



Interested in learning more about security?

SANS Institute InfoSec Reading Room

This paper is from the SANS Institute Reading Room site. Reposting is not permitted without express written permission.

Data-Centric Quantitative Computer Security Risk Assessment

A quantitative risk assessment strategy is outlined with brief discussions of threat, risk categories and data classification. The differences between quantitative and qualitative assessments are specified with the conclusion that both methods have significant strengths and weaknesses. A quantitative method that spans both assessment types is then presented with rigorous analysis of impact of individual risk factors upon the overall risk to information. A method of easily organizing risk factors according to the quanti...

Copyright SANS Institute
Author Retains Full Rights

AD

An advertisement banner for Watchfire. On the left, there is a graphic of a globe and a login form with fields for "login" and "password". The text "Testing Web applications for vulnerabilities?" is written in white on a dark blue background. To the right is the Watchfire logo, which consists of a red flame icon and the word "watchfire" in a lowercase, sans-serif font.

Testing Web applications for vulnerabilities?

Data-Centric Quantitative Computer Security Risk Assessment

Brett Berger

August 20, 2003

GSEC Practical Version 1.4b (Option1)

© SANS Institute 2003, Author retains full rights

Data-Centric Quantitative Computer Security Risk Assessment

Brett Berger

August 20, 2003

Abstract

A quantitative risk assessment strategy is outlined with brief discussions of threat, risk categories and data classification. The differences between quantitative and qualitative assessments are specified with the conclusion that both methods have significant strengths and weaknesses. A quantitative method that spans both assessment types is then presented with rigorous analysis of impact of individual risk factors upon the overall risk to information. A method of easily organizing risk factors according to the quantitative method called a Risk Assessment Orgchart is explained and demonstrated. Careful manipulation of the method can make the analysis very sensitive to data classification and thus data-centric. A discussion on how to assign values to individual risk factors (scoring) should help users of the method be successful. Finally, a simple sample assessment is presented to tie all the analysis elements together and to further clarify the method.

Introduction

Current computer security theory and practice is divided into the categories of confidentiality, integrity, and availability. But within each of these areas, all policies, assessments, controls and procedures focus on securing critical resources whose compromise would cause harm to people, business advantage or national security. Critical computer resources can be divided into the subcategories of infrastructure and data. Infrastructure is defined as individual computers, purchased software (both for enterprise and local installation), and networks. The infrastructure is a crucial aspect of the data system, especially when planning for catastrophic failure, but consideration of securing infrastructure is beyond the scope of this discussion.

Data can be defined as computer customizations, user-created documents/files, and in-house-developed software; it is essentially, the value added to the computer system by the everyday efforts of users and administrators. The value of data routinely transcends the value of computers and infrastructure by many orders of magnitude. Some data are actually irreplaceable or may represent an intolerable degradation of competitive advantage or national security.

Risk to data is represented as the possibility of something adverse happening to the data. Risk management is the process of assessing risk, taking steps to reduce risk to an acceptable level and maintaining that level of risk.¹ A data-centric approach to managing risk skews the magnitude of perceived risk chiefly according to the classification level of the data. By taking this approach in a quantitative risk analysis method, some of the vagaries, complexities and subjectiveness of quantitative risk assessment can be diminished.

This paper will describe the difficulties with quantitative risk assessment. A discussion of threat with an eye towards minimizing the subjective nature of threat analysis will be presented. The importance of data classification and its application to a quantitative risk assessment method is at the core of the paper's purpose. The method will be detailed out with a mathematical description of risk factor averaging and a sample risk assessment scenario.

Steps To Risk Assessment

1. Develop a threat list
2. Compile a set of risk categories based upon the threat outline
3. Develop data classifications
4. Build a Risk Assessment "Orgchart"
5. Develop numerical scoring method
6. Conduct assessment pilot survey
7. Analyze pilot data; apply some subjective judgement
8. Adjust Risk Assessment Orgchart and numerical scoring method
9. Conduct final survey
10. Analyze risk data to obtain final overall information risk (OIR) scores

