Data Collection and Processing Framework for Smart Transportation Based on IoT

A dissertation

Submitted in Partial fulfilment of the Requirements

for the award of the Degree

of

Master of Technology

in

Computer Science and Engineering

Submitted By

Suhani Sharma

2K16/CSE/16

Under the Supervision of

Mr. R.K. Yadav

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering), Shahabad Daulatpur, Main Bawana Road, Delhi-110042

June 2018

CERTIFICATE

This is to certify that Project dissertation entitled **"Data Collection and Processing Framework for Smart Transportation Based on IoT"** submitted by **Suhani Sharma** (roll no. 2K16/CSE/16) in partial fulfilment of the requirement for the award of degree Master of Technology (Computer Science and Engineering) is a record of the original work carried out by her under my supervision. To the best of my knowledge this work has not been submitted in part or full for any degree or diploma to this university or elsewhere.

Mr. R.K. Yadav

Date **Assistant Professor**

Place : Delhi Department of Computer Science & Engineering

Delhi Technological University

DECLARATION

I, Suhani Sharma, Roll No. 2K16/CSE/16 student of M. Tech(Computer Science and Engineering), hereby declare that the project dissertation titled **"Data Collection and Processing Framework for Smart Transportation Based on IoT"** which is submitted by me to the department of Computer Science and Engineering, Delhi Technological University, Delhi in partial fulfilment of requirement for the award of the degree of Master of Technology is a bonafide report of Major Project-II carried out by me. I have not submitted the matter embodied in this dissertation for the award of any other Degree or Diploma.

Place: Delhi **Suhani Sharma**

Date:

ACKNOWLEDGEMENT

First of all, I would like to express my deep sense of respect and gratitude to my project supervisor Mr. R.K. Yadav for providing the opportunity of carrying out this project and being the guiding force behind this work. I am deeply indebted to him for the support, advice and encouragement he provided without which the project could not have been a success.

Secondly, I am grateful to Dr. Rajni Jindal, HOD, Computer Science & Engineering

Department, DTU for her immense support. I would also like to acknowledge Delhi Technological University library and staff for providing the right academic resources and environment for this work to be carried out.

Last but not the least I would like to express sincere gratitude to my parents and friends for constantly encouraging me during the completion of work.

Suhani Sharma

Abstract

The efficiency of any ITS depends upon two crucial aspects- To gather the information from multiple-hop- event-driven nodes and to store it in a way that the database extracts relevant information in a time bounded manner. In my research work, I was able to find out the various connectivity glitches that led to ITS not performing relevant actions as and when required. I broke down my work into three modules viz-a-viz data collection, data processing and data storage. By doing this I was able to focus on the different levels of working of a smart transportation and find out the glitches. In my work, the functions follow as follows; first the data is gathered from the sensors using a dedicated server-client connection, then the data is processed in a way that the system doesn't stop working even if data is not available for a while due to loss in connectivity and then stored in a way that it can be accessed using keywords. This will lead to less dependence on technically skilled workforce and help the unskilled workers to secure jobs.

Chapter-1
Introduction

In today's age, one cannot think life without transportation. As urban development is on its peak, people are migrating to cities and urban population is continuously rising. Roads are heavily loaded with traffic. This makes smart transportation need of the hour. The objective of smart transportation is effective utilization of road space available, hence saving ones time and energy.

Through my research I found out that problem of road congestion is mainly due to lack of real-time information. To go in further depth of my work I firstly introduce you with what a smart city is and in further chapters the functioning will follow.

1.1 Smart city

- To manage a city's so much variable and distinct technologies using IOT is a vision held to bring efficiency and reliability in fast growing systems.

1.1.1 Fundamental characteristics

- **Dynamism:** A smart city has both the dimensions : static as well as dynamic. Static questions are invariable and hence don't need regular attention once set-up. However they require a timely maintenance. The problem arises with the dynamic questions like traffic, disaster management, accidents, crime etc.

