International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 29595-29608

http://www.ripublication.com

Understanding Oil & Gas Pipeline Failures and Mitigation Measures Using Bayesian Approach

G. Unnikrishnan[1], Nihal Siddiqui[2] and Shrihari Honwad[3]
[1]*Department of Health, Safety and Environment, University of Petroleum and Energy Studies, Dehradun, India, ukrishnan77@yahoo.com*
[2] *Department Health, Safety and Environment, University of Petroleum and Energy Studies, Dehradun, India, nihal@ddn.upes.ac.in*
[3]*Department of Chemical Engineering, University of Petroleum and Energy Studies, Dehradun, India, shrihari@ddn.upes.ac.in*

Abstract

Pipelines carrying oil and gas poses a serious risk to public and environment and therefore its safety is of paramount importance. Several agencies have been collecting failure data of oil and gas pipelines with the objective of studying the same to improve its performance. Improvements in mitigation measures have resulted in reduction of failures. However the impact of such mitigation measures on pipeline failures have not been studied in detail. This paper presents a Bayesian approach to the situation and by the use of conditional probability statements the cause and effect of mitigation measures have been modelled as a Bayesian Network. The network can be easily tuned for site specific data and simulation of the same can provide a better insight in to the mitigation measures and its impact of the probability of pipeline failure. The diagnostic feature of the network also offers the most probable causal mechanism during incident investigations and helps to arrive at root causes.

Keywords: Pipeline failures, mitigation measures, Bayesian Network, conditional probability

1. Introduction

Pipelines are efficient means of transportation for liquid and gases products. Oil and gas industries have been using pipelines for nearly 100 years. Pipelines carrying hydrocarbons are of particular interest due to the hazardous nature of the products and the impact of the consequences that may occur in case of Loss Of Containment (LOC). Therefore monitoring of pipelines, its parameters and failures are of prime importance to the designers as well as owners. Several agencies have been monitoring pipeline failures for more than 40 years by now and good data is available in the public domain. These data has been used to gain better understanding of pipeline failures and develop measures to reduce and prevent them. There are a number of publications describing analysis of the data by various means including descriptive and probability statistics. The causes of pipeline failures have been identified and the category-wise number of failures are available. However linking of the mitigation measures and its effects have remained a difficult subject due to the complex nature of the system as well as due to the uncertain and qualitative nature of the measures. The main objective of the paper is to develop a causal model of impact of mitigation measures against pipeline failures using Bayesian approach, since by its very nature; conditional probability statements are ideally suited for analyzing causes and effects. It can be seen that Bayesian approach offers a model of the impacts of mitigation measures on pipeline failures which can be used for a better understanding of risks and means of focusing on the right mitigation approaches.

Section 2 of the paper briefly describes the important databases available in the public domain and summary of pipeline failures. Section 3 summarizes the fundamentals of Bayesian Network (BN), Section 4 presents the causal model of mitigation measures on pipeline failures, and Section 5 gives the BN and simulation for the model. Section 6 contains the conclusions and future work.

2. Important Pipeline Databases

2.1 CONCAWE Database:

CONCAWE was established in 1963 by a group of leading oil companies in Europe to carry out research on environmental issues relevant to the oil industry. CONCAWE has been collecting facts and statistics since 1971on incidents and spills related to European cross-country pipelines carrying crude oil or petroleum products. Results are published yearly in a report including a full historical analysis. The database is available in public domain [1].

2.2 EGIG Database

European Gas pipeline Incident data Group (EGIG) was set up in 1982 by a joint effort by six European gas transmission operators to gather data on the unintentional releases of gas in their pipeline transmission systems. Now EGIG is a co-operation between a group of seventeen major gas transmission system operators in Europe and is the owner of an extensive gas pipeline-incident database. The latest report on gas pipeline incidents is for the period from 1970 to 2013 [2].

2.3 US Dot- PHMSA

US Department of Transportation Pipeline and Hazardous Material Safety Administration (PHMSA) collects, maintains and analyzes all types of data related to pipelines in USA and the extensive database is available to public [3]. The data is for gas, Liquefied Natural Petroleum Gases (LNG) and liquid pipelines and starts from 1970.

