International Journal of Engineering Science Invention Research & Development; Vol. IV, Issue IV, OCTOBER 2017
www.ijesird.com, E-ISSN: 2349-6185

A STOCHASTIC MODEL TO ENHANCE INFORMATION SECURITY IN SOFTWARE DEVELOPMENT THROUGH RISK MANAGEMENT

Dr Rajesh Kapur
Associate Professor, TIMSCDR, Mumbai

ABSTRACT- **This paper develops a model of deployment of risk management techniques; to determine the extent to which effort should be expended by scarce resources to deter threats in the hostile information security environment. The data for analysis is taken from public data available with the National Criminal records Bureau (NCRB) and the Indian Computer Emergency Response Team (ICERT). It extrapolates the Weibull distribution to assess the weight-age to be given to each threat by a stochastic process; and indicates the most effective controls to be deployed to counter a specific threat during software development prior to the initiation of the development process. The paper also details the risk management process used.**

I. INTRODUCTION

In the competitive and dynamic environment of the contemporary Indian software industry, information security concerns are often relegated to the background during software application development. Many Indian software companies do not integrate security concerns with mainstream development activities during initial stages of software development [1]. This leads to increased efforts and resource consumption to maintain normal operations when information security related issues arise subsequent to software deployment. Service design and service performance gaps that impact service delivery and quality result when security concerns are not mapped in the initial stages of software quality models [2].

I posit that the reasons for this are twofold: 1) the delivery pressure that makes software teams defer integrating security concerns till the software is deployed in the production environment; and 2) the lack of clarity regarding which specific aspects of information security should be addressed during the design and development stages of software delivery. There is ambiguity about what specific issues are to be addressed - more so because of the limited availability of resources; project managers are not sure of how to optimally deploy scarce resources to implement controls to act as safeguards against future security threats. Adherence to a structured and consistent risk management process will empower program and project managers with tools on which to base scarce resource deployment decisions that incorporate security management controls during the preliminary stages of software development [3].

II. MOTIVATION

The prime motivation for this paper is to develop a structured and consistent model of risk management to be used for basing software development effort; and to foster control protection against information security threats to be embedded in the initial stages of the software development process. The model is based on stochastic risk management techniques.

III. CONTRIBUTION

The model if applied consistently will result in the following benefits:

1. Overall savings in cost and resources in dealing with information security related issues as the deployed controls ensure that safeguards are built into the software ab-initio.
2. Security aspects are embedded into the software by integration of information security requirements into the 'requirement analysis' and 'design' stages of software development. This will result in enhanced cohesion in the code and coordinated response during the emergence of information security threats.
3. The use of stochastic risk management techniques facilitates optimal consumption of the resources for incorporating information security controls during the development process.

IV. REVIEW OF LITERATURE

The process of risk assessment is useful in identifying complex modules that require detailed inspection, estimating potentially troublesome modules, and estimating testing effort [1]. There is however a lack of an established measurement method for the business model for information security, and future research should be directed towards this area [3]. In [5], the difficulties of obtaining estimates of risk probability are highlighted while suggesting a risk management methodology based on group decision making and analytic process hierarchy models. The model is illustrated with a test case. In [6], an overview of the risk management and security management is presented as well as the difference between the two process-domains. It is explained how the two processes-domains are interwoven and cannot be conceived or operated separately. In the execution of processes, there is a need to objectively compare alternatives. This need is fulfilled by determining and analyzing security and related risk metrics as necessary [7]. While lamenting on the lack of sufficient research on the process of developing new, or improving existing, information security risk management methods, [8] proposes "a systematic process for the development of new, or improvement of existing, information security risk management methods" by operating within the paradigm of design science; by emphasizing the effective utilization of pre-existing and new knowledge on information security risk management created throughout the process.

In the context of information security, there are scarce references on how to assess the maturity of a risk management process. Organizations should implement risk management in a consistent, systematic manner in order to achieve compliance with current laws, standards, and regulations; as well as to meet mandatory requirements for the certification of an information security management system (ISMS) [9]. The structure of a model for the assessment of the maturity level of the risk management process in the realm of information security is enunciated. In reference [10], information security risk management (ISRM) is termed a 'major worldwide concern'. It is felt that in spite of the high number of existing ISRM methodologies, the accurate capture of risks for complex information systems for crucial knowledge intensive processes is still carried out in an ad hoc manner. There is a need for systematic and consistent approaches for the development of improved ISRM methodologies that would enhance the effectiveness of the process. They propose a meta-process by laying down specifications for a collaborative and knowledge-sharing platform to support virtual intra-organizational and cross disciplinary teams.