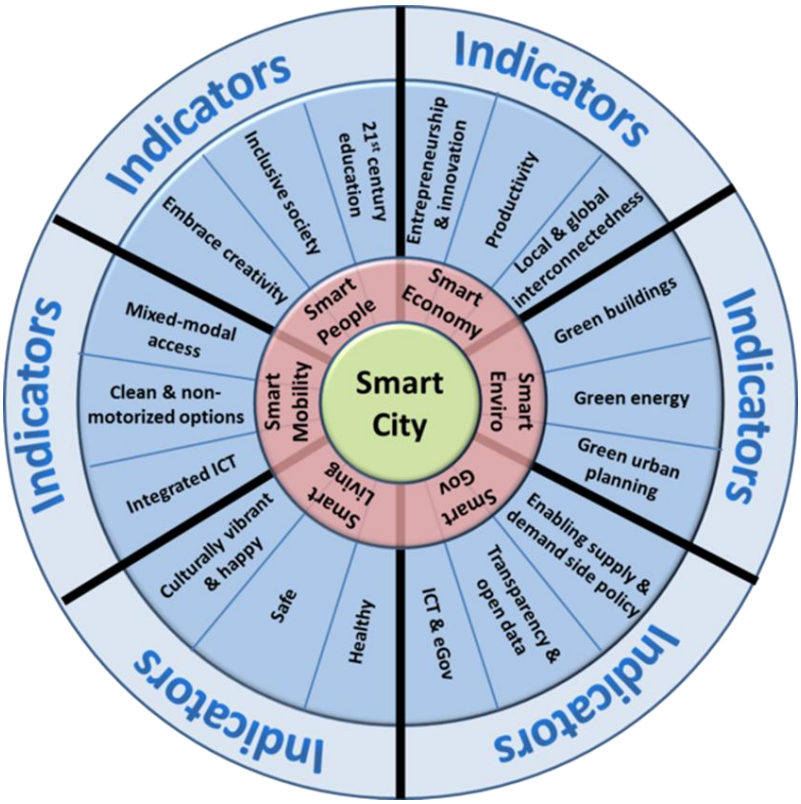

Figure 1.1 Boyd Cohens smart city wheel

- **Uncertainty**: It refers to events that are very less obvious. It leads to change in the course of action i.e. systems start working against the previous predictions.
- **Real-time:** Without any doubt, a smart city requires a real time data, like the people commuting, crimes etc.
- **Mobility:** It requires to physical and control layers to communicate. The data set is mostly redundant. It also requires a high range within connectivity.
- **Duration and scale of effect:** As every function is related to the other in one or the other way, it can lead to effecting those systems which were not faulty in the beginning. Sometimes small bug fixing can lead to a major change and sometimes a major check may produce small results.
- **Efficiency:** It has many aspects like cost, time and desired output. So if we enable systems to work according the targeted efficiency, it will increase manifolds.

1.1.2 Key Challenges:

- **IT Infrastructure:** Lack of infrastructure, will prove to be a barrier in ICT systems.
- **Security and privacy:** A community's trust and confidence is of utmost importance. An individual will only opt for technology if only his safety is not at stake. Smart cities hold very crucial data of an individual and hence it becomes important to make it secure.
- **Big data Management:** A system with high reliability and scalability with less downtime is required to handle the huge volumes of data involved in ICT.
- **Cost:** Sensors, networking equipments require a huge amount of investment.
- **Heterogeneous environment and Interoperability:** Networks, devices, platforms heterogeneously connected is required for smart cities.
- **Efficiency, Availability and Scalability:** As data at every moment is important, so the devices should be always available to record the data and reliable enough for capturing the correctness.
- **Social Adaption:** Digital divide, inequality and cultural divide can be reduced.

1.1.3 Proposed smart city architecture:

The overall architecture has the following objectives:

- M2M communication has become the need of the hour. So this shall form the basis.
- Mixing a new establishing system with already established one in order to make the new one efficient.
- The resources that are limited and less efficient can be clubbed with delay tolerant devices.
- The base to every system can be integrated in order to keep competition alive in terms of efficiency and output.
- Pilot projects can be put in place in order to check the loopholes and save cost.

1.2 Internet of Things

- Due to coming up of wireless technologies, IOT has gained recognition recently. In IOT, the objects act like humans and have certainly all the senses a human has like hearing seeing etc. Like human beings have conversations, similarly, with IOT objects can interact with each other by assigning them certain addresses.

1.2.1 Connecting the dots

identification, sensing, communication, computation, services and semantic are the components of IOT. Services are mostly demand driven hence their identification becomes very important as to when and where one is required. Addressing is mostly done using today's IPV4 and IPV6. Collecting data and forwarding it for processing falls mainly under sensing.

1.2.2 Applications:

- **Smart parking:** In order to save time and fuel of a visitor, smart parking has come into place. The incoming and outgoing cars are tracked and the system is updated with the free spaces available. The visitor can simply go to the place and park.
- **Augments maps:** This can be done in order to provide nearest locations to the user according to his search interests. This will save his hassle of searching it separately on the maps. Like if a user searches for a particular doctor, his results can be augmented with few more suggestions that will help him make a better choice.
- **Logistics**: Retail chain monitoring has many advantages when used with IOT. Sale and purchase of every commodity can be recorded and later viewed using NIFC or RIFC
- **Data collection:** In order to come to a quick diagnosis of a disease, doctors may use the previously recorded data of the patient. This will save the time lapse that may lead to worsening of patients health.
- **Smart water supply:** To provide adequate water to community members, wireless systems can help keep a track on water supply.

1.3 Some of the important domains to be addressed in smart city

1.3.1 Waste management:

- Waste generated if not treated well can cause various diseases that can destroy a cities lifestyle and economy. From its generation to its disposal, waste has to be given different treatments at different level so that it disposes off as soon as possible from the environment. Also it has various legality involved when it comes to its disposal.
- Waste is of many types and this term is inclusive of all such types. Like waste generated in processing raw materials in manufacturing sector, human waste, animals and plants waste etc.

1.3.2 Water management:

- Water system is the critical of all the resources and that too with the ever growing population which will lead to increase in its consumption too. When we look at the wastage of water at all the levels, we see that from its transportation to end usage, water is being wasted at every level. If we get the data about its pressure, flow and amount its leakages can be checked and tracked. For eg. if a place is expected to receive a certain amount of water, a smart system we alert using its sensors that it has not received water

accordingly and hence early detection and fixing of leakages is possible.

