Edge Intelligence in Industrial Internet of Things(IIoT)

Bachelor Thesis Report submitted
by

Vaishnavi Munghate (18CS30045)

Under the guidance of Prof. Sudip Misra

Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur - 721302
West Bengal, India

Spring 2021

ABSTRACT

We propose here a schema and its rudimentary implementation using the concept of edge intelligence which will be useful in industries. Nowadays IoT has become an important part of technology and also in industries. Thus industrial IoT is separately defined. We will discuss IoT and IIoT and their differences. We will also talk about edge computing, fog computing and edge intelligence. Some related works in Edge Intelligence and IIoT are also mentioned. Finally, we will propose a probable solution and its rudimentary implementation that can be applied in industrial IoT. In the proposed solution, we will talk about two types of solutions and their architecture. For the implementation part, we will take a dataset, analyse it and apply machine learning algorithms to get results and make predictions. Finally, limitations and further extensions to the proposed idea are discussed in the conclusion section.

Contents

1 Introduction

Internet of things is a network of large number of electronic/physical devices. The devices are used to collect, process, filter, analyse and share data. These devices have sensors embedded in them which helps to collect, process, filter, analyse and share data. These devices like mobile phones, laptops are the basic elements of IoT. IoT provides facilities like analytics and web processing into the world of physical devices. As these physical devices are connect through a large network, using this network they recieve and send instructions to perform the actions that they are supposed to do without any human intervention. They are self programmed.

The IoT devices have their own IP addresses as they are connected through a network. They are internet-connected. Some devices are simple like mobile phones, laptops, barcode readers in various shops, simple sensors that monitor the temperature in buildings etc. Some devices are very complicated like self-driving cars,autonomous vehicles that move products around the factory floors. Also certain IoT devices are for personal use like fitness tracker, moblie phones, laptops and computers. Several people use them in their day to day life. The IoT network handle these devices in several ways so that they can perform their task successfully. All the factory robots, the autonomous vehicles that move products around industrial settings and warehouses, all of them are examples of IoT devices. There are also some generic devices which are used to build our own IoT endpoints. Examples of such devices are Raspberry Pi and Arduino.

The data sent by the IoT devices is collected at a data gathering point. The data transmission is done through wired networks or wirelessly employing a range of technologies that is wireless network. The place where this data is sent, stored, processed, shared and analysed is called the knowledge centre or the cloud. The cloud has high storage capacity and compute power so it is used for this purpose.

The processing of data is done by the cloud that is the knowledge centre, but we can also think of processing or computing small amount of data at the source. Let us take the case where we need quick responses, like fire alarm or a critical device in industry got shut off. In this case response time or latency becomes an important factor. If in such cases data is collected then sent to cloud, processed and then sent back, then this might lead to loss of several resources. The factory might suffer from a great loss. To prevent such situations we can think of a method that does data processing in the device itself or something near to that. So, the method of Edge-computing is used here. It plays an important in computation of data at the endpoint. It also facilitates sharing or data so that it can be computed in the further sent place. This is called upstream connectivity. As the distance of processing is decreased as it is closer relatively, latency is decreased by a very large amount. Such devices are called smart edge or edge devices.

Moving towards our topic Industrial Internet of Things that is IIoT, It is a subcatagory of IoT. This means that IoT used in Industries or used for Industrial purposes is called as IIoT. IIoT is an important part of industries which contributes in further enhancement of Technology in Industrial sector. Industies these days focuses on data automation, smart technology, interconnectivity, Artificial Intelligence(AI) and several other recent technologies. One of these technologies is IIoT.

The advantages of IIoT is similar to IoT but in IIoT we deal with large amount of data at a time. IIoT has long period of life cycles than IoT. One more advantage of IIoT over IoT is that it is more reliable, more secure and protects data for a long period of time. IIoT also deals with large scale networks and data. IIoT devices varies from sensors, barcode readers and detectors to large product moving autonomous devices. Use of these devices helps the industrial sector to uplift their efficiency, productivity and employ safety and many more.

IIoT increases machine-to-machine communication and decreases human intervention, the manager of the edge has gets all the data where it can process and send it back to them. They interpret the data and work accordingly with least human intervention. These also decrease the chances of errors. Through the continual collection of granular data, industrial companies can keep a closer eye on the energy, water and other resources they are using, when their machinery is running and how much they are producing.