Other pipeline databases in public domain are from Canada's Alberta Energy Board [4] and British Columbia [5]. Private and subscribed databases are not considered in this study.

The creation and availability of these extensive pipeline-incident databases has helped the pipeline designers and operators to improve safety of the pipeline transmission systems.

2.4 Causes For Pipeline Failures

The paper will not attempt to present the descriptive statistics on pipeline failure data since they are already available in the databases referred above. However, the databases contain category-wise causes for pipeline failures and summary of the same is presented below.

Main causes and sub causes for pipeline failures from the databases are given in Table 1 below:

Table 1: Main Causes and Sub Causes For Pipeline Failures

CONCAWE-Liquid pipelines (1971-2012)		**US Dot PHMSA (2006-2010)**	**EGIG (2009-2013)**
Main causes	**Sub-causes**	**Main causes**	**Main causes**
1.Mechanical failure	Construction fault	1. Mechanical/weld / Equipment failure	1. Construction defects /Material failures
	Materials fault		2. Hot tap
2. Operational	System malfunction	2. Incorrect Operation	
	Human error		
3. Corrosion	External	3. Corrosion	3. Corrosion
	Internal		
	Stress cracking		
4.Natural hazard	Ground movement	4. Natural force damage	4. Ground movement
	Other		
5.Third party activity	Accidental	5. Other outside force damage	5. External interference
	Malicious	6. Excavation damage	
	Incidental	7. All other causes	6. Other /unknown

The proportion of causes for liquid and gas pipeline failures for various periods are given in Figure 1. a, b, c.

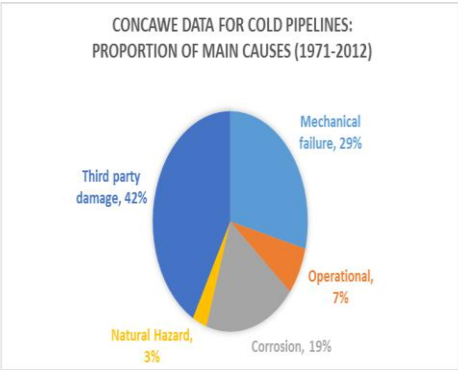

Figure 1a: Proportion of causes-CONCAWE [1]

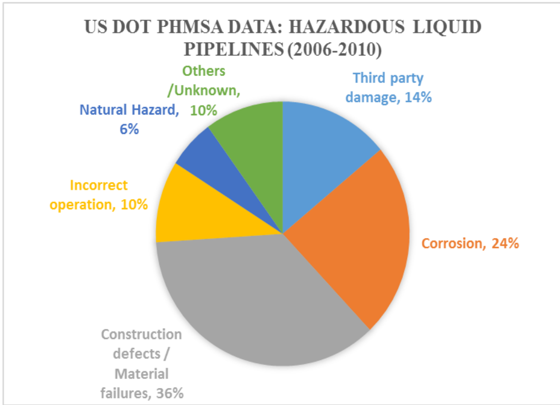

Figure 1b: Proportion of causes-US DoT [3]

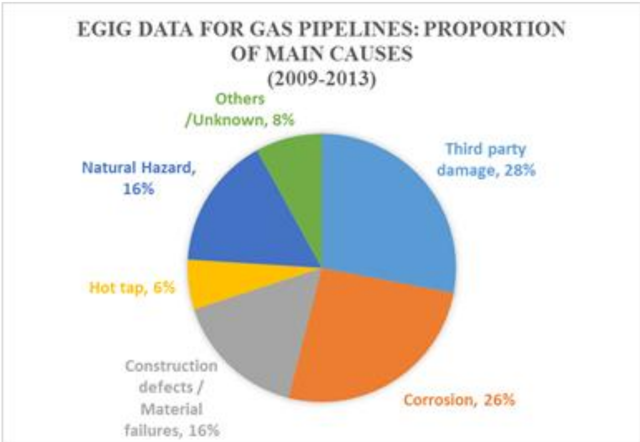

Figure 1c: Proportion of causes-EGIG [2]

It can be seen that all the databases have identified more or less the same number of causes for pipeline failures. The percentage contribution of the causes vary to some extent. The range of individual contribution is shown below. Pipeline carrying hot liquids are not considered.

