ALGORITHM STEPS

This section breaks down the algorithm into step-by-step actions to achieve the profiling and vulnerability assessment objectives. Below are the detailed steps:

1. **Data Collection and Preprocessing**
   - **Source Domain Data Preparation**: Collect and preprocess data from the source domain (e.g., CIC IoT 2022 dataset). Normalize the data, handle missing values, and prepare it for feature extraction.
   - **Target Domain Data Preparation**: Collect data from the target domain (e.g., IMC 2019 Payload or IoT Sentinel dataset) for testing, ensuring similar preprocessing steps as applied to the source data.
2. **Feature Extraction**

- Use statistical methods, such as ANOVA, to identify and select the most relevant features that effectively differentiate device types.
- Transform the raw data into a feature vector comprising important features like packet counts, DNS queries, and protocol types.

3. **Model Training in Source Domain**

- **Model Selection and Initialization**: Choose an appropriate machine learning model based on the source data characteristics (e.g., Random Forest, SVM).
- **Training and Cross-Validation**: Split the source data into training and validation sets. Train the model on the source data, using cross-validation to avoid overfitting and to ensure generalizability.
- **Feature Selection Tuning**: Fine-tune feature selection using methods like SelectKBest to ensure only the most significant features are used.

4. **Transductive Transfer to Target Domain**

- **Model Application**: Apply the trained model to the target domain data for classification. As the target domain data is unlabeled, predictions are made directly.
- **Inference Generation**: For each device type prediction in the target domain, calculate the inference percentage (IP) to evaluate the model's confidence level for each device type. This helps confirm the reliability of predictions in the target domain.

5. **Vulnerability Assessment**

- **Keyword Search**: Based on the predicted device type, conduct a keyword search in vulnerability databases (e.g., NVD, Vulners, IBM X-Force) to assess associated security risks.
- **Data Harvesting**: Gather search results and store them in a local database for further analysis.
- **Standardization and Aggregation**: Standardize vulnerability data by matching results across different databases and aggregate these by device type, creating a unified view of vulnerabilities.

6. **Visualization and Reporting**

- **Dashboard Integration**: Visualize device profiles and vulnerabilities on a dashboard. Use visualizations such as sunburst charts and bar graphs to represent device categories and corresponding vulnerabilities.
- **Final Output**: Generate a report summarizing device types, profiling results, and vulnerability assessments for easy interpretation.

The algorithm provides a structured approach to identifying, assessing, and visualizing IoT devices across domains, ensuring reliable transferability and comprehensive security evaluation.

6. EXPERIMENTS AND RESULTS

DATASET DESCRIPTION

In this study, three key datasets are utilised: **CIC IoT**, **IMC 2019**, and **IoT Sentinel**.

- **CIC IoT Dataset**: Developed by the Canadian Institute for Cybersecurity, this dataset contains a wide variety of IoT device traffic data, which includes benign and malicious traffic. It is significant for transfer learning as it serves as a rich source domain, helping to understand the behavior of IoT devices under different network conditions.
- **IMC 2019 Dataset**: Collected during the Internet Measurement Conference (IMC) 2019, this dataset consists of network traffic data from IoT devices in various scenarios. Its significance lies in providing a diverse target domain that allows for effective transfer learning applications, reflecting real-world usage patterns.

- **IoT Sentinel Dataset**: This dataset is designed to evaluate the security of IoT devices and applications. It captures traffic patterns from IoT devices under attack scenarios, making it a critical resource for assessing the transferability of models trained on benign traffic to those under attack.

