Prediction of Fire Propagation in Forest Areas using Genetic Algorithm

Dhumal Rashmi Amol[1]

Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India. 400706

Abstract. Forest fire causes significant damages to environment and living organisms. It becomes troublesome to recover from the consequences of forest fire. The early and accurate prediction of forest fire propagation in forest area will help in fire management efforts. The models required for prediction of forest fire spread requires many parameters and which are not known precisely during the emergency. So there is a need of intelligent system to model and predict the propagation of forest fire to reduce the consequences of forest fire and minimize the management efforts. In this work, we have discussed different techniques available in literature for forest fire prediction. This paper also describes the implementation and analysis of prediction of forest fire propagation using Genetic Algorithm and Geometric Semantic Genetic Programming.

Keywords: Forest Fire Danger Index, Genetic Algorithm, Geometric Semantic Genetic Programming

1 Introduction

According to United Nations Environment Program, Forest covered 30% of earths land area in 2014. Forest land is covered with trees and a place to live and survive for animals. Forests are an ecosystem for various animals and plants. Forest fire is most common and severe hazard in forests especially occur during summer. It causes a significant damage to forest wealth and also disturbing the ecology and bio-diversity of affected region. The main two causes of forest fire are natural causes and man-made causes. Man-made fires are occurred due to negligence behaviour of human being during the use of cigarette or bidi, naked flame or electric equipment's etc. Whereas natural causes like lightening put trees on fire. These kinds of fires occur in rainy season and are under control due to rain. During summer, forests become dry and lead to forest fire ignited by slight spark. Even changes in climate also influence frequency and intensity of forest fire. The increase in atmospheric temperatures and dryness creates favourable circumstances for ignition of fire. 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) and an ignition source (natural or man-made). The fire can't be ignited in the absence of any one of these elements.

There is a noticeable increase in forest fire across India. According to Forest Survey of India, there is huge increase in forest fire from November 2018 to February 2019.

In India most of the forest fires occur due to human activity. In the Bandipur National Park of Karnataka state of India, broke out with massive forest fire in the month of February 2019 and last for 5 days. This mainly influences the life of small animals, reptiles and birds as they are burnt and various plant species also destroyed.

The change in climate causes rise in Earth's temperature results in increase in forest fire in the regions that have no such history in past. During recent years, the fires happened in Australia in September 2019 and in California in 2018 have brought attention to wildfires across the world. These two fires destroyed millions of hectares of forests and which results in loss of habitats and food sources. This is majorly influence on the bio-diversity and ecology of affected area and need to focus on forest fire management.

The various factors influence the wild fire are listed in Table 1.

Table 1. Parameters affecting the forest fire

Parameters	Description
Wind	Change the direction of fire
Slope	Moves faster in uphill direction than downhill
Temperature	High temperature capture moisture from fuel
Humidity	Amount of water vapour in air

* rashmi.dhumal@gmail.com

Times and Seasons	Summer provides higher oxygen than winter
Fuels	Dry leaves and grass, resins, vegetation with high oil
Space between Fuels	Distance between fuels

2 Related Work

In [1], the author suggested data mining technique to predict the forest burned area. The data set used for fire spread prediction is consists of four meteorological factors such as rain, temperature, wind and humidity, which are collected through local sensors. The data mining techniques, Support Vector Machine (SVM) and given climatological factors were tested on real world data collected through local sensors. The work done in this paper was able to predict the burned areas of frequent small fires which occur in northeast region of Portugal. This was a first attempt made towards prediction of forest fire based on meteorological data. The cost of data collection is low and is real time as compared to satellite or scanner approaches. But this approach is failed in prediction of large fires. Even the data is collected on real time basis the data mining technique is applied in offline ie. after collection of data.

Perminov et al. [2] discussed the theoretical investigation was carried out for the fire spread in windy condition for crown forest. The crown forest properties like moisture content of forest fuel, bulk density, etc. are mainly contribute to estimate the spread area of forest fire. This model gives an idea about the temperature change, component concentration with respect to time and effect of other conditions on the initiation of forest fire.

In [3], the forest fire spread area was estimated using Multilayer Perceptron, Support Vector Machine, Fuzzy Logic and Radial Basis Function Networks using historical data. The total nine factors are considered which includes geographic condition, meteorological data and date and time of forest fire. The different models are implemented and evaluated their performance. The best model out of tem is Multilayer Perceptron with 65% accuracy with humidity and wind speed as an input and outputs the size of fire as small, medium and large.

