SANS | GIAC CERTIFICATIONS

WHITE PAPER

A Preamble Into Aligning Systems Engineering and Information Security Risk

Craig Wright

A preamble into aligning Systems engineering and Information security risk

GIAC (GCPM) Gold Certification

Author: Craig S Wright, craig.wright@Information-Defense.com
Advisor: Rodney Caudle

Accepted: 11th Oct 2011

Abstract

For many years information security and risk management has been an art rather than a science. This has resulted in the reliance on experts whose methodologies and results can vary widely and which have led to the growth of fear, uncertainty and doubt within the community. At the same time, the failure to be able to effectively expend resources in securing systems has created a misalignment of controls and a waste of scare resources with alternative uses. This paper aims to introduce a number of models and methods that are common in many other areas of systems engineering, but which are only just starting to be used in the determination of information systems risk. This paper introduces the idea of using neural networks of hazard data to reliably model and train risk systems.

1. Introduction

This paper presents and extends the major statistical methods used in risk measurement and audit, and extends into other processes that are used within systems engineering (Elliott, Jeanblanc, & Yor, 2000). Security risk assessment is fundamental to the security of any organization (Grandell, 1991). It is essential in ensuring that controls and expenditure are fully commensurate with the risks to which the organization is exposed. The paper starts with defining risk and the terms; next it explores a few of the methods used.

The equations presented in this paper can be used by organizations in order to quantify the relative risk of various solutions and systems and hence to assign risk strategies using the historical data and the organization as well as that of third parties providing a means to optimize audits and system reviews in a manner that detects an incident in the most economical fashion. Projects are all risk consequential endeavors and if our profession can better manage and calculate risk, society will be at an advantage.

The paper defines processes as being the methods that are utilized in order to achieve a set of desired objectives. What is needed is to know just how these processes are implemented within an organization. An objective on the other hand is a goal or something that people desire to have accomplished. It is important to ask just who sets these objectives and how they are designed if risk management solutions are to be achieved effectively and economically.

Controls are the mechanisms through which an individual or group's goals are achieved. Controls are useless if they are not effective. As such, it is important to ensure that any control that is implemented is both effective as well as being justifiable in economic terms. Controls are the countermeasures for vulnerabilities but they need to be economically viable to be effective. There are four types:

1. Deterrent controls reduce the likelihood of a deliberate attack
2. Preventative controls protect vulnerabilities and make an attack unsuccessful or reduce its impact

3. Corrective controls reduce the effect of an attack
4. Detective controls discover potential (attempted) or successful attacks and trigger preventative or corrective controls.

1.1. Identifying classify risk.

A risk analysis is a process that consists of numerous stages. Some of these are defined below:

- Threat analysis,
- Vulnerability analysis,.
- Business impact analysis,
- Likelihood analysis (the probability of an event),

The Risk analysis process should allow the organization to determine the risk for an organization based on threats and vulnerabilities. From this point, the auditor will be able to classify the severity of the risk and thus assign an overall importance to each risk. It should be feasible to use this information to create a risk management plan (Wright, 2008). This should consist of:

- Preparing a risk treatment plan using a variety of control methods.
- Analyzing individual risks based on the impact of the threats and vulnerabilities that have been identified from the risks.
- Rate the individual risks from highest to lowest importance.
- Create a risk treatment plan that categorizes each of the threats and vulnerabilities in order of its priority to the organization, together with some possible controls.

1.1.1. Monte Carlo method

A number of stochastic techniques have been developed to aid in the risk management process. These are based on complex mathematical models that use stochastically generated random values to compute likelihood and other ratios for our analysis model (Corcuera, Imkeller, Kohatsu-Higa, & Nualart, 2004).

Monte Carlo methods can aid in other risk methodologies such as Time-based analysis (Curtis, et al 2001). There is a good introduction to Monte Carlo methods available at http://www.chem.unl.edu (Woller, 1996). This technique further allows for the determination of the range of possible outcomes and delivers a normalized distribution of probabilities for likelihood. Combining stochastic techniques with Bayesian probability and complex time series analysis techniques such as Heteroscedastic mapping is mathematically complex, but can aid in situations where accuracy is crucial (Dellacherie, & Meyer, 1982).

These methods are truly quantitative. They help predict any realistic detection, response and thus exposure time in a manner that can be differentiated by the type or class of attack. These types of statistical methods are known to have a downside in that they are more expensive than the other methods. The level of knowledge needed to conduct a true quantitative type of analysis is not readily available and the level of knowledge of the organization needed by the analyst often excludes using an external consultant in all but the smallest of risk analysis engagements.

2. System Survival

When assessing network reliability, it is necessary to model the various access paths and survival times for not only each system, but for each path to the system. This requires the calculation of the following quantitative fields

- $R(t)$ The Reliability function
- MTBF Mean Time Between Failures
- MTTF Mean Time to Repair/Fix
- λ The expected survival rate (Therneau et. Al. 1994)

Other measures will be introduced later. The expected survival or failure rate λ is used throughout this paper and is detailed further in EQ 3.6 and EQ 3.7. Where possible, the standard systems reliability engineering terms have been used. In the case of a measure such as the MTTF, this represents the time both to discover and recover a

compromised system. The true value estimate for the system comes as a measure of the applications on the system, this may be estimated for a less economically expensive (though less accurate) estimate. In this calculation, the compromise measure, MTBF is best thought of as the mean time to the first failure.

This can be modelled with redundancy in the design. Here, each system is a parallel addition to the model. Where a system is required to pass another, a serial measure is added. For instance, if an attacker has to:

- bypass system A (the firewall) to
- compromise system B (an authentication server) which allows
- an attack against a number of DMZ servers (C, D and E) where
- systems C and D are connected to the database through
- a secondary firewall (system F) to (Not in figure 2.1)
- the database server G (as displayed in figure 2.1).

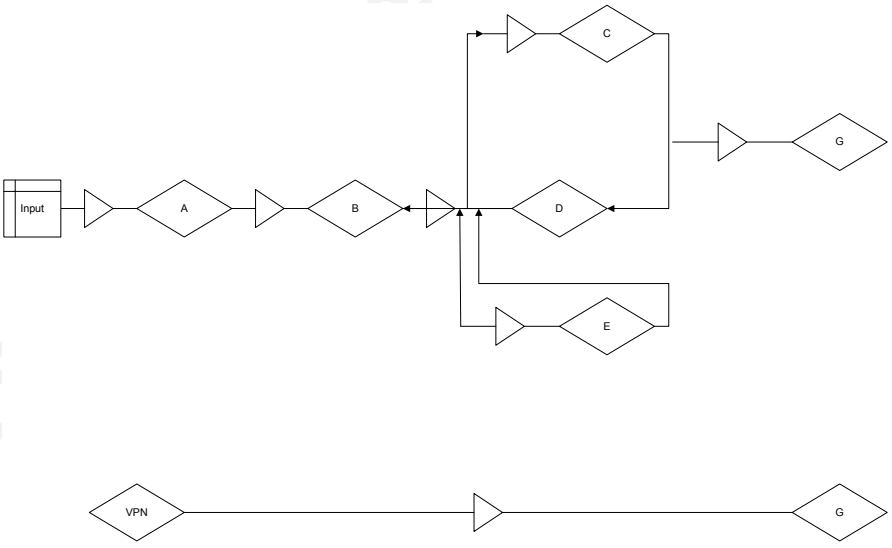

Figure 2.1 Attacking a series of systems

The attacker can either attack the system directly through the VPN or by following the attack path indicated by the systems. If the firewall system A restricted the attacker to a number of IP addresses, the attacker may do 1 of a number of things in attacking this system (in order to gain access as if the attacker was one of these

IPs):

1. Compromise the input host
2. Spoof an address between the input IP address (such as through a router compromise at an ISP or other system)
3. Compromise the VPN

Other options, such as spoofing an address without acting as a MITM (Man In The Middle) will leave some attacks possible that can not result in a compromise of system G. These could have an economic impact that would be calculated separately. Such an event that can be calculated would be a DDoS (Distributed Denial of Services) attack on the server.