As technology is getting updated day by day, the use of IIoT is increasing that helps companies to save a significant amount of time and energy. It helps companies to achieve their increasing demands and manufacturing goals.

Now talking about Edge Computing, it means processing, analyzing, and computing data at the edge. The edge is a place where computation is done other than the cloud and it is at a lesser distance from the devices than the cloud. The edge devices are also called as edge nodes. They can be mobile phones, laptops, sensors, several autonomous industrial machines. Edge computing is used as it decreases latency that is delay in giving the results. Edge computing came into role when the number of IoT devices were increased enormously, the cloud could not manage such large number of devices thus increasing the latency.

When we think of doing everything that cloud can do at the edge,edge intelligence came into role. When Edge computing is combined with machine learning intelligent algorithms it is called Edge Intelligence. Edge intelligence can be said to be a step ahead of edge computing where we perform actions after analysis at the edge itself involving Artificial Intelligence, Machine learning or deep learning. This deviates from cloud analytics and cloud intelligence where we send all the data over the network to the centralized data store that is the central edge and perform analysis and decisions.

As the name of the paper suggests that is Edge Intelligence in IIoT, will be using Edge Intelligence in IIoT. After knowing everything about IIoT and edge intelligence we will be proposing a solution that can be used in the industrial sector that uses Edge Intelligence.

1.1 Motivation

As the use of IIoT is increasing day by day, we would like to propose an edge intelligent scheme for the better use of IoT in Industrial sector.

1.2 Contributions

This report contributes an edge intelligent scheme that can be used in Industries. We used Machine learning techniques that are implemented at the edge. The scheme that we proposed can be used in two ways. Both the architectures are described in the coming sections.

- A two-layered architecture in which an edge and the cloud is present and the computation is done in the cloud.
- An three-layered architecture in which an edge, fog and the cloud is present and the computation is done by the fog.

2 Related Works

In this section, we will introduce some related works about Edge Intelligence and the Internet of Things (IoT).

The article Edge computing in industrial IoT framework for cloud-based manufacturing control [1] talks about using embedded intelligence as a solution for their problem statement. The solution is for data aggregation, data collection and smart processing from shop floor resources and products with embedded intelligence. It was successfully implemented and tested with two IoT device networks. The crucial part of the solution is that the aggregation node collects information from different devices and writes the records on a database located on a private cloud. They also performed experiments that proved that the Message Queuing Telemetry Transport (MQTT) protocol is one of the best choices for data transfer from embedded systems to the cloud. This is because of the following characteristics: First, it is a published or subscribed solution that allows the realization of infrastructures that can be communicated in the many-to-many case(three cases were discussed one-to-one, one-to-many and many-to-many). Second, this article proposed a solution for shop-floor device interoperability for intelligent products and resources which are present in a production control system. Third, the provided solution is based on aggregation nodes that gather data from many devices which are directly connected and forward the gathered data to the cloud control platform hosting the high-level Manufacturing Execution System (MES) control system. This control system is in charge of operation optimization, execution, and monitoring.

The article Mobile edge cloud-based industrial internet of things: Improving edge intelligence with hierarchical SDN controllers [2] proposed a realistic solution with IIoT architecture having a hierarchical control structure in the mobile edge cloud (MEC) that is a set of devices and governing software is arranged in the form of a hierarchical tree. The architecture separates the data plane and the control plane based on software-defined networking (SDN). As hierarchical control structure is used, the hierarchical controllers refined the intelligence and the flexibility of the control plane. Also, the remote radio head (RRHs) and servers in the same cluster, formed a MEC-based radio access network (RAN) and the supported RAN function in a split. It also improved the scalability and cooperative gain of the data plane. To enhance edge intelligence, a deep learning algorithm is implemented in the MEC. They also designed two control schemes, the first is centralized and the other is distributed. These schemes provided a trade-off between performance and overhead. They aimed to lessen the system delay that is as little as possible, so they formulated a joint optimization problem of RRH allocation, RAN function, task scheduling and virtual machine assignment split as an example. To search for a solution, an efficient algorithm was proposed based on submodular function maximization. Finally, they created a delay minimization problem and estimated that a greedy algorithm can find sub-optimal solutions with an approximation of 0.5.

