Non-Independent Errors

by M. E. Kabay, PhD, CISSP-ISSMP
Associate Professor, Information Assurance
Norwich University, Northfield VT

In teaching students about how to compute the likelihood of failure of complex systems in which all components must function correctly for the system to work (or, put another way, where failure of one or more components results in system failure), statisticians reason as follows:

- Let P be the probability that the system consisting of “n” components will fail in a given period under study.
- Let p(i) be the probability of failure of component “i”.
- Then the probability that component “i” will not fail (i.e., will work) is [1 – p(i)].
- So the probability that all components will work is the product (usually represented as a capital PI) of all terms [1 – p(i)] if the failure of components is random and independent of each other.
- Therefore the probability P that the system will fail is P = 1 – {[1 – p(1)]* [1 – p(2)]*… [1 – p(n)]}
- If we use a system of “n” components where all the probabilities are the same (i.e., p(i)=p) then the formula simplifies to
- P = 1 – {[1 – p]^n}.

For example, if a memory array has 2 GB consisting of 2048 1 MB chips and the likelihood of failure of each 1 MB chip is 1 in a million (1 x 10^6) per year then the likelihood that the array will fail because at least one chip has failed is

P = 1 – {1 – 10^-6}^2048
= 1 – (0.999999)^2048
= 1 - 0.997954095
= 0.002 or 0.2% per year.

Remember, all this depends on independence of failure rates; i.e., we assume for these calculations that failures of chips are not correlated. The fact that one chip fails is not supposed to influence the probability that another chip will fail.

And there’s the problem that reader Paul Schumacher has identified in this standard description of failure rates. Mr Schumacher served in the US Army as an area communications chief many years ago and is now a retired electrical engineer with a reputation in spread spectrum communications. He currently monitors and contributes to discussions of counter-terrorism issues. In particular, he has often contributed to online discussions of items in Bruce Schneier’s CRYPTO-GRAM newsletter <http://www.counterpane.com/crypto-gram.html>. Today, I pass
on a thoughtful and interesting analysis of error rates for anyone interested in risk management. Here are his edited comments about evaluating risk of failure for complex systems that depend on multiple components.

* * *

The equation \([1 - (1 - p)^n]\) is good for independent error rates. However, having a background in high-reliability (jamming-resistant) communications, I have learned that many, if not most, errors are not independent of each other.

If the error is larger in 'volume' than a single bit, it will affect that bit and the bits next to it. Radio communications can be looked upon as a lossy storage medium. If a bit has a duration of 1 microsecond, and the cause of the error lasts 1 nanosecond, then that 1 bit is upset (loses integrity). The chances of the error cause overlapping into the following bit (it occurred just as the bit was about to close), is 1:1000.

Reverse this, so that the bit endures 1 nanosecond, and the cause of the error lasts 1 microsecond, the ratio of neighboring bits being affected becomes 1000:1, or simply, a block of 1000 bits is upset.

On a disk surface, if a cosmic ray, or other physical effect, upsets a single bit, it is likely to also upset the surrounding bits, as it is likely to have a zone of physical effect greater than that of the zone of a single bit.

Error-Correcting Codes (ECC) can correct for large error rates. But when they are unable to correct for more errors than they are designed for, they fail. The smallest ECC I know of is the Hamming 4.7.1 (it encodes 4 data bits into 7 transmitted bits, and can handle 1 error). If non-Gaussian errors occur (i.e., two successive bits are upset), then it is unable to correct them, and possibly does not even allow us to recognize that errors have occurred (two bits in sequence within the same ECC block). To correct this, a technique called interleave is used. The bits of each ECC block are woven with the bits of other blocks so that they are well separated from each other. Using all this, I was able to bring a communications channel from 1:100 error rate (raw) to better than 1:3*10^12 error rate at an overhead of half the bandwidth, which was acceptable.

The proper use of ECC can increase the file reliability tremendously. However, this is only true of random errors. What happens during a systemic error can be very different.

Another kind of problem occurs when the unexpected happens.

At one then-large defense corporation I worked at several decades ago, back-up tapes of the entire computer memory were kept on magnetic tape on a rotating basis. Daily tapes were rotated every week, with only the Saturday tape being kept as the weekly back-up. Beyond a year, only the end-of-month tapes were kept.

When an old back-up tape needed to be referred to, it was discovered that it was totally useless. The janitor had destroyed it, and many others. The destroyed tapes were all located on the bottom shelf of the storage racks: each time the janitor waxed the floors, his buffer's motor generated magnetic fields that slowly degraded the tapes, even inside the metal cans.
There are many lessons to be learned from this accident; the one I found most useful is to use a back-up that is not only physically separate from the main data well-spring, but also physically different in its properties. The corporation recovered from the data loss using paper and microfilm copies.

* * *


M. E. Kabay, PhD, CISSP-ISSMP is Associate Professor in the Division of Business and Management at Norwich University in Northfield, VT. Mich can be reached by e-mail at <mailto:mkabay@norwich.edu>; Web site at <http://www.mekabay.com/index.htm>.

Copyright © 2006 P. Schumacher & M. E. Kabay. All rights reserved.

Permission is hereby granted to Network World to distribute this article at will, to post it without limit on any Web site, and to republish it in any way they see fit.