Valuing Data Security and Privacy using Cyber Insurance

Anand Shah
Tata Consultancy Services
TRDDC, Hadapsar,
Pune, Maharashtra
+91 20 66086378
shah.anand@tcs.com

Shishir Dahake
Tata Consultancy Services
MIDC SEZ, Hinjewadi,
Pune, Maharashtra
+91 20 67940972
shishir.dahake@tcs.com

Sri Hari Haran J
Tata Consultancy Services
TRDDC, Hadapsar,
Pune, Maharashtra
+91 20 66086204
sriharihara n.j@tcs.com

ABSTRACT

What should be the minimum value of data security or privacy to a customer? We reason that at a minimum this value should be equal to the premium charged by an insurer for cyber insurance that compensates the customer for the claims resulting from the data security and privacy breaches. We calculate the premium for cyber insurance and the percentage coverage availed by a customer using Monte Carlo simulations.

General Terms

Management, Measurement, Design, Economics, Security

Keywords

Cyber insurance pricing, data security, data privacy, utility principle, operational risk, value at risk, Basel regulations, pricing of contingent claims in incomplete markets, Monte Carlo simulations

1. INTRODUCTION

Cyber insurance is a hedge to reduce the liabilities resulting from data security and privacy breaches. The importance of this possible hedge to the cyber "risk" becomes apparent when one considers that fine up to 5% of global turnover of the defendant is stipulated under the forthcoming EU General Data Protection Regulation (GDPR). Between January 2013 and October 2014 the Information Commission's office in UK issued 66 enforcement notices and levied penalties of 2.17 million British Pounds under the Data Protection Act, UK [1]. The benefit of cyber insurance is simply not just as a liability mitigating contingent claim but also as a risk capital reducing measure.

Basel regulations are developed by the Basel Committee on banking supervision, to strengthen the regulation, supervision and risk management of the banking sector. Basel regulations define operational risk as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events [2]. Under one of the prescribed approaches for the operational risk measurement (Advanced Measurement Approach), a bank is allowed to recognize the risk mitigating impact of the insurance while calculating the minimum regulatory capital requirement [2]. This offset though is limited to 20% of the total operational risk capital charge. Thus a bank could buy cyber insurance and get a reduction in the operational risk capital requirement [2]. Apart from the above mentioned benefits, a systematic cyber risk assessment by insurers and corresponding differentiation in the insurance premiums charged is likely to provide an impetus to the development of newer data security and privacy technologies [3]. Savings due to reduction in the actuarial premiums attributable to newer and better security and privacy technologies could be a good indicator for pricing these new technologies and tools. At a given time, the minimum value of data security and privacy is the premium charged for cyber insurance to cover all the liabilities that could result from data security and privacy breaches. One could argue that the bigger losses from such breaches are the reputation loss and lost business (opportunity cost) rather than the liabilities that result from the claims. Unfortunately, these bigger losses may not be covered adequately by any cyber insurance hence our modest aim in this work is to at least put some lower bounds on the value of data security and privacy by estimating the premium of cyber insurance.

2. RELATED WORK

Böhme [4] models the individual cyber risk using a Bernoulli distribution. Gritzalis *et al.* [5] propose a probabilistic model for the cyber insurance but the model is specific to an outsourcing environment. Yannacopoulos *et al.* [3] use the classical loss distribution approach (LDA) methodology for modeling the cyber insurance but do not factor in the extent of data loss during a breach in their LDA model. Furthermore, the impact of operational risk capital savings due to cyber insurance is not factored in the premium calculation. Berthold and Böhme [6] value privacy with option pricing theory but we do not adopt such an approach as they use Shannon's "information" as the underlying factor process. Böhme and Schwartz [7] survey the existing models of cyber insurance and propose a network model. We, in this work, model cyber insurance using a modified LDA methodology because LDA methodology is well accepted in the Basel operational risk framework [8].