Threat

A threat is a potential violation of security. The word "potential" is key here in that an actual violation need not occur in order for a threat to be present. Threats can be divided into four broad classes: disclosure, deception, disruption and usurpation.² There are very complete lists of threats to be found in various security documents and it is advised that these be consulted when developing the threat list.^{3,4} Some quantitative risk assessment methods assign a rate of occurrence (RO) to a threat. In this, an attempt is made to make a guess at the percentage rate that a threat might occur. This is then extrapolated to a rate of occurrence during one year or an annualized rate of occurrence (ARO). This factor can then be combined with other cost factors such as single loss expectancy (SLE) to give annualized loss expectancy (ALE)⁵:

$$\text{SLE} \times \text{ARO} = \text{ALE}$$

The problem with this analysis is that it is only as good as the estimate of the rate of occurrence of a given threat. Single loss expectancy can also be difficult to estimate as loss does not always express itself in convenient or compatible units. Loss of brand confidence or losses due to schedule impact are examples of this. When dealing with the failure of hardware, statistical data about the potential for a given failure are compiled and analyzed. Tests can be accomplished that will give a good estimate of mean time between failures and this information can be extrapolated to predict potential future failures. Unfortunately, the information security world does not have equivalent statistical data available on the rate of occurrence of computer security threats. In fact, many businesses and

governments do not want to draw attention to successful attacks upon their systems for fear that other attackers will exploit similar vulnerabilities.⁶ In addition, statistics of past events break down completely when trying to predict a hostile act by a determined, intelligent, adversary. An attack could be completely original or it might target some previously unknown vulnerability. “For example, what was the probability that two airliners would strike the World Trade Center on September 11th, 2002? There were no precedents. What is the probability today?”⁷

Because of the uncertainty of predicting the likelihood of a given threat occurrence, the presented method sidesteps this completely by heavily weighting overall risk based upon the classification level of data being protected. All possible threats are evaluated but some subjectivity is possible based upon how the risk is calculated. Regardless of the likelihood of a threat, it is assumed that if a vulnerability exists the threat of the data being protected is proportional to the value of the data.

For the purposes of explaining this method the following abridged list of threats will be used:

- a) Access to Computer Left Logged In
- b) Theft of Entire Computer
- c) Alternate Boot Floppy
- d) Account Switching/Escalation of Privileges
- e) Unauthorized Mount of Network Shared File System
- f) War Dialing Modems

This is not in any way meant to be a complete list of threats but is merely a simplified list that will be used to demonstrate this risk analysis method.

Risk Categories

Once a complete list of threats is obtained they need to be categorized into risk categories. These categories will form the questions of the risk assessment survey. Each risk category must contain one or more threats. For example, the risk category of physical access encompasses the risk for the threat categories: Computer Left Logged In and Theft Of Entire Computer. This is an example list of risk categories and the threats above that they attempt to encompass.

- 1) Physical Access – Lack of physical access makes possible threats a, b and to a lesser extent, c.
- 2) Shared User Accounts – Makes threat d possible.
- 3) BIOS Password – Lack of a BIOS password makes possible threat c.
- 4) Network Shared File System – Makes threat e possible.
- 5) Modem Attached – Makes threat f possible.

Evaluation of risk categories through survey will point in the proper direction for risk mitigation. It is essential that this portion of a security program be given a high priority and adequate resources.

Data Classification

Since this method is data-centric, data classification takes a central role. There are numerous schemes that are used to classify data. The US Department Of Defense has very clear rules regarding the categorization of classified documents which are described in many security classifications guides. Unfortunately, the index to those guides is “for official use” only as are many of the guides. Obviously, needing an index for security classification guides implies that there are a lot of guides and a lot of rules for classifying documents. Anyone that has worked with the Department Of Defense has experienced a complex data classification scheme first hand. The table below is a simplified example of a military classification method.⁸

Classification	Definition	Examples
Top Secret	If disclosed could cause grave damage to national security	<ul style="list-style-type: none"> - Blueprints of new wartime weapons - Spy satellite information - Espionage data
Secret	If disclosed could cause serious damage to national security	<ul style="list-style-type: none"> - Troop deployment plans - Nuclear bomb targets
Confidential	If disclosed could seriously affect national security	<ul style="list-style-type: none"> - Technical specifications on older deployed weapons - Reserve troop mobilization plans
Sensitive but unclassified	If disclosed could cause serious damage	<ul style="list-style-type: none"> - Medical data - Answers to test scores
Unclassified	Data is not sensitive or classified	<ul style="list-style-type: none"> - Computer manual and warrantee information - Recruiting information

Controls on data in the different classifications of a military method can vary greatly yet the complex method of classification can cause problems in selecting the correct classification. Data classification is key to any information security program so a policy needs to be able to be known and followed by those creating and storing data. This means the audience of any security classification method is anyone who creates, distributes or modifies data; so administrative assistants, engineers, computer programmers, scientists, presenters, pitch creators, analysts and a host of other disciplines must understand and follow data security classification policies. A more simple classification method that could be used in a business is given below.

