Knowledge Elicitation for Predictive Maintenance Modelling with Bayesian Networks

K.R. McNaught, A. Zagorecki and A. Garcia Perez

Decision Analysis and Risk Modelling Lab, Department of Informatics and Systems Engineering, Cranfield University, Defence Academy, Shrivenham, Swindon SN6 8LA, UK. k.r.mcnaught@cranfield.ac.uk, a.zagorecki@cranfield.ac.uk, a.garcia-perez@cranfield.ac.uk

Abstract. Predictive maintenance is closely connected to condition monitoring, health and usage monitoring systems (HUMS), prognostic modeling and condition-based maintenance (CBM). It is concerned with identifying what maintenance actions should be chosen and when, based on the predicted state of the item of interest over time. It can be viewed as a logical development of CBM. The authors of this paper believe that dynamic Bayesian networks (DBNs) provide a useful computational approach to predictive maintenance modeling. However, DBNs require a large number of probabilities to populate a model. While these numbers can come from collected data or domain experts, one of the key challenges in creating prognostic models is dealing with a lack of the data required, particularly for new systems. Therefore, human expertise can serve as an initial source of knowledge and can come from designers, maintainers and users of older systems. In this paper, we discuss the various types of knowledge required to build such models, how it might be elicited and the practical considerations involved.

1. Introduction

In recent years the dynamic development of affordable sensing technologies has led to the situation where modern engineering systems, such as aircraft, land vehicles, engineering systems in buildings, etc. are equipped with a large number of sensors that constantly collect data about their operation. At the same time the development of affordable mass storage devices and fast and reliable onboard communication, both wired and wireless, have allowed the implementation of sophisticated monitoring systems on many platforms. Newly designed platforms are now routinely equipped with Health and Usage Monitoring System (HUMS). The role of the HUMS is to collect data on the status of the system, possibly the environment in which it operates, and to identify any potential abnormal situations that can be linked to impending failure of the system. In theory, the availability of such data should allow for Condition Based Maintenance (CBM) – the maintenance regime that aims at repairing the right equipment at the right time – in other words, to use knowledge of early indicators of upcoming failures (possibly learned using HUMS) to perform maintenance actions before the failure occurs. In theory, this approach should reduce costs related to unnecessary preventive maintenance and costs related to experienced failure: equipment downtime, unscheduled repair, and possible costs of failure propagation.

In the early years of HUMS development, it was believed, somewhat naively, that collection of sensor data would be sufficient to implement CBM. Unfortunately, the reality showed that there are fundamental problems with the data collected by HUMS: (1) the amount of data generated by the sensors is far beyond our capabilities of drawing conclusions without automated analysis, (2) the data does not cover all aspects required to implement CBM (for example repair records), and (3) the data needs additional, often very sophisticated, analysis to turn it into actionable information. Another factor is lack of understanding of the physics of failure – in many situations it is not clear which measurements are most relevant to estimation of the residual life in the system and so many of these might simply not be collected.

This paper focuses on investigating the implementation of an approach to predictive maintenance which combines prognostic modelling with CBM. Prognostic modelling attempts to address the problem of transforming the data generated by HUMS into information that can be used to help predict the likely remaining useful life of a component or equipment and in more sophisticated form to subsequently determine the optimal maintenance schedule. The optimal maintenance schedule should be based on a combination of factors such as costs of maintenance and failures, availability of data, uncertainties related to predictions,
planned usage of equipment, etc. We emphasise, that under certain conditions, letting the equipment run to failure may be the optimal maintenance option. However, both predicting when maintenance will be required and determining the best maintenance option requires additional information that is not available through HUMS and must be obtained from other sources. These sources may include repair records, accounting information, domain expertise, etc., and may vary from well-documented digital sources, through sparse records, to human expertise.

We propose the use of dynamic Bayesian networks (DBN) as a tool for combining all relevant data and information and supporting prognostic modelling. A Bayesian network (BN) can also be extended into a decision-theoretic model in order to suggest optimal maintenance actions. In this paper we attempt to address the problem of acquiring data and knowledge for the purpose of creating prognostic models using Bayesian network models.