Table 6.1 Devices in the different labs and the corresponding Device Type

	CIC IoT dataset 2022	
No.	**Device Name**	**Device Type**
01.	Amazon Echo Dot	Audio
02.	Amazon Echo Spot	Audio
03.	Amazon Echo Studio	Audio
04.	Google Nest Mini	Audio
05.	Sonos One	Audio
06.	Amcrest Camera	Camera
07.	ArloQ Camera	Camera
08.	Borun Camera	Camera
09.	DLink Camera	Camera
10.	HeimVision Camera	Camera
11.	HomeEye Camera	Camera
12.	Luohe Camera	Camera
13.	Nest Camera	Camera
14.	Netatmo Camera	Camera
15.	SimCam	Camera
16.	Arlo Base Station Camera	Camera
17.	Amazon Plug	Home Automation
18.	Globe Lamp	Home Automation
19.	Gosund Plug	Home Automation
20.	Heimvision Lamp	Home Automation
21.	Teckin Plug	Home Automation
22.	Yutron Plug	Home Automation
23.	D-Link Water sensor	Home Automation
24.	Philips Hue	Smart Hub
25.	Ring Basestation	Smart Hub
26.	Eufy Homebase	Smart Hub
27.	Atomi coffee maker	Appliance
28.	iRobot roomba	Appliance
29.	Smart board	Appliance

	IMC 2019 Payload dataset	
No.	**Device name**	**Device type**
01.	Amcrest Cam	Camera
02.	Blink Cam	Camera
03.	Lefun Cam	Camera
04.	Luohe Cam	Camera
05.	Microseven Cam	Camera
06.	Wansview Cam	Camera
07.	Yi Cam	Camera
08.	Sengled hub	Smart Hub
09.	Smartthings	Smart Hub
10.	Blink Hub	Smart Hub
11.	Insteon hub	Smart Hub
12.	Wink 2 hub	Smart Hub
13.	Philips Hue	Smart Hub
14.	Flux Bulb	Home Automation
15.	TP-Link Bulb	Home Automation
16.	Wemo Plug	Home Automation
17.	Amazon Echo Dot	Audio
18.	Amazon Echo Spot	Audio
19.	Amazon Echo Plus	Audio
20.	Google Home Mini	Audio
21.	Behmor Brewer	Appliance
22.	Samsung Washer	Appliance
23.	Xiaomi Cleaner	Appliance
24.	Xiaomi Rice Cooker	Appliance

	IoT Sentinel	
No.	**Device name**	**Device type**
01.	D-Link Cam	Camera
02.	D-Link Day Cam	Camera
03.	EdimaxCam 1	Camera
04.	EdimaxCam 2	Camera
05.	EdnetCam 1	Camera
06.	EdnetCam 2	Camera
07.	D-Link Switch	Smart Hub
08.	Hue Switch	Smart Hub
09.	Ednet Gateway	Smart Hub
10.	Max Gateway	Smart Hub
11.	WeMo Insight Switch 1	Smart Hub
12.	WeMo Insight Switch 2	Smart Hub
13.	WeMo Switch 1	Smart Hub
14.	WeMo Switch 2	Smart Hub
15.	Lightify	Smart Hub
16.	HomeMatic plug Switch	Smart Hub
17.	D-Link Home hub	Smart Hub
18	TP-LinkPlugHS210	Home Automation
19	TP-LinkPlugHS110	Home Automation
20	Edimax Plug 1	Home Automation
21	Edimax Plug 2	Home Automation
22	Smarter coffee	Appliances

Table 6.2 Number of records for each device type in the CIC data set after data preprocessing and the number of sampled records used for training the ML model

No.	Device Type	Total no. of records
1	Camera	191812
2	Audio	19580
3	SmartHub	12711
4	HomeAutomation	8593
5	Appliances	1635

No.	Device Type	Sampled no. of records
1	Camera	1635
2	Audio	1635
3	SmartHub	1635
4	HomeAutomation	1635
5	Appliances	1635

6.2 E

VALUATION METRICS

To evaluate the performance of the machine learning models, we employed several key metrics:

- **Accuracy**: The ratio of correctly predicted instances to the total instances. It provides a general idea of the model's performance but may not reflect its effectiveness in imbalanced datasets.

$$ACCURACY = \frac{\text{TruePositive+TrueNegative}}{True\ Positive+TrueNegaive+FalsePositive+FalseNegative}$$

- **Precision**: The ratio of true positive predictions to the total positive predictions made by the model. This metric is crucial for understanding the model's ability to minimize false positives.

$$PRECISION = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}$$

- **Recall**: The ratio of true positive predictions to the actual positives in the dataset. It helps assess the model's ability to capture all relevant instances.

$$RECALL = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}$$

- **F1-Score**: The harmonic mean of precision and recall, providing a balance between the two metrics. It is particularly useful when dealing with imbalanced datasets.

$$F1 - SCORE = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Inference Percentage**: This metric assesses how well the model can apply learned knowledge from the source domain to make predictions in the target domain, thus evaluating the transferability of the learning.

