

**Tribhuvan University**

**Institute of Science and Technology**

A Final Year Project proposal

on

**Traffic Analysis System Using K-means**

**Submitted to:**

**Department of Computer Science and Information Technology**

**Ambition College**

Baneshwor, Kathmandu, Nepal

**In partial fulfilment of the requirements**

**for the Bachelors of Science in Computer Science and Information Technology**

**Submitted by:**

Arish Nepal(3108/070)

Manish Khatiwoda

Suraj Mukhiya

Udichi luitel

December 2016

**Under the supervision of**

**Mr. Chiranjibi Sitaula**

**Table of Contents**

1. Introduction ..........................................................................1

2. Problem definition ...........................................................1

3. Objectives ............................................................................1

4. Proposed Work ..............................................................2

5. Research methodology ....................................................3

5.1 Literature Review ......................................................3

5.2 Data Collection ................................................................4

5.3 Implementation.................................................................4

5.4 Testing and Verification ............................................4

5.5 Expected Result .......................................................4

6. Working Schedule .........................................................4

Reference ...............................................................................5

1. **Introduction**

Traffic analysis system is the system that keeps record of all the road accidents that occur in all parts of the country. This system allows the general people to view the areas in the nation where maximum number of accidents occur. Through this the areas can be divides as safe areas where least number of accidents occur, risky or high alert areas where maximum number of accidents occur. People from any part can view the high risky accidental areas with the help of this system. In this system the areas will be divided on the basis of 14 zones. People will also be able to view in which age group people maximum accidents occur, the time during which most accidents take place and also the major causes due to which accidents occur.

k-means algorithm will be implemented in this system. Data can be clustered with the help of this algorithm and can be represented in desired form. It is one of the data mining algorithm.

1. **Problem Definition**

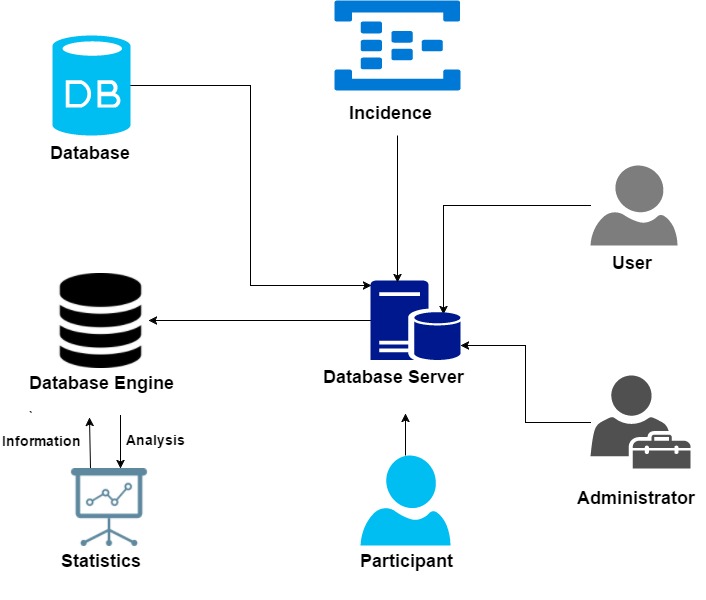
In today’s world the number of vehicles have increased rapidly. Accidents occur in almost all places in country. Increased numbers of vehicles have also resulted increased number of accidents these days. General public do not know about the highly risky areas where most of the accidents occur. Similarly, the internal or international tourists also do not know which place is safe and which one is risky. The people who want to conduct any survey or do any research need to visit several places for data collection. Similarly, it is the world of digitalization. People want everything online. They want to view everything on their laptops, computers and mobile phones. This system will be able to eliminate above discussed problems.

1. **Objectives**

The main objectives of this system is to make people aware about accidental areas and keep them safe. Some of the major objectives of this system are:

1. To make people aware about high accidental and risky areas.
2. To let the people know about the major causes due to which accidents occur.
3. To provide data to anyone who want to conduct any kinds of survey and research works.
4. To help tourists or any people who want to visit any new places to know about that particular place.
5. To display high risky areas on the map so that it can be easily understood by general people also.
6. **Proposed work**

The diagram below shows the block diagram of the system.