To address the new challenges of anomaly detection in IoT because of the massive use of IoT applications and cyber attacks in the article Edge Intelligence (EI)-Enabled HTTP Anomaly Detection Framework for the Internet of Things (IoT) [3], a realistic anomaly detection framework based on recent advances in edge intelligence has been proposed. They used machine learning algorithms like running sequential clustering and classification on the HTTP traffic, which can efficiently and effectively discover unauthorized and unknown network intrusions. To eliminate redundant data they presented a data processing method with field detection, which divided the header of HTTP traffic data into several multiple fields used for detection. As some redundant data was removed for detection, the procedure accelerated the detection speed and improved the detection performance. The proposed solution and framework has better performance in terms of precision, recall, accuracy and F-measure that is the detection speed, as well as the accuracy test. Also, the proposed framework can efficiently relieve computing pressures of centralized servers and network congestion, just like releasing the potential of Edge Intelligence in IoT. The proposed framework does not talk about the protection-based security schemes, like continuous authentication which may be a drawback.

The article Toward edge intelligence: multi-access edge computing for 5G and internet of things [4] talks about multi-access edge computing(MEC) in IoT. MEC supports stringent requirements, like the high-bandwidth requirement and the low-latency requirement. It also provides cloud computing capabilities at the edge network. MEC enables a very wide variety of applications because of these superior characteristics of low latency, high bandwidth, proximity awareness, etc. They conducted a comprehensive survey on MEC in the context of 5G and IoT in their article as MEC plays an important role in 5G and IoT. They summarized the features of MEC and compared them with the traditional mobile-cloud computing (MCC) in terms of energy-saving, context awareness, and latency. They used technologies like information-centric networking (ICN), software-defined networking (SDN), cloud computing, network function virtualization (NFV), virtual machine (VM) and containers, computation offloading, smart devices and networking slicing that facilitate MEC to be integrated into 5G and IoT. They also elaborated on the role that MEC played in 5G and IoT. Some of the roles are, they gave an overview of the MEC-enabled 5G and IoT applications, like intelligent factory automation, healthcare, transportation systems, smart grid, smart city, smart home, smart retail and smart farming, and also proposed several promising research directions when MEC is applied into 5G and IoT. Also, they talked about the challenges and open issues in the solution provided, such as edge intelligence, deployment of MEC servers and systems, green MEC, mobility management, privacy, pricing, and security. They proposed a use case that uses MEC to realize edge intelligence in IoT scenarios. They did not provide any heuristic algorithms as a solution, they just provided a use case and applications of MEC in IoT and 5G.

In the article Edge intelligence in the cognitive internet of things: improving sensitivity and interactivity [5], the writer introduces us to the evolution of IoT technologies caused by cognitive computing and edge computing. They also talked about the two key features, cognitive services and collaborative sensing which are inspired by edge intelligence in the Cognitive Internet of Things (CIoT) are. Moreover, some open issues and challenges of CIoT were presented and four representative cognitive services and applications applied using edge intelligence in CIoT were presented. It is important that cognitive computing should be developed from time to time, to meet the technical challenges, like efficient computing/storage at the CIoT edge, integration of multiple data sources and types, and generation of big sensory data. Also, to get evolved with the upcoming computing and communication paradigms, the CIoT ecosystem should remain updated by absorbing new capabilities like deep learning, machine learning, a CIoT sensing system, data analytics, and cognitive capability to provide human-like intelligence. This article talks about the usage of edge intelligence in CIoT, its applications and also possible future usage.

In the article, A re-configurable method for intelligent manufacturing based on industrial cloud and edge intelligence [6], an edge-decision-making and a cloud-assisted manufacturing architecture is presented. This architecture is called the CASOMA-IPE, which is made up of the cloud, an intelligent production edge (IPE), and field devices. The IPE provided inter-operability between the cloud and the field devices as an IIoT data bridge. The cloud converted a particular process into a task and then send it to the field devices by using a negotiation process. The IIoT-based elements of the CASOMA-IPE can be modelled as agents which builds a multi-agent system (MAS). This mechanism makes sure that the production logic could be reconfigured from the cloud to the field devices by the CASOMA-IPE taking the presumption of guaranteeing the global system performance. Their results of the experiments showed that the dynamic scheduling that is done based on the CASOMA-IPE is more efficient and effective than the traditional scheduling and it is easy to implement in intelligent manufacturing systems (IMS). Also, it is suitable for handling mixed-flow production tasks based on random orders. Moreover, the article did not provide in-depth information about the formal modelling of production lines which could be a limitation.

