Reducing cybersecurity risk

Better metrics and measurement are the keys to addressing a growing threat.

BY DOUGLAS A. SAMUELSON

Hillary Clinton's emails, Wikileaks and Edward Snowden, among others, have directed muchattentiontocybersecurity. New work indicates that the most serious component of cybersecurity risk is probably the use of inferior methods to assess that risk.

Douglas W. Hubbard, familiar to many *Analytics* readers for his books on measurement (2007, 2014) and risk management (2009), has now entered the discussion with a new book (2016), coauthored by corporate cybersecurity expert Richard Seiersen. They cited surveys of professionals in the cybersecurity field to reach a similar conclusion to the earlier results for more general corporate decision-making (Hubbard and Samuelson, 2009): many decision-makers choose methods that increase their feeling of confidence without actually improving performance. They then neglect to measure how good their forecasting methods are. In business, such sloppiness costs large amounts of money. In cybersecurity, it can cost lives – for instance, undercover intelligence and law

Douglas W. Hubbard

Richard Seiersen

Photo Courtesy of 123rf.com | Benoit Daoust

Risk is almost always understated, effort is misapplied and unrealistic models will remain undiscovered.

enforcement agents exposed, military missions compromised, transportation safety diminished.

The widespread lack of empiricism about the performance of forecasts appears in three ways: 1) reliance on subjective inputs from over-confident (uncalibrated) experts, 2) not going back and improving models as new data become available (and analyzing which new data would be most meaningful), and 3) not measuring the performance of forecasts. Hence, risk is almost always understated, effort is misapplied to the issues that are less likely to have bearing on ultimate outcomes, and unrealistic models will remain undiscovered for much longer periods than should be necessary to observe model performance.

PARTICULAR CHALLENGES

The cybersecurity domain entails additional challenges that make measurement, validation and assessment more difficult. Among these are:

Cultural diffusion. People know (or think they know) things without realizing it and without checking sources. I

often illustrate this phenomenon by asking audiences to name the first two ironclad warships to do battle. Most name the Monitor and the Merrimac. However, the Monitor never fought the Merrimac; she fought the CSS Virginia. The Merrimac was a Union ship that was sunk in an earlier battle, reconstructed by the Confederates and renamed the CSS Virginia. Still, so many people "know" it was the Merrimac that even many historical sites generally use both names to reduce confusion.

The current political campaigns, as they are reflected in postings on social media, also reflect a large amount of information people think they know, without verification – and, in many cases, without openness to fact-checking. Hubbard hammers at the practice of statistical modeling without checking assumptions – another version of harm resulting from "knowing" things that aren't so.

Indirect effects. Hubbard's earlier studies of forecasting in business followed phenomena that could be observed: production, costs and sales. Some of the metrics were obscured, but most could be observed directly by applying some ingenuity. In contrast, the most important thing we could know about a surveillance system, in cybersecurity or other applications, is what it can't detect. This can only be estimated by controlled and carefully monitored "white hat" challenges to the system. (Wargaming such challenges with bold, creative adversary players can generate some of these estimates – partially.)

Thoroughly tracking what the system did do and what it captured is thus a very poor substitute for the metric of true interest, yet this is typically the focus in many organizations. They could do much better by applying indirect detection and inferential methods to estimate the right metric.

False positives. Detecting anomalous behavior, to identify possible insider threats, is a critical element of cybersecurity. Scientifically, such detection is quite similar to detecting suspected fraud and abuse in medical claims. In that field, it is well known that a simple anomaly detection method, such as statistical pattern recognition or unsupervised data mining, will readily generate many anomalies. However, the vast preponderance of such anomalies turns out to be coding error or unusual but legitimate patterns of practice. A second round of pattern detection, to discard identified anomalous claims that strongly resemble claims previously found to be of no interest, greatly improves the proportion of identified cases that merit further investigation.

Another useful activity is to look for providers whose claims never fail a range check or any other naïve form of pattern recognition – indeed, they often have smaller variation than other providers' claims. This can be an indication of a provider who is not performing the services at all and is relying on "looking normal" to avoid detection.

Another approach is to combine multiple areas of activity, wherein each area looks normal but the combination does not – such as hospital claims for surgery without associated claims for a surgeon and anesthetist. Similar analytical logic has reportedly been useful in detecting security insider threats, as well.