Hence, the effective attack paths are:

- Input, A, B, C, F, G
- Input, A, B,D, F, G
- Input, A, B, E, C, F, G
- Input, A, B, E, D, F, G
- VPN, G

In this instance, it is necessary to calculate conditional probabilities as these paths are not independent. Here the options need to consider first include (the paper will use the term I to define an attack on the Input system and S to refer to a spoofed attack of the input system):

- The conditional probability (Annis, 2010) of compromising system A given a successful spoof attack on the Input system, $P(I \cup A_1) = P(I).P(A_1 \mid I)$ (where A_1 refers to an attack on system A using path No. 1, or Input, A, B, C, F, G)
- The conditional probability of attacking system A
- The probability of attacking system G, $P(V \cup G_5) = P(V).P(G_5 \mid V)$

Each of the attack paths are able to be treated as independent. Hence, the overall probability of an attack is a sum of the conditional probabilities from each attack path. As a consequence, the attacker will most likely come over the lowest cost path, but the probabilistic result allows for an attack from any path. The high and low probability attack measures are jointly incorporated in the process.

Presuming no other paths (such as internal attacks etc) it is feasible to model the alternate probability as not possible (or at least feasible). Here $P_6 = 0 = \in$. Additionally, the probability of an attack over path 5 (the VPN) can be readily calculated without further input as:

$$P_5 = P(V \cap G_5) = P(V).P(G_5 \mid V)$$

Here:

$P(V) = e^{-\lambda_V t} + \left(\lambda_V t\right) e^{-\lambda_V t}$ (Blanchard & Fabrycky, 2006) and

$P(G_5) = e^{-\lambda_G t} + \left(\lambda_G t\right) e^{-\lambda_G t}$

$P(G_5 \mid V) = \prod P(G_5)$ EQ 2.0

Here there exists a single system behind the VPN. Where more than one system exists, it is necessary to calculate the joint probability as is detailed below. In the example, with only a single system:

$P(G_5 \mid V) = \prod P(G_5) = P(G_5) \quad 1$

$\therefore P(G_5 \mid V) \rightarrow P(G_5)$ EQ 2.1

Equation 2.1 holds as the probability of the attacker compromising system G when the VPN has been compromised approaches 1. This is as the attacker has a single target with the VPN and the utility of attacking the VPN and no more is negated as no other systems exist and the VPN offers no other utility for the attacker alone.

The values, λ_V and λ_G are the expected survival time or the mean time to compromise for the VPN and database respectively as configured and *t* is the amount of time that has passed from install and represents the current survival time of the system

(Jeanblanc, & Valchev, 2005).

Here:

$$P_5 = P(V \cap G_5) = P(V).P(G_5 \mid V)$$
$$P(G_5 \mid V) = \qquad \text{EQ 2.2}$$
$$= e^{-(\lambda_G + \lambda_V)t} - (\lambda_G \quad \lambda_V)te^{(\lambda_G \quad \lambda_V)t}$$

On the other hand, the probability of compromise to system I is based on the number of systems and as $n_I \to \infty, P(I) \to 1$. Basically, as more systems are allowed to connect to system A, the closer the probability of compromise tends towards a value of P=1. That is, as the systems available to be compromised increase, the probability of compromise approaches certainty. This is generalised as each addition to the system adds a positive probability of compromise when added to an existing system (Blanchard & Fabrycky, 2006). This occurs as no system can be shown to have a probability of compromise P=0. Hence, for each additional component added into the system, the chance of compromise approaches P=1 (where it is finally reached at n=∞).

Where there are only a limited number of systems, the probability can be computed as a sum of the systems. Where there are a large number of systems with equivalent (or at least similar) properties, these can be calculated through the sum of the systems. If in the above example, system E is replaced with a series of systems (each with the same configuration), it is possible to calculate the probability of a compromise of one of the "E" systems as follows:

$$P(E) = R(E) = 1 - \prod_{i=1}^{n} \left[1 - P\left(E_i\right)\right] \qquad \text{EQ 2.3}$$

Here, P(E) is a multiplicative and not additive function. As such, if system "E" is defined as a DNS server with a single BIND service and SSH for management of the host, an attacker has two means to compromising the system;

- Attack SSH
- Attack BIND

The probability can be considered as independent in this case if there are no

restrictions. In the example, DNS is an open service, that is, P(I)=1. The SSH service may or may not be open and could be restricted. If this is the case 0<P(I)<1. In the simple case where no restrictions have been imposed on SSH, the probability can be calculated as a standard independent probability formula. This is:

$$P(SSH) = e^{-\lambda_{SSH}t} + \left(\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}$$
$$P(DNS) = e^{-\lambda_{DNS}t} + \left(\lambda_{DNS}t\right)e^{-\lambda_{DNS}t}$$
$$P(E) = 1 - P(SSH).P(DNS)$$
$$\therefore P(E) = 1 - \left[e^{-(\lambda_{SSH}+\lambda_{DNS})t} + \left(\lambda_{DNS}\lambda_{SSH}\right)t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} + e^{-\lambda_{SSH}t}\left(-\lambda_{DNS}t\right)e^{-\lambda_{DNS}t} + e^{-\lambda_{DNS}t}\left(-\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right]$$
$$P(E) = 1 - \left[e^{-(\lambda_{SSH}+\lambda_{DNS})t} + \left(\lambda_{DNS}\lambda_{SSH}\right)t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - \left(\lambda_{DNS}t\right)e^{-(\lambda_{SSH}+\lambda_{DNS})t} - \left(\lambda_{SSh}t\right)e^{-(\lambda_{SSH}+\lambda_{DNS})t}\right]$$
$$P(E) = 1 - \left[e^{-(\lambda_{SSH}+\lambda_{DNS})t} + \left(\lambda_{DNS}\lambda_{SSH}\right)t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - \left(\lambda_{DNS}-\lambda_{SSh}\right)te^{(\lambda_{SSH}+\lambda_{DNS})t}\right]$$

EQ 2.4

The complication comes where one of the services has been restricted as a further control. This is a combination of the probability of compromising the restrictions on the service (that is spoofing or otherwise bypassing IP address controls) and the compromise of the service itself. This can be represented by:

$$P(SSH) = \left[e^{-\lambda_{SSH}t} + \left(\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right].P(I)$$
$$\therefore P(SSH) = \left[e^{-\lambda_{SSH}t} + \left(-\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right].\left(1 - \prod_{i=1}^{n}\left[1 - P\left(I_i\right)\right]\right)$$
$$= \left[e^{-\lambda_{SSH}t} + \left(-\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right] - \left(\prod_{i=1}^{n}\left[1 - P\left(I_i\right)\right]\right).\left[e^{-\lambda_{SSH}t} + \left(\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right]$$

In this case, there exists a probability $P(I) = 1 - \prod_{i=1}^{n}\left[1 - P\left(I_i\right)\right]$ where the allowed source systems (I) are limited to a total of "n" IP addresses (or keys). The probability $P\left(I_i\right)$ of any source system being compromised will vary, but may be estimated based on the type and location of each system. As more systems are added into the equation, the polynomial equation becomes more complex. In the event that similar systems are also accessing this, these can be calculated and the equation simplified.