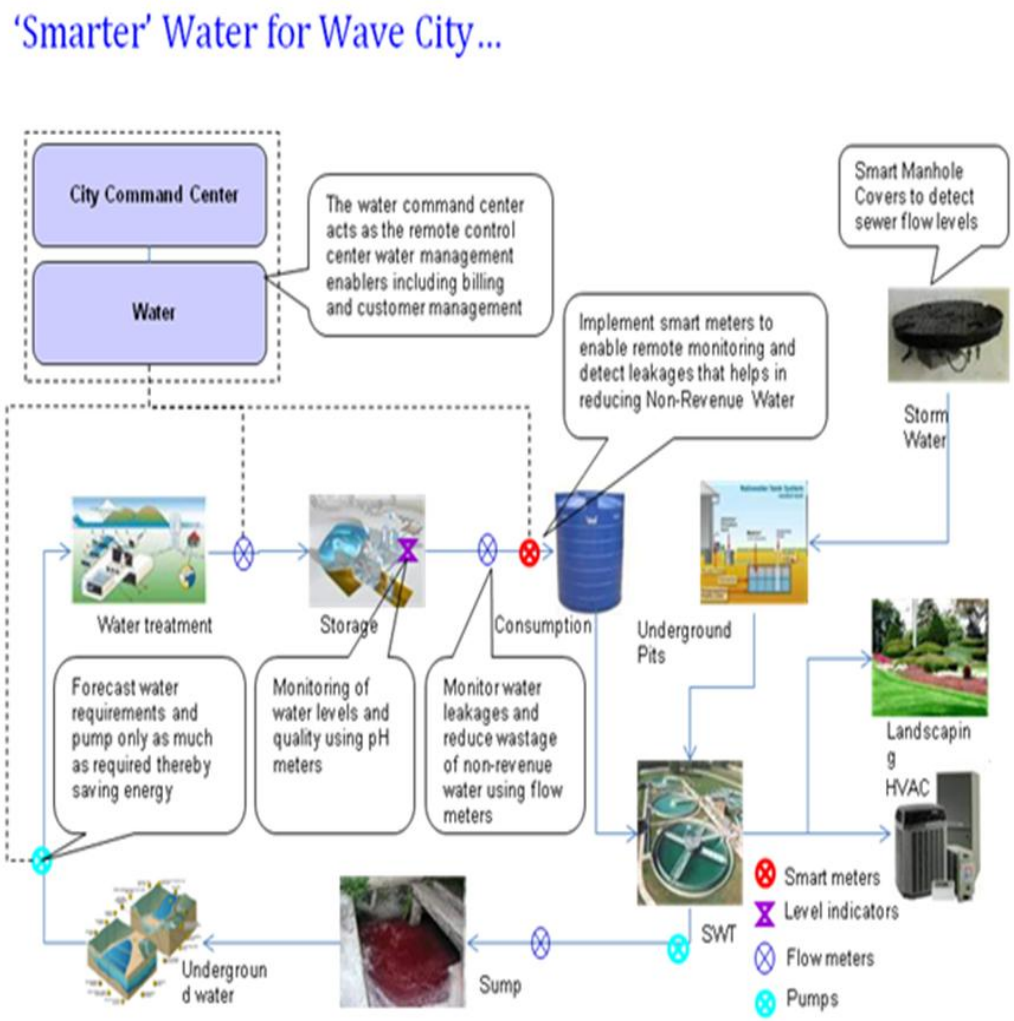

Figure 1.2 Wave city water management example

1.3.3 Energy management:

- Energy being a sensitive resource needs a special attention. From its production to its consumption it has to be used wisely and wastage should be minimum. Businesses, private use both have to be dealt differently as both have a huge difference in amount of use. For eg. a household can still afford a loss in power connection for certain time but for businesses this will amount to a huge loss.

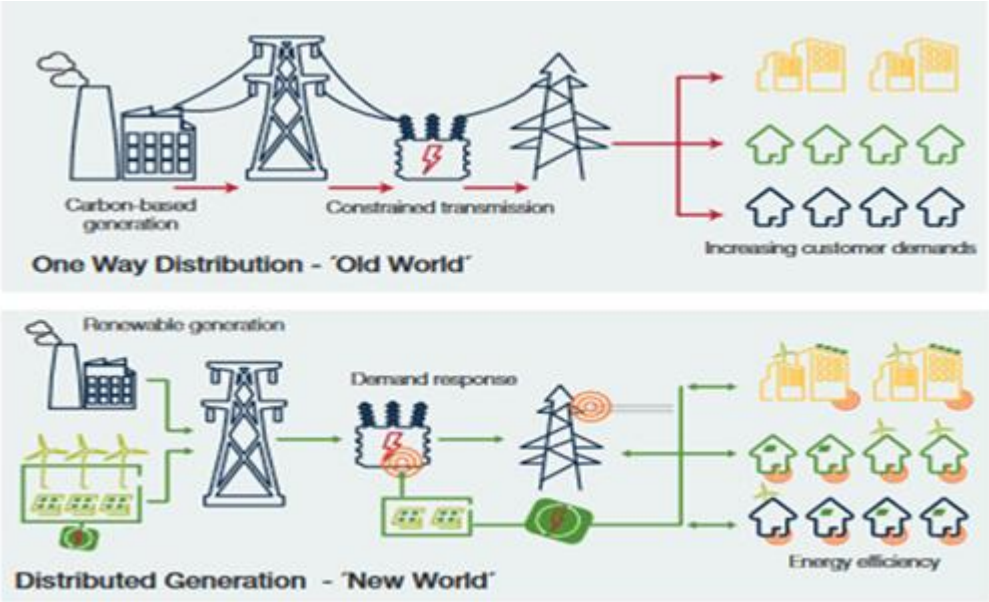

Figure 1.3 Different ways of energy distribution

Without effecting the total cost of production, energy needs to be dealt with very specifically as to how to maximise its usage per unit.