Classification	Definition	Examples
External	Security and handling requirements are given by another entity outside of company	- Data from a government program - Controlled information from a business partner
Private	If disclosed could cause serious harm to business	- Specifications or drawings of products - Business plans/strategies
Sensitive	If disclosed could cause moderate harm to business or personnel	- Salary information - Sales figures - Organization charts
Public	Data is not sensitive	- Company picnic plans - Sales literature

One of the most important aspects of any classification method is that users of the method know how to classify data they create or modify. Therefore, a detailed procedure for data classification is needed. This procedure should at least accomplish the following:

1. Identify information sources that need to be protected
2. Clarify the criteria for putting information sources into each classification category
3. Provide instructions for labeling each classification category so that its classification can be identified

In addition, policy should identify a data custodian or other classification authority who will be responsible for making difficult decisions regarding specific classification issues.

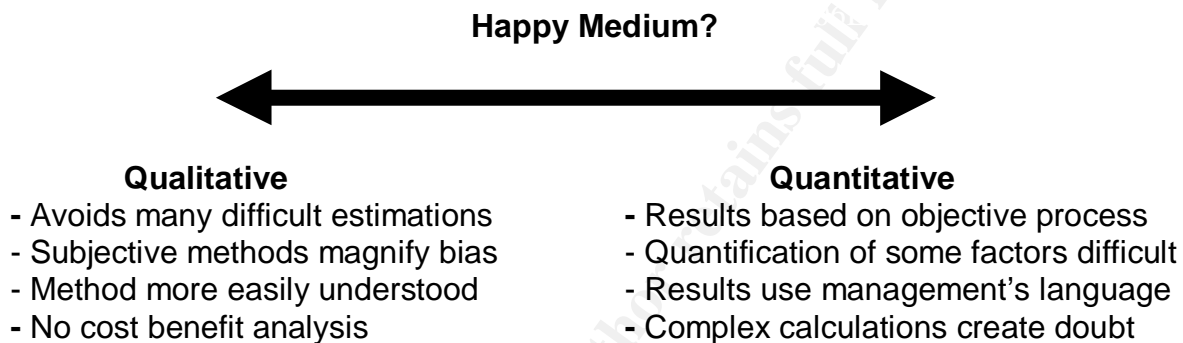
Quantitative Method

There are two techniques that are commonly discussed which are used to analyze risk. These are known as Quantitative Risk Assessment and Qualitative Risk Assessment. Qualitative risk assessment attempts to evaluate risk based upon relative risk levels evaluated in a subjective sense. Quantitative attempts to assign values to individual risks and to mathematically combine them in a way that demonstrates their relative or absolute effects. These effects can be generic numbers that only have bearing on each other or they can be converted into costs or time deficits. Both methods can be said to conform to the equation:

$$\text{Risk} = \text{Impact} \times \text{Likelihood} \times (\text{Threat} \times \text{Vulnerability})$$

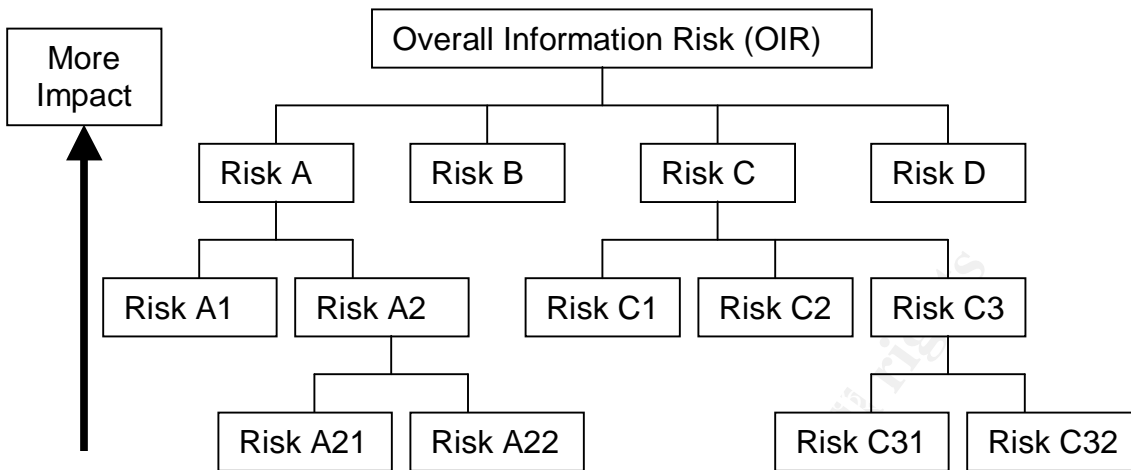
Both qualitative and quantitative risk assessments are “conducted using similar techniques. Their differences lie in the robustness and preciseness of the values used in the risk equation described previously, and in the resulting measurement

