Infosys®
POWERED BY INTELLECT
DRIVEN BY VALUES

Improve your odds with Infosys Predictability

Right Questions | Business Insights | Integrated Solutions | Value Multipliers

White Paper

Operational Risk and Probabilistic Networks – An Application to Corporate Actions Processing

– Shilpa Ramamurthy, Harpreet Arora, Anirbid Ghosh

Introduction

With the Basel deadlines fast approaching, most affected banks are scrambling to get their operational risk management act together. Among many factors that are impeding the development of stable operational risk management infrastructures, unavailability of loss data is probably the most common. Without a credible internal loss history database, most of the advanced risk analysis and measurement techniques (e.g. Loss Distribution Approach) cannot be implemented. The Basel defined requirements for loss history is fairly stringent – to use any one of the actuarial approaches, organizations would need to have at least three years (preferably five) of operational loss history. Most banks and financial institutions are not likely to have this historical data, especially for some business lines – risk category combinations.

November 2005

Bayesian Belief Networks (or BBN's as they are commonly known) provide an elegant solution to this problem. They combine both qualitative and quantitative information for arriving at loss estimates. They are particularly appropriate for modeling operational risks with little or no historical losses - most low frequency-high severity operational losses fall in this category. Moreover, BBN's are causal networks, unlike other approaches like simulations, and are particularly useful for analyzing causes that contribute to operational risk.

This paper provides a reference implementation of a BBN with an application to Corporate Actions Processing. It tries to model operational losses from "missed corporate action announcements", a fairly prevalent problem in the custody business. Though the implementation is for a specific risk category, the technique can be easily extended to any other category.

Risk Category	Frequency	Severity
Internal fraud	Low	High
External fraud	Low	High
Employment practices & workplace safety	Low	Low
Clients, products, & business practices	Low/medium	High/medium
Damage to physical assets	Low	Low/medium
Business disruption & system failures	Low/medium	Low/medium
Execution, delivery, & process management	High	Low

Table 1:
Frequency & severity of Basel defined operational risk categories

Bayesian Networks – A Brief Overview

A BBN, also known as qualitative or causal probabilistic network, is a technique that helps model, measure and manage operational risk using prior knowledge of the causal risk factors and probabilistic reasoning. It is represented in the form of an acyclic graph consisting of nodes and directed arcs. Each node represents a variable that impacts or determines operational risk. The arcs denote causal or influential relationships between variables. Each variable in the network is assigned an underlying probability distribution based on subjective prior beliefs. Bayesian analysis involves improving the prior estimates in the light of additional information about one or more variables in the network.

Application of Bayesian Networks to Corporate Actions Processing

The Business Problem

Corporate action processing is one of the most risky and labor-intensive back-office processes in the securities industry. A corporate action is defined as any action taken by the issuer of a security that affects the structure or financial status of the security. An end-to-end corporate action processing involves data capture, event certification, entitlement determination, event notification, voluntary response tracking, settlement, and reconciliation.

There are more than 150 types of corporate action events that can be classified into three broad categories – mandatory, mandatory with options and voluntary. About a million corporate actions take place every year worldwide over and above the three million fixed-rate interest payments and redemptions. A single event may involve many different market participants including custodians, investment managers, broker-dealers and depositories. Each of these parties face high operational risks because corporate action processing is, to a large extent, non-standardized and manual. Failure in handling a single corporate action has the potential to result in huge losses running into millions of dollars. The global fund management industry alone incurs actual losses of approximately US$ 400million – US$ 900million every year due to corporate action processing failures. Indirect losses, arising from incorrect interpretation of corporate action information, can be significantly larger - such losses cost the securities industry between US$2 billion and US$ 9.6 billion annually. *Source of estimates: DTCC and Oxera*

Our article constructs a BBN to model operational loss arising from missed corporate action announcements (sent by an issuer/vendor). Announcements are typically sent in multiple formats by data vendors and collected by financial organizations for further processing. Since it is largely a manual & non-standardized process, there are chances of missing an announcement, which can result in missed notifications and subsequent settlements, thereby leading to an opportunity/cash loss.