1.3.4 Transportation:

In order to make commuting time efficient and less hectic, it has become important to analyse the data of traffic movement and develop certain patterns. The manner in which a community travels is very dynamic hence we can only deduce certain patterns based on everyday data and assume the rest.

Some of the technologies that hold relevance for smart cities:

- **Bicycle sharing system:** Bicycles are environment and user friendly. For small distances one can travel on a bicycle and save the environment. The best way can be a sharing support where users can share bicycles as and when required.
- **Geospatial TS:** Based on real time data, it provides information on traffic etc.
- **Car Sharing** : Drivers are linked with commuters. It is a cost saving way of commuting. People travelling on a similar route can pair up and travel..

- **GPS tracking :** Public transport can be checked upon by continuously keeping an eye on it by tracking through GPS. This increases its credibility.
- **Integrated transit hubs:** There are few places that can't be reached using a single mode. So mixing all the modes of transportation and allowing people to change from one to another under the same roof makes travelling efficient and easy.
- **Smart parking:** Often we see cars wrongly parked. This can ensure efficient parking.

- **Smart toll:** The vehicles number can be scanned and allowed entry without any hassle and manpower involved.
- **Smart traffic lights:** will help in a good flow of traffic and lesser traffic jams.
- **Electric vehicles:** Pilot systems should be put in place in order to check the market value and efficient consumption of electricity using EV. Making it a public transport is the best alternative.

A huge amount of data is to be transferred from one node to other as huge functions are involved. Not only the information sharing and delivery is important but also delivering the correct information becomes important. So system's mechanisms have to deal with:

- In ITS not only data is very uncertain but also a huge amount of data is exchanged due to ever increasing traffic. So storage becomes important.
- The ITS data is highly vulnerable to delays. In the applications where safety is not a concern, one can still accept the delays. But with ITS, users trust is at stake. This forms the basis of usage of a system.

Data has an importance assigned to it. Like a video streaming or audio streaming may be delayed but an expected natural expected has to be informed on time. Systems should be made smart enough to Prioritize and send data.

1.4 Research Questions

While I was going through my research, and following the work previously done in this area, I came across following questions that have not been touched in this area yet.

Table 1.1 Research questions

S no.	Research Questions	Tag
1	Can a UDP connection made reliable? And can it replace a TCP connection for gathering data in smart transportation?	RQ 1
2	Is it better to use JSon file or a CSV file for storing the collected data? Can a CSV file be converted to JSon file format?	RQ 2
3	How can the filtering and be done if a M2M connection is lost ?	RQ 3

Chapter 2
User Perspective and ITS

Mobility has been the basis of day-to-day living of any being. For human beings it is required for studies, health, education purposes and many more. It has become a necessity of modern age. It can be tangible and intangible; tangible in the sense physical mobility and intangible refers to movement of thoughts, ideas and information.

Holding a great potential to change the people's daily experiences, ITS can be exploited to bring in concepts that can uplift weaker sections of the society. For example, Real-time detection of crop productivity with respect to monsoon can help farmers in having a broader perspective of his farm's yield Tracking and monitoring capabilities of an ITS system can help in better preparedness against disasters.

2.1 Assurance in mobility

Security and safety becomes a challenge when it comes to providing human mobility in remote areas. These areas are usually disconnected or have poor network connections. This can be an obstacle in inclusive health facilities, education, leisure activities etc. ITS helps in overcoming these barriers.

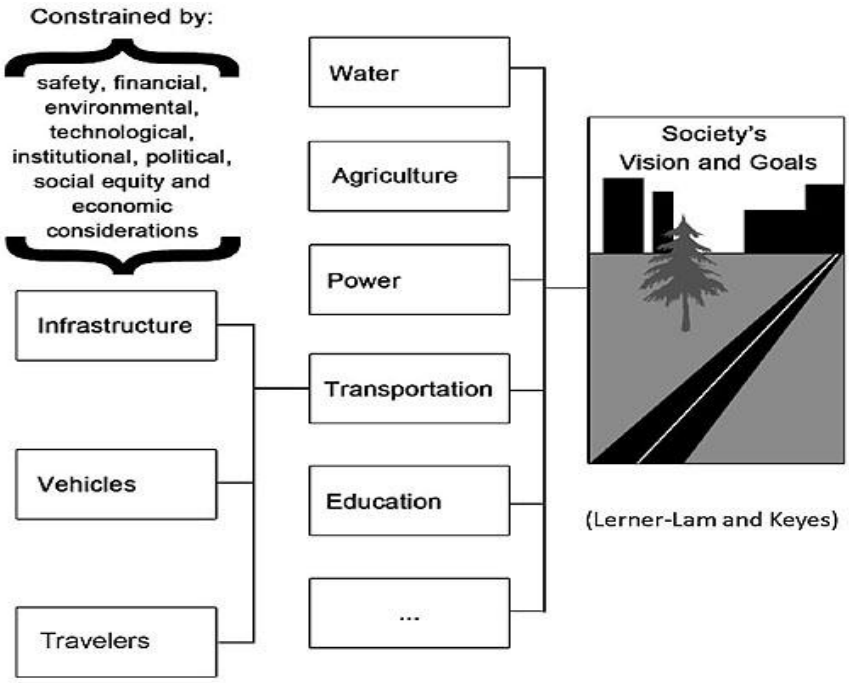

Figure 2.1 ITS and user expectations

While overcoming these barriers, it is important to know the purpose of whole journey. Here context becomes important as it will guide in the routes to be taken, frequency of travel and user can be made available with suggestions whether to travel or not. Here information becomes very important as it is a remote and unfamiliar location. For vulnerable groups like elderly and women this acts as a decisive factor and encourages them to travel if their safety is assured.

