1. PROJECT SYNOPSIS

The International Traffic Safety Data & Analysis Group (IRTAD) conducts surveys to get aware of the several reasons leading to the accidents.

Few of the very concerning problems leading to accidents are mentioned below:

1. A very recent headline in an English newspaper revealed that only three of England's 32 ambulance services reach a large majority of 'immediately life-threatening' call-outs within eight minutes, according to the latest statistics. There are more than thousands of people who lose their lives struggling in ambulances to reach the hospitals every day.

2. Most of the hospitals and schools are located at some random places in the country. This makes the schools to be located in the areas which are more prone to the accidents and where the most of the fatalities are related to kids.

3. The process of notifying the police is very slow and it usually takes a lot of time for the police to reach the spot after the occurrence of the accidents.

4. An article dated November 22, 2011, published in The Hindu in India states that 70 per cent of the road accidents happened in India in 2011 was due to drunken driving.

5. There is no process to notify the concerned car insurance agent, life insurance agent as well as the blood banks in case of medical emergencies.

Our idea is to solve many such problems by performing an analysis on the collected data. This analysis would use machine learning and data mining concepts to reveal certain hidden information.

The proposed solutions for each of these problems are as follows:

1**.** The drivers of ambulances would get notifications about accidents so that they can reach the spot faster. These ambulance drivers, instead of being at random places in the city, can be suggested some of the places which are more likely to have more accidents. This would reduce the time taken for the ambulance to reach the spot.

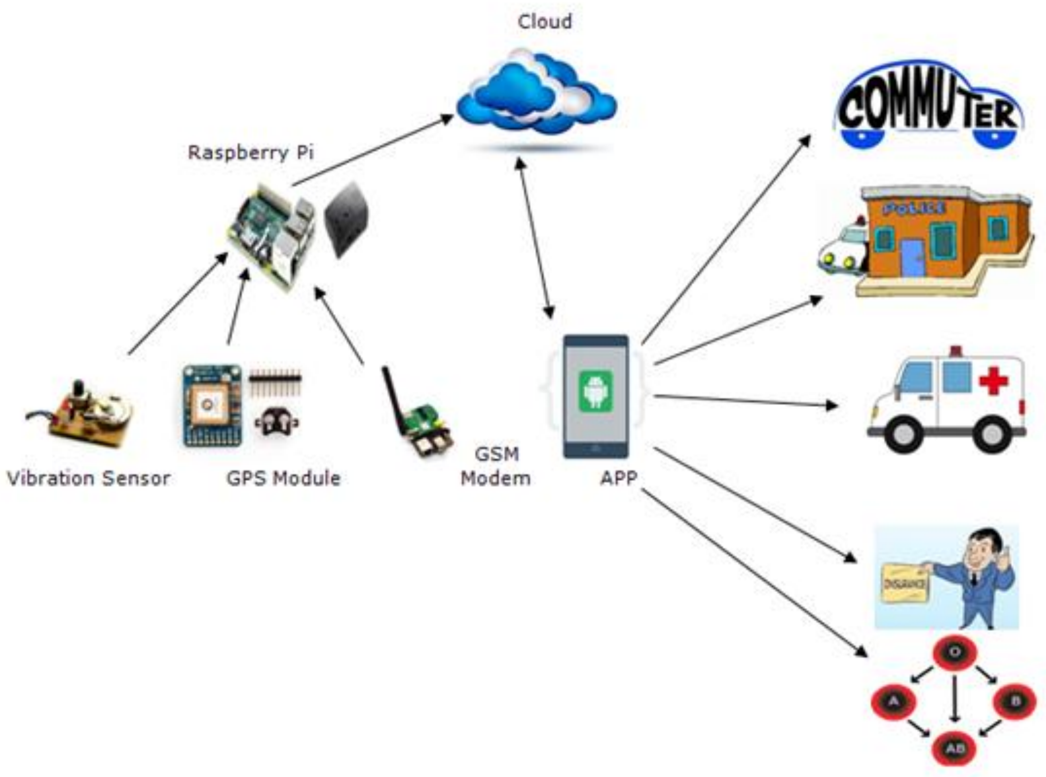
2. The setup and relocation of hospitals and schools easier by suggesting them the best place in the city based on the results of analysis. They can obtain the locations where the maximum and the minimum number of accidents are happening.

3. This spot of accident is tracked and sent to the police for the further investigation. The *GPS module* for Raspberry Pi allows to connection to the *Raspberry Pi* board to get position, altitude, speed etc. The GSM module is used for sending the location of the accident to the police.

4. We can find out the locations where most of the accidents are caused due to the consumption of alcohol. This feature would be helpful for the cab companies to know the areas where people are more likely to book their cabs after consuming alcohol so that they can always keep some cabs waiting in those areas. Another idea could be to introduce cabs meant only for the picking and the dropping of the drunken people.

5. The concerned car insurance and life insurance organizations can also be intimated with the notifications regarding the accidents. The blood banks can also be notified in case of any medical urgencies leading to the requirement of blood.

We plan to detect the accident, analyze it and notify the concerned people for further actions. The vibration sensors will be fit into the cars, which will help in the collection of the real time accident data. The most commonly used vibration sensors are piezoelectric sensors which use the piezoelectric effect to measure changes in pressure, acceleration, temperature, strain or force by converting them to an electrical charge. The location could be tracked using the GPS module with Raspberry Pi and the data is stored onto the cloud. The GSM module is used to send messages, notifications or alerts. The app communicates with the data stored on to the cloud.



The modules for this project are mentioned below:

1. **Sensing Module** – This module consists of the following :

   a) **Vibration Sensors -** The project would use vibration sensors for the detection of accidents, which are piezoelectric sensors and use the piezoelectric effect to measure changes in pressure, acceleration, temperature, strain or force by converting them to an electrical charge. The sensors are in definite numbers, 10 to 12, the more they would be, the better would be the performance. These sensors are placed at various places in the vehicle to ensure the detection of accidents at each and every part of the vehicle.

   b) **GPS Module** - The sole idea of using GPS module is to track the location of the vehicle in case of accidents. This location is further sent to the police for the further investigation. The location is also sent to the ambulance drivers, insurance agents and blood banks. The *GPS module* for Raspberry Pi allows to connection to the *Raspberry Pi* board to get position, altitude, speed etc.

c) **GSM Modem** – The GSM module is used for sending the location of the accident to the police, ambulance drivers, insurance agents and blood banks. GPRS/*GSM Module* (SIM900) offers GPRS connection to *Raspberry Pi*, send SMS, make calls and create TCP / UDP sockets.

2. **Monitoring Module:** The sensors send the real-time data to cloud and the cloud can communicate with the app also.

3. **Analysis Module (Cloud Infrastructure)**: The huge data collected from the vehicles using the sensors are analyzed by using data analysis concepts and certain predictions are done using the machine learning techniques. The decision tree algorithm could be used to determine the severity of the accidents. The clustering algorithms could be used to obtain better insights. The mispredictions could be minimized by using pruning techniques to trim the levels of the decision tree and produce optimum results.

4. **Android Application Modules**

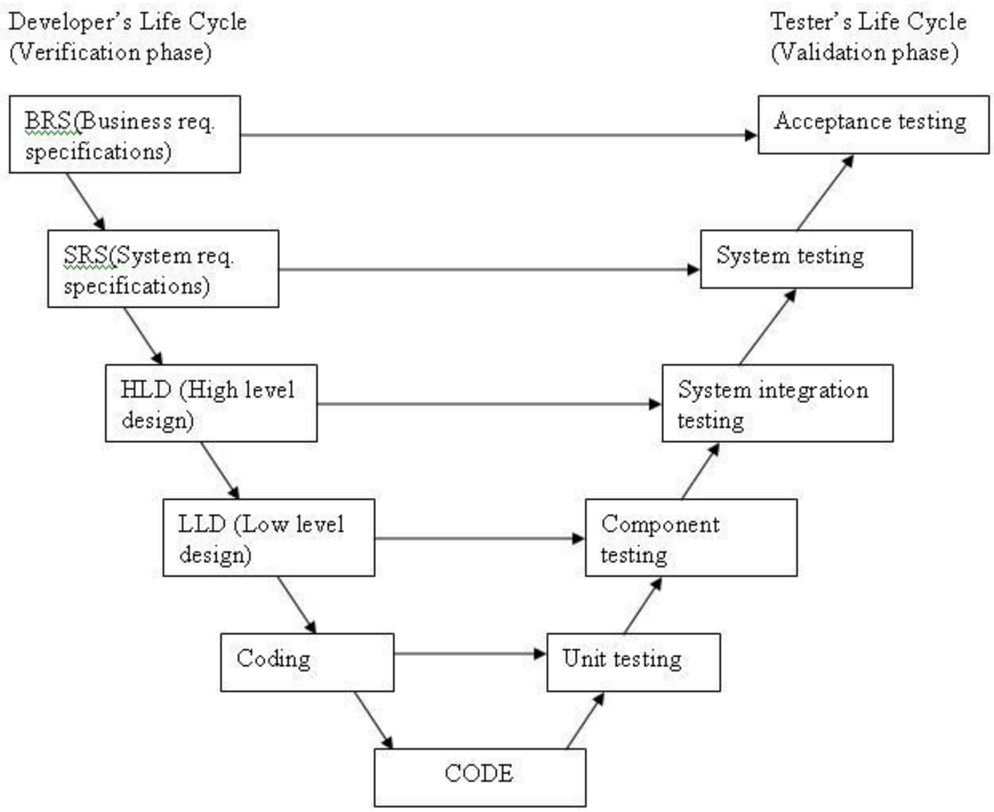
   The android app would contain lots of modules, each module dedicated to a specific functionality. Following are some of the modules listed:

   a) Redirection of traffic
   b) Alerting using App
   c) Cab drivers
   d) Notify Ambulances
   e) Inform Insurance Agents

2. PROJECT PLAN

2.1 PROCESS MODEL: V MODEL

V- Model means Verification and Validation model. Just like the waterfall model, the V-Shaped life cycle is a sequential path of execution of processes. Each phase must be completed before the next phase begins. Testing of the product is planned in parallel with a corresponding phase of development.



The various phases of the V-model are as follows:

1. **Requirements** like BRS and SRS begin the life cycle model just like the waterfall model. But, in this model before development is started, a system test plan is created. The test plan focuses on meeting the functionality specified in the requirements gathering.
2. **The high-level design (HLD)** phase focuses on system architecture and design. It provides overview of solution, platform, system, product and service/process. An

integration plan is created in this phase as well in order to test the pieces of the software systems ability to work together.

3. **The low-level design (LLD)** phase is where the actual software components are designed. It defines the actual logic for each and every component of the system. Class diagram with all the methods and relation between classes comes under LLD. Component tests are created in this phase as well.
4. **The implementation** phase is, again, where all coding takes place. Once coding is complete, the path of execution continues up the right side of the V where the test plans developed earlier are now put to use.
5. **Coding:** This is at the bottom of the V-Shape model. Module design is converted into code by developers.

Advantages of V-model:

- Simple and easy to use.
- Testing activities like planning, test designing happens well before coding. This saves a lot of time. Hence higher chance of success over the waterfall model.
- Proactive defect tracking – that is defects are found at early stage.
- Avoids the downward flow of the defects.
- Works well for small projects where requirements are easily understood.

Disadvantages of V-model:

- Very rigid and least flexible.
- Software is developed during the implementation phase, so no early prototypes of the software are produced.
- If any changes happen in midway, then the test documents along with requirement documents has to be updated.

Why V-model:

- The V-shaped model should be used for small to medium sized projects

- Requirements are clearly defined and fixed.
- The V-Shaped model should be chosen when ample technical resources are available.