of risk.”⁹ In fact, it can be said that a quantitative analysis is just a refining or a separation of a qualitative analysis. It breaks down the qualitative issues into smaller factors that can have quantities ascribed to them. A further linkage between qualitative and quantitative analysis is created by the fact that many smaller qualitative decisions make up large portions of the quantitative analysis. For example, the impact of malicious modification of a company’s web site has no fixed cost associated that can account for embarrassment or loss of respect to a brand; therefore, a qualitative decision must be made to account for the effect. So it can be stated that no risk assessment strategy is completely qualitative or quantitative but they form a continuum.



As an IT security professional attempts a risk assessment, they must decide how much emphasis must be put on quantification. In a company that has a very precise budget process based on hard return-on-investment (ROI) numbers, a more exact quantitative method is called for. The danger in this can be in multiplying numbers together that have little fundamental or widely accepted validity. As risk is assessed and analyzed, the security professional should consider the audience and make sure they communicate risk factors in a way that will be understood and accepted. The method in this paper attempts to form a happy medium by quantifying risks in a relative way. The main working document resembles an organization chart and is known as a “Risk Assessment Orgchart”. This provides a way to organize the risk factors and to combine them while maintaining a rationale that is easily described to management. Instead of representing an individual in an organization, each box on this chart represents a risk factor such as use of a BIOS password. This factor would be assigned a value reflecting compliance with a standard. Non-compliance would receive a low value while complete compliance would receive a high mark. Factors located on the same horizontal level in the same branch are averaged together. Then that average is averaged with others on the branch above and so on until an overall information risk (OIR) value is obtained. This is done for each computer or computer system as desired.

Risk Assessment Orgchart



Describing risk in this way provides an easy method to emphasize the relative importance of different risk behavior and to also communicate this to management. Factors located at the bottom to have less of an effect than factors towards the top. Also, factors can be decreased in effect by adding other factors to their branch. The equations below show the relative effects of the factors in the Risk Assessment Orgchart above given the technique of averaging the factors on each horizontal level and carrying values up the tree.

Overall Information Risk:

$$OIR = \frac{A}{4} + \frac{B}{4} + \frac{C}{4} + \frac{D}{4} \quad (1)$$

Risk A breakdown:

$$A = \frac{A_1}{2} + \frac{A_2}{2} \quad (2)$$

$$A_2 = \frac{A_{21}}{2} + \frac{A_{22}}{2} \quad (3)$$

Risk C breakdown:

$$C = \frac{C_1}{3} + \frac{C_2}{3} + \frac{C_3}{3} \quad (4)$$

$$C_3 = \frac{C_{31}}{2} + \frac{C_{32}}{2} \quad (5)$$

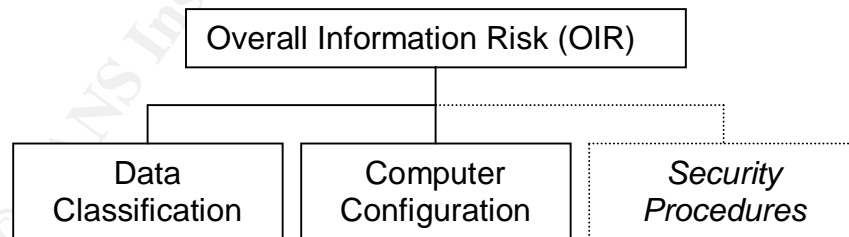
Now equations 2, 3, 4 and 5 can be substituted into equation 1 so that the relative effect of each individual term can be assessed. This assumes that each risk factor uses the same grading scale (such as 1 = low, 2 = medium, and 3= high).