2.2 Mode of transportation.

Demographic groups prefer different modes of travel. Like men prefer private travel, on the other hand women prefer a public mode of transport as it is associated with large number of people travelling together. This brings a sense of security in them. But as huge number of people are travelling together it also leads to increase in crime rate . ITS helps in reducing the uncertainty as these groups will be assured that there travel is being tracked and even a delay of shorter spans will lead to alarming the authorities that will reach on the spot. Hence this will increases there travelling and lead to their social up-liftment.

2.3 Challenges before ITS:

No doubt ITS has given many opportunities for technological development but it still has its weaknesses. Firstly collecting enormously huge amount of data and then augmenting it with different user's requirements. This makes processing data user dependent. Hence, ITS is to be mixed with artificial intelligence. The already fragmented data that has been collected to track user movement has to be read and drawn conclusions from. As privacy and ITS both are very dynamic this leads to circular relationship between two. Although an individual has given guaranteed anonymity but this works only in groups. The truth is that individual's data is being recorded and his privacy is at stake. The users have become so dependent on these technologies on daily basis that their decision-making capability is compromised

Chapter-3
Literature Review

In paper [3] Smart city transportation and its various aspects are discussed. It also talks about problems dealing with its efficiency.

Paper[4] talks about various data collection techniques. High quality of data is required from the vehicular movement in order to provide efficient and most reliable results.

Paper [5] presents about art in data gathering and the diversity of data that can be encountered in smart transportation. It speaks on why it becomes necessary to understand the forms of data and errors that occur in data gathering and how can they be overcome.

Paper [6] talks about the real timeliness of the entire ITS. It discusses how already existing technologies in this area can be augmented and connected to advanced architectures. This will help in building efficient systems that are cost-effective

Paper [8] proposes a GMM based fusion technique which can be used in order to keep the connectivity and reliability alive. The presenters have done a great research in this field and have clearly mentioned the strengths and weaknesses of this method.

Through [15] and [17], I found out that data collected from the sensors can best be collected in CSV file format. As CSV file format is platform dependent and can only be exploited in not more than 2 ways, it becomes important to think of a different file format.

[16] talks about accuracy of pneumatic road tube counters. It gave me a brief idea on how no. of vehicles are counted and what happens if sensor stops working for a while.

[18] and [19] gave me an idea about how loop detectors function. It gave me the information about the data format that is transferred and why is it that a particular data format is chosen over the rest.

[20] suggested a new method of traffic detection through video streaming. It is an efficient way but it takes your focus towards the cost efficiency.

Summary and Problem analysis: After surveying all the papers we can say that huge amount of heterogenous data from various sensors is generated in smart transportation environment which need to be converted into common format and some unified framework is required. Paper [3] presents smart city, smart transportation scenario,[4] discuss about possible data collection methods.[7][10][8] discuss about various ways of filtering data and fusion of all collected data via various approaches.[14][13][12] all these papers discuss about the repository to be used for storage of processed and unprocessed real time data coming from various sensors.

The unified framework should cover all functionality required starting from data collected from sensors up to storage in database. And also the framework should less total processing time and better filtration accuracy so as correct data is available faster for further data analysis.

Table 3.1 Literature Review of data collection module

Sr no	Title	Authors	Publication & year	Summary
1	The Internet of Things : A survey	L. Atzori, A. Iera, and G. Morabito	Elsevier, 2010	Address IoT with integration of several technologies and communications solutions, different visions of IoT paradigm are reported and enabling technologies reviewed.
2	Overview of Data Collection Methods for Intelligent Transportation Systems	R. Prabha and M. G. Kabadi	IEEE, 2016	Here paper provides the overview of data collection methods for Intelligent Transportation Systems
3	A Survey : Internet of Things (IOT) Technologies , Applications and Challenges	S. H. Shah	IEEE ,2016	This paper discusses IoT in wider sense, its architecture, technical aspects, protocols and applications related issues with comparison of other survey papers.
4	Internet of things : Vision , applications and research challenges	D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac	Elsevier, 2012	Survey on various technologies, applications and research challenges for IoT are discussed.
5	A literature survey on smart cities	Y. I. N. Chuantao, X. Zhang, C. Hui, W. Jingyuan, C. Daven, and D. Bertrand	Springer, 2015	Smart city vision, its domain and technologies required.

Table 3.2 Literature review of Processing module

Sr no	Title	Authors	Publication year	Summary
6	Data Fusion for ITS: Techniques and Research Need	N.-E. El Faouzi and L. A. Klein	Elsevier, 2016	A survey of Intelligent Transportation System DF applications and directions for future research are discussed.
7	Kalman Filter and Its Application	Q. Li, R. Li, K. Ji, and W. Dai	IEEE, 2015	Briefly surveys recent developments about Kalman filter (KF), Extended Kalman filter (EKF) and Unscented Kalman filter (UKF) and basic theories of Kalman filter are discussed.
8	Fusing Heterogeneous Traffic Data by Kalman Filters and Gaussian Mixture Models	C. Wang, Q. Zhu, Z. Shan, and Y. Xia	IEEE , 2014	Here a fusion method of heterogeneous traffic data based on the Kalman filters (KF) and Gaussian mixture models (GMM) is proposed. The noise in collected raw data is reduced by the KF to improve the quality of input data for fusion.
9	Performance Comparison of Filtering Techniques for Real Time Traffic Density Estimation Under Indian Urban Traffic	B. Dhivyabharathi, S. Fulari, R. Amrutsamanvar	IEEE, 2015	This paper discuss, model based approaches for the estimation of traffic density are discussed. The non-linear model equations are based on the conservation principle and the fundamental traffic flow.