For example, if two (2) classes of systems exist (Linux and Windows Vista) that comprise the set of systems I_i for a total of 4 systems (2x Windows and 2x Linux) these

can be defined using:

$$P(I_1);\ P(I_2) = e^{-\lambda_{Win}t} + \left(\lambda_{Win}t\right)e^{-\lambda_{Win}t}$$
$$\&$$
$$P(I_3);\ P(I_4) = e^{-\lambda_{Linux}t} + \left(\lambda_{Linux}t\right)e^{-\lambda_{Linux}t}$$

In the case where $P(I_1);\ P(I_2) = 0.25$ and $P(I_3);\ P(I_4) = 0.2$ due to the system configurations and patch status, it is possible to calculate P(I):

$$P(I) = \left(1 - \prod_{i=1}^{n}\left[1 - P\left(I_i\right)\right]\right)$$
$$= 1 - \left[(1-0.25)^2.(1-0.2)^2\right]$$
$$= 1 - \left[(0.75)^2.(0.8)^2\right] \quad 1 - \left[0.5625 \times 0.64\right] \quad 1 \quad 0.36$$
$$= 0.64 \qquad \text{EQ 2.5}$$

In this case, the probability of a compromise due to SSH would become:

$$P(SSH) = \left[e^{-\lambda_{SSH}t} + \left(\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right].P(I)$$
$$= 0.64\left[e^{-\lambda_{SSH}t} + \left(\lambda_{SSh}t\right)e^{-\lambda_{SSH}t}\right] \qquad \text{EQ 2.6}$$

With the details from the example at 2.2, it is possible to calculate the survival function for system E:

$$P(E) = 1 - 0.64\left[P(SSH).P(DNS)\right]$$
$$\therefore P(E) = 1 - 0.64\left[e^{-(\lambda_{SSH}+\lambda_{DNS})t} + \left(\lambda_{DNS}\lambda_{SSH}\right)t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - \left(\lambda_{DNS} - \lambda_{SSH}\right)e^{(\lambda_{SSH}+\lambda_{DNS})t}\right] \qquad \text{EQ 2.7}$$

Thus, there exists a method to calculate the probability of each system as well as the conditional probability of that system.

The addition of a device (such as an IDS) changes or otherwise impacts *t* and adds additional complexity to the calculations. An IDS for instance can limit the value of *t* through a probabilistic feedback process. The more effective the IDS is, the quicker an attack or other incident will be intercepted. In this instance, *t* becomes a probabilistic function based on how effective the IDS itself is. This becomes a combination of the following factors:

- The inherent accuracy of the IDS (which is a trade-off between TYPE I and TYPE II errors (Rogers, 1996) and it is a cost function in itself)
- The missed detection rate (even where an incident is noted, the analyst may miss the detection. As more false negatives are seen, the missed detection rate increases (Ikeda, & Watanabe, 1962). As a result, increasing false negatives to capture all possible attacks ends in a limit where the IDS is no longer effective).

A Type I error is often denoted as a "false positive". This involves incorrectly rejecting the null hypothesis in favor of the alternative. Where an IDS is involved, a false positive would involve detecting and alerting on an event that did not actually happen to be an incident or attack. A Type II error is the opposite of a Type I error. A Type II error in an IDS system involves the false acceptance of the null hypothesis and is commonly referred to as a false negative. It would imply that the packet or traffic is not an attack and is safe when it is in fact malicious or otherwise dangerous.

The IDS forms a cost function as the increase in reporting results in a greater number of false positives that need to be investigated. In limiting the false positives, the likelihood of missing an incident of note also increases. Each validation of a false positive takes time and requires interaction from an analyst. Hence the tuning of an IDS is balanced on maximizing the detection against cost.

In the event that the IDS does not detect the attack, the function mirrors that of the system without the IDS in effectiveness. Note that the cost of the system with the IDS is greater than the system without IDS. As a result, the addition of IDS is a limiting function. An increase in cost adds to the power of the IDS. This is, more analyst time and more detection capability lowers the false negative and false positive rate through an increase in cost. Each IDS system has an expected TYPE I and TYPE II error rate that will vary as the system is tuned to a particular environment. The result of this is an individualistic function for the organisation that can only be generally approximated for other organisation (even when the same IDS product is deployed).

For a given probability of survival, it is possible to calculate the expected survival time (t) of the system. This process becomes computationally infeasible in large systems

with numerous inputs. For instance, on system E (as defined in EQ 2.3) it is feasible to rearrange the equation of the expected probability of system E being compromised. For instance, if a calculation of the expected function of survival time for a set survival probability P is desired, rearrange the equations in EQ 2.3 as follows.

$$P(E) = 1 - 0.64\left[e^{-(\lambda_{SSH}+\lambda_{DNS})t} + (\lambda_{DNS}\lambda_{SSH})t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - (\lambda_{DNS} - \lambda_{SSh})te^{(\lambda_{SSH}+\lambda_{DNS})t}\right]$$

$$\therefore if\ P = 0.99,\quad 0.99 = 1 - 0.64\left[e^{-(\lambda_{SSH}+\lambda_{DNS})t} + (\lambda_{DNS}\lambda_{SSH})t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - (\lambda_{DNS} - \lambda_{SSh})te^{(\lambda_{SSH}+\lambda_{DNS})t}\right]$$

$$or\quad e^{-(\lambda_{SSH}+\lambda_{DNS})t} + (\lambda_{DNS}\lambda_{SSH})t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - (\lambda_{DNS} - \lambda_{SSh})te^{-(\lambda_{SSH}+\lambda_{DNS})t} = 0.0015625 \qquad \text{EQ 2.8}$$

$$\ln(e^{-(\lambda_{SSH}+\lambda_{DNS})t} + (\lambda_{DNS}\lambda_{SSH})t^2 e^{(\lambda_{SSH}+\lambda_{DNS})t} - (\lambda_{DNS} - \lambda_{SSh})te^{(\lambda_{SSH}+\lambda_{DNS})t}) = 6.4615$$

This result is in the form of:

$$At + B\ln(t) = C$$

From EQ 2.8 it is clearly seen that as $t \to \infty\ [t + \ln(t)] \xrightarrow{t\to\infty} t$. From these equations, as long as *t* is large, an approximation can be deployed to obtain a lower limit estimate of $At + B\ln(t) = C$ as $At\ ;\ C$. As such an approximate for the lower limit of time for system E's survival is defined as:

$$t = \frac{6.4615 + \ln(\lambda_{DNS} + \lambda_{SSH})}{2(\lambda_{SSH} + \lambda_{DNS})} \qquad \text{EQ 2.9}$$

In EQ 2.4, it is demonstrated that the lower the value of t, the greater the error. Measuring "*t*" in seconds and substituting normal system values of λ allows for the use of Monte Carlo simulations to approximate the expected value of *t*.

For simplicity, let R represent reliability and Q the unreliability (hence, $1 - R = Q$).

For each application, a possibility exists to use Bayes' theorem to model the number of vulnerabilities and the associated risk. A mathematical introduction to Bayes' Theorem is available online from Weisstein and in more detail from Joyce (2008). For open ports, the person evaluating risk can use the expected reliability of the software

together with the expected risk of each individual vulnerability to model the expected risk of the application. For instance, it is conceivable to model $P(SSH)$ using this method.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad \text{EQ 2.10}$$

alternatively;

$$P(A \cap B) = P(B)P(A \mid B) = P(A)P(B \mid A)$$

Over time, as vulnerabilities are uncovered and fixed (assuming that new vulnerabilities have not been introduced), fewer issues will remain. Hence, the confidence in the software product increases. This also means that mathematical observations can be used to produce better estimates of the number of software vulnerabilities as more are uncovered.

It is thus possible to observe the time that elapses (Guo, Jarrow, & Zeng, 2005) since the last discovery of a vulnerability. This value is dependent upon the number of vulnerabilities in the system and the number of users of the software. The more vulnerabilities, the faster the discovery rate of bugs. Likewise, the more users of the software, the faster the existing vulnerabilities are found (through both formal and adverse discovery).

