Ramrao Adik Institute Of Technology
(Affiliated to the University of **Mumbai)**
Dr. D. Y. Patil Vidyanagar, Sector-7, Nerul, Navi Mumbai 400706.

Forest Fire Prediction

Internship Project Group

Submitted in partial fulfillment of the requirements of

Software Development Group (SDG)-RAIT

By

Chetan S. Kuckian [15CE1075]
Viraj V. Khedekar [15CE1075]

Supervisor
Rashmi Dhumal

AUGUST-2018

Ramrao Adik Institute of Technology
(Affiliated to the University of **Mumbai)**
Dr. D.Y. Patil Vidyanagar, Sector7, Nerul, NaviMumbai400**706.**

CERTIFICATE

This is to certify that,the internship project titled

Forest Fire Prediction

is a bonafide work done by

Chetan S. Kuckian
Viraj V. Khedekar

and is submitted in the partial fulfillment of the requirement for the Internship Programmer under SDG-RAIT

Co-supervisor
(Varsha Patil)

Supervisor
(Rashmi Dhumal)

Vice-Chairman SDG-RAIT
(Dr. Leena Ragha)

Chairman SDG-RAIT
(Dr. Ramesh Vasappanavara)

Table Of Contents

LIST OF FIGURES

Abstract

Forest fires are significant hazards that cause significant damages around the world every year. To fight these hazards and use the available resources in the best possible way, it is necessary to have an accurate prediction of their evolution beforehand. So, propagation models have been developed to determine the expected propagation of a forest fire. Such propagation models have been implemented in computer simulators. In recent years, many efforts have been oriented towards enhancing both the quality of those simulation tools and the feasibility of putting them into operational environments. However, the presence of complete and reliable fire spread prediction frameworks is still scarce, and they rarely include a real-time computational simulation module. The main reasons for this lack of confidence in computational simulations at the time of making decisions during a forest fire are the uncertainty involved in two key aspects: the accuracy of the input data with respect to the actual scenario and the execution time of the simulation tool. To deal with these two problems, a system based on the Genetic Algorithmic Approach is proposed. This system aims at reducing execution time as well as providing accurate results despite the uncertainty of the input parameters. This is achieved by the nature of Genetic Algorithm itself, through its principle of generational optimisation of the population. Thus, results can be predicted quickly and accurately so that apt measures can be taken in real case scenarios.

CHAPTER No. 1

INTRODUCTION

1.1 Motivation

1.1.1 What are Forest Fires?

The most common hazard in forests is forests fire. Forests fires are as old as the forests themselves. They pose a threat not only to the forest wealth but also to the entire regime to fauna and flora seriously disturbing the bio-diversity and the ecology and environment of a region. During summer, when there is no rain for months, the forests become littered with dry senescent leaves and twinges, which could burst into flames ignited by the slightest spark.

In Norway, an average of about 1100 forest fires occur each year. Most of these are small and relatively easy to control. Only two per cent of the registered forest fires in Norway are larger than 100 decares (100 000 m2).

1.1.2 Reasons for Forest Fires

The three basic elements required for a fire to start and continue to burn are:

- Fuel to burn (wood, brush, lichen or any combustible material in forest)
- oxygen (from the air)
- ignition source (natural or man-made)

In the absence of any one of these three elements, there can be no fire. The basic principal of wild land firefighting, therefore, is to remove one or more of these elements in the quickest and most effective way.

1.1.3 Reasons for spread of forest fires

The primary factors that influence the spread of fires are:

- Fuels
- Weather
- Topography

1.1.4 Why Use AI?

Artificial Intelligence refers to the ‘intelligence’ of machinery; the ability of the machine has in understanding its environment and then the corresponding actions it takes. Artificial intelligence has been on the Radar of technology leaders for decades and these autonomous machines can help us making decisions or carry out complex tasks.

Advantages of Artificial Intelligence:

– Machines can do complex and stressful work
– Faster than Human
– Multiple task at a time.

So in order to get appropriate data and accurate result, we decided to use AI in our project.

1.2 Purpose and Problem Definition

Forest fires are a major environmental issue, creating economical and ecological damage while endangering human lives. The forest fires importantly influence our environment and lives. Fast detection is a key element for controlling such phenomenon. Powerful computational tools are needed for predicting the amount of area that will be burned during a forest fire. There is a need of a system for predicting the area that may be involved in a forest fire event.