The article Edge computing gateway of the industrial internet of things using multiple collaborative microcontrollers [7] proposed the design of a realistic multi-microcontroller (MCU) IIoT gateway architecture. In this architecture or the gateway, a high-speed parallel bridge controller was integrated into a multi-MCU hardware architecture. This gave local collaborative and computation advantage. The IIoT system saved the bandwidth costs of transferring the data by using the local computing of the gateway, to the database systems and the remote cloud servers. The other feedback of using the local computing of the gateway is that it shortened the response time of machine-to-machine communication. The multi-MCU gateway architecture that is used here used low-cost and low-clock MCUs to perform collaborative and distributed computing, which reduced the power consumption of IoT systems. The designed IIoT gateway solves the expandable peripherals services problem, the high scalability problem, and the flexibility problem which a single MCU can not achieve. At last, an application of the given solution to monitor an IoT-based smart machining system is proposed.

In the article, Serious challenges and potential solutions for the industrial Internet of Things with edge intelligence [8], some potential solutions and serious challenges are discussed in edge intelligent IIoT. These serious challenges involved data security and energy consumption and their potential solutions based on CS, chaos, edge intelligence etc, were designed accordingly. Some serious challenges like security and energy consumption in Edge intelligent IIoT were discussed. Some potential solutions were presented to address the serious challenges that were mentioned. Several surveys were done and a simple experiment analysis is also done that verified the feasibility of the solutions that were proposed.

3 Reco-Edge: The Proposed Scheme

We propose *Reco-Edge*, a reconfigurable edge network for industrial IoT. Our solution deploys the required analytical module to the edge node, which is configurable using different analytical parameters. The centralized cloud does the programming part, make reconfigurable decisions and provide required datasets. But when we consider real-time decision making and real-life problems such as fire prediction, where low latency is required, our proposed solution could be helpful. As because of edge computing we can quickly respond so it is implemented in our solution.

We propose a two-layered architecture. The architecture that we propose consists of an edge node and the cloud.

- **The Cloud:** The cloud is the architecture layer which stores a massive amount of data and has high latency compared to the edge node. The role of the cloud in our proposed scheme is that the cloud will provide the required dataset for computation and prediction. Let us consider a case of temperature prediction in the industry to maintain a constant temperature in a room, if we think of doing all the computation in the cloud then it will be time taking and may decrease the efficiency and increase product defects. So, instead, we will do the computations and prediction in the edge node. This will decrease latency and increase the efficiency of the product.

- **The Edge node:** The edge is the architecture layer which does all the computing using different analytical parameters and methods. The role of the cloud in our proposed scheme is that the edge node will be taking the required dataset from the cloud and will perform the computing that is all the analytical and calculation part. It will also apply machine learning algorithms for data modelling and prediction. As the prediction is done it will respond accordingly. Let us go back to the case of temperature prediction in the industry, as the edge node does the computation and prediction part the response will be faster. Thus making the production efficiency. As the prediction is dependent on the dataset provided and also the amount of model training the system becomes reconfigurable. Thus the scheme proposed is a reconfigurable edge network.

