Discovering Rules for Fault Management

Discovering Rules for Fault Management

Roy Sterritt
School of Information and Software Engineering, Faculty of Informatics, University of Ulster
r.sterritt@ulst.ac.uk

Abstract

At the heart of the Internet revolution is global telecommunication systems. These systems, initially designed for voice traffic, provide the vast backbone bandwidth capabilities necessary for Internet traffic. They have built-in redundancy and complexity to ensure robustness and quality of service. To facilitate this, this requires complex fault identification and management systems. Fault identification and management is generally handled by reducing the amount of alarm events (symptoms) presented to the operating engineer through monitoring, filtering and masking. The ultimate goal is to determine and present the actual underlying fault. While en-route to automated fault identification it is useful to derive rules and techniques to attempt to present less symptoms with greater diagnostic assistance. With these objectives in mind computer-assisted human discovery and human-assisted computer discovery techniques are discussed.

1. Introduction

1.1. 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 net 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 [1]. 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 [2] 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.

1.2. Event Correlation

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 [3]. 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 meet 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.

1.3. Towards Intelligent Fault Management

A technique that can suggest the fault and not just deal with the sheer volume of alarm event traffic when a fault occurs would be ideal. AI offers that potential and has been and still is an active and worthy area of research to assist in fault management.

Yet telecommunication manufacturers have shown reluctance in incorporating AI, in particular those techniques that have an 'uncertainty' element, directly into their critical systems. Rule-based type systems have achieved acceptance largely because the decisions obtained are deterministic, they can be traced and understood by domain experts

1.4. Rule Development and Maintenance

The time to market and the R&D lifecycle of these products are continuously being squeezed while at the same time market demands for features and functionality

0-7695-1086-8/01 $10.00 © 2001 IEEE

increase with each release. It is also the nature of the domain that customers expect legacy support with the new systems as well as multi-vendor support.

This not only creates challenges for rule-based systems development but also creates a substantial rule-base maintenance burden [4]. As such techniques to assist in discovery and development of rules in heterogeneous network environments is also essential.

1.5. The Challenge

There is a predicament. AI would seem to offer the potential of achieving automated fault diagnosis for these complex systems while there is doubt if it would be accepted as the engine within the fault management system.

While en-route to automated fault identification and with the previously mentioned domain challenges in mind it is useful to utilise AI to derive rule discovery techniques to attempt to present less symptoms with greater diagnostic assistance.

As such computer-assisted human discovery and human-assisted computer discovery techniques are discussed.

2. Human and Computer Discovery

It may be proposed that a flaw in data mining or Knowledge Discovery (KD) is that it is not user-centered. It would be helpful to visualise the data at all stages to enable the user to gain trust in the process and hence have more confidence in the mined patterns. The transformation from data to knowledge requires interpretation and evaluation, which also stands to benefit from multi-stage visualisation of the process [5].

2.1. Computer-aided Human Discovery

The aim is to discover hidden knowledge, unexpected patterns and new rules from data mountains. Visualisation techniques of vast amounts of data allow the remarkable perceptual abilities that humans possess to be utilised, such as the capacity to recognise images quickly, and detect the subtlest changes in size, colour, shape, movement or texture - and thus potentially discover new event correlations in the data.

2.2. Human-aided Computer Discovery

Data mining (discovery algorithms) may reveal hidden patterns and new rules yet these require human interpretation to transform them into knowledge.

The human element attaches a more meaningful insight into the decisions allowing the discovered correlations to be coded as useful rules for fault identification and management.

2.3. A Three-Tier Discovery Process

Computer-assisted human discovery and human-assisted computer discovery can be incorporated together via a three tier process, specifically providing a mechanism for discovery and learning of rules for fault management.

The tiers are;

Tier 1 - Visualisation Correlation
Tier 2 - Knowledge Acquisition or
Rule Based Correlation
Tier 3 - Knowledge Discovery Correlation

The top tier (visualisation correlation) allows the visualisation of the data in several forms. The visualisation has a significant role throughout the knowledge discovery process, from data cleaning to mining. Therefore allowing analysis of the data with the aim of identifying other alarm correlations (knowledge capture). The second tier (knowledge acquisition or rule-based correlation) aims to define correlations and rules using more traditional knowledge acquisition techniques - utilising documentation and experts. The third tier (knowledge discovery correlation) mines the TMN (Telecommunications Management Network) data to produce more complex correlation candidates.