The Genetic algorithm has many advantages like it produce “closer” to optimal results in a “reasonable” amount of time, suitable for parallel processing, so proposed system is based on Genetic Algorithm to predict and categorize the burned area of forest and fire spread in the event of fire.

1.3 Scope

Forest fire propagation prediction is a crucial issue when fighting these hazards as efficiently as possible. Several propagation models have been developed and integrated in computer simulators. Such models require a set of input parameters that, in some cases, are difficult to know or even estimate precisely beforehand. Therefore, a recently defined Genetic Programming (GP) based system was introduced to reduce the uncertainty in input parameters values and improve the accuracy of the predictions. We have implemented an intelligent system based on genetic programming for the prediction of burned areas, using only data related to the forest under analysis and meteorological data. It may help in optimizing fire management efforts.

CHAPTER No. 2

LITERATURE SURVEY

Previous Research:

In [7], the author suggested the mathematical for predicting rate of spread and intensity in a continuous stratum of fuel that is contiguous to the ground. The initial growth of a forest fire occurs in the surface fuels (fuels that are supported within 6 feet or less of the Ground). Under favourable burning conditions, if sufficient heat is generated, the fire can grow vertically into the treetops causing a crown fire to develop. The nature and mechanisms of heat transfer in a crown fire are considerably different than those for a ground fire. Therefore, the model developed in this paper is not applicable to crown fires. An exception can be made for brush fields. Brush, such as chemise, is characterized by many stems and foliage that are reasonably contiguous to the ground, making it suitable for modeling as a ground fire.

In [10], the author proposed that a PC-based fire spread prediction application that can be used by fire control officers as an aid in strategic planning and resource allocation. This computer application, called Siro Fire, is a DOS protected mode application that incorporates several fire spread models that can be applied to the two major fuel types found in Australia-grass and forest. It uses GIS-derived geographic maps and digital elevation models to calculate the probable spread of a fire across the landscape and displays the results on a graphical representation of the map of the area of concern. The application is being used in the South Australian Country Fire Service as both an operational and training tool. This paper outlines the evolution of previous fire spread simulation software, the development of SiroFire and how SiroFire carries out fire spread simulations using existing fire spread prediction algorithms.

In [11], the author proposed the Mathematical model of forest fire was based on an analysis of known experimental data and using concept and methods from reactive media mechanics. In this paper the assignment and theoretical investigations of the problems of crown forest fire spread in windy condition were carried out. In this context, a study mathematical modeling-of the conditions of forest fire spreading that would make it possible to obtain a detailed picture of the change in the temperature and component concentration fields with time, and determine as well as the limiting condition of fire propagation in forest with fire break.

In [12], the author proposed that the Forest fires have environmental impacts that create economic problems as well as ecological damage. Developing a means to predict the possible size of a fire shortly after it first breaks out has the potential to guide proper resource allocation for improved fire control and was the main motivation of this research. In this study, the burned areas resulting from possible forest fires were estimated using historical forest fire records which contained parameters like geographical conditions of the existing environment, date and time when the fire broke out, meteorological data such as temperature, humidity and wind speed, and the type and number of trees in a unit area. The output from the estimation methods implemented in this work predicted the size of the area lost due to the fire and the corresponding fire size, i.e. big, medium, or small fire. Some of the estimation methods investigated were Multi-layer Perceptron (MLP), Radial Basis Function Networks (RBFN), Support Vector Machines (SVM) and fuzzy logic. The results of these estimates are presented and compared to similar studies in literature.

In [2], the author proposed that the Wildland fires have consumed acres of land and affected natural habitat in ways beyond a common man's intuition. More often than not, rescue operations including evacuation of surrounding urban areas have failed in saving the damage to life and property. Case studies of historical fires hold lack of situational awareness the biggest obstacle in forest fire-fighting. Mathematical algorithms can be applied to real time environmental and spatial information to predict the spread of fire-perimeters and intensities. Systems like this can be used with surveillance-based unmanned aircraft and enable fire-fighters plan on-field fire-attacks and air-drops. The following is a step towards building such a system. The work uses topographical data from the GAP project for the West Virginia Land Cover. A decision making tool is developed using fuzzy logic to designate a fuel model for forest- fires. This fuel model is then subjected to surface fire-spread techniques provided by Huygens's Principle and Rothermel's equation to develop a real time fire-predicting system. Eliminating the considerable lack of accurate information about fire-spread behaviour can help fire managers enhance safety of fire-personnel during on-field attacks. Fire-behaviour study tools like FARSITE provide a good platform for study of historical fire and help in better understanding.