Agarwal et. al. [4] discussed about the fuzzy logic based decision making technique for fuel model based forest fire. This system is similar to fire-behavioural study tool used for historical data, but can be used for real-time environmental and spatial data for fire spread prediction. Based on study of fire-behavioural model and mathematical equations the required dataset can be easily available from national databases. These databases help in appropriate selection of fuel model for a particular area. The tables associated with each image gives the essential characteristics of fuel and are used in fuzzy system for prediction of forest fire spread. This system can also be used for real-time fire spread

prediction as data is available online and can be easily accessed through MATLAB commands. This system predicts the fire spread in the span of next 10 hours. This enables fire-personnel to easily handle the situation and take preventive measures. This is a first step towards real-time fire spread prediction.

In [5], author addressed the problem with propagation models based on input parameters which are difficult to identify or estimate in advance. Hence to reduce the uncertainty in the values of input parameters and to improve the prediction accuracy, a genetic algorithm based calibration technique is proposed here. The several iterations are performed to calibrate the values which increase the execution time. The multicore capabilities are applied to reduce the execution time. Three different strategies are implemented to limit the generations. The proposed work in this paper mainly focuses on time aware feature and dynamic population size. The parallel GA allows to coexist the execution of individuals from different generations improves the prediction accuracy.

In [6,9], the author discuss about the forest fire spread prediction based on intelligent system. The geometric semantic genetic programming used for fire spread prediction using meteorological data. The prediction of amount of fire spread area during forest fire is based on the relationship between climatological data, forest related data and burned area. The results obtained shows that the geometric semantic genetic algorithm is better than standard genetic algorithm.

Table 2. Parameters affecting the forest fire

	Technology Used	Data Set Used	Number of Parameters	Predicted Spread Area
[1]	Support Vector Machine	Real-world data collected through sensors	>12	Small
[2]	Multilayer Perceptron (MLP), Radial Basis Function Networks (RBFN), Support Vector Machines (SVM) and fuzzy logic	Historical Data	9	Small and medium
[3]	Fuzzy Logic	Topographical Data	40	Small, medium and big
[4]	Genetic Algorithm	Meteorological data	--	Small, medium and big
[5]	Geometric semantic Genetic Algorithm	Historical Data	>12	Small, medium and big

3 Description of Proposed Method

3.1. Genetic Algorithm

Genetic Programming [7,8] is a evolution technique where computer programs are represented as a set of genes that are then modified using an evolutionary algorithm known as Genetic algorithm. Genetic programming is an optimization technique which helps to produce results significantly better than traditional methods. The iterative behaviour of genetic algorithm generates better quality of results based on number of iterations and the settings of genetic algorithm.

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 like crossover and mutation.

The individuals in genetic algorithm are represented by chromosomes/individuals, which is a collection of parameters within a specified range of parameters. The encoding techniques are used to represent a chromosome. There are two types of encoding techniques: value encoding and binary encoding. The initial population of individuals is generated randomly. The size of initial population is N. This size remains same for every next generation. The fitness value of each chromosome is calculated using fitness function.

3.2. Geometric Semantic Genetic Programming (GSGP)

In the genetic programming, genetic operators are applied on randomly selected individuals to evolve new child individuals. Hence these operators operate on structure or syntax of the individuals and ignore its semantic information. The syntax based genetic algorithm generates solution which is completely ignores the knowledge associated with input data which leads to incorrect design of model for particular domain where semantic of data plays an important role. The geometric semantic genetic algorithm [10] focuses on the semantic information of individuals while performing genetic operations on them.

In GSGP, the initial population is randomly generated by selecting individuals from the existing database similar to genetic algorithm. The two parent individuals P1 and P2 are randomly selected from the current generations along with random function R, whose value ranges in [0,1]. The Geometric semantic crossover generates new individual by evaluating eq.(1)

$$P_{new} = (P_1 * R) + (1-R)*P_2 \qquad (1)$$

The Geometric semantic mutation operator generates new individual by selecting one random parent from current population along with random functions R1 and R2 whose value ranges in [0,1] with mutation rate 0.06