The application of the 3-tier process is iterative and flexible in nature. The visualisation tier may require the knowledge acquisition tier to confirm its findings. Likewise visualisation of the Knowledge Discovery (KD) process could facilitate understanding of the patterns discovered.

3. Correlations and Discoveries - The Process

3.1. Correlation Discovery via Visualisation

Knowledge Discovery is considered to be "the non-trivial extraction of implicit, previously unknown, and potentially useful information from data" [6]. This implies a focus only on the discovered information, yet the current opinion is that KD means more than this. KD refers to the over-all process of discovering useful knowledge from data, while data mining refers to the application of algorithms for extraction purposes [7]. Brachman and Anand present a process that includes human intervention [8]. Although autonomous KD may be desirable in the long run this is not the current state of affairs. It has therefore been highlighted that KD

researchers need to place more emphasis on the overall KD process and on tools to support its various stages [9].

It is prevalent in the literature that upon engaging in real-world discovery tasks it has been found that they can be extremely complex [8]. Adding visual tools to the process can reduce this complexity by facilitating understanding of the data and patterns discovered.

Visualisation applied to KD can offer 'human-assisted computer discovery' and 'computer-assisted human discovery'. Such a visual environment, by reducing the time to understand complex data, would enable practical solutions to many real world problems to be developed far more rapidly than either human or computer operating independently [9].

As such there are two distinct roles for the visualisation tier; (1) specifically to facilitate human discovery of correlations/potential rules and (2) visualising the KD process (tier 3).

Figure 1 portrays how visualisation tools and techniques can enable discovery of potentially useful correlations between events in the telecom data. Those that hold up to scrutiny could then be developed into rules for a fault management system or other diagnostic tool.

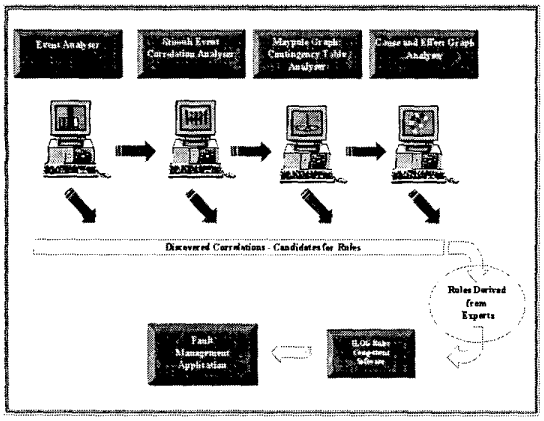

Figure 1. Discovering Rules through Visualisation - generic view of tier 1

The tools developed are;

- Event Analyser - to facilitate the study of the frequency of events
- Stimuli-Event Correlation Analyser - to facilitate the study of the events over time
- Contingency Table Analyser - to facilitate the study of the frequency of events occurring within the same time period
- Cause and Effect Graph Analyser - to facilitate the study of the mined probabilistic network of events.

3.2. Correlation Discovery via Knowledge Acquisition

Many of the problems encountered using traditional Knowledge Acquisition (KA) and Rule-Based Systems (RBS) such as; the KA bottleneck, their inability to handle uncertainty well and the "maintenance burden" have emphasised the success of data mining. Yet there is still a place for it in a discovery process. To move from discovered patterns (event correlations), be they through visualation or data mining, to knowledge (interruption, validation and coded rules) will require consultation with experts and/or documentation.

KA also offers the potential of discovering implicit hidden knowledge from experts or documentation that can also form the basis of rules for a fault management system (Figure. 2).