2.1 Issues in Current System

As we already discussed in above section in recent years, many efforts have been taken towards enhancing both the quality of simulation tools and the feasibility of putting them into operational environments. However, the presence of complete and reliable fire spread prediction frameworks is still scarce, and they rarely include a real-time computational simulation module. The main reasons for this lack of confidence in computational simulations at the time of making decisions during a forest fire are the uncertainty involved in two key aspects: the accuracy of the input data with respect to the actual scenario and the execution time of the simulation tool. Regarding the first aspect and the accuracy of the input data, it is well known that describing the forest fire scenario is a complex task because it involves data coming from very different sources such as satellites, weather stations, databases, and so on, and it also requires data computed by complementary models.

2.2 Benefits of the Proposed System

The solution to the problem defined above requires a system based around a model that can predict and categorize the forest fire spread with improved accuracy despite the uncertainty in the input parameters involved. The Genetic Algorithm does exactly that. Using only data related to forests under analysis and also meteorological data, the Genetic Algorithm can produce results in a reasonable amount of time, and is also suitable for parallel processing.

The main advantages of the Proposed System are:

1. Improved accuracy of Results

The Genetic Algorithm has the ability to produce "closer to optimal" results. This leads to a better intuition of the possible spread of fire.

2. Quicker Predictions

The system can make predictions in a "reasonable amount of time" while also "supporting parallel processing", which results in quicker predictions and timely actions to respond to the situation with a proper strategy.

3. Aiming for the Worst

The Genetic Algorithm is based around optimizing the results which in our case is maximum fire spread area. Thus the resultant prediction obtained is a worst case area spread with respect to given input parameters. This is very beneficial in perspective of firefighting team.

CHAPTER No. 3

METHODOLOGY

3.1 Module Description

Here we propose the Geometric Semantic-Genetic Programming (GS-GP) method on fire frequent region. We are using ForestFires dataset.

3.**1.1 Genetic Algorithm:**

Genetic Programming is a technique whereby computer programs are encoded as a set of genes that are then modified using an evolutionary algorithm known as Genetic algorithm. Genetic programming is able to produce results significantly better than traditional methods. GAs constitute a technique that works in an iterative way, that is, the quality of its results directly depends on the times it is iterated as well as its specific configuration settings.

These are evolutionary algorithms which find approximate solutions to optimization and search problems. The algorithm starts with randomly selected solutions, and then repeatedly evolves new population from it by applying various operations. Proposed inputs for project are wind speed and wind direction, moisture content of the live fuel, moisture content of the dead fuel and type of vegetation.