Modeling the Network

The Variables

The first step in the Bayesian process is construction of the operational loss model by identifying the key variables and their cause-effect relationship. The key BBN variables for a missed corporate action (CA) announcement are given below. These variables, in combination, determine operational loss from a missed announcement.

Data Sources – A number of sources including registrar, custodian/sub-custodian and data vendors provide corporate action announcement data in multiple formats using multiple delivery methods. Further, there is no standard way in which the events are announced by issuers, there is no single securities identification system that is universally accepted and the processing terms and details are often specific to the particular market or financial instrument. Data sources may, therefore, be considered as good or bad.

CA Volumes – A large number of corporate actions are announced every year on both equity and debt instruments. Volumes tend to surge during the corporate earnings season, thereby straining the efficiency of the corporate action staff and increasing the risk of operational loss. The volume of corporate action announcements to be processed may be low or high.

CA Type Complexity – In general, voluntary corporate actions and mandatory actions with options are considered to be more complex than mandatory actions as they are dead-line driven and require processing investor responses. CA type complexity may be in either of the two states – low or high.

CA Processing System – The amount of automation in corporate action processing varies amongst organizations. Even the more successful organizations have not managed to automate the entire lifecycle. A mosaic of heterogeneous platforms also adds to the chaos. The CA Processing System may be assumed to be either good or bad.

Staff Efficiency – Corporate actions processing continues to be labor-intensive and requires a high degree of manual intervention to resolve exceptions. The risk of manual error is accentuated by the complexity and volume of corporate actions to be processed. Staff efficiency may be considered to be in either of the two states - low or high.

The Network

The network below shows the inter-linkages between all the BBN variables. The nodes for data sources, CA volumes, CA processing system, CA type complexity and staff efficiency represent the causal risk factors or parameters.

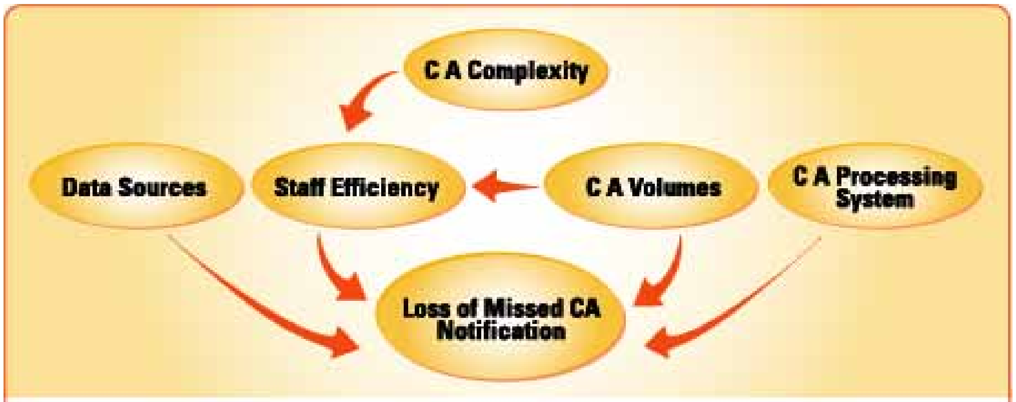

Figure 1:
BBN for a Missed CA Announcement

The node for “Loss due to missed CA announcement” shows the evidence of operational loss. The magnitude of operational loss due to a missed corporate action announcement is impacted by the data source quality, CA system maturity, staff efficiency and corporate action volumes.

Assigning Probability Distributions to BBN Variables

The next step in the Bayesian process involves assigning probability distributions to each of the variables in the network. The probability distributions of the variables are provided on the basis of prior knowledge about the behavior of parameters before operational loss data is observed (For the sake of simplicity, discrete distributions are considered; however, the technique for assigning continuous distributions is very similar). In practice, this would involve gathering inputs from the operations staff.