Table 2: Range of values for causes of pipeline damage

Sl. No.	Main cause	Range
1	Mechanical failure	16% to 36 %
2	Operational	7% to 10% (no data for EGIG)
3	Corrosion	19% to 26%
4	Natural Hazard	3% to 16%
5	Third party damage	14% to 42 %
6	Others	8% to 10% (no data for CONCAWE)

However none of databases contain description or details of the protective measures installed or practiced in order to ascertain a measure of mitigation methods used. Reports [1] & [3] illustrate that the pipeline failures have come down over the years and the reasons attributed are mainly due to the mitigation measures taken. However maintaining the standard and quality of such measures is a challenge in the face of increasing length of pipelines and complex nature of the system. Thus pipelines failures have to be always viewed with respect to the protective measures employed. [6] and [7] describe the general protective measures employed in pipeline industries. [7] describes a system to access the risk of pipeline based on a comprehensive scoring system. [8] contain models using fault trees for gas pipeline failures and specific models for failures due to third party activities considering mitigation measures.

3. Bayesian Network Basics

A detailed description of Bayesian Network (BN) is not envisaged in the paper. Interested readers can go through any of the several books on the subject [9], [10], [11]. Briefly BN is a directed acyclic graph (DAG) in which the nodes represent the system variables and the arcs symbolize the dependencies or the cause–effect relationships among the variables. A BN is defined by a set of nodes and a set of directed arcs. Probabilities are associated with each state of the node. The probability is defined, a priori for a root (parent) node and computed in the BN by inference for the others (child nodes). Each child node has an associated probability table called Conditional Probability Table (CPT).

The computation of the net is based on the Bayes Theorem which states that if P (B) is probability of B happening, then P (A/B) is probability of A happening given that B has happened. P (B) not equal to zero.

Following gives the most common form of Bayes equation

$$P (A|B) = (P (B|A) * P (A)) / (P (B)) \quad 1$$

$$\text{Where } P (B) = P (B|A) * P (A) + P (B| A^{\wedge\prime})*P (A') \quad 2$$

A’ stands for A not happening

In equation 1, right hand side represents the prior situation –which when computed gives the left hand side –called posterior values. The value P (A) is the prior probability and P (B | A) is the likelihood function –which is data specific to the situation. P (B) is the normalizing function or the unconditional probability of B, which is calculated from equation 2 (Total law of probability).

The cause and effect relationship as conditional probability is shown in Figures 2a and 2b for 1 cause and 1 effect and 4 causes and 1 effect.

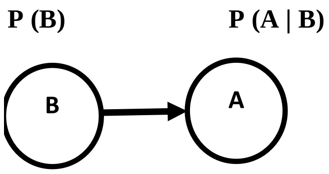

Figure 2a: BN for 1 cause & 1 effect

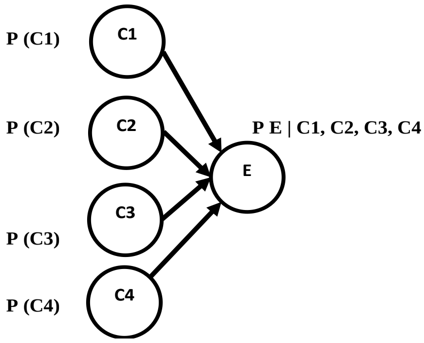

Figure 2b: BN for 4 causes and 1 effect

Essentially Bayesian Network (BN) is an efficient way of encoding all information as joint probability distribution about a set of random variables defined by the user. From the BN user can compute any value in the joint probability distribution. It can capture both quantitative and qualitative information between variables. One of the most important application of BN is in analyzing cause to effect (predictive reasoning) and effects to cause (diagnostic) models. While computing full joint probabilities is time consuming, in actual applications there are conditionally independent variables (for example in Figure 2b, C1 is conditionally independent of C2 to C4) and only these need to be computed. The graphical representation of cause and effect immediately conveys information about which variables are dependent and which are not.

The above features of BN can be used to model the impact of mitigation measures on pipeline failure. Once the causal model is built from cause (parent) to effect (child), it is populated with initial probability values (prior) for the parents and values or equations for the Conditional Probability Tables (CPTs) at each child node.

