PhD school: Techniques for Increasing Trust and Reliability in IoT Systems

Codruta Maria SERBAN
Department of Computer Science, Technical University of Cluj-Napoca, Romania
Codruta.Serban@cs.utcluj.ro
Career stage: 2nd year PhD student

Abstract

Centralized data processing from IoT devices may generate privacy issues as sensitive information would be transmitted through less secure networks to cloud systems. As the number of IoT devices connected to the Internet is increasing rapidly, it leads to huge amounts of collected data that are the target of malicious attackers.

A trade-off between privacy and performance (e.g. accuracy) should always be considered when designing an IoT system. Global modeling generates more precise classification, but without privacy preservation; integrating several privacy preservation techniques may provide similar results without exposing sensitive data.

CCS Concepts

• **Computer systems organization → Embedded and cyber-physical systems**; • **Security and privacy → Privacy protections**; • **General and reference → Cross-computing tools and techniques**.

Keywords

IoT systems, sensitive data, privacy preservation, reliability

1 Introduction

The introduction of IoT determined significant changes in how data is handled in real time systems. Since its beginning in 1990 [3], IoT has been integrated in a large variety of domains: Industry 4.0, healthcare, transportation, military, smart homes, cities, and vehicles [6], [2].

1.1 Motivation

One of the reasons behind choosing to study trusted and reliable IoT systems is the accelerated growth of the number of IoT devices. As more devices are connected to the Internet every day, they surround us everywhere and it results in an increased volume of data (Cisco estimates that by 2030, there will be 500 billion connected devices [6]). In the same time, more and more attacks are reported against them, affecting user confidence in this technology.

Industry 4.0 describes the trend toward automation and data exchange between technologies and manufacturing processes that include Cyber-Physical Systems (CPS), Internet of Things (IoT), Industrial Internet of Things (IIOT), cloud computing and artificial intelligence. In this context, the security and reliability of systems is very important, because in industrial processes, unforeseen incidents can have serious consequences, both materials, but which can also involve human victims.

1.2 IoT problems

The following problems are common in IoT system:

- IoT heterogeneity: device from different vendors, having various architectures, with different operating systems and software application programming interfaces [5].
- resource limitations: IoT devices are limited in terms of computational power, storage capacity, battery lifespan and data access [5].
- trust in IoT systems: the expectation that the system outcome and process won't harm the user(privacy leakage or incorrect outcomes) [1].
- reliability: a critical factor for ensuring seamless operation

Those problems can be reformulated as research questions:

- How can lightweight privacy-preserving methods be designed to provide strong protection for resource-constrained IoT devices?
- How can IoT data access mechanisms be designed to provide flexible, secure, and efficient access control across heterogeneous devices and dynamic network conditions?
- What decentralized processing techniques can be used to ensure secure and reliable interactions in IoT systems without relying on a central authority?

1.3 Research objectives

My work research objectives aim to address critical challenges in enhancing trust and reliability in IoT systems:

- Enhancing trust in IoT systems through privacy preservation mechanisms
- Develop flexible IoT data access mechanisms
- IoT Data Processing in a decentralized distributed environment

2 Research Approach and Methodology

In the beginning of my PhD, I studied about IoT systems, the existing methodologies and challenges.

Currently, my focus is on privacy preservation mechanisms. Differential privacy, homomorphic encryption, secure multi-party computation and digital twin approaches are applied to potentially protect model and data privacy during execution and data transfer in IoT networks. I intend to conduct experiments with these methods to enhance privacy preservation in a system able to detect anomalies using federated clustering; but having in mind the balance between privacy preservation, performance, and the operational overhead associated with them.

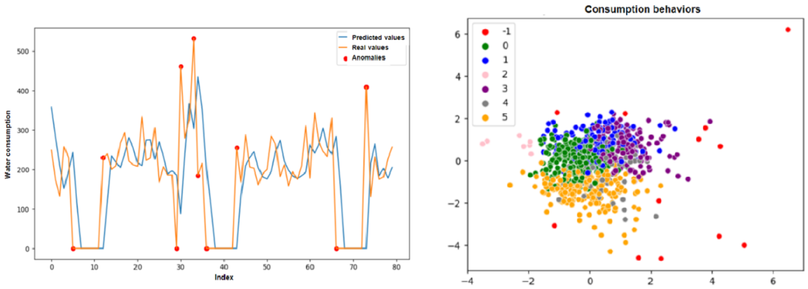

Figure 1: Anomalies detected using prediction and evaluation based method (left), anomalies (red dots) detected using clustering and classification based method (right)

Next, I will direct my efforts toward flexible IoT data access mechanisms. Various data handling techniques have been developed throughout the entire process, from device identification to data collection. The biggest challenge would be to ensure reliability along the way, as there are many possible places where failure and faults can occur. Flexible access mechanisms can allow more controlled, real-time verification and validation of data, ensuring that only authorized entities can access or modify data, thus enhancing trustworthiness and integrity. I want to integrate content-based addressing in the system that I will develop so that the user can search for certain information without knowing the location of the device that collects the necessary data.

