

# Strategic Management of Intellectual Property: R&D Investment Appraisal Using Dynamic Bayesian Network

L . O. Nwobodo

Enugu State University of Science and  
Technology  
Enugu State, Nigeria

H.C. Inyama

Nnamdi Azikiwe University[NAU]  
Awka, Anambra State, Nigeria

## Abstract :

Company executives are under increasing pressure to proactively evaluate the benefits of the huge amounts of investment into intellectual property (IP). The main goal of this paper is to propose a Dynamic Bayesian Network as a tool for modeling the forecast of the distribution of Research and Development (R&D) investment efficiency towards the strategic management of IP. Dynamic Bayesian Network provides a framework for handling the uncertainties and impression in the qualitative and quantitative data that impact the effectiveness and efficiency of investments on R&D. This paper specifies the process of creating the graphical representation using impactful variables, specifying numerical link between the variables and drawing inference from the network.

**KEY TERMS:** IP Dynamic Bayesian Network, R&D Investment Efficiency.

## 1 INTRODUCTION

Intellectual property (IP), which is now being regarded [1][2] as constituting the core assets of companies, is centered around innovation. The very basis for innovation is developing new products and services, which logically means that more IP will be generated and thus increases the need for protection [3].

The research and development (R&D) Investment of major manufactures has reached an annual level of several billion dollars [4]. Questions are being raised as to whether this investment amount is efficient from the perspective of its effectiveness, such as whether such investment is made efficiently and whether the results of R&D are appropriately contributing to company profits.

Essentially, activities for R&D and intellectual properties must be dealt with in the management cycle shown in figure 1. It is important to realize that R&D strategy and the IP strategy exactly represent the business strategy of a company. There are many

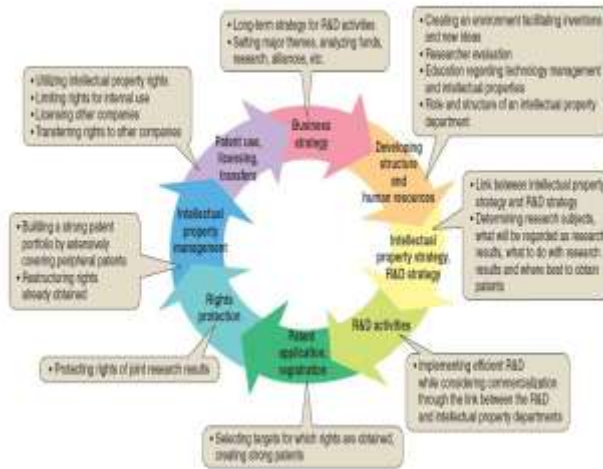
companies that spend amounts for R&D that exceed their amount of capital investment. In response to these moves, institutional investors have started to show an interest in the content of R&D investment and the effectiveness and efficiency of such investment. However, due to uncertainties and imprecise data, it is extremely difficult to proactively demonstrate explicitly the efficiency and effectiveness of R&D investment.

In an automated Bayesian Network decision support system, the probability judgment analysis make over time can be captured, compared to actual data when it becomes available, and then provide feedback on a timely basis [5]. Based on the work by Alpart and Raiffa [6], this feedback should improve analysts' probability assessment which should lead to improved future performance.

[7] observed: “whether or not they recognize it, virtually all decisions that investors make are exercises in probability: For them to succeed, it is critical that their probability statement combines the

historical record with the most recent data variable. And that is Bayesian analysis in action”.

Hence this paper proposes a Bayesian Network model for the forest of the distribution of R&D investment efficiency.



