Engineering Applications of Artificial Intelligence 24 (2011) 888–894

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Efficient forest fire occurrence prediction for developing countries using two weather parameters

George E. Sakr [a,*], Imad H. Elhajj [a], George Mitri [b]

[a] Electrical and Computer Engineering Department, American University of Beirut, Beirut, Lebanon
[b] Institute of the Environment, University of Balamand, Tripoli, Lebanon

ARTICLE INFO

Article history:
Received 13 December 2010
Accepted 18 February 2011
Available online 12 March 2011

Keywords:
Forest fire occurrence prediction
Support vector machines
Artificial neural networks
Feature Reduction
Weather data

ABSTRACT

Forest fire occurrence prediction plays a major role in resource allocation, mitigation and recovery efforts. This paper compares two artificial intelligence based methods, artificial neural networks (ANN) and support vector machines (SVM), utilizing a reduced set of weather parameters. Using a reduced set of parameters results in an efficient and reduced cost prediction system especially for developing countries. In this paper the aim is to predict forest fire occurrence by reducing the number of monitored features, and eliminating the need for weather prediction mechanisms. The reason is to reduce errors due to inaccuracies in weather prediction. The challenge is to choose a limited number of easily measurable features in the aim of reducing the cost of the system and its maintenance. At the same time, the chosen features must have a high correlation with the risk of fire occurrence. A literature review of forest fire prediction methods divided into systems/indices, and artificial intelligence is provided. The two fire danger prediction algorithms utilize relative humidity and cumulative precipitation to output a risk estimate. The assessment of these algorithms, using data from Lebanon, demonstrated their ability to accurately predict the risk of fire occurrence on a scale of four levels.

1. Introduction

Artificial neural networks (ANN) and support vector machines (SVM) have been showing impressive diversity in their applications. In traffic engineering, Mazloumi et al. (2011) showed that ANN have high ability to predict bus travel time. While in biomedical engineering, Übeyli (2008) used SVM to classify ECG signal between normal and epileptic. His reported accuracy of prediction was as high as 99.5%. SVM has also shown high accuracy in agitation detection for people with dementia (Sakr et al., 2010). In this paper we show that ANN and SVM can be used for ecological purposes to predict forest fires occurrence, based on weather data.

Forest fires are an integral part of many terrestrial ecosystems such as boreal forests, temperate forests, Mediterranean ecosystems, savannas and grasslands, among others. Fires in the Mediterranean basin account for a significant percentage of total fires occurring worldwide (Mitri and Gitas, 2004). Forest fire occurrence prediction, prevention and management measures have become increasingly important. Systems for forest fire danger prediction represent an essential tool to predict forest fire dangers, back up the forest fire monitoring and extinction phase, and it assists in the fire control planning and resource allocation (Alonso-Betanzos et al., 2003). Most industrial countries have in place prediction systems of forest fire occurrence involving a large number of monitored features and comprising weather prediction mechanisms. However, the majority of developing countries and countries in transition cannot afford the use of such systems due to lack in technical resources. This leads to a new trend in supporting new approaches in forest fire occurrence prediction and considering a reduced number of monitored features. The discussion about forest fire occurrence prediction can be divided into forest fire occurrence prediction systems and indices and artificial intelligence for forest fire occurrence prediction.

1.1. Forest fire occurrence prediction systems and indices

The first fire danger system was installed in Australia and North America. Its aim was to protect citizens and forest owners from the danger of a forest fire spreading into their properties. The system has the ability to inform local authorities and thus reducing the damage of a fire, or even preventing it (Fiorucci et al., 2008). The most famous system is used by the Canadian forest services and was first implemented in 1968. It is the Canadian Forest Fires Danger Rating System (CFFSDRS) (Stocks et al., 1989; Van Nest and Alexander, 1999). In order to reduce the complexity

* Corresponding author.
E-mail address: ges07@aub.edu.lb (G.E. Sakr).