4. The Casual Model of Mitigation Measures and Pipeline Failures

Table 4 below gives the main and sub-causes of failures and mitigation measures usually used in the industry to prevent the causes from escalating to failures. The causes and sub-causes are primarily based on CONCAWE database. Some mitigation measures are listed in [6] and [7]. [12] discusses certain interesting aspects as to why Australian pipeline failures are low. Rest are from industry practices and authors' experience.

Table 3: Mitigation measures for prevention of pipeline failures

Sl. No.	Main cause	Sub-cause	Mitigation Measure
1	Mechanical failure	Construction fault	Procedures and implementation
			Supervision
		Materials fault	Design factors

			Design, review and inspection procedures
2	Operational	System	SCADA
			Overpressure protection
			Safety systems (HIPPS)
			Hazard identification
			Risk assessments
			Composition monitoring
		Human error	Training
			Operations & Maintenance manual and its review
			Up-to-date drawings
			Positive safety culture
3	Corrosion	External	Intelligent pigging
			Cathodic protection
			External coating
		Internal	Intelligent pigging
			Internal lining
			Corrosion inhibitor
		Sulphide Stress cracking	Closed interval survey / Intelligent pigging
4	Natural hazard	Ground movement / subsidence	
		Flooding	
		Other	
5	Third party	Accidental Malicious Incidental	Increase in wall thickness
			Pipeline safety zones
			Depth of cover minimum 1 M
			Warning marker posts
			Plastic marker tapes
			Concreate slabbing
			Physical barrier
			Vibration detection
			Right of Way patrolling
			Video cam monitoring
			Site survey before construction
6	Failure due to other causes		

5. Bayesian Network (BN) and simulation

Once these mitigation measures are identified they are used to construct a Bayesian Network in terms of probability of the mitigation measures not being in place and its effect on the main causes for pipeline failures. The impact probabilities of mitigation are based on authors and expert's experience and general industry observations since no data except [6] and [7] are available to access the impact. Further to a great extent, these are site specific. But Bayesian methods offer a practical way of updating these generic probabilities with site specific probabilities and the particular company expert's assessment.

The node states for all Parent nodes and Child nodes are all binary with 'Yes' or 'No' (that is with 0 & 100 % value) except that for 'Procedures and implementation'. Since there are several parents and child nodes, it is best to do the computations using a software. Several software / commercial programs are available now [13] [14] to help the user focus on the domain knowledge since the probabilistic computational part is fully taken care of by the software.

The Conditional Probability Tables (CPT) for the some nodes are given below in Figure 3a, b & c below:

Table 3a: CPT for node MechFailure

MechFailure:			
Yes	No	ConstFailure	DefectiveDesignOrMat
1	0	Yes	Yes
0.9	0.1	Yes	No
0.9	0.1	No	Yes
0	1	No	No

Table 3b: CPT for node OperationalFailure

OperationalFailure:			
Yes	No	SystemMalfunction	HumanError
0.99	0.01	Yes	Yes
0.8	0.2	Yes	No
0.8	0.2	No	Yes
0.1	0.9	No	No

Table 3c: CPT for Failure Due To Natural Hazards

FailureDueToNatuaralHazards:				
Yes	No	Subsidence	Flooding	Other
0.6	0.4	Yes	Yes	Yes
0.5	0.5	Yes	Yes	No
0.5	0.5	Yes	No	Yes
0.4	0.6	Yes	No	No
0.4	0.6	No	Yes	Yes
0.5	0.5	No	Yes	No
0.4	0.6	No	No	Yes
0.01	0.99	No	No	No

In the case of node 'Failure Due To Third Party activity' the number of parents are 11. As the number of parents goes up there is a corresponding increase in the number of entries required in the CPT, which will be tedious and time consuming if done manually. For 11 parents with binary states, the number of entries in the CPT will be 2 ^{11} * 11 = 22528 entries. In such cases the use Noisy-(logical) equation is suggested which is described below in the following section. Portion of the node CPT 'Failure Due To Third Party activity' is shown in Figure 4 to indicate the large number of entries required.

