27

Handling Imprecise Information in Risk Management

Love Ekenberg and Mats Danielson

The DECIDE Research Group
Department of Computer and Systems Sciences, Royal Institute of Technology and
Stockholm University, Electrum 230, S-164 40 KISTA, SWEDEN

Abstract

Cost evaluation often constitutes a substantial part of the total risk analysis. Often, models that use decision theoretical methods at different levels in the risk evaluation process are unable to take into account situations where the available information concerning the consequences of different incidents is vague or numerically imprecise. Based on a more general theory for decision analysis, a method for cost evaluation is suggested. It includes well-founded and computable procedures that enable a risk manager to work with interval statements and comparisons. The method is easy to implement in a computer system and does not require the use of numerically overprecise statements of probability and cost. The evaluation results in an interval that expresses the maximum and minimum expected cost with respect to the estimations of the risk manager. The interval can be further investigated with respect to the range of values consistent with the estimations. The method extends a risk evaluation process currently in use in Telia AB (formerly Swedish Telecom).

1. INTRODUCTION

In a society of ever-increasing complexity, the needs for more sophisticated risk management tools are stronger than ever. Risk analysis lies at the heart of every enterprise, be it a government or an organisation, profit or non-profit making. At all levels, they have to face and deal with risk because uncertainty is built into every facet of their operations. Risk management is carried out all the time, more or less consciously, using more or less advanced models of risk. Because of this, risk analysis has been *"considered an essential part of an effective approach to information systems security. Properly used, it raises the management awareness of security exposures, provides a practical mechanism for understanding the magnitude of these exposures, and assists in the evaluation and selection of appropriate safeguards."* [ESF 1991, p.1].

To acquire a satisfactory understanding of the risk situation, management usually desires some kind of structured approach to the analysis. Thus, a risk analyst, conducting a risk analysis, has often access to standard procedures for identifying and assessing threats and for identifying and valuating assets. The basic steps of risk management could be considered to be something like the following:

J. H. P. Eloff et al. (eds.), *Information Security — the Next Decade*

1. Identify the assets.
2. Determine the vulnerabilities.
3. Estimate the likelihood of exploitation.
4. Compute the annualized loss expectancy (or alternatively the expected cost of a particular incident).
5. Survey current and possible safeguards.
6. Determine the savings according to the safeguards.[1]

When performing a risk analysis according to these steps, the following problems are important to a risk manager when the assets have been identified:

A. Given an incident, which are the relevant sets of consequences to consider?
B. How should the costs of the different consequences be evaluated?
C. Which are the relevant sets of safeguards to consider?
D. How should the effects of the possible safeguards be evaluated?
E. How should the choice of safeguards be optimised in the light of previous analyses?

In this paper, we have chosen to focus on item B, and we propose a method suitable for different stages of the risk management process. The method extends our earlier results in [Ekenberg 1994c], that presents a working procedure constituting an integral part of the risk management process of Telia AB (formerly Swedish Telecom).[2]

To make it easier to grasp the ideas behind the method; to compare it with traditional approaches; and to indicate some disadvantages of these; a brief survey of some current approaches to risk management is included. Ensuing this, an informal overview of the method we propose is given, followed by a description of its theoretical background.

2. TWO CURRENT APPROACHES TO RISK ANALYSIS

Different decision theoretical methods are often used in risk management. They are typically used in several steps to identify and evaluate assets, such as properties and information, and to identify and evaluate threats, such as fire, burglary, and industrial espionage. Such analyses are also used to verify the current protection, and to evaluate the effects of modifying it. The following two sections focus on two common techniques used in such evaluations.

When analysing the consequences of an incident, we are not always interested only in monetary costs. Thus, we will use the concept of cost in a general sense, meaning both quantitative and qualitative values. Utilities could have been used instead, but, in this context, we feel that cost is a more natural concept than utility. Note that monetary cost is a special case of cost.

---

[1] For a more detailed account, see, e.g., [Pfleger 1989].

[2] Techniques for handling problems with D can be treated in a similar way. To our knowledge, no attempts at a formal treatment of A or C exist, and it is not obvious what such a treatment would look like. Problems related to E can be treated as decision problems by a method similar to what we will describe here. (Procedures for the treatment of such decision problems have been treated by us in, e.g., [Ekenberg 1994b]).