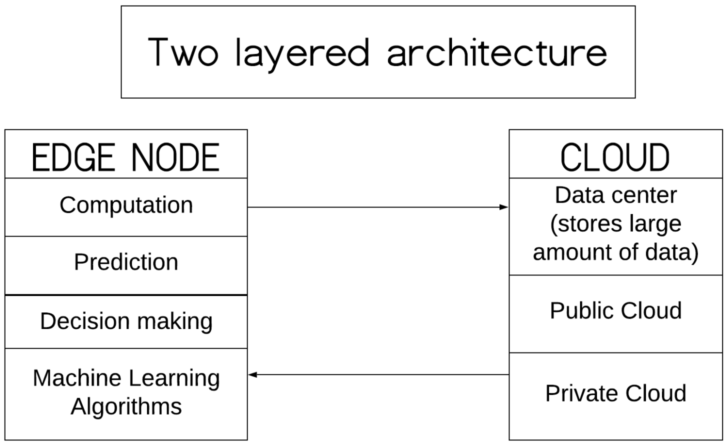

Figure 1: IIoT Reconfiurable edge network

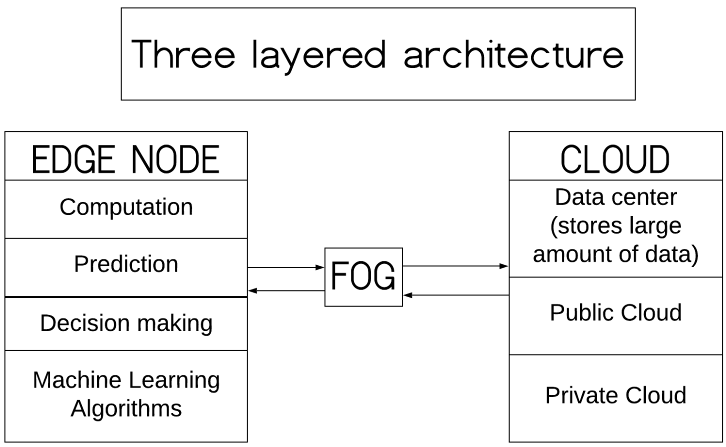

Figure 2: Distributed analytical network

We also propose another method that is more accurate and efficient. Instead of the edge node doing all the computing and decision making we can distribute the data analytics and computing parts between the cloud and the edge node if the processes are independent of each other. This is called distributed edge analytics. This can help in making more accurate predictions thus making the system better. Also, fog computing or fog networking can be done. Fog computing is a decentralized computing infrastructure in which storage, data, compute, and applications are located somewhere between the data source and the cloud. Fog computing has the similar advantages as of edge computing and also the power of the cloud closer to where data is created and acted upon. Some people use the terms fog computing and edge computing interchangeably because both involve bringing intelligence and processing closer to where the data is created. This is often done to improve efficiency, though it might also be done for security and compliance reasons. So, with Fog computing, the data is processed inside a fog node or IoT gateway which is situated within the LAN. As for edge computing, the data is processed on the device or sensor(edge node) itself without being transferred anywhere.

3.1 Advantages

The advantages of the scheme that we proposed are:

- **Performance:** Performing computing and analytics at the edge lowers latency as compared to the cloud, and therefore it greatly increases performance. This is because in the case of the cloud it takes time to transfer data as it is at a large distance from the device. While in the case of edge node it is not.

- **Availability:** As edge is used to compute for nearby networks only so number or computations are less over all thus the availability of the edge is more. A critical system should always be operated and at a relevant time, as the rush is less we the problem will be evaluated faster. Also, connectivity is more at the edge and the chances of loosing connection is less.

- **Data security:** As the data to be analysed never leaves the edge where it is collected and used for computation so the data becomes more secure as we do not have to transmit the data to other places. We only have to make the edges secured in order to make the data secured.

4 Performance Evaluation

We have implemented the idea which is discussed in the previous section using socket programming and some machine learning algorithms. The work of the reconfigurable edge node is done by the server and the data is sent by the client which acts as the cloud in our implementation. All the analysis is done and decisions are made by the server and sent back to the client.

We took a dataset that contained the temperature readings from IoT devices installed outside and inside of an anonymous Room (say - admin room). The

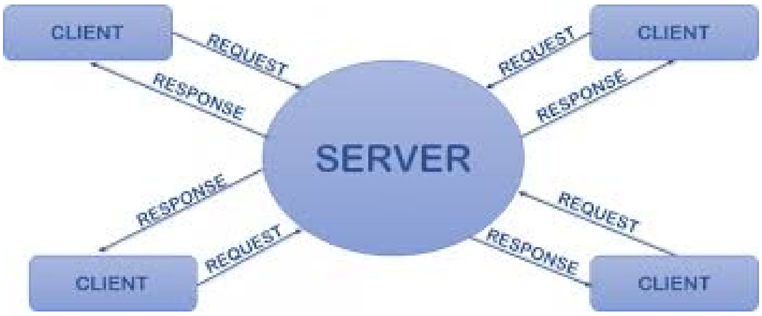

Figure 3: Client-Server socket programming (used in implementation)

