MIS Quarterly

RESEARCH ARTICLE

CORRELATED FAILURES, DIVERSIFICATION, AND INFORMATION SECURITY RISK MANAGEMENT

Pei-yu Chen
Department of Management Information Systems, Fox School of Business and Management, Temple University,
1801 N. Broad Street, Philadelphia, PA 19122 U.S.A. {pychen@temple.edu}

Gaurav Kataria
Booz & Co., 127 Public Squire, Suite 5300, Cleveland, OH 44114 U.S.A. {gkataria@google.com}

Ramayya Krishnan
School of Information Systems and Management, The Heinz College, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213 U.S.A. {rk2x@cmu.edu}

Appendix A

Proofs

Proof of Observation 1

$$
\begin{aligned}
\frac{\partial \rho_{12}}{\partial c} &= \frac{1+m-2c}{\sqrt{m(m-c)}} + 1/2\,\frac{\big((c(1\_m-c)=m\big)m}{\big(m(m-c)\big)^{3/2}} \\
&= 1/2\,\frac{2m+1-3c}{\sqrt{-m(-m+c)}} \\
&> 0 \because m > c \;\&\; c < 1
\end{aligned}
$$

Since $\frac{\partial \rho}{\partial c} > 0$, increase in shared vulnerabilities increases correlation of failure. QED.

Proof of Proposition 1

Loss distribution under diversity second order stochastically dominates homogeneity if the cumulative area under its cumulative distribution function (CDF) is lower than under homogeneity (see Figure A1), that is,

$$2 \times P_H(2F) > 2 \times P_D(2F)\text{, } and$$

$$2 \times P_H(2F) + 1 \times \left[P_H(2F) + P_H(1F)\right] > 2 \times P_D(2F) + 1 \times \left[P_D(2F) + P_D(1F)\right]$$

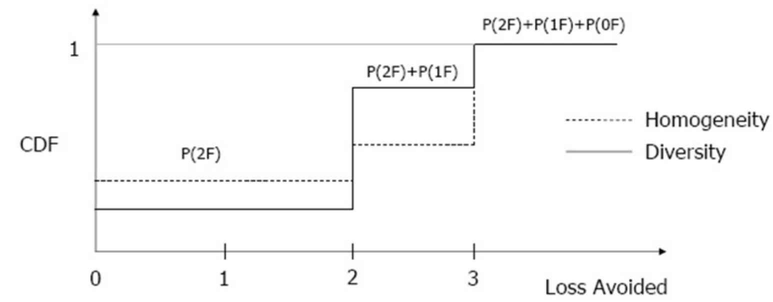

Figure A1. Availability Loss: Cumulative Distribution of Function (CDP)

Which implies that

$$\frac{1+c(\pi+\rho-\pi\rho)}{1+\pi+\rho-\pi\rho} < m < c+(1-c)(1+\pi+\rho-\pi\rho)$$

This takes into account both the software homogeneity scenarios (i.e., all nodes have software configuration 1 or all nodes have software configuration 2). ρ, as before, is the failure correlation for software (configuration) 1 and software (configuration) 2. The aforementioned condition states that when the two software (configurations) have comparable attack rates, software diversity second order stochastically dominates software homogeneity. QED.

Proof of Proposition 2

We can get the results by differentiating the lower bound and upper bound of Proposition 1 to c, π, and ρ.

Derivation of E[Y] and E[Y²]

$E[Y] = E_{attack}E[Y/attack]$

An attack can be one of three types:

1. Specific to vulnerability in software 1.
2. Specific to vulnerability in software 2.
3. Exploiting a vulnerability common to both software 1 and 2.

Therefore, the expected number of failures under diverse deployment is given by

E[Y] = Prob(attack = 1) * E[Y/attack = 1] +Prob(attack = 2) * E[Y/attack = 2] +Prob(attack = common) * E[Y/attack = common]

Now, a is the rate of attacks on software 1, and $m \cdot a$ is the rate of attacks on software 2, where m is related to relative market shares. $c \cdot a$ is the rate of attacks which are common to both software configurations. Then,

$$Prob(attack = 1only) = \frac{1-c}{1+m-c}$$

$$Prob(attack = 2only) = \frac{m-c}{1+m-c}$$

$$Prob(attack = common) = \frac{c}{1+m-c}$$

Therefore,

$$E[Y] = \frac{1-c}{1+m-c} * E[Y_1] + \frac{m-c}{1\_m-c} * E[Y_2] + \frac{c}{1+m-c} * E[Y]$$
$$= \frac{1-c}{1+m-c} * Nx_1\pi + \frac{m-c}{1+m-c} * N(1-x_1)\pi + \frac{c}{1+m-c} * N\pi$$

We know that $E[Y^2] = V[Y] + E[Y]^2$, and $V[Y] = E_{attack}[V[Y/attack]] + V_{attack}[E[Y/attack]]$, using the two we get $E[Y^2] = E_{attack}[V[Y/attack]] + V_{attack}[E[Y/attack]] + E[Y]^2$. Where variance $V[Y]$ for a beta–binomial distribution is given by $V[Y] = N\pi(1-\pi)\rho(1/\rho - 1 + N)$. Therefore, $E[Y^2]$ can be expanded as

$$\begin{aligned}
E[Y^2] = & \frac{(1-c)Nx\pi(1-\pi)\rho(\rho^{-1}-1+Nx)}{m-c} \\
& + \frac{(m-c)N(1-x)\pi(1-\pi)\rho\left(\rho^{-1}-1+N(1-x)\right)}{1+m-c} \\
& + \frac{cN\pi(1-\pi)\rho(\rho^{-1}-1+N)}{1+m-c} \\
& + (1-c)\left(Nx\pi - \frac{(1-c)Nx\pi}{1+m-c} - \frac{(m-c)N(1-x)\pi}{1+m-c} - \frac{cN\pi}{1+m-c}\right)^2 (1+m-c)^{-1} \\
& + (m-c)\left(N(1-x)\pi - \frac{(1-c)Nx\pi}{1+m-c} - \frac{(m-c)N(1-x)\pi}{1+m-c} - \frac{cN\pi}{1+m-c}\right)^2 (1+m-c)^{-1} \\
& + c\left(N\pi - \frac{(1-c)Nx\pi}{1+m-c} - \frac{(m-c)N(1-x)\pi}{1+m-c} - \frac{cN\pi}{1+m-c}\right)^2 (1+m-c)^{-1} \\
& + \left(\frac{(1-c)Nx\pi}{1+m-c} + \frac{(m-c)N(1-x)\pi}{1+m-c} + \frac{cN\pi}{1+m-c}\right)^2
\end{aligned}$$