Let us now introduce some terminology. We will first consider *simple incidents* (resulting only in direct consequences), and later in the paper extend the concept to *incidents* (resulting in both direct consequences and new incidents).

Def: A set of consequences $\{c^i_1,\ldots,c^i_t\}$ is a *simple incident* (denoted H_i).

Def: The *expected cost of a simple incident* H_i is expressed by the formula $E_i = p^i_1e^i_1 +\ldots+ p^i_te^i_t$, where e^i_j denotes the cost of the consequence c^i_j, and p^i_j denotes the probability of the consequence c^i_j occuring given that H_i occurs.

2.1. Point Scales and Expected Cost

Often, when evaluating the expected cost of a simple incident, the probability and cost variables above are instantiated with numerically precise data. A main problem is that in real-life analysis it is often impossible for any analyst to explain the difference between closely proximate probabilities, for example, 72% and 75%. This problem is also emphasized by the inability to express varying reliabilities for different kinds of data. Which data are based on long experience, and which are mere guesses? In models using numerically precise information, this kind of expressibility is severely limited.[1]

One attempt to overcome the unrealistic and time-wasting assumption of numerically precise information is to be more imprecise, even in making the estimations. J.F. Broder writes: *"it is neither necessary nor desirable to make precise statements of impact and probability. The time needed for the analysis will be considerably reduced and its usefulness will not be decreased, if impact (i) and frequency (f) correlations are given in factors of 10."* [Broder 1984, p.22]. Then he proposes the following scale:[2]

Loss valuation of an incident:		**Estimated frequency of occurrence:**	
\$10	i = 1	Once in 300 years	f = 1
\$100	i = 2	Once in 30 years	f = 2
\$1,000	i = 3	Once in 3 years	f = 3
\$10,000	i = 4	Once in 100 days	f = 4
\$100,000	i = 5	Once in 10 days	f = 5
\$1,000,000	i = 6	Once per day	f = 6
\$10,000,000	i = 7	10 times per day	f = 7
\$100,000,000	i = 8	100 times per day	f = 8

The annualized loss expectancy is then approximated by $\frac{10^{(f+i-3)}}{3}$.

A problem with this approach is that the possible values are spaced too far apart. Needless to say, this can be solved by using decimal numbers for *i* and *f*, but then we are back where we began. Furthermore, an important feature of a method that allows imprecise data is that it should enable detection of critical variables and the study of what effects modifications to the given data will have. This is not least important when the possible values are spaced far apart. Also, a risk manager is still unable to express varying degrees of reliability for the different kinds of data.

---

[1]Methods for estimating the monetary cost of a simple incident by using numerically precise data in an expected cost model can be found in, e.g., [Dixon 1990, p.86] ff.

[2]The method was originally suggested in [Courtney 1977], and is recommended to prospective government suppliers by NIST.

2.2. Risk Level Models

One way to partially overcome the problems with point scale models is to allow the analyst to express the different values in non-monetary terms. In Sweden, for instance, a relative three-level model has been used for example by [Hamilton 1988], [SAF 1986], [Wermdalen 1991]. The probabilities and utilities involved (somewhat misleadingly called consequences in this approach) are expressed as shown in Figure 1. Variants of the three-level model are also frequently used. For example, [Statskontoret 1989–91] uses a four-level model, and so does the Swedish SBA method [Wrede 1984]. Not infrequently, even more rudimentary models are proposed.[1]

	Probability	Value	Risk level
1	**Low/Small** Seldom occurs	**Small** Low cost, little damage or loss	**Acceptable** Can be allowed Should be remedied
2	**Medium** Occurs neither often nor seldom	**Medium** Greater cost Greater damage or loss	**Unacceptable** Not allowed Must be remedied
3	**Great/High** Often occuring	**Great** Cost cannot be borne Total loss	**Catastrophic** Must be remedied immediately Unforgivable

Figure 1 From [Hamilton 1988, p.76].