**Fig 1: Block Diagram**

The above diagram shows two types of users, administrator and general user. The admin stores the record of total number of incidences or accidents and the participants which may either be the victim or the suspect. The records are maintained in the database which can be viewed through the database server. Similarly, there is the database engine with the help of which analysis is done

1. **Research Methodology**

The methods that are used for the research works are described in this chapter.

* 1. **Literature Review**

In China, along with increasing motorization, there is high traffic congestion issues especially in urban areas. For urban traffic management there is development of Intelligent Traffic System (ITS) which provides several traffic information and trip advises to the users. [1]

The major objectives of ITS is to develop, evaluate, analyse and integrate new sensor, information, and communication technologies and concepts to achieve traffic efficiency, improve environmental status, save energy and enhance safety and comfort of drivers and pedestrians, and other traffic groups also. ITS have been developed in some part of world like in Canada, Japan, USA, Middle East, Europe. This system is in developing phase in India also.

Other type of traffic prediction systems are also under research or construction. Some of them are Traffic Flow Prediction Model, Urban Road Traffic Changing Trend Analysis etc. In past few years real time traffic prediction have been studied and developed in certain places based on simulation and traffic data detection.[2]

k-means clustering is a partition based cluster analysis method in which firstly ‘k’ data are selected and distance is calculated between each data value and each data is assigned to its nearest cluster. This process is repeated until accurate centroid point is not determined.

K.A Abdul Nazeer et. al[5] proposed an enhanced algorithm to enhance the accuracy and efficiency of the k-means clustering algorithm. In this algorithm two methods are used, i.e finding the better initial centroids and efficiently assigning data points to appropriate clusters with reduced time complexity. So in this algorithm we are produces good effective clusters in minimum computational time.

[6] in 2009 proposed the initial centroids algorithm based on k-means that have avoided alternative randomness of initial center.

**5.2 Data Collection**

The data about victims, suspects, accidents, high risky areas will be collected with the help of Nepal Traffic Police. Other data will be collected from some general people and internet as well.

**5.3 Implementation**

This system will be implemented using

a. PHP

b.HTML

c.CSS

d. JavaScript

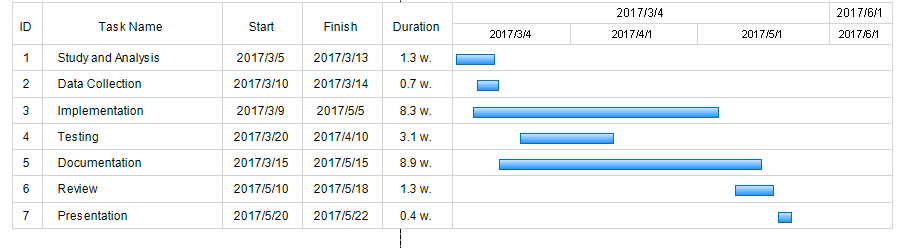
**5.4 Testing and Verification**

Testing is the most important part in any system development. After implementation of the algorithm and completion of the system unit testing and integration will be performed.

**5.5 Expected Result**

After the completion of this project and implementation of this algorithm an accurate and efficiently working system will be achieved.

1. **Working Schedule**



**Fig 2: Gantt Chart**

**References**

1. Ruimin Li, Hongliang Ma, Huapu Lu, and Min Guo, “Research and Application of the Beijing Road Traffic Prediction System,” Discrete Dynamics in Nature and Society, vol. 2014, Article ID 316032, 8 pages, 2014
2. B. S. Kerner, H. Rehborn, M. Aleksi´c, and A. Haug, “Traffic prediction systems in vehicles,” in Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems
3. B. S. Kerner, H. Rehborn, M. Aleksic, and A. Haug, “Recognition and tracking of spatial-temporal congested traffic patterns on freeways,” Transportation Research Part C, vol. 12
4. 3C. Alecsandru and S. Ishak, “Hybrid model-based and memory-based traffic prediction system,” Transportation Research Record
5. K.A.Abdul Nazeer and M. P. Sebastian, “Improving the accuracy and efficiency of the k-means clustering algorithm,” in International Conference on Data Mining and Knowledge Engineering (ICDMKE), Proceedings of the World Congress on Engineering (WCE-2009)
6. Chen Zhang and Shixiong Xia, “ K-means Clustering Algorithm with Improved Initial center,” in Second International Workshop on Knowledge Discovery and Data Mining (WKDD).
7. E.Alan Calvillo, Alejandro Padilla, Jaime Munoz”Se archins Research papers using clustering and text minig”,IEEE 2013.