**Figure 1:** Ideal management cycle for activities for R&D and intellectual properties [4]

## 2. BAYESIAN NETWORKS

Bayesian networks provide a formalism for reasoning about partial beliefs under conditions of uncertainty. These parameters are combined and manipulated according to the rules of probability theory [8][9]. Let us consider  $n$  discrete random variables  $x_1, x_2, \dots, x_n$ , a discrete acyclic graph with  $n$  nodes, and suppose the  $j^{\text{th}}$  node of the graph is associated to the  $x_j$  variable. Then the graph is a Bayesian network, representing the variables  $x_1, x_2, \dots, x_n$ , if

$$P(x_1, x_2, \dots, x_n) = \prod_i p(x_i / \text{parents}(x_i)),$$

Where  $\text{parents}(x_j)$  denotes the set of all variables  $x_i$  such that there is an arc from node  $x_i$  to  $x_j$  in the graph. The probability terms in the product are

described by conditional probability Tables (CPT) which may be set by hand or learned from data. Standard algorithms such as junction tree [9] [10] exist to perform inference networks.

Dynamic Bayesian Networks (DBNs) allow the modeling of entities in a changing environment where the values of variables change over time [10] [11]. Functionally, DBNs capture the process of variable values changing over time by representing multiple copies of network nodes with one copy for each time step [10]. Visually, they may be displayed using two copies of each recurrent node representing the current  $t$ , and previous,  $t-1$  states. A Bayesian network is a tool to help expert represent uncertainty, ambiguous or incomplete knowledge. Bayesian networks use probability theory to represent uncertain knowledge. A Bayesian network consists of two parts—a qualitative graphical structure of the relationships in the model and a quantitative structure represented by the probability distributions that are indicated by the graph. In a Bayesian network, historical information can provide a framework or baseline model to develop prior distribution. New quantitative information, qualitative information, or evidence can be added to the network as appropriate to develop posterior probabilities.

Decision makers in many different contexts combine qualitative data and qualitative information. Bayesian networks have been applied in a wide variety of decision-making contexts. Some examples are venture capital financing [12], auditing [13], medical diagnosis [14], and software design [15], among many.

### 3 DEVELOPING THE BAYESIAN NETWORK

There are three important steps in building the Bayesian network. The first step is developing the graphical model. This step includes the relevant variables and specifying whether they are independent, or not. The second step is the specification of the numerical relationship between the variables that are not independent. The third step is making inferences or decisions based on new evidence.

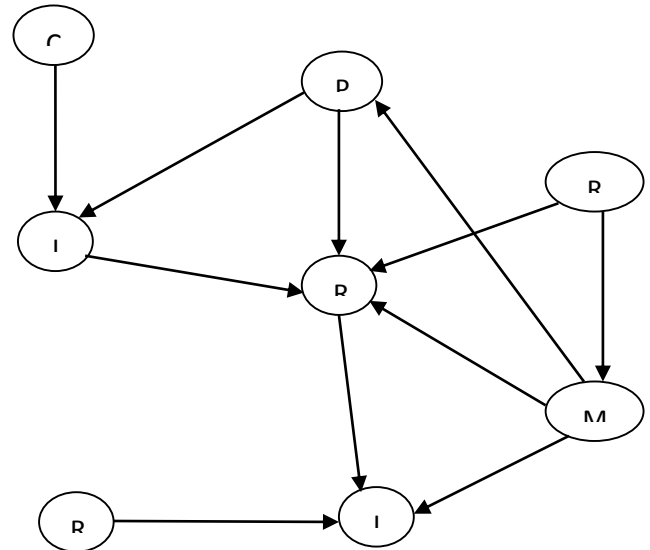
#### 3.1 Graphical Representation

As mentioned in the previous section, the first step in construction the Bayesian network is the graphical model. The graphical model is a directed, acyclic graph where nodes represent variables and directed arcs (arrows) represent the conditional probability relationship assumed in the model. The variables and the arcs between the variables are the inputs to the graph.

In this paper, as part of R&D strategies, the forecast of R&D investment efficiency is to be made using the variables: infrastructure, capability of science and technology, Risk, R&D efficiency management performance, total R&D personal capability, total R&D expenditure. These are variables that are assumed by management (for the sake of analysis) that impacts Return on Investment (ROI) on IP R&D.

The relationship among the IP variables may be represented by the Bayesian network of figure 2. The network consists of four discrete variables: Capability of Science and Technology (CST), Risk (RSK), R&D Efficiency (RDE), Management

Performance (MP), Total R&D Personnel Capability (PC), Total R&D Expenditure (RDEP), and Investment Efficiency (IE).