The *risk level* is a function of the sum PV = *probability* + *value*. If PV $\in$ {2}, the risk level is 1, if PV $\in$ {3, 4}, the risk level is 2, and if PV $\in$ {5, 6}, the risk level is 3. A major problem with this approach is that the intervals are too wide, with no discrimination within them. Therefore, most risks evaluate to risk level 2, with no indication of how to order the risks within that level. An ordinary risk analyst is capable of differentiating between disastrous, unacceptable, and acceptable risks without the aid of decision tools. The problem is to decide the order and the extent of the reduction needs of different unacceptable risks. Hence, when the risk situation is obvious, there is little need for a model, and when it is not, the models offer little help.

Also, the choice of the formula above for evaluation seems peculiar, and it is obvious that what results from this differs from evaluations using the expected utility. We will not argue extensively in favour of the latter principle, but rather be content with the expected utility model being indisputably the most well-founded of the two, and the literature concerning the formal justifications of the principle is vast (see, e.g., [Fishburn 1981]). Naturally, the expected utility model is not the only reasonable candidate, since the principle in itself is not uncontroversial already on formal grounds [Malmnäs 1994b], but it is not clear that the risk level model could be a serious candidate when compared to the expected utility model.

---

[1]Many practitioners abandon the concept of probability altogether. For instance, insurance advisors often find it too hard to make estimates of the frequencies of accidents because of low levels of repetition, and they sometimes erroneously draw the conclusion that all kinds of probability based reasoning should be abandoned. In, e.g., [Green 1992] a five-level model without probabilities is suggested. In [ESF 1991] probabilities are also ignored.

3. A METHOD FOR COST ANALYSIS

We will now discuss the method with which we are proposing to solve the problem raised above: How should a risk manager evaluate the cost of different incidents? The method we propose extends the traditional approaches to risk analysis and to some extent resolves the problems mentioned in the treatment of the traditional approaches above. Section 3.1 presents some theories for handling imprecise statements. Section 3.2 contains a treatment of simple incidents, and sections 3.3 and 3.4 generalise the method.

3.1. Theories for Handling Imprecise Statements

As indicated above, there are several difficulties with providing adequate estimations of probabilities and costs in risk evaluations. Thus, there is a need for relaxing the pointwise quantitative nature of probability and cost estimations. One proposal to extend the probability theory is to represent belief states defined by interval-valued probability functions by means of classes of probability measures [Smith 1961], [Good 1962]. In [Choquet 1953/54] the concept of capacities was introduced, and later these ideas were studied in connection with probability theory [Huber 1973a], [Huber 1973b]. Also different kinds of generalisations of first-order logic have been used for providing methods for how to deal with sentences with upper and lower probabilities [Nilsson 1986]. Another approach was suggested by Arthur Dempster, who investigated the properties of multivalued mappings from a space X, with a known probability distribution over subsets of X, to a set S, defining upper and lower probabilities for subsets of S [Dempster 1967]. Glenn Shafer further developed a non-Bayesian approach to quantifying subjective judgments [Shafer 1976], and the Dempster-Shafer formalism has subsequently become rather popular, especially within artificial intelligence.[1]

Thus, mathematically sound bases for the integration of probability intervals with classical probability theory have been developed, but the particular concern here is how problems modelled with numerically imprecise probabilities and costs can be *evaluated.* A general approach to this problem is investigated by [Levi 1974]. However, his theory has some counter-intuitive implications, and is also problematic in some respects when confronted with empirical results [Gärdenfors 1982]. Another suggestion to extend classical analysis is the application of fuzzy set theory, i.e., fuzzy theories have also relaxed the requirements for numerically precise data, and it provides a more realistic model of the vagueness in subjective estimations of utilities and probabilities. Early attempts in decision analysis are [Bellman 1970], [Freeling 1980], and in [Zimmermann 1984]. These approaches also allow a decision (or risk) manager to model the problem in vague linguistic terms, and membership functions can be defined according to the statements involved. Despite the success of these theories, our approach attempts to conform to traditional statistical reasoning, and in that way avoid problems emanating from difficulties with providing set membership functions and defining set operators with a satisfying intuitive correspondence.

---

[1] As has been pointed out by, e.g., [Weichelberger 1990], the Dempster-Shafer representation seems to be unnecessarily strong. Further, the possibility to state that, e.g., one consequence is worse than another is very useful particularly when handling qualitative costs. Therefore, in addition to using interval statements, we use inequality statements to express comparisons between probabilities and costs, a feature lacking for instance in the Dempster-Shafer approach.