Then, I want to explore the advantage of decentralized computation, eliminating the single point of failure. Implementing the FL architecture in a decentralized manner can increase even more the trust and reliability of the system.

Lastly, I will assemble my theses from the results and papers published along the day.

3 Preliminary Results

Until now, my work has focused on enhancing trust in IoT systems through privacy preservation mechanisms. I conducted experiments in two directions: anomaly detection in daily water consumption patterns using prediction and clustering approaches [4], and comparison of Federated Learning (FL) techniques in IoT context.

Anomaly detection techniques help identify abnormal behavior or events that may indicate a fault or failure within the system. Attack detection is typically categorized as a form of anomaly detection, as it involves identifying abnormal or malicious behavior in system metrics, network traffic, user activities, or other data sources. An attack is considered an anomalous event because it deviates from the expected behavior of the system.

The water consumption of a home can be monitored using smart water meters, able to transmit the readings to other devices. Behavior anomalies are characterized by either a sudden shift in the user's consumption tendency or a deviation from the dominant consumption behaviors. I developed two mechanisms to detect those unusual changes in daily water consumption for a single home. Sudden significant increases or decreases in consumption are targeted by the method based on prediction (ARIMA) and evaluation (comparison with a computed threshold). For identifying the anomalies represented by those records that fit hardly or not at all into dominant consumption behaviors, I combined clustering (KMeans) with classification (Isolation Forest). KMeans identifies the behavioral patterns, and later, Isolation Forest determines if a new record is part of those patterns or not. During experiments, cases were observed where the same consumption was detected as abnormal by both approaches.

Data from the same building was input for both methods and the results are shown in Figure 1. The graph on the left characterizes the method based on prediction and evaluation. The evaluation for a day takes place after the water reading was collected. This way the predicted value is compared with the real value, and based on a computed threshold, that reading is labeled as an anomaly or normal consumption.

On the right side of Figure 1, the red dots are the anomalies identified using the method based on clustering and classification. Using the elbow method, I found that 6 is the optimal value for the number of clusters representing consumption behaviors. Principal component analysis was used to represent consumptions as points. Anomalies were labeled with "-1" while normal consumption is labeled with 0-5. Anomalies are records from clusters with very few elements (less than 1% of the total number of values), or if the distance from the point to the centroid of the cluster it belongs to exceeds a certain value. If a cluster contains few records, it means that it describes an inferior consumption pattern. For threshold computing, histograms for each cluster were used to analyze the distribution of points. The further a point is from the centroid, the harder it is to fit into that cluster.

These results were published in [4], it also includes the methodology used for mitigating measurement errors (negative, missing and duplicated values), augmented by algorithms for computing the thresholds applied in labeling the data set. While they cannot be evaluated due to the used unlabeled dataset, these methods demonstrate an ability to detect consumption changes that can later be verified by the user.

For the second experiment, the use case is myocardial infarction detection in ECG signal using one-dimensional convolutional neural network. I replaced the classic centralized architecture with one based on FL. By default, FL improves privacy as it eliminates the need to send all the data to a central node to train the classification model. The focus was on the aggregation step of federated training in a setup with multiple nodes (six and three) and three learning iterations. We performed experiments with three categories of federalization methods for computing the global model:simple average, selective average and best candidate.

For the first category, we calculated the parameters of the global model in two ways; first with a simple arithmetic mean. Another approach is updating the parameters of the local model with the mean value of its parameters from the last iteration and the ones of the global model, computed with the simple mean.

The selective average approach refers to selecting a subset of the local nodes that are used for computing the global model using the simple mean. For selection, we used three criteria: the first four models with the best accuracy, the first three/four most similar models, and the first two most similar models.

The similarity between models is determined by the sum of the absolute differences in their weights: a smaller sum indicates greater similarity. For the last two criteria, we begin by identifying pairs of similar models and ranking them. From these, we select either the two most similar pairs or just the top pair. If the two pairs share a common model, the average is calculated using three models; otherwise, four models are used.

In the case of best candidate, the global model is the local model with the best accuracy.

Experiments were performed in both centralized and FL setups. 5-second ECG inputs of multiple leads are used for training and testing the models. The purpose was to observe the impact of the FL and the aggregation methods on the classifier performance. The best results after ten runs are captured in Table 1. In the centralized architecture, the best accuracy is 99.3%.