This paper proposes a model for embedding information security controls in the initial stages software development based on stochastic risk management. Probability distributions are used as a basis for analyzing forecasting information security breaches; more specifically the Weibull distribution.

V. RISK MANAGEMENT PROCESS

The methodology of risk management as relevant to this study is

1) Identify the basic purpose, functions, and capabilities of the software application and missions/business processes supported; and how the operation of the software when exploited is vulnerable to information security threats.
2) Identify the threats that can result in an infringement to those rights. The identification of threats (and their respective threat vectors) has been through reports issued by Indian Computer Emergency Response Team (ICERT).
3) Mapping the elucidated threats to specific software functionality. This is subjective. It is also quite likely that a specific threat or a threat vector might lead to the disruption of multiple functionalities.
4) Determination of the probability curves that 'best fits' the data gathered from the National Crime Records Bureau (NCRB) and ICERT sites. For this study the Weibull distribution was identified as the best fit distribution.
5) Identifying the controls (countermeasures) required to counter online threats; assign the effectiveness and cost of each control.
6) Enunciating a decision based stochastic framework which leads to executable action by various stakeholders. As the Weibull distribution has been determined as the best fit for the data gathered, framework will base the decision on linear mapping of heuristically assigned values of the Weibull distribution shaping factor parameters and the effectiveness of the controls.
7) Determining the controls to be deployed at during the development stages of the software.
8) Testing the framework and estimation of review periodicity once the software has been deployed in the production environment.

VI. ANALYSIS OF THREATS & CONTROL DEPLOYMENT

A. ASSUMPTIONS.

In the analysis of threats and control deployment vis-à-vis the meta-model, certain heuristic assumptions have been made. These are:

1) All damages, effectiveness parameters and cost of controls are quantified in terms of standard value units which may be on a scale from 0-10;

2) The likelihood of occurrence is determined by the percentage of occurrence of those threats as part of the cumulative occurrence of threats as given in the statistics by the National Crime Records Bureau (NCRB) and Computer Emergency Response Team-India (CERT-In);

3) the damage that is caused by an infringement to the basic threats is assigned heuristically, as also the effectiveness of each category of controls deployed to mitigate the threat;

4) the cost of each category of controls is relative, and is set on a uniform scale from 0 -10;

5) the range of the 'k' parameter in the Wiebull distribution is from 0.25 to its value as determined from the Weibull plot for each right; and

6) the cumulative value of effectiveness of controls listed for deployment to '0' is linearly mapped to a value of 'k' in the range of .25 to the value of 'k' determined corresponding to the right as determined from the Weibull plot for that right.

Data in respect of software security infringements (obtained from NCRB and CERT-In) is shown tabulated in Appx 'A' and Appx 'B' respectively. The initial phase of the risk management process involves listing the objectives and an analysis of the associated threats. This paper assumes the likelihood concept for probability estimates; likelihood differs from that of a probability in that a probability refers to the occurrence of future

events, while likelihood refers to past events with known outcomes. Consequently, the probability of future occurrence is extrapolated from the frequency of previous occurrences. This approach has been used by [12]. The analysis involves adopting frequency of occurrence as a basic concept to guide the strategy of assessing threat severity, and consequently of the controls to be deployed. Estimation of their effectiveness against a specific threat severity is subjective.

B. THREATS.

The aim of the threat analysis is to determine their relative severity so that the priority of allotment of resources can be set by the development teams. The result of the threat analysis is shown in Table 1. The general function for the calculation of the residual risk, $M_{R,}$ is

$$M_R = M_T – (M_C - M_G) \quad \text{…… Eqn 1}$$

where, M_T = Magnitude of threat (severity)

M_C = Magnitude of reprieve (overall effectiveness) of the controls applied by the framework

M_G = Value of risk introduced regressively by application of controls

Also, $M_T = \sum m_i * p_i$. …….…... Eqn 2

Where, m = severity of the breach (initially assigned heuristically)

p = likelihood of the occurrence of the infringement (obtained from data of NCRB and ICERT)

i = the index number of the threat (to a specific right)

Relative severity of a specific threat = $(M_i / M_T)*100 = (m_i*p_i / \sum m_i*p_i)*100$…. Eqn 3

Another factor that is introduced is the need of dealing with threat – that refers to the feasibility and optimality perspectives of negating the threat during the development stages of the software. For example though phishing has the greatest relative severity of the threats recorded as per the conventional analysis, anti-hacking features are the most appropriate for building in the development stages of the software. That is why negating the hacking threat is given the highest priority for deploying information security controls during software development.

The results of the analysis are tabulated in Table 1.

