A Risk Assessment approach based on Information Theory

Luis E. Alfie (luis.alfie@gmail.com)

2018/11/15

Abstract

One of the most critical steps in the risk analysis is evaluate the probability and impact of defined scenarios that could impact negatively in organizational processes. In this regard, one of the main objectives of the present paper is to try to achieve, by using information theory, a quantitative approach (this is, a unit measure well-defined) for performing evaluations that incorporates the inherent and residual risk into one probability (and in the final evaluation, as consequence), and the margin of error as result of a wrong evaluation at the moment of assigning probabilities and/or in case of an event not considered as (sufficiently) probable occurs.

Keywords: Risk Assessment, Probability, Information Theory, Shannon.

1. Introduction

Claude Shannon, in a famous paper[1] of 1948, laid the foundation of the information theory assuming that the information *I(E)* of a source is a logarithmical function of the probability of an event *E*, this is $I=I(p)$ where $0 \leq p \leq 1$, and, at the same time, this information function is inversely proportional to the occurrence of the event $I=1/p$, this is, the more certainty we have about the result of an event, less information we have.

In consequence, we define information as: $I=-\log_b p$, where being b=2, the results are expressed in bits.

For example, if we'd want to know how much information has a tail of the coin, we obtain: $I=-\log_2 0.5=1$ bit. In the same way, if we'd like to know the information associated to two faces of a die, the information adds up[2]: $I=-\log_2(1/6)-\log_2(1/6)=5{,}17$ bits.

2. Probability and Information

Now, let see the way that information behaves. To do that, we will assume that the likelihood is in function of the days per year.

Under this assumption, we obtain:

E	Frequency	$P(E_i)$	$I(E_i)$ bits
1	Daily	0,99000000000	0,01450
2	Weekly	0,14246575342	2,81131
3	Biweekly	0,07123287671	3,81131
4	Monthly	0,03287671233	4,92679
5	Quarterly	0,01095890411	6,51175
6	Semiannual	0,00547945205	7,51175
7	Annual	0,00273972603	8,51175
8	Bi-Annual	0,00136986301	9,51175
9	Quinquennial	0,00054794521	10,83368
10	Decennial	0,00027397260	11,83368
11	Twenty-year	0,00013698630	12,83368
12	Fifty-year	0,00005479452	14,15561

Table 1: Information value examples.

As we can observed, the value of the information grows quickly as the probability of occurrence decrease, and it has so much sense because if an event that we think that it's very improbable that happens occurs, we will obtain much information than the information that we could obtain if an event that we know is very probable, happens. Moreover, if the value of $P(E_i)$ is equal to 1 (absolute certainty of occurrence) or 0 (the event never happens), the value of the information becomes 0 because we have absolute certainty that the event is/isn't going to occur.

So, as first approach, we have converted a probability into a more significant unit: *bits*, the elemental unit of information.

3. Inherent and Residual Information

As we already know, the traditional risk methodology suggests the definition of risk as probability by impact where, if we don't take into consideration the efficiency and effectiveness of the controls in place, we obtain the inherent risk, and, if we do, we obtain the residual one. Although the details under this classical methodology, I'd like to propose a new way doing use of the information unit (bits). The approach supposes, previous

[1] https://culturemath.ens.fr/sites/default/files/p3-shannon.pdf

[2] There are several reasons for using a logarithmic function, but looking into this in a deeply way exceed the purpose of the present work. Just say that one of its properties implies that $\log_2(p_1) + \log_2(p_2) = \log_2(p_1 . p_2)$. This mathematical relationship has remarkable relevance from an information theory point of view.

incorporating the impact associated to an event, relate the information associated to the likelihood of *occurrence* (let's called *inherent* probability P_{inh}) and the one associated to the likelihood of *success*, this is, the likelihood that the event occurs and in the successfully way (let's called *residual* probability P_{res}), this is: $P_t=P_{inh}.P_{res}.$ Now, following the idea of becoming probabilities in bits, we define the residual information as $I_{res} = -log_2P_t$. So, for example, if we know that an event can occurs 1 time per month, and our unit is the year, and also we know that the likelihood that the event materializes is only 5 % (because that, for example, we performed a revision of the controls in place recently with an adequate grading), we can determine the residual information as $I_{res} = -log_2((12/365).0,05)=9,2487$ *bits.* It's to notice the way that the residual information increased, as we can observe in the table 1, a monthly inherent information is equal to 4,9268 bits. Just for comparison, if we had taken a residual likelihood of 0,95 (because there are not any controls in place, for example) the result had been of 5,0008 bits. Indeed, closer to 1 the residual likelihood, closer to the inherent information we will be[3] (in consequence, *less* information added to the event).

Now, we can evaluate the way and results of incorporating this new value into the impact.

4. Impact

As the impact is not in relation with probabilities, it cannot be converted to bits, so, in principle, is not the main concern of the present work an exhaustive exam of this risk component. However, with the purpose of evaluating the potential behavior and final evaluations, I'd like to propose the following approach: in first place, in order to have an information-as-asset constant for evaluating the Risk Exposure, we will define an information impact Π_{inf} as the result of the product of the level ($0 > l \leq 1$) of the availability (a), integrity (i), and confidentiality (c) of the information affected in case that the event occurs.

In second place, a Π_{fin} as the result of some financial weighing. Finally, in third place, another Π_o as the result of pondering others impacts (reputational, compliance, legal, etc.) and, finally, a $\Pi_{\%}$ ($-1 < \Pi_{\%} < \infty$) as a "correction" constant that can affect the percentage of the sum of the previous Π's.

This is: $I=\sum\Pi_i.(1-\Pi_{\%})$

5. Risk Exposure

At this point, and thinking in establishing the Risk Exposure of an event/scenario, we could be tempted to relate the residual information and the impact as the quotient between them, this is: $RE=(I_{res}/I)$ in order to obtain the quantity of information per unit impact *(bits/impact)*.

