

# Exploring Dynamic Bayesian Belief Networks for Intelligent Fault Management Systems

R. Sterritt, A.H. Marshall, C.M. Shapcott, S.I. McClean  
University of Ulster at Jordanstown  
Shore Road, Newtownabbey, Northern Ireland BT37 0QB

## Abstract

Systems that are subject to uncertainty in their behaviour are often modelled by Bayesian Belief Networks (BBNs). These are probabilistic models of the system in which the independence relations between the variables of interest are represented explicitly. A directed graph is used, in which two nodes are connected by an edge if one is a 'direct cause' of the other.

However the Bayesian paradigm does not provide any direct means for modelling dynamic systems. There has been a considerable amount of research effort in recent years to address this. In this paper, we review these approaches and propose a new dynamic extension to the BBN.

Our discussion then focuses on fault management of complex telecommunications and how the dynamic bayesian models can assist in the prediction of faults.

*Keywords:* dynamic bayesian belief networks, telecommunication networks, fault management, intelligent systems.

## 1 Introduction

Systems that are subject to uncertainty in their behaviour are often modelled by Bayesian Belief Networks (BBNs). These are probabilistic models of the system in which the independence relations between the variables of interest are represented explicitly. A directed graph is used, in which two nodes are connected by an edge if one is a 'direct cause' of the other.

However BBNs provide no direct mechanism for representing temporal dependencies [1],[2],[39]. In certain domains such as medicine, planning and control, and industrial environments, the incorporation of a temporal aspect into the model is crucial if the model is to achieve an effective and accurate representation of the system in question. The time that symptoms appear and their duration, the time that observations/measurements are made and the time that faults are induced can significantly affect the formulation of hypotheses used. The model must be

able to update the system given that observations and evidence can be made over time, that is capture the evolution of the system as it changes over time.

This paper is organised as follows. Section 2 introduces the fault management domain. Section 3 describes Bayesian Belief Networks. Section 4 summarises the published research into adapting BBNs with a dynamic or temporal dimension. Section 5 proposes an alternative approach combining BBNs with survival analysis. Section 6 explores how this can be used for fault management and section 7 finally ends the paper with a conclusion and future work.

## 2 Fault Management

High-speed broadband telecommunication systems are built with extensive redundancy and complex management systems to ensure robustness. The presence of a fault may not only be detected by the offending component and its parent but the consequence of that fault discovered by other components. This often results in a nett effect of a large number of alarm events being raised and cascaded to the element controller.

The behaviour of the alarms is so complex it appears non-deterministic [5]. It is very difficult to isolate the true cause of the fault. Failures in the network are unavoidable but quick detection and identification of the fault is essential to ensure robustness. To this end the ability to correlate alarm events becomes very important.

The major telecommunication equipment manufacturers deal with alarm correlation through alarm monitoring, filtering and masking as specified by ITU-T [17] and other international standard bodies, with rule-based type systems for assistance to the operator. Yet often it is left to the operator's expertise to determine the actual fault or multiple-faults from the filtered set of alarms reported.

At the heart of alarm event correlation is the determination of the cause. The alarms represent the symptoms and as such, in the global scheme, are not of general interest once the failure is determined [14]. There are two real world

concerns: (1) the sheer volume of alarm event traffic when a fault occurs; (2) the cause not the symptoms.

Alarm monitoring, filtering and masking meets criterion (1), which is vital. They focus on reducing the volume of alarms but do not necessarily meet criterion (2) to determine the actual cause - this is left to the operator to resolve from the reduced set of higher priority alarms. Ideally, a technique that can tackle both these concerns would be best.