Table 4: Portion of the large CPT entries for node 'Failure Due To Third Party activity'

FailureDueToThirdPartyactivity:								
Yes	No	IncreaseInWallTkNotAdeq	PipellineSafetyZonesNotIdent	DepthOfCover1MNotProvid	WarningMarkerPostsNotAvail	PlasticMarkerTapeNotInstall	ConcreateSlabbingNotProvided
0.877455	0.122545	Yes	Yes	Yes	Yes	Yes	Yes
0.846819	0.153181	Yes	Yes	Yes	Yes	Yes	Yes
0.871005	0.128995	Yes	Yes	Yes	Yes	Yes	Yes
0.838757	0.161243	Yes	Yes	Yes	Yes	Yes	Yes
0.863839	0.136161	Yes	Yes	Yes	Yes	Yes	Yes
0.829799	0.170201	Yes	Yes	Yes	Yes	Yes	Yes
0.856673	0.143327	Yes	Yes	Yes	Yes	Yes	Yes
0.820841	0.179159	Yes	Yes	Yes	Yes	Yes	Yes
0.863839	0.136161	Yes	Yes	Yes	Yes	Yes	Yes
0.829799	0.170201	Yes	Yes	Yes	Yes	Yes	Yes
0.856673	0.143327	Yes	Yes	Yes	Yes	Yes	Yes
0.820841	0.179159	Yes	Yes	Yes	Yes	Yes	Yes
0.84871	0.15129	Yes	Yes	Yes	Yes	Yes	Yes
0.810888	0.189112	Yes	Yes	Yes	Yes	Yes	Yes
0.840748	0.159252	Yes	Yes	Yes	Yes	Yes	Yes
0.800934	0.199066	Yes	Yes	Yes	Yes	Yes	Yes
0.846819	0.153181	Yes	Yes	Yes	Yes	Yes	Yes
0.808524	0.191476	Yes	Yes	Yes	Yes	Yes	Yes
0.838757	0.161243	Yes	Yes	Yes	Yes	Yes	Yes
0.798446	0.201554	Yes	Yes	Yes	Yes	Yes	Yes

5.1 Noisy-Or Distribution

The software used for this paper (Netica) has a facility to use Noisy-Or or Noisy-And distributions and this can be used to provide values to the CPTs more realistically. Sometimes when the system is complex, the noise in the system cannot be adequately modelled. As an example, any amount of protective measures may not be able to prevent an incident. Noisy-(logical) distribution essentially can model such noise. It can also include probability factors for impact of one cause acting independently to produce the effect (independence of causal interaction).

Noisy-Or distribution can be used when there are several possible causes for an event, any of which can cause the event by itself, but only with a certain probability. Also, the event can occur spontaneously (without any of the known causes being true), which can be modelled with probability 'leak'. (This can be zero if it cannot occur spontaneously).

$$\text{Noisy Or Dist (e, leak, b1, p1, ... bn, pn)} \quad 3$$

Where e is the effect node, 'leak' factor is the probability of the effect even when all causes are zero, b1 is the node name for the cause, p1 is the probability of that cause impacting the effect node. The above can be written as in equation 4 for better understanding.

$$\text{(Effect | Cause1, Cause2) =NoisyOrDist (Effect, 0.1,Cause1, 0.2, Cause2, 0.4)} \quad 4$$

Using Noisy-Or, the equation to fill the probability values in the CPT for 'FailureDueToThirdParty activity ' is written as below.

5.1.1 *Failure due to Third party activity*

P (Failure Due To Third Party activity | Increase In Wall Tk Not Adeq, Pipelline Safety Zones Not Ident, Depth Of Cover1MNotProvid, Warning Marker Posts Not Avail, Plastic Marker Tape Not Install, Concreate Slabbing Not Provided, Physical Barrier Not Provided, Vibration Detection Not Avail, ROW Patrolling Not Done, Video Cam Monitoring Not Avail, Site Survey Before Constr Not Done) =

Noisy Or Dist (Failure Due To Third Party activity, 0.04, Increase In Wall Tk Not Adeq, 0.20, Pipelline Safety Zones Not Ident, 0.10, Depth Of Cover1M Not Provid, 0.20, Warning Marker Posts Not Avail, 0.25, Plastic Marker Tape Not Install, 0.25, Concreate Slabbing Not Provided, 0.20, Physical Barrier Not Provided, 0.20, Vibration Detection Not Avail, 0.10, ROW Patrolling Not Done, 0.10, Video Cam Monitoring Not Avail, 0.05, Site Survey Before Constr Not Done, 0.20) 5

