For Good Measure
The Denominator

DAN GEER

Dan Geer is the CISO for In-Q-Tel and a security researcher with a quantitative bent. He has a long history with the USENIX Association, including officer positions, program committees, etc. dan@geer.org

Let's start with a recent paper that is very much worth your time to read: "Global Cyberspace Is Safer than You Think: Real Trends in Cybercrime" by Eric Jardine and released by Chatham House this past July [1]. Its message is exactly that given by its title: that cyberspace is getting better—not getting worse, that cyberspace is getting more safe—not getting more dangerous. The argument for that message is that thinking cyberspace is ever worse, ever more dangerous comes from failing to properly normalize whatever measures of safety you've heretofore been paying attention to. It is only fair to quote its front matter directly:

> Information technology (IT) security firms such as Norton Symantec and Kaspersky Labs publish yearly reports that generally show the security of cyberspace to be poor and often getting worse. This paper argues that the level of security in cyberspace is actually far better than the picture described by media accounts and IT security reports. Currently, numbers on the occurrence of cybercrime are almost always depicted in either absolute (1,000 attacks per year) or as year-over-year percentage change terms (50 percent more attacks in 2014 than in 2013). To get an accurate picture of the security of cyberspace, cybercrime statistics need to be expressed as a proportion of the growing size of the Internet (similar to the routine practice of expressing crime as a proportion of a population, i.e., 15 murders per 1,000 people per year)....In particular, the absolute numbers tend to lead to one of three misrepresentations: first, the absolute numbers say things are getting worse, while the normalized numbers show that the situation is improving; second, both numbers show that things are improving, but the normalized numbers show that things are getting better at a faster rate; and third, both numbers say that things are getting worse, but the normalized numbers indicate that the situation is deteriorating more slowly than the absolute numbers. Overall, global cyberspace is actually far safer than commonly thought.

In short, Jardine is saying that the denominator matters, i.e., that reporting counts of anything is poorer decision support than reporting rates and proportions, that counts of events per unit time will, and must, mislead. It is incorrect to talk about how much mayhem there is without talking about how much opportunity for mayhem there is.

Jardine's line of critique is entirely straightforward, and cyberspace is not the only place that such arguments about the validity of inference are taking place. As a prominent example, consider Stephen Pinker's 2012 book *The Better Angels of Our Nature: Why Violence Has Declined*. In a synopsis in the *Wall Street Journal*, he wrote:

> We tend to estimate the probability of an event from the ease with which we can recall examples, and scenes of carnage are more likely to be beamed into our homes and burned into our memories than footage of people dying of old age. There will always be enough violent deaths to fill the evening news, so people's impressions of violence will be disconnected from its actual likelihood.

This is, again, an argument for looking at rates and proportions rather than counts. But in a direct cross, Nassim Nicholas Taleb responded with a paper, "On the Super-Additivity and Estimation Biases of Quantile Contributions" [2], in which he argues that when a distribution is fat-tailed, estimations of parameters based on historical experience will inevitably mislead:

> When I finished writing *The Black Swan*, in 2006, I was confronted with ideas of "great moderation," by people who did not realize that the process was getting fatter and fatter tails (from operational and financial leverage, complexity, interdependence, etc.), meaning fewer but deeper departures from the mean. The fact that nuclear bombs explode less often than regular shells does not make them safer. Needless to say that with the arrival of the events of 2008, I did not have to explain myself too much. Nevertheless people in economics are still using the methods that led to the "great moderation" narrative, and Bernanke, the protagonist of the theory, had his mandate renewed.

And to highlight his central point:

> [We are] undergoing a switch between [continuous low grade volatility] to ... the process moving by jumps, with less and less variations outside of jumps.

You will possibly find Taleb's paper difficult, but he is speaking to our interest in cybersecurity—are we getting worse or are we getting better? Is there anything we are currently measuring that is leading us to conclude that we are doing the right thing(s) as inferred from measurements of what we believe to be outcomes? Are our inferences confounded with little understood assumptions about thin tails (Gaussian) when we are actually in a fat-tailed (power law) situation? Are we moving into a world where, as Taleb suggests, we are switching from continuous low grade volatility to less frequent but much larger jump changes in the state of the (our) world?

The present author asked this question in a naive form in the spring of 2008 at SOURCE Boston:

Everyone but everyone classifies the 9/11 attack as out-of-nowhere—a black swan to use Taleb's terminology. That attack changed everything because it was not foreseen. It was a physical attack, but we, here, deal in digital attacks. Many of us have heard the phrase "Digital Pearl Harbor," and many of us here have wished we hadn't. If we talk with a member of the general public, we are likely to hear something like, "Look, you paranoid worry-warts keep predicting a big problem and if it was all that likely it would have happened by now. In fact, every day that goes by without something like that happening makes it more likely that it never will. Would you just stop bothering me?"