### 3 Bayesian Belief Networks (BBNs)

Bayesian Belief Networks (BBNs) offer a potential solution. BBNs consist of a set of propositional variables represented by nodes in a directed acyclic graph. Each variable can assume an arbitrary number of mutually exclusive and exhaustive values. Directed arcs (arrows) between nodes represent the probabilistic relationships between nodes. The absence of a link between two variables indicates independence between them given that the values of their parents are known. In addition to the network topology, the prior probability of each state of a root node is required. It is also necessary, in the case of non-root nodes, to know the conditional probabilities of each possible value given the states of parent nodes or direct causes. A good illustration of a BBN and its related joint probability distribution is contained within Lauritzen and Spiegelhalter's paper [25] where they consider an example based on doctors' diagnoses of patients suffering from shortness of breathe (*dyspnoea*), Figure 1.

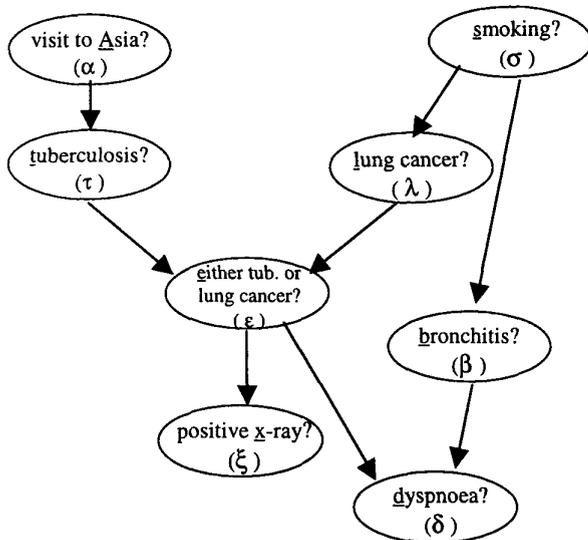


Figure 1: Lauritzen and Spiegelhalter's illustration of a BBN

The graph can be considered as representing the joint probability distribution for all the variables. In the above example this is  $P(\alpha, \tau, \epsilon, \delta, \lambda, \beta, \sigma)$ . The chain rule can re-express this joint probability as the product of the conditional probabilities which need to be specified for each variable or node.

$$P(\alpha, \tau, \epsilon, \delta, \lambda, \beta, \sigma) = P(\alpha)P(\tau|\alpha)P(\xi|\epsilon)P(\epsilon|\tau, \lambda)P(\delta|\epsilon, \beta)P(\lambda|\sigma)P(\beta|\sigma)P(\sigma) \quad (1)$$

The chain rule is given below:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | pa(X_i)) \quad (2)$$

where  $pa(X_i)$  is the parent set of  $X_i$ .

Each node has associated with it a conditional probability table that quantifies the effects that the parents have on the node. Taking the graph as a whole, the conditional probabilities and the structure can be used to determine the marginal probability or likelihood of each node holding one of its states.

The power of the BBN comes to light whenever we change one of these marginal probabilities. The effects of the observation are propagated throughout the network and the other probabilities updated. In simple networks the marginal probabilities or likelihood of each state can be calculated from the knowledge of the joint distribution, shown earlier, using the product rule and Bayes' theorem. This simply means that the directed acyclic graph (DAG) is singly connected; each link is a bridge where the removal of one leads to a disconnected network.

However, cycles often occur and the calculation is much more complex. Algorithms have been devised to cope with the complication of such cycles. Some calculate the marginal probabilities exactly but by doing so introduce calculations which are NP-hard [6]. Therefore many researchers have developed algorithms which approximate the marginal probabilities. This may lead to the compromise of accuracy over a lower computational overhead.

The BBN can be used for deduction in the fault management domain. Given alarm data it will determine the most probable cause(s) of the supplied alarms, thus enabling the system to act as an expert system.

In previous work [37] have developed an exact algorithm approach to deduce the marginal probabilities for their BBN application based on that developed by Lauritzen and Spiegelhalter [25].