Of particular interest is filling the gap in obtaining information that is not readily available from some database, but resides with domain experts. Examples of such knowledge include identification of failure modes, failure patterns, indirect costs and technical constraints.

2. Condition Based Maintenance and Prognostic Modelling

Condition Based Maintenance has become an aspiration of modern maintenance managers. For example, the US Department of Defence has implemented a comprehensive program called CBM⁺ for developing infrastructure, processes and new technologies required to implement CBM (DoD, 2005). It emphasizes a holistic approach by addressing several aspects of the problem which are required for successful CBM on a large scale. The aspects addressed include: data collection and storage, fleet management, predictive models, and the development of procedures and processes to benefit from the results of the data analysis.

In this paper we focus discussion on the prognostic modelling aspect of predictive maintenance and in particular the associated elicitation of expert knowledge to support that modelling. Prognostic modelling is needed to inform the maintenance chain when a particular system is expected to fail. It is primarily concerned with situations where failure predictions are not obvious and are often beyond the capabilities of a human being to derive them from multi-dimensional and temporal data. Machine learning techniques have proven to be successful in analysing large sets of HUMS data to provide useful information (Dubrawski et al., 2007). BNs provide one of the most successful machine learning approaches. They are capable of learning models efficiently from highly dimensional data and providing estimates of failure times, anomaly detection, etc.

Unfortunately, in practical applications not all data relevant to failure prediction is available in the form of digital records. Typical examples of data not available include: failure modes, repair action records, intended future usage patterns of equipment. Some of these aspects can be obtained from the maintenance experts.

An additional challenge is posed by new systems, that have not yet been fielded, but where there is an aspiration to provide CBM from the very beginning of their service life. Often it is more than an aspiration – the contractual agreements with customers (for example in the defence sector) often require the equipment supplier to include CBM as part of the delivery requirement. In this situation a viable approach is to elicit initial models from domain experts but as service time increases, gradually replace them with models derived from real usage data. However, to build these initial models, tools for the efficient elicitation of knowledge from domain experts are needed.

3. Dynamic Bayesian Networks

The traditional BN is a static model, representing variables at a point in time (Pearl, 1988). However, DBNs also contain temporal arcs and permit the representation of system evolution over time (Dean and Kanazawa, 1989). This allows traditional Markov Chain modelling and more, while retaining the flexibility of the BN framework. Essentially, variables are replicated in a series of time slices. While arcs between variables in the same time slice have the same meaning as in a normal BN, arcs between variables in different time slices correspond to temporal dependencies. A DBN-based prognostic model is described by Muller et al. (2008) in
the context of manufacturing metal bobbins. Dynamic variables track a number of degradation mechanisms and the effect of alternative maintenance policies can then be estimated.

Figure 1. Illustrative prognostic dynamic Bayesian network.

An example prognostic DBN is presented in Figure 1. This illustrative model contains nodes/variables of several different types. Nodes $S1$, $S2$, $P1$ and $P2$ represent the true condition of components within a system of interest. Components $P1$ and $P2$ make up parallel sub-system $PSS1$. In terms of a reliability block diagram, $PSS1$ is connected in series with components $S1$ and $S2$. The overall system condition is represented by the node $Sys$. $Env1$ is an environmental variable (e.g. ambient temperature) which influences components $S1$ and $S2$ in this case. $Op1$ is a variable concerned with the operational running of the equipment (e.g. normal running or high speed). In this example, it influences each of the four components. $CM1$ to $CM4$ represent condition monitoring measurements which are imperfect indicators of the true condition of each component. Finally, $Maint1$ to $Maint4$ represent the possible maintenance actions which can be performed on each of the four components. Information from the relevant condition monitoring node is assumed known when the maintenance action is taken and in this case it is assumed that the maintenance action affects the true condition of the component in the next time slice (hence, the lag value of 1 shown on the arcs from the maintenance nodes). Several variables in Figure 1 also exhibit autocorrelation with lag 1.