Data Sources	Probability
Good	70
Bad	30

CA Complexity	Probability
Low	80
High	20

CA Volumes	Probability
Low	30
High	70

CA Processing System	Probability
Good	70
Bad	30

CA Complexity	Low		High	
CA Volumes	Low	High	Low	High
Efficiency - Low	25	40	40	55
High	75	60	60	45

Volumes	Low							
Processing	Good				Bad			
Staff Efficiency	Low		High		Low		High	
Data Sources	Good	Bad	Good	Bad	Good	Bad	Good	Bad
Loss $ 0 - 1 million	50	20	60	30	30	10	50	20
1 - 2 million	30	50	30	40	40	50	40	50
2 - 3 million	20	30	10	30	30	40	10	30

Volumes	Low							
Processing	Good				Bad			
Staff Efficiency	Low		High		Low		High	
Data Sources	Good	Bad	Good	Bad	Good	Bad	Good	Bad
Loss $ 0 - 1 million	40	10	50	20	20	10	30	10
1 - 2 million	40	60	40	50	50	40	50	50
2 - 3 million	20	30	10	30	30	50	20	40

Bayesian Inference

Based on the probability distribution of the causal variables, likelihood estimates for the operational loss is calculated. The results are provided below. It should be noted that the probability distribution of "staff efficiency" is recalculated since it is dependent on two other variables – CA complexity and CA volumes.

Staff Efficiency	Probability
Low	38.5
High	61.5

Operational Loss US$	Probability
0 - 1million	34.32
1 - 2 million	43.49
2 - 3 million	22.19

Using the above loss estimates, an organization can estimate the expected loss arising from missing a CA announcement. The expected loss is –

$$(.5*.3432) + (1.5*.4349) + (2.5*.2219) = .1716 + .6524 + .5548 = 1.3787 \text{ million}$$

Analysis

There are two common techniques of analyzing a Bayesian Network – scenario analysis and causal analysis.

Scenario Analysis

Scenario Analysis involves calibrating one or more causal risk factors in the network and analyzing its impact on the loss estimate. For example, an operations manager might be interested in estimating operational losses under heavy processing volumes (all other conditions remaining unchanged). In such a situation, the estimated operational loss is given below.

CA Volumes	Probability
Low	0
High	100

Operational Loss US$	Probability
0 - 1million	30.42
1 - 2 million	46.06
2 - 3 million	23.53

$$(.5^*.3042) + (1.5^*.4606) + (2.5^*.2353) = .1521 + .6909 + .5884 = 1.4314 \text{ million}$$

It can be seen that, according to the model, processing volumes do not have a significant impact on operational losses.

Causal Analysis

Conceptually, causal analysis is the exact opposite of scenario analysis. Under causal analysis, new evidence of operational losses is used to calculate updated probabilities (also referred to as posterior probabilities) of all the causal factors. In other words, additional loss information is propagated to all the nodes in the network. This technique of evidence propagation is extremely useful for analyzing the causes that impact operational losses.

An example is provided below to clarify this concept. If an operations manager is most concerned with large losses (2-3 million) and wants to mitigate this risk, understanding the causes that typically contribute to such losses is important to design a better control infrastructure. From the results below, it is evident that the quality of data sources has the maximum impact on

Operational Loss US$	Probability
0 - 1million	34.32
1 - 2 million	43.49
2 - 3 million	22.19

Data Sources	Probability
Good	53.58
Bad	46.42

CA Complexity	Probability
Low	79.02
High	20.98

CA Volumes	Probability
Low	25.79
High	74.21

Staff Efficiency	Probability
Low	48.47
High	51.53

CA System	Probability
Good	50.54
Bad	49.46

operational losses. Consequently, an operations manager, on the basis of these results, might strive to improve the quality of source data – probably by subscribing to more automated feeds.

Refining the Network

A BBN may be extended to include decision nodes and utility nodes (the enhanced network is referred to as an Influence Diagram), as shown below, in order to facilitate the management of operational risk. A decision node represents a variable controlled by a risk manager in order to manage operational risk. In the above example, the risk manager may decide to train his staff in order to improve their efficiency. Since the decisions are controlled by the risk manager, they do no have conditional probability tables.

A utility node represents the expected utility from the decision. For instance, the utility node, cost, gives information about the cost associated with training while the utility node, payoff, represents the payoff from increased staff efficiency.