2.2 PROJECT SCHEDULING: PROJECT LIBRE TOOL

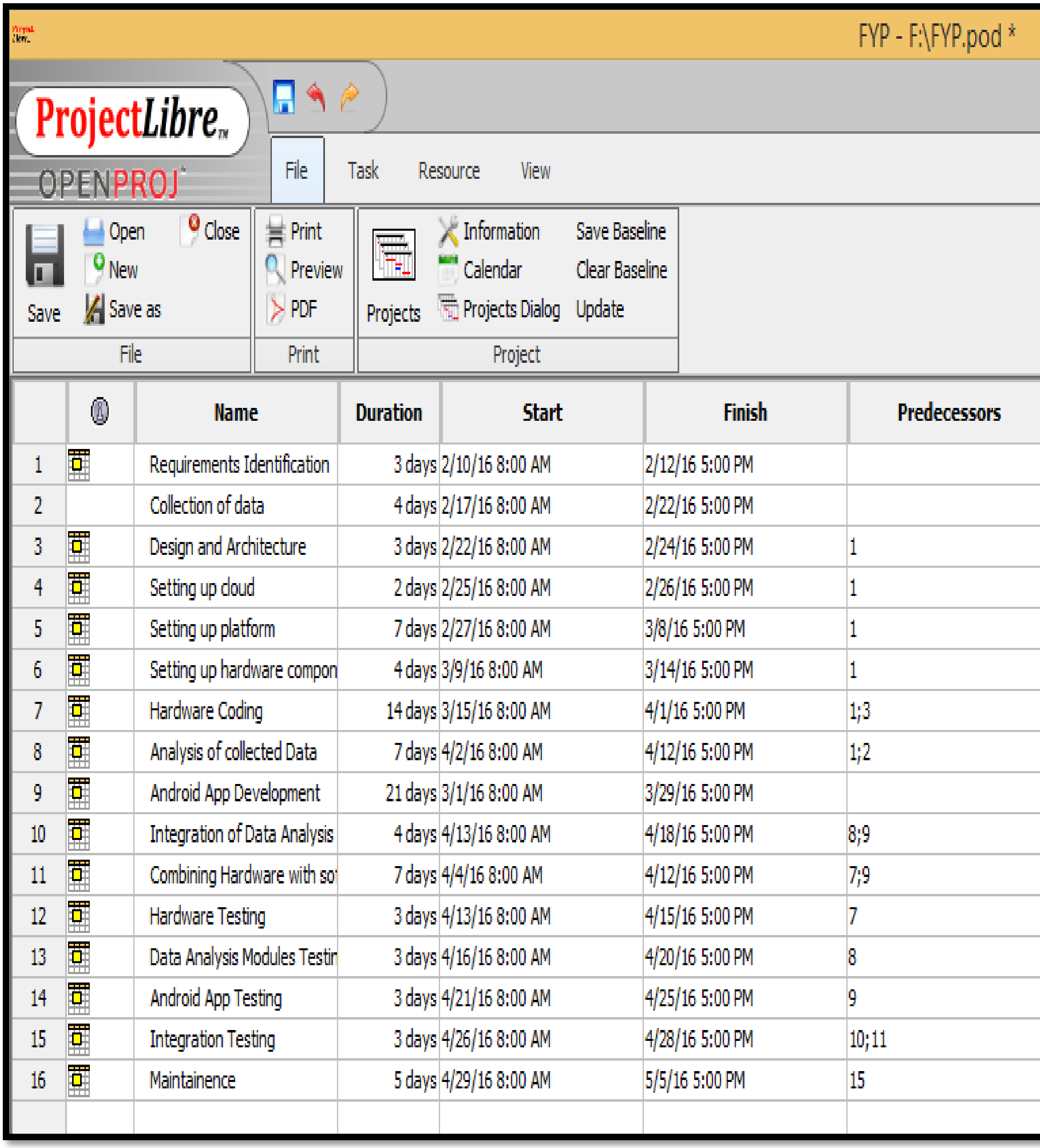
The project schedule is the tool that communicates *what work* needs to be performed, *which resources* of the organization will perform the work and the *timeframe*s in which that work needs to be performed. The project schedule should reflect all of the work associated with delivering the project on time.

Project Libre:

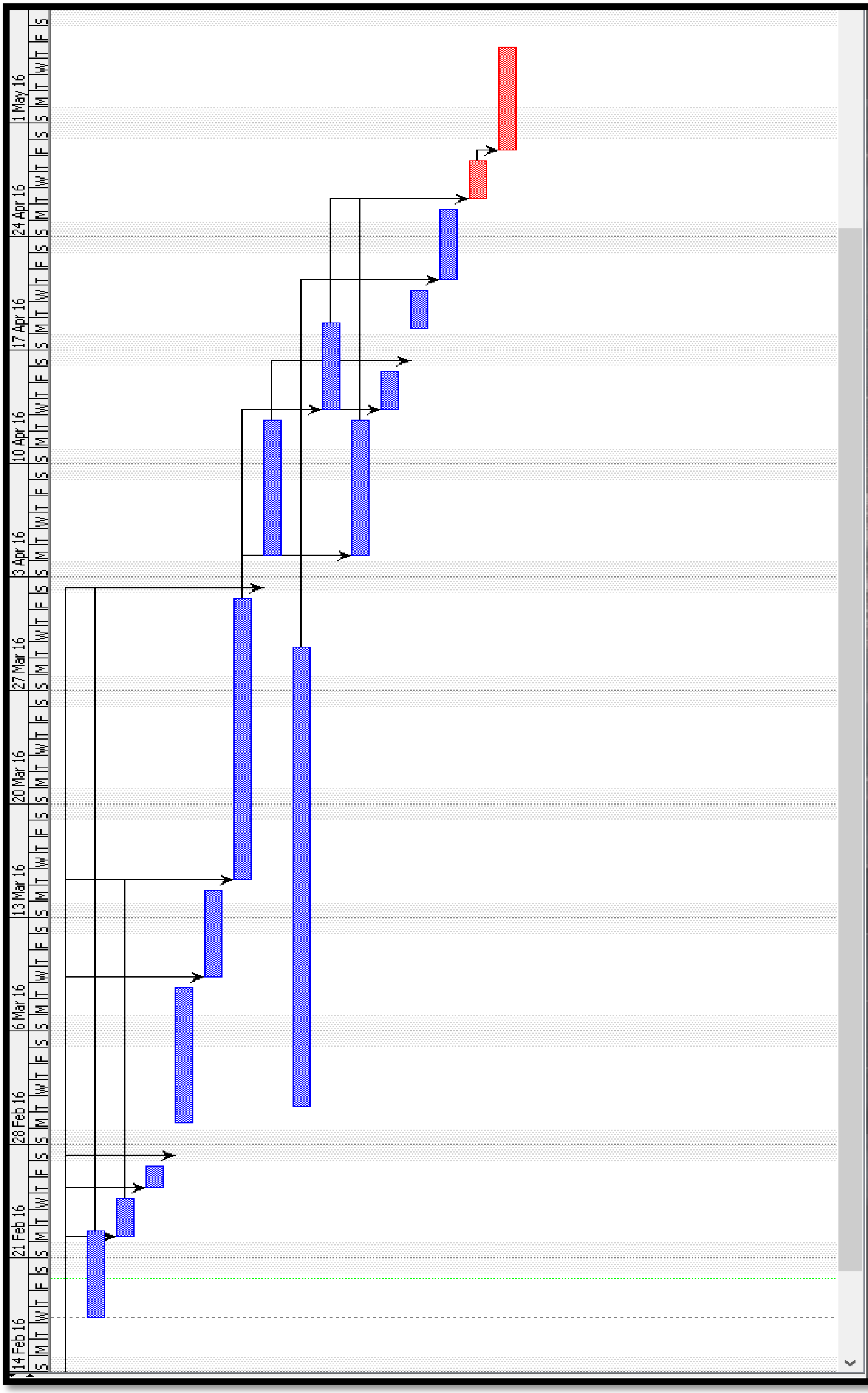
ProjectLibre is a project management software system. ProjectLibre runs on the Java Platform, allowing it to run on the Linux, Mac OS or MS Windows operating systems. It is released under the Common Public Attribution License (CPAL) and thus qualifies as Free Software according to the Free Software Foundation.

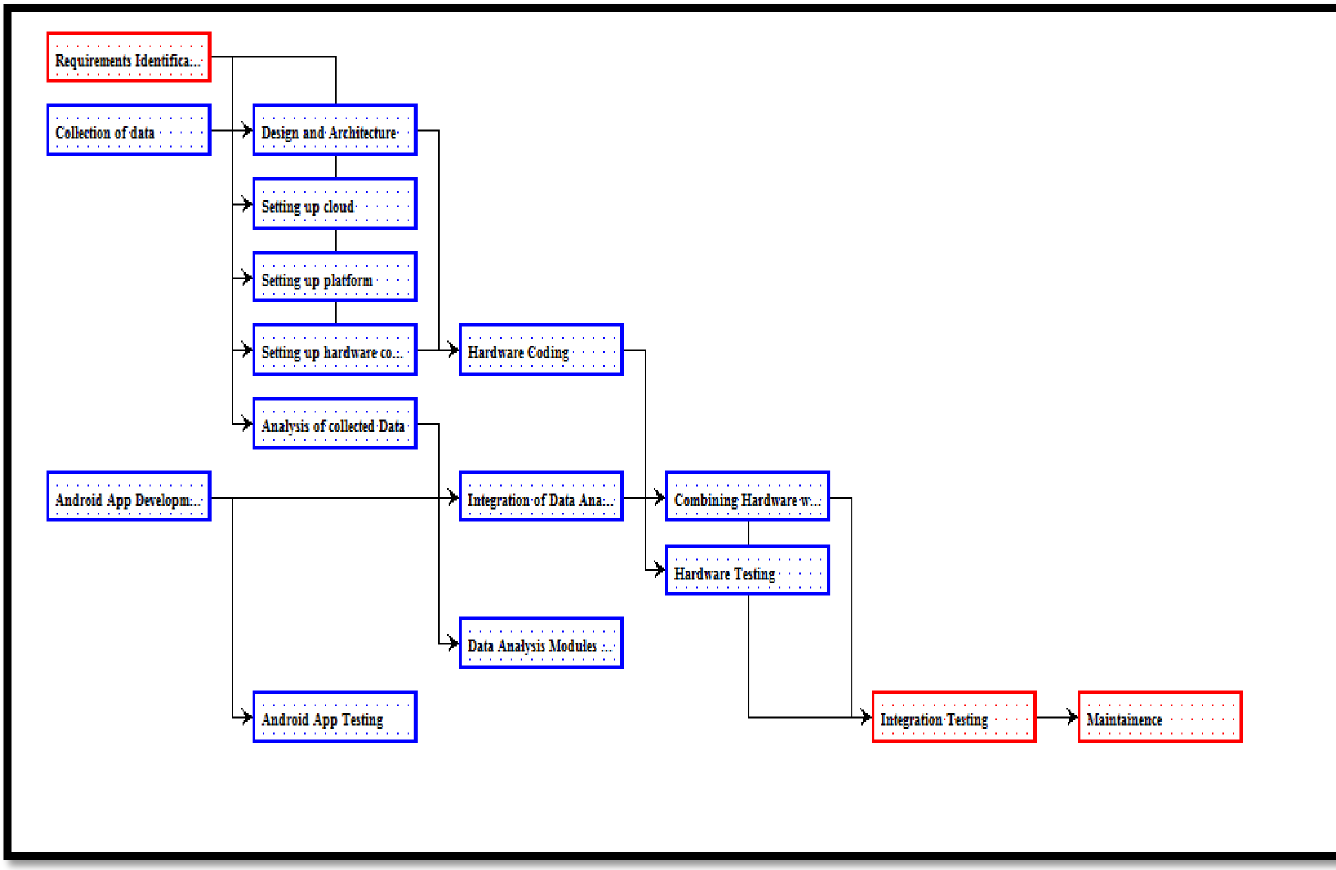
Gantt Chart:

Gantt charts allow us to visualize a project's start and end date along with each element or task that needs to be completed. Because many project tasks are dependent on previous tasks, these charts also allow us to see these dependencies and schedule around them.



	Name	Duration	Start	Finish	Predecessors
1	Requirements Identification	3 days	2/10/16 8:00 AM	2/12/16 5:00 PM	
2	Collection of data	4 days	2/17/16 8:00 AM	2/22/16 5:00 PM	
3	Design and Architecture	3 days	2/22/16 8:00 AM	2/24/16 5:00 PM	1
4	Setting up cloud	2 days	2/25/16 8:00 AM	2/26/16 5:00 PM	1
5	Setting up platform	7 days	2/27/16 8:00 AM	3/8/16 5:00 PM	1
6	Setting up hardware compon	4 days	3/9/16 8:00 AM	3/14/16 5:00 PM	1
7	Hardware Coding	14 days	3/15/16 8:00 AM	4/1/16 5:00 PM	1;3
8	Analysis of collected Data	7 days	4/2/16 8:00 AM	4/12/16 5:00 PM	1;2
9	Android App Development	21 days	3/1/16 8:00 AM	3/29/16 5:00 PM	
10	Integration of Data Analysis	4 days	4/13/16 8:00 AM	4/18/16 5:00 PM	8;9
11	Combining Hardware with so	7 days	4/4/16 8:00 AM	4/12/16 5:00 PM	7;9
12	Hardware Testing	3 days	4/13/16 8:00 AM	4/15/16 5:00 PM	7
13	Data Analysis Modules Testir	3 days	4/16/16 8:00 AM	4/20/16 5:00 PM	8
14	Android App Testing	3 days	4/21/16 8:00 AM	4/25/16 5:00 PM	9
15	Integration Testing	3 days	4/26/16 8:00 AM	4/28/16 5:00 PM	10;11
16	Maintainence	5 days	4/29/16 8:00 AM	5/5/16 5:00 PM	15


14 Feb 16
21 Feb 16
28 Feb 16
6 Mar 16
13 Mar 16
20 Mar 16
27 Mar 16
3 Apr 16
10 Apr 16
17 Apr 16
24 Apr 16
1 May 16



Advantages:

1. Allows for efficient organization
2. Helps establish timeframes.
3. Highly Visual

Disadvantages:

1. Potentially overly complex – If you've ever worked on a complex project, and looked at the Gantt chart, you know that these charts can be large and hard to read.
2. Need to be updated – Gantt charts are developed early in the planning stages of a project, there is a good chance that the project will change, thus the chart will need to be updated.
3. Don't show the whole picture.

2.3 EFFORT ESTIMATION: COCOMO Model

Software development *effort estimation* is the process of predicting the most realistic amount of effort (expressed in terms of person-hours or money) required to develop or maintain software based on incomplete, uncertain and noisy input. Effort estimates may be used as input to project plans, iteration plans, budgets and investment analyses, pricing processes and bidding rounds.