In the above equation the first value on the RHS after the 'Failure Due To Third Party activity' is the leak factor that determines the probability of the event, even if all causes are not true. This has been set to 0.04. Rest of the values after each cause (parent) node name represents the probability of that particular cause (parent) impacting the effect (child) node

5.1.2 *For System Malfunction*

P (System Malfunction | SCADA Not Available, Over Pr Protection Not Avail, Safety Systems HIPPS Not Avail, Hazard Identification Not Done, Risk Assessment Not Done, Composition Monitoring Not Done) =

Noisy Or Dist (System Malfunction, 0.0, SCADA Not Available, 0.05, Over Pr Protection Not Avail, 0.20, Safety Systems HIPPS Not Avail, 0.10, Hazard Identification Not Done, 0.10, Risk Assessment Not Done, 0.10, Composition Monitoring Not Done, 0.15) 6

CPT for nodes 'Failure Due To Corrosion' and 'Pipeline Failure' have been have similarly filled using Noisy-Or equation with assigned probabilities for each of the parent node.

5.1.3 *For Failure due to Corrosion*

P (Failure Due To Corrosion | Internal Corrosion, External Corrosion, SCC, Intelligent Pigging Not Avail) =

Noisy Or Dist (Failure Due To Corrosion, 0.0, Internal Corrosion, 0.50, External Corrosion, 0.50, SCC, 0.30, Intelligent Pigging Not Avail, 0.60) 7

5.1.4 *For Pipeline failure*

P (Pipe Line Failure | Failure Due To Corrosion, Failure Due To Natuaral Hazards, Failure Due To Third Party activity, Operational Failure, Mech Failure, Failure Due To Other Causes) =

Noisy Or Dist (Pipe Line Failure, 0.0, Failure Due To Corrosion, 0.22, Failure Due To Natuaral Hazards, 0.05, Failure Due To Third Party activity, 0.30, Operational

Failure, 0.07, Mech Failure, 0.22, Failure Due To Other Causes, 0.14)
8

The values in the above are customizable depending on site specific conditions.

Netica Help file at Norsys [14] provides further details of the syntax for the above distribution.

5.2 Simulation the Bayesian Network (BN)

The nodes for mitigation measures (parent nodes) have been given probabilities 0 or 100 based on 'yes' or 'no' and rest of the values in CPTs are automatically computed by the model. For the given probabilities the probability of pipeline failure is calculated as 0.4010. This is the predictive mode of BN. See Figure 3 for the full BN of the mitigation measures and pipeline failure. Bars and values in each node depicts the states and probability values in percentage for that node.