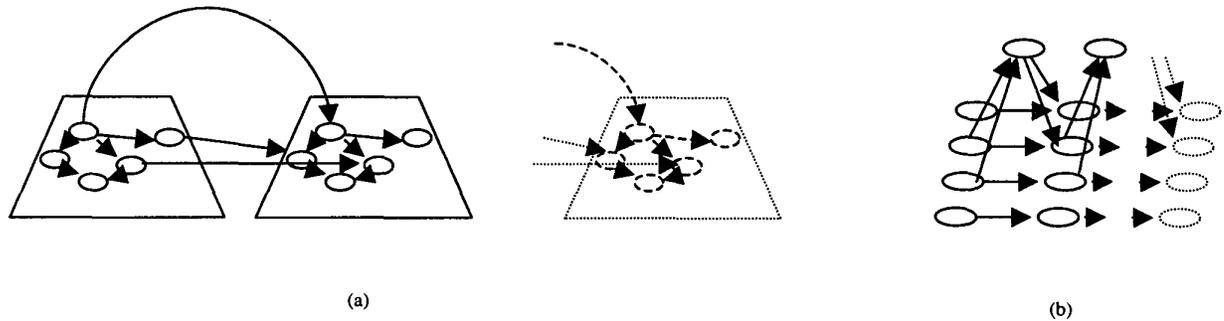


Figure 2: Diagram showing the difference between models based on time slices (adapted from Hanks [13]).

#### 4 Dynamic Belief Networks (DBNs)

A problem with the standard theory of belief networks is that there is no natural mechanism for representing time [1], [2], [39]. There have been various efforts to extend the theory to allow time to be modelled. For example, where probability of movement from one state to another has a temporal dependency, survival analysis [27] can be used. Constraints on the behaviour of the system can be expressed using the formalism of temporal logic.

Dynamic Belief Networks (DBN) model a system that is dynamically changing or evolving over time [10],[21],[22], [30]. This model will enable the user to monitor and update the system as time proceeds.

The literature tends to use the terms dynamic and temporal interchangeably. Dynamic is the opposite of static, it concerns a motive force. Changing the nature of the static BBN to model 'motive forces' can then be classified as adapting it to a dynamic model. Although systems with change involve time, we differentiate between the two terms dynamic and temporal in that temporal models explicitly model time as opposed to other changes in the system such as the change in state. As such temporal models would be a sub-class of dynamic.

Bayesian Belief Networks were not designed to model temporal relationships explicitly; they are a static model. The prediction and deduction made do not vary depending on when the observations or predictions are made.

This standard theory of belief networks has been further developed by researchers to incorporate a temporal feature or time element into the model. This has been approached in various different ways. Aliferis and Cooper [1] summarise just some of the extensions of belief networks for time modelling presented over the last few years. These include temporal influence diagrams [33], Dynamic Belief Networks (DBNs) [7], temporal models of endogenous

change [13], Temporal Bayesian Networks (TBNs) [39], Temporal Nodes Bayesian Network (TNBNs) [3], embedded Markov processes [4], logic and time nets [19],[20], Modifiable Temporal Belief Networks (MTBNs) [1], as well as specific applications [4], [30].

An obvious way of classifying the literature is to use the authors' individual terms (as above) to describe the various approaches. However, in most cases these terms, mainly dynamic and temporal, are interchangeable. For instance, if the time slices of a temporal model were applied so that the movement between the slices was based on a change in state instead of time, we could then classify them as belonging to a dynamic model. Likewise if Markov-chain approaches, dynamic models, were implemented that each state was a point in time, we could classify the applications as temporal models. Therefore we can say that the difference is primarily dependent on the application of the model.

Another approach of classifying the literature is to consider how the methods actually model the time/temporal element. This classification has been used by Palmer et al. [31] who divide the temporal approaches into two main categories of time representation, namely those models which represent time (1) as points or instances or (2) as time intervals.

Within category (1) the models, based on points in time, require that events occur instantaneously where each event considered occurs at an instant in time. These are basically the time slice models and temporal reasoning models reviewed by Hanks et al. [13] and illustrated in Figure 2.

