AN INTEGRATED APPROACH TOWARDS A PENETRATION TESTING FOR CYBERSPACES

Hanaa. M. Said, Mohamed Hamdy, Rania El Gohary and Abdel-Badeeh M. Salem

Ain Shams University
Faculty of Computing, Information Science
Abbassia,Cairo, Egypt

ABSTRACT: *The attack on a computer system with the intention of finding security weaknesses are becoming increasingly frequent and evermore sophisticated, potentially gaining access to it, its functionality and data. Organizations wishing to ensure security of their systems may look towards adopting appropriate tests to protect themselves against potential security breaches. One such test is to hire the services of penetration testers (or "pen-tester") to find vulnerabilities present in the case study for "Cairo Cleaning and Beautification Agency", and provide recommendations as to how best to mitigate such risks. By using series of the standards built on the application of data mining methods specifically decision tress model, Logistic regression, association rules model, Bayesian network for making reference penetration testers. This paper discusses the definition and role of the modern pen-tester and summaries current standards and professional qualifications. The paper further identifies issues arising from pen-testers; their motivation is to improve security.*

KEYWORDS: *Penetration testing, cyber security, vulnerability assessments decision tress model, Logistic regression, association rules model, Bayesian network,*

INTRODUCTION

With the ever growing technology, its advantages and disadvantages are increasing; Computer related crime is on the rise too. Technology is producing several negative impacts on society Internet hacking is worth noting. Cyber Security is the most serious issue around the world. Penetration testing is one of the oldest methods for assessing the security of a computer system. In the early 1970's, the Department of Defense used this method to demonstrate the security weaknesses in computer systems and to initiate the development of programs to create more secure systems. Penetration testing is increasingly used by organizations to assure the security of Information systems and services, so that security weaknesses can be fixed before they get exposed [1].

Commonly deployed security tests include firewalls, intrusion detection systems and anti-virus software, but security-conscious organizations go one step further by trying to understand the possible weaknesses of their deployed network, rather than just a paper-based analysis of the documented system. This can be achieved by employing a highly skilled security specialist to attempt to "break-in" to the network and related systems to determine what vulnerabilities are present. This service would typically include recommendations for mitigating the vulnerabilities and/or re-configuration to block these potential holes in the network. These security specialists are referred to as penetration testers or pen-testers.

A penetration test can therefore be defined as the process of systematically and actively testing a deployed network to determine what vulnerabilities may be present and to create a report with recommendations to mitigate or resolve these vulnerabilities, a penetration test is a method of evaluating the security of a computer system or network by simulating an attack from a malicious source [2].

Hacking is the act of breaking into another system with or without the owner’s knowledge. A penetration test is an in-depth information risk analysis exercise to assess the security of the systems from a hacker’s perspective. Penetration Testing and Web Application testing service simulate a hacker or an attacker like environment to conduct the exercise so as to match the hacker’s thought process. Penetration testing can be done both from the Internet and local area network depending on the placement and operational usage of the system such as: Web Application Penetration Test (Application discovery, Data Mining, Cryptography, Database Listener and Business Logic Testing).Data mining is a relatively new approach for intrusion detection. Data mining is defined as [Gsr98] “the semi-automatic discovery of patterns, associations, changes, anomalies, rules, and statistically significant structures and events in data”.

We Needs penetration testing, it is important for companies that want to guarantee the best product before distributing it. The results are used to find out security defects and to patch them before it will be too late. However companies' lack of time and resources, computer related crime is usually on the rise. Consequently penetration testers have reduced amount of resources. This brings testers to adopt automatic tools widely, as it is demonstrated by the continuous release of platforms finalized to automate this process, Discover gaps in compliance, Find holes now before somebody else does, Verify secure configurations, Test new technology and Report problems to management.
Collaborative processes oriented on large data sets are presented.

Security requirements are defined. Metrics for collaborative processes, for working with large data sets and for minimization of the effects generated by vulnerabilities resulting from the use of distributed applications are proposed.