2.1. Mapping Vulnerabilities within software

Now let *E* stand for the event where a vulnerability is discovered within the Times *T* and *T*+*h* for n vulnerabilities in the software

$$P(E \mid n) = \int_{T}^{T+h} n\alpha e^{-n\alpha t} dt \approx n\alpha e^{-n\alpha T} h$$

Where a vulnerability is discovered between time *T* and *T*+*h* use Bayes' Theorem to compute the probability that *n* bugs exist in the software:

$$P(n_{vu\ln erabilities} \mid E) = \frac{ne^{-(n\alpha T+\beta)}\frac{\beta^n}{n!}}{\sum_{n=0}^{\infty}\left[ne^{-(n\alpha T+\beta)}\frac{\beta^n}{n!}\right]} \quad \text{EQ 2.11}$$

From this it can be seen that:

$$P(n_{vu\ln erabilities} \mid E) = \frac{\frac{\left(\beta e^{-\alpha T}\right)^{n-1}}{(n-1)!}}{\sum_{n=0}^{\infty}\left[\frac{\left(\beta e^{-\alpha T}\right)^{n-1}}{(n-1)!}\right]} \quad \text{EQ 2.12}$$

EQ 2.12 will apply for all versions of software (Wright & Zia, 2011). As patches and updates are applied to the software, existing vulnerabilities will be rectified and removed, but new flaws related to how many new lines of code have been added in the patching process will be introduced and will also need to be calculated.

By summing the denominator it can understood that in observing a vulnerability at time T after the release and the decay constant for defect discovery is α, then the conditional distribution for the number of defects remaining is a Poisson distribution with expected number of defects $\beta e^{-\alpha T}$.

Hence:

$$P_{\beta e^{-\alpha T}}(n) = e^{\beta e^{-\alpha T}}\frac{\left(\beta e^{-\alpha T}\right)^n}{n!} \quad \text{EQ 2.13}$$

This can be extended to create a method to calculate the expected failure of a system based on the interaction of multiple software products.

3. Exponential Failure

The reliability function (also called the survival function) represents the probability that a system will survive a specified time *t*. Reliability is expressed as either MTBF (Mean time between failures) or MTTF (Mean time to failure). The choice of

terms is related to the system being analysed. In the case of system security, it relates to the time that the system can be expected to survive when exposed to attack. This function is hence defined as:

$$R(t) = 1 - F(t) \quad \text{EQ 3.1}$$

The function F(t) in EQ 3.1 is the probability that the system will fail within the time 't'. As such, this function is the failure distribution function (also called the unreliability function). The randomly distributed expected life of the system *'t'* can be represented by a density function, $f(t)$ and thus the reliability function R(t) can be expressed as:

$$R(t) = 1 - F(t) \quad \int_t^{\infty} f(t)dt \quad \text{EQ 3.2}$$

The time to failure of a system under attack can be expressed as an exponential density function:

$$f(t) = \frac{e^{-t/\theta}}{\theta} \quad \text{EQ 3.3}$$

where θ is the mean survival time of the system when in the hostile environment and t is the time of interest (the time that the user wishes to evaluate the survival of the system over). Together, the reliability function, R(t) can be expressed as:

$$R(t) = \int_t^{\infty} \frac{e^{-t/\theta}}{\theta} dt = e^{-t/\theta} \quad \text{EQ 3.4}$$

The mean (θ) or expected life of the system under hostile conditions can hence be expressed as:

$$R(t) = \int_t^{\infty} e^{-t/M} dt = e^{-t\lambda} \quad \text{EQ 3.5}$$

Where M is the MTBF of the system or component under test and λ is the instantaneous failure rate (Brémaud, 1981) where Mean life and failure rate are related by

the formula:

$$\lambda = \frac{1}{\theta} \qquad \text{EQ 3.6}$$

The failure rate for a specific time interval can also be expressed as:

$$\lambda = \frac{\#Failures}{\sum Operating \quad Hours} \qquad \text{EQ 3.7}$$

Failure rates are generally expressed in terms of failures per hour, percentage of failures per each 1,000 hours or the rate of failures per million hours. For instance, if a system has a 90 day patch cycle (the total mission time) and that the total number of software failures in that time is expected to be (or is later measured to be) 6 vulnerabilities, it is conceivable to calculate the failure rate per hour as:

$$\lambda = \frac{6}{90 \times 24} = \frac{6}{2{,}160} = 0.002778 \qquad \text{EQ 3.8}$$

In the case of an exponential distribution for the system mean survival under attack, the MTBF can be defined as:

$$MTBF = \frac{1}{\lambda} = \frac{1}{0.002778} = 360 hours \qquad \text{EQ 3.9}$$

Hence, it is expected the system to survive 15 days before a vulnerability is discovered. This does not return when a system will actually be exploited, simply the expected probabilistic time that can be used to project and plan future expenditure.

4. Modeling System Audit as a Sequential test with Discovery as a Failure Time Endpoint

Combining hazard models (http://www.statsoft.com) with SIR (Susceptible-Infected-Removed) epidemic modeling (Altmann, 1995) provides a means of calculating the optimal information systems audit strategy. Treating audit as a sequential test allows for the introduction of censoring techniques (Chakrabarty, & Guo, 2007) that enable the estimation of benefits from divergent audit strategies (Benveniste, & Jacod, 1973). This

process can be used to gauge the economic benefits of these strategies in the selection of an optimal audit process designed to maximize the detection of compromised or malware infected hosts.

Computer systems are modeled through periodic audit and monitoring activities. This complicates the standard failure and hazard models that are commonly deployed (Newman, et. al. 2001). A system that is found to have been compromised by an attacker, infected by malware or simply suffering a critical but unexploited vulnerability generally leads to early intervention. This intervention ranges from system patching or reconfiguration to complete rebuilds and decommissioning.

Audits and reviews of computer systems usually follow a prescribed schedule in chronological time. This may be quarterly, annually or to any other set timeframe. Further, periodic reviews and analysis of systems in the form of operational maintenance activities also provide for a potential intervention and discovery of a potential system failure or existing compromise.

Using a combination of industry and organizational recurrence rates that are stipulated from a preceding failure and covariate history as derived from the individual organization introduces a rational foundation in modeling current event data. An incident as defined for the purposes of this paper is an event leading to the failure of the system. This can include a system compromise from an attacker or an infection process of malware (such as a scanning worm). By denoting the number of incidents within the organization as $N(t)$ by follow-up time *t* and $N(t)$ as the corresponding observed incidents in $(0,t]$ with regards to absolute continuous event times, the hazard or intensity process $\lambda(t)$ for the intervention time *t* using the covariate data $X(t)$ can be expressed as:

$$\lambda(t) = P\left[dN(t)\right] = 1\left[N(u), 0 \le u \quad t, X(t)\right] \qquad \text{EQ 4.1}$$

Taking the assumption that the administrative and audit staff are not the direct cause of an incident, a point process $(T_1, T_2, T_3, \ldots)$ will usually be observed for the system

being examined. A system is defined by an isolated and interactive grouping of computers and processes. This could be a collection of client and server hosts located at a specific location isolated by a common firewall.

Due to censoring through the audit process, $N(t)$ can be greater than $\hat{N}(t)$. Equation (4.1) has an assumption that only a single incident has occurred, that is, $\hat{N}$ increments by units. Live systems can and do experience multiple incidents and compromises between detection events. Hence it is also necessary to model the mean increments in $\hat{N}$ over time

$$d\Lambda(t) = E\left[\hat{N}(t) \mid \hat{N}(u), 0 \le u < t, X(t)\right] \quad \text{EQ 4.2}$$

with the cumulative intensity process Λ.

In the case of a continuous-time process with unit jumps, expressions (1) and (2) can be expressed as

$$\Lambda(t) = \int_0^t \lambda(u)\,du \quad \text{EQ 4.3}$$

Independent censorship requires that $C \ge t$ [15]. This assumption of independent censorship allows the preceding covariate histories to be incorporated into the model. In defining $Y(t) = 1(0 < t \le)$, it is now necessary that

$$E\left[dN(t) \mid N(u), Y(u); 0 \le u < t, X(t)\right] = Y(t)\ (t) \quad \text{EQ 4.4}$$

for all times ($t \le C$) prior to the audit or review.