3.2. The Proposed Method

We have earlier developed a theory for supporting real-life decision situations [Malmnäs 1994a], [Ekenberg 1994a], [Ekenberg 1994b]. The theory is developed to handle general decision problems involving different criteria, alternatives, and consequences, and it enables a decision maker to work with an imprecise basis for decisions and still reach a conclusive result. A decision maker is allowed to define a full mapping of natural language terms into numerical values, or express her uncertainty with intervals. By a procedure primarily based on the principle of maximising the expected utility, the alternatives are evaluated. The theory has been implemented in a Macintosh computer program that has been used in real-life applications [Malmnäs 1995].

As has been pointed out, when evaluating information from a consequence analysis, risk analysts often use a formula expressing the expected cost of an incident, and we will now demonstrate how our earlier theory can be extended to evaluate the expected cost when the probability and cost estimations are vague or numerically imprecise.

We choose to treat a set of simple incidents simultaneously, since much can be gained from studying several interrelated incidents at the same time. First we consider the representation of *probabilities*, formalise the interpretations of admissible statements, and describe this for four types of possible probability statements. The *cost statements* are translated in a similar way.

1. The consequence c^i_j is probable, c^i_j is possible, c^i_j is improbable, etc.
2. The probability of c^i_j equals the real number m, is greater than m, is less than m.
3. The probability of c^i_j lies between the real numbers m and n.
4. The probability of c^i_j is equal to the probability of c^r_s, is approximately equal to the probability of c^r_s, is more probable than c^r_s, etc.

We now suggest interpretations of the sentences in 1 to 4. The essential point is that statements as above are translated into a system of linear inequalities that makes them particularly easy to handle. Thus, for instance, the k_i's below are positive real numbers in the interval [0,1], used for expressing different qualitative statements. Suppose, for example, that a risk manager decides that in order for c^i_j to be probable, the probability for it to occur must be greater than 50%, but less than 90%. In this case, k_1 and k_2 in 1.1 below would be 0.5 and 0.9 respectively. Note that the upper bound (k_2) is not necessarily 100%, because it is not certain that a risk manager's attitude to the statement "*c^i_j is probable*" implies that she believes the probability of c^i_j could be 100%.

Naturally, such a translation is not uncontroversial. What rational grounds do we have for accepting a particular translation? One answer is that this is not necessarily critical as we proceed, since we can study the effects of different translations by means of sensitivity analyses during the process. If a risk manager still is averse to the use of qualitative statements, she can use only interval statements and weak inequalities instead.

1.1. $(p^i_j \geq k_1, k_2 \geq p^i_j)$
1.2. $(p^i_j \geq k_3, k_4 \geq p^i_j)$
1.3. $(p^i_j \geq k_5, k_6 \geq p^i_j)$
2.1. $p^i_j = m.$
2.2. $(p^i_j \geq m, k_7 \geq p^i_j)$
2.3. $(p^i_j \geq k_8, m \geq p^i_j)$
3.1. $(p^i_j \geq m, n \geq p^i_j)$
4.1. $p^i_j = p^r_s$
4.2. $p^i_j + k_9 \geq p^r_s$ and $p^r_s + k_{10} \geq p^i_j$
4.3. $p^i_j \geq p^r_s$

We call the conjunction of expressions of the four types above, together with the equalities $\Sigma\, p^i_j = 1$ for each simple incident involved, the *probability base* S(p). The corresponding list expressing statements of costs is called the *cost base* V(e). We will also refer to the expressions within parenthesis as *intervals*. The probability base and the cost base are linear systems and constitute the *information base*. Thus we can define our problem as a structure:

Def: An *information base* is a structure $(\{H_1,\ldots,H_h\}, S(p), V(e))$, where each H_j is a simple incident $\{c^j_1,\ldots,c^j_{h_j}\}$, and S(p) (V(e)) is a finite list of linear equalities, linear inequalities, and intervals in the probability (cost) variables.

Note that, at this stage of the presentation, it is assumed that the consequences are disjoint. This restriction will be relaxed later, but the incorporation of non-disjoint consequences will restrict the cost functions involved.

To simplify checks for consistency used in the definitions below, we adapt a theorem from [Malmnäs 1994a]. (In the remainder of this paper, we assume an information base $(\{H_1,\ldots,H_h\}$, S(p), V(e)) underlying the definitions and results.)