This paper represents the analysis; studies securing one of the minor cyber space's which is the cyberspace of the Cairo Cleaning and Beautification Authority, Arab Republic of Egypt (www.ccba.gov.eg).it is one of the important cyberspaces in the frame of the e-government services mechanism. Also we test Cyber space security provided by e-government systems which enables the province using the strategies of **"Penetration testing model"** present and explain the testing and security of data in the e-government systems [3].

This paper mainly discusses the issue of Penetration testing, pen test, cyber security, vulnerability assessments and security of the e-Government Information Systems.
The paper is divided into the following sections: Section one is the completely introduction: a general review of Penetration testing, overview of security of cyberspace in e-governments. Section two describes the related work, Section three present's discussions of the different results of this study. Finally, section four closes this paper with a summary and conclusions.

RELATED WORK

In this section we will see a real-life example of penetration testing that involved the civilian government Agency FBI. The example is taken from an article in Computer world [2] those talks about the penetration tester Chris Goggans who has been working as a penetration tester since 1991. One of his latest exploitations was against the FBI. It only took him six hours to break into a crime database without permission. This is how he acted: he discovered a series of unpatched vulnerabilities in the civilian government agency's Web server, used a hole in the Web Server to pull down usernames and passwords that were reused on a host of enterprise systems, By so doing, he got Windows domain administrator privileges gaining full access to almost all Windows based system in the enterprise, including workstations used by police officers. Finally, remotely controlling them he found programs on their desktops that automatically connected the workstations to the FBI's crime database. This vulnerability could have been eliminated by clear separation of domains such as between the police network and the enterprise network.

In April 2007, Estonia [4] suffered a major cyber-attack, after which Estonia is contributing to securing cyber space worldwide. According to Joak AAVIKSOO, Minister of education and Research of Estonia, they analyzed weak point in their infrastructure. [4].As per their conclusions their law enforcements, border line do not hold in cyberspace [4], most of the infrastructure is not under single body and, 80% of web infrastructure is in private hands [4].

In 2008, Estonia [4] formulated a National Cyber Security Strategy. The objective of National Cyber Security Strategy is to ensure cyber security and help private sectors develop highly secured standards [4].In Malaysian primary schools, cyber bullying and hacking are the major occurring crimes [5].There is an Adaptive Information Security Model that was developed to lessen the gap between what we can do and control with ICT[6].There are five critical systems that ensure the highly secured and prospered network[6].Forty-one41 internet crimes have been analyzed[6].The analyses show that victims were missing in these five security tests[6] .A penetration test on internet service provider was conducted in Sweden[7].In Burma just before country's first national elections in twenty years, the internet was shutdown[1]. Offenders usually use public places to commit crimes which hides their identity and where there is no effective legislation. Internet gave birth to terrorist propaganda. Radicalization can be done using internet [4]. Mis configuration of websites causes search engines to penetrate into website and causes illegal access to data [8].Search engines need to obey some rules to disallow, some folders, files and images [8].

Halfond et al [5], [6] presented a technique for penetration testing which that involves static and dynamic analysis to increase the efficiency of the information gathering and response analysis phase. The author implemented static and dynamic analysis to improve penetration testing. To discover the input vector, the static analysis technique of automatic response that analyzes the dynamic analysis technique is used. The main objective of dynamic analysis is to find error while running the program. To test the effectiveness of these techniques, an experiment was conducted for static and dynamic analysis based penetration testing on nine web applications [7].

Halfond et al [8] developed Amnesia (Analysis for Monitoring and Neutralizing SQL Injection Attack).In this paper the author proposed a model based technique that combines the static and dynamic analyses. In this paper the tool first identifies hotspot, where SQL queries are issued to database engines. Non-deterministic finite automata are used at each hot spot to develop query model.