COCOMO Model:

The Constructive Cost Model (COCOMO) is an algorithmic software cost estimation model developed by Barry W. Boehm. It makes use of various parameters and a defined formula to estimate effort. It is a *Parametric Model* which accepts as input quantative and qualitative weighted characteristics and produces effort estimation.

COnstructive COst MOdel II (COCOMO II) is the latest major extension to the original COCOMO (COCOMO 81) model published in 1981.

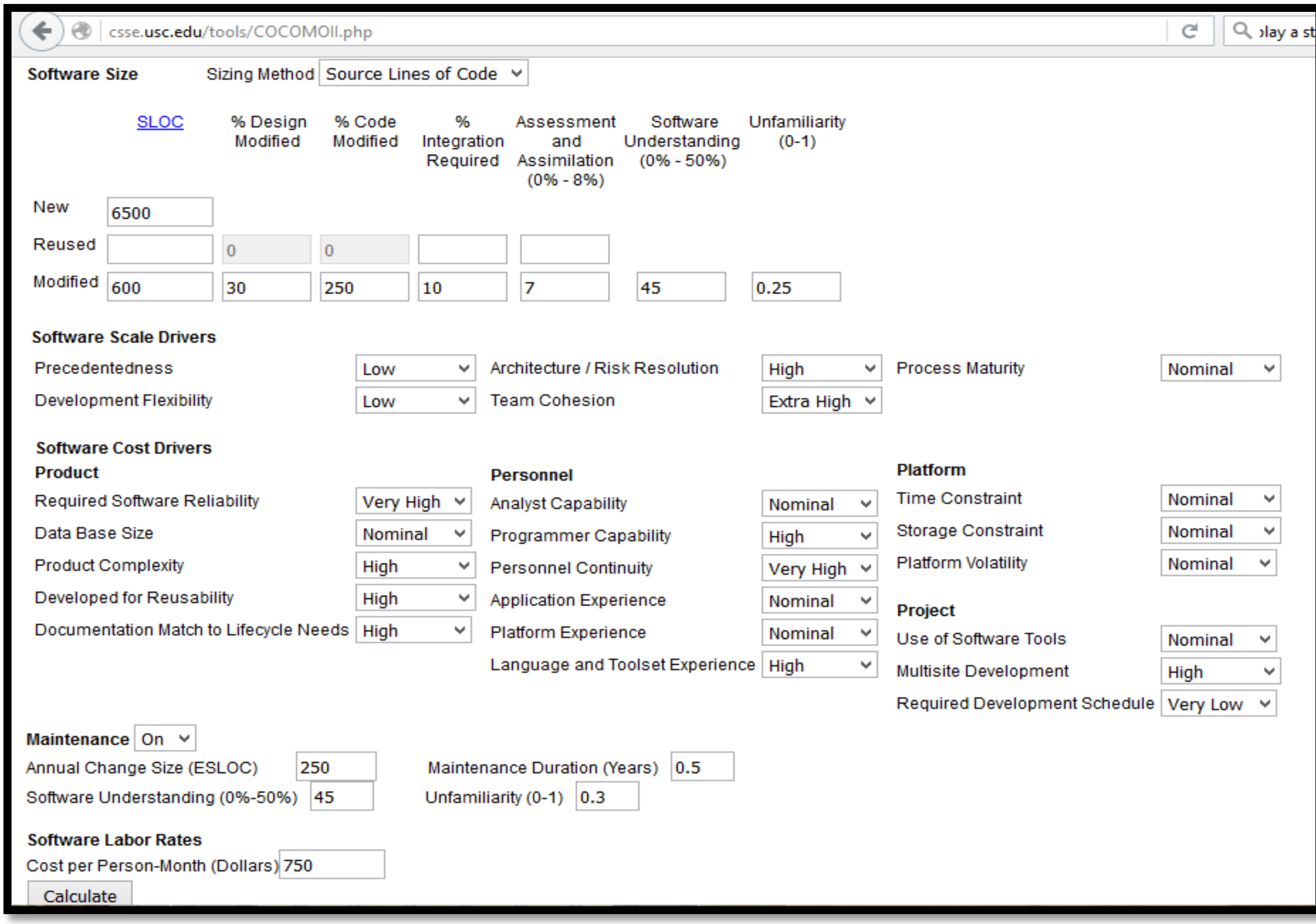


Figure: Input to COCOMO Effort Estimation Tool csse.usc.edu/tools/COCOMOII.php

Results

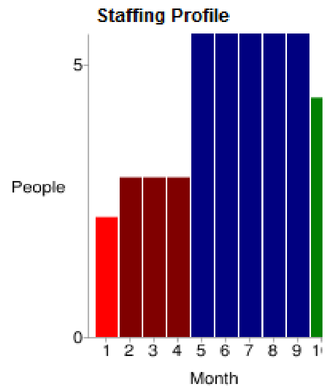
Software Development (Elaboration and Construction)

Effort = 36.8 Person-months
Schedule = 8.0 Months
Cost = $27606

Total Equivalent Size = 7149 SLOC

Acquisition Phase Distribution

Phase	Effort (Person-months)	Schedule (Months)	Average Staff	Cost (Dollars)
Inception	2.2	1.0	2.2	$1656
Elaboration	8.8	3.0	2.9	$6626
Construction	28.0	5.0	5.6	$20981
Transition	4.4	1.0	4.4	$3313



Software Effort Distribution for RUP/MBASE (Person-Months)

Phase/Activity	Inception	Elaboration	Construction	Transition
Management	0.3	1.1	2.8	0.6
Environment/CM	0.2	0.7	1.4	0.2
Requirements	0.8	1.6	2.2	0.2
Design	0.4	3.2	4.5	0.2
Implementation	0.2	1.1	9.5	0.8
Assessment	0.2	0.9	6.7	1.1
Deployment	0.1	0.3	0.8	1.3

Maintenance
Annual Maintenance Effort = 0.0 Person-Months
Annual Maintenance Cost = $0
Total Maintenance Cost = $0

Your output file is http://csse.usc.edu/tools/data/COCOMO_February_19_2016_11_03_57_647895.txt

Created by Ray Madachy at the Naval Postgraduate School. For more information contact him at rjmadach@nps.edu

Figure: Output of COCOMO Model

Advantages:

1) COCOMO is factual and easy to interpret. One can clearly understand how it works.
2) Accounts for various factors that affect cost of the project.
3) Works on historical data and hence is more predictable and accurate.
4) The drivers are very helpful to understand the impact on the different factors that affect the project costs.

Drawbacks:

1) COCOMO model ignores requirements and all documentation.
2) It oversimplifies the impact of safety/security aspects.
3) It ignores hardware issues
4) It is dependent on the amount of time spent in each phase.

2.4 RISK ASSESSMENT: USING RISK MATRIX 220

What Is Risk?

Risks are events or occurrences that prevent a program from achieving its cost, schedule, or performance objectives.

Baseline Risk Assessment Process

A good risk assessment and management process is essential to the success of any program. The process summarized here consists of

- Planning for Risk Assessment
- Identifying Program Objectives or Requirements
- Defining Program Risks
- Ranking Program Risks
- Managing Program Risks
- Managing Action Plans
- Continuously Assessing Program Risks.

Planning for Risk Assessment:

The first stage of a risk assessment is to plan the activity. We begin this process by selecting the risk assessment team, setting forth ground rules, and determining the supporting risk management tools, such as the Risk Matrix application. The risk assessment team should include representatives from all areas of the program, not just technical experts. In addition, a facilitator and recorder should be selected to assist the team. The Risk Matrix application can be used at any time.

So in our project we are using the Risk Matrix application to assess the risks involved and to identify their type, probability of occurrence, impact and mitigation plan.

Identifying Program Objectives or Requirements:

Once we have identified the risk assessment team and tools, the team Identifies the key program objectives or requirements. The program objectives and Requirements should assist the team in identifying risks.

Defining Program Risks:

A team facilitator leads the team through a structured brainstorming process to identify the Program risks. For example, each team member individually writes down risk ideas. Next, the facilitator asks each person to present one idea in sequence or pass in a rotation until all candidate risks are offered. Once all ideas are heard, an affinity diagram is created to Group, merge, and eliminate duplicate risks, and to identify dependent risks.

Following the identification of the risks, we assign various attributes to each risk. At a minimum, relevant time frame, impact, and probability of occurrence are assigned. Time frame is the beginning and end dates of when a risk may occur.

Then the team sets impact definitions. The Risk Matrix impact definitions are:

C (Critical): If the risk event occurs, the program will fail. Minimum acceptable requirements will not be met.

- **S (Serious):** If the risk event occurs, the program will encounter major cost/schedule increases. Minimum acceptable requirements will be met. Secondary requirements may not be met.
- **Mo (Moderate):** If the risk event occurs, the program will encounter moderate cost/schedule increases. Minimum acceptable requirements will be met. Some secondary requirements may not be met.

- **Mi (Minor):** If the risk event occurs, the program will encounter small cost/schedule increases. Minimum acceptable requirements will be met. Most secondary requirements will be met.
- **N (Negligible):** If the risk event occurs, it will have no effect on the program. All requirements will be met.

Probability of occurrence is the team's assessment of the likelihood of a risk happening.
For this purpose, we have estimated probabilities using a relative scale:

- **0-10%:** very unlikely the risk will occur
- **11-40%:** unlikely the risk will occur
- **41-60%**: even likelihood the risk will occur
- **61-90%**: likely the risk will occur
- **91-100%**: very likely the risk will occur.

We have identified the various risks, their impact and probability of occurrence and depicted in the risk matrix.

Risk No.	RISK	Event Type	I	Po (%)
1	Durability and quality of sensors	Performance and Cost	S	30%
2	Connectivity at accident site	Performance	S	30%
3	Personnel shortfalls	Schedule	S	40%
4	Unrealistic schedules and	Schedule and Cost	C	20%
5	Develop wrong software functions	Performance, Cost and Schedule	C	20%
6	Develop wrong user interface	Performance,Cost and Schedule	S	30%
7	Frequent requirement changes	Schedule	Mo	20%
8	Shortfalls in software tools	Performance	Mo	30%
9	Crash or inconsistency in database	Performance	C	20%
10	Crash in website or app	Performance	S	20%
11	Programmatic risks	Performance, Cost	S	20%
12	Storage and memory constraints	Cost	Mo	5%
13	Failure of sensors	Performance	Mo	10%

Ranking Program Risks

After the above processes, we have all the information needed to rank the risks. Since we are using the Risk Matrix tool, this process becomes simple and automated. Risk rankings are calculated using the Borda voting method. The Borda method *ranks* risks from most-to-least critical on the basis of multiple evaluation criteria.

We have identified the various risks, their rankings and ratings depicted in the risk matrix.

Risk No.	RISK	Event Type	I	Po (%)	Borda Rank	R
1	Durability and quality of sensors	Performance and Cost	S	30%	1	H
2	Connectivity at accident site	Performance	S	30%	1	H
3	Personnel shortfalls	Schedule	S	40%	0	M
4	Unrealistic schedules and	Schedule and Cost	C	20%	2	H
5	Develop wrong software functions	Performance, Cost and Schedule	C	20%	2	H
6	Develop wrong user interface	Performance,Cost and Schedule	S	30%	1	M
7	Frequent requirement changes	Schedule	Mo	20%	8	M
8	Shortfalls in software tools	Performance	Mo	30%	5	M
9	Crash or inconsistency in database	Performance	C	20%	2	H
10	Crash in website or app	Performance	S	20%	6	M
11	Programmatic risks	Performance, Cost	S	20%	6	M
12	Storage and memory constraints	Cost	Mo	5%	9	L
13	Failure of sensors	Performance	Mo	10%	2	M

Managing Program Risks:

All risks need some form of management, whether it involves a plan for handling risks (action plan) or merely keeping watch.

After the risks are ranked, we identify the risks that are high priority, need to be managed, and require resources (we identify the top *N* risks).

Decisions for handling the top *N* risks will vary. Some risks will be eliminated because the requirements changed; others will be transferred to other organizations (like a contractor) for action because the program team does not have the proper resources to handle the risks or because it's more appropriate for the other organization to handle them; and others will require mitigation strategies. The remaining risks (the ones not in the top *N*) should be watched.

Managing Action Plans:

We develop detailed action plans and enter an initial status, assign the Office of Primary Responsibility (OPR), and determine exit criteria for the top *N* risks. The status of each action plan should be reviewed and assessed periodically (approximately each week or month if possible), and the risks rankings adjusted accordingly.

The Risk Matrix tool provides an Action Plan worksheet in the Advanced Mode for tracking risk action plans and adjusting the risk rankings based on the action plan status.

We have identified the various risks manage and mitigation plans in risk matrix.