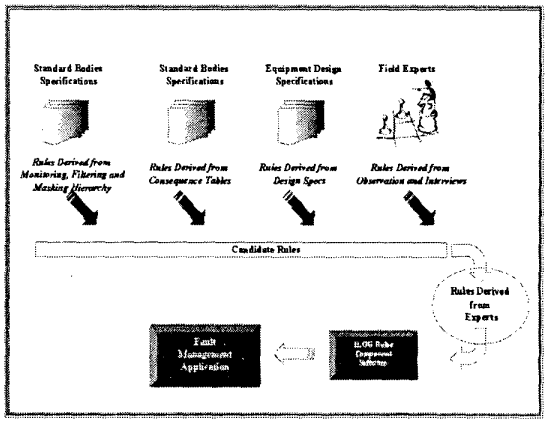

Figure 2. Discovering Rules through Knowledge Acquisition - generic view of tier 2

3.3. Correlation Discovery via Knowledge Discovery and Data Mining

As has been stated data mining deals with the discovery of hidden knowledge, unexpected patterns and new rules from large databases and that it is now generally considered as the discovery stage in a much larger process [9][7] – Knowledge Discovery (KD). These discoveries could be harnessed for a fault management system (Figure 3). Adriaans and Zantinge presents a comprehensive introduction for undertaking Data Mining and KD,

including all the stages; Data selection, Cleaning, Enrichment, Coding, Data Mining, and Reporting [10].

Rules could be written from mined results from such tools as Clementine. For example; *when a Comms fail alarm occurs it is likely a Qecc-Comms_fail alarm will be injected into the network.* This information may then be encoded as a rule.

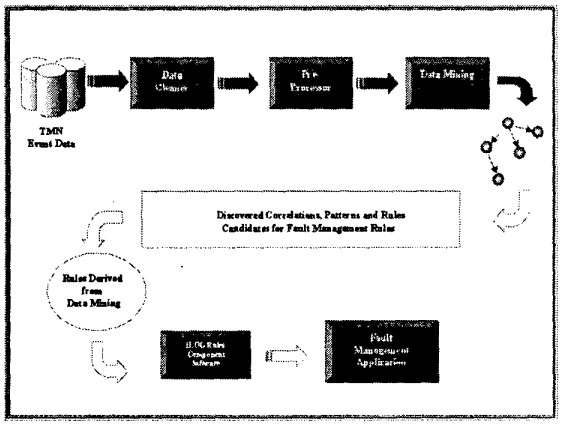

Figure 3. Discovering Rules through Data Mining - generic view of tier 3

3.3.1 Mining rules for other uses. The emphasis so far has been on using the discovered and developed rules for inclusion in a fault management system. Yet the approach is open to developing rules for different implementations for instance a rules system for the testing environment.

Complementary research [11] which data mines probabilistic networks / cause and effect graphs [12] (as opposed to rules) does deal with the previously mentioned criteria (volume of alarms and cause not the symptoms). Yet the approach does have its problems that could benefit from a pre-processing alarm correlator [13].

The cause and effect graph can be considered a complex form of alarm correlation. The alarms are connected by edges that indicate the probabilistic strength of correlation. Yet the cause and effect network can contain more than just alarms as variables - actual faults can be included as variables.

Data Mining is used to produce the probabilistic network by correlating offline alarm event data, and deducing the cause using this probabilistic network from live alarm events.

In this case, as in many cases, the structure of the graphical model (the Bayesian belief network - a specialised form of probabilistic network) is not known in advance, however a database is available which includes information concerning the frequencies of occurrence of combinations of different variable values (the alarms). Therefore the problem is that of induction – to induce or learn the structure from the data. Heckerman details a good description of the problem [14][15].

There has been a lot of research involved with the induction of probabilistic networks for example Cooper and Herskovits' algorithm [16]. Unfortunately the general problem is NP-hard [17]. For a given number of variables there is a very large number of potential graphical structures which can be induced. To determine the best structure, then in theory, one should fit the data to each possible graphical structure, score the structure, and then select the structure with the best score. Consequently algorithms for learning networks from data are usually heuristic, once the number of variables gets to be of reasonable size.

In practice, when it comes to learning the cause and effect graph, the volume of event traffic and correlation of alarms can be reduced by simple first stage correlation (the discovered rules defined from all tiers). The expert system approach (in this case the deduction from the probabilistic network) could then handle the remaining more complex problems, taking advantage of the much reduced and enriched stream of events.

3.3.2 Developing rules from the probabilistic network. Another source of rules is to actually extract correlations from the induced probabilistic network for those variables that have an exceedingly high probability of cause and effect.