Difficulty of assessing forecasts. Actual proven insider threat incidents are quite rare. The ones that turn out to be verified take a long time to investigate and prosecute, and investigators, prosecutors and courts usually withhold much of the important information until the prosecution is resolved. This compounds the problems of assessing quality of forecasts.

For these and other reasons, many professionals in cybersecurity resist quantitative models, even for those components for which data are available and some metrics can be tested. However, as Hubbard and others have noted before, decision-makers trying to manage risk-responsive activities also ultimately want quantitative measures, such as probability of occurrence, estimated consequential costs, and costs to mitigate. Having all these seemingly quantitative metrics derive from qualitative assessments, often no better than guesses by experts who may not be as expert as they think, poses the most serious and fundamental risk in the whole process.

MAKING RISK FORECASTS MORE QUANTITATIVE

Hubbard has described in detail previously, and reiterates in the new book, how to train experts to calibrate their risk assessments, so that, for instance, what they say is 90 percent probable does, in practice, occur about 90 percent of the time upon verification. Similarly, he has elaborated his Applied Information Economics method, which begins with an estimate of how good a forecast would be with perfect information, notes where the imperfect information actually in hand is most severely affecting accuracy, and calculates the expected value of obtaining more information in various areas. This highlights the types of information whose acquisition (or better estimation) would most improve the forecast. Typically, the aspect of the situation about which we know the least is the one in

which even a few more observations would make the biggest difference.

Supporting his claim that "you can measure anything," Hubbard offers three observations: 1) something like what you're doing has been done before, generating data; 2) you have more data than you realize; and 3) you need less data than you think. Even a few observations, well below the number needed to pass a test of statistical significance, can provide a substantial improvement in your appreciation of the situation. And what you don't see is also meaningful: While the absence of proof is not proof of absence, absence of evidence is evidence of absence.

In addition to these improvements in data collection, of course, Hubbard also strongly urges much more and better measurement, over time, of how the resulting assessments track actual events.

ORGANIZING TO DO BETTER

Hubbard and Seiersen urge that organizations, both government and private, that are serious about cybersecurity establish a cybersecurity risk management (CSRM) function, reporting directly to the CEO or CIO and the board or the government equivalents (agency director and top executive council). This function would review all major initiatives for technology risk; monitor and analyze existing controls investments; use proven quantitative methods to understand and communicate risk; maintain organizational risk tolerances in coordination with the chief financial officer, general counsel and the board; manage and monitor exception-management programs that violate established risk tolerances; and maintain cyber insurance policies, in conjunction with legal and finance.

They go on to explain that the auditors under the CSRM's direction can "avoid killing better methods" by auditing all models, including the informal or judgmental ones; auditing models with awareness of their larger context, not just within their own assumptions; refusing to assume that something cannot be measured simply because the model output is ambiguous; asking for more research backing up the relative performance of the model versus alternatives; and being skeptical of claims about the levels of complexity the decision-makers will accept. (Over-simplification to "help" the decision-makers is, in their experience, usually harmful.)

SUMMARY

Cybersecurity risk assessment poses unusual challenges because the events

of interest are rare, because these events often cannot be observed directly, and because experienced analysts in this field tend to distrust quantitative models. Nevertheless, more and better use of quantitative metrics, especially to assess the quality of forecasts, is both feasible and likely to be highly beneficial. ❙

Douglas A. Samuelson *is president and chief scientist of InfoLogix, Inc., an R&D and consulting company in Annandale, Va. He has worked in cybersecurity and cyber-counterintelligence in national security, among other practical applications. He is a frequent contributor to* Analytics *and* OR/MS Today, *and a longtime member of INFORMS.*

REFERENCES

1. **Hubbard, D. W. and R. Seiersen, 2016, "How to Measure Anything in Cybersecurity Risk," John Wiley & Sons.**
2. **Hubbard, D. W. and D. Samuelson, 2009, "Modeling Without Measurements: How the Decision Analysis Culture's Lack of Empiricism Reduces its Effectiveness," *OR/MS Today*, October 2009.**
3. **Hubbard, D. W., 2009, "The Failure of Risk Management: Why It's Broken and How to Fix It," John Wiley & Sons.**
4. **Hubbard, D. W., 2014, "How to Measure Anything: Finding the Value of Intangibles in Business," John Wiley & Sons, 2007; 3rd edition, 2014.**