$$OIR = \frac{A_1}{8} + \frac{A_{21}}{16} + \frac{A_{22}}{16} + \frac{B}{4} + \frac{C_1}{12} + \frac{C_2}{12} + \frac{C_{31}}{24} + \frac{C_{32}}{24} + \frac{D}{4} \quad (6)$$

Obviously, the terms with the lowest denominators would have the greatest impact on the final calculated OIR. This can be easily demonstrated by assigning a value of 1 to all variables calculating the value of OIR and then changing a value and recalculating. With all values set to 1, since OIR is an average, the value of OIR is 1. If the value of Risk B is changed to 3 then the value of OIR becomes 1.5. However, if the value of Risk C31 is changed to 3, then the value of OIR is only changed to ~1.08. This makes OIR over 5 times more sensitive to changes in Risk B than to those in Risk C31. This more rigorous analysis of the Risk Assessment Orgchart is not necessary in order to apply it to a given assessment situation. The chart is meant to simplify the prioritization of risks and does so graphically. The mathematics follows the rules set forth by the structure of the chart. These rules are: 1) factors nearer the top of the chart have more individual effect on the final total, and 2) factors alone on a horizontal line have more individual effect on the final total.

Making Assessment Data-centric

Using the Risk Assessment Orgchart, the analysis can be made more data-centric by putting the factor for data classification at the top of the chart. The OIR can then be tuned easily by adjusting the way the data classification number is mitigated by the other factors on its horizontal row.



The example above shows the two main factors of data classification and computer configuration. This means that the only way that the overall risk to highly classified data could be reduced would be to have a very low risk (low vulnerability) computer configuration. Using the 1, 2, 3 scale that was used in the example above, if data classification was seen as a 3 (very high importance) then the best the OIR could be with a computer configuration of 1 would be 2. If it was determined that this OIR was still too high, then security procedures like audits or

incident response could be put into place. This would make it possible to lower the OIR further to 1.7. Alternately, a computer with very low impact data (with a score of 1) would inherently have a lower OIR regardless of computer configuration. This method prioritizes resources towards solutions that affect the most important data.

Scoring

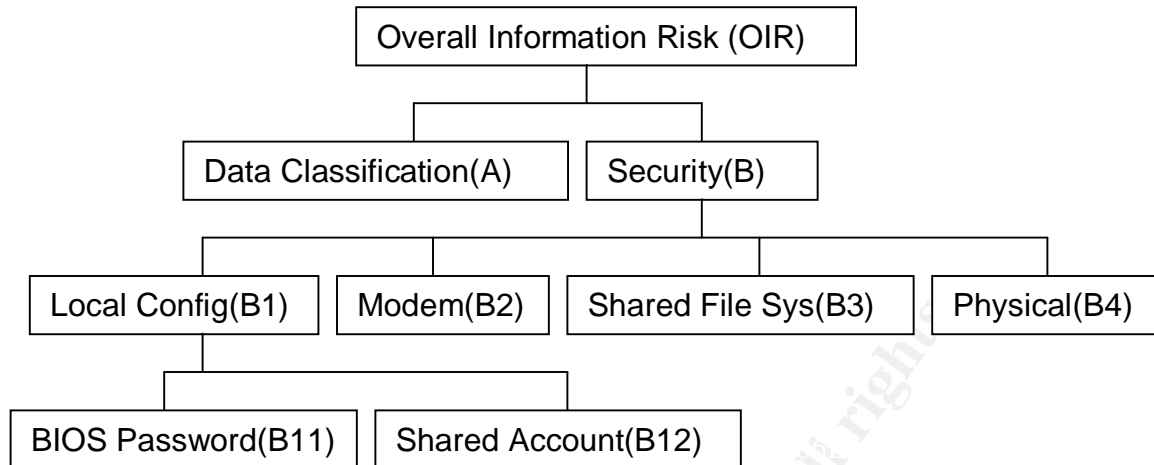
When scoring individual risk factors, care must be taken to ensure that the method used is both simple and consistent. The first step is to pick an appropriate scale. The simplest scale is the “1, 2, 3” scale already used in the examples above. Each factor is assigned a number from 1 to 3 that quantifies its compliance with a standard—the higher the number, the lower the compliance. Picking the scale might seem like a minor factor but it profoundly impacts the entire analysis. Larger scales make the OIR more sensitive to individual factors as the deviation from the average is potentially larger. It is best for survey purposes to keep the scale smaller as it is difficult to be consistent in describing a specific security factor with a large scale. Taking physical security as an example, a scale of 1 to 3 could be assigned as maximum, some and none. It would be more difficult with a scale of 1 to 10. The number 1 could still be assigned as “none” and 10 as “maximum” but what would differentiate between the numbers 3, 4, 5, 6 and 7? Generally, the smaller the scale, the easier it is to actually quantify the factors in a consistent manner.