**Figure 2:** Bayesian Network for forecasting R&D Investment efficiency.

In the network, a node with arcs leading out only indicates a marginal probability distribution. For example CST, RDEXP, RSK are marginal probability distribution. A node with arcs leading into it indicate a conditional relationship. For example, the factors (variables) that determine the value for infrastructure (IF) is the capability of science and technology (CST) and total R&D Personnel Capability (PC). The node infrastructure (IF) is a conditional node.

The absence of a directed arc from a node is also meaningfully because the absence indicates independence assumption. The absence of a directed arc denotes conditional independence between nodes. Thus the lack of an arc from CST to RSK signifies that it is assumed that RSK is independent of CST;

the lack of an arc from IF to MP signifies that it is assumed that MP is independent of IF.

The analyst benefits in several ways from the graphical construction of a Bayesian network. First, the construction of the graphical portion of the network helps the analyst clarify and refine his view of the relationships among the variables [5] Next, the analyst may not always have a good understanding of how a decision is reached. They fully understand which variables are used, but how the variables are combined and heuristics to come up with a decision is not always well understood or systematic [5]. In [16], almost all analyst agree that qualitative information was important, but when questioned about how it was incorporated in a decision, most analyst could not be specific.

### 3.2 Determining Numerical links between variables

Each node in the Bayesian network is a variable that is described either as a constant value, a probability distribution, or as a function of other variables [17]. In a Bayesian network, the primary focus is on determining the probability distribution of the relevant nodes. A Bayesian network model is represented at two levels, qualitative and quantitative. At the qualitative level a directed graph is used (as done here in figure 2) in which nodes represent variables and directed arcs describe the conditional independence relations embedded in the model. At the level, conditional probability distributions are specified for each variable in the network. Each variable has a set of possible values called its state space that consists of mutually exclusive and exhaustive values of the variable.

There are two primary ways to find probability distribution, for the nodes in the network. One way is historical data. The other is to use subjective probability judgments. The two methods can also be combined.

For the model in this paper capability science and technology (CST) is specified as having two states: “High” and “Low”; infrastructure (IF) has two states: “Good” and “Bad”, total R&D Personnel Capability (PC) has two states: “High” and “Low”, total R&D Expenditure (RDEXP) has two states: “High” and “Low”; Management Performance (MP) has two states: “Good” and “Bad”; R&D efficiency (RDE) has two states: “High” and “Low”; Risk (RSK) has two states: “High” and “Low” and investment efficiency (IE) has two states: “High” and “Low”.

A fundamental assumption of a Bayesian network is that when the conditionals for each variable is multiplied, the joint probability distribution for all variables in the network is obtained.

Suppose a sequence of the variables in a Bayesian network is picked such that for all directed arcs in the network, the variable at the tail of each arc precedes the variable at the head of the arc in the sequence. Since the directed graph is acyclic, there always exist one such sequence. In figure 1, one such sequence is CST IF RDE MP RSK IE.

From figure 2, this shows that the model makes the assumption:

$$P(\text{CST, IF, RDE, PC, RDEXP, MP, RSK, IE}) = P(\text{CST}) \otimes P(\text{IF/CST, PC}) \otimes P(\text{RDE/IF, PC, RDEXP, MP}) \otimes P(\text{PC/RDE}) \otimes P(\text{RDEXP}) \otimes P(\text{MP}) \otimes P(\text{RSK}) \otimes P(\text{IE})$$

$$P(\text{MP/RDEXP}) \otimes P(\text{RSK}) \otimes$$

$$P(\text{IE/RDE,MP,RSK}),$$

Where  $\otimes$  denotes point wise multiplication of tables.

For each variable, a table of conditional probability distribution has to be specified, one for each configuration of states of its parents.

For the model, tables 1(a) – (h) gives the tables of conditional distributions- $P(\text{CST})$ ,  $P(\text{IF/CST,PC})$ ,  $P(\text{RDE/IF, PC,RDEXP,MP})$ ,  $P(\text{PC/MP})$ ,  $P(\text{RDEXP})$ ,  $P(\text{MP/RDEXP})$ ,  $P(\text{RSK})$ ,  $P(\text{IE/RDE,MP,RSK})$ .