3.1.2 Genetic Operators:

- Selection
- Crossover
- Mutation

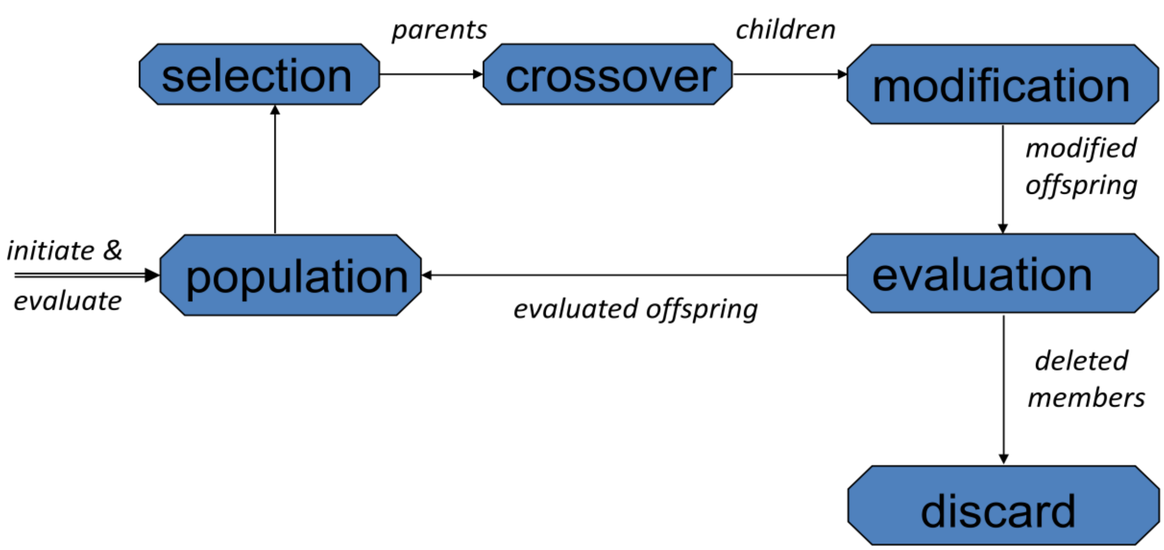


Fig. 3.1 Genetic Algorithm

Example:

- **Chromosome:** Genetic representation of a solution. It is a collection of parameters within specified range for each parameter.

  Parameters in this case are:

  Drought factor:

  Drought factor is a measure of fuel availability or a condition of dryness in the duff and upper soil layers that progress from total moisture saturation to an absence of available moisture (fuel dryness). Drought factor value range from 0 – 10, with zero value of DF the potential of fire is very low due to the fuel is fully of moisture as the value of DF increases the potential of fire increase. With DF value of 10 it indicate that the fuel is very dry and has no moisture and the potential of fire is very high.

  Temperature:

  A temperature is a numerical measure of hot or cold. It's measured in degrees and is the heat of the air mass. High temperature is responsible to increase the temperature of grass and forest fuels to their ignition points. In addition temperature affects the atmospheric stability and wind speed.

  Relative Humidity:

  Relative humidity is the amount of water vapor in the air and it is measured as a percentage. The amount of the moisture in the atmosphere will affect the amount of moisture in the grass and forest fuel. When humidity is high bushfires will not spread because the grass and forest fuel contain have a high amount of moisture.

  Wind:

  Wind speed is the rate of the movement of wind in distance per unit of time. It can be affected by terrain and obstacles such as buildings or trees. Cup-anemometer has been used to sense wind speed.

- **Encoding:** The technique used to represent a chromosome.

  The encoding techniques are of two types one is **value encoding** where every chromosome represents sequence of integer values.

  e.g. Chromosome C= [5 2 3 4 1]

  And another is **binary encoding** where every chromosome represents sequence of binary values.

  e.g. Chromosome C= [0 1 0 1 1]

  Initial population is generated randomly. The size of initial population is N. This size remains same for every next generation.

- **Fitness function:** Evaluates and rates a particular chromosome. Based on fitness value the chromosome is selected for reproduction.

  The McArthur Forest Fire Danger Indices (FFDI) [6] have been used in Australia since was developed in 1960s and used by fire authority to measure the level of fire danger. Its value depend on the drought factor which is based on dryness of the fuel and other weather variable factors such as temperature, relative humidity, rainfall, wind speed and wind directions. The accuracy measurement of FFDI output depends on the accuracy of the input variables, inaccuracy and uncertainty of these inputs may result in incorrect of the FFDI value. The value of FFDI increased from the inaccuracy of the wind speed, temperature and humidity. Bushfires prediction systems sensitivity depends on the instruments accuracy used to measure the input data.

Fig 3.2 Forest Fire Danger Index Formulation

$$FFDI = 2\,EXP^{(-0.45+0.978\ \ln(DF)-0.0345\ RH+0.0338\ T+0.0234\ U)}$$

- **Selection:** some of the individuals of existing population are selected on basis of their fitness to breed new population which is represented in the form of decision tree. Two random individuals are chosen from existing population and the fittest one is selected among these five. Elitism ensures that the fittest individual is included in every next generation.
- **Crossover:** Exchanges portions of chromosome according to some criteria.
  The Exchange sub-trees of parents is performed.
- **Mutation:** Does the random change in some values of selected chromosome Replace Sub-tree with a random tree.

The stopping criterion used is the maximum number of generations. The algorithm terminates after the 1000 generations.

3.1.3 Parameters:

We created a database of wildfire activity within the boundaries of the park which contains

- TM: Temperature(^{o}C)
- RH: Relative humidity (%)
- WN: Wind speed (km/hr)
- DF: drought factor
- FFDI: forest fire danger index

Input Data:

RH= 60 30 70
WN= 20 50 60
TM= 30 20 34
DF= 7 6 9

3.1.4 Initial Population:

Generate initial population of *N* random set of data.