An example will clarify this concept. Training might entail an initial outlay of US$10,000 and would result in an expected payoff of US$30,000. The influence diagram would highlight the decision which maximizes the expected utility. In other words, the risk manager would decide to spend on training if the expected utility from training exceeds the utility from no training for a given value of operational loss.

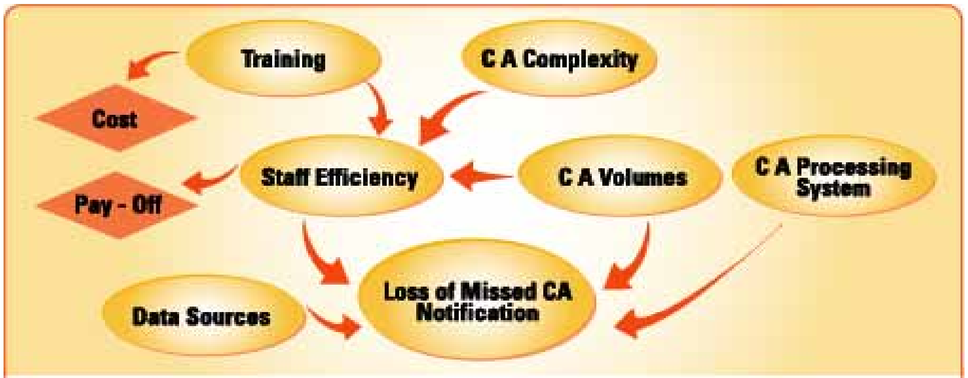

Figure 2:
Influence Diagram for a Missed CA Announcement

Conclusion

As has been highlighted throughout this paper, BBN's provide an effective technique for modeling operational losses. Though they can be used to model almost all operational risk types, they are more appropriate for situations where loss data availability is low. Unlike many statistical techniques, BBN's are investigative in nature – they try to analyze the causes rather than focus solely on the effects. In that respect, it is a forward-looking technique and does not depend entirely on historical losses – this feature makes BBN's particularly effective when past is not the best predictor of the future.

On the downside, BBN's are somewhat subjective in nature. They are a modeler's view of reality and hence there can be multiple models representing the same operational loss type. Moreover, since business landscapes are dynamic in nature, BBN's involve some amount of maintenance – they need to be regularly updated to incorporate changes in the business.

Acknowledgements

All examples in this paper have been developed using a BBN modeling tool (Hugin Lite) from Hugin Expert A/S. We would like to extend our thanks to Hugin for making this tool freely available for educational purposes.

References

1. F.V.Jensen. An Introduction to Bayesian Networks, Springer-Verlag, 1996
2. Marcelo G. Cruz. Modeling, Measuring and Hedging Operational Risk, Wiley Finance, 2002
3. Nidel Da Costa Lewis. Operational Risk with Excel and VBA – Applied Statistical Methods for Risk Management, Wiley Finance, 2004
4. Carol Alexander. Bayesian Methods for Measuring Operational Risk – Discussion Paper in Finance, 2000-2002
5. Microsoft Research – Carl M.Kadie, David Hovel, Eric Horvitz. A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks, 2001
6. Yuji Yasuda. Application of Bayesian Inference to Operational Risk Management, Doctoral Theses, 2003
7. Norman Fenton. Incorporating Expert Judgement in Operational Risk Quantification, Critical Systems Conference, 2002

About the Authors:

Shilpa Ramamurthy works with the Banking and Capital Markets Consulting & Solutions Group at Infosys to design and develop solutions in the areas of process automation, risk management and compliance, data integration and quality for clients in the financial services industry.

Harpreet Arora works as a Principal with Infosys' Banking & Capital Markets practice. Harpreet works with financial services clients from both a domain and business consulting standpoint. Specific areas of expertise include business process redesign and domain focus areas like investments, accounting, trading, settlements, global custody & corporate actions. Harpreet has prior experience with Citigroup in the capital markets space.

Anirbid Ghosh specializes in delivering technology enabled solutions for banks and securities firms. His key focus areas include trading platforms, analytics, risk management, compliance & data management.