(a)

P(CST)	High Low
	0.25 0.75

(b)

P(IF/CST,)PC	Good Bad
High, High	0.80 0.2
High, Low	0.06 0.40
Low, High	0.25 0.75
Low, Low	0.50 0.50
High, High	0.01 0.90

(c)

P(RDE/IF, PC,RDEXP,MP)	High Low
High, High, High, Good	0.85 0.35
High, High, High, Bad	0.60 0.40
High, High, Low, Good	0.80 0.20
High, High, Low, Bad	0.55 0.45
High, Low, High, Good	0.67 0.33
High, Low, High, Bad	0.70 0.30
High, Low, Low, Good	0.68 0.32
High, Low, Low, Bad	0.66 0.34
Low, High, High, Good	0.50 0.50
Low, High, High, Bad	0.49 0.51
Low, High, Low, Good	0.40 0.60
Low, High, Low, Bad	0.75 0.25
Low, Low, High, Good	0.30 0.70
Low, Low, High, Bad	0.20 0.80
Low, Low, Low, Good	0.10 0.90
Low, Low, Low, Bad	0.25 0.75

(d)

P(PC/MP)	High Low
Good	0.15 0.85
Bad	0.70 0.30

(e)

P(RDEXP)	High Low
	0.65 0.35

(f)

P(MP/RDEXP)	Good	Bad
High	0.10	0.90
Low	0.70	0.30

(g)

P(RSK)	High	Low
	0.70	0.30

**Table 1:** Tables of conditional probabilities for the Bayesian network of figure 2.

P(IE/RDE, MP, RSK)	High	Low
High, Good High,	0.90	0.10
High, Good, Low	0.85	0.15
High, Bad, High	0.02	0.98
High, Bad, Low	0.65	0.35
Low, Good, High	0.70	0.30
Low, Good Low	0.55	0.45
Low, Bad, High	0.5	0.5
Low, Bad, Low	0.1	0.9

### 3.3 Making inference

The ultimate goal is to model the probability distribution of the investment efficiency (IE) for IP portfolio.

Once a Bayesian network is constructed, it can be used to make inferences about the variables in the model. The conditionals given in Bayesian network representation specify the prior joint distribution of

the variables. If the values of some are observed (or learnt), then such observations can be represented by tables where 1 is assigned for observed values and 0 for unobserved values. Then the product of all tables (conditionals and observations) gives the posterior joint distribution of the variables. Thus the joint distribution of variables changes each time new information is learnt about item.

Often the interest is on some target variables. In this case, inference is made by computing the marginal of the posterior joint distributions for the variables of interest. Consider the situation described by the Bayesian network in figure 1. The interest is in the true state of the R&D investment efficiency. Given the prior model (as per the probability tables given): table 1(a) – (b)), the marginal distribution is computed (giving probability values for “High” and “Low”). Now suppose it is learnt that (i.e new observation) risk (RSK) is “Low” and management performance (MP) is “Good”. The posterior marginal distribution of R&D investment efficiency changes.

### 4 CONCLUSION

The main goal of this paper is to propose a Bayesian network as a tool for modeling the forecast of the distribution of R&D investment efficiency towards the strategic management of IP.

The improvement of investment efficiency by reviewing R&D and the reorganization of a business portfolio have become urgent issue to manage. Many companies are realizing that if they are going to spend any money on IP, it better be IP that has value to the business. So companies are now developing strategic plans as to (a) where they want inventions (b) what

resultant IP they want to create and (c) the efficiency of their R&D investment. Company executives now face a variety of opportunities that require sophisticated analysis and decisions. These requirements are better met by advanced decision support tools.