Theorem: Let $p^i_1e^i_1 +\ldots+ p^i_te^i_t > 0$ be a bilinear inequality in probability and cost variables in S(p) and V(e). Then the system $A(p,e) = \{p^i_1e^i_1 +\ldots+ p^i_te^i_t > 0\} \cup S(p) \cup V(e)$ can be reduced to a disjunction of linear systems $L = L_1 \vee\ldots\vee L_N$ in only the probability variables, with the following property: A(p,e) has a solution iff L has a solution.

Thus, if $p^i_1e^i_1 +\ldots+ p^i_te^i_t > 0$ is a bilinear inequality in probability variables in S(p) and cost variables in V(e), then we can transform the system A(p,e) to linear systems. This simplifies further processing of the bases.

Now we can define what cost is, given the imprecise information in the information base. Actually, two kinds of cost need to be considered with respect to the information base: A greatest possible cost, and a least possible cost.

Def: If E_i denotes the expected cost of a simple incident H_i (i.e., $p^i_1e^i_1 +\ldots+ p^i_te^i_t$), and if k is a positive real number, then *I-cost*(E_i,k) = *true* iff the system $\{E_i - k < 0\} \cup S(p) \cup V(e)$ is consistent.

Def: If E_i denotes the expected cost of a simple incident H_i, and if k is a positive real number, then *S-cost*(E_i,k) = *true* iff the system $\{E_i - k > 0\} \cup S(p) \cup V(e)$ is consistent.

To obtain the adequate interval, calculate the lower bound of k, such that I-cost(E_i,k) = true, and the upper bound of k, such that S-cost(E_i,k) = true. This is readily done by using standard linear programming algorithms.

When confronting a real-life problem, the risk manager is encouraged to be deliberately imprecise. Values close to the boundaries of the intervals seem to be the least reliable ones. Another problem is that, in nontrivial evaluations, the cost intervals are too large to give any significant information by themselves. The question therefore is how a better estimation of the expected cost could be obtained, not only for extreme points in the probability base, but also for inner points of the different intervals. The most immediate solution to this is to study in how large parts of V(e) and S(p) an expected cost is consistent. One way to do this is by using Monte-Carlo methods [Malmnäs 1994a]. A problem with this approach is that it considers boundary points as equally essential as points close to the middle of the intervals.

As a complement to such a procedure, we suggest the use of the concept of *proportion*. The idea behind this concept is to investigate how much the different intervals can be decreased before the cost estimation becomes inconsistent. By this procedure we can study the stability of a result by gaining a better understanding of how important the boundary points are for the result. By integrating the concept of proportion with the procedures for handling expected cost above, we can create a process taking account of the amount of consistent instances of probability and cost variables where the different costs are consistent. (A suggestion for a procedure for calculating proportions is given in the appendix.)

3.3. Non-Disjoint Consequences

Further, the use of the method is not restricted to problems where the consequences of an incident are not disjoint. The following observation can be made (proven in [Ekenberg 1994a]): If the consequences of an incident are not disjoint, then every cost function involved is required to be a linear function.

Theorem: Let H be an incident with the consequences c_1 and c_2, and let v be a cost function. Also let H* be an incident with consequences c_1*, c_2*, and c_3*, where c_1* is (c_1 *and not* c_2), c_2* is the consequence (*not* c_1 *and* c_2), and c_3* is the consequence (c_1 *and* c_2). The expected cost of H (denoted E) is equal to the expected cost of the incident H* (denoted E*), identical to H but modelled with disjoint consequences iff $v(c_1 \text{ and } c_2) = v(c_1) + v(c_2)$.

Example: A certain division of a business company uses two personal computers (A and B). A burglary might result in either one or both personal computers being stolen. There are backup copies of all programs, but on one hard disk (in computer B) there is a file with 300 records that would have to be restored. The monetary cost of one computer is exactly \$5,000. The company has access to statistical information on the direct monetary cost of restoring the file, and according to this information, restoring it will have a monetary cost of about \$10,000. They also have the following information concerning the operations of the company:

- The cost of the monetary loss of \$5,000 is 1, the cost of the loss of \$15,000 is 4, and the cost of the loss of \$20,000 is 6.
- The probability that computer A will be stolen in the event of burglary is 0.9, the probability that computer B will be stolen is 0.8, and the probability that both computers will be stolen is 0.7.