4. Case Study

4.1. An experiment - inducing simple commands to simulate faults

The following case study demonstrates the simulation of simple faults into a test network via a command line user Interface (CLUI) on the element controller. The network consists of two multiplexers named Enfield and Acton. The faults were induced on Enfield. The sample commands shown demonstrate the disconnection of a tributary port 1 (in slot 2 of the multiplexer) then its reconnection after a time period, followed by the disconnection and reconnection of port 2. In the sample test, ports 3 - 8 were also disconnected then after a time period reconnected in the same way.

```
Cmd=c/n/d S6-k111&S7-k111 S2-1
Cmd=c/n/c S7-1-J1-K111&S6-1-J1-K111 S2-1
Cmd=c/n/d S6-k112&S7-k112 S2-2
Cmd=c/n/c S7-1-J1-K112&S6-1-J1-K112 S2-2
```

Thus in total 16 commands were performed (8 sets of disconnection and reconnections). Table 1 displays a

breakdown of the event types that were recorded in the event log on the element controller during this test. Note no other activity was occurring on the network during the experiment.

Table 1. Breakdown of recorded events during experiment

Event Type	Total
Alarm Events	476
Login Events	106
User Action Events	16
Message Tool Events	159
System Error Events	1
Total number of events	758

For 16 commands (recorded as user action events) 476 alarms instances occurred, yet only 5 actual alarm types transpired which are shown in table 2.

Table 2. Alarm types that occurred during the experiment

Alarm Event Type	Explanation
PPI-AIS	PDH Physical Interface - Alarm Indication Signal
PPI-Unexp_Signal	PDH Physical Interface - Unexpected Signal
LP-PLM	Lower order Path - Path Label Mismatch
INT-TU-AIS	Internal - Tributary Unit - Alarm Indication Signal
INT-TU-LOP	Internal - Tributary Unit – Loss of Pointer

4.2. Visual Correlation - Computer-assisted Human Discovery

In this simple experiment, approximately 10,100 lines were recorded in the event log as 758 event records. It can be easily envisaged that the event log grows into a data mountain over a relatively short time period and thus requires visualisation and mining techniques.

Figure 4 displays a possible correlation discovery through using one of the developed visualisation tools [5]. The alarms' life span (horizontal Gantt bars) is between the commands disconnect and connect (vertical command bars) indicating a causal effect. Also note that when PPI-Unexp_Signal is active on Enfield so is LP-PLM on Acton presenting a possible correlation.

On investigating the standards specifications it is found that a PPI-Unexp-Signal causes no impacts nor has no consequent actions. LP-PLM affects traffic and can have consequent actions of injecting an AIS and LP-RDI alarm depending on configuration (consequent actions for LP-PLM can be enabled/disabled, the default being disabled). Thus there is no explicit connection defined for these two alarms.

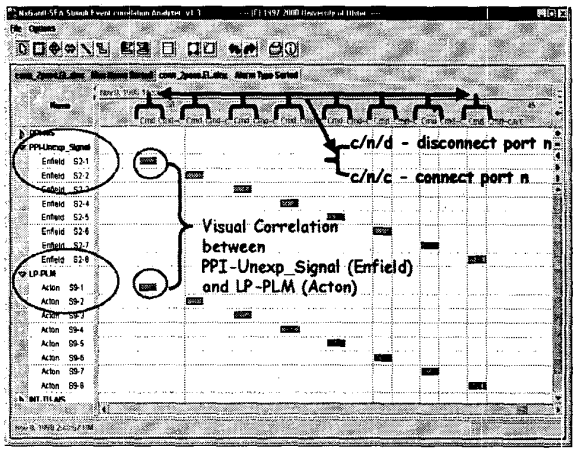

Figure 4. Screenshot of NxGantt with comments displaying a human discovered correlation

4.3 Defining a rule from the discovered Correlation

Since this human discovered correlation is an unexpected pattern it may be considered of interest and be coded as a rule for a commercial fault management system or other diagnostic tool. The First stage Alarm Correlator (FAC) [13] is a part of a prototyped fault management system where simpler correlations that initially tended to be defined from knowledge acquisition (tier 2 - figure 2) are handled. In the prototype more complex correlations would be passed onto the deduction component (inference engine) that has causal information about the alarms (a Bayesian Belief Network -BBN) at its core. The causal information is initially produced and updated via data mining (tier 3 - figure 3). The rule below demonstrates how the discovered correlation can be coded for the FAC.

```
rule portDisabled
{
 when
 {
  ?x Form(alarm==PPI-Unexpl_Signal;
port==?p; mux==?a);
  ?y Form(alarm==LP-PLM; port==?p;
mux==?b);
 }
 then
 {
  retract ?x
  retract ?y
  assert ?z Form(JigAlm-portDisabled,
port==?p, mux==?int);
 }
}
```