device was in the alpha testing phase. So, It was uninstalled or shut off several times during the entire reading period (28-07-2018 to 08-12-2018). These random interval recordings and few misreadings (outliers) make it more challenging to perform analysis on this data. We used this dataset to predict the temperature of a room at the edge node. This data could be considered as the one sent by a sensor to the edge node where the dataset is analysed. Socket programming is used to exchange data in the form of text files between the server and the client. Then in the server as the file sent are text files, we first converted them into CSV format for analysis. We used Logistic Regression which is a machine-learning algorithm to classify the temperature readings and predict whether the given temperature is the temperature in the room or temperature outside the room. The dataset contained 97605 rows and 5 columns which are given below:

- **id:** These are unique IDs of every reading.
- **roomid/id:** These are room ids in which device was installed (either inside and/or outside). Currently 'admin room' only is taken for example purpose.
- **noteddate:** These are dates and time when the readings are taken.
- **temp:** This is the value of the temperature readings.
- **out/in:** This tells us whether the reading was taken from the device installed inside or outside of the room.

As we analysed the data, we found many variables but we are not concerned about the variables like "id", "room id/id", "noted date" because this is classified as noisy or we can say not useful data because they don't have any impact on the class of our output. We can consider "noted date" as a parameter but we did not consider it in our example as it might increase the complexity of the algorithm. So we used the "out/in" values and the "temp" values in our algorithm.

Our algorithm goes like, first we checked the data having the same ids as it was not found so we moved further. Then we found that the value of room id/id was the same for all, so we removed the id and room id columns. Thus segregating the useful parts of the dataset. As the value of out/in was a text that is 'In'

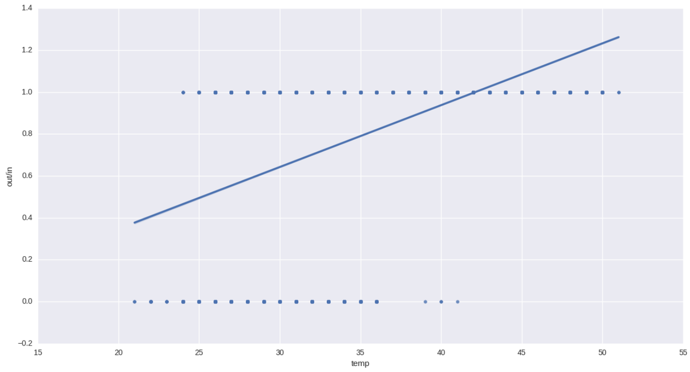

Figure 4: line plot of 'temp' vs 'in/out'

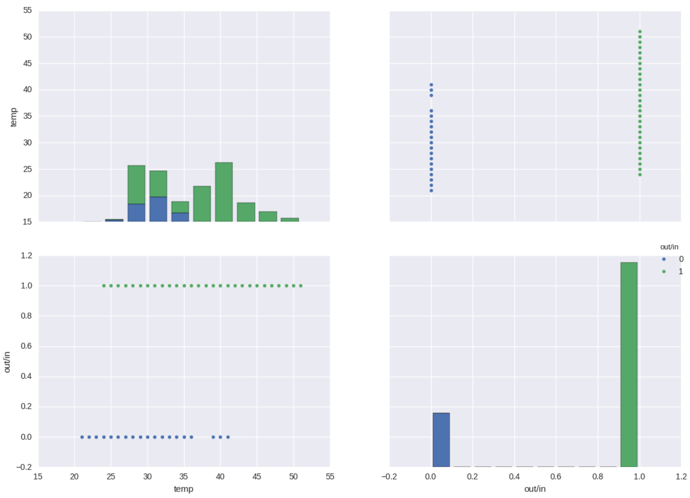

Figure 5: pair plot of 'temp' vs 'in/out'

or 'Out', we cannot work on this data. So, we converted it using LabelEncoder which is a part of *sklearn*. This made the value of In as 0 and Out as 1. Now, we can use the 'In/Out' column as Y for modelling. As there was diversity between input and output variables or we can say dependent and independent variables, we standardized the data to get better results using StandardScaler which is also a part of *sklearn*. Then we split and train data using *sklearn*. Then Training is done on Logistic Regression as this is a classification problem and we don't have any diverse data. Finally, we predict the value and check accuracy using an accuracy score. The value of accuracy came as 0.77(approx).