Figure 2(a) represents an approach where a *time slice* is used to represent a snapshot of the evolving temporal process [21]. The belief network consists of a sequence of sub-models each representing the system at a particular point or interval in time (time slice) and which are interconnected by temporal relations. Kjaerulff [22], Dagum et al.[7][8], Provan[32], Berzuini[4], Lekuona[26] are just some of the researchers currently using the time slice approach.

Figure 2(b) represent models where the network is composed of sub-models and duplicated over time slices, as before. However links between state variables within a time slice are disallowed. Dean and Kanazawa [9] and Kanazawa [19][20] use this approach in their research.

Hanks et al. [13] proposed a modification to the time slice approach where they take into account the system as it changes over time, both due to exogenous and endogenous influences.

Category (2) of the classification approach in Palmer et al. [31] considers interval representations of time. Allen's interval algebra and its 13 relations were used to provide the temporal basis for the model [2]. This may be more broadly thought of as a dynamic model as an interval in time represents an event or process during which a property (either true or false) holds uniformly throughout. Examples of work in this area are the Temporal Abduction Problem (TAP) [35] and the Probabilistic Temporal Network (PTN) [39].

An alternative to the above approaches is introduced in the next section.

## 5 A Dynamic Bayesian Belief Network Approach (DBBNs)

A new approach currently being researched and applied to geriatric patient management [28] is that of combining BBNs and Survival Analysis to create a Dynamic Bayesian Belief Network (DBBN).

DBBNs are described as generalising the concept of BBNs to include a time dimension. The approach represents a stochastic (or probabilistic) process along with causal information [9],[34]. Heckerman et al. [16] has also introduced a temporal component to BBNs by providing a temporal definition of causal dependence where he associates a set of variables indexed by time with each cause and with an effect.

In statistical theory, Markov models are often used to represent stochastic processes. Structured phase-type (Ph) distributions [29] characterise a type of latent Markov model which provide an intuitive and robust way of describing probabilistic processes. Such models describe duration until an event occurs in terms of a process consisting of a sequence of latent phases - the states of a latent Markov model. For example, duration of stay in hospital can be thought of as a series of transitions through phases such as: acute illness, intervention, recovery or

discharge. The representation of such a process in terms of latent phases is realistic, as that is how a domain expert conceptualises the process. It is also mathematically suitable since we can prove that any such statistical distribution may be represented arbitrarily closely by one of phase-type form [10].

In this approach we combine the advantages of BBNs in incorporating prior knowledge and causation into the model with the elegant and intuitive process representation of phase-type distributions (Figure 3).

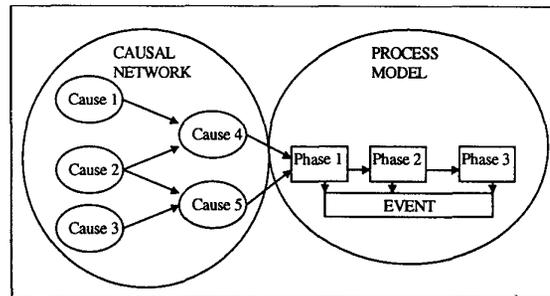


Figure 3: The underlying representation describes a DBBN in terms of a number of interrelated causal nodes which temporally precede and predetermine (in a probabilistic sense) the effect node(s) which constitute the process. The effect node(s) here are characterised by a continuous positive random variable(s), the duration, described by a phase-type distribution

The Causal Network is modelled as a BBN. The Process Model may be defined in a manner similar to that of [9], [13] where we consider an event  $\langle E \rangle$  which initiates a process  $P$  at time zero and  $\langle P, t \rangle$  indicates that the process  $P$  is active at time  $t$ . Then  $\text{prob}\langle P, t \rangle$  is the probability that the process is still active at time  $t$ . In statistical terminology,  $\text{prob}\langle P, t \rangle$  is known as the *survivor function*, denoted by  $F(t)$  and, for a continuous time representation, its derivative  $f(t)$  is the *probability density function* (p.d.f.) of the time for which the process is active. Here we define  $f(t)$  by  $f(t)\Delta t = \text{prob}(\text{process terminates in } (t, t+\Delta t) \mid \text{process is still active at } t)$ .