4.1. NHPP, Non-homogeneous Poisson Process

Poison processes have been used to model software (Zhu et. al 2002) and systems failures (Marti, 2008), but these models are too simplistic and it is necessary to vary the intensity (rate) based on historical and other data in order to create accurate risk models for computer systems. The non-homogeneous Poisson process (NHPP) can be used to model a Poisson process with a variable intensity. In the special case when $\lambda(t)$ takes a

constant value λ, the NHPP is reduced to a homogeneous Poisson process with intensity $\lambda(t) = \lambda$.

In the heterogeneous case, an NHPP with intensity $\lambda(t)$, the increment, $N_t - N_u, 0 \le u \quad t$ has a Poisson distribution with an intensity of $\lambda(t) = \int_u^t \lambda(x)dx$. Hence the distribution function of the incident discovery can be expressed as:

$$\begin{aligned} \Lambda_u &= 1 - P(N_{u+t} - N_t = 0) \\ &= 1 - \exp\left(-\int_u^{u+t} \lambda(x)dx\right) \\ &= 1 - \exp\left(-\int_0^t \lambda(u \quad v)dv\right) \end{aligned}$$

The NHPP format is better suited to information systems risk modeling than is the homogeneous Poisson Process as it can incorporate changes that occur over time across the industry.

This can also be modeled as the Poisson process with parameter λ. Here $\left(N_t^{(\lambda)}, t \ge 0\right)$ represents the unique (in law) increasing right continuous process with independent time homogeneous increments. Each $t > 0, N_t^{(\lambda)}$ has a Poisson distribution with rate λt. The process $\left(X_t^{(\lambda)}, t \ge 0\right)$ is also stationary with independent time increments.

With $t_0 = 0$ and $(t_1, ..., t_n) \in °_+^n$, $t_1 < t_2 < ... < t_n$ the r.v.'s $\left[\left(X_{t_1}^{(\lambda)}\right), \left(X_{t_2}^{(\lambda)} - X_{t_1}^{(\lambda)}\right), ..., \left(X_{t_n}^{(\lambda)} - X_{t_{n-1}}^{(\lambda)}\right)\right]$ are independent and for each $k = 1, ..., n, \left(X_{t_n}^{(\lambda)} - X_{t_{n-1}}^{(\lambda)}\right)$ has the same distribution as $\left(X_{t_k - t_{k-1}}^{(\lambda)}\right)$.

The characteristic function of $\frac{1}{\sqrt{\lambda}}\left(X_{t_k - t_{k-1}}^{(\lambda)}\right)$ can be computed for any $\lambda_m \in °$ as:

$$E\left[e^{\frac{i\lambda_m}{\sqrt{\lambda}}\left(X^{(\lambda)}_{t_k-t_{k-1}}\right)}\right] = e^{\left(-i\lambda^{1/2}\lambda_m(t_k-t_{k-1})-\lambda(t_k-t_{k-1})\right)}\sum_{n=0}^{\infty}\left(\frac{\lambda^n\left(t_k-t_{k-1}\right)^n}{n!}e^{\left(\frac{i\lambda_m n}{\sqrt{\lambda}}\right)}\right)$$

$$= e^{\left(-\lambda(t_k-t_{k-1})\left(1-e^{\left(\frac{i\lambda_m}{\sqrt{\lambda}}\right)}\right)-i\lambda_m(t_k-t_{k-1})\sqrt{\lambda}\right)}$$

as $\lambda \to \infty$ the expression converges to $Exp\left(\frac{-\lambda_m^2}{2}(t_k-t_{k-1})\right)$, which is the characteristic function of a Gaussian variable with variance $\sigma^2 = (t_k - t_{k-1})$.

4.2. Recurrent Events

In many cases, audit and review processes are limited in scope and may not form a complete report of the historical processes that have occurred on a system (Revuz, & Yor, 1999). The audit samples selected systems and does not check neighboring systems unless a failure is discovered early in the testing. In these instances, the primary interest resides in selected marginalized intensities that condition only on selected parts of the preceding histories. Some marginal intensity rates drop the preceding incident history all together

$$d\Lambda_m(t) = E\left[dN_m(t) \mid X(t)\right] \qquad \text{EQ 4.5}$$

A common condition for the identification of Λ_m is that

$$E\left\{dN(t) \mid \left[Y(u); 0 \le u < t\right], X(t)\right\} = Y(t)d\Lambda_m(t). \qquad \text{EQ 4.6}$$

For (EQ 4.6) to be valid, censoring intensity cannot depend on the preceding incident history for the system $\left[N(u); 0 \le u < t\right]$. The process of randomly selecting systems to audit makes it unlikely that particularly problematic systems will be re-audited on all occasions. This would include the exclusion of targeting client systems that have been compromised several times in the past or which have suffered more than one incident in recent history. The result is that covariates that are functions of $\left[N_m(u); 0 \le u < t\right]$ will also have to be excluded from the conditioning event. Here

$$E\left[dN(u) \mid X(u)\right] = E\left[dN(u) \mid X(t)\right], \qquad \forall \quad t \geq u. \qquad \text{EQ 4.7}$$

When this occurs

$$\begin{aligned} E\left[dN(u) \mid X(u)\right] &= \int_0^t E\left[dN(u) \mid X(t)\right] \\ &= \int_0^t E\left[dN(u) \mid X(u)\right] \\ &= \Lambda_m(t) \end{aligned}$$

$\Lambda_m(t)$ models the expected number of incidents that have occurred in the system over $(0,t]$ as a function of $X(t)$.

4.3. Cox Intensity Models

Using a Cox-type model

$$d\Lambda(t) = d\Lambda_0(t) e^{[Z(t)'\beta]}, \qquad \text{EQ 4.8}$$

with $Z(t)' = \left[Z_1(t), ..., Z_p(t)\right]$ having been created using functions of $X(t)$ and $\left[N(u); 0 \leq u < t\right]$, inference differs little to univariate failure time data. The log-partial likelihood function, score statistic and the integral notation for the information matrix may be written respectively as

$$\mathrm{l}(\beta) = \log L(\beta) = \sum_{i=1}^{n} \left[\int_0^{\infty} \left\{Z_i(t)'\beta - \log\left[S^{(0)}(\beta, t)\right]\right\} dN_i(t)\right] \qquad \text{EQ 4.9}$$

$$U(\beta) = \frac{\partial \mathrm{l}(\beta)}{\partial \beta} = \sum_{i=1}^{n} \left[\int_0^{\infty} \left\{Z_i(t) - \xi(\beta, t)\right\} dN_i(t)\right] \qquad \text{EQ 4.10}$$

and

$$I(\beta) = \frac{-\partial^2 \mathrm{l}(\beta)}{\partial \beta \partial \beta'} = \sum_{i=1}^{n} \left[\int_0^{\infty} \left\{V(\beta, t)\right\} dN_i(t)\right] \qquad \text{EQ 4.11}$$

where,

$$S^{(j)}(\beta,t) = \sum_{i=1}^{n} Y_i(t) Z_i(t)^{\otimes j} e^{Z_i(t)'\beta}$$

$$\xi(\beta,t) = \frac{S^{(1)}(\beta,t)}{S^{(0)}(\beta,t)}$$

and

$$V(\beta,t) = \frac{S^{2}(\beta,t)}{S^{0}(\beta,t)} - \xi(\beta,t)^{\otimes 2}.$$

By defining $Z(t)$ in terms of fixed or external time varying covariates, (8) can be further defined by adding additional elements $1\left[N\left(t^{-}\right)=1\right]\gamma_1 + 1\left[N\left(t^{-}\right)=2\right]\gamma_2 + \ldots$ to $Z(t)'\beta$. This would allow the intensity to be altered by a multiplicative factor e^{γ_j} following the j[th] incident on an individual system when compared against another system without any incidents at the same point in time.