3. PRICING OF CYBER RISK

Any contingent claim based on the cyber risk is currently priced in incomplete markets as the underlying factor process i.e. the loss over time from the claims resulting from data security and privacy breaches, is not a tradable asset and hence a replicating portfolio cannot be formed using such an underlying factor process. As there are multiple $\boldsymbol{Q}$ measures in an incomplete market, there is no unique price for the cyber insurance. But if the markets were liquid, in order to avoid arbitrage opportunities, the contingent claims in markets may satisfy some internal consistency relations between them. But currently cyber insurances are tailored to the customer's specific needs and thus the insurance market is not liquid. Hence the market price of cyber risk or aggregate risk-aversion in the market is not \easily known [9, 10, 11]. If the cyber liability risk were assumed to be not diversifiable, the insurer would price the cyber insurance using the standard utility principle [3, 9]; under this principle the total premium for issuing cyber insurance will be such that the utility (U) of the initial wealth of the insurer (W) is same as the expectation (E with respect to a $\boldsymbol{P}$ measure) of utility of summation of initial wealth

and premium charged (π) less the expected payoff (L) under the insurance [3].

$$E_P[U(W - L + \pi)] = U(W)$$

4. METHODOLOGY

We follow the classical LDA methodology [8] albeit with some cyber risk specific changes. The loss distribution from the claims resulting out of data security and privacy breaches is estimated by combining the loss frequency distribution with the loss severity distribution. Loss frequency distribution estimates the number of losses observed during the policy period whereas loss severity distribution is the distribution of the size of losses. These distributions are assumed to be independent of each other. The severity of loss is assumed to be dependent on the extent of data loss and the extent of data loss may vary during each loss event. Every loss event obviously may not result in a 100% data loss. Thus we first determine a loss severity distribution for 100% data loss and then estimate the loss severity for the given extent of data loss using a polynomial function. Our rationale is that the loss severity may not be directly proportional to the extent of data loss.

Following are the steps in a simple Monte Carlo simulation that we propose:

As there are n number of simulations, for each simulation:

1. We estimate loss frequency using a Poisson distribution. We sample from this Poisson distribution to determine the number of loss events N(t) where t = 1 year (policy period).
2. For calculating loss (l) for each loss event in the simulation:
   - We sample the loss severity for 100% data loss from a log normal distribution ($l_{100\%}$).
   - The extent of data loss (between 0% - 100%) is sampled using a beta distribution which is bounded between 0 and 1.
   - Then we map this extent of data loss sampled from the beta distribution on to loss severity using a second degree polynomial.
   - Polynomial relationship between severity of loss l_i and extent of data loss θ_i is given by (A and B are constants):

$$l_i = (A\theta_i^2 + B\theta_i)\, l_{100\%}$$

3. Thus the total loss $L(t)$ for that simulation is a simple sum of losses in each loss event.

$$L(t) = \sum_{i=0}^{N(t)} l_i$$

Once the loss distribution is estimated by this Monte Carlo procedure, we calculate operational risk capital using risk measure Value at Risk (VaR) for a given quantile directly from the loss distribution. For a survey and limitation of VaR as a risk measure we direct the reader to work by McNeil *et al.* [8].

We stated in the previous section that the utility pricing by an insurer is given by:

$$E_P[U(W - L + \pi)] = U(W)$$

We assume an exponential utility function [3]

$$U(X) = -e^{-bx}$$

Where b is the coefficient of risk aversion of the insurer and $b > 0$

Thus, it is easy to show that the premium is given by (risk free rate is assumed to be zero).

$$\pi = \frac{1}{b} \ln E(e^{bL})$$

Now given a premium a customer may choose the % coverage (q). This % coverage can be calculated by solving simple maximization:

$$Max_q\, E[U(W_c - L + qL - \pi q + r_\Delta qVaR)]$$

$$Max_q\, E\left[-e^{-b_c(W_c - L + qL - \pi q + r_\Delta qVaR)}\right]$$

Where r_Δ is the incremental return over the risk free rate, VaR is the risk measure for operational risk capital requirement, W_c and b_c are the initial wealth and the coefficient of risk aversion of the customer respectively.

5. RESULTS

We demonstrate the above methodology using a simple example – a base case for pricing the cyber insurance and then analyze the impact of varying risk aversions for both the insurer and the customer on this base case. Following table summarizes the parameters and the results for the base case:

Table 1. Base Case

Parameter	Value
Lambda (Average number of loss events per year)	20
Average Loss size for 100% data lost (in USD millions)	0.2
Deviation of Loss size for 100% data lost (in USD millions)	0.05
Beta Parameter (a)	2
Beta Parameter (b)	5
Functional relationship between the loss severity and the extent of data loss (A =0.95, B = 0.05)	Polynomial
Risk aversion of the insurer	5
Risk aversion of the customer	5
Number of customers (Homogenous)	50
Number of simulations	50,000
Initial wealth of the insurer and that of each customer	0
Policy period (T)	1 year

Table 2. Results for the Base Case

Parameter	Value
Total Premium for all 50 customers (in USD millions)	0.53
Probability of ruin for the insurer	31%
% Coverage by each customer	83%
VaR (99%)	0.86