These metrics enable a comprehensive assessment of model performance, highlighting strengths and weaknesses in classification tasks.

Table 5.3 Performance of the several classifiers on the CIC IoT data set 2022 acting as the source domain and used for training

Metric	RF	DT	GBC	KNN	SVM	MLP
Accuracy	99.87	99.46	**100**	95.18	98.50	98.52
Precision	99.62	99.48	**99.92**	95.47	96.10	98.53
Recall	99.63	99.49	**99.93**	95.23	95.69	98.58
F1-score	99.63	99.48	**99.93**	95.28	95.81	98.55
Time(sec)	1.20	**0.06**	65.97	0.98	2.20	9.33

PERFORMANCE ANALYSIS

The performance analysis of various machine learning algorithms showed distinct results across both the source and target domains. Key findings include:

- **Gradient Boosting Classifier**: Demonstrated the highest accuracy and F1-score in both the CIC IoT and IoT Sentinel datasets, indicating strong performance in distinguishing between benign and malicious traffic.
- **Multi-Layer Perceptron (MLP)**: Achieved competitive results, particularly in recall, suggesting its effectiveness in identifying all relevant instances in the target domain.
- **Support Vector Classifier (SVC)**: While not the top performer, it provided reasonable results and may be beneficial in scenarios requiring clear decision boundaries.

Overall, these results indicate that transductive transfer learning enhances model performance, particularly when adapting to new environments characterized by different traffic patterns.

VULNERABILITY ASSESSMENT

Vulnerability assessment is a systematic process used to identify, evaluate, and prioritize vulnerabilities in a system, network, or application. In the context of IoT devices, this assessment is crucial for understanding potential security risks that could be exploited by attackers. It enables organizations to proactively address vulnerabilities, ensuring the integrity, confidentiality, and availability of their IoT systems.

The vulnerability assessment conducted in this study aims to analyze potential security risks associated with the identified IoT device types through a systematic four-step process. This assessment ensures that the identified vulnerabilities are comprehensively documented, facilitating better security measures for IoT devices.