We thus assume that the model includes variables, some of which are qualitative (the causal variables) and some quantitative (the survival variables). Some previous work has been done on data of this sort, mainly involving the introduction of conditional Gaussian (CG) distributions [12], [24]. We here introduce the idea of Conditional Phase-type (C-Ph) distributions which are more appropriate for process data.

## 6 A Discussion of Potential Fault Management Applications

Downtime in a network not only results in loss of revenue but can lead to serious financial contractual penalties. It is therefore not surprising that network operators are extremely keen to remedy faults as quickly as possible. To this end not only is identification of the fault critical but an estimation of a fault's likely life span would greatly assist in managing and assessing maintenance strategies.

Fault management is an important but difficult area of telecommunications network management. Networks produce large amounts of alarm information that must be analysed and interrupted before the faults can be located [23]. As has been stated earlier alarm correlation is the central technique in fault identification [18].

The instance of a fault can cause numerous alarm events to be raised from an individual network element (NE), this means that the alarms are often interrelated. Also a fault may trigger numerous similar and different alarms to be generated in different NE's up or down stream on the network. For example, the Comms fail alarm, an alarm raised by the management system if it cannot maintain a communications channel to the indicated NE, may cause other alarms such as RS-LOS, RS-LOF, Qecc-Comms\_fail, MS-EXC or even laser alarms depending on the fault and configuration.

Correlation serves to diminish the number of alarms presented to the operator, yet ideally the approach should be able to facilitate fault prediction;

- Fault identification/diagnosis - prediction of the fault(s) that have occurred from the alarms present
- Behaviour prediction – warn the operator before hand of severe faults from the alarms that are presenting themselves.

The Bayesian paradigm and its extensions offer the machinery to achieve these ideals. Although methods such as Artificial Neural Networks (ANNs) have been proven to obtain good predictive performance, they do not meet one important goal; that of comprehensibility. Telecommunication companies do not wish to install 'black boxes' into their fault management systems therefore ruling out ANNs [15]. BBN's graphical structure more than meet the need for 'readability'.

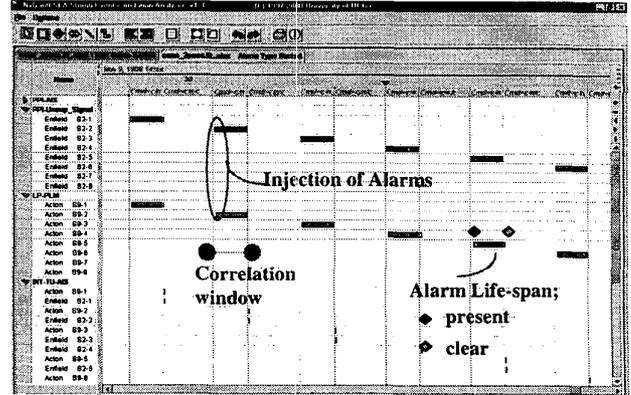


Figure 4: SDH Alarm data viewed over time. Screenshot of NxGantt [38] with comments displaying an alarm's lifespan (horizontal Gantt bars), how close the injection of 2 alarms may occur in time and the correlation window.

BBNs, DBNs and DBBNs can all be applied to fault management of telecommunications. Below, we discuss how BBNs may be developed, refined by DBNs and further enhanced by DBBNs.

When an alarm occurs in a network it is "present" until its accompanying "clear" arrives thus implying a temporal life span (Figure 4) and a correlation window.