Xiong et al [9], [10] presented an approach of model driven framework that integrates the software development life cycle phases with penetration testing process, so vulnerability can be easily detected and testing can be done repeatedly by the expert personnel, To test the cost effectiveness, systematic and fully integrated into systematic and fully integrated into a security oriented software development life cycle, security experts are still required to maintain knowledge. In this paper the test cases are derived from models.

Stepien et al [6] presented an approach to penetration testing inherent to penetration testing of web application the approach consists inherent features of TTCN-3 languages. This paper derives the functional test cases and has taken an example of a malicious bank website. This paper described a message sequence diagram of a malicious bank website to show the XSS attacks. It generates the functional test cases.

Pietraszek et al [11],[12] presented an approach of Taint based Technique in which the author modified PHP interpreter to track taint information at the character level, Context sensitive analysis is used in this technique to reject SQL queries if an entrusted input has been used to create certain types of SQL tokens. The advantages of this approach are that they require modifications to the run time environment, which decreases the portability.

Arkin, Stender and McGraw (Arkin, B. et al 2005) [13] investigate the importance of the subject from the software pen-testers perspective, concentrating on where the role of the tester lies when flaws are assessed during software Development. Within the software development life cycle, Arkin et al. suggest without proper and timely Assessment, organizations "...often find that their software suffers from systemic faults both at the design level and in the implementation" (Arkin, B. et al, 2005).
The same can be said for the network security of organization; without proper and rigorous assessment, the network design of an organization will lead to unknown flaws inherent in the network implementation. The same can be said for the network security of organization; without proper and rigorous assessment, the network design of organization will lead to unknown flaws inherent in the network implementation.

Pierce, Jones and Warren (Pierce, J. et al, 2007) [14] in their paper provide a conceptual model and taxonomy for penetration testing and professional ethics. They describe how integrity of the professional pen tester may be achieved by "...avoiding conflicts of interest, the provision of false positives and false negatives, and finally do the legally binding testers of their ethical obligations in [their] contract" This is certainly noteworthy and should be expected of an individual working with potentially sensitive information; however, this appears more of a personal "ethical code of conduct" than something that can be enforced and assessed. Pierce et al (Pierce, J. et al, 2007) also discuss the provision by universities "...toward offering security testing courses". Additionally, in 2006 .

McRue(McRue, A., 2006)[15] Commented on the "first U.K. University to offer a dedicated degree course in hacking" This has certainly shown an emerging trend in the educational sector for penetration testing courses; however these tend to be degree classifications and not necessarily an industry recognized Certification standard.
Information gathering section of the penetration test is important for the penetration tester. Assessments are generally limited in time and resources. Therefore, it is critical to identify points that will be most likely vulnerable, and to focus on them. Even the best tools are useless if not used appropriately and in the right place and time. That's the reason why experienced testers invest an important amount of time in information gathering [15].

PROPOSALS

We have developed a framework, MADAM ID (for Mining Audit Data for Automated Models for Intrusion Detection); we apply a penetration for evaluating the security of a computer System or network by simulating an attack from a malicious source.

We Build Penetration test model for evaluating the security state of a system or network by simulating an attack from a malicious source. This process involves identification and exploitation of vulnerabilities in real world scenario which may exist in the systems due to improper configuration, known or unknown weaknesses in hardware or software systems, operational weaknesses or loopholes in deployed safeguards. Attacks on the computer infrastructures are becoming an increasingly serious problem. Computer security is defined as the protection of computing systems against threats to confidentiality, integrity, Objectivity, timeless and availability.

We will use how strategy of inferring and analyzing the data, searching for them in the cyberspace by one of the technology tools (data mining) this paper shows the vision of the insurance. and the general arrangement for extracting the required data, through the cyberspace, enabling fighting terrorism to limit the harms in advance by making the relief arrangements from the view of comprehensive security and through the analysis of the results for the data survey. as it depends on using the models of test to assess the extent of the correctness and safety of the data , identifying the methods of research identifying the standards of test that can exceed the limitations of the available data , such as using the proposed model in the Figure 1" as the follows.