0952-1976/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2011.02.017

of the system it has been divided into two sub systems: the fire weather index (FWI) (Van Wagner and Pickett, 1985; Van Wagner, 1987) and fire behavior prediction (FBP) (Canada, 1992). The main weather features that are monitored and recorded daily at noon are: wind speed, temperature, relative humidity, and precipitation. The FWI will generate a danger index related to the rate of fire spread, and most importantly it only relies on meteorological data recorded during the day. It is also noticeable that the FWI is totally independent of any topographic information. The index corresponds to three main fuel types: litter and other fine fuel, loosely compacted organic layers of moderate depth, and deep compact organic layer. So for the same weather parameters the index changes as the nature of fuel changes. The computation of the FWI index is split into two parts: first from an estimation of moisture in the three types of fuels the initial spread index (ISI) is computed. Then based on the fuel available, the build up index (BUI) is computed. These two indices are merged together to give the FWI. Topography information is used in the second subsystem (FBP) as well as the type of vegetation that covers the studied area. Once a fire is ignited, the FBP has the ability to predict the behavior of the fire using a semiphysical model. The American model NFDRS (National Fire Danger Rating System) (Deeming et al., 1977) is a similar approach to the CFFSDRS. This system has the ability to predict the risk of fire occurrence for the next 24 h. It outputs a series of indices that takes into account fuel moisture, weather data, and the slope of the forward rate of spread. A main difference between NFDRS and CFFSDRS is that NFDRS can also estimate the available energy per area unit and hence generate the probability that a fire will start if a heat source is placed near the fuel. It also tries to incorporate human causes or natural causes by generating an index associated with the connection of a fire to human or natural cause (lightning). The Australian fire danger index is a statistical based system (McArthur, 1966). The main relationships are between fire occurrence and temperature, wind speed, relative humidity soil moisture and the amount of fuel of the studied area. The models were generated and tested on the data provided from the dry land of Australia during dry winter and a specific equation was estimated for each type of vegetation. The system also estimates a spread index that corresponds to the danger of a fire spreading in a specific type of vegetation. Empirical systems also exist and are found mainly in the Mediterranean countries (Viegas et al., 2000). Weather data play a key role in these systems, as well as the fuel moisture. A spread index that depends on the specified location is then generated. By assuming uniform vegetation it can provide an overall index. A recent model has emerged in Italy named RISICO (RISchio Incendi e COordinamento) (Fiorucci et al., 2008). It presents a prediction over a 72 h period on the basis of available information including meteorological forecast data provided by a limited area model (LAM). This system provides a dynamic assessment of fire danger over all the Italian territory. It is based on statical information concerning the vegetation cover as well as the topography, and on dynamic data like the meteorological weather forecast over three days. The risk is obtained by combining two models; the fuel moisture model and the potential fire spread model. The Italian territory is split into cells and these two models are applied on each cell to get their risk index. The system is relatively sensitive to its numerous parameters and a careful calibration is carried out to minimize errors.

1.2. *Artificial intelligence for forest fire occurrence prediction*

Forest fire databases started being developed about 10 years ago, and their main target was to record forest fires, burnt area and associated weather parameters. At present many fire models make use of these databases to construct and assess probabilistic models. Brillinger et al. (2003) proposed a model for each specific location based on its fire history, its elevation above sea level and the corresponding dates of fire days and non-fire days. The system tries to fit different probabilistic models to data from different locations. When the fitting is accomplished the system is used to estimate the probability of a forest fire taking place at a particular location and time. No weather parameters are used and the main output is the probability of the number of fires being greater than a certain threshold. This probability estimation is helpful for fire fighters' resource allocation. It is also notable the importance of the elevation feature of the considered terrain. Experiments show that the assessed risk is accurate for a specific area. A shortcoming of this model is its close dependency over the trained area and its inability to generalize the model to nearby areas. Jaiswal et al. (2002) implemented a GIS-based forest fire danger model, to study the relationship between vegetation, climate, topography and their associated factor to cause forest fires. A forest fire danger zone map was constructed using a four-category risk scale. The resulting map was found to have a strong correlation with the highly fire affected sites. This method is also dependent on the studied area and the model cannot be generalized to other lands. It is also notable that the use of GIS is imperative in the construction of the models. Estimation of forest fire danger on a global scale was introduced by Iliadis (2005) who developed a decision support system (FFIR-EDESSYS) that implements fuzzy logic and fuzzy algebra concepts. The system was implemented in Greece, and it showed a good estimation of the forest fire danger areas but it has no indication on the estimated area of the fire which might compromise the resource allocation problem. Some important clustering techniques using partitioning methods like k-means or density based clustering like (DBSCAN) that are normally used on spatial data, can be extended to spatio-temporal data. Many models implement spatio-temporal data mining (STDM) techniques, and demonstrate that these techniques have great potential in forest fire occurrence prediction. Their application to forest fire is described by Cheng and Wang (2006) as follows:

1. Forecasting and trend analysis: basically one can make use of historical data related to burnt area to predict future forest fires. These methods provide also the ability to forecast the burnt area and the length of fire field.
2. Association rule mining for prediction of ongoing forest fire development: this is based on spatial forest data like the slopes or position of the slope as well as weather data (precipitation, wind speed and direction, temperature) and fuel type, to predict the spreading of the fire. It will add the ability to create logical conditions such as: if a fire occurs in *W* then it is very likely to spread towards *S* and allows forest fire fighters to obtain an optimal plan.
3. Pattern detection for sequence of fire events: using spatio-temporal data to discover sequential patterns that occur frequently, and hence having the ability to generate logical rules such as: if the weather parameters have a certain value in town *X*, then it is most probable that forest fire will occur in town *Y*.
4. Cluster analysis and identification of fire spots: spatio-temporal clustering may discover the cells that have a high probability of starting a fire. These cells constitute hot spots. The discrimination of fire spots, will have a direct implication on the probability of forest fires.

Li et al. (2001) developed a system for automatic identification of fire smoke using artificial neural networks applied to advanced very high resolution radiometer imagery, while Wiering and