Risk No.	RISK	Event Type	I	Po (%)	Borda Rank	R	Manage/Mitigate
1	Durability and quality of sensors	Performance and Cost	S	30%	1	H	1)Use good sensors which are ressistant to damage
2	Connectivity at accident site	Performance	S	30%	1	H	1)Using a good network provider
3	Personnel shortfalls	Schedule	S	40%	0	M	1). Adequate training 2).Rotate jobs among team members 3).Team building sessions 4).Documentation for each team member
4	Unrealistic schedules and	Schedule and Cost	C	20%	2	H	1).Identify critical path and convince clients
5	Develop wrong software functions	Performance, Cost and Schedule	C	20%	2	H	1).Constant feedback from clients 2).Proper monitoring strictly according to SRS
6	Develop wrong user interface	Performance,Cost and Schedule	S	30%	1	M	1).Regular feedback from clients and users
7	Frequent requirement changes	Schedule	Mo	20%	8	M	1).Convince client about the severe impact on schedule 2).Allow changes upto a particular phase
8	Shortfalls in software tools	Performance	Mo	30%	5	M	1).Detailed analysis of currently available software tools with advanced features 2).Selection of the most reliable and popular tool
9	Crash or inconsistency in database	Performance	C	20%	2	H	1).Proper designing of database 2).Estabilish back-up database
10	Crash in website or app	Performance	S	20%	6	M	1).Quick fixing of technical bugs 2).Proper designing of database
11	Programmatic risks	Performance, Cost	S	20%	6	M	1).Choosing of appropriate programming language 2).Choosing of appropriate database 3).Frequent and accurate testing
12	Storage and memory constraints	Cost	Mo	5%	9	L	1).Accurate estimation of LOC 2).Use latest hardware
13	Failure of sensors	Performance	Mo	10%	2	M	1)Regular inspection of sensors

Continuously Assessing Risks

We have to continuously assess risks is essential to develop a good program risk management. As a program progresses, risks should be reassessed periodically (perhaps monthly or quarterly) to determine whether their level of importance has changed or whether new risks have developed that should be identified, assessed, ranked, and managed.

3. SOFTWARE REQUIREMENT SPECIFICATIONS

3.1 PRODUCT/PROJECT OVERVIEW

Hundreds of accidents go unnoticed and unattended by the police all over the world. This is due to the absence of a mechanism, which can notify all the concerned authorities such as the police, nearest hospitals, insurance agents. Things haven't changed much in the context of accidents in the last few decades.

Our product is *Accident Detection Kit,* which has many sensors, each dedicated to the sensing of certain parameters which help in the detection of accidents. The kit has a Raspberry Pi as System on Chip (SoC) and some of the sensors such as Vibration or Shock sensors, Tilt sensors, Fire and Smoke sensors, Pressure sensors etc.

On the occurrence of accident, the location of the vehicle is tracked in terms of latitude and longitude using *GPS Module*. The nearest police station and hospital is identified and the location value is sent to them so that they can timely arrival of help can be ensured. The GSM module sends notification to the concerned authorities. An android application eases the process of sending the notifications to the hospitals and police station in case of failure of *GSM module* sometimes. The *AWS* cloud is deployed to collect the data from the Raspberry Pi, implements the data analytics modules and notify. They are provided with a web interface where they can choose the options to generate the reports. Reports are generated weekly as well as monthly and are sent to the respective authorities for further actions using automated e-mail system.

3.2 EXTERNAL INTERFACE REQUIREMENTS

3.2.1 **User Interfaces**

i. Web Interface

The following Web Interfaces will be created as a part of this project:

a) *For Hospital Staff:*
   - The hospitals located in the vicinity will get immediate notifications about an accident.

- They have to make sure that the doctors and resources are available and send a response back as soon as possible, so that the ambulance drivers are notified.

a) *For Insurance Companies:*

All accident details and logs will be maintained on this website, which will act as proof in order to avail insurance.

b) *For police Inspectors, Ambulances, Hospitals, Cab Companies, Schools, NGO's and other interested organizations:*

- All statistics about accidents, such as the most prone areas, most common reason for the accident, most affected age groups, etc. are provided on this website on this website for all the relevant organizations.
- The trends for a given time periods will also be displayed.
- All details will be updated in real time. The statistics can be viewed for different time durations, such as daily, weekly, monthly, quarterly, half yearly, yearly, whole history.
- However, to ensure privacy and prevent misuse of data, there will be different access rights for each type of user.

ii. *Mobile Interface:*

The following mobile interfaces, in the form of android app will be developed in this project:

a) *For traffic police staff:*

The following 2 functionalities are implemented:

- Get the location of the accident immediately, as a pop up notification
- User friendly form to log all possible details about the accidents.

b) *For ambulance drivers:*

- The ambulance drivers located closest to the accident spot will get immediate notifications of the accident.
- If the driver is available and can reach the spot without delay, he can send his confirmation.

- He’ll get a notification about which hospital is the closest and has its doctors on stand -by to attend to the accident victims.

The app will be later extended to other platform such as iOS, Blackberry, etc.

iii. *Emails*

Automated emails will be sent to the Traffic Police Inspectors and Sub Inspectors *periodically*, that is, weekly, monthly, quarterly, half yearly and yearly. This will help them make better decisions in deploying their manpower and utilizing their resources. The mails will contain informative maps, graphs, pie charts which will highlight the most severe issues to be tackled. It’ll also contain the website link to get the detailed analysis.

3.2 .2**Hardware Interfaces**

Our project has the following hardware interfaces:

1) *Interfacing the Sensors and SoC:*
   We have written code in Python to periodically get sensing information from the various sensors.
2) *Interfacing the SoC and Cloud:*
   We make use of the ‘Nimbus’ platform.

3.2.3 **Software Interfaces**

i. *Operating System*

The Raspberry pi runs a ***Raspbian OS*.**

The AWS cloud server runs an ***Ubuntu OS.***

ii. *Web Server*

The software is being designed to run on ***Apache Web Server.***

iii. *Database*

We make use of ***MySQL*** Database for storing and querying our data.

iv.*Page Layout Tools*

For the web pages we plan to use twitter ***Bootstrap 3*** to ensure that are pages are responsive and mobile friendly.

v. Data Visualization Tools:

To provide an informative and comprehensive dashboard, containing bar graphs, pie charts, city maps, etc. to illustrate various stats, we plan to make use of ***Tableau*** tool.

vi. API's

We make use of ***Google Map API*** to get the location details.

3.2.4 **Communication Interfaces**

Our project shall use the HTTP protocol for communication over the internet. It makes use of GSM technology for communication between SoC (Raspberry Pi) and Cloud Infrastructure (AWS).

3. FUNCTIONAL REQUIREMENTS

Functional Requirements are the description of the facility or feature required. Functional requirements deal with what the system should do or provide for users. They include description of the required functions, outlines of associated reports or online queries, and details of data to be held in the system.
The functional requirements for our project are as follows:

3.3.1 Detection of Accidents

The detection of the accidents is done by placing the kit in the vehicle. The part of the vehicle to set the kit up depends on the place where the measurements of the required parameters can be done precisely. The sensing of the accidents is done by using the following sensors:

a) The accident is detected using *Vibration Sensor* which senses vibrations. During the normal running of vehicles, the magnitude of vibration will be lesser in comparison to that of the vibration in case of accidents.

b) The *Tilt Sensor* will measure the orientation of the vehicle with respect to the ground. Usually when the accidents occur, the vehicles aren't in the normal position as they run on the roads.

c) The *Fire Sensor* will sense the presence of fire or smoke caused due to the burning of fuel.

d) The *Pressure Sensor* will sense the amount of force with which the collision has taken place. This can be used to determine the severity of the accident. The accident would be generally more severe when the value of force/pressure applied during the collision is high.

3.3.2 Storing in the Cloud

The data generated from the sensors are collected in the cloud. AWS is used for the same. Further, data analytics is performed on the data stored in the cloud to analyze the behavior of accidents and uncover the trends it follows.

3.3.3 Notifying the nearest Police Station

The nearest police station can be tracked and the vehicle number, latitude and longitude will be sent to them. This will make sure that every accident is attended by the police and gets registered. The analysis will help in identify the areas which are very prone to accidents. This will help in the dynamic assignment of police, so that more number of police can be present near the areas which are more prone to accidents.

3.3.4 Notifying the nearest Hospital

The GPS Module can track the location of the vehicle in terms of latitude and longitude. The nearest hospital will be identified and the location value will be sent to them so that they can send their ambulance. The identification of the most prone areas will be done by the analysis. Further, the ambulances can be suggested to always be nearby some of the prone areas to reduce the time it takes for the ambulance to reach the accident spot.

3.3.5 Determine the suitable place for relocation/construction of Schools / Hospitals

The identification of the most prone areas will be done by the analysis. We can help the upcoming schools and colleges to choose a site for construction or relocation with least amount of traffic to make sure that the number of deaths caused due to the presence of schools in the prone to accident areas gets reduced.

3.3.6 Obtaining Confirmation

The Android app lets the ambulance driver, police and the hospital people confirm their arrival to the spot. Under circumstances when they don't confirm it for some reasons, other options would be searched for such as trying to contact the second nearest hospital or police station.

3.3.7 Sending mails to the Police Station

The web interface can be used by the concerned person in the police station to generate and display the accident reports in an interactive manner. The interface also lets the user choose the options for the generation of the reports. The weekly and monthly reports are mailed to the concerned person for planning and strategizing things in a better way.

3.3.8 Data Analytics Aspects

The huge data collected from the vehicles using the sensors are analyzed by using data analysis concepts and certain predictions are done using the machine learning techniques. The decision tree algorithm could be used to determine the severity of the accidents. The clustering algorithms could be used to obtain better insights. The mispredictions could be minimized by using pruning techniques to trim the levels of the decision tree and produce optimum results.

3.4 SOFTWARE SYSTEM ATTRIBUTES

3.4.1 Reliability

Reliability is the ability of a system to continue operating in the expected way over time. Reliability is measured as the probability that a system will not fail and that it will perform its intended function for a specified time interval. The key issues for reliability are:

- The sensors and board crashes and becomes unresponsive.
- The system fails due to unavailability of other externalities such as systems, networks, and cloud access.
- Log performance and auditing information about accessing the accurate geographical location and inform the nearest police station, hospitals and other services

3.4.2 Availability

Availability defines the proportion of time that the system is functional and working. Availability will be affected by system errors, infrastructure problems, malicious attacks, and system load. The key issues for availability are:

- A physical tier such as the cloud server or application server can fail or become unresponsive, causing the entire system to fail.
- A network fault or the low bandwidth internet can cause the application to be unavailable or slow down the response time.
- Inappropriate use of resources can reduce availability.

3.4.3 Security

Security is the capability of a system to reduce the chance of malicious or accidental actions outside of the designed usage affecting the system, and prevent disclosure or loss of information.

- Use authentication and authorization to prevent spoofing of user identity. Identify trust boundaries, and authenticate and authorize users crossing a trust boundary to access the cloud data

- Damage caused by malicious activity such as DOS attack, cross-site scripting and infiltrate the sensitive data stored in cloud

3.4.4 Portability

Portability is the usability of the same system in different environments. The prerequisite for portability is the generalized abstraction between the application logic and system interfaces. When software with the same functionality is produced for several computing platforms, portability is the key issue for development cost reduction. The key issues are:

- The Android/ios app should work in all the environments like Linux or windows
- The Source code should give the same accurate results in all the running platforms

3.4.5 Maintainability

Maintainability is the ability of the system to undergo changes with a degree of ease. Maintainability affects the time it takes to restore the system to its operational status following a failure or removal from operation for an upgrade. Improving system maintainability can increase availability and reduce the effects of run-time defects. There are measures by which the quality of a design of modules and their interaction among them can be measured. These measures are called coupling and cohesion.

3.4.6 Performance

Performance is an indication of the responsiveness of a system to execute specific actions in a given time interval. It can be measured in terms of latency or throughput. Latency is the time taken to respond to any event. Throughput is the number of events that take place in a given amount of time. An application’s performance can directly affect its scalability, and lack of scalability can affect performance.

3.5 PERFORMANCE REQUIREMENTS

The system must be very interactive and the delays involved must be very less. So, in every action-response of the system, there are no immediate delays experienced by the system or user. In case of opening or accessing the cloud, searching the nearest hospital/ police station, popping error messages and saving the settings or sessions there is very minimal delay. In case of opening databases, sending mails/SMS and analyzing there are no delays and the operation is performed. When connecting to the cloud server the delay is based on the bandwidth speed and the configuration between the systems. The probability that there will be or there will not be a successful connection in less than the timeout period should be less for sake of good communication.

Information transmission should be securely transmitted to server without any changes in information

As the system provides the right tools for discussion and problem solving, it must be made sure that the system is reliable in its operations and for securing the sensitive details stored in the cloud.

3.6 DATABASE REQUIREMENTS

We would require a consistent and normalized database for the system. The database will be stored in the cloud for easy and secure access.

We have the following tables:

Vehicle_Data

Attribute Name	***Attribute Type***	***Attribute Size***
VehicleID	String	30
Make	String	10
OwnerName	int	10
Specification	String	30

Sensor_Data

Attribute Name	***Attribute Type***	***Attribute Size***
Type	String	30
Analog_Output	int	10
Digital_Output	int	10
Specification	String	30
State	String	2
VehicleID	String	10

Accident_Data

Attribute Name	***Attribute Type***	***Attribute Size***
AccidentNo	String	30
AccidentDate	date	10
AccidentTime	String	10
AccidentType	Int	1
PoliceAttended	Int	1
VehicleID	String	10
SpeedZone	Int	3
Longitude	float	10
Latitude	float	10
TotalPerson	int	2
NoOfVehicles	int	2

Accident_Type_Data

Attribute Name	***Attribute Type***	***Attribute Size***
AccidentType	int	2
Specification	String	30

Ambulance_Data

Attribute Name	***Attribute Type***	***Attribute Size***
Amb_ID	String	10
Address	String	30
Longitude	float	10
Latitude	float	10
Phone	String	10

Police_Data

Attribute Name	***Attribute Type***	***Attribute Size***
Police_Station_ID	String	10
Address	String	30
Longitude	float	10
Latitude	float	10
Phone	String	10

3.7 DESIGN CONSTRAINTS

One of the major constraints is that the system is prone to damage during accidents and thus we have to design in a way that it can tolerate the damage caused in accident. To overcome this we can place the system in secured way. We would also require sensors with good sensing ability to eliminate false accident alerts.