With the capability to model the forecast of the efficiency of R&D investment given different scenarios of the combination of resources (R&D personnel, management performance, R&D budget), risk and market demand, companies are able to make effective decisions regarding their IP portfolio. This enables companies to evaluate the business benefit of any IP before its creation. This means for example they can strengthen the few patents they will file by focusing on “inventing around their own IP” before filing it. An advanced analytic tool, such as the Bayesian network, helps to work the process of inventing around” to systematic and robust.

There are many off-the-shelf software systems that allow Bayesian networks to be constructed graphically by end-users, for example Bayesialab ([www.bayesi.com](http://www.bayesi.com)), Netica ([www.norsys.com](http://www.norsys.com)) and Hugin([www.hugin.com](http://www.hugin.com)). These tools allow the user to enter the graph and specify numerical relationships among the variables. The software in use calculates the inference based on these inputs. The inference results are shown graphically as probability distribution for the network. The analysis would help executives make better strategic decision regarding their IP portfolio.

## 5. REFERENCES

[1] Dino Isa, Pter Blanchfield, Zh.yuan 2009“Intellectual Property Management System

for the super-capacitor Pilot plant “workshop on advances on advances intelligent computing.

- [2] Benintend, S. 2003 “Intellectual property valuation one important step in a successful asset management system “Payton: University of Dayton School of Law, (2003) PP 12, 14; 16-20
- [3] John Cronin 2010, “The cause for Developing IP an “Executable” IP strategy in 2010” IP Capital Group, Inc, PP 1.
- [4] Masayuki Miyake, Yuji Mune, and Keiichi Himeno, Dec 2004 “Strategic Intellectual Property Management: Technology Appraisal by using the “Technology Heetmap”, NRI papers No. 83.
- [5] Riza Demiser, Roreld R. Mom, Catherine Shenoy, June, 2005, “Bayesian Networks. A Decision Tool to improve portfolio Risk Analysis”, PP 9-11 .
- [6] Alpert, M., and H. Raiffa, 1982 “A Progress Report on the Training of Probability Assessment. Judgment under uncertainty: Heuristics and Biases. ed. D. Kahneman, P. Slovic, A. Tversky. New York: Cambridge University Press.
- [7] Hagstrom, R.G. 1999 “The Warren Buffett Portfolio. Wiley and Sons, New York.
- [8] Pearl .J., 1988 “probability reasoning in Intelligent systems: Networks of plausible Inference. Morgan Kaufmann Publishers, 2<sup>nd</sup> edition on.
- [9] Mark Taylor, Charles fox “Inventory Management with Dynamic Bayesian Network Software systems” Adaptive Behavior Research Corp, Sheffield, UK.
- [10] Arbib M.A, editor 2003 “The handbook of brain theory and neural networks, chapter Bayesian networks. Press.
- [11] K.P. Murphy 2002 “Dynamic Bayesian networks: Representation, Inference and learning “Technical report, University of California, Berkeley.
- [12] Kemmeser, B.S. Mishra, and PP. Shenoy 2001 “Bayesian Casual Maps as Decision Aids in venture capital decision making: Methods and Applications” University of kensas, working Paper.
- [13] Gillete, P and R.P. Srivestava 2000. “Attribute Sampling: A Belief Function Approach to Statically Audit Evidence”. Auditing: A journal of practice and theory vol.19, no 1. (Spring) 145 – 155.
- [14] Bielza, C., S. Rios-Insua, and M. Goniez 1999 “Influence Diagrams for Neorata Jaundice

management” In lecturer notes in Artificial Intelligence 1620, (ed. Werror Horn, Yural Shahar, G. Lindberge, Steen Andresen, J. Wyatt) springing-Verley Berlin Heride/berg 1999, 138-142.

- [15] Horvitz, E., J. Breese, D. Heckerman, D. Hovel, and K. Rommelse 1998. “The Lumiere Project: Bayesian user modeling for Inferring the Goals and Needs of software Users” proceeding of the fourteenth conference on uncertainty in Artificial intelligence, July.
- [16] Catherine Slonoy, Pracash P. Sloney 1998. “Bayesian network models of portfolio risc and return”, School of business working paper No.279.
- [17] Slonoy 2001. “ A Description of Security Analysis” Decision making process.” University of Kansas, working paper.