This evaluates as follows: $E = 0.9 \cdot 1 + 0.8 \cdot 4 = 4.1$ and $E^* = 0.2 \cdot 1 + 0.1 \cdot 4 + 0.7 \cdot 6 = 4.8$. H and H* are exactly the same incident, but the calculations of the expected costs give different values. By the proposition above, the inconsistency is caused by the nonlinearity of the cost function. Hence, when using the method with consequences that are not disjoint, the cost function must be linear in order for calculations of the expected cost to remain consistent.[1]

3.4. Modelling Tree Problems

We will now modify the description of a simple incident resulting in a set of consequences. The new model allows an incident to generate both new incidents *and* consequences, which in

[1] It is also worth mentioning that when there is no requirement for the probabilities to sum up to 1, the procedures for solving bilinear inequalities can be simplified further, as demonstrated in [Ekenberg 1994a].

turn can generate even more incidents and consequences, and so on, see Figure 2. The H's in the figure denote incidents, and the C's different consequences. The P's denote the probabilities involved.

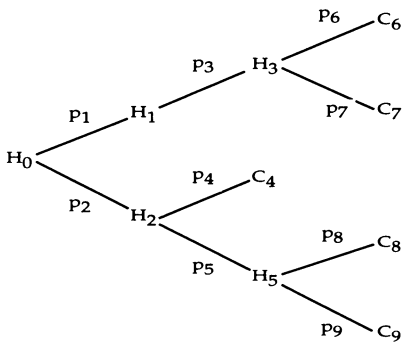

Figure 2 An extended consequence analysis

We extend the definition of expected cost. Note that in the following definitions, we use the fact that an incident is formally a set of consequences and incidents.

Def: A set of consequences $\{c^i_1,\ldots,c^i_t\}$ is a *simple incident* (denoted H_i). A set of incidents and simple incidents $\{H_1,\ldots,H_r\}$ is an *incident*.

Def: The *expected cost of an incident* H_i $\{H_1,\ldots,H_r\}$ is expressed by the formula $E_i = p^i_1E_1 + \ldots + p^i_rE_r$, where E_j denotes the expected cost of the incident H_j (or simple incident H_j), and p^i_j denotes the probability of the incident H_j (or simple incident H_j), given H_i.

Example: Consider Figure 2. The incident H_5 can result in c_8 and c_9, and only these. Hence, H_5 is a simple incident, and the expected cost of it is equal to $p_8E_8 + p_9E_9$. The incident H_2 generates a new incident H_5 and can also result in c_4. The *expected cost of the incident* H_2 is therefore equal to $p_4E_4 + p_5E_5$. E_4 is the cost of the simple incident consisting of the single consequence c_4, and E_5 (= $p_8E_8 + p_9E_9$) is the expected cost of the simple incident H_5.[1]

The discussion in the preceding section was based on a one-level description, i.e., an incident did not generate new incidents. This does not cause any real restriction, because a tree problem (where an incident generates new incidents) can always be transformed into a one-level problem. To understand this transformation, consider the case modelled in Figure 2. H_2 generates H_5 with probability p_5, and generates c_4 with probability p_4. H_5 generates c_8 with probability p_8, and generates c_9 with probability p_9. The supremum of the expected cost interval of H_2 is calculated by first calculating the supremum of H_5, and then calculating the supremum of H_2, with respect to this result. The infimum of H_2 is calculated in a similar way. By repeated application, a one-level problem is received. It can then be evaluated by the method proposed in section 3.2.

[1]For clarity, the indices have been simplified in the example.

4. CONCLUSIONS

In this paper, it is argued that many models for risk analysis are too weak to carry out a thorough cost evaluation of threats. A general method is proposed to overcome some of the problems connected with evaluating real world problems without idealising conditions. The approach models a typical incident in a structured way, and it is important to note that the probabilities and costs can be entered into the model via informal and numerically imprecise sentences. These are then translated into a suitable representation. The evaluation method used is based on calculating the expected cost of an incident. Very often this cost, in nontrivial situations, varies within a large interval. Thus, the concept of proportion is useful. The proportion shows how much the different intervals can decrease in the probability base for different cost intervals and yet remain valid.

