Why We Keep Getting Caught with Our Pants Down

Michael Barwise

According to a recent BBC news report, the managers of the ill-fated Fukushima nuclear facility "had difficulty realising they were facing a worst-case scenario". I beg to differ - they weren't facing a worst-case scenario at all. They were facing an event more extreme than their worst-case scenario - one that exceeded the capacities of the defences and controls they had built on the basis of that scenario. That's why the incident was such a disaster.

I meet this problem constantly in infosec risk practice, but why does it happen so often? Firstly I think it's because operational risk decision-makers generally have a very shaky understanding of the first principles of probability - particularly of the circumstances under which it can mislead one badly. So instead of going back to basics, they use tools and methodologies provided by others, blindly - or at least without real insight.

Secondly, many of the tools and methodologies are seriously flawed. A widely accepted one that comes to mind (but shall out of courtesy remain nameless) assumes a normal (Gaussian) distribution of hazardous events and categorises as "low" and "high" those events that fall one standard deviation either side of the mean. That means the central 68 per cent of events are classed as "medium", the top 16 per cent are "high" and the bottom 16 per cent are "low". This has several failings in the context of operational risk. It's much too crude a categorisation, the "medium" category is far too wide, and the bell shape of the curve excessively de-emphasises extreme events at the "high" end where it really matters. There's absolutely no warranty that the probability distribution of real hazardous events is Gaussian, or indeed even that it's stable over time. Particularly in the outlier regions it usually isn't either.

In the real world of commerce and engineering, probability distributions of hazardous events are mostly extremely skewed, with frequent trivial events and occasional whoppers that can wipe you out. "Medium" events are much less common than the bell curve suggests. And furthermore, the rarer the event the less certain we can be about its statistical properties. That's because statistics only offer a truly reliable indication of what's actually going to happen as the number of samples you base them on approaches infinity. But most risk professionals are unwilling to accept that number of disasters just to find out what their probability distribution was, and I can't say I blame them.

However, almost all methodologies I have encountered have an even more fundamental failing - premature recourse to these crude "high - medium - low" categorisations. The reduction of each input parameter to one of just three (or if they're going overboard - five) categories as soon as it's been identified prevents the possible cumulative effects of individually minor events being recognised. This matters hugely, as an extreme adverse event almost always results from the chance coincidence of numerous less extreme events - there's very rarely an identifiable sole or even primary causal factor.

But it gets worse. Thus crudely categorised, the inputs are then applied to some non-mathematical (indeed often arbitrary and non-linear) matrix to look up an "answer" expressed in the same primitive terms. Often, in order to handle a greater number of input parameters, this procedure is iterated several times, using the output of previous lookups as inputs to the next matrix.

This approach has two - and only two - advantages. It's quick, and it requires little or no expertise to arrive at a semblance of a result. But it's seldom recognised that the result may in fact be meaningless. Even assuming the matrix is well-conceived for the specific problem in hand (which is particularly questionable in infosec, where we usually apply a standardised matrix to a constantly changing hazard space), the output can never be even as good as the inputs.

The practitioner's overriding problem, though, is to assign an event to the right category in the first place. The matrix not only gives no guidance here, but as we perform successive iterations, any errors tend to be compounded by the repeated process of simplification. Coupled with the excessive width of the "medium" category, this causes poorly controlled inputs to drive the ultimate results inexorably towards "medium". The outcome is the all-too-familiar excess optimism exhibited by many operational risk decision-makers.

But both because it's really hard to detect a consistent rationale to the process as a whole, and due to real data being discarded in favour of arbitrary labels too early in the process, it's effectively impossible to generate even an approximate cross-check of one's results. So it's "take it or leave it"

time, with no warranty that the final risk judgements accord with reality. And as we keep finding out - more and more frequently they don't.

Although it's certainly not the only factor involved, the primitive and uncontrolled way we perform what we like to believe are risk assessments is the main contributor to the constant stream of accidents and infrastructure failures we experience across the board - from nuclear incidents to online service breaches. To stem this flood before we drown we need to start getting risk judgement right, and that means we need to start understanding what we're doing from first principles rather than merely relying on "dashboard knowledge" and blind faith.

Originally appeared on the Infosecurity Network, June 2011