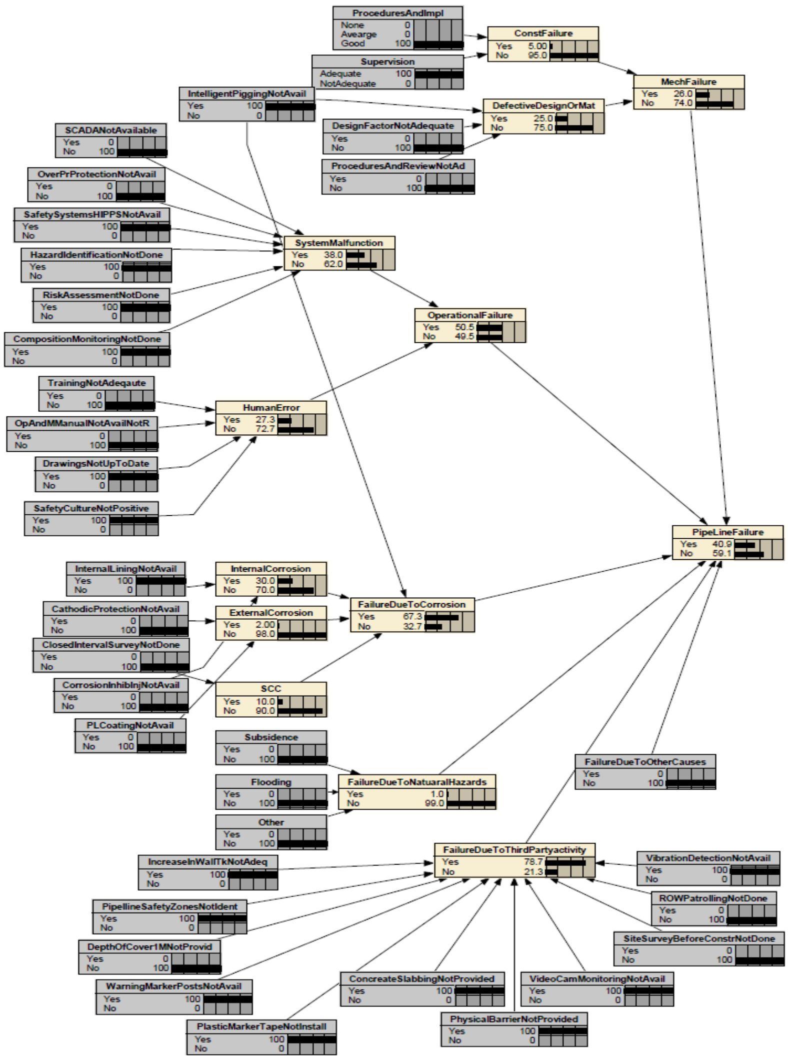

Figure 4: BN for pipeline failure. Predictive mode

Parameters for the parents can be fine-tuned to the site specific situation and the pipeline failures will be computed accordingly.

The user can change any of the parameters and see the impact on the pipeline failure probability which illustrates the flexibility of the BN. Further given a fine tuned BN model for a pipeline system, if there is a pipeline incident, the failure probability can be set to 100% and the BN will calculate the probabilities backward. This feature of diagnostics mode is very useful to find out the most probable cause during incident investigations. See Figure 4. The BN is predicting that the most probable causes as failure due to Third party activity (0.882) followed by failure due to Corrosion (0.756).