Ser	Threat Vectors	Freq	Lkelhd (L)	Dmg (D)	Sev (L*D)	Rel Sev	Need	Overall Priority
1	Hacking	956	0.07	90	6.3	14	1	1
2	Obtaining licence or Digital Signature Certificate by misrepresentation/suppression of fact / False publishing	81	0.01	55	0.55	1	.5	7
3	Obscene publication / transmission in electronic form	1657	0.13	40	5.2	11	.2	4
4	Un-authorized access / attempt to access to protected computer system	21	0.01	75	0.75	2	.6	5
5	Phishing	9035	0.7	40	28	60	.2	2
6	Breach of confidentiality/privacy	103	0.01	60	0.6	1	.9	6
7	Tampering computer source documents	366	0.03	85	2.55	5	.8	3
8	Loss due to piracy in software, films, illegal publishing	605	0.05	50	2.5	5	.8	3

Table 1 – Generation of Priorities for Treatment of Threats

C. CONTROL DEPLOYMENT.

As per ISO 27001: 2005 a control is "any administrative, management, technical, or legal method that is used to manage risk. Controls are safeguards or countermeasures. Controls include things like practices, policies, procedures, programs, techniques, technologies, guidelines, and organizational structures." In this paper, control deployment is based on the basic criteria of effectiveness and cost.

The model has to be based on established stochastic probability distributions for credible inferences on which to base decisions vis-à-vis control deployment. The best probability fit for the data has to be investigated, and applied as per the nature and behaviour of the various components of the risk management process. A critical aspect in deciding the deployment of controls deals with the costs and the resource investment that the software development agency is willing to make to mitigate the risk of damage caused by various threats. This is where the process of risk management is relevant. Though the controls are accepted and deployed by a stochastic process, estimation of their effectiveness against a specific threat severity is subjective. In [11], the authors feel that in situations where past data is not available, the decision-maker or the risk-assessment team can subjectively estimate the parameters (or assess prior distributions for the parameters); though this can be a difficult task. Heuristics assignment may be made, as long as the methods work well in practice.

The nature of the response (countermeasures) in respect of the controls to be deployed is based on the Wiebull[1] analysis. Wiebull analysis has many applications in industry, and though it has not been used in the reliability of information security controls, the data for information security incidents for hacking fits the criteria for application of Wiebull analysis as mentioned in (Chapter 2 of) the Wiebull Analysis Handbook [13]; this is that the security incident data when plotted on a Weibull (log-log) graph be in a straight line. The data is tabulated in Table 2 and plotted in Figure 1on a Weibull chart. The year on year data from 2008-14 has been plotted. Since as per Table 1, hacking is the threat with the highest priority, the meta-model shall discuss the risk management process to mitigate the threat of hacking during the software development process. As it falls on a straight line (on the Weibull graph) it can be approximated by a Wiebull distribution which is given by

$$\left.\begin{aligned} f(x;\lambda,k) &= (k/\lambda(x/\lambda)^{k-1})/\exp(x/\lambda)^{k} \quad \text{for } x>=0 \\ &= 0 \text{ for } x < 0 \end{aligned}\right\}$$

Weibull distribution maybe viewed as an extension of the exponential distribution. In Equation 5.1 (above), 'k' is called the shape (or the slope) parameter; when k=1, it reduces to the exponential distribution. It is inferred that the rate of failure of a specific control varies as per the value of 'k' (as determined from the Weibull plot)

If $k < 1$ indicates that the failure rate decreases over time.

$k = 1$ indicates that the failure rate is constant over time.

$k > 1$ indicates that the failure rate increases with time

Let c_i, and e_i be the cost and efficacy of an instance of the control

Then the cost efficacy indicator, CEI is = c_i / e_i

As per the Weibull graph, the value of the slope parameter is 2. The range of .25-2 (for the sloping parameter 'k') has thus to be linearly mapped to a range of 37-0 (for effectiveness of the control). If the aim is to decrease the rate of failure (eg .9), we calculate the required effectiveness as

$$\text{Effectiveness}_{\text{required}} = (37*.9)/1.75 = 19.03 \text{ (approx)}$$

This implies that to lessen the rate of occurrence of hacking breeches, the cumulative total of effectiveness has to greater than 20 (as per Table 1). We start with the control with the lowest value of cost index. We see that as controls at Ser No 4, 5, 7, 8 &3 have the lowest CI correspond to an effectiveness index of 4,5,6,4 &5; and correspond to the minimum value that will lessen the rate of increase of breeches. It would thus be optimal to deploy these controls in the initial stages of the software development.