The rule states that when PPI-Unexpl_Signal and LP-PLM occur together on the same port number but different multlplexers then correlate these alarms raising an internal alarm JigAlm-portDisabled. The internal alarm in this prototype is a trigger to provide diagnostic assistance.

Since the rule is coded using ILOG Rules it is feasible that it could be adapted and used in any of the commercial fault management systems that use ILOG Rules as their chosen component library.

4.4. Data Mined Correlation - Human-assisted Computer Discovery

The induction of a probabilistic network from the set of data produced the following results (Figure 4 & Table 3).

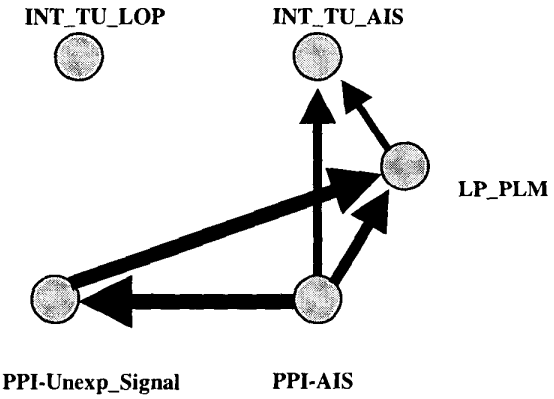

Figure 5. Screenshot of NxGantt with comments displaying a human discovered correlation

The algorithm used was based on Chow and Liu [18] the mutual information between pairs of variables is calculated and those variables with the highest value are connected. The algorithm continues with successive elimination of variables. It has the advantage of simplicity but generates only tree structures.

Table 3. Probabilistic network induction results

Frequencies of Alarm Occurrence
0, PPI-AIS, 192
1, INT-TU-LOP, 0
2, PPI-Unexp_Signal, 8
3, LP-PLM, 8
4, INT-TU-AIS, 15
Strength of Edges (mutual information score)
0 >:2, 0.121383
2 >:3, 0.113237
0 >:3, 0.0832668
0 > 4, 0.0676941
3 > 4, 0.052712
Probabilistic Connections
p(PPI_Unexp_Signal | PPI_AIS)
p(LP_PLM | PPI_AIS, PPI_Unexp_Signal)
p(INT_TU_AIS | PPI_AIS, LP_PLM)

Likely candidates for rule development based on the strength of connection are PPI-Unexplained-Signal & LP_PLM (which matched the visual correlation) and PPI-Unexplained-Signal & PPI-AIS.

Upon consultation with engineers it was discovered that the PPI-AISs strong connections are specific to the test configuration set-up for the experiment since the signal was unstructured and tributaries daisy chained and as such may not be of generic interest.

5. Summary and Conclusion

This paper presented the complexities and open research issues in fault management and identification namely that commercial systems tend to present a reduced set of symptoms of the fault not the actual fault. It then presented a three-tier process to assist in the discovery of new rules that could potentially reduce that set of symptoms further, en-route to automated fault identification and management.

The case study has demonstrated how a simple experiment, which simulated a port fault, raised a relatively sizable amount of event data and as such demonstrated how visualisation, data mining and in general discovery processes (be they computer assisted human discovery or human assisted computer discovery) can assist in fault identification and management.

The study went on to show a discovered correlation via visualisation of the data. Since correlating events is at the heart of fault identification and management this process is significant in that newly discovered knowledge, unexpected patterns and rules can be used to greatly assist

the human operator. Finally it was demonstrated how a rule could then be coded in ILOG from the discovered correlation.

5.1. Future work

The trend is towards automation; but there is reluctance in the telecommunications industry to utilise, say for example UAI (Uncertainty in AI) techniques which may achieve this, directly in the fault management system. The aim here is still to achieve automation but by utilising the UAI techniques along with visualisation in the discovery of higher order rules and correlations for prediction.