The inducing of this alarm data into a static BBN (Section 3) then provides the 'guts' of an expert system, for answering "if then" questions exploring the effects of changing variable values. For example, if Alarm type LP-PLM is observed, this alters the probability (among others) that alarm PPI-Unexpl\_Signal will be observed. [36] and [37] describes an architecture that induces a BBN from this data inferring from it the likely alarm behaviour (Figure 5).

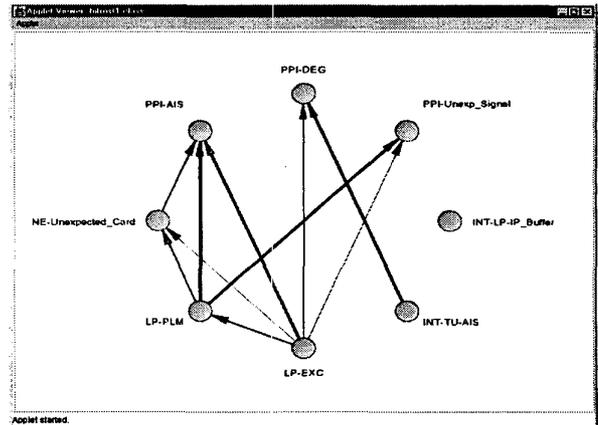


Figure 5: Alarm BBN for Fault Management

TBNs or DBNs (Section 4) offer an opportunity to be more precise when predicting the fault by adding a temporal dimension to the model, since the alarms have a lifespan (Figure 4) and the network changes in state under fault conditions. The alarms that are correlated to produce a higher priority alarm may be correlated with other alarms in a later slice – narrowing to a prediction of likely faults.

The DBBN (Section 5) could offer the previously mentioned expert system in the form of the BBN with additional benefits of extra predictions of how long until the fault occurs (in the case of behaviour prediction) or how long until the fault is repaired (in the case of fault identification). These additional predictions come from the inclusion of survival analysis into the model.

Once the phase-type distribution has been modelled from suitable available data it may be possible to adapt the DBBN model for more precise fault prediction. The incorporation of phase-type variables within the BBN could contribute to a more realistic network where identification of the phase containing the evolving fault episode, would strengthen/weaken the time based prediction options.

## 7 Conclusion

This paper has deliberated the Bayesian paradigm and reviewed the literature on its dynamic extensions. It has proposed a new dynamic approach by incorporating survival analysis as part of the model.

Included is a brief discussion on the potential applications of these models for intelligent fault management in complex telecommunications systems.

The paper has demonstrated the potential power of the dynamic approach for fault identification and behaviour prediction, for instance the ability to determine the likelihood of an alarm being set off at a particular point in time due to a fault occurring at a precise moment in the past.

In fault management there are two real world concerns: (1) the sheer volume of alarm event traffic when a fault occurs; and (2) the cause not the symptoms. The rule-based type systems (monitoring, filtering and masking) used in telecommunication systems address the first. The approaches discussed in this paper would address both concerns paving the way to true intelligent fault management.

## Acknowledgements

We are very grateful to NITEC, Nortel Networks for their support, in particular to the people that are more closely involved in the project for the efforts they make in this collaboration. We also acknowledge DHFETE's support for Adele Marshall.