Chapter-4
Proposed Work

4.1 Overview of proposed work:

After conducting literature survey and problem identification in this section I have proposed a framework (shown below) which can provide proper data collection and processing of smart transportation data.

Here in this framework key issues like handling data heterogeneity by converting it to common format, proper way of collecting data, filtration of data and its storage are addressed.

4.2 Description of above framework modules:

Sensors: This module comprises of various sensors used in the smart transportation environment. The various sensors used as shown above are Pneumatic tube, Inductive loop, Co2, Fog, GPS, Video camera, RFID.

1) Pneumatic tube: The data received is in CSV format. It generally gives information on number of vehicles.
2) Inductive loop: It outputs data in CSV format. The accuracy of inductive loop is ±2.96%. It is mainly used for counting number of vehicles.

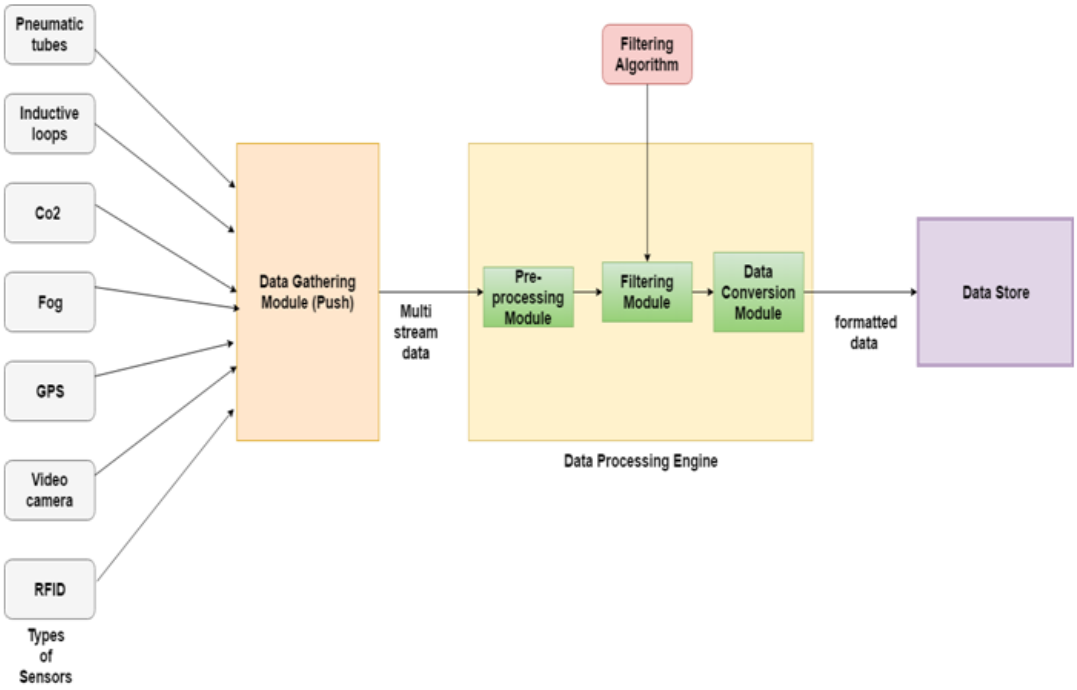

Figure 4.1 Framework for data collection and its storage in processed manner for smart transportation.

3) Carbon dioxide sensor: It sends data to data gathering unit in PPM (particles per minute) format.. It is mainly used to find amount of co2 in atmosphere.
4) Fog sensor: It gives data to data gathering unit in PWM (pulse width modulation) format. The accuracy of fog sensor is ± 20%.
5) GPS: It gives data to the data gathering module in NMEA format. The accuracy of GPS sensor ±7.8 meters. Used to measure speed, know current location of device.
6) Video camera: It gives data to data gathering module in mp4 or frames per second etc. The accuracy of video camera is ±3 %. Used to measure speed of vehicle, traffic density.
7) RFID: It gives data to data gathering module

From above details, we can say that data coming from sensors is of varied format and heterogeneous in nature.

Data Gathering Module:

This module collects data from various sensors. Push method is used to collect data in which time to time data is sent to the module by sensors when some data is generated. Hence data gathering module do not pull data from sensors.

Data Processing Engine:

This module serves as the core part of the framework which does the whole processing job. It is divided into three sub modules i.e. Preprocessing module, Filtering module and Data conversion module.

The **preprocessing module** is used to make data ready for filtration. As we know data is coming from numerous sensors so there are chances that data may come with some delay for example buffering in data coming from video camera. So, in such cases preprocessing module holds that data until complete data is arrived. Now this preprocessed data is suitable for filtering.

The **filtering module** takes the preprocessed data as input and filters out the unwanted data or outliers from the present data before it gets converted into a common format. For purpose of filtering an algorithm as shown in the diagram is given as input to the filtering module. The choice of algorithm is important because filtering efficiency of that algorithm is an important factor to be considered. The algorithm must also be feasible enough to filter varied heterogeneous data coming from all sensors (means should be generic).