Figure 1: Beta Distribution for the extent of data loss (Base Case)

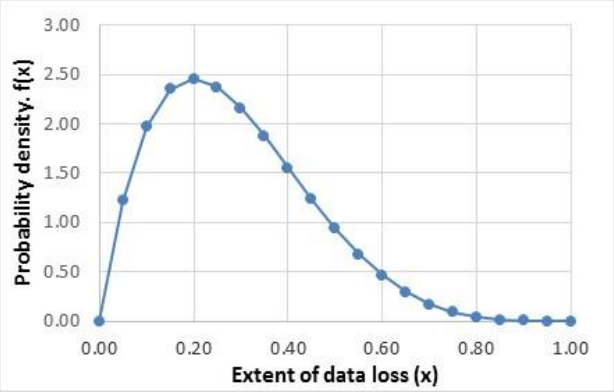


Figure 1 shows the beta distribution used in the base case. As b>a the probability density is skewed to the left, i.e. smaller data loss values are more likely.

Figure 2: Quadratic relationship between the extent of data loss and % of loss severity (Base Case)

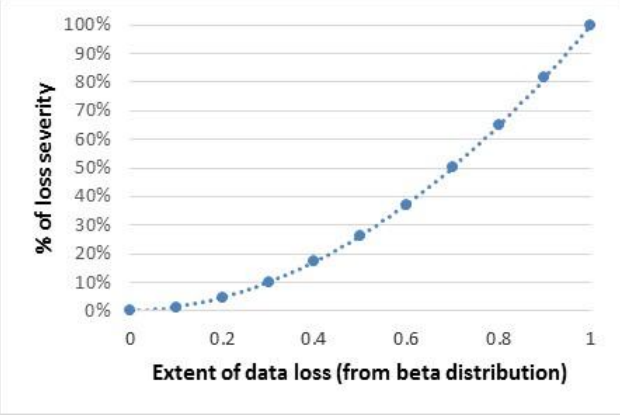
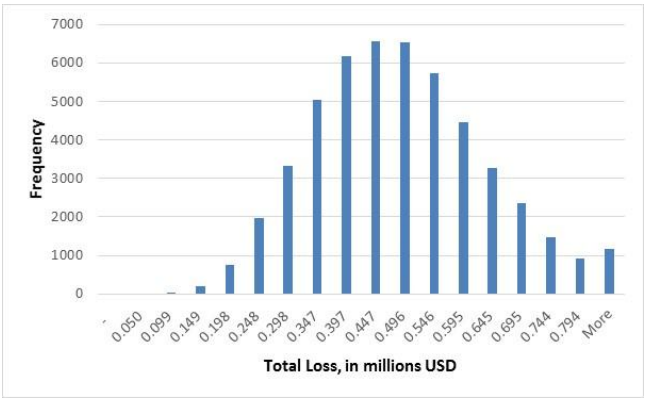


Figure 2 depicts the quadratic relationship between the extent of data loss and the percentage of the loss severity. The loss severity may not increase linearly with the increase in the extent of data loss.

Figure 3: Total loss distribution (Base Case)



Figures 3 captures the simulated total loss distribution (for all 50 customers) for the base case. The input parameters result in total loss distribution that has a right skew and a heavy right tail.

Now for the base case we vary the coefficient of risk aversion of the insurer and access its impact on the total premium charged by the insurer to all 50 customers. As the customers are assumed to be homogenous in terms of risk aversion and data security technologies employed, the premium charged is same for all the customers. The contention here could be that as the customers employ similar data security technologies, the data security and privacy breaches could be highly correlated across customers. But at this juncture we wish to estimate only the minimum bound on the value of data security and privacy hence we assume these breaches to be independent.

Figure 4: Impact of the coefficient of risk aversion of the insurer on the total premium charged and the probability of ruin (all other parameters same as in the base case)

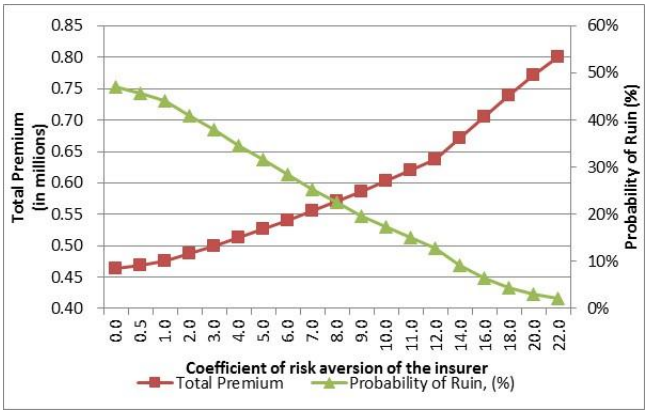


Figure 4 shows that as the coefficient of risk aversion of the insurer increases the total premium charged for the coverage increases and the probability of ruin of the insurer decreases.