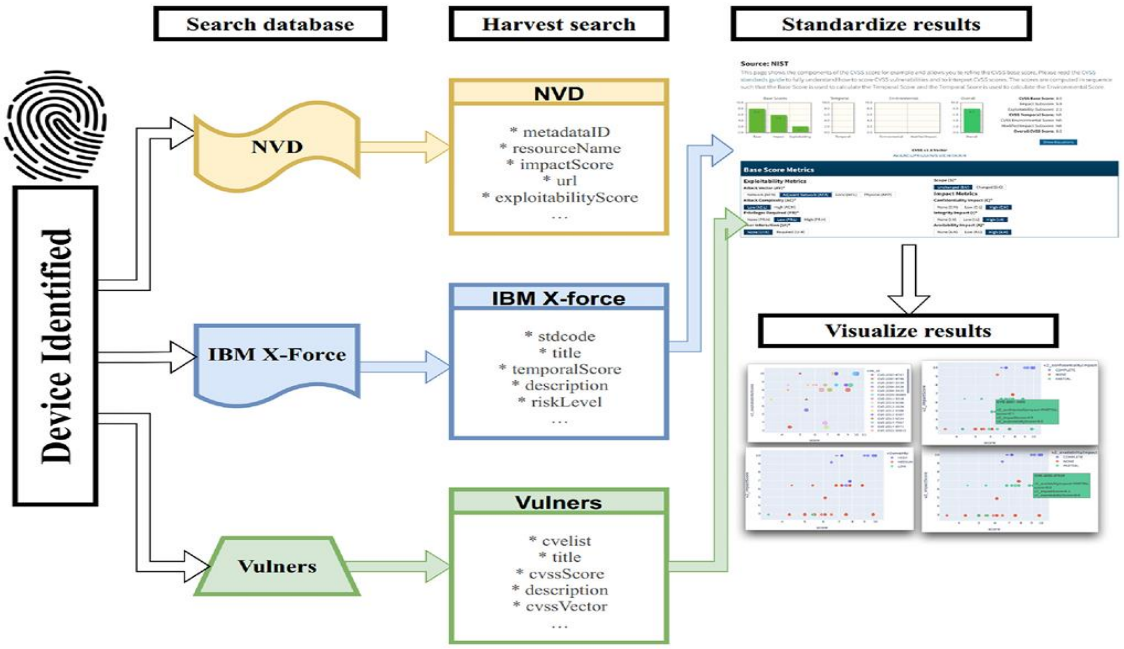

Fig. Vulnerability assessment overview

1.

Vulnerability Keyword Search

Following the identification of device types by the machine learning classifier, each predicted device type is used as a keyword to search three prominent vulnerability databases: the National Vulnerability Database (NVD), Vulners, and IBM X-Force. This step involves:

- Performing API calls to the IBM X-Force and Vulners databases using the predicted device types as keywords.
- For the NVD, the data stream is scraped annually, allowing for real-time updates of vulnerability information.
- A cron job is scheduled to synchronize the local database with the NVD website, ensuring that our dataset remains current with the latest vulnerabilities reported.

The output from these searches is stored in respective local databases for further processing.

2. Vulnerability Harvesting

Each of the three vulnerability databases is queried for vulnerabilities related to five predicted device types (audio, appliance, camera, home automation, and smart hub). This results in a comprehensive analysis consisting of 15 iterations (3 databases × 5 device types). Key steps include:

- Conducting targeted searches to gather vulnerabilities based on specific keywords that encompass a range of individual devices within each type. For instance:
  - **Audio**: Audio and Speaker
  - **Camera**: Camera and Video
  - **Home Automation**: Plugs, Bulb, and Lamp
  - **Appliances**: Brewer, Kettle, Microwave, Vacuum, Washer, Dryer
  - **Smart Hub**: Hubs, Smarthub, and Smartthings
- Normalizing the search results to facilitate storage and management, ensuring that the data gathered is comprehensive and relevant.

3. Vulnerability Standardization

Standardization of the results is critical for ensuring uniformity across the datasets obtained from various sources. The process involves:

- Extracting the CVE-ID from the results harvested from Vulners and IBM X-Force, which are then matched against the NVD database.
- The NVD serves as the "master source," given its comprehensive verification and scoring of CVEs.
- This step guarantees that the standardized results from Vulners and IBM X-Force align with the attributes provided by the NVD, allowing for a more reliable assessment.

The standardized results are stored in a local database, enabling efficient access and management.

4. Vulnerability Aggregation

In the final step of the vulnerability assessment, standardized results from each database are aggregated into a single, coherent dataset for each device type. The aggregation process includes:

- Combining the standardized vulnerabilities obtained from NVD, Vulners, and IBM X-Force for each device type. For example, vulnerabilities related to the **Camera** device type would be represented as a combined data source: NVDcamera, Vulnerscamera, and XForcecamera.
- This unified view facilitates a more straightforward analysis and understanding of vulnerabilities associated with each device type, enabling stakeholders to prioritize vulnerability management activities effectively.

VISUALIZATIONS

To effectively communicate the results of our performance analysis, we included several visualizations:

- **Bar Graphs**: These graphs illustrate the accuracy, precision, recall, and F1-score of different algorithms across the datasets, facilitating easy comparison.
- **Scatter Plots**: Used to depict the relationship between inference percentages and model performance metrics, highlighting trends and insights into transferability.

These visualizations not only enhance the interpretability of the results but also provide a clear representation of the vulnerabilities associated with each model in the context of IoT device profiling.