Now as discussed, several times throughout this research heterogeneous data is required to be converted into common format. The next module **Data conversion module** does the above task of converting data into common format. For conversion purpose, different logic is applied to different type of data, means logic differs as per data formats.

Data store:

After getting the processed data there is requirement of storing data efficiently so as it can be retrieved at the time of data analysis. So using this module proper storage of data can be done.

4.3 Measuring Parameters:

Mainly the measuring parameters that we have focused are related to time required for operation completion. As per the requirement and various constraints of IoT environment, the parameters to be considered are:

1) Filtration Accuracy: It is the measure of correctness of the filtration algorithm used for filtering sensor data after preprocessing.
2) Total Processing time: It is the measure of total time taken to process raw data up to storage in the database.
   It can be calculated by taking sum of individual time required in each phase of the framework.

So, focusing on above parameters we will try to get better filtration accuracy and less total processing time.

Chapter- 5
Implementation and Results

Each module has been implemented certainly and explained as follows. The implementation has been done using java in eclipse and references are added in Microsoft word.

5.1 Implementation:

Here I have discussed step by step logic followed in each module.

5.1.1 Client-Server

As I have proposed replacing TCP with a UDP connection, I have tried to make UDP reliable. I have not made my system dependent clearly on UDP as I have applied filters later, but to much extent UDP has been made reliable. Fig. 5.1 talks about the client side.

Fig. 5.2 is a

```
Scanner scanner = new Scanner(System.in);
    System.out.println("Want to  send more from GPS Sensor(Yes/No) ??");
String sendMore = scanner.next();
while(sendMore.equals("Yes")){
    System.out.print("Enter Data :: ");
    String str12 = scanner.next();
    System.out.println("String data"+str12);
        ps.println(str12);
        System.out.println("Want to  send more from GPS Sensor(Yes/No) ??");
    sendMore=scanner.next();
}
```

Figure 5.1 Client Side

Fig 5.2 is about the server side.

```
import java.io.BufferedReader;

public class ServerSide {

    public static void main(String args[]) throws Exception {
        ServerSocket serverSocket = new ServerSocket(7777);
        System.out.println("Hello");
        Socket socket = serverSocket.accept();
        System.out.println("Connection Established !!");
        OutputStream obj = socket.getOutputStream();
        PrintStream ps = new PrintStream(obj);
        String str = "Hello Client";
        String str2 = "I am ready to listen ";
        BufferedReader br = new BufferedReader(new InputStreamReader(socket.getInputStream()));
        while (true) {
            String str1;
            while ((str1 = br.readLine()) != null) {
                System.out.println(str1);
                break;
            }
        }
    }
```

Figure 5.2 Server Side

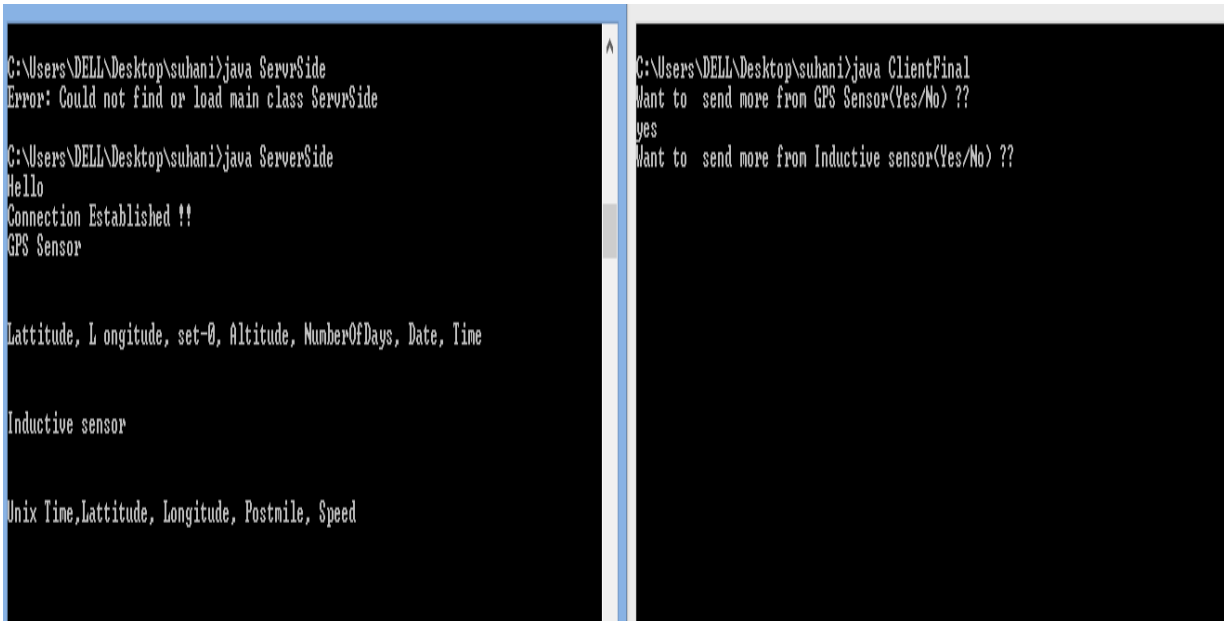

Figure 5.3 Client-Server connection

Once the connection is established, server will keep on asking if there is more data to send. This adds the reliability factor. If client wants to send more data it can simply send the data without first acknowledging. This decreases latency.