4.4. SIR (Susceptible-Infected-Removed) epidemic modeling of incidents during Audit

Allowing that a compromised or infected system remains infected for a random amount of time τ, the discovery of an incident by an auditor will be dependent on a combination of the extent of the sample tested during the audit and the rate at which the incident impacts individual hosts. Here, it is assumed that the audit is effective and will uncover an incident if an infected host is reviewed. When a host in a system is infected, any neighboring hosts are attacked and infected at a rate *r*. The sample size selected in the audit is set as κ and the total number of hosts in the system being audited is defined by K were $\kappa \leq \mathrm{K}$. The time between audits (the censor time) is defined by C.

If $C \leq \tau$, an infected or compromised system will be undiscovered and attacking other hosts within the system when the audit occurs. At the end of the time τ, the system is removed as it is either 'dead' - that is decommissioned and reinstalled or it has been patched against the security vulnerability.

A NSW (Newman, Strogatz, and Watts, 2001) random graph is obtained by

investigating the neighboring systems in the SIR model. From this, the thresholds can be computed.

4.4.1. Calculations with a constant τ.

First, consider the case where τ is a constant value, and without loss of generality scale time to make a constant. Start with letting p_k be the degree of distribution.

Starting with a single infected host in a system, it leads to the ability to compute the probability that j of k neighboring hosts will be infected is given by:

$$\hat{p}_j = \sum_{k=j}^{\infty} p_k \binom{k}{j} \left(1 - e^{-r}\right)^j e^{-(k-j)r} \qquad \text{EQ 4.12}$$

Setting μ is the mean of p then the mean of $\hat{p}$ is $\hat{\mu} = \mu\left(1 - e^{-r}\right)$

With the network constructed as an NSW random graph, systems that are compromised in the first and subsequent iterations will each have k neighbors. The value k includes subsequently compromised machines and the host that compromised the existing system. The probability of a compromise associated with these neighboring hosts is given by:

$$q_k = \frac{(k+1)\, p_{k+1}}{\mu} \quad \forall \quad k \geq 0 \qquad \text{EQ 4.13}$$

This allows us to calculate the probability that *j* neighboring hosts also become infected:

$$\hat{q}_j = \sum_{k=j}^{\infty} q_k \binom{k}{j} \left(1 - e^{-r}\right)^j e^{-(k-j)r} \qquad \text{EQ 4.14}$$

Setting ν to represent the mean of q, leads to the mean of $\hat{q}$,

$$\hat{\nu} = \nu\left(1 - e^{-r}\right).$$

From this it is not too difficult to see that for the attack or malware to propagate and infect other systems, it is necessary have the condition where;

$$\nu\left(1-e^{-r}\right). \qquad \text{EQ 4.15}$$

Using this condition provides the capability to calculate the probability that a particular attack or type of malware could result in an outbreak (Thomson, 2007). Setting $T = 1 - e^{-r}$ (Newman, Strogatz, and Watts, 2001) returns:

$$\begin{aligned} \hat{G}_0(z) = & \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} p_k \binom{k}{j} T^j (1-T)^{k-j} z^j \\ = & \sum_{k=0}^{\infty} p_k \sum_{j=0}^{\infty} \binom{k}{j} (Tz)^j (1-T)^{k-j} \\ = & G_0\left(Tz + (1 \quad T)\right). - \end{aligned} \qquad \text{EQ 4.16}$$

Similarly, it is simple to prove that $\hat{G}_1(z) = G_1\left(Tz + (1 \quad T)\right)$.

As such, the probability that an incident behaves as an epidemic is $1 - \hat{G}_0(\xi)$ where $\hat{G}_1(\xi) = \xi$ is the smallest fixed point in $[0,1]$. This is important for where incident can be demonstrated to behave in a manner that lies between [0,1] it is possible to utilize the joint probability substitutions from section 2.

4.4.2. Calculations with a variable τ.

Next it is necessary to consider the effects of a variable or random value of τ. This is the probability that a compromised system causes a compromise in its neighboring system:

$$T = 1 - \int_0^{\infty} dtP(\tau \quad t) e^{-rt}. \qquad \text{EQ 4.17}$$

Again, T is the transmissibility factor.

Newman (Newman, Strogatz, and Watts, 2001) asserted that the infection of neighbors was independent. This does not hold as valid for malware as neighboring systems are commonly linked to form workgroups and domains that share files and even execute code across network boundaries, but it gives a good approximation where most systems are not openly connected as a grid. Interactions in systems and the ability of software to rescan the same systems carry a degree of dependence. Here the time to

compromise may be modeled exponentially with mean λ (again the failure rate), such that $P(\tau = t) = e^{-\lambda t}$.

From this it can be shown that the probability of a host not being compromised is,

$$1 - T = \int_0^\infty e^{-\lambda t} e^{-rt} dt = \int_0^\infty e^{-(\lambda + r)t} dt$$
$$= \frac{1}{(\lambda + r)} \quad . \qquad \text{EQ 4.18}$$

Likewise, the probability that n hosts in a system are not compromised is,

$$1 - T = \int_0^\infty e^{-\lambda t} e^{-nrt} dt = \int_0^\infty e^{-(\lambda + nr)t} dt$$
$$= \frac{1}{(\lambda + nr)} \qquad \text{EQ 4.19}$$

For cases where τ is not constant, Jensen's inequality (Dempster, Laird, and Rubin, 1977) implies that for neighboring hosts, the probability of escaping compromise is positively correlated as

$$E\left(e^{-nrt}\right) > \left(Ee^{-rt}\right)^n . \qquad \text{EQ 4.20}$$

It is now viable to compute the expected number of systems that will be compromised by substituting r_k for p_k and q_k which gives us,

$$\hat{G}(z) = \int_0^\infty P(\tau = 1)\, dt \sum_{j=0}^{\infty} z^j \sum_{j=0}^{\infty} r_k \binom{k}{j} \left(1 - e^{-rt}\right)^j e^{-r(k-j)t} \qquad \text{EQ 4.21}$$

if $r_0 + r_1 > 1$, G is strictly convex as $\hat{\nu} = \nu T$, $\hat{G}_i = G_i\left(1 + (z - 1)T\right)$.

4.5. Applications to audit and review

The first case where $C > \tau$ has a host in the system being discovered as having been infected or compromised before the audit. Here, the rate of infection r determines the chance of other systems being uncovered during an audit. If the time between audits exceeds the time to compromise the first host is insufficient for the incident to spread (i.e.

$\tau \le C, C < r$), then only the initial host will have been compromised and this will be known prior to the audit.

If $C \le \tau$ a compromised host in the system will be undiscovered and attacking other hosts within the system when the audit occurs. At the end of the time τ, the system is found. As such, if $(C + c) \le \tau$ (where c is the average time taken to conduct an audit), a compromised host is discovered during the audit through a process independent of the audit.

The alternative scenario and that which is of most interest is where $(C + c) < \tau$. In this case, the incident will not be discovered independent of the audit. In this instance, the calculation of the probability that an auditor will discover a compromised system during the audit process may be determined as there exists a time limited network function which is coupled with a discovery process that is formulated using the Bayesian prior. That is a risk professional can calculate the probability that all systems are not compromised given a selected audit strategy that finds that none of the audited hosts have been compromised.

So, though it is never feasible to absolutely know if systems that are outside the audit have been compromised, there is a level of relative risk that can be calculated within confidence bounds and used as a means of calculating the expected loss for a variety of risk mitigation strategies. The security professional can then use hypothesis tests to determine if one strategy is significantly better than another and select risk strategies based on expected loss.