To test the extent of the data correctness for the cyberspace, and that the infrastructure of the propped model of cyberspace for the Cairo Cleaning and Beautification Authority, a model will be built in steps represented in 2 states as follows
•The first stage (hybrid of auto regression and decision trees)
•The second (Bayesian network and association rules) to enable the decision maker to know interact with the features of the value traits. And the data extraction tools will be adapted with data mining.

Penetration testing was among the first activities performed when security concerns were raised many years ago [3]. The basic process used in penetration testing is simple: attempt to compromise the security of the mechanism undergoing the test. In earlier years, computer networked operating systems, with their access control mechanism, were the most suitable components for penetration testing, because O.S. is the core component of the machine, so it is more exposed to security threats [3].

The major advances in penetration testing technologies have been made by the IT Security practitioners, while indeed software tools were written with the IT Security Professionals as its principal users [3]. The earliest penetration testing processes were highly and manually intensive, while later automatic processes started to be clearly utilized for cost reduction [3].We need to determine how the attacker is most likely to go about attacking a network or an application. Locate areas of weakness in network or application defenses ,Determine how an attacker could exploit weaknesses , Locate assets that could be accessed, altered, or destroyed , Determine whether the attack was detected , Determine what the attack footprint looks like and Make recommendations.

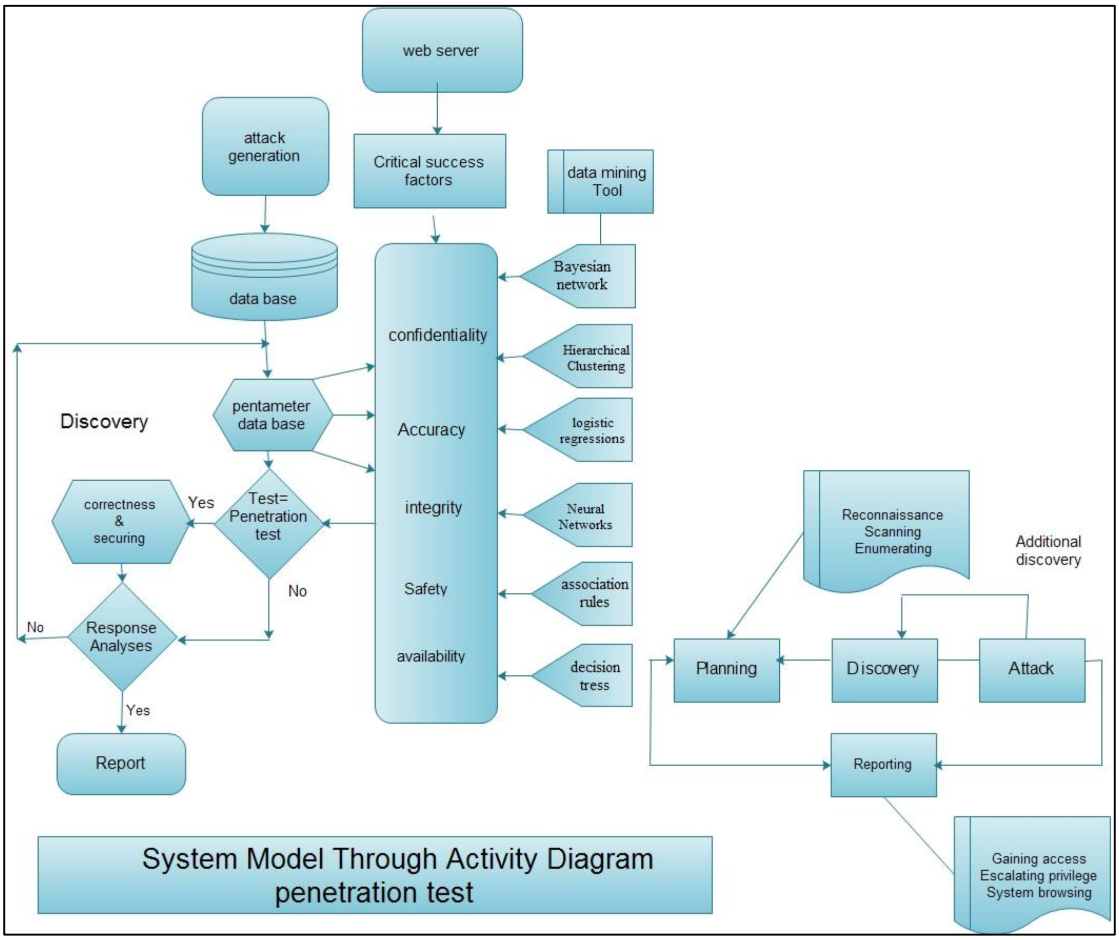


Figure 1: Proposals of penetration test model

Security aspects

Finding the factors that have an effect on the rate of data security , information security and safety of the government service rendered , which is identifying the difficulties that hinder the execution is one of the main cases that encounters the government officials to achieve the accuracy of the data and the range of its safety, in addition to watching the levels of securing them through data collection, and analysis and recognition of the major obstacles. This led to decline of the government performance, level.

Table 1: Titrated Test how Secure the Minor Cyberspace Key

1-Availability	The continuous operation of the system that needs different levels of availability.
2-Timeliness	The available information for the user must be in the suitable time not in late time with which it will be difficult to utilize them.
3-Feedback Value	It means the ability of the information to help decision-maker to make the procedures of prediction and correcting them,.