5.1.2 Filtration of preprocessed data:

For filtration of preprocessed data suitable filtering algorithm is the key decision. Currently as per the study conducted by me up till now and discussed in some papers [9] Kalman filter can be a suitable . Kalman filter is iterative mathematical process that uses set of equations and consecutive data input to quickly estimate true value, position velocity etc. this is chosen as it is better for real time smart transportation environment.

5.1.3 CSV to JSon:

In the diagram below code to convert inductive loop data from CSV format to Json is shown. CSV file format has certain limitations as talked about in literature review. It can't be used

```
import java.io.File;

public class CSVTOJSON {
    private static String csvFileName = "resources/Data.csv";  //Change as per need.
    private static String jsonFileName = "resources/Data.json"; //Change as per need.

    public static void main(String[] args) throws IOException {
        File csvFile = new File(csvFileName);
        File jsonFile = new File(jsonFileName);

        List<Map<?, ?>> data = readFromCsvFile(csvFile);
        writeInJsonFile(data, jsonFile);
    }

    public static List<Map<?, ?>> readFromCsvFile(File csvFile) throws IOException{
        CsvSchema bootstrap = CsvSchema.emptySchema().withHeader();
        CsvMapper csvMapper = new CsvMapper();

        MappingIterator<Map<?, ?>> mappingIterator = csvMapper.reader(Map.class).with(bootstrap).readValues(csvFile);

        List<Map<?, ?>> data = new ArrayList<>();
        Map<?, ?> innerData = null;

        while(mappingIterator.hasNext()) {
            innerData = mappingIterator.next();
            data.add(innerData);
        }

        return data;
    }

    public static void writeInJsonFile(List<Map<?, ?>> data, File jsonfile) throws IOException {
        ObjectMapper mapper = new ObjectMapper();
        mapper.writeValue(jsonfile, data);
    }
}
```

Figure 5.4 CSV to JSon

Chapter 6
Conclusion

As proposed in introduction and literature review, I was able to find answers to the glitches. A UDP connection can replace a TCP connection if we apply suitable filters to the system. Filters will do the work where UDP lags in TCP i.e. whenever the connectivity is lost for a certain time, it will provide data to the system by predicting it from the previous value. Also the problems that accompanied the CSV file format produced by the sensors can be repealed by converting it to JSon file format.

6.2 Answers to Research Questions

Through my work, I was able to conclude the research questions asked in the introduction.

6.2.1 RQ 1 : Can a UDP connection replace a TCP connection for gathering data in smart transportation?

Yes. Through this research I found that a TCP connection can be replaced by a UDP connection. A UDP connection lacks only in reliability. What if a communication is lost? This has been answered by using kalman filters in place. Whenever the connection will be lost, kalman filters can take over and use previous data to speculate the further results till the connection is restored.

As long as connection can be restored in just a matter of time as is the case with UDP, kalman filters will do its job. With UDP, we get access to a faster network and

a simpler network. In ITS, it becomes important to have a fast communication because of the liabilities attached to it.

6.2.2 RQ 2 : Is it better to use JSon file or a CSV file for storing the collected data? Can a CSV file be converted to JSon file format in M2M communication?

It is always better to use a Json file format because of following reasons:

1. It is an efficient way of communicating with the machines.
2. The managers posted in the nodes of smart transportation may not always be smart enough to write and communicate in CSV file format. So they can just right in JSon format and communicate. This will help the investors to save cost by hiring unskilled workforce and in turn provide employment opportunities to the unskilled
3. The JSon file format can be used on any platform as it is platform independent.
4. It is used for sending serialized and structured data over the networks. Using JSon keeps the networks structure alive and at the same time maintains the simplicity.

Coming to our next question that whether a CSV file format be converted to a JSon file format for M2M communication? Through my research the answer luckily came out to be yes. In my implementation part, I have implemented the same.

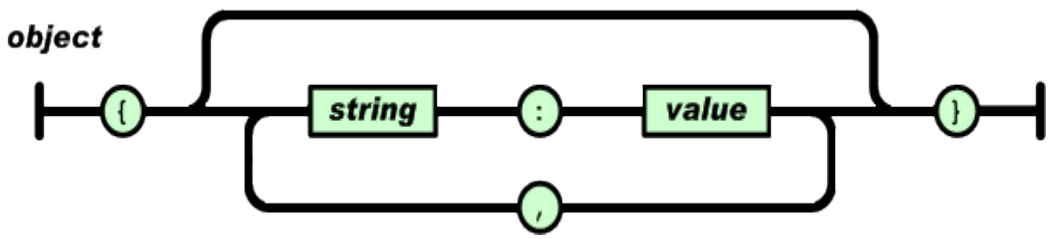

Figure 6.1 Json file format

6.2.3 RQ 3 : How can the filtering and be done if a M2M connection is lost ?

The problem in smart transportation arises when communication is lost. The ITS still has to keep working to keep its credibility alive. Through my research I found that there are only two ways to do so :

- To speculate the results further. This is mere guesswork done by the system and has no link to the ongoing real situation.
- To come to definite conclusion based on previous data. This makes the results closer to what a situation looks like.

For this I found out that kalman filters did a great job as :

- It is a statistical technique. Hence it makes the random data act in an organised way.
- It takes the variance of initial state and model error. This has been neglected in all other filters.
- It executes in real time i.e. there is no hassle of storing the previous to previous data. Data storage of only one previous step is required.