Another constraint would be connectivity. The device should always be connected to network to give alert signals to desired teams assigned. So we have to use good network providers and devices.

The notification alerts given on app should be quick enough to reduce the damage caused due to accidents.

3.8 OTHER REQUIREMENTS

- Good connectivity.
- 24 X 7 availability.
- Good components and better design to produce effective results.
- Better UI to make it more user friendly

4. LITERATURE SURVEY

4.1 Introduction:

Accidents are one of the most devastating global tragedies with the ever-rising trends. The situations get worse due to the void created by absence of a mechanism which tracks, registers and informs the concerned authorities. Moreover, ensuring the arrival of medical response to the spot within seconds or minutes would definitely be a miracle for many victims. The International Traffic Safety Data & Analysis Group (IRTAD) conducts surveys to get aware of the several reasons leading to the accidents. Some of the very concerning problems leading to the accidents worldwide are mentioned below:

1. A very recent headline in an English daily revealed "*Only 3 of the England's 32 ambulance services reach a large majority of 'immediately life-threatening' call-outs within eight minutes*". There are more than thousands of people who lose their lives every day, struggling in the ambulances to reach the hospitals. These ambulances find it extremely difficult to find the hospitals in case of accidents happening at places like highways.

2. Most of the hospitals and schools are located at some random places in the country. This result in the schools being situated in the areas which are more prone to the accidents and also where the most of the fatalities are met by the kids. On the other hand, the hospitals are not always near the accident spots so that medical response could be sent to the accident spot immediately post the occurrence of the accidents.

3. In the present scenario, the process of notifying the police entirely relies on a third person who witnesses the accident and informs, but this usually takes a lot of time for this information to reach the police and also for the police to reach the spot after the occurrence of the accidents. Also, there are cases where the accidents remain unwitnessed by the passersby and eventually lead to demise of the victims.

4. An article dated November 22 in 2011, published in The Hindu says, "*70% of the road accidents in India in 2011 was due to drunken driving*". The trends uncovered by the analysis of the collected data provide us the locations where most of the accidents are caused due to the consumption of alcohol. This app can be used by the cab companies to know the areas where people are more likely to book their cabs after consuming alcohol so that they can always keep some cabs waiting in those areas. Another idea could be to introduce cabs meant only for the picking and the dropping of the drunken people.

5. There is no process to notify the concerned car insurance agent, life insurance agent as well as the

blood banks in case of medical emergencies. We plan to contact the concerned insurance company/agent post the detection as well as the confirmation of accident to ease and speed up the further process.

4.2 Main Body

4.2.1 Definition of IoT

The Internet of Things (IoT)[17] is the network of physical devices, vehicles, buildings and other items embedded with electronics, software, sensors, and network connectivity that enables these objects to collect and exchange data. The IoT allows objects to be sensed and controlled remotely across existing network infrastructure, creating opportunities for more direct integration of the physical world into computer-based systems, and resulting in improved efficiency, accuracy and economic benefit; when IoT is augmented with sensors and actuators, the technology becomes an instance of the more general class of cyber-physical systems, which also encompasses technologies such as smart cities, smart homes etc.

4.2.2 History of IoT

The concept of a network of smart devices was discussed as early as 1982, with a modified Coke machine at Carnegie Mellon University becoming the first internet-connected appliance, able to report its inventory and whether newly loaded drinks were cold. Mark Weiser's seminal 1991 paper on ubiquitous computing, "*The Computer of the 21st Century*", as well as academic venues such as UbiComp and PerCom produced the contemporary vision of IoT few years back when this concept was very novel.

In 1994 Reza Raji described the concept in IEEE Spectrumas *"moving small packets of data to a large set of nodes, so as to integrate and automate everything from home appliances to entire factories*". Between 1993 and 1996 several companies proposed solutions like Microsoft's at Work or Novell's NEST [19]. However, only in 1999 did the field start gathering momentum. Bill Joy envisioned Device to Device (D2D) communication as part of his "Six Webs" framework, presented at the World Economic Forum at Davos in 1999.

The concept of the Internet of Things [17] first became popular in 1999, through the Auto-ID Center at MIT and related market-analysis publications. Radio-frequency identification (RFID) was seen by Kevin Ashton (one of the founders of the original Auto-ID Center) as a prerequisite for the Internet of Things at that point. If all objects and people in daily life were equipped with identifiers, computers could manage and inventory them. Besides using RFID, the tagging of things may be achieved through such technologies as near field communication, barcodes, QR codes and digital watermarking.

British entrepreneur Kevin Ashton first coined the term in 1999 while working at Auto-ID Labs (originally called Auto-ID centers, referring to a global network of objects connected to radio-frequency identification, or RFID). Typically, IoT is expected to offer advanced connectivity of devices, systems, and services that goes beyond machine-to-machine (M2M) [5] communications and covers a variety of protocols, domains, and applications. The interconnection of these embedded devices (including smart objects), is expected to usher in automation in nearly all fields, while also enabling advanced applications like a smart grid and expanding to the areas such as smart cities.

"Things," in the IoT sense, can refer to a wide variety of devices such as heart monitoring implants, biochip transponders on farm animals, electric clams in coastal waters, automobiles with built-in sensors, DNA analysis devices for environmental/food/pathogen monitoring or field operation devices that assist firefighters in search and rescue operations. Legal scholars suggest to look at "Things" as an "inextricable mixture of hardware, software, data and service". These devices collect useful data with the help of various existing technologies and then autonomously flow the data between other devices. Current market examples include smart thermostat systems and washer/dryers that use Wi-Fi for remote monitoring.

In its original interpretation, one of the first consequences of implementing the Internet of Things by equipping all objects in the world with minuscule identifying devices or machine-readable identifiers would be to transform daily life. For instance, instant and ceaseless inventory control would become ubiquitous. A person's ability to interact with objects could be altered remotely based on immediate or present needs, in accordance with existing end-user agreements. Each thing is uniquely identifiable through its embedded computing system but is able to interoperate within the existing Internet infrastructure. Experts estimate that the IoT will consist of almost 50 billion objects by 2020.]

A paper[2] about "*Automatic Vehicle Accident Detection and Messaging System Using GSM and GPS Modem*" published in "International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering" states that GSM is used as a media which is used to control and monitor the transformer load from anywhere by sending a message. It has its own deterministic character. Thereby, here GSM is used to monitor and control the DC motor, Stepper motor, Temperature sensor and Solid State Relay by sending a message through GSM modem. Hence there is no need to waste time by manual operation and transportation. Hence it is considered as highly efficient communication through the mobile which will be useful in industrial controls, automobiles, and appliances which would be controlled from anywhere else. It is also highly economic and less expensive; hence GSM is preferred most for this mode of controlling. GPS is used in vehicles for both tracking and navigation. Tracking systems enable a base station to keep track of the vehicles without the intervention of the driver where, as navigation system helps the driver to reach the destination. Whether navigation system or tracking system, the architecture is more or less similar. When an accident occurred in any place then GPS system tracks the position of the vehicle and sends the information to the particular person through GSM by alerting the person through SMS or by a call.

In April 2012 International Journal of Scientific and Research Publications [8], K-nearest-neighbor was suggested to measure the distance between a query scenario and a set of scenarios in the data set. The k-nearest neighbor algorithm (KNN) is a method for classifying objects based on closest training examples in the feature space. KNN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor.

As of 2013, the vision of the Internet of Things has evolved due to a convergence of multiple technologies, ranging from wireless communication to the Internet and from embedded systems to micro-electromechanical systems (MEMS) [3]. This means that the traditional fields of embedded systems, wireless sensor networks, control systems, automation (including home and building automation), and others all contribute to enable the Internet of Things (IoT).

In February 2014, a paper on *Intelligent accident identification system using GPS and GSM modem* published in International Journal of Advanced Research in Computer and Communication Engineering [1], states that every vehicle should have vehicle unit. The vehicle unit consists of a vibration, controller, MEMS sensor, GPS system, GSM module. The vehicle unit installed in the vehicle every vehicle should have a vehicle unit. The vehicle unit consists of a vibration sensor, controller, MEMS sensor, GPS system and a GSM module. The vehicle unit installed in the vehicle senses the accident and sends the location of the accident the main server. The vibration sensor used in the vehicle will continuously sense for any large scale vibration in the vehicle. The sensed data is given to the controller. GPS module finds out the current position of the vehicle which is the location of the accident and gives that data to the GSM module. The GSM module sends this data to the control unit whose GSM number is already there in the module as an emergency number. The main server discovers the nearest ambulance to the accident place and also the shortest route between the accident spot, ambulance and the nearby hospital. Then the server sends this path to the emergency vehicle. Ambulance unit also using this information the controller controls all the traffic signals in the path of emergency vehicles and makes it ready to provide a free path to the ambulance, which ensures that the ambulance reaches the hospital without delay. At the same time, the ambulance section turns ON the RF transmitter [4]. This is used to communicate with the traffic department. Whenever a traffic signal section receives the information about the accident, the RF receiver in this section is turned ON to search for ambulance nearing the traffic signal. Control the traffic signal automatically with the help of RF module. Whenever the emergency vehicle reaches near to the traffic signal (approximately 100m), the traffic signal will be made of green via RF communication. Thereby the ambulance is recommended to attain the hospital without delay.

In April 2014, *International Journal of Engineering Research and Development* [10] published Vehicle Accident Detection and Reporting System Using GPS and GSM which implements automatic accident detection and reporting system. When accident occurs, it is sensed by Accelerometer. Short message including location of accident obtained using GPS, is sent via GSM network. It provides more than 70% safety for four wheelers. It is the fact that implementation of system will increase cost of vehicle but it is better to have some percent safety rather than having no percent of safety. This system consists of AVR microcontroller. Accelerometer sensor is used whose output values will be along X, Y and Z axes. Output of Accelerometer is input to the microcontroller. Output of microcontroller is given to relays. Load shown in diagram refers to motor of vehicle and airbag. Relays used are SPDT (single pole double throw) [5] . Normally closed contact of 1st relay is connected to motor of car that means supply is applied to motor and engine of vehicle is on. When accelerometer detects the collision of vehicle and if it's values are above specific limits then output of microcontroller is high, supply of motor goes to the ground which opens the relay contact and engine will stop. 2nd relay contractor is connected to air compressor which is further connected to air-bag. Initially this relay contact is open. When output of microcontroller is high, this activates air compressor and airbag blows. GPS receiver gives location of vehicle to microcontroller in each second. Message with location of accident is sent using GSM to the pre-programmed numbers. The GSM is already connected to microcontroller through MAX232 to get the location.

In 2015, A paper on *IoT Based Accident Prevention & Tracking System for Night Drivers* was published in International Journal of Innovative Research in Computer and Communication Engineering [9], which implements EBM (Eye Blink Monitoring Technique) to detect drowsiness of night drivers and preventing accidents. The other technologies that detect Drowsiness are EEG or Brain waves monitoring technique [6]. Such a technique requires sophisticated system to map or monitor the brain of subject and determine the state of drowsiness based on the neurological sleep cycle. Though EEG technique is accurate to a larger extent, yet it is not cost effective and has a difficult implementation. This project involves measurement of eye blink using IR sensor and head movement using accelerometer. The IR transmitter is used to transmit the infrared rays in our eye. The IR receiver is used to receive the reflected infrared rays of eye. If the eye is closed then the output of IR receiver would be high, otherwise the IR receiver output is low. To know whether the eye is in closing or opening position, the output is provided to a logic circuit for alarm indication and status will displayed on LCD display. Accelerometer is placed on driver fore-head it measures tilt angle of the drivers in vertical either forward or backward direction and left or right direction from the driver knee. If tilting angle exceeds certain threshold range, this output is given to the logic circuit to indicate the alarm/buzzer and the status is finally displayed on the LCD.

4.2.3 Automotive within the IoT

For a car to be truly connected, and to really be part of an Internet of Things, it needs to be connected not just to other cars but to the wider infrastructure. The web of connected devices known as the Internet of Things is set to “explode” this year, reaching almost 5bn items, a new report has found. The number of connected devices in on track to increase by 30 % in 2015 to 4.9 billion things before growing fivefold to 25bn by 2020 [20], according to projections from the technology research firm Gartner. Cars will be a “major element” of the expanding Internet of Things, with one in five vehicles having some sort of wireless network connection by 2020, accounting for more than a quarter of a billion cars on global roads.

Garbus [18] cites the transformation of the software-defined cockpit as a key megatrend driving Intel’s activities in the automotive industry. This involves the need to bring together connected experiences, both inside and outside the car, into a seamless environment that integrates centre stack and instrument cluster and a heads-up display is critically important to delivering a compelling, safer capability. Integrated and extending into that are advanced driver assistance systems, or ADAS. All of these screens create an opportunity for the appropriate level and place for visualizations, to help keep us safer.

