M S RAMAIAH INSTITUTE OF TECHNOLOGY

(Autonomous Institute affiliated to VTU, Belgaum)

MSR Nagar, Bangalore-560054

****

January to June 2016

Department of Computer Science & Engineering

**LITERATURE SURVEY  
ON   
REDUCING ACCIDENTS USING TRAFFIC DATA ANALYSIS AND DETECTION**

Submitted by

**ABHINAV SHIKHAR 1MS12CS002  
 NAVEEN JS 1MS13CS416  
 SANDEEP KUMAR 1MS12CS097  
 T HEMA 1MS12CS139**

Under the guidance of

Sowmya BJ

Assistant Professor, MSRIT

**Literature Survey**

**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.

**2. Main Body**

**2.1 Definition of IoT**

The Internet of Things (IoT)[17] is the network of physical devices, vehicles, buildings and other items [embedded](https://en.wikipedia.org/wiki/Embedded_system) with [electronics](https://en.wikipedia.org/wiki/Electronics), [software](https://en.wikipedia.org/wiki/Software), [sensors](https://en.wikipedia.org/wiki/Sensor), and [network connectivity](https://en.wikipedia.org/wiki/Internet_access) 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](https://en.wikipedia.org/wiki/Cyber-physical_system), which also encompasses technologies such as [smart cities](https://en.wikipedia.org/wiki/Smart_grid), smart homes etc.

**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](https://en.wikipedia.org/wiki/Carnegie_Mellon_University) becoming the first internet-connected appliance, able to report its inventory and whether newly loaded drinks were cold. [Mark Weiser](https://en.wikipedia.org/wiki/Mark_Weiser)'s seminal 1991 paper on [ubiquitous computing](https://en.wikipedia.org/wiki/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 Spectrum](https://en.wikipedia.org/wiki/IEEE_Spectrum)as *"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](https://en.wikipedia.org/wiki/Microsoft)'s [at Work](https://en.wikipedia.org/wiki/At_Work) or [Novell](https://en.wikipedia.org/wiki/Novell)'s [NEST](https://en.wikipedia.org/wiki/Novell_Embedded_Systems_Technology) [19]. However, only in 1999 did the field start gathering momentum. [Bill Joy](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/Auto-ID_Labs) at [MIT](https://en.wikipedia.org/wiki/Massachusetts_Institute_of_Technology) and related market-analysis publications. Radio-frequency identification ([RFID](https://en.wikipedia.org/wiki/RFID)) was seen by Kevin Ashton (one of the founders of the original [Auto-ID Center](https://en.wikipedia.org/wiki/Auto-ID_Labs)) 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](https://en.wikipedia.org/wiki/Tag_(metadata)) of things may be achieved through such technologies as [near field communication](https://en.wikipedia.org/wiki/Near_field_communication), [barcodes](https://en.wikipedia.org/wiki/Barcodes), [QR codes](https://en.wikipedia.org/wiki/QR_codes) and [digital watermarking](https://en.wikipedia.org/wiki/Digital_watermarking).

[British](https://en.wikipedia.org/wiki/United_Kingdom) entrepreneur [Kevin Ashton](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/Radio-frequency_identification), or RFID). Typically, IoT is expected to offer advanced connectivity of devices, systems, and services that goes beyond [machine-to-machine](https://en.wikipedia.org/wiki/Machine_to_machine) (M2M) [5] communications and covers a variety of protocols, domains, and applications. The interconnection of these embedded devices (including [smart objects](https://en.wikipedia.org/wiki/Smart_objects)), is expected to usher in automation in nearly all fields, while also enabling advanced applications like a [smart grid](https://en.wikipedia.org/wiki/Smart_grid) and expanding to the areas such as [smart cities](https://en.wikipedia.org/wiki/Smart_city).

"Things," in the IoT sense, can refer to a wide variety of devices such as heart monitoring implants, [biochip](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/End-user) agreements.Each thing is uniquely identifiable through its embedded computing system but is able to interoperate within the existing [Internet](https://en.wikipedia.org/wiki/Internet) infrastructure. Experts estimate that the IoT will consist of almost 50 billion objects by 2020.[]](https://en.wikipedia.org/wiki/Internet_of_Things#cite_note-9)

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](https://en.wikipedia.org/wiki/Embedded_system) to [micro-electromechanical systems](https://en.wikipedia.org/wiki/Microelectromechanical_systems) (MEMS) [3]. This means that the traditional fields of embedded systems, [wireless sensor networks](https://en.wikipedia.org/wiki/Wireless_sensor_network), [control systems](https://en.wikipedia.org/wiki/Control_system), [automation](https://en.wikipedia.org/wiki/Automation) (including [home](https://en.wikipedia.org/wiki/Home_automation) and [building automation](https://en.wikipedia.org/wiki/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.