4-security (Representational Faithfulness)	The information must be in safety and secured form free of any intended tricking- using security policy for the information (protection).
5- Objectivity	Being far from the personal estimations and depending on the trusted evidences.
6- Verifiability	High degree of agreement by using similar measuring methods, and attaining the same results with the same accuracy.
7- Integrity	Using the policy to guarantee the safety of the information when conducting the necessary measurements to ensure the safety of all data.
8- Backup	Save data using backup by using the available policy and means of storage.
9- Recovery	Recover data that has been lost for any reason by using the available means of storage.
10- Agree	Achievement of general and complete satisfaction for operation of the cyberspace.

3.1. Testing and verification of availability (Availability) & (Timeliness) Using association rules

The terminology of availability is also used in communications. It means the degree at which the cyberspace works in abiding state. It usually represents fraction like, 9998 for the simple available A, It is the percentage of the expected value in the state of working and the stopping time (time down,)

$$A = \frac{E[UPtim]}{E[UPtim] + E[Downtim]}$$

As follows: x (t) if you define a function Status

1 = sys Function at Tim t

$X(T) = \begin{cases}1\\0\end{cases}$ 0 = other wise

The savings is as:

$$A\ (t)\ =\ Pr|x(t) = 1|$$

$$\langle \mathrm{E}|\mathrm{x(t)}| = \mathrm{X.Pr}|\mathrm{x(t)} = 1| \qquad \mathrm{t} > 0\rangle$$

And must know the average availability in the on-line and real-time as a static and random, the average saving is expressed in the formula

$$A_{\mathrm{C}=\frac{1}{\mathrm{C}}}\int_0^C \mathrm{A(t)dt}, \qquad \mathrm{c} > 0$$

This represents the forbidden availability (steady state) as

$$A = \lim_{t\to\infty} A(t)$$

The average of availability in the period on the real time line is considered field and random. So the average of availability can be expressed in the following figure

$$A_{\infty = \lim\limits_{\mathrm{C}\to\infty} \mathrm{A_C}} = \lim_{\mathrm{C}\to\infty} \frac{1}{\mathrm{C}} \int_0^{\mathrm{C}} \mathrm{A(t)dt} \qquad \mathrm{c} > 0$$

It is important to follow the state of the protection arrangements for periodical control, to ensure the remaining cyberspace is highly protected with all its effectiveness, we can eliminate the weakness points and identify the strengths taking into account the arrangements for the protection and applying the previous equations to the sample of the data. As we find the availability which is achieved by the rate of 96% we can find the timeliness which achieve the rate of 94%, the systems can reveal, prevent espionage, make controls and applies the reactions for the suitable deeds.