4.1 Calculations

In our implementation, we are using TCP network and both server and client are using same IP address that is the local host of the PC(127.0.1.1). The latency was found to be 0.620 s(maximum RTT) that is the Round Trip Time(RTT). Generally server and clients are operated in different IPs so the value will be greater than what is calculated.

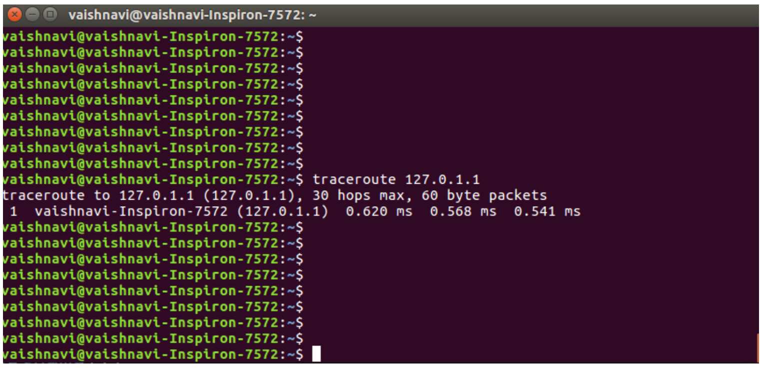

Figure 6: sum of RTTs of all hops gives Latency

The Throughput of the network can be calculated as,

$$Throughput(inbitspersecond) = \frac{TCPwindowsize(inbits)}{latency(inseconds)}$$

Taking maximum TCP window size that is found to be 6291456 bits(by using the command 'cat /proc/sys/net/ipv4/tcp_rmem'). Therefore, throughput is found to be 10147509.7 bits per second. As TCP minimizes data loss through re-transmission of lost data packets so it is efficient.

5 Conclusion and Further Work

This report presented a scheme that can be implemented in IIoT. The proposed scheme used logistic regression, a machine learning technique used for classification which gave an accuracy of 77 percent(approx.). We conclude that the proposed scheme is useful in industrial sectors for prediction purpose. Although the accuracy can be increased by using better machine learning or deep learning techniques.

References

[1] S. Raileanu, T. Borangiu, O. Morariu, and I. Iacob, "Edge computing in industrial iot framework for cloud-based manufacturing control," in *2018 22nd International Conference on System Theory, Control and Computing (ICSTCC)*. IEEE, 2018, pp. 261–266.

[2] W. Xia, J. Zhang, T. Q. Quek, S. Jin, and H. Zhu, "Mobile edge cloud-based industrial internet of things: Improving edge intelligence with hierarchical sdn controllers," *IEEE Vehicular Technology Magazine*, vol. 15, no. 1, pp. 36–45, 2020.

[3] Y. An, F. R. Yu, J. Li, J. Chen, and V. C. Leung, "Edge intelligence (ei)-enabled http anomaly detection framework for the internet of things (iot)," *IEEE Internet of Things Journal*, 2020.

[4] Y. Liu, M. Peng, G. Shou, Y. Chen, and S. Chen, "Toward edge intelligence: multiaccess edge computing for 5g and internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6722–6747, 2020.

[5] Y. Zhang, X. Ma, J. Zhang, M. S. Hossain, G. Muhammad, and S. U. Amin, "Edge intelligence in the cognitive internet of things: improving sensitivity and interactivity," *IEEE Network*, vol. 33, no. 3, pp. 58–64, 2019.

[6] H. Tang, D. Li, J. Wan, M. Imran, and M. Shoaib, "A reconfigurable method for intelligent manufacturing based on industrial cloud and edge intelligence," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4248–4259, 2019.

[7] C.-H. Chen, M.-Y. Lin, and C.-C. Liu, "Edge computing gateway of the industrial internet of things using multiple collaborative microcontrollers," *IEEE Network*, vol. 32, no. 1, pp. 24–32, 2018.

[8] Y. Zhang, H. Huang, L.-X. Yang, Y. Xiang, and M. Li, "Serious challenges and potential solutions for the industrial internet of things with edge intelligence," *IEEE Network*, vol. 33, no. 5, pp. 41–45, 2019.