APPENDIX: ALGORITHMS FOR CALCULATING PROPORTIONS

To simplify the definitions below, we first define what is meant by a feasible (probability or cost) base. A feasible base contains the minimal representation of the solution set of the original base. Not all bases are feasible since they consist of translations of the risk manager's subjective statements.

Def: Given a base X, containing a sublist Z of all intervals $[(E \geq a_1, E \leq b_1),\ldots,(E \geq a_n, E \leq b_n)]$ in X, *Feasible(X)* is the list $(X - Z) \cup Y$, where Y is a list of intervals, constructed from the list of expressions in X, where each interval $(E \geq a_i, E \leq b_i)$ in X has been substituted by $(E \geq \inf(\{p : (p \geq a_i, p \leq b_i)$ is consistent with $X\})$, $E \leq \sup(\{p : (p \geq a_i, p \leq b_i)$ is consistent with $X\}))$.[1]

The concept of proportion is defined in three steps. First we define a procedure for decreasing the size of all the intervals in a base while maintaining consistency. The size of the steps could for instance be proportional to the size of the original interval.

Def: Let X be a system of linear equations, inequalities, and intervals, containing variables $\{x_1,\ldots,x_n\}$, and let $D = [d_i]$ be a list of *n* numbers. A *D-reduction of X to X** is achieved as follows:

Let $X_0 = X$.

For all $x_j \in \{x_1,\ldots,x_n\}$:

Let $X'_i = X_{i-1}$, and replace the interval $(x_i \geq a_i, x_i \leq b_i)$ in X'_i by $(x_i \geq a_i + d_i, x_i \leq b_i - d_i)$, where $d_i \in D$. Call this X_i and check if it is consistent. If not, let $X_i = X_{i-1}$.

Let $X^* = X_n$.

We now iterate this procedure as far as possible. (The number δ in the procedure below should be suitable for the given situation.)

---

[1]The supremum and the infimum used in the definition are easily calculated by using standard linear programming algorithms [Ekenberg 1994a].

Def: Let X be a system of linear equations and intervals. Let δ be a number in the interval [0, 0.5[. Let $[(x_i \geq a_i, x_i \leq b_i)]$ be the sublist of intervals in Feasible(X), and let $D = [d_i]$ be a list where $d_i = \delta(b_i - a_i)$.

A *δ_0-reduction of X to Y_0* is to let Y_0 = Feasible(X).

Given a δ_{i-1}-reduction of X to Y_{i-1}. A *δ_i-reduction of X to Y_i* is the D-reduction of Y_{i-1} to Y_i, where $Y_{i-1} \neq Y_i$.

The concept of proportion can now be used to find out in how large parts of S(p) I-cost(E_i,k_1) and S-cost(E_i,k_2) are true.

Def: The *δ_t-^{i}proportion of $E_i - k_1 < 0$* is the number t = sup({s : there is a δ_t-reduction of $\{E_i - k_1 < 0\} \cup S(p) \cup V(e)\}$).

Def: The *δ_t-^{s}proportion of $E_i - k_2 > 0$* is the number t = sup({s : there is a δ_t-reduction of $\{E_i - k_2 > 0\} \cup S(p) \cup V(e)\}$).

Needless to say this is one of several possible approximations, but if we do not have any more information available and want an estimation with respect to the information base, then this seems to be a rational way of obtaining a deeper insight into the problem. By integrating the concept of proportion with the procedures for handling intervals of estimated cost above, a procedure that takes account of the ranges of consistent instances of probability and cost variables can be defined.

REFERENCES

Bellman, R., and Zadeh, L.A.: 1970, Decision Making in Fuzzy Environment", *Management Science* 17, pp.B-144-B-164.

Broder, J.F.: 1984, *Risk Analysis and the Security Survey*. Butterworth Publishers.

Choquet, G.: 1953/54, "Theory of Capacities", *Ann. Inst. Fourier* 5, pp.131-295.

Courtney, R.H.: 1977, "Security Risk Assessment in Electronic Data Processing", *AFIPS NCC 46.*

Dempster, A.P.: 1967, Upper and Lower Probabilities Induced by a Multivalued Mapping", *Annals of Mathematical Statistics*, XXXVIII, pp.325-339.