Another good technique that adds more precise control to the security factors is the use of modifiers. In this technique, a base score is assigned to the factor and then modified by some desirable or undesirable characteristic of the factor. Returning to the example of physical security, a number of 2 might be assigned to an office with a door typically locked after business hours. A modifier of -1 could be used if the office is under close observation in a heavily traveled area during the day. A modifier of $+1$ could be assigned if key management is poor and there are many keys distributed or available in a common lock box.

When base scores and modifiers are selected it is important to remember that overall goal is to prioritize application of security mitigation. The basis of scores and modifiers should be carefully selected so that they accurately reflect security and provide a possible means for some kind of mitigation.

Sample Assessment

A simple example assessment will now be used to clarify the assessment method. Several candidate risk assessment factors were identified above. Now a Risk Assessment Orgchart can be constructed and base scores and modifiers can be assigned to the factors.



For this example, a 3 point scale will be used. Data classification will be assigned 3 for proprietary, 2 for sensitive and 1 for public with no modifications. The other factors will be assigned as below:

Physical Access

- 1 – Maximum : cipher lock with automatic logging of entries
- 2 – Some : closed door, observed area, card into building
- 3 – None

Modifier : -1 if video camera records area (but total can't go below 1)

Shared User Accounts

- 1 – None
- 2 – Some

Modifier : +1 if shared account is administrative or root

BIOS Password

- 2 – Password
- 3 – No password

Modifier : None

Network Shared File System

- 1 – None
- 2 – Shared with access control
- 3 – Shared with no controls

Modifier : None

Modem Attached

- 1 – None
- 2 – Attached but no phone line in area
- 3 – Attached with phone line

Modifier : -1 if no drivers installed (assumes no administrative access)

Note that there need not be a 1, 2 and 3 in every case as in BIOS password above. The assessor may not want to assign a 1 to this factor as the BIOS password can be reset to null by moving a jumper on the motherboard. This can be used as another technique to reflect situations in which security controls are difficult or impossible. These then become accepted risks. Also, modifiers should not be selected that actually reflect other factors; a modifier that assigns a -1 for a locked office would not be used for the BIOS password factor as there is already a factor for physical security.

The equation for the OIR is:

$$OIR = \frac{A}{2} + \frac{B_{11}}{16} + \frac{B_{12}}{16} + \frac{B_2}{8} + \frac{B_3}{8} + \frac{B_4}{8} \quad (7)$$

Now let's suppose that three computers are surveyed with the following results:

Computer 1 – Proprietary data, closed door with no video, shared user account with administrative access, no BIOS password, no shared file system, modem with phone line but no drivers.

Computer 2 – Proprietary data, card into building with video observation of computer, no shared accounts, BIOS password, no shared file systems, no modem attached.

Computer 3 – Public data, observed area, shared user account with administrative access, no BIOS password, shared file system with access control, modem with phone line.

Now apply equation 7 to each computer:

$$\text{Computer 1 } OIR = \frac{3}{2} + \frac{3}{16} + \frac{3}{16} + \frac{2}{8} + \frac{1}{8} + \frac{2}{8} = 2.50$$

$$\text{Computer 2 } OIR = \frac{3}{2} + \frac{2}{16} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = 2.06$$

$$\text{Computer 3 } OIR = \frac{1}{2} + \frac{3}{16} + \frac{3}{16} + \frac{3}{8} + \frac{2}{8} + \frac{2}{8} = 1.75$$

Note that computers 1 and 2 both share the same data classification level but their OIR numbers show the effects of the better security controls on computer 2. Computer 3 has poor security controls but its data classification level is also low making it a lower overall risk. If this method were used for more computers, a review could be made as to the relative security of the computers based upon the final OIR numbers. Risk categories, the Risk Assessment Orgchart and scoring could then be adjusted if necessary. Once the assessment survey is complete

and analysis accomplished, numerical thresholds should be set for OIR that will determine the corrective course of action. If the OIR equation is computerized and run against a survey database, blanket changes in individual scores can be made and OIR recalculated to see the effects of a given control. Future security efforts could be integrated into the chart such as continuity of business plans, security procedures, and other strategies.