4. Sources of Expert Knowledge

The process of eliciting and analysing experts’ judgements in the context of uncertainty and decision making has been studied by authors such as Meyer and Booker (2001) and O’Hagan et al. (2006). They have found a number of areas of concern in relation to planning and implementing the elicitation, including the selection of experts and the tools used to elicit and record unbiased knowledge. For example, outcomes of existing methods have been limited by issues such as the following:

- the resources, time and communication skills required by the knowledge elicitation technique (Young, 2006, p33).
- The perceived validity of the elicited knowledge, which can affect individuals’ perception of its value and applicability (Gaines, 2004).
- Experts’ motivation to provide their views, which may affect the quality of the knowledge elicited (King et al., 2002).

Additionally, existing elicitation techniques rely on interaction between a limited number of individuals and are therefore criticised for providing potentially biased results (Christel and Kang, 1992) and also for allowing policy makers to craft judicious responses to emerging problems (Rizak and Hrudey, 2005).
We consider that quantifying the uncertainty relating to failure modes occurring within particular equipment would benefit from information and expertise available in different, sometimes complementary areas. These areas include:

- the team dealing with customers’ queries. Often called the Service team, it becomes the primary point of contact between the organisation and its customers throughout the period of use of the product in service.

  Although on certain occasions members of the Service team may not be fully aware of the technical details of a specific product, they are expected to successfully handle within a reasonable time (and in most cases they do) any failures reported. In doing so, Service engineers develop significant expertise in relation to the different failure modes, degradation mechanisms, and failure rate patterns related to particular equipment.

- Maintainers. Their responsibilities often provide them with a good understanding of the cost-benefit information related to the product’s maintenance.

  Over time, maintainers develop expertise in relation to secondary effects of specific failure modes, which becomes valuable knowledge when analysing the overall cost of corrective maintenance compared to the cost of planned interventions.

- Design and automation specialists. Engineers from different backgrounds who have been involved in equipment design and automation are often able to understand the nature of a failure mode and its potential symptoms, effects and solutions. They are, in many cases, also capable of presenting this information to others in a clear fashion.

- Customers. They may be able to provide valuable usage data and also information related to the level of past or predicted usage of the equipment. Such data and information are essential to understand and predict failure modes. Customers’ views also need to be taken into account when considering specific solutions which affect running costs and availability of the equipment.

5. Knowledge Required to Develop Prognostic DBNs

BN models are very versatile in terms of the types of data that can be incorporated in a model – they allow models to be built directly by humans (Druzdzel and van der Gaag, 1995), automated learning from available data (machine learning approach), or a combination of the two. In this section we focus on elicitation of models from human experts. A BN can be viewed as a causal model where links between variables are given a causal interpretation – for example a link between a failure mode and a symptom indicates that the failure causes the presence of that symptom. Such interpretation of a BN is appealing to domain experts and has been successfully exploited in diagnostic applications of BNs (Shwe et al., 1991).

Another feature of BNs is decomposition of a problem into local dependencies. In practice this means that if some aspects of a model are changed (for example, an expert determines that the model has a missing link, or some probabilities should be revised), the change does not propagate to require re-evaluation of the rest of the model. This property is exploited to focus the elicitation on one aspect of the model at a time without distracting the expert with the whole model. An example of such a model elicitation tool for diagnostic BN models is GeNiRate (Kraaijveld and Druzdzel, 2005).

Diagnostic BN models are defined to represent statistical dependencies between failures and observations. They are not concerned with how the system operates under normal conditions. In other words – they are not system blueprints. They capture the probability distribution over possible patterns of failures. In a similar way, when eliciting knowledge for prognostic models, the focus should be on capturing the failure processes, not the regular operation.