$$P_{new} = P_1 + m(R1-R2) \qquad (2)$$

The implementation of Genetic algorithm is compared with geometric semantic genetic algorithm to get the best values of input parameters in the prediction of forest fire spread.
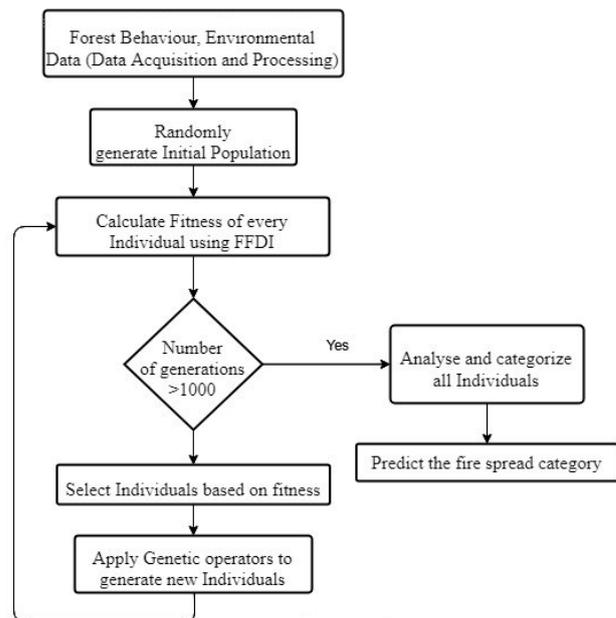
3.3. Architecture of Proposed Genetic Algorithm

Fig. 1. Working of Genetic Algorithm.

The working of proposed system is shown in figure 1. The environmental data collected and used to randomly generate individuals of initial population. The fitness of each individual is calculated and the fittest individual is propagated to the next generation using elitism. Randomly select two individuals of current generation and apply crossover and mutation operation on them to generate new child chromosomes for next generations. The above procedure is repeated till the stopping criteria have been not reached. The stopping criterion is either the number of generations or the threshold for fitness parameter.

The proposed system, consider following parameters for forest fire spread prediction:

1. Drought factor:
Drought factor measures the availability of fuel or dryness in upper layers of soil. The value of drought factor ranges from 0 to 10. The potential of fire is low for low values of drought factor and increases the potential of fire with increased values of it. If value of drought factor is 10 indicates that there is no moisture and fuel is dry enough to initiation of fire is very high.

2. Temperature:
A temperature is a measure of heat in air mass and measured in Celsius. The increase in temperature also

increases the temperature of vegetation and fuel, which results in initiation of fire. The temperature also impact on the speed of wind and atmospheric stability.

3. Relative Humidity:

The amount of water vapour in air is a measure of relative humidity. It is measured in terms of percentage. The presence of the moisture in the atmosphere will affect the amount of moisture in the grass and forest fuel. The possibility forest fire spread is low when humidity is high and it increases with decrease in moisture in atmosphere.

4. Wind:

The rate of movement of air is measured as wind speed. It is measured in terms of movement of wind in distance per unit of time. It is mainly influenced by obstacles like buildings, trees etc.

In the initial population individuals/chromosomes are defined using values of above four parameters, for example [5, 26.4, 21, 4.5]. The dataset used in this study is collected from UCI, machine learning repository. The size of each population considered here is 200. The individuals from initial population are randomly selected to generate new individuals for next generation using tournament selection and modified to find best out of them to form new individuals. In every generation the fitness function to calculate Forest Fire Danger Index (FFDI) is applied to every individual and the fittest individual with maximum FFDI value, is carried forward to next generation. The stopping criterion used here is the number of generations. The system halts after 1000 generations as there are no major changes found in the initial parameters and corresponding FFDI values.

The McArthur Forest Fire Danger Indices (FFDI) [6] is used to measure the fire danger level.

$$FFDI = 2EXP\text{-}^{(-0.45 + 0.978\,ln(DF)\,-0.0345\,RH\,+\,0.0338\,T\,+0.0234\,W)} \quad (3)$$

The value of FFDI is measured using weather variable factors such as temperature (T), relative humidity (RH), wind speed (W) and drought factor (DF). The accuracy of FFDI value depends on the accuracy of the input variables. Inaccuracy and uncertainty of these inputs may result in incorrect of the FFDI value.

4 Results and Discussion

The genetic algorithm and geometric semantic genetic algorithm are implemented on four parameters: drought factor, temperature, relative humidity and wind. In genetic algorithm, the FFDI values for fittest individual maxed out at 27 for given dataset whereas in GSGP it is maxed to 16. 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.