Now, a statement like, "That we have gone this long without anything big happening" is precisely the kind of statement that expects stability to continue, and which is necessary but not sufficient for a punctuation of that stability. If we look at 9/11 as digital security people, we might remember that the Nimda virus appeared the evening of September 18, 2001, i.e., a week later. Until that point, we'd never seen a virus that had carried more than one method of attack, and Nimda had five. So, to begin with, even if we had known everything about each of those five methods, including population statistics for the numbers and connectivity of vulnerable machines, we would not have predicted the ability of Nimda to spread as it did as we had not yet thought to model the union of multiple vulnerabilities.

That, however, is not all. For writers of classic virus attacks, the measure of their success is the energy differential between the first entry into a given target and the second, i.e., the bigger the difference in how hard it is to break in the first time and how easy it is to break in the second time, the bigger the win. The lowest energy way to maximize this energy differential is to install a new back door. Nimda followed this pattern and installed such a back door.

Because it had five methods for propagation and because it was evidently written with speed in mind, Nimda was also the fastest spreading virus we had yet seen. That rate of spread is known among infectious-disease people as virulence, and we'll return to that in a moment.

As you know, nearly all malware in the wild persists there. An older virus called E911 was such an example. E911 would cause your modem to dial 911 repeatedly; that is all it did. Now when I call you on the phone, the circuit stays up until the calling party disconnects. When I call 911, however, the circuit stays up until the called party disconnects, a difference that is done at the switch for the obvious reason that you do not want the intruder to cut the phone line and the police dispatcher to have to say, "Now whom was I talking to?" For the police to hang up on a 911 call when the calling party has gone away requires a human decision, made under uncertainty, done at human time scales. Because of this, it is possible to saturate a 911 console and that is precisely what the E911 virus was crafted to do.

The E911 virus was old and forgotten on September 18, 2001, but it was still available on the Net, and, of course, the Internet in the fall of 2001 was still dominated by dial-up connections. We got lucky in the simplest, stupidest, dumb luck kind of way. No jackass had the imagination to grab the E911 virus and re-target it at the back door Nimda was busy installing at warp speed everywhere while we all were preoccupied with watching CNN 24x7. If someone had done that, then everyone in America would have gotten up the morning of September 19 only to find that there was no emergency service available nationwide; it would have been turned off everywhere and all at once, like a light switch. While that would not have been a disaster of a physical sort, I submit that it would have been a grand mal seizure of the public confidence. Clinically, that defines terror; it would have required no planning just opportunistic reaction, and it would have been an unpredictable event whose downstream influence was out of all proportion to its concrete effects. It would parallel the Treasury's position that money lost or banks folded is a private tragedy of no importance, but that public loss of confidence in the financial system must be avoided.

On September 18, 2001, we escaped a public loss of confidence by luck and luck alone. As such, the next time someone tells you that the absence of a major Internet attack to date makes the absence of one tomorrow more assured, you can remind them that this proof (that we escaped such an attack by dumb luck) puts to bed any implication that every day without such an attack makes such an attack less likely. It does not make it less likely, but what it does most assuredly do is make it more surprising when it does come.

So is cyberspace getting worse or getting better? Jardine asks us to normalize how many events did occur to the size of how many events could have occurred, not how many did occur in an interval of unit time. He is correct that the possible event space is expanding dramatically, accelerating in its expansion by all accounts. Part of that is network extent, which I've estimated as having a 35% compound annual growth rate [3]. Part of that is the question of attack surface, per se [4]. In any case, Jardine is right that when we count events, we are misleading ourselves as to whether we are getting better or getting worse. But does changing the divisor alone really make the correction we need?

There is a power law here, to be sure. Wikipedia's concise reminder (under "Power Law") is that "Power-laws have a well-defined mean only if the exponent exceeds 2 and have a finite variance only when the exponent exceeds 3; most identified power laws in nature have exponents such that the mean is well-defined but the variance is not, implying they are capable of black swan behavior." That, my friends, is our situation—cyberspace does not have a well-defined variance for what can go wrong and hence cyberspace is unarguably capable of black swan behavior.

Elroy Dimson famously suggested that the definition of risk is that "more things can happen than will happen" [5], and our rate of growth in interdependence is absolutely making the number of things that can happen larger. Unfortunately, complexity prevents us from counting the number of things that can happen, and hence Jardine's argument that we divide the number of things that did happen by the number of things that could have happened is correct in spirit but would be irrelevant if our estimate of the number of things that could have happened were to be wrong.