Bayesian network models are quantified using probabilities. Even though elicitation of probabilities has been studied in the fields of psychology, statistics and intelligent systems, we strongly believe that there is a lack of proper tools to support it in the context of hardware prognostics. In particular the types of elicitation needed are:

- Identification of failure modes and degradation mechanisms (if not available from Failure Mode and Effects Analysis)
- Elicitation of failure patterns (e.g. bath-tube curve, constant, slow rise, etc).
- Elicitation of failure rates and frequencies (parameterization of the failure pattern)
- Elicitation of event probabilities, including transition probabilities between different levels of degradation from one time slice to the next. Different elicitation approaches are required for: moderate values and extreme values (important, as the models are highly sensitive to them).
- Identification of failure symptoms and condition monitoring/HUMS measurements which are relevant to the identified failure modes
- Identification of possible maintenance actions for the various failure modes, together with their costs and benefits, and the costs associated with the failure modes

Referring to Figure 1, the conditional probability tables (CPTs) required to specify a component node’s dependence on its parents will capture the degradation process and the effects of any maintenance actions. The state space for each component node may consist of several levels of wear and the main failure mode(s) associated with that component. The CPTs of the Condition Monitoring/HUMS nodes will reflect the correlations between the observable measurements and the true condition of the component. If a fixed maintenance policy is assumed, the CPTs of the Maintenance nodes will simply correspond to the logical choice of a maintenance action (including the choice of taking no action) given the observed value of the parent Condition Monitoring/HUMS variable. It is also possible, however, to have a probability distribution over the choice of actions to reflect the attitudes of different maintainers, for example. Another possibility is to extend the representation to that of a decision network or influence diagram and treat the Maintenance nodes as decisions to be optimized given the relative costs and benefits of the various possible outcomes.

One of the lessons learned during the development of diagnostic BNs for practical applications is the observation that models which have a less elaborate structure can perform better than models with additional information (Wang et al., 2006). There is a danger of illusory improvement of a model by increasing the number of variables in it – however, one should remember that this comes at a price – introducing more assumptions and requirements for more data to reliably learn its parameters. But in terms of BN models the main risk comes from making implicit assumptions on statistical independencies between variables at the time when a graphical part of the model is assumed. Therefore we postulate a simplistic approach to model creation – questioning and limiting variables that, even if they seem reasonable to include, often represent unverified beliefs about causal dependencies in the modelled domain. This especially applies when introducing risk factors to the model.

Running time may not always be the best measure of usage for predicting degradation or failures. For example, for aircraft landing gear a more useful measure of usage is take-offs and especially landings. We call these key usage measures. This translates directly into modelling practice for DBNs where the model is defined in terms of discrete time slices – these time slices do not necessarily correspond to clock time, but may be defined in terms of a key usage measure (e.g. engine starts, miles accumulated, landings). We also note that for different components within a system, different key usage measures may be appropriate and therefore the problem of unifying the results can be an additional challenge.

A key piece of information needed to determine the process of deterioration and failure is the future usage of the system. This problem touches a very fundamental challenge of prognostics in general – prognostics is concerned with future events and as such knowledge of future usage and environmental conditions to which the system will be exposed is only possible with a degree of uncertainty. This degree depends on the type of system and it can range from near certainty to high unpredictability. Some systems are used in very regular and predictable patterns (for example, manufacturing equipment) while some other are characterised by highly unpredictable usage patterns – e.g. military equipment that is subject to extreme variations and uncertainties related to future usage ranging from lengthy storage to very intensive and unpredictable usage.

One of the solutions to this problem is identification of the key classes of usage patterns. For example, for military equipment these can include: storage, stand-by, training missions (with seasonal distinctions), overseas deployment. For each class, the usage pattern would be defined. One should keep in mind that BNs are probabilistic models and usage patterns would be specified by means of likely probability distributions of key usage measures (such as engine starts, miles accumulated, landings, etc). Knowledge of how and where equipment is intended to be used may be represented by evidence which can be entered for the Operational and Environmental type variables at several time slices.
6. Conclusions and Future Work

In this paper we have presented a generic predictive maintenance model in the form of a dynamic Bayesian network and discussed the various types of knowledge required to develop such a model in order that it can be used to support decision making in relation to predictive maintenance. We conclude that a practical approach to knowledge elicitation is needed which produces a model of failure modes based on the views of experts from different areas. Our next steps are to develop both the approach to knowledge elicitation and the modelling methodology further, to test them in practical applications and then to implement them in specialist software tools. We are grateful to the KT-Box knowledge transfer programme funded by EPSRC for supporting the development of these tools.

References