4.6. False Negatives in an Audit

False negatives result (Jacod, 1975) in an audit where an incorrectly reported result is supplied noting the organization as safe when it is not (i.e. no compromise was detected where hosts have been compromised). By letting A represent the condition where the organization has been compromised and let B represent the positive evidence of a compromise being reported:

$$P\left(A \mid \overline{B}\right) = \frac{P\left(\overline{B} \mid A\right)P\left(A\right)}{P\left(\overline{B} \mid A\right)P\left(A\right) + P\left(\overline{B} \mid \overline{A}\right)P\left(\overline{A}\right)} \quad \text{EQ 4.22}$$

Here it is practicable to model the actual rate of compromise in the system, $P(A)$. Given a network compromise model (EQ 4.21) it is realistic to substitute the censored time:

$$\hat{G}(z) = \int_0^C P(\tau \le C)dt \sum_{j=0}^{\infty} z^j \sum_{j=0}^{\infty} r_k \binom{k}{j} \left(1 - e^{-rt}\right)^j e^{-r(k-j)t} \quad \text{EQ 4.23}$$

From this the results show that the probability of a host not being compromised in the censor time is,

$$\begin{aligned} P\left(\overline{A}\right) = 1 - T &= \int_0^C e^{-\lambda t} e^{-rt} dt = \int_0^C e^{-(\lambda + r)t} dt \\ &= \frac{e^{-(\lambda+r)t-}\Big|_{-0}^{C+}}{-(\lambda + r)} = \frac{e^{(\lambda\ r)C}}{(\lambda + r)} + \frac{1}{(\lambda + r)} \\ &= \frac{1 - e^{-(\lambda+r)C}}{(\lambda + r)} \end{aligned} \quad \text{.EQ 4.24}$$

Equation (EQ 4.24) also derives the probability of any single host being compromised between the audits

$$P\left(A\right) = 1 - \frac{1 - e^{-(\lambda+r)C}}{(\lambda + r)} \quad \text{EQ 4.25}$$

Depending on whether τ is constant or varies; the process can calculate the expected number of hosts that will be compromised in the period between audits as a fixed or variable function. In either event, each calculation is an exercise in Bayesian estimation of the type where a random sample is selected and the defects or failures are analyzed

$$f\left(x\right) = \binom{n}{x} p^x q^{n-x} \quad \text{EQ 4.26}$$

This binomial distribution is simplified in the case of a false negative (no failures

or x=0 from a sample of n hosts)

$$f(x) = \binom{n}{0} p^x q^{n-x} = p^x q^{n-x} \quad \text{EQ 4.27}$$

Based on the types of systems, the audit periods can be selected to create an economically optimal choice. A baseline audit frequency is frequently set by regulatory guidelines. Here, the baseline requirement would be to have a minimum audit process designed around the most effective returns within the regulatory regime. In this, using an equation that calculates the expected number of compromised or infected hosts within a censor time allows for the selection of either a fixed audit schedule, C = Constant, or vary C over the course of the system life in order to maximize the detection, C=C(t).

In conducting this exercise, the cost of the audit, and differences that occur would also need to be modeled. The required effort for an audit of 10 hosts in a 1 month period is not necessarily linearly related to the audit of 60 hosts in a 6 month period. In each case, the individual constraints faced by the selected organization also need to be incorporated.

5. Automating the process

The main advantage to a systems engineering approach is the ease with which it can be automated. The various inputs and formula noted throughout this paper can become inputs into a neural network algorithm (Fig. 5.1). Equation (2.1) could be modeled in three layers (Fig 5.2).

Here, an input layer with one neuron for each Input (system or application) could be used to map for IP Options, Malware and Buffer overflow conditions, selected attacks etc. The system of perceptrons would be processed using a hidden neuron layer in which each neuron represents combinations of inputs and calculates a response based on current data coupled with expected future data, a prior data and external systems data. Data processed at this level would feed into an output layer. The result of the neural network would supply the output as an economic risk function.

In this way, a risk function can be created that not only calculates data based on

existing and known variables (He, Wang, & Yan, 1992), but also updates automatically using external sources and trends. Many external sources (http://www.dshield.org) have become available in recent years that provide external trending and correlation points. Unfortunately, most of these services have clipped data as the determination of an attack is generally unclear and takes time to diagnose where much otherwise useful data is lost. When monitoring the operation of a system or the actions of uses, thresholds are characteristically defined above or below which alerting, alarms, and exceptions are not reported. This range of activity is regarded as baseline or routine activity.

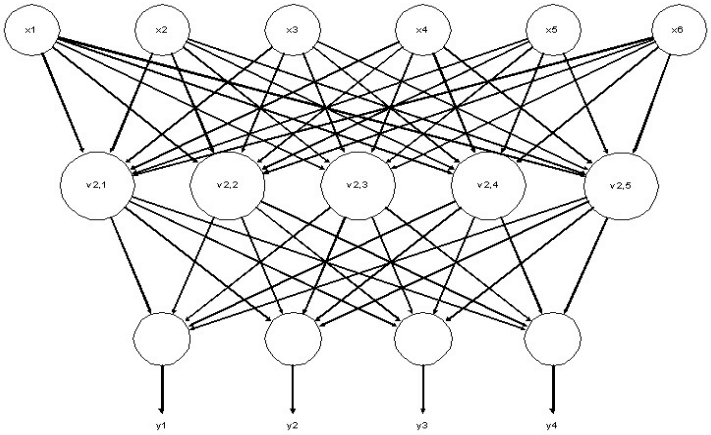

Fig 5.1 A depiction of a Multi-Layer layer topology neural network

Multi-Layer layer topology neural networks can be used to accept data from risk models and automatically update the risk profile of an organization. In modeling risk, each application and system can be modeled using a perceptron.

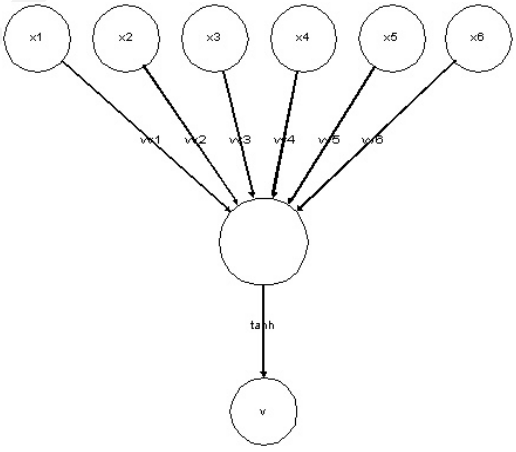

Fig 5.2 Inputs being fed into a perceptron.

The perceptron is the computational workhorse in this system. In this it is reasonable to model the selected risk factors for the system and calculate a base risk that is trained and updated over time. The data from multiple organizations can be fed into a central system (Kay, 1977) that can be distributed to all users. This could be integrated and sold as a product enhancement by existing vendors or independent third parties could maintain external datasets.

$$v_{i,j} = f(\sum_{k=0}^{n} w_{i,j,k} \cdot x_k \ ) \qquad \text{EQ 5.1}$$

EQ 5.1 defines the input variables as,

- $x_1 \ldots x_n$ are the inputs of the neuron,
- $w_{i,j,0} \ldots w_{i,j,n}$ are the weights,
- f is a non-linear activation function,
- hyperbolic tangent (tanh),
- $v_{i,j}$ is the output of the neuron.

A large vendor such as Microsoft could create an implementation model. In place of offering stale recommended security settings (such as currently occurs with Microsoft's MBSA), the risk application could automatically collect data from user systems on patch levels and group policy configurations and utilize these in order to calculate and report on an estimated level of risk and an expected survival time for the system in a number of different scenarios. For instance a notebook computer could have a set of risks. This would include the risk when connected to the corporate network, when connected to a wireless hotspot etc.

The training of the network would require the determination of the correct weights for each neuron. This is possible in selected systems, but a far larger effort would be required to enable this process for more generalized deployment. The data needed for such an effort already exists in projects such as DShield, the Honeynet Project and in

many similar endeavors. The question is whether there truly exists a will as a community to move from an art to a science.

6. Conclusion

The equations presented in this paper allow organizations to compare the deployed risk strategies against both their own historical data and that of third parties. In this manner, strategy can be formulated in order to optimize audits and system reviews in a manner that detects an incident in the most economical manner. Projects are all risk derived exercises and if our profession can better manage and calculate risk, society will benefit.