Conclusions

Construction of a Risk Assessment Orgchart is an effective way to organize and combine the numerous complex factors associated with computer security risk assessment. It is often easy to breakdown a given factor (such as network configuration) into many sub-factors (type of connection, file sharing, services running, etc.). The Orgchart also facilitates explanation of complex security interactions to those that may not be familiar with security or computer terminology. The analysis that stems from the Orgchart could be captured in a computer program so that analysis could be more easily accomplished and solutions optimized for greatest overall impact on data security.

© SANS Institute 2003, Author retains full rights.

1. National Institute of Standards and Technology, "An Introduction To Computer Security: The NIST Handbook" (Special Publication 800-12), Oct. 1995, pg. 59.
URL: <http://csrc.nist.gov/publications/nistpubs/800-12/handbook.pdf>
2. Matt Bishop, Computer Security: Art and Science, Addison-Wesley, 2003, pg. 6-7.
3. CACI International Incorporated, "Computer Security Threats"
URL: <http://www.caci.com/business/ia/threats.html>
4. Chaddock, Mary M., "A Breakdown of SAN's Top Ten Threats", SANS GIAC Practical Repository, Oct. 2000, pg. 1-10.
URL: http://www.giac.org/practical/Mary_Chaddock_GSEC.pdf
5. Harris, Shon, CISSP All-In-One Certification Exam Guide, McGraw-Hill, 2002, pg. 79-82.
6. Moore, Andrew P., Ellison, Robert J., Linger, Richard C., "Attack Modeling for Information Security and Survivability", Technical Note CMU/SEI-2001-TN-001 Carnegie-Mellon, Mar. 2001, pg. 1.
URL: <http://www.sei.cmu.edu/pub/documents/01.reports/pdf/01tn001.pdf>
7. "Creating Secure Systems Through Attack Modeling", Amenaza Technologies Limited, Jun. 2003, pg. 3.
URL: http://www.amenaza.com/downloads/docs/5StepAttackTree_WP.pdf
8. Harris, Shon, CISSP All-In-One Certification Exam Guide, McGraw-Hill, 2002, pg. 100-103.
9. Smock, Robert, "Reducing Subjectivity In Qualitative Risk Assessments", SANS GSEC Practical Repository, Jun. 2002, pg. 2.
URL: http://www.giac.org/practical/robert_smock_GSEC.doc



Upcoming SANS Training

[Click Here for a full list of all Upcoming SANS Events by Location](#)

SANS London 2009	London, United Kingdom	Nov 28, 2009 - Dec 06, 2009	Live Event
SANS WhatWorks in Incident Detection Summit 2009	Washington, DC	Dec 09, 2009 - Dec 10, 2009	Live Event
SANS CDI East 2009	Washington, DC	Dec 11, 2009 - Dec 18, 2009	Live Event
SANS WhatWorks in Data Leakage Prevention and Encryption Summit 2010	New Orleans, LA	Jan 07, 2010 - Jan 12, 2010	Live Event
SANS Security East 2010	New Orleans, LA	Jan 10, 2010 - Jan 18, 2010	Live Event
SANS AppSec 2010 and WhatWorks in AppSec Summit	San Francisco, CA	Jan 29, 2010 - Feb 05, 2010	Live Event
SANS Phoenix 2010	Phoenix, AZ	Feb 14, 2010 - Feb 20, 2010	Live Event
SANS Tokyo 2010 Spring	Tokyo, Japan	Feb 15, 2010 - Feb 20, 2010	Live Event
SANS Geneva CISSP at HEG 2009 Autumn	OnlineSwitzerland	Nov 23, 2009 - Nov 28, 2009	Live Event
SANS OnDemand	Books & MP3s Only	Anytime	Self Paced