Ser No	Year	No of Incidents	Cumulative No of Incidents	Cumulative No of Months
1	2009	82	82	12
2	2010	118	200	24
3	2011	164	364	36
4	2012	157	521	48
5	2013	435	956	60
6	2014	550	1506	72

Table 2 - Hacking Details

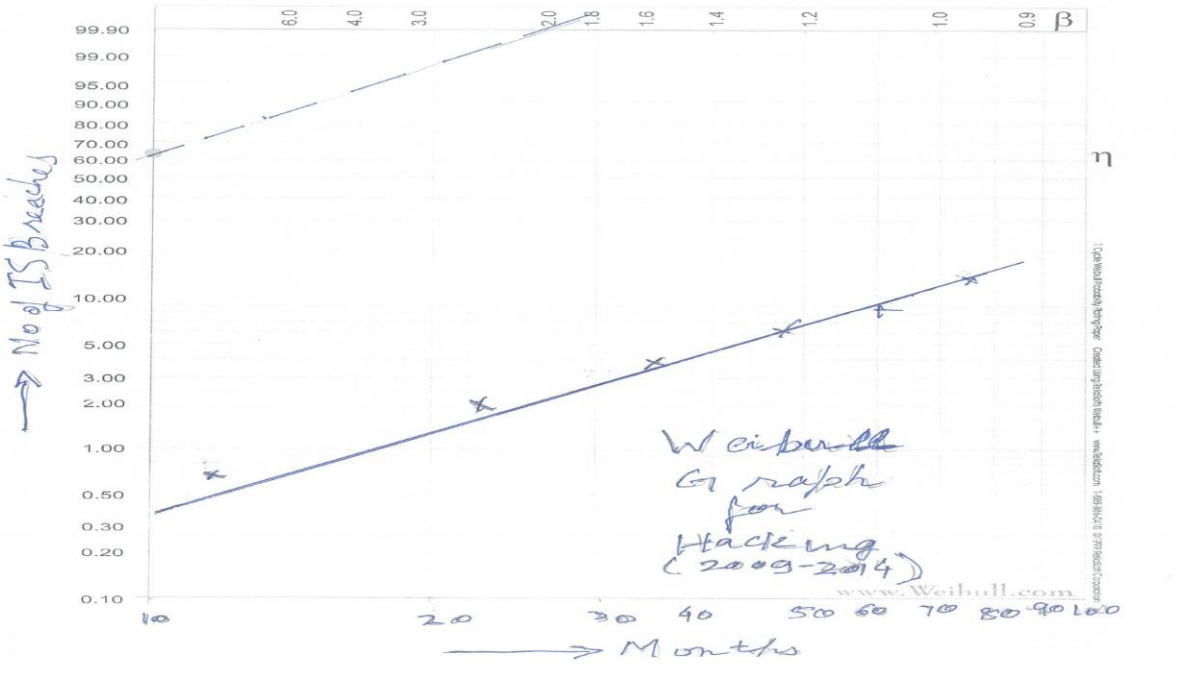

Figure 1 – Weibull Chart for Hacking

Serial No	Illustrative Controls to be deployed during development	Perceived Effectiveness	Cost	Cost Index
1	Design for SQL injection protection	4	8	2
2	Design for cross-scripting (XSS) code protection	4	8	2
3	Design for server side validations	5	9	1.8
4	Force regular password change /strong passwords	4	4	1
5	Design for blocking file uploads / allowing only after malware check	5	8	1.33
6	Design for role based server access	3	6	2
7	Design for In-built check-sum facility for check of data integrity during storage	6	8	1.33
8	Design for default closure of open ports	4	6	1.5

Table 3 - Controls Deployment Analysis

[1]*Waloddi Weibull delivered his hallmark paper on this subjectl in 1951 (Wiebull,, Waloddi (1951), A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, Pg 293-297.) He claimed that his distribution, or more specifically his family of distributions, applied to a wide range of problems; he illustrated this point with seven examples ranging from the yield strength of steel to the size of adult males born in the British Isles, while claiming that the function "-.may sometimes render good service". Time has shown that Waloddi Weibull was correct in all of those statements ; though it has it has many applications in many industries and in particular the aerospace industry, it is being used for control deployment protection on an experimental basis.*

VII. CONCLUSION

This paper has demonstrated the following:

- Information security breaches can be controlled by planning in advance the controls to be built in ab-initio in response to existing threats.

- The Weibul distribution can be used to model a response to the information security threats; however a periodic review to gauge the efficacy and put in place the correction is mandatory.