Dixon, G.: 1990, *Riskanalys*, SBF – Svenska Brandförsvarsföreningen.

ESF: 1991, A Risk Analysis Method which is Easy to Understand and Simple to Apply, Draft Method, European Security Forum.

Ekenberg, L.: 1994a, Decision Support in Numerically Imprecise Domains, Ph.D. thesis, Report 94-003-DSV, Department of Computer and Systems Sciences, Stockholm University.

Ekenberg, L., and Danielson, M.: 1994b, "A Support System for Real-Life Decisions in Numerically Imprecise Domains", Proceedings of the International Conference on Operations Research '94, Springer Verlag, 1994.

Ekenberg, L., Oberoi, S., and Orci, I.: 1994c, "A Cost Model for Managing Information Security Hazards", Proceedings of 10th IFIP SEC Conference, Elsevier, North-Holland, 1994.

Fishburn, P.: 1981, "Subjective Expected Utility: A Review of Normative Theories", Theory and Decision 13, pp.139-199.

Freeling, A.N.S.: 1980, "Fuzzy Sets and Decision Analysis", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-10, No.7, pp.341-354.

Good, I. J.: 1962, "Subjective Probability as the Measure of a Non-measurable Set", *Logic, Methodology, and the Philosophy of Science*, eds. Suppes, Nagel, Tarski, pp.319-329, Stanford University Press.

Green, B.: 1992, "Vad kan bankerna lära sig av en entrepenör som utvecklas till organisationsforskare", *Riskbedömning – kunskap om risker*, NUTEK, Stockholm, pp.121-126.

Gärdenfors, P., and Sahlin, N-E.: 1982, "Unreliable Probabilities, Risk Taking, and Decision Making", *Synthese* 53, pp.361-386.

Hamilton, G.: 1988, *This is Risk Management*, Studentlitteratur, Chartwell-Bratt.

Huber, P.J.: 1973a, "The Case of Choquet Capacities in Statistics", *Bullentin of the International Statistical Institute*, Vol. 45, book 4, pp.181-188.

Huber, P.J., and Strassen, V.: 1973b, "Minimax Tests and the Neyman-Pearsons Lemma for Capacities", *Annals of Statistics 1*, pp.251-263.

Levi, I.: 1974, "On Indeterminate Probabilities", *The Journal of Philosophy*, Vol. 71, pp.391-418.

Malmnäs, P-E.: 1994a, "Towards a Mechanization of Real Life Decisions", *Papers in Logic, Methodology, and Philosophy of Science*, eds. Prawitz and Westerståhl, Kluwer Academic Publishers, Dortrecht.

Malmnäs, P-E.: 1994b, "Axiomatic Justifications of the Utility Principle", *Synthese*, Vol. 99, No.2.

Malmnäs, P-E., Danielson, M., and Ekenberg, L.: 1995, *Decision Analysis of the Spent Nuclear Fuel Issue in Sweden*, forthcoming.

Nilsson, N.: 1986, "Probabilistic Logic", *Artificial Intelligence* 28, pp.71-87.

Pfleger, C.P.: 1989, *Security in Computing*, Prentice-Hall, Inc.

SAF: 1986, *Riskanalys*, Näringslivets Beredskapsbyrå.

Shafer, G.: 1976, *A Mathematical Theory of Evidence*, Princeton University Press.

Smith, C.A.B.: 1961, "Consistency in Statistical Inference and Decision", *Journal of the Royal Statistic Society*, Series B, XXIII, pp.1-25.

Statskontoret: 1989-91, *Vägledning i ADB-säkerhet 1-8.*

Weichelberger, K., and Pöhlman, S.: 1990, *A Methodology for Uncertainty in Knowledge-Based Systems*, Lecture Notes in Artificial Intelligence, Springer-Verlag.

Wermdalen, H.: 1991, *Securitas – Säkerhetsboken 1992*, Studentlitteratur.

Wrede, R.: 1984, "The SBA Method: A Method for Testing Vulnerability", *IFIP/Sec '84*, pp.313-320.

Zimmermann, H.J., Zadeh, L.A., and Gaines, B.R.: 1984, *Fuzzy Sets and Decision Analysis*, TIMS Studies in the Management Sciences, Vol. 20, North-Holland.