Figure 5: Impact of the coefficient of risk aversion of the customer on the % coverage bought by the customer for different coefficient of risk aversions of the insurer (all other parameters same as in the base case)

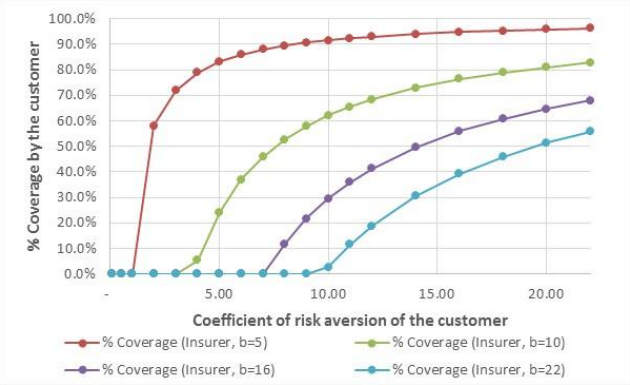


Figure 5 shows that for a given coefficient of risk aversion of the insurer and the corresponding premium charged by the insurer, as the coefficient of risk aversion of the customer increases, the % of coverage bought by the customer increases.

6. CONCLUSION

The minimum bound on the value of the data security and privacy to a customer is determined not just by the premium charged for the cyber insurance by an insurer but also by the customer's own risk aversion. Thus the minimum bound on the value would be the percentage coverage bought by a customer multiplied with the insurance premium charged by an insurer. This approach could also be used to suggest a minimum price of a data security and privacy tool. The minimum price of the tool could be equated to the change in the premium of a cyber-insurance due to the usage of such a tool multiplied by the percentage coverage bought by the customer.

7. ACKNOWLEDGMENTS

Our sincere gratitude to Dr. Sachin Lodha, Dr. Shirish Karande and Mr. K Padmanabhan for support during the work.

8. REFERENCES

[1] Website: IT Governance UK, http://www.itgovernance.co.uk/dpa-penalties.aspx#.VJJ84tKUdps

[2] Basel Committee on Banking Supervision, *International Convergence of Capital Measurement and Capital Standards- A Revised Framework*, June 2006

[3] Yannacopoulos, A. N., Lambrinoudakis, C., Gritzalis, S., Xanthopoulos, S. Z., and Katsikas, S. N., 2008, *Modeling Privacy Insurance Contracts and Their Utilization in Risk Management for ICT Firms,* ESORICS 2008, pp. 207-222, 2008.

[4] Rainer Böhme, 2005, *Cyber-Insurance Revisited,* Workshop on the Economics of Information Security (WEIS) 2005.

[5] Gritzalis, S., Yannacopoulos, A. N., Lambrinoudakis, C., Hatzopoulos P., and Katsikas, S. N., 2007, *A probabilistic model for optimal insurance contracts against security risks and privacy violation in IT outsourcing environments,* International Journal of Information Security (2007) 6 pp. 197-211

[6] Stefan Berthold, Rainer Böhme, 2009, *Valuing privacy with option pricing theory,* Workshop on the Economics of Information Security (WEIS) 2009.

[7] Rainer Böhme, Galina Schwartz, 2010, *Modeling Cyber-Insurance: Towards A Unifying Framework*, Workshop on the Economics of Information Security (WEIS) 2010

[8] Alexander McNeil, Rüdiger Frey, Paul Embrechts, (2005) *Quantitative Risk Management: Concepts Techniques and Tools.* Princeton University Press, 2005

[9] Paul Embrechts, 1996 *Actuarial versus financial pricing of insurance*. Paper presented at the conference on Risk Management of Insurance Firms, The Wharton School of the University of Pennsylvania, 1996.

[10] Tomas Bjork, 2009, *Arbitrage Theory in Continuous Time*, Third Edition, Oxford University Press, 2009

[11] Delbaen, F., Haezendonck, J., 1989, *A martingale approach to premium calculation principles in an arbitrage free market*. Insurance: Mathematics and Economics 8 (1989) 269-277.