Initial Population:

1: 60 20 30 7

2: 30 50 20 6

3: 70 60 34 9

3.1.5 Calculating Fitness and Selection:

Fitness:

1: 60 20 30 7= 4

2: 30 50 20 6= 5

3: 70 60 34 9= 9

3.1.5 Crossover:

Crossover:	Result:	Fitness:
1: 30 | 50 20 5	1: 30 60 34 5	34
2: 70 | 60 34 9	2: 70 50 20 9	4

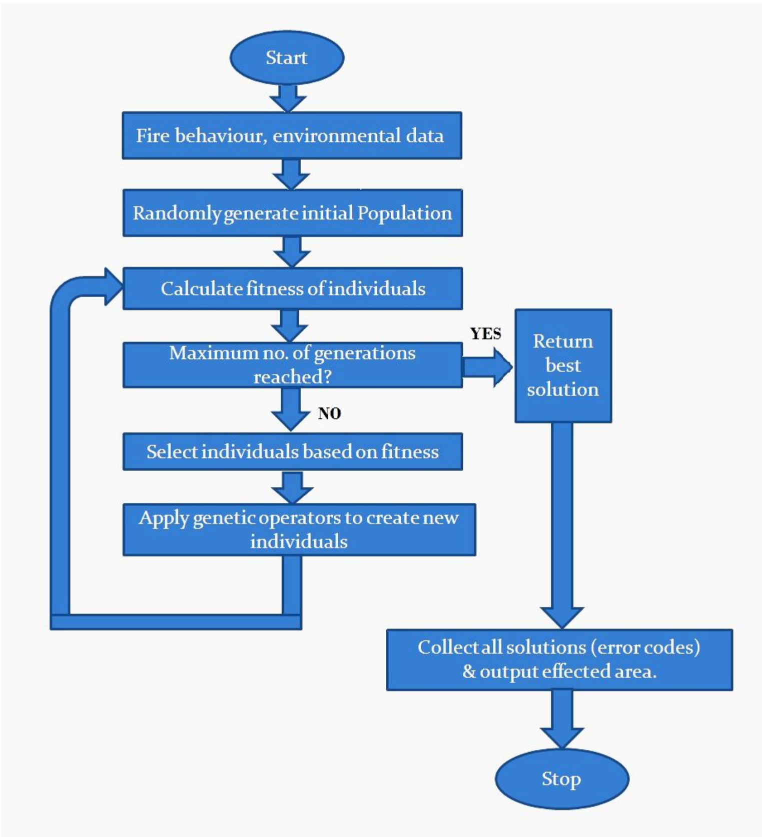
Result:

Optimum Solution:

RH=30 WN=60   TM=34

Fire intensity: 34 VERY HIGH

Fig 3.3 Flowchart of Proposed System



3.2 Hardware and Software Requirements

Software Requirements :

Operating System : Any OS with parallel processing and Python support

Platform : Python 3.5 or above pre-installed with data analysis modules

Hardware Requirements:

The hardware required for data collection is out of the scope of the project:

Computational hardware required:

Minimum 8 GB RAM on system

Processor with high clock rate

CHAPTER No. 4

Project Results

The given dataset was provided as input to our implementation of the Genetic Algorithm to run for 1000 generations. The FFDI value for fittest individual maxed out at 27 for given dataset. The population produced after given iterations was then divided into three sets according to McArthur FFDI ranges:

Low : FFDI <= 10
Medium : 10 < FFDI <= 20
High : 20 < FFDI

However the system takes a step ahead and calculates the optimal range of values for all the parameters and for all the three categories based on the above FFDI ranges defined by McArthur's formulations:

Table 4.1 Parameter Ranges

	Drought Factor (DF)	Temperature (T)	Relative Humidity (RH)	Wind(W)	FFDI
Low	DF <= 7.357	T <= 26.307	RH >= 33.55	W <= 6.278	FFDI <= 9.55
Moderate	7.357 < DF <= 9.174	26.307 < T <= 30.579	18.249 < RH <= 33.55	6.278 < W <= 6.757	9.55 < FFDI <= 19.265
High	DF > 9.174	T > 30.579	RH < 18.566	W > 6.757	FFDI > 19.265

CHAPTER No. 5

Benefits of Proposed System

The system in its current form can accurately categorize the current situation depending on values provided into one of the three categories namely: low threat, moderate threat and high threat, considering only forest damage and not the human casualties. The major benefits can be listed as follows:

1. Accuracy

The predicted values are result of thousandth generation of initial population. The McArthur index also proves as a basis to predict the results correctly.