**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](http://www.telegraph.co.uk/finance/newsbysector/industry/11358628/Industrial-Internet-of-Things-to-boost-UK-economy-by-531bn-by-2030.html), 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.

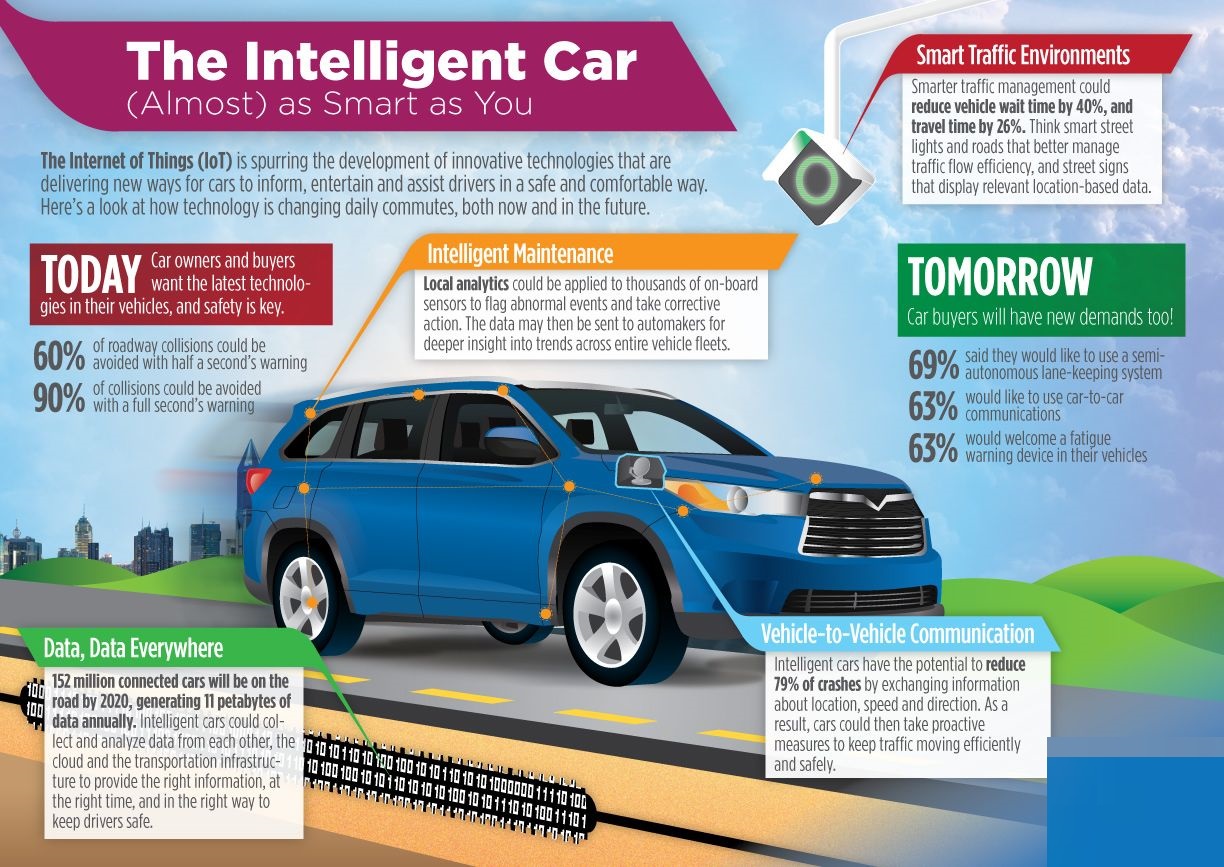


Fig 1. IoT Enabled Smart Car [20]

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](http://www.informationweek.com/it-life/ford-is-adding-apple-carplay-android-auto-to-sync-3-vehicles/d/d-id/1323746) 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](http://www.informationweek.com/software/operating-systems/microsoft-confirms-windows-10-now-on-200m-devices/d/d-id/1323753) [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.

**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](https://en.wikipedia.org/wiki/Forensic_scientists), [forensic engineers](https://en.wikipedia.org/wiki/Forensic_engineering#Forensic_engineering) or [health and safety](https://en.wikipedia.org/wiki/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.

**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.

**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.

4. **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](http://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](http://gyti.techpedia.in/project-detail/automatic-vehicle-accident-detection-and-%20%20%20%20%20%20%20%20%20%20%20%20%20%20messaging-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](http://www.ijarcce.com/upload/2014/march/IJARCCE6I%20%20%20S%20%20nit%20%20%20%20%20%20%20%20%20%20%20%20%20in%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"](http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx). 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/](http://www.automotiveworld.com/analysis/interview-elliot-garbus-president-internet-things-%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20solutions-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>