## References

- [1] C. F. Aliferis, G. F. Cooper, "A Structurally and Temporally Extended Bayesian Belief Network Model: Definitions, Properties, and Modeling Techniques", *Proc. of the 12<sup>th</sup> Conference on Uncertainty in Artificial Intelligence*, pp.28-39, 1996.
- [2] J.F. Allen, "Maintaining Knowledge About Temporal Intervals", *Comms of the ACM*, Vol. 26(11), pp.832-884, 1983.
- [3] G. Arroyo-Figuerao, L. E. Sucar, A. Villavicencio, "Probabilistic Temporal Reasoning and its Application to Fossil Power Plant Operation", *Expert Systems with Applications*, Vol. 15, pp.317-324, 1998.
- [4] C. Berzuini, R. Bellazzi, S. Quaglioni, "Temporal Reasoning with Probabilities", *Proc. of Workshop on Uncertainty in Artificial Intelligence*, pp.14-21, 1989.
- [5] A. T. Bouloutas, S. Calo, A. Finkel, "Alarm Correlation and Fault Identification in Communication Networks", *IEEE Transactions on Communication*, Vol. 42, No 2/3/4, 1994.
- [6] G. F. Cooper, "The Computational Complexity of Probabilistic Inference Using Belief Networks", *Artificial Intelligence*, Vol. 42, pp.393-405, 1990.
- [7] P. Dagum, A. Galper, E. Horvitz. "Dynamic Network Models for Forecasting", *Proc. of the 8<sup>th</sup> Workshop on Uncertainty in Artificial Intelligence*, pp.41-48, 1992.
- [8] D. Dagum, A. Galper, "Forecasting Sleep Apnea with Dynamic Network Models", *Proc. of Uncertainty in Artificial Intelligence*, pp.64-71, 1993.
- [9] T. Dean, K. Kanazawa, "A Model for Reasoning about Persistence and Causation", *Computational Intelligence*, Vol. 5(3), pp.142-150, 1989.
- [10] M. Faddy, "Examples of Fitting Structured Phase-Type Distributions", *Applied Stochastic Models and Data Analysis* Vol. 10, pp.247-255, 1994.
- [11] N. Friedman, K. Murphy, S. Russell, "Learning the Structure of Dynamic Probabilistic Networks", *Proc. of the Conf. on Uncertainty in Artificial Intelligence*, 1998.
- [12] A. Gammerman, Z. Luo, C.G.G. Aitken, M.J. Brewer, "Computational Systems for Mixed Graphical Models. Adaptive Computing and Information Processing". *UNICOM Seminars Ltd, Brunel Conference Centre, London*, 1994.