Modeling the failure rate of systems and the propagation rate of an attack, allows us to calculate an expected number of hosts that are anticipated to have been compromised in the time between an audit given a specified survival function or threat. Past data and comparisons from similar systems (such as survival data from http://www.dshield.org/reports.html) allow for the modeling of alternative systems where a reported number of events have been reported against those deployed.

Dependence, variation, randomness, and frailty add to the risk toolset of multivariate failure event analysis. Using frailty theory to model information system risk allows us to better predict risk and to more effectively allocate scarce resources through selecting the most economically viable targets to defend as well as choosing the optimal detection strategies. The properties of censoring-handling and frailty modeling have turned multivariate survival analysis into an exceptional tool for the determination of system risk.

For decades, information security practitioners have engaged in qualitatively derived risk practices due to the lack of a scientifically valid quantitative risk model. This has led to both a misallocation of valuable resources with alternative uses and a corresponding decrease in the levels of protection for many systems. Using a combination of modern scientific approaches and the advanced data mining techniques that are now available provides the technologies and data to create a new approach to information systems risk and security.

The optimal distribution of economic resources across information system risk allocations can only lead to a combination of more secure systems for a lower overall cost. The reality is that, like all safety as an issue, information security is based on a set of competing trade-offs between economic constraints. The goals of any economically based quantitative process are to minimize cost and risk through the appropriate allocation of capital expenditure. To do this, the correct assignment of economic and legal liability to the parties best able to manage the risk (this is the lowest cost insurer) is essential and needs to be assessed. This will allow insurance firms to develop expert systems that can calculate risk management figures that can be associated with information risk. This will allow for the correct attribution of information security insurance products that can be provided businesses generally.

Externality or the quantitative and qualitative effects on parties that are affected by, but who are not directly involved in a transaction is likewise seldom quantified, but is an integral component of any risk strategy. The costs (negative) or benefits (positive) that apply to third parties are an oft overlooked feature of economics and risk calculations. For instance, network externality (a positive effect that can be related to Metcalfe's law; value of a network = 2 times the network's number of users) attributes positive costs to most organizations with little associated costs to themselves. In these calculations, the time-to-market and first-mover advantages are critical components of the overall economic function with security playing both positive and negative roles at all stages of the process.

The processes that can enable the creation and release of actuarially sound threat risk models that incorporate heterogeneous tendencies in variance across multidimensional determinants while maintaining parsimony already exist in rudimentary form. Extending these though a combination of Heteroscedastic predictors (GARCH/ARIMA etc) coupled with non-parametric survival models will make these tools more robust. Effort needs to be expended in the creation of models where the underlying hazard rate (rather than survival time) is a function of the independent variables (covariates). Cox's Proportional Hazard Model with Time-Dependent Covariates would be a starting point, with a number of non-parametric methods available

where cost allows.

As we move further into the 21st century, it is time we as a profession started to model risk as a scientific process and move away from the art based cottage industry that exists right now. This paper has presented a number of methods that can be used to gauge the expected failure events in a system of computer hosts.

7. References

Altmann, M., (1995) "Susceptible-infected-removed epidemic models with dynamic partnerships", Journal of Mathematical Biology Volume 33, Number 6, 661-675

Annis, C. (2010) "Joint, Marginal, and Conditional Distributions" Retrieved October 11, 2011, from http://www.statisticalengineering.com/joint_marginal_conditional.htm

Benveniste, A. and Jacod, J. (1973). Systèmes de Lévy des processus de Markov. *Invent. Math.* **21** 183--198. Mathematical Reviews

Blanchard, B. and Fabrycky, W. (2006) "Systems Engineering and Analysis", 4th Ed. Prentice Hall International Series in Industrial and Systems Engineering, USA

Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York. Mathematical Reviews

Chakrabarty, A. and Guo, X. (2007). A note on optimal stopping times with filtration expansion. Preprint, Univ. California, Berkeley.

Corcuera, J. M., Imkeller, P., Kohatsu-Higa, A. and Nualart, D. (2004). Additional utility of insiders with imperfect dynamic information. *Finance and Stochastics* **8** 437--450.

Dellacherie, C. and Meyer, P. A. (1982). *Probabilities and Potential. B*. North-Holland, Amsterdam.

Dempster, A.P. and Laird, N.M. and Rubin, D.B. (1977).Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, B39.

Elliott, R. J., Jeanblanc, M. and Yor, M. (2000). On models of default risk. *Math. Finance* **10** 179--195.

Grandell, J. (1991) "Aspects of Risk Theory". Springer, New York

Guo, X., Jarrow, R. and Zeng, Y. (2005). Modeling the recovery rate in a reduced form model. Preprint, Cornell Univ.

He, S. W., Wang, J. G. and Yan, J. A. (1992). *Semimartingale Theory and Stochastic Calculus*. Science Press, Beijing.

Ikeda, N. and Watanabe, S. (1962). On some relations between the harmonic measure and the Lévy measure for a certain class of Markov processes. *J. Math. Kyoto Univ.* **2**

79--95.

Internet Storm Center StormCast. Retrieved October 11, 2011, from http://www.dshield.org

Jacod, J. (1975). Multivariate point processes: Predictable projection, Radon--Nikodým derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31** 235--253.

Jeanblanc, M. and Valchev, S. (2005). Partial information and hazard process. *Int. J. Theor. Appl. Finance* **8** 807--838.

Joyce, James, (2008) "Bayes' Theorem", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), *Retrieved October 11, 2011,* from http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem

Kay, R. (1977), "Proportional Hazard Regression Models and the Analysis of Censored Survival Data" *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 26, No. 3 pp. 227-237 Blackwell Publishing.

Marti, K. (2008) "Computation of probabilities of survival/failure of technical, economic systems/structures by means of piecewise linearization of the performance function",Structural and Multidisciplinary Optimization, Vol35/3, Pp 225 - 244.

Newman, M.E.J., Strogatz, S.H., and Watts, D.J. (2001) "Random graphs with arbitrary degree distributions and their applications". Phys. Rev. E 64. paper 026118

Revuz, D. and Yor, M. (1999). *Continuous Martingale and Brownian Motion*, 3rd ed. Springer, Berlin.

Rogers, T. (1996) "*Type I and Type II Errors - Making Mistakes in the Justice System*" *Retrieved October 11, 2011, from* http://www.intuitor.com/statistics/T1T2Errors.html

Survival/Failure Time Analysis. Retrieved October 11, 2011,from http://www.statsoft.com/textbook/survival-failure-time-analysis

Therneau, T.; Sicks, J.; Bergstral, E. and Offord, J. (1994) "Expected Survival based on Hazard Rates", Technical Report No. 52, Mayo Foundation, *Retrieved October 11, 2011, from* http://mayoresearch.mayo.edu/biostat/upload/52.pdf

Thomson, I. (2007) *"Google warns of web malware epidemic" Retrieved October 11, 2011, from* http://www.securecomputing.net.au/News/81027,google-warns-of-

web-malware-epidemic.aspx

Weisstein, Eric W. "*Bayes' Theorem" Retrieved October 11, 2011* from MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/BayesTheorem.html

Woller, J (1996) "The Basics of Monte Carlo Simulations" *Retrieved October 11, 2011,* from http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html

Wright, C. (2007) "*The IT Regulatory and Standards Compliance Handbook: How to Survive Information Systems Audit and Assessments*" Syngress

Wright, C. & Zia, T. (2011)"A Quantitative Analysis into the Economics of Correcting Software Bugs" CISIS Spain

Wright, C. & Zia, T. (2011) "*Modeling System Audit as a Sequential test with Discovery as a Failure Time Endpoint*" in the proceedings of 2011 International Conference on Business Intelligence and Financial Engineering (ICBIFE 2011) December 12-13, 2011, Hong Kong

Zhu, H; Zhang, Y; Huo, Q & Greenwood, S; (2002)"Application of Hazard Analysis to Software Quality Modelling" Computer Software and Applications Conference, Annual International, p. 139, 26th Annual International Computer Software and Applications Conference