Testing and verification of security Cyberspace (Security) & (Integrity) using A simple decision tree model (Chaid1 Algorithms)

Using a simple decision tree model Chaid algorithm security rating for classifying the data including the fields of entry or the variables, Decision tree is the structure of the tree on the shape of tree branches that represents sets of decisions. These decisions generate rules for classification of the set of the data. It includes limited forms for the branches of the branches, which includes the decision of classification, or decline, it includes the space of the automatic discovery of the mistakes.

Table 2: Field name Description

Field name	Description,
Input variable	$x_1 ,x_2,x_3,x_4 x_5$& Y
Security rating	Security rating : 0= attack 1= security
Data risk	Number of test range of security 1=< 88.00 , 0>88.00

Coding Input data (0, 1) and the independent variables [x_(1 ,x_2,x_3,x_4 x_5) & Y]

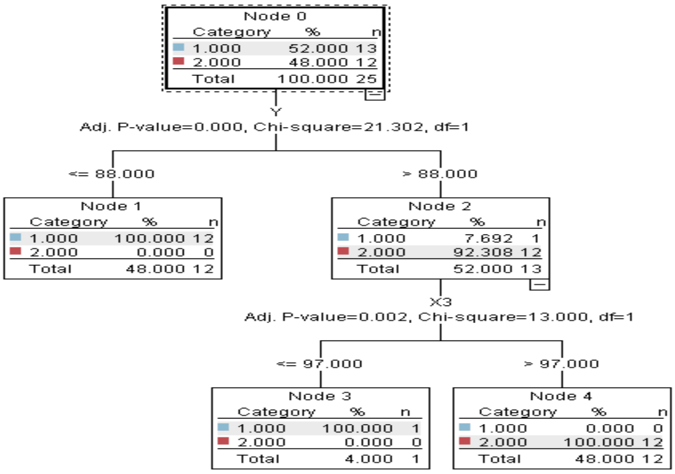


Figure 2: A simple decision tree model

When looking to figure 2 the upper part of the first node of the tree, It gives us a summary of all records in the set of the data. We can find that the rate of 48% represents 12 scores only in the cases of the set for the data sample representing the secured data that has protection, with the rate of 52%,and represents 13% scores for the risk and not secured, it needs to improve the performance for protection and security, it is exactly the first part of the analysis , so let us see if each tree can give us any evidence to what are the factors that may be responsible We can see that the first division is according to the level of the input data. So it will be possible to assign or to determine the scores on terms that the income level in allow class to (2node) it is not surprising to see that this classification contains the highest rate of the non secured data/ it is a clear indicator of the data of this class, to contain high risks and needs a solution thus the rate of 52% for the data of this class represents a risk actually, if not supposedly ,consequently, the prediction model cannot in practice respond but that the model must be good and allow us to expect and to respond more likely to each score based on the available data by the same way if we looked to, the data which the node 2 refers to. We could find that the vast majority (92308%) appears unsecured represents a risk and needs to set a new mechanism security. Therefore, we can improve the standards of security in this set of data to reduce the risk, so we learned that each score is an indicator of this model. We will identify the weaknesses by assigning certain node.

Assigning the new predictions either good or bad, depending on the most common response of this node, this process is known for assigning the predictions of the individual scores as it is the objective, by recording the same scores used for assessing the model. We can assess the extent of accuracy for the training data that know the result, this model is used for the tree of decisions that classifies the scores and expects the response by using series of rules for taking the decision.