- [13] S. Hanks, D. Madigan, J. Gavrin, "Probabilistic Temporal Reasoning with Endogenous Change", *Proc. of the 11<sup>th</sup> Conf. on Uncertainty in Artificial Intelligence*, pp.245-254, 1995.
- [14] K. Harrison, "A Novel Approach to Event Correlation", *Intelligent Networked Computing Lab, HP Labs*, HP-94-68, pp. 1-10, 1994.
- [15] K. Hatonen, M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, "Knowledge Discovery from Telecommunication Network Alarm Databases", *Proc. 12<sup>th</sup> Int. Conf. on Data Engineering (ICDE'96)*, pp.115-122, 1996.
- [16] D. Heckerman, J. S. Breese, "A New Look at Causal Independence", *Proc. of the 10<sup>th</sup> Conf. on Uncertainty in Artificial Intelligence*, pp.286-292, 1997.
- [17] ITU-T Recommendations M.3030 "Principles for a Telecommunications Management Network", 1988.
- [18] G. Jakobson, M.D. Weissman, "Alarm correlation", *IEEE Network*, 7(6), pp52-59, November 1993.
- [19] K. Kanazawa, "A Logic and Time Nets for Probabilistic Inference". *Proc. of the 10<sup>th</sup> National Conference on Artificial Intelligence, AAAI* pp.360-365, 1991.
- [20] K. Kanazawa, "Reasoning about Time and Probability" Thesis, *Brown University*, May 1992.
- [21] K. Kanazawa, D. Koller, S. Russell, "Stochastic Simulation Algorithms for Dynamic Probabilistic Networks", *Proc. of the 11<sup>th</sup> Annual Conference on Uncertainty and Artificial Intelligence*, 1995.
- [22] U. Kjaerulff, "A Computational Scheme for Reasoning in Dynamic Probabilistic Networks", *Proc. of the 8<sup>th</sup> Conference on Uncertainty in Artificial Intelligence*, pp. 121-129, 1992.
- [23] M. Klemettinen, "A Knowledge Discovery Methodology for Telecommunication Network Alarm Databases", PhD Thesis, University of Helsinki, Finland, 1999
- [24] S. L. Lauritzen, N. Wermuth, "Graphical Models For Associations Between Variables, Some of Which are Qualitative and Some Quantitative", *Annals of Statistics*, Vol. 17, pp.31-57, 1989.
- [25] S. L. Lauritzen, D. J. Spiegelhalter, "Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems", *J. R. Statist. Soc. B*, Vol. 50(2), pp.157-224, 1988.
- [26] A. Lekuona, B. LaCruz, P. Lasala, "On Graphical Models for Dynamic Systems", *Proc., AI and Statistics*, pp.317-323, 1995.
- [27] A. H. Marshall, S. I. McClean, C. M. Shapcott, P.H. Millard, "Using Bayesian Belief Networks to Predict the Survival of Stroke Patients", *Proc. of the IX International Symposium on Applied Stochastic Models and Data Analysis (ASMDA-99)*, pp.112-117, 1999.
- [28] A. H. Marshall, S. I. McClean, C. M. Shapcott, P.H. Millard, "Learning Dynamic Bayesian Belief Networks using Conditional Phase-Type Distributions", *Conference on Principles and Practice of Knowledge Discovery in Databases - PKDD 2000* (to appear).
- [29] M. Neuts, "Structured Stochastic Matrices of M/G/1 Type and Their Application", *Marcel Dekker*, New York, 1989.
- [30] A. E. Nicholson, J. M. Brady, "Dynamic Belief Networks for Discrete Monitoring". *IEEE Trans. on Systems, Man and Cybernetics*, Vol. 34(11), pp.1593-1610, 1994.
- [31] F. L. Palmer, R. Sterritt, C. M. Shapcott, E. P. Curran, K. Adamson, "Exploring Dynamic Belief Network Visualisation", *Conference on Artificial Intelligence and Soft Computing - ASC 2000*, to appear.
- [32] G. M. Provan, "Model Selection for Diagnosis and Treatment using Temporal Influence Diagrams" *Proc International Workshop on AI and Statistics*, pp.469-480, 1993.
- [33] G. M. Provan, "Tradeoffs in Constructing and Evaluating Temporal Influence Diagrams", *Proc. of the 9<sup>th</sup> Conference on Uncertainty in Artificial Intelligence*, pp. 40-47, 1993.
- [34] S. Russell, J. Binder, D. Koller, K. Kanazawa, "Local Learning in Probabilistic Networks with Hidden Variables", *Proc. 14th Int. Joint Conf. on AI*, pp.1146-1152, 1995.
- [35] E.(Jr.) Santos, "Unifying Time and Uncertainty for Diagnosis", *J. Experimental and Theoretical Artificial Intelligence*, Vol. 8, pp.75-94, 1996.
- [36] M. Shapcott, R. Sterritt, K. Adamson, E. P. Curran, "NETEXTRACT - Extracting Belief Networks in Telecommunications Data", *Proceedings of the ERUDIT Workshop on Application of Computational Intelligence Techniques in Telecommunication*, pp63-71, 1999
- [37] R. Sterritt, K. Adamson, C. M. Shapcott, D. A. Bell, "An Architecture for Knowledge Discovery in Complex Telecommunication Systems", (eds) Adey R.A., Rzevski G., Nolan P., *Applications of Artificial Intelligence in Engineering XIII*, Computational Mechanics Publications, pp.627-640, 1998.
- [38] R. Sterritt, E. P. Curran, K. Adamson, C. M. Shapcott, "Visualisation for Data Mining Telecommunications Network data", *Data Mining II*, (eds.) F.F. Ebecken, C.A. Brebbia, A. Weigend, WIT Press, Southampton UK, (2000).
- [39] J. D. Young, E. Santos, "Introduction to Temporal Bayesian Networks", Presented at the 7<sup>th</sup> Midwest AI and Cognitive Science Conf., 1996.