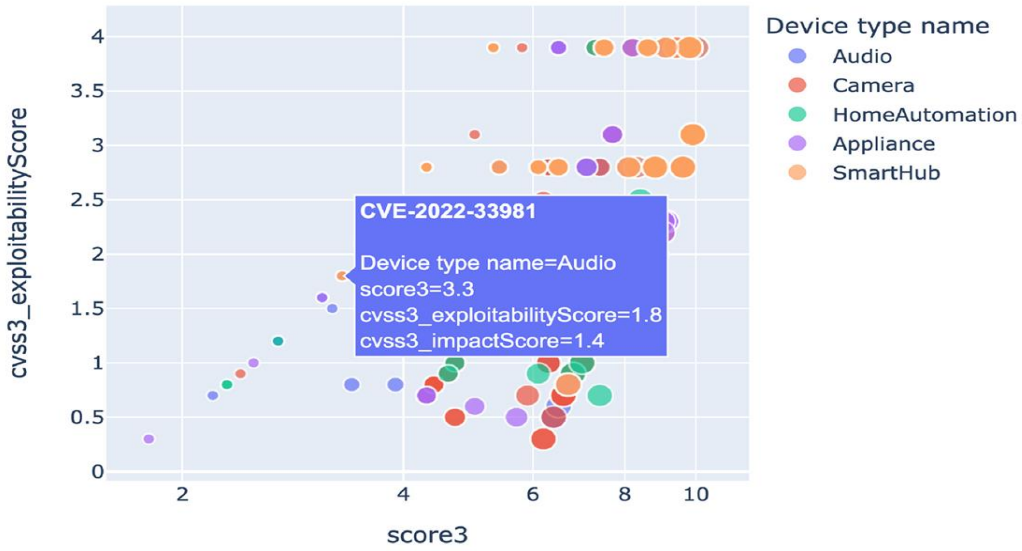

Fig. Scatter chart of device types with score3 on the x-axis and exploitability score on the y-axis.

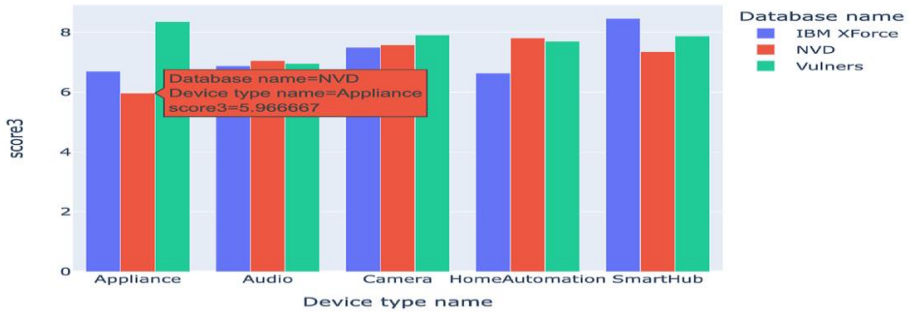

Fig. Bar chart showing the average CSSV score3 of the different predicted device types in the three different databases.

7- CONCLUSION

This study explored the use of transductive transfer learning to enhance IoT device classification accuracy by transferring knowledge from the CIC IoT 2022 dataset to the IMC2019 dataset. The results showed reliable identification of device types, even across different environments, with strong performance for core IoT devices. The findings highlight the potential for improving IoT device security by tailoring security mechanisms to specific vulnerabilities. However, challenges remain with new device types and environment shifts. Future work could focus on integrating unsupervised learning to improve adaptability and fine-tuning anomaly detection for better security in specific deployments.

REFERENCES

1. Danso, P. K., Dadkhah, S., Pinto Neto, E. C., Zohourian, A., Molyneaux, H., Lu, R., & Ghorbani, A. A. (2024). Transferability of Machine Learning Algorithm for IoT Device Profiling and Identification. *IEEE Internet of Things Journal*, *11*(2), 15 January 2024.
2. A. M.Almomani and M.N.A.Rahman,"Aliterature review of the adop tion of Internet of Things: Directions for future work," Int. J. Contemp. Manage. Inf. Technol., vol. 2, no. 2, pp. 15–23, 2022.
3. L. Fan et al., "An IoT device identification method based on semi supervised learning," in Proc. 16th Int. Conf. Netw. Service Manage. (CNSM), 2020, pp. 1–7.