Let's see it with an example (a) the behavior of this relation:

- Threat: Informatics malware.
- Vulnerability: Absence of antimalware installed.
- Scenario: a final user executes, without intention, a malware.
- Likelihood of Occurrence: Monthly.
- Likelihood of Success: 25% (according to registered incidents).
- Π_{inf}(a,i,c) = 0,8.0,2.0,2 = 0,0320 (the availability of the information is the main affectation).
- Π_{fin} = 0,02 (lowest financial impact).
- Π_o = 0,02 (lowest other impacts).
- $\Pi_{\%}$ = -0,2 (we reduce the impact in a 20 % because the internet access is restricted to certain websites, among others mitigaters).

Under this scenario, we obtain the following Risk Exposure:

RE = $I_{res}/I = -log_2 (P_{inh.}P_{res)}) / (\Pi_{inf} + \Pi_{fin} + \Pi_o).(1+ \Pi_{\%})$ = 6,9268/0,0648 = 106,8949 bits/impact.

As we can observe, the quotient indicates the quantity of information per impact unit. In consequence, we could assume that those event/scenarios with less level of information have priority and should be (re)evaluated in first place.

Finally, two more examples[4]:

b)

- Threat: Misuse of sensitive information of management laptops.
- Vulnerability: Absences of mechanisms of encryption.
- Scenario: Theft or loss of laptop.
- Likelihood of Occurrence: Semiannual.
- Likelihood of Success: 50% (audit department in its last revision detected that this percentage of notebooks weren't encrypted).
- Π_{inf}(a,i,c) = 0,6.0,6.0,9 = 0,3240 (the confidentiality of the information is the main affectation, but the others information dimensions also).
- Π_{fin} = 0,30 (medium-low financial impact).
- Π_o = 0,50 (reputational, compliance and even legal impacts).
- $\Pi_{\%}$ = 0 (there is not any additional changes).

RE: 7,5727 bits/impact.

c)

- Threat: Earthquake.
- Vulnerability: Danger zone.
- Scenario: Earthquake that affects the main Datacenter.
- Likelihood of Occurrence: Each-Fifty-years.
- Likelihood of Success: 90%.
- Π_{inf}(a,i,c) = 0,95.0,95.0,02 = 0,0181.
- Π_{fin} = 0,90 (high financial impact).
- Π_o = 0,30 (compliance impacts).
- $\Pi_{\%}$ = -0,70 (there is an alternative Datacenter and a mature BCM).

RE: 39,1544 bits/impact

[3] Just for the records, is for highlighting that this approach ensure that the residual information never can be lower that the inherent one. This is in line with classical methodologies (and the common sense) that affirm that the residual risk value must be "less risky (or equal)" than the inherent one. This is, if we have a medium inherent risk, we can never have a high residual one. Under the proposed approach, we can say that we can never obtain less (residual) information than the inherent one.

[4] The values incorporated in these examples are merely descriptive, with the purpose of show and analyze the behavior of this approach.

6. Heat Map

Now we have established the RE, we should be in conditions of incorporate the information into a risk heat map.

Just for the purpose of showing one way to do it, follow an illustration in basis on the scenarios used as examples. I chose just 3 (three) categories of risk with their respective thresholds[5]: High (0 - 20 bits), Medium (20 - 60 bits), Low (60 - ∞ bits):

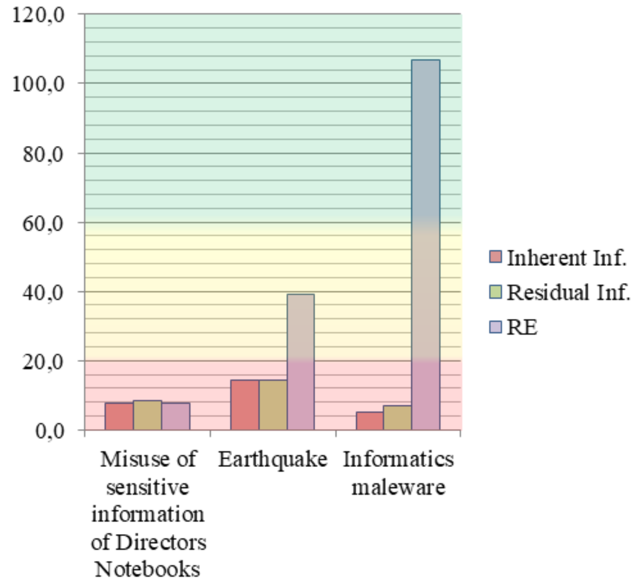

Figure 1: Heat Map ordered by RE

As we can see, this way of representing the RE into a heat map offers a good and intuitive way for analysis and establishing priorities.

Conclusions

According to the objectives that I have proposed, I think that the results of the present analysis are very encouraging for future developments of risk analysis due to the successful incorporation of the level of information related to the events, either about the likelihood of an event occurs (inherent component), as well about the likelihood that it occurs in successfully way (residual component). As we could observe, lower likelihoods (i.e. earthquake example) transfer more information to final results than higher ones, becoming harder take them toward lower risk areas.

On the other hand, and other innovative approach, is the fact of not to multiply the information by impact, but obtaining the quotient between them in order to obtain a quantitative measure with a well-defined unit (bits/impact), allowing, among other, perform objective comparisons between results to the extent that the methodology for calculating the impacts is the same.

[5] The thresholds keep direct relation with the distribution of the RE that, at the same time, depends (only) directly on the way for calculating the impact (as we saw, the information component I_{res} is a quantitative and objective value). For this reason, the thresholds values should be defined and could vary (substantially) from one company to other.