In the FL architecture, the experiments showed that the best-performing methods are those using the simple average and the best candidate. The worst results were from the similarity-based methods because there were cases when bad-performing models that were very similar ended up being used for global model aggregation.

Table 1: Accuracy results of the six-nodes FL architecture using the best performing local model as the new global one

Accuracy	First iteration	Second iteration	Third iteration
Local model 1	95.14%	97.08%	97.08%
Local model 2	92.23%	97.08%	100.00%
Local model 3	93.20%	95.14%	98.05%
Local model 4	94.17%	96.11%	96.11%
Local model 5	96.11%	95.14%	97.08%
Local model 6	93.20%	96.11%	97.08%
Average accuracy	94.00%	96.11%	97.56%
Global model	96.11%	97.08%	100.00%

Table 1 contains the accuracy of the classifier in the FL architecture with six nodes and using the best candidate method for obtaining the parameters of the global model. In the first iteration, all local models were initially the same. They were trained with different learning datasets and then used to calculate the global model. In the next two iterations, the local models' parameters are updated with the ones of the global model previously obtained. The table values are obtained after testing each local model and the global ones with the same testing dataset. This way, the models can be compared.

It can be noticed that the accuracy of the local binary classifiers is slightly lower compared to the centralized approach. Additionally, it is evident that the number of learning iterations plays a significant role in enhancing the performance of the global model. This improvement is due to the greater diversity and volume of records used for training the local models.

An important aspect that needs to be mentioned is that the dimensions of the training and testing datasets used by a single model in the centralized and FL architectures are significantly different. The same dataset was used for the experiments; in the first scenario there are around 8000 records used for training and 2000 for testing; Compared to 555, respectively 103 in the 6-node FL architecture.

In most cases, the accuracy achieved in the FL architecture is acceptable, given its key advantage: privacy preservation. This trade-off between accuracy and privacy should be carefully considered when implementing real-world applications involving sensitive data.

4 Current Research Challenges as PhD Student

While working on my PhD, I encountered several challenges, both technical and non-technical. They are areas where I would appreciate insights and advice from experts and colleagues.

4.1 Technical Challenges

Another challenge would be developing trust models that can scale efficiently across large, heterogeneous IoT networks with diverse devices and platforms. Additionally, IoT devices often have limited computational power, storage, and energy, making it difficult to implement complex privacy preservation techniques or reliability mechanisms without affecting performance. Lastly, the deployment process is a real challenge. A robust communication network is required, and it should maintain consistent connectivity, especially in remote or hard-to-reach areas, where the IoT devices are located. The infrastructure should be able to handle and process in real-time large volumes of sensitive data.

4.2 Non-Technical Challenges

The IoT field evolves quickly, with new devices, standards, and threats emerging regularly. Staying current with these trends and ensuring your research remains relevant can be difficult. Balancing the demands of research, publishing, and potentially teaching or working in the industry can stretch your time and energy.

References

[1] Hanan Aldowah, Shafiq Ul Rehman, and Irfan Umar. 2021. Trust in iot systems: a vision on the current issues, challenges, and recommended solutions. *Advances on Smart and Soft Computing: Proceedings of ICACIn 2020* (2021), 329–339.

[2] Enrique Tomás Martínez Beltrán, Mario Quiles Pérez, Pedro Miguel Sánchez Sánchez, Sergio López Bernal, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez, and Alberto Huertas Celdrán. 2023. Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges. *IEEE Communications Surveys & Tutorials* (2023).

[3] Rui Hou, YanQiang Kong, Bing Cai, and Huan Liu. 2020. Unstructured big data analysis algorithm and simulation of Internet of Things based on machine learning. *Neural Computing and Applications* 32, 10 (2020), 5399–5407.

[4] Codruta Maria Serban, Gheorghe Sebestyen, and Anca Hangan. 2024. Anomaly Detection in Water Consumption Patterns Using Prediction and Clustering Approaches. In *2024 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*. IEEE, 1–6.

[5] Tuo Zhang, Lei Gao, Chaoyang He, Mi Zhang, Bhaskar Krishnamachari, and A Salman Avestimehr. 2022. Federated learning for the internet of things: Applications, challenges, and opportunities. *IEEE Internet of Things Magazine* 5, 1 (2022), 24–29.

[6] Yousaf Bin Zikria, Rashid Ali, Muhammad Khalil Afzal, and Sung Won Kim. 2021. Next-generation internet of things (iot): Opportunities, challenges, and solutions. *Sensors* 21, 4 (2021), 1174.