Faults are rare and as such fault data is limited from an operational network thus offering a challenge for this research approach. As such the data under study is gathered from a manufacturers' R&D captive office test rig where numerous faults are induced. Even when using only data from verification testing, changes can be expected before customer final release and as such represent a difference in captured behaviour. Therefore it is planned to utilise a new private STM-4 radio network that has been set-up for academic research purposes, providing the certainty of a final product environment.

6. Acknowledgements

The author is greatly indebted to our industrial collaborators Northern Ireland Telecommunications Engineering Centre (NITEC), Nortel Networks, who have supported our research for many years now. We would also like to thank the Industrial Research and Technology Unit (IRTU) [Start-187 - The Jigsaw Programme 1999-2002] for funding this work jointly with Nortel Networks. This paper is based on a presentation [19].

7. References

[1] 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.

[2] ITU-T Receommendations "M.3030 Principles for a Telecommunications Management Network", 1988.

[3] K. Harrison, "A Novel Approach to Event Correlation", HP, Intelligent Networked Computing Lab, HP Labs, Bristol. HP-94-68, July, 1994, pp. 1-10.

[4] I. Bratko, S. Muggleton, "Applications of Inductive Logic Programming", *Communications of the ACM*, Vol. 38, no. 11, 1995, pp. 65-70.

[5] 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,pp445-454.

[6] W. Frawley, G. Piatetsky-Shapiro, C. Matheus, "Knowledge Discovery in Database: An Overview", *AI Magazine* 14(3), 1992, pp57-70.

[7] U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, "From Data Mining to Knowledge Discovery: An Overview", *Advances in Knowledge Discovery & Data Mining*, AAAI Press & The MIT Press: California, 1996, pp1-34.

[8] R.J. Brachman, T. Anand, "The Process of Knowledge Discovery in Databases: A Human-Centered Approach"., *Advances in Knowledge Discovery & Data Mining*, AAAI Press & The MIT Press: California, 996, pp37-57.

[9] R. Uthurusamy, "From Data Mining to Knowledge Discovery: Current Challenges and Future Directions", *Advances in Knowledge Discovery & Data Mining*, AAAI Press & The MIT Press: California, 1996, pp 561-569.

[10] P. Adriaans, D. Zantinge, *Data Mining*, Addison-Wesley: England, 1996.

[11] R. Sterritt, K. Adamson, M. Shapcott, D. Bell, F. McErlean, "Using A.I. For The Analysis Of Complex Systems", *Proc. IASTED International Conference on Artificial Intelligence and Soft Computing* (in co-operation with AAAI), 1997, pp113-116.

[12] R. Sterritt, M. Daly, K. Adamson, M. Shapcott, D.A. Bell, F. McErlean, "NETEXTRACT: An Architecture For The Extraction Of Cause And Effect Networks From Complex Systems", *Proc. of the 15th Int. IASTED Conf. on Applied Informatics*, 1997, pp55-57.

[13] R. Sterritt, C.M. Shapcott, K. Adamson, E.P. Curran, "High Speed Network First-Stage Alarm Correlator", *International Conference Intelligent Systems And Control*, Hawaii, USA, 2000, pp.391-397.

[14] D. Heckerman, "Bayesian Networks for Data Mining", *DM&KD* 1, 1997, pp.79-119.

[15] D. Heckerman, "Bayesian Networks for Knowledge Discovery" (eds). Fayyad UM, Piatetsky-Shapiro G, Smyth P and Uthurusamy R, *Advances in Knowledge Discovery and Data Mining*, AAAI Press / The MIT Press, 1996 pp.273-305.

[16] G.F. Cooper, E. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data". *Machine Learning*, 9, 1992 pp. 309-347.

[17] D.M. Chickering D. Heckerman, "Learning Bayesian networks is NP-hard", *MSR-TR-94-17*, MS Research, Microsoft Corporation, 1994.

[18] C. J. K. Chow, C. N. Liu., "Approximating discrete probability distributions with dependence trees", *IEEE Trans. Information Theory*, Vol. 14(3), pp.462-467, 1968.

[19] R. Sterritt, "Discovering Rules for Fault Management", presentation given at the Jigsaw Research Symposium, Queens University Belfast, April 2000.