Table3: Case processing summary

Data security		***N***	**Marginal Percentage**
Wor	Y	*13*	52.0%
	No	*12*	48.0%
variable	X1	*6*	24.0%
	X2	*11*	44.0%
	X3	*8*	32.0%
Valid		*25*	100.0%
Missing		*0*	
total		*25*	
subpopulation		*25(a)*	
a. The dependent variable has only one value observed in 25(100.0%)			

Preliminary calculations for Verifiability control using Logistic Regression Algorithms

In the logistic decline, and that every field accrues two s for each class or the value of original field except for the last class, it is known as the reference class for each score, the value of the field in advance to the class of the attendee will be assigned at1.0 and all the other fields will be assigned stemmed from the field 0, 0.Thesefields which are derived are called the false fields; they are named **the false recoding**

For example the following data as x is a symbolic field with the possible value of C, B, A

Record #	X	X1’	X2’
1	B	0	1

In this data the accrual of the original group field, x in the two fields derivate from $X1', x2' \& X1, '$it is the indicator of the class A, and X2 is the indicator of class B the last class IS C is the reference class. the scores which belong to this class equally x1,x2, the group of 0,0 model it found the threat each h score passed through the model of two sides of the logistic decline.

The expected value and reliability are calculated.

The expected value " the probably z =1 the value of scores is calculated as follows:

Table 4: variable and Parameter Description

The following notation is used through this state :	
n	The number of observed cases
p	The number of parameters
Y	n x 1 vector with element y .the observed value of the jth case of the dichotomous dependent variable.
X	n x p matrix with element x_{ij}, the observed value of the ith case of the PARAMETER

B	PX 1 vector with element Bi , the coefficient for the jith parameter
W	n x 1 vector with element w, the weight for the jth case
L	Likelihood function
L	Log –Likelihood function
I	Information matrix

- Possibility of a ["flexible value feedback " &"Objectivity "&" accuracy verifiability"]Confidence

$$\hat{\pi}_i = \frac{\exp(\hat{\eta}_i)}{1 + \exp(\hat{\eta}_i)}$$

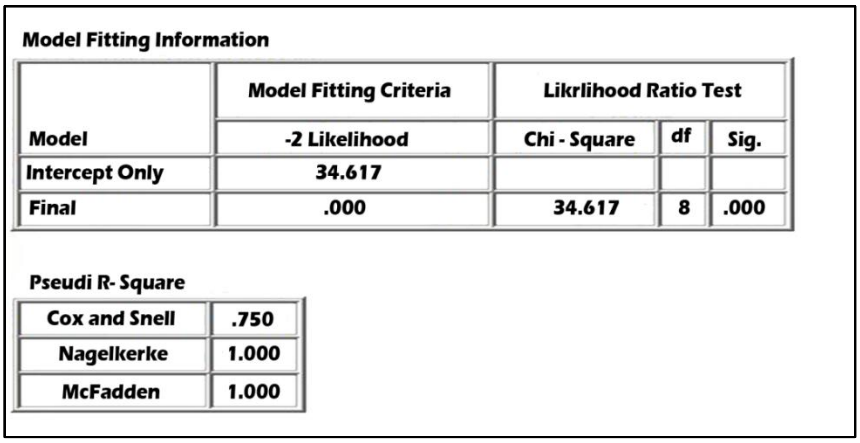
$$\hat{\eta}_i = X_i^{'}\hat{\beta}$$

$IF\hat{\pi} > 0.5,$ The predicted value is 1: otherwise, the predicted value is 0.Confidence
For records with predicted value of y = 1, the confidence value is $\hat{\pi}$, for records with a predicted value of y = 0, the confidence value is (1- $\hat{\pi}$)
$\sum_{i=1}^{n} w_i \,(y_{i-}\pi_i)x_{ij} = 0$, for the parameter
Where $x_{i0=1}\, for i = 1, \ldots\ldots, n.$

Can we rely on the use of Newton Rafsson algorithm for Verifiability? No, we cannot rely on. Because we cannot depend on the proximity,
1. The absolute difference estimated between the frequencies.
2. The difference percent in the probability.
3. Between the successive frequencies

The maximum limit for the number of the set frequencies to identify the accuracy of the data (verifiability) through the frequency is the smallest ranging from 10-8 for all cases. The probable occurrence of the frequency is very close to zero. stopping of the frequency or the message, expectation of all the values either 1-zero will be issued , getting the maximum limit for the estimates of probabilities and the matrix of the variation is the proxy estimated since before , The reverse of the first information matrix.