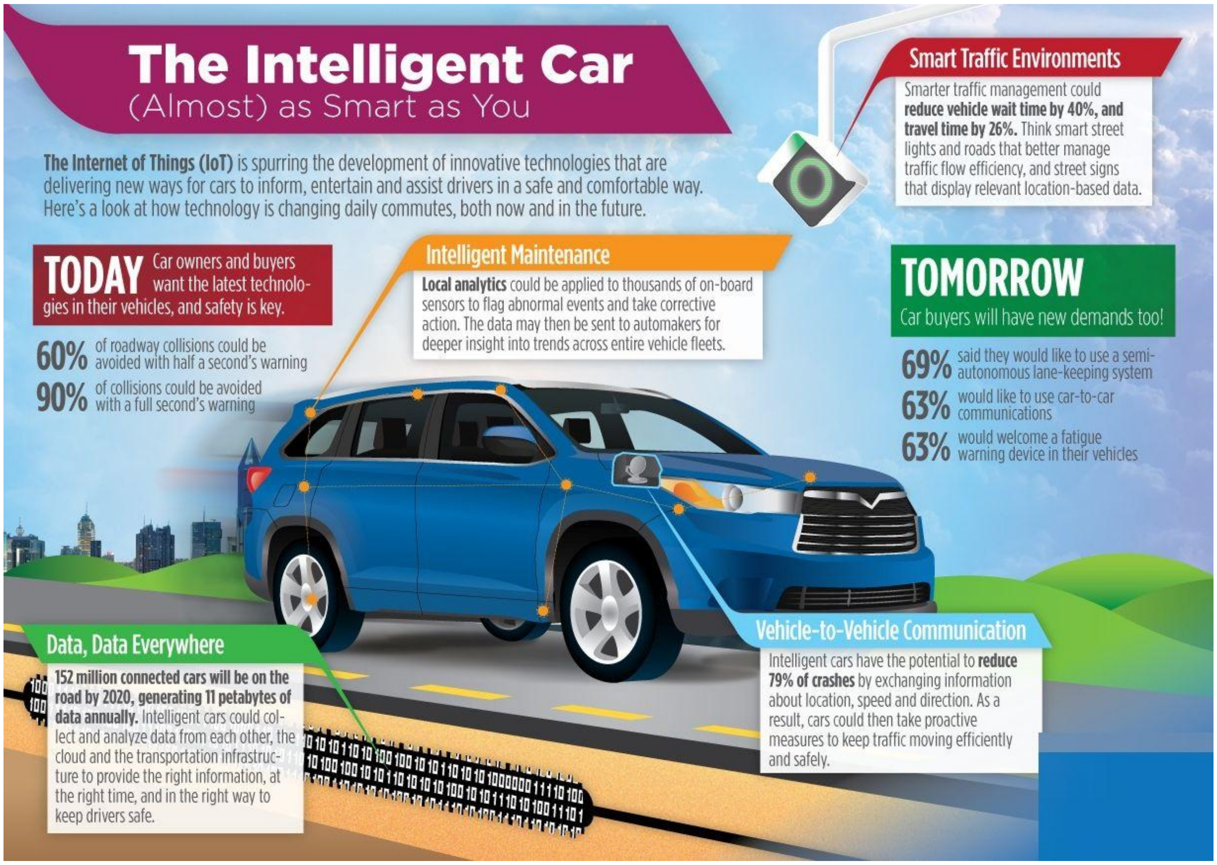


Fig 1. IoT Enabled Smart Car [20]

A recent report from McKinsey [18] found that the “dramatic increase in vehicle connectivity” that is “transforming the automotive sector” could boost the value of the global market for connectivity components and services to €170bn (£127bn) by 2020, more than five times higher than today’s €30bn. Microsoft is deepening its foray into connected cars, as indicated by updates on its partnerships with Volvo, Nissan, Harman, and IAV. Announcements came from the 2016 Consumer Electronics Show (CES) taking place this week in Las Vegas. The future of cars is a core trend of this year's show, along with home automation technology and next-generation health wearables. As connected car tech continues to evolve and driverless cars consistently garner attention, consumers will begin to demand more.

Microsoft [19] isn't new to the connected car space. The tech giant has also partnered with Toyota, Ford, Qoros, Delphi, and others to integrate its products and services into automobiles. “In the near future, the car will be connected to the Internet, as well as to other cars, your mobile phone and your home computer," Microsoft’s executive vice president for business development Peggy Johnson said.

4.2.4 Definition of Accident Data Analysis

Accident analysis is carried out in order to determine the cause or causes of an accident or series of accidents so as to prevent further incidents of a similar kind. It is also known as accident investigation. It may be performed by a range of experts, including forensic scientists, forensic engineers or health and safety advisers. Accident analysis is performed in four steps:

1. ***Fact gathering***: After an accident happened a forensic process starts to gather all possibly relevant facts that may contribute to understanding the accident.

2. ***Fact Analysis***: After the forensic process has been completed or at least delivered some results, the facts are put together to give a "big picture." The history of the accident is reconstructed and checked for consistency and plausibility.

3. ***Conclusion Drawing:*** If the accident history is sufficiently informative, conclusions can be drawn about causation and contributing factors.

4. ***Counter-measures:*** In some cases the development of counter-measures is desired or recommendations have to be issued to prevent further accidents of the same kind.

4.2.5 History of Accident Data Analysis

]

The topic of crash severity has been of interest to traffic safety community because of the direct impact on occupants involved. Weather is frequently cited and found as one of the factors contributing to either a more or less severe crash. The approaches used to model injury severities vary from one to another, depending on the purpose of the study, scope of the study and the data availability also.

In 2004, Ulfarsson and Annering found that rainy weather significantly affected the increase of property damage only level in female single UV/minivan accidents but not in male driver single SUV/minivan accidents.

In 2006, Hill and Boyle [11] utilized a logistic regression model in the fatality and incapacitating injury prediction. In their study, females in the older age groups (age of 54 or older) were more likely to suffer severe injuries in poor weather Driver characteristics such as age or gender also play important roles in the likelihood of injury severity associated with weather conditions.

In 2006, Weather impact was also evaluated in crash count-based models with the emphasis on severity counts. Abdel-Aty et al [12] used a seemingly unrelated Negative Binomial regression model to estimate the number of property damage only and injury crashes, respectively. The result showed that crash severity in adverse weather conditions causing wet pavement surface was more likely to increase at curves or ramps.

Chang and Chen [12] analyzed national freeway-1 data from Taiwan using CART and negative binomial regression model. Abellan et al. [15] analyzed two lane rural highway data of Granada, Spain using decision rules extracted from decision tree method. Depaire et al. [2] applied latent class clustering on two road user traffic accident data from 1997 to 1999 of Belgium which divides the accident data into seven clusters. Rovsek et al. [16] analyzed crash data from 2005 to 2009 of Slovenia with classification and regression tree (CART) algorithm. Kashani et al. [17] uses CART to analyze crash records obtained from information and technology department of the Iran traffic police from 2006 to 2008.

In 2007, Savolainen and Mannering predicted motorcyclists' injury severities in single and multiple crashes using nested logit and multinomial logit models, respectively. Wet pavement was significant to increase no injury severity only in single-vehicle motorcycle crashes while none of the weather related factors were found to be significant to motorcyclists' injury severities in multi-vehicle crashes involving motorcyclists.

In 2007, Caliendo et al [12] grouped crashes by total, fatal and injury crashes on curves and tangent roadways sections, and compare them using Poisson, Negative Binomial and Negative Multinomial regression models. In their study, rain was found to be a highly significant variable increasing the expected number of severe crashes for curves by a factor of 3.26 and for tangents by a factor of 2.81. Their study suggests wet-skidding for the higher number of severe crashes on curves. Compared with previous studies, our study applied a rain related crash dataset and included microscopic data at the crash moment to predict crash severity outcomes. To be specific, variables used in this study were real-time information at the crash moment, such as momentary weather and traffic data, and other non-weather data such as driver characteristics and roadway geometries. Additionally, rain-related single-vehicle crash severity models were compared to clear weather models to identify the common factors that contributed.

In 2009, Malyshkina and Mannering explained unobserved heterogeneity related to variant weather conditions over time for single- and two-vehicle crash severity potentials using a Markov switching multinomial logit model. In their study, daily averaged or maximal weather data over one week were used as follows: rain precipitation, temperature, snowfall, visibility, gust wind, and fog/frost. Weather variables such as rain precipitation, low visibility, gust wind were key factors generating time-related two-state nature of severities in single vehicle accidents on high-speed roads, but not in two-vehicle accidents.

Accident analysis is conducted to discover the reasons why an accident occurred and to prevent future accidents. Safety professionals have attributed 70-80% of aviation accidents to human error. Investigators have long known that the human and organizational aspects of systems are key contributors to accidents, yet they lack a rigorous approach for analyzing their impacts. Many safety engineers strive for blame-free reports that will foster reflection and learning from the accident, but struggle with methods that require direct technical causality, do not consider systemic factors, and seem to leave individuals looking culpable. An accident analysis method is needed that will guide the work, aid in the analysis of the role of human and organizations in accidents and promote blame-free accounting of accidents that will support learning from the events.

Current hazard analysis methods, adapted from traditional accident models, are not able to evaluate the potential for risk migration, or comprehensively identify accident scenarios involving humans and organizations. Thus, system engineers are not able to design systems that prevent loss events related to human error or organizational factors. State of the art methods for human and organization hazard analysis are, at best, elaborate event-based classification schemes for potential errors. Current human and organization hazard analysis methods are not suitable for use as part of the system engineering process. Systems must be analyzed with methods that identify all human and organization related hazards during the design process, so that this information can be used to change the design so that human error and organization errors do not occur. Errors must be more than classified and categorized, errors must be prevented in design. A new type of hazard analysis method that identifies hazardous scenarios involving humans and organizations is needed for both systems in conception and those already in the field. This thesis contains novel new approaches to accident analysis and hazard analysis. Both methods are based on principles found in the Human Factors, Organizational Safety and System Safety literature. It is hoped that the accident analysis method should aid engineers in understanding how human actions and decisions are connected to the accident and aid in the development of blame-free reports that encourage learning from accidents.

The goal for the hazard analysis method is that it will be useful in:

1) Designing systems to be safe

2) Diagnosing policies or pressures and identifying design flaws that contribute to high-risk operations

3) Identifying designs that are resistant to pressures that increase risk

4) Allowing system decision-makers to predict how proposed or current policies will affect safety.

Cluster analysis which is an important data mining technique can be used as a preliminary task to achieve various goals. Karlaftis and Tarko [13] used cluster analysis to categorize the accident data into different categories and further analyzed cluster results using Negative Binomial (NB) to identify the impact of driver age on road accidents. One of the key objectives in accident data analysis to identify the main factors associated with a road and traffic accident. However, heterogeneous nature of road accident data makes the analysis task difficult. Data segmentation has been used widely to overcome this heterogeneity of the accident data. In this paper, we proposed a framework that used K-modes clustering technique as a preliminary task for segmentation of 11,574 road accidents on road network of Dehradun (India) between 2009 and 2014 (both included). Next, association rule mining are used to identify the various circumstances that are associated with the occurrence of an accident for both the entire dataset (EDS) and the clusters identified by K-modes clustering algorithm. The findings of cluster based analysis and entire data set analysis are then compared. The results reveal that the combination of k mode clustering and association rule mining is very inspiring as it produces important information that would remain hidden if no segmentation has been performed prior to generate association rules. Further a trend analysis have also been performed for each clusters and EDS accidents which finds different trends in different cluster whereas a positive trend is shown by EDS. Trend analysis also shows that prior segmentation of accident data is very important before analysis.

In 2015, Kumar and Toshniwal published a paper in Journal of Big Data [11] which used clustering as their first step to group the data into different segments and further they used Probit model to identify relationship between different accident characteristics. Poisson models and negative binomial (NB) models have been used extensively to identify the relationship between traffic accidents and the causative factors. It has been widely recognized that Poisson models outperform the standard regression models in handling the nonnegative, random and discrete features of crash counts [10, 11]. Regression analysis (such as linear regression models, negative binomial regression models and Poisson regression models) has been the most popular technique in crash analysis because the connection between accidents and factors affecting them can be evidently identified. Using such information, the accident-prone locations can be located by the traffic engineers, and facilities such as illumination and enforcement, can then be effectively applied. However, they have limited capacity to discover new and unanticipated patterns and relationships that are hidden in conventional databases, [12] demonstrates that certain problem may occur while using traditional statistical analysis to analyze datasets with large dimensions such as an exponential increase in the number of parameters with an increase in number of variables and there could be some invalidity of statistical tests as a due to sparse data. Also, Regression models usually have their own model specific assumptions and predefined underlying relationships between dependent and independent variables. Violation of these assumptions may lead the model to provide erroneous results [13]. Hence, we need a different technique that can be used to analyze road accidents properly and can extract better results. Data mining [14] can be described as the set of techniques used for the extraction of implicit, previously unknown and hidden information from the huge amount of data. Data mining is an upcoming area that is being used by the researchers worldwide for the analysis of various types of transportation data. Several data mining techniques such as clustering, classification, association rule mining have been used to analyzed road safety data.