REFERENCES

1. Sanjay Bahl, O P Wali & Ponnurangam Kumaraguru I*nformation Security Practices Followed in the Indian Software Services Industry: An Exploratory Study*, Cyber security Summit (WCS), 2011 Second Worldwide
2. Sherif M. Yacoub and Hany H. Ammar, *A Methodology for Architecture-Level Reliability Risk Analysis* IEEE Computer Society IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 28, NO. 6, JUNE 20
3. Charlie C. Chen, Chuck C. H. Law, and Samuel C. Yang, *Managing ERP Implementation Failure: A Project Management Perspective*, IEEE Transactions on Engineering Management, Vol. 56, NO. 1, February 2009, pp 157-170
4. Weerahandi, S. and Hausman, R.E., *Software quality measurement based on fault-detection data*, IEEE Transactions on Software Engineering Volume:20 Issue:9 Sep 1994 pp 665 - 676
5. Zhang Xinlan, Huang Zhifang, Wei Guangfu, Zhang Xin, *Information Security Risk Assessment Methodology Research: Group Decision Making and Analytic Hierarchy Process*, Software Engineering (WCSE), 2010 Second World Congress on (Volume:2), 19-20 Dec. 2010, pp 157 – 160
6. Tashi, I. , Solange Ghernouti-Helie, *Information Security Management is Not Only Risk Management* , Internet Monitoring and Protection, 2009, 24-28 May 2009, pp 116 – 123
7. Mehmet Sahinoglu, *An Input–Output Measurable Design for the Security Meter Model to Quantify and Manage Software Security Risk*, IEEE Transactions on Instrumentation and Measurement, Vol. 57, No. 6, JUNE 2000, pp 1251-1260
8. Papadaki, Katerina, Polemi, N. *Towards a Systematic Approach for Improving Information Security Risk Management Methods*, Personal, Indoor and Mobile Radio Communications, 2007, Date 3-7 Sept. 2007 pp1-4
9. Mayer, J. , Lemes Fagundes, L., *A model to assess the maturity level of the Risk Management process in information security* , Integrated Network Management-Workshops, 2009, 1-5 June 2009, pp 61-70
10. Papadaki, E., Polemi, D., Damilos, D.K., *A Holistic, Collaborative, Knowledge-Sharing Approach for Information Security Risk Management* , Internet Monitoring and Protection, 2008, June 29 2008-July 5 2008, pp125-130
11. Clemen, Robert T & Winkler, Robert L, *Combining Probability Distributions From Experts in Risk Analysis*, Risk Analysis, Vol. 19, No. 2, 1999
12. Salvati D, Management of Information System Risks (PhD Dissertation), http://www.dissertation.de/FDP/ds1600.pdf
13. Abernethy RB, Breneman JB, Medlin CH, Reinman GL, *Weibull Analysis Handbook*, Pratt and Whitney Aircraft Government Products Division, United Technologies Corporation, http://www.dtic.mil/dtic/tr/fulltext/u2/a143100.pdf

Apendix 'A'

NCRB DATA

Ser	Activity	2015	2014	2013	2012	2011	2010
1	Tampering computer source documents (Section 65 I T Act)	137	161	94	64	21	26
2	Loss / damage to computer resource / utility (Section 66 (1) I T Act)	1966	1440	826	346	115	56
3	Hacking (Section 66 (2) I T Act)	550	435	157	164	118	82
4	Obscene publication / transmission in electronic form (Section 67 I T Act)	1203	589	496	328	139	105
5	Failure of compliance / orders of Certifying Authority (Section 68 I T Act)	13	6	6	2	3	1

6	Failure to assist in decrypting the information intercepted by Govt Agency (Section 69 I T Act)	6	3	3	3	0	0
7	Un-authorized access / attempt to access to protected computer system (Section 70 I T Act)	27	3	5	3	7	3
8	Obtaining license or Digital Signature Certificate by misrepresentation / suppression of fact (Section 71 I T Act)	9	6	6	9	1	0
9	Publishing false Digital Signature Certificate (Section 73 I T Act) /Frraud	75	11	15	5	5	3
10	Breach of confidentiality / privacy (Section 72 I T Act)	26	45	26	15	10	6
11	Others	274	176	157	30	1	4

APPENDIX 'B'

ICERT DATA

Ser No	Activity	2014-15	2013-14	2012-13	2011-12	2010-11
1	Security Incidents Handled	16078	5431	4727	6828	4003
2	Indian Website Defacements Tracked	16643	14603	10,953	5639	5859
3.	BOT Infected Systems Tracked	5435062	12,798,761	1,635,212	261,1087	1,044,975
4.	Open Proxy Servers Tracked	1807	2558	1,585	2149	2766
5.	Vulnerability Notes Published	97	126	142	117	237
6.	Security Alerts Issued	9	19	22	21	57