2. Speed

Once trained through 1000 generations, the parameter values for maximum spread are identified. Thus the output for real life input values can be predicted quickly and without excessive computational costs.

3. Maximum Damage Assumption

Due to the nature of Genetic Algorithm itself, the system is enabled with consideration of worst case conditions, in other words the assumption of maximum damage for given input values.

CHAPTER No. 6

Conclusion

The aim of the proposed system has been achieved. The system can now generate an optimal population and produce a set of ranges for each parameter considered. On the basis of the ranges, the system can predict the spread category of affected area with high accuracy. The generation of optimal population was made possible due to implementation of Genetic Algorithm. Trained for maximum value of the fitness function, which in this case is the Forest Fire Danger Index, it provides an accurate and pessimistic basis for calculation of range of values. The fire spread thus, based on a reliable dataset, can be predicted accurately and quickly to enable proper firefighting techniques.

CHAPTER-7

References

Papers and websites referred:

[1] Castelli, M. L. Vanneschi, And A. Popovic. 2015. Predicting Burned Areas of Forest Fires: An Artificial Intelligence Approach. Fire Ecology 11(1):106-118.

[2] Fuzzy Logic Based Real-Time Prediction Model for Wild-Land Forest Fires: JutshiAgarwal, Kelly Cohen in August 2013

[3] Castelli, M., S. Silva, and L. Vanneschi. 2014. A C++ framework for geometric semantic genetic programming. Genetic Programming and Evolvable Machines 15: 19.Castro, J., T. Figueiredo, F. Fonseca, J.P. Castro, S. Nobre, and L. Pires. 2010.

Montesinho Natural Park: general description and natural values. Pages 119-132 in: N. Evelpidou, T. Figue-iredo, F. Mauro, V. Tecim, and A. Vassilopoulos, editors.Natural Heritage from East to West.Springer, Berlin Heidelberg, Germany.

[4] Certini, G. 2005. Effects of fire on properties of forest soils: a review. Oecologia 143(1): 1-10. doi: 10.1007/s00442-004-1788-8

[5] Cortez, P., and A.d.J.R. Morais. 2007. A data mining approach to predict forest fires using mete-orological data. Pages 512-523 in: J.M. Neves, M.F. Santos, and J.M. Machado, editors. Proceedings of the Portuguese conference on artificial intelligence (EPIA 2007). Associação Portuguesa paraaInteligência Artificial, 15–18 Dec 2007, Guimarães, Portugal. Springer, Berlin Heidelberg, Germany.

[6] Koza, J.R. 1992. Genetic Programming: on the programming of computers by means of natural selection. MIT Press, Cambridge, Massachusetts, USA.

Koza, J.R. 2010.Human-competitive results produced by genetic programming. Genetic Programming and Evolvable Machines 11(3-4): 251-284. doi: 10.1007/s10710-010-9112-3

[7] Rothermel, R.C. 1972. A mathematical model for predicting fire spread in wildland fuels. USDA Forest Service, Intermountain Research Station, Ogden, Utah, USA.

[8] StuartRusssell, Peter Norvig. “Artificial Intelligence –A modern approach”.

[9] Herbert Schildt. “Java The Complete Reference”

[10] A real-time computer application for the prediction of fire spread across the Australian landscape: J.R. Coleman, A.L. Sullivan in October 1996.

[11] Mathematical Modeling of Crown Forest Fire Spread: 2012

[12] Estimation of the Burned Area in Forest Fires Using Computational Intelligence Techniques: A. Murat Özbayoğlu in 2012

Acknowledgment

We would like to express our gratitude towards the Chairman, the Vice Chairman and members of Software Development Group, RAIT for providing us with this amazing opportunity to work on the project through the SDG internship programme. Moreover, our sincere and most heartfelt thanks towards Rashmi Salvi, faculty of Department of Computer, for providing satisfactory guidance us throughout the project. All the support and encouragement of the aforementioned has been an important factor towards completion of the project.