The values for input parameters in genetic algorithm are changes drastically as it randomly exchanges the values of parameters of parents based on their syntax. In geometric semantic genetic programming, the parameter values are varying smoothly to preserve the semantics of the individuals. The FFDI values calculated in GSGP are below 20 and are able to detect low to medium fire spread area affected due to forest fire.
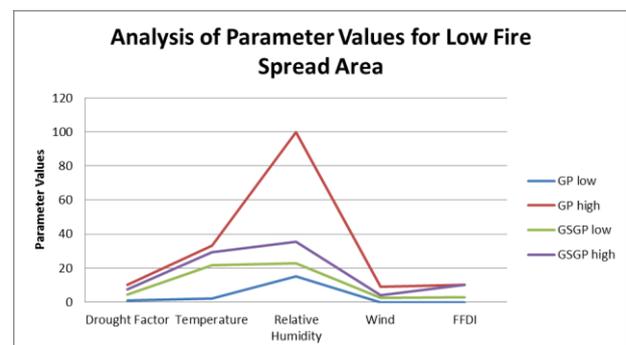

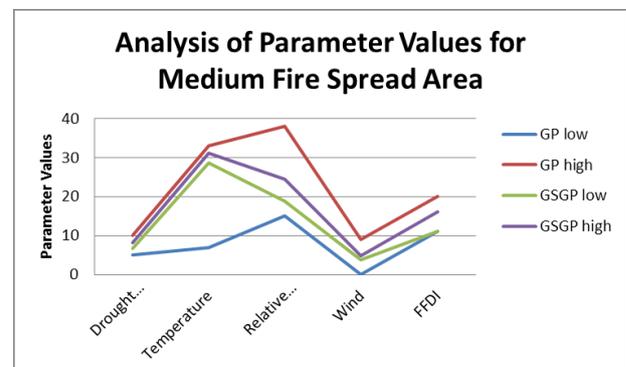**Fig. 2.** Analysis of Parameter Values for Low Fire Spread Area

Fig. 3. Analysis of Parameter Values for Medium Fire Spread Area

Conclusion

The main aim of the proposed system is to predict fire spread area during forest fire. A genetic algorithm and geometric semantic genetic algorithm based fire spread prediction system is implemented and the result obtained shows the ranges of values for input parameters that contribute towards the initiation of fire. 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 Geometric Semantic Genetic Algorithm. The fire spread thus, based on a reliable

dataset, can be predicted accurately and quickly to enable proper fire fighting techniques. The proposed system can be implemented on real time data set captured through the sensors for more accuracy and can be compared with other predictions techniques by incorporating other factors influencing the fire spread.

References

1. P. Cortez and A. Morais. A Data Mining Approach to Predict Forest Fires using Meteorological Data. In Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence, December, (2007), pp512-523

2. Perminov, Valeriy, Mathematical Modeling of Crown Forest Fire Spread. Open Journal of Forestry. 2. 17-22. 10.4236/ojf.2012.21003, (2012).

3. Ozbayoglu, Murat & Bozer, Recep. Estimation of the Burned Area in Forest Fires Using Computational Intelligence Techniques. Procedia Computer Science. **12.** (2012). 282–287. 10.1016/j.procs.2012.09.070.

4. Agarwal, Jutshi & Cohen, Kelly & Kumar, Manish. (2013). Fuzzy Logic Based Real-Time Prediction Model for Wild-Land Forest Fires. AIAA Infotech at Aerospace (I at A) Conference. 10.2514/6.2013-5060.

5. Artes, Tomas & Cencerrado, Andrés & Cortés, Ana & Margalef, Tomàs. Time aware genetic algorithm for forest fire propagation prediction: exploiting multi-core platforms: Time aware genetic algorithm for forest fire propagation prediction: exploiting multi-core platforms. Concurrency and Computation: Practice and Experience. 29. 10.1002/cpe.3837. (2016).

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

7. Koza, J.R. Genetic Programming: on the programming of computers by means of natural selection. MIT Press, Cambridge, Massachusetts, USA. (1992)

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

9. 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)

10. Vanneschi, Leonardo & Castelli, Mauro & Manzoni, Luca & Silva, Sara. A New Implementation of Geometric Semantic GP and Its Application to Problems in Pharmacokinetics. 205-216. 10.1007/978-3-642-37207-0_18. (2013).