Yet if the denominator is the number of things that could have happened and we severely underestimate that, doesn't that make the news even better? Taleb says "no" emphatically; the fat tails of power law distributions enlarge the variance of our estimates, leading to less frequent but more severe failures (The Black Swan). The best one could say is that most days will be better and better but some will be worse than ever. Everything with a power law underneath has that property (think earthquakes and whether one is overdue in California), and cyberspace's interconnectivity and interdependence are inherently power law phenomena.

Put differently, are pessimists getting the right answer for the wrong reasons? Is what Pinker said about the memorableness of televised violence making violence seem more prevalent than it is both true and yet misleading? Is what Jardine said about how looking at time series of cybersecurity failures is inherently misleading when the numbers of failures are not normalized in some way? Is what Taleb describes as the trivializing of risk when a power law is mistaken for a Gaussian the heart of what is in play?

In an article in the *San Francisco Chronicle* [6], Thomas Lee recounted how

> I found myself at a dinner in a fancy Menlo Park hotel to discuss cybersecurity with the executives of top Silicon Valley firms. [T]he mood was decidedly grim.
>
> "A devastating cyberattack is likely to occur in the next five years," said a top HP exec. Companies are nowhere near prepared for it. Neither are the feds. There were plenty of comparisons to hurricanes and earthquakes.
>
> "A slow-moving train wreck," one executive said.
>
> There [is] a kind of collective cognitive dissonance in Americans' thinking about tech. We'll eagerly pursue new innovations like the Internet of Things and electronic health records even as we're increasingly aware how vulnerable such technology makes us to terrorists and criminals.
>
> What struck me about the dinner, attended by executives from Hewlett Packard, Cloudera and PayPal, along with academics and investors, was the naked pessimism expressed by those in the room. Nobody even tried to put a happy face on the situation.

Are those executives, academics, and investors getting the right answer for the wrong reasons? Are Jardine and/or Pinker getting the wrong answer for the right reasons? Is it a truism that when risk cannot be estimated it will therefore be underestimated [7]? How do we tell if we are getting better or getting worse, and how can we explain this to citizens, regulators, and reinsurers? Is Taleb right that fat-tailed distributions and asymmetry are where risk lies and, which is more, that the apparent suppression of small failures is "balanced" by yet-to-be-observed black swan excursions?

Resources

[1] Eric Jardine, "Global Cyberspace Is Safer than You Think: Real Trends in Cybercrime," Centre for International Governance Innovation, Chatham House: http://www.cigionline.org/sites/default/files/no16_web_0.pdf.

[2] Nassim Nicholas Taleb, "On the Super-Additivity and Estimation Biases of Quantile Contributions": http://www.fooledbyrandomness.com/longpeace.pdf.

[3] Dan Geer, "T. S. Kuhn Revisited": http://geer.tinho.net/geer.nsf.06i15.txt.

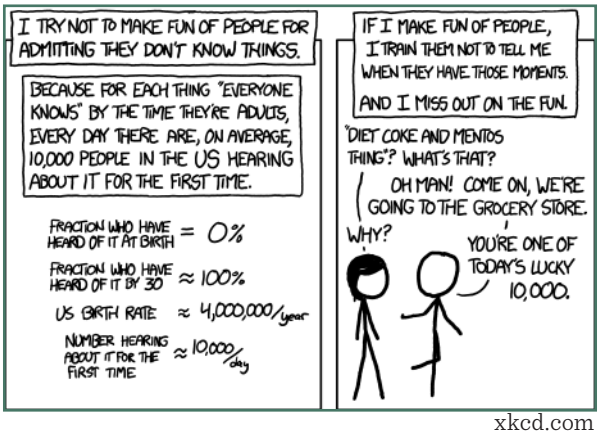
[4] Dan Geer, "Attack Surface Inflation": http://geer.tinho.net/geer.secot.7v14.txt.

[5] Peter L. Bernstein on Risk (Flash video): http://www.mckinsey.com/insights/risk_management/peter_l_bernstein_on_risk.

[6] Thomas Lee, "Forget Target, Ashley Madison Hacks; a Bigger Online Threat Looms" (paywalled): http://www.sfchronicle.com/business/article/Forget-Target-Ashley-Madison-hacks-a-bigger-6395645.php.

[7] Dan Geer, "Cybersecurity as Realpolitik": http://geer.tinho.net/geer.blackhat.6viii14.txt.

XKCD



xkcd.com