4.3 Conclusion:

The existing product, though practically used on a very small scale can only detect accidents using the specific sensors which would try to detect the variations in the parameters using which the accidents could be measured. Our product detects accidents using sensors such as vibration sensors, tilts sensors etc, tracks the location of the accidents and informs the nearest police station. On arrival at the spot, the police can use the android app to register the certain required parameters such as the severity of accident, light condition etc. For every occurrence of accident, the associated real-time data from the sensors is stored onto the centralized cloud. The setup and relocation of hospitals and schools can be made easier by suggesting them the most suitable place in the city based on the results of analysis, by the drivers of ambulances to get notifications about accidents and to be directed to the spot faster. These ambulance drivers, instead of being at random places in the city, can be suggested some of the places which are more likely to have more accidents. This would reduce the time taken for the ambulance to reach the spot.

As many as 134,000 fatalities in road accidents happen in India every year and a vast 70 per cent of them are due to drunken driving. The analysis also provides us the locations where most of the accidents are caused due to the consumption of alcohol. This app can be used by the cab companies to know the areas where people are more likely to book their cabs after consuming alcohol so that they can always keep some cabs waiting in those areas. Another idea could be to introduce cabs meant only for the picking and the dropping of the drunken people. The concerned car insurance and life insurance organizations can also be intimated with the notifications regarding the accidents. The blood banks can also be notified in case of any medical urgencies leading to the requirement of blood. The request for the arrangement of blood could be sent by the police after reaching the spot and examining the situation.

The additions to the existing prototypes and the existing products are as follows:

1) ***Redirection of Traffic* -** The accidents in India happen and have almost nil chances of being predicted. This app would reroute the traffic dynamically based on the results obtained after applying machine learning techniques and analyzing the collected data, which reduces the chances of encountering severe accidents.

2) ***Alerting using App* -** There are several apps made for the traffic police but none of them aim to reduce the occurrences of accidents. On the entrance of the commuter in the high risk zones, the alerts are generated on the app, along with some of the safety instructions too, which would be helpful to ensure the minimal losses caused

3) ***Reduce Drunk and Drive cases* -** Our app provides the locations where most of the accidents are caused due to the consumption of alcohol. This feature would be helpful for the cab companies to know the areas where people are more likely to book their cabs after consuming alcohol so that they can always keep some cabs waiting in those areas. Another idea could be to introduce cabs meant only for the picking and the dropping of the drunk people.

4) ***Help Hospitals and Schools* -** This app can make the setup and relocation of hospitals and schools easier by suggesting them the best place in the city based on the results of analysis.

5) ***Notify Ambulances* -** The app can also be used by the drivers of ambulances to get notifications about accidents and to be directed to the spot faster. These ambulance drivers, instead of being at random places in the city, can be suggested some of the places which are more likely to have

more accidents. This would reduce the time taken for the ambulance to reach the spot.

6) ***Inform Insurance Agents* -** The concerned car insurance and life insurance organizations can also be intimated with the notifications regarding the accidents.

7) ***Faster and Easier Blood Donation* -** The blood banks can also be notified in case of any medical urgencies leading to the requirement of blood. The request for the arrangement of blood could be sent by the police after reaching the spot and examining the situation. The users of the app need to specify their blood group while registering for this app. In case of accidents, requests to donate blood could be sent to some of the nearest people and after they confirm to donate, the further details can be sent to them.

5. DESIGN

5.1 Introduction

5.1.1 Description of the organization of the document

The design document gives us an insight into how the requirements stated in the SRS will be met. The system is divided into a number of modules, each module focusing on the delivery of a particular functionality. The design document starts with details of the modules used, the architecture diagram of the system, the various user interfaces. Further, various modeling diagrams such as class diagram, sequence diagram and data flow diagram are provided to get a better understanding of the system.

5.1.2 Number of Modules:

The project can be broadly divided into 4 modules, which are:

1) Sensing Module
2) Monitoring Module
3) Computing and Analysis Module (Cloud Infrastructure)
4) Web and Mobile Interface Module

5.1.3 Modules Description:

A detailed description of the various modules used in the project is as follows:

Sensing Modules: This module consists of the following :

a) ***Vibration Sensor:*** The project would use vibration sensors for the detection of accidents, which are piezoelectric sensors and use the piezoelectric effect to measure changes in pressure, acceleration, temperature, strain or force by converting them to an electrical charge. The sensors are in definite numbers, 10 to 12, the more they would be, the better would be the performance. These sensors are placed at various places in the vehicle to ensure the detection of accidents at each and every part of the vehicle.

b) ***Tilt Sensor***: The tilt sensor will measure the orientation of the vehicle with respect to the ground. Usually when the accidents occur, the vehicles aren't in the same level as the road.

c) ***Fire Sensor:*** The fire sensor will sense the presence of fire or smoke caused due to the burning of fuel.

d) ***Pressure Sensor***: The pressure sensor will sense the amount of force with which the collision has taken place. This can be used to determine the severity of the accident. The accident would be generally more severe when the value of force/pressure applied during the collision is high.

e) ***GPS Module***: The sole idea of using GPS module is to track the location of the vehicle in case of accidents. This location is further sent to the police for the further investigation. The location is also sent to the ambulance drivers, insurance agents and blood banks.

f) ***USB Dongle Module:*** This module is used to send data in real time from the monitoring module to cloud infrastructure.

Monitoring Module: The sensors send the real-time data to the monitoring module, which in our project is *Raspberry Pi Model 2*. Each vehicle is equipped with a monitoring module and various sensors. The data received from sensors is sent in real time to the cloud infrastructure.

Computing and Analysis Module (Cloud Infrastructure):

a) Detecting an Accident: The values from various sensors are analyzed based on many parameters by using our algorithm to detect the accident. Based on previous accident information, the algorithm keeps learning to detect an accident more accurately.

b) ***Obtaining various Statistics:*** The huge data collected from the vehicles using the sensors are analyzed by using data analysis concepts and certain predictions are done using the machine learning techniques. The decision tree algorithm could be used to determine the severity of the accidents. The clustering algorithms could be used to obtain better insights.

5.1.4 Algorithm design

Step 1: Import the libraries.

Step 2: Extract the dataset in a R variable.

Step 3: Identify the dependant and independent variables present in the dataset.

Step 4: Analyze the relation between dependent and independent variable based on certain conditions.

Step 5: Attach the new observation column corresponding to the results obtained in Step 4 to the dataset.

Step 6: Remove the dependant column from the extracted dataset.

Step 7: Apply Machine learning techniques on the training data sample.

train_var <- training(dataset) #training of dataset.

Step 8: Apply Machine learning techniques on the testing data sample

test_var <- testing(dataset) #testing of dataset.

Step 9: Measure the accuracy of the prediction based on the results obtained from Step 7 and Step 8.

predicted_var <- mean(test_var!=train_var) # Mean of the differences between both

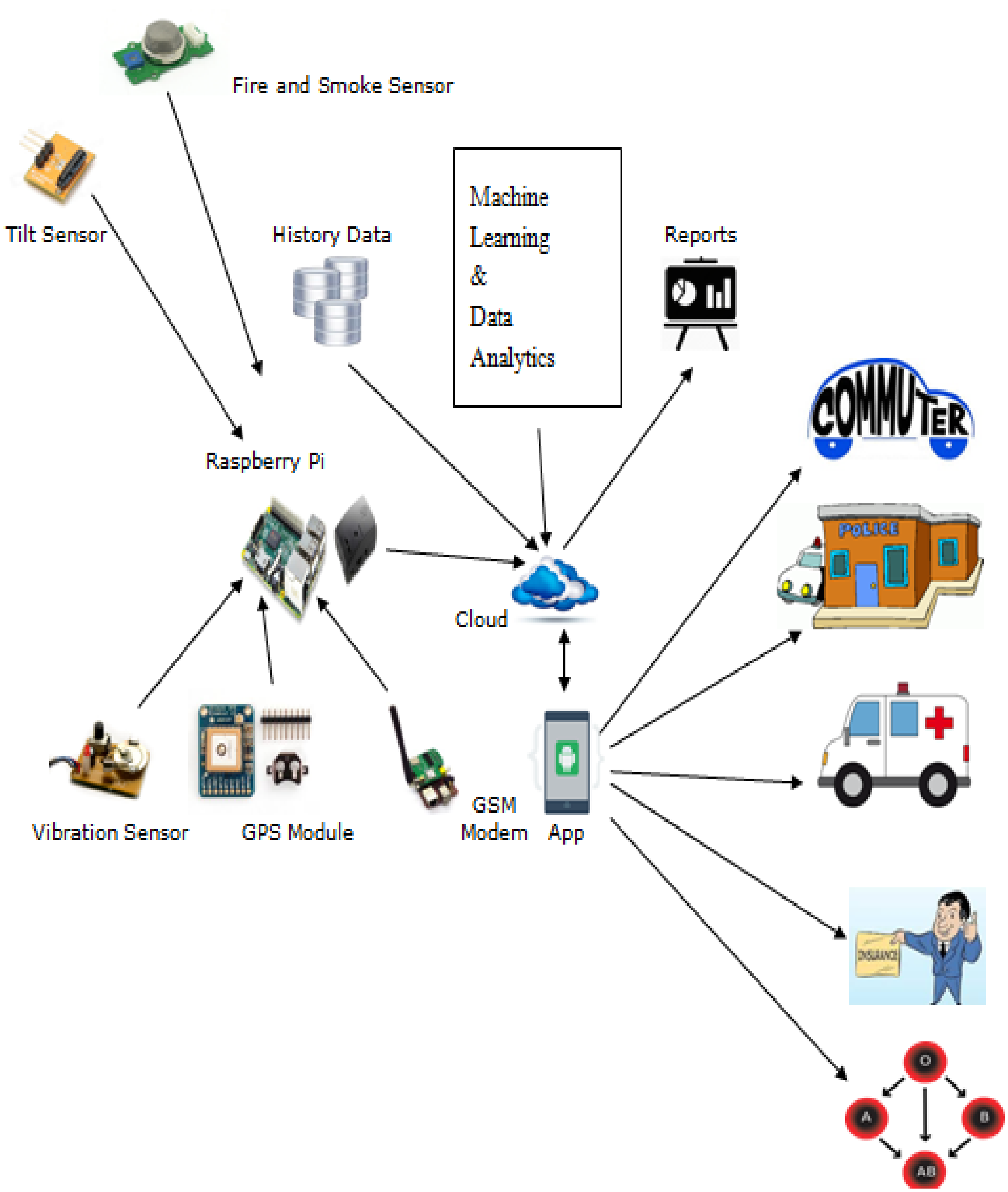
tree <- plot_tree(predicted_var) # Plotting the decision tree

Step 10: Optimize the prediction by performing the prune the decision tree obtained from Step 9.

val <- prune(tree) # pruning is performed

Repeat Step 9 until the value of the predicted_var is minimum.

5.2 Architecture Design



5.3 Graphical User Interface

5.3.1 Web Interfaces:

The following Web Interfaces will be created as a part of this project:

c) *For Hospital Staff:*

- The hospitals located in the vicinity will get immediate notifications about an accident.
- They have to make sure that the doctors and resources are available and send a response back as soon as possible, so that the ambulance drivers are notified.

d) *For Insurance Companies:*

All accident details and logs will be maintained on this website, which will act as proof in order to avail insurance.

e) *For police Inspectors, Ambulances, Hospitals, Cab Companies, Schools, NGO's and other interested organizations:*

- All statistics about accidents, such as the most prone areas, most common reason for the accident, most affected age groups, etc. are provided on this website for all the relevant organizations.
- The trends for a given time periods will also be displayed.
- All details will be updated in real time. The statistics can be viewed for different time durations, such as daily, weekly, monthly, quarterly, half yearly, yearly, whole history.
- However, to ensure privacy and prevent misuse of data, there will be different access rights for each type of user.

5.3.2 Mobile Interfaces:

The following mobile interfaces, in the form of android app will be developed in this project:

a) For traffic police staff:

The following 2 functionalities are implemented:

- Get the location of the accident immediately, as a pop up notification

- User friendly form to log all possible details about the accidents once he reaches the accident location.

