Expert Systems with Applications 40 (2013) 5021–5026

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Review

Decision trees and genetic algorithms for condition monitoring forecasting of aircraft air conditioning

M. Gerdes *

Philotech GmbH, Brauereiweg 4, D-21614 Buxtehude, Germany

ARTICLE INFO

Keywords:
Decision tree
Forecasting
Expert system
Machine learning
Time series
Maintenance
Genetic algorithm

ABSTRACT

Unscheduled maintenance of aircraft can cause significant costs. The machine needs to be repaired before it can operate again. Thus it is desirable to have concepts and methods to prevent unscheduled maintenance. This paper proposes a method for forecasting the condition of aircraft air conditioning system based on observed past data. Forecasting is done in a point by point way, by iterating the algorithm. The proposed method uses decision trees to find and learn patterns in past data and use these patterns to select the best forecasting method to forecast future data points. Forecasting a data point is based on selecting the best applicable approximation method. The selection is done by calculating different features/attributes of the time series and then evaluating the decision tree. A genetic algorithm is used to find the best feature set for the given problem to increase the forecasting performance. The experiments show a good forecasting ability even when the function is disturbed by noise.

1. Introduction

Unscheduled maintenance costs are a significant cost factor in aircraft operation (Gerdes et al., 2009). Aircraft operators have significant increased costs, if an aircraft departure is delayed or canceled. One source of unscheduled maintenance is when a part of an aircraft needs to be replaced or repaired before the scheduled replacement time. The aircraft air conditioning and filtering system is such a system. The air filters clog faster or slower depending on the environment conditions where the aircraft is mainly operating. Filters clog faster in a moister environment and slower in a dry environment. A clogged filter system may not only cause a delay but also passenger discomfort. Currently the air condition system is monitored by using pressure sensors to detect changes in the air pressure. Forecasts are done by comparing the measurement against an empirical curve, which relates the pressure difference to the probability of clogging and a forecast for the mean time to clogging (Anonyms, 2008).

This paper describes a method to use machine learning to forecast the condition of a system. The method uses a decision tree to decide what the best method to forecast a future data point is and a genetic algorithm to adapt the decision tree to the current problem to improve the performance. The decision of the best forecasting method is based on learned patterns from past data. Motivation for this approach was to have a simple method to forecast the condition of the air conditioning system, which can adapt itself to different time series based on the operation of the aircraft and handle the influence of events on time series. The method should be easy to understand by an operator and should be able to adapt itself to different problems without much need for human interaction/expert. A time series of the system condition may be constant or linear for a long time, but suddenly an event happens and the time series changes significantly. Forecasting such a time series is difficult. The use of a decision tree enables the proposed method to use the best available forecasting method based on learned experience to adapt better to the new condition. An advantage of this method is, that it can use currently existing sensors and forecasting concepts to calculate the forecast. A genetic algorithm is used to increase the performance of the forecasting by searching for the optimal features set which gives generates the best decision tree for the given problem.

1.1. Time series

A time series is a chronological sequence of observations on a particular variable (Bowerman and O'Connell, 1993). This can be the production of a company, the DOW, a temperature, a pressure difference or an system condition. The history of a system condition can be seen as a single or multi dimensional time series. If the condition of a system is represented only by a single variable then the resulting time series is a one dimension time series. If the condition is represented by two or more different variable then the resulting time series is a multidimensional time series. Prediction of future events and conditions is called forecast, the act of

* Tel.: +49 4161 5020 128; fax: +49 4161 5020 10.
E-mail address: mike.gerdes@philotech.de

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.03.025

making such a prediction is called forecasting (Bowerman and O'Connell, 1993). Common methods for time series forecasting are:

- Simple linear regression (Montgomery et al., 1990)
- Polynomial Regression (Bowerman and O'Connell, 1993)
- Multi Regression (Montgomery et al., 1990)
- Moving average (Montgomery et al., 1990)
- Exponential smoothing (Montgomery et al., 1990)
- Autoregressive integrated moving average (ARIMA) (Montgomery et al., 1990)

Marin Goluband Andrea Budin Posavec propose the use of genetic algorithms for adapting the approximation functions for forecasting (Golub and Posavec, 1997), Animesh Chaturvedi and Samanvaya Chandra use quantitative data and a neural network to forecast financial time series (Chaturvedi and Chandra, 2004).