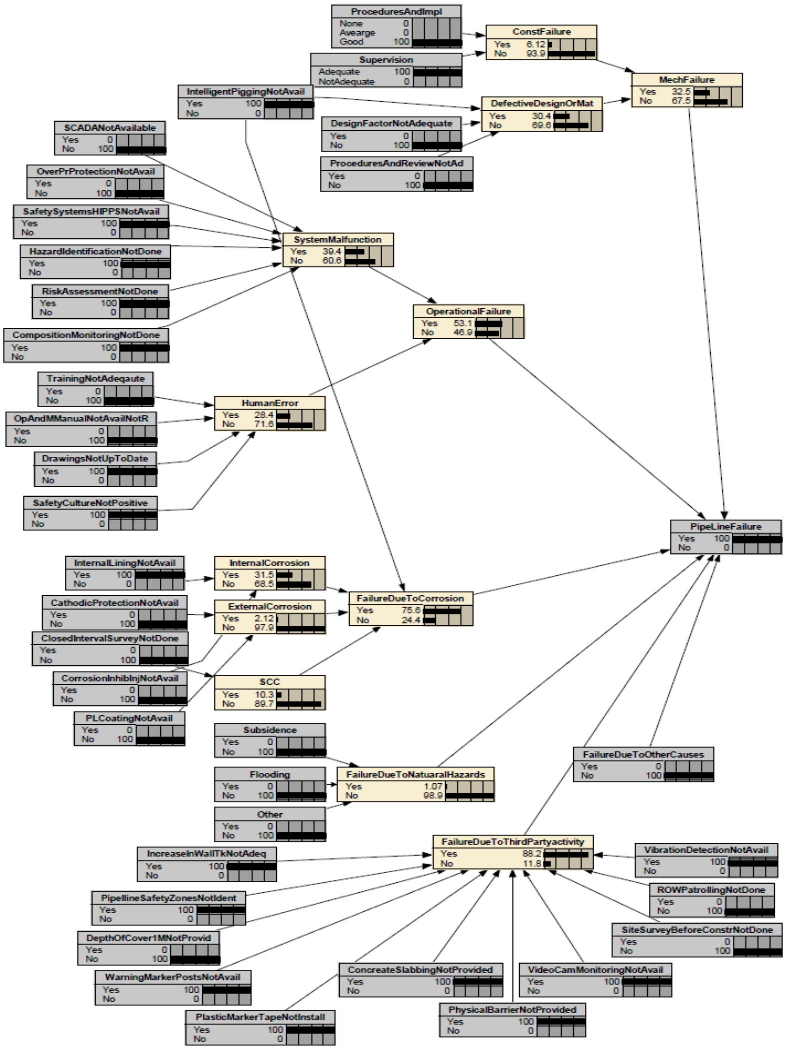

Figure 4: BN for pipeline failure. Diagnostic mode

This is logical since in the initial predictive model itself there is an increased chance for failure due to Third party activities since only 3 of the 11 mitigation measures are in place. Also, the lack of Intelligent pigging facilities have skewed the probability of failure to Corrosion.

In the diagnostic mode, the backward calculations will not change the parent nodes for which deterministic values have given as input.

6. Conclusions and Future Work

The power and flexibility of Bayesian Network is clear from the above example. The model can be easily customized to a site specific situation and if required made more mathematically complex by using advanced statistical techniques. Apart from the immediate visual impact of the causes and effects that is understandable to the all concerned, including operating staff, the BN can give clear warning about possibility of incidents. A site specific model can be put to very productive use by the industry. The pipeline failure in 27 June 2014 at Nararam village in India (22 fatalities) is an example. This gas pipeline did not have the protective measures of Corrosion inhibitor even though Intelligent pigging had revealed corrosion spots. Also there was change in composition in the gas that flowed through the line. Though full details of the protection measures are not publically available, BN would have focused on which of the mitigation measures were inadequate and would have predicted high probability of failure.

BN is a complement to conventional Quantitative Risk Assessment (QRA). QRA requires considerable amount of time whereas BN can be developed quite fast. Updating of QRA is also time consuming. Currently work is ongoing for developing models that are more sophisticated with better predictive capability that can be easily used by the industry.

References

[1] Davis, P.M., Diaz, J-M., Gambardella, F., and Uhlig, F., Performance of European cross country oil pipelines, Report No 12/13, 2013, Concawe, Brussels. www.concawe.org

[2] European Gas Pipeline Incident Data Group., 9th Report of the European Gas Pipeline Incident Data Group (period 1970-2013), 2014, EGIG, Netherlands. www.EGIG.eu

[3] Stover, R., “Review of the US Department of Transportation Report: The State of National Pipeline Infrastructure”, 2013, Accessed from Internet, www.icog.com/~oildrop/PHMSA_report_analysis.pdfitate

[4] Alberta Energy Regulator., Report 2013-B: Pipeline Performance in Alberta, 1990-1012,https://www.aer.ca/about-aer/spotlight-on/pipeline-safety-review

[5] BC Oil & Gas Commission., 2011 Pipeline Performance and Activity Report, 2011, www.bcogc.ca

[6] Pettitt, G., and Morgan, B., “ A tool to estimate the failure rates of cross country pipelines”, Symposium Series, Hazards XX1 Process Safety and

Environmental Protection in a Changing World, 2009, Symposium Series No. 155, IChemE
[7] Muhlbauer, Kent., Pipeline Risk Management Manual, Third Edition: Ideas, Techniques, and Resources, 2004, Gulf Publishing. Massachusetts, USA.
[8] Healthand Safety Executive., An assessment of measures in use for gas pipelines to mitigate against damage caused by third party activity, Contract Research Report 372/2001, 2001, Accessed from Internet, http://www.hse.gov.uk/research/crr_pdf/2001/crr01372.pdf
[9] Pourret, O., Naim, P., Marcot, B., Eds., Bayesian Networks: A Practical Guide to Application, 2008, Wiley, West Sussex.
[10] Kjærulf, U.B., and Anders, M. L., Probabilistic Networks -An Introduction to Bayesian Networks and Influence Diagrams, 2008, Springer, New York.
[11] Neapolitan, R. E., Learning Bayesian Networks, 2003, Prentice Hall, New Jersey.
[12] Tuft, P., 2014, "Comparing International Pipeline Failure Rates", The Australian Pipeliner., 124-134, April, www.pipeliner.com.au
[13] HUGIN Expert., www. hugin.com
[14] Norsys Software Corporation., 2015, Netica, www. norysys.com.