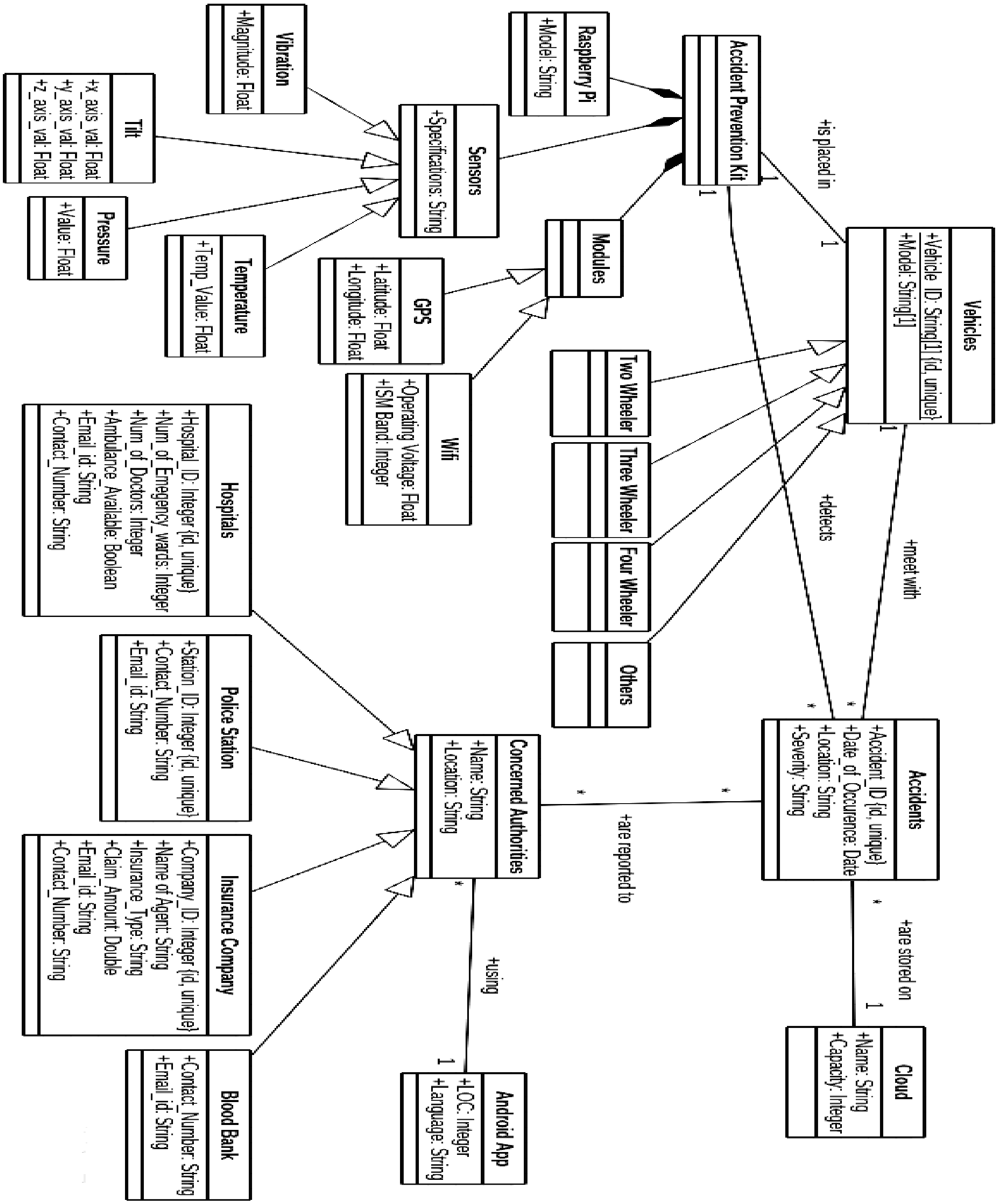
b) *For ambulance drivers:*

- The ambulance drivers located closest to the accident spot will get immediate notifications of the accident.
- If the driver is available and can reach the spot without delay, he can send his confirmation.
- He'll get a notification about which hospital is the closest and has its doctors on stand -by to attend to the accident victims.

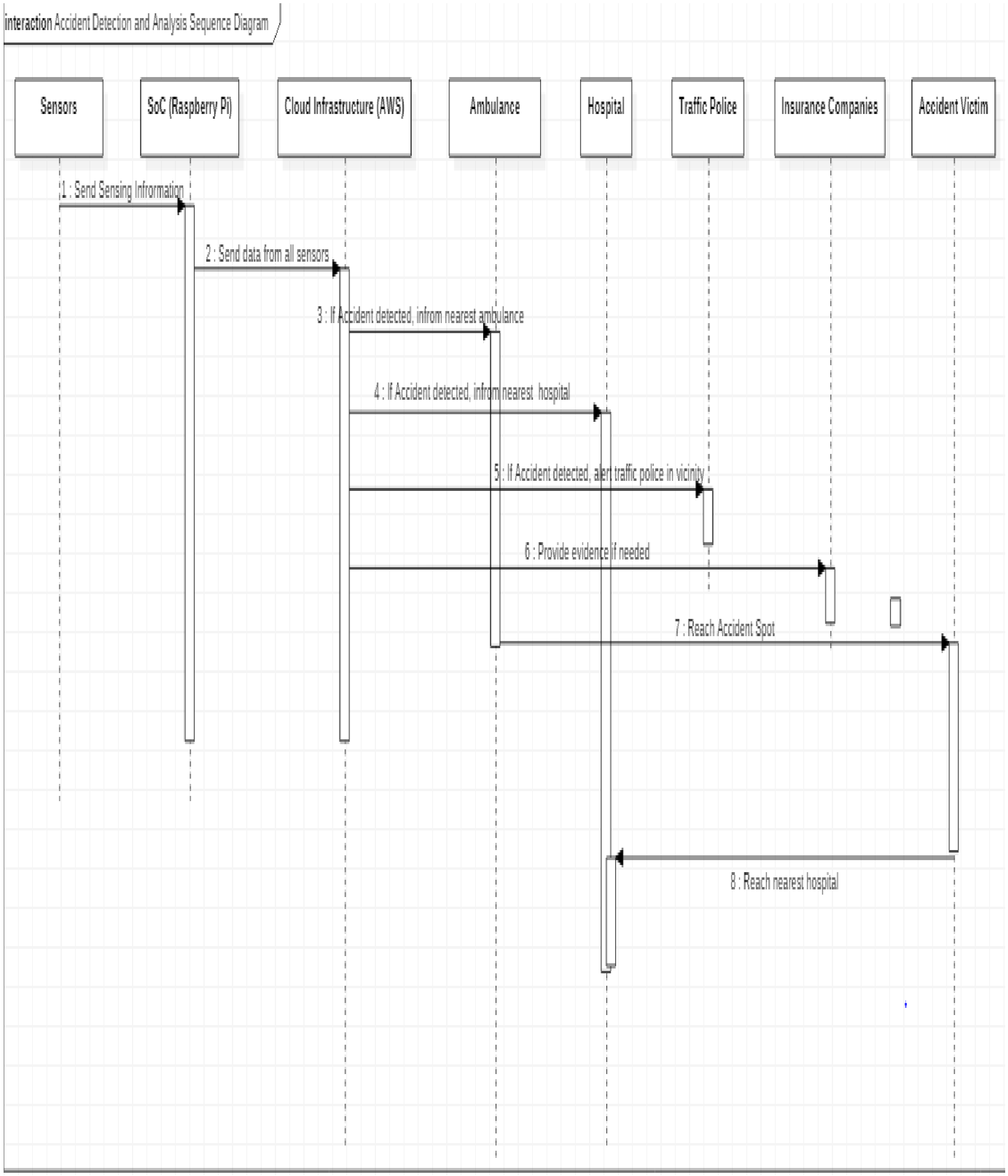
5.3.3 Emails

Automated emails will be sent to the Traffic Police Inspectors and Sub Inspectors *periodically*, that is, weekly, monthly, quarterly, half yearly and yearly. This will help them make better decisions in deploying their manpower and utilizing their resources. The mails will contain informative maps, graphs, pie charts which will highlight the most severe issues to be tackled. It'll also contain the website link to get the detailed analysis.

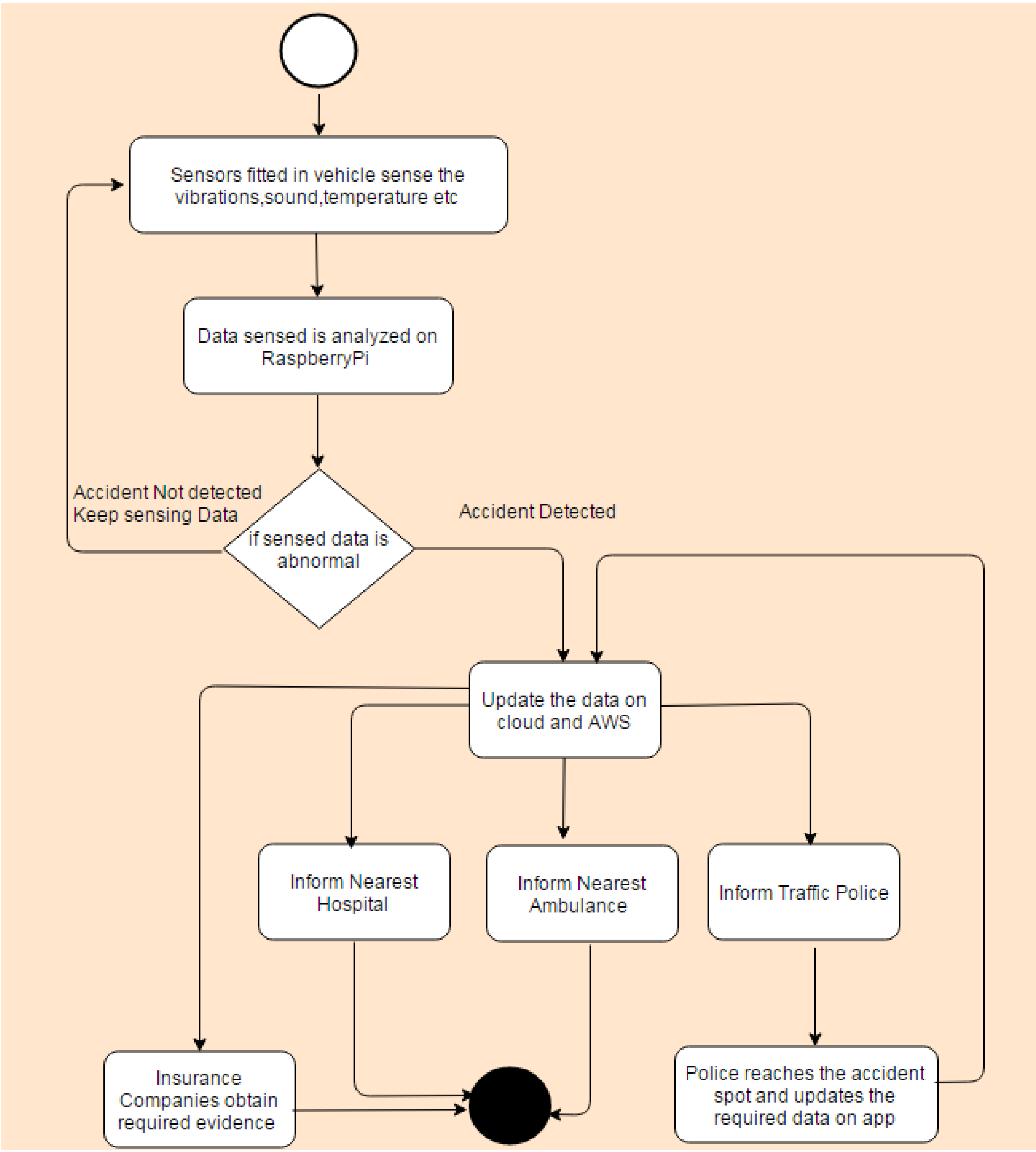
5.4 Class Diagram



5.5 Sequence Diagram



5.6 Data flow diagram



REFERENCES

[1] http://www.ijarcce.com/upload/2014/february/IJARCCE3H__s_sonika_Intelligent.pdf

[2] http://www.ijareeie.com/upload/2014/july/33_Automatic.pdf

[3] http://www.projectsof8051.com/vehicle-accident-detection-with-gps-and-gsm-modem/

[4] Design of vehicle positioning system based on arm---- IEEE-2011. Zhang Wen ; Dept. of Phys.& Electron. Inf. Eng., Neijiang Teachers Coll., Neijiang, China ; Jiang Meng may 2011.

[5] www.ijircce.com/upload/2015/april/105_49_Accident.pdf

[6] gyti.techpedia.in/project-detail/...accident-detection-and...system.../2741

[7] http://gyti.techpedia.in/project-detail/automatic-vehicle-accident-detection-and-messaging-system-using-gsm-and/2741

[8] http://www.ijsrp.org/research_paper_apr2012/ijsrp-apr-2012-83.pdf

[9] http://www.ijarcce.com/upload/2014/march/IJARCCE6I%20%20%20S%20%20nit in%20Design%20and%20Development%20of%20Automatic.pdf

[10] http://www.ijerd.com/paper/vol10-issue4/Version_4/B1042528.pdf

[11] http://link.springer.com/article/10.1186%2Fs40537-015-0035-y

[12] Heinz Hautzinger, Claus Pastor, Manfred Pfeiffer, Jochen Schmidt, Analysis Methods for Accident and Injury Risk Studies.

[13] Chang LY, Chen WC. Data mining of tree based models to analyze freeway accident frequency. J Saf Res Elsevier.2005;36.

[14] Chen Lei, Xu Nuo, Analyzing Method of Traffic Accident Causation through Experts Method and Statistical Analysis http://www.whiznets.com/iot-solution/iot-device-cloud.html/

[15] "Internet of Things Global Standards Initiative". *ITU*. Retrieved 26 June 2015.

[16] *"A list of IoT and M2M platforms* http://monblocnotes.com/node/1979

[17] https://en.wikipedia.org/wiki/Internet_of_Things#cite_ref-1

[18] http://www.automotiveworld.com/analysis/interview-elliot-garbus-president-internet-things- solutions-group-general-manager-automotive-solutions-intel-corporation/

[19] http://www.informationweek.com/iot/microsoft-partnerships-drive-connected-cars-ces-2016/d/d-id/1323768

[20] http://www.telegraph.co.uk/finance/newsbysector/industry/engineering/11372205/How-connected-cars-are-driving-the-Internet-of-Things.html