1.2. Decision trees

Decision trees are a method from the area of artificial intelligence and are used for machine learning (Russel and Norvig, 2003). They are often binary trees, where each node has an if-then-else function on an attribute of the sample data. More complex versions with more than two branches use a switch function. Training of the tree can be done with the ID3 algorithm (Iterative Dichotomiser 3, published by Quinlan (1986)). ID3 was the first algorithm to construct decision trees. ID3 had some problems and was improved. The improved version of ID3 is C4.5 (Quinlan, 1993). There are other algorithms to construct a decision trees available, including random trees. Decision trees are easy to understand and a decision/classification can be calculated fast. A sample decision tree can be seen in Fig. 1.

1.3. Genetic algorithms

Genetic algorithms below to the class of heuristic local search algorithms. They evaluate multiple valid solutions (population), chooses (selection) the best (firness) solutions and creates new variations of those by combining (crossover) and changing (mutation) the solution. The new set of solutions is now evaluated and the best ones are selected again and combined and changed. Each iteration of these steps is called generation. The search is finished when the algorithm calculated a certain number of generations or when an abort criteria is reached (Russel and Norvig, 2003).

2. Method

The method proposed in this paper for time series forecasting is based on decision trees. The inputs to a decision tree are time series characteristics (e.g. maximum value, gradient) and the output is an approximation function/method for forecasting based on training data. Forecasting quality is increased by using a genetic algorithm (Russel and Norvig, 2003) for optimizing the process parameters. Optimization of the process parameters allows the process also to adapt itself to different problems without human interaction. Training enables the forecasting process to experience of past data to predict the future data points in a much more reliable way then without training. This is obtained due to the fact that the process can learn when to use a different forecasting function then the obvious, because of irregularities in the time series. These irregularities can be triggered by the occurrence of certain events, that change the future data points of the time series significantly e.g. switching from a simple linear behavior to an exponential behavior.

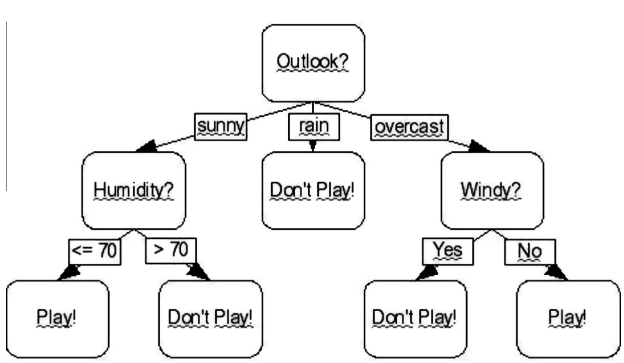

Fig. 1. A simple decision tree.

The process is divided into two parts, one part for training of the algorithm and optimizing the decision tree and one part for forecasting the time series. In the training part training samples are created, time series features are calculated, a forecasting method is selected and the decision tree is generated. Data points are forecasted, after the decision tree is generated and the process parameters are optimized. Each iteration of the forecasting process calculates a single future data point. With multiple iterations it is possible to calculate more data points. Variations of the default process can calculate multiple data points and are shown later in this section.

2.1. Training process

The learning process takes much more time than the forecasting process and is only executed when new training samples are available and during the initial training. Goal of the training process is to find an ideal set of features of the times series that give the most information for finding the optimal extrapolation algorithm. Input to the training process is a data set with different features and the best extrapolation algorithm for the time seres that the features represent. The learning process does have seven steps. The first four steps are executed only once to generate the input for the last three steps. All steps except the last one (process parameter optimization) are iterated multiple times to generate a random base population of decision trees for the parameter optimization with a genetic algorithm (last step).

Samples. Decision trees and most other concepts from artificial intelligence need many data samples for learning and finding patterns. For a time series this means, that a time series should not be to short and/or that multiple time series are available. Sample time series should include all relevant conditions and events. The algorithm can only learn from past data and thus it cannot predict events that were not in the sample data.

Process parameters. The training process is controlled by multiple parameters. These parameters control how samples and time series characteristics are calculated. Process parameters are:

- Window size [numeric]. This parameter defines how many data points each data sample contains. These data points include past data and the data points to forecast. This can be a fixed or a varying number, which is different for each data sample.
- Window shift [numeric]. This parameter defines by how many data points the sampling window should be shifted to generate a new data sample from the time series. Window shift and window size define how many training samples are generated.
- Forecast horizon [numeric]. The parameter controls how large the forecasting horizon is. This means for how many future data points the forecasting method should be calculated. The forecasting horizon can be from one data point up to all remaining data points in the time series. In the remaining document is a forecasting horizon of one data point used. Forecasting horizon cannot be larger than window size.