As Where: Remark: this example shows how to get the total percent as evidence in designing the model. It can be accuracy. In some cases it may be detective. The original; zero model was 72.6% accuracy in general, Meanwhile the final model and the expectation to have full accuracy of 79.1 %, but as we saw, among the accuracy of the predictions the actual individualism is of different class to a large extent,
Table 5: Model Fitting Information



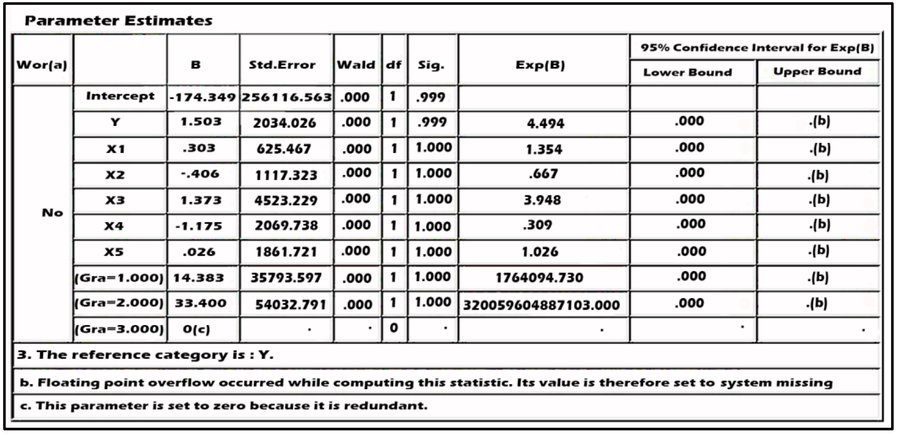
Model Fitting Information

Model	Model Fitting Criteria	Likrlihood Ratio Test		
	-2 Likelihood	Chi - Square	df	Sig.
Intercept Only	34.617			
Final	.000	34.617	8	.000

Pseudi R- Square

Cox and Snell	.750
Nagelkerke	1.000
McFadden	1.000

Table 6: expected values for parameter estimates



Parameter Estimates

Wor(a)		B	Std.Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
No	Intercept	-174.349	256116.563	.000	1	.999			
	Y	1.503	2034.026	.000	1	.999	4.494	.000	.(b)
	X1	.303	625.467	.000	1	1.000	1.354	.000	.(b)
	X2	-.406	1117.323	.000	1	1.000	.667	.000	.(b)
	X3	1.373	4523.229	.000	1	1.000	3.948	.000	.(b)
	X4	-1.175	2069.738	.000	1	1.000	.309	.000	.(b)
	X5	.026	1861.721	.000	1	1.000	1.026	.000	.(b)
	(Gra=1.000)	14.383	35793.597	.000	1	1.000	1764094.730	.000	.(b)
	(Gra=2.000)	33.400	54032.791	.000	1	1.000	320059604887103.000	.000	.(b)
	(Gra=3.000)	0(c)	.	.	0	.	.	.	.

3. The reference category is : Y.

b. Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing

c. This parameter is set to zero because it is redundant.

The expected value for the scores is the output class with the biggest value,
Logarithmic prospects

$$r_{ij} = \log\left(\frac{\pi_{ij}}{\pi_{ij}}\right) = x_i'\beta_j$$

$for j = 1, \dots, j - 1. The$ log $itfor$reference categoryj, r, I, is 1.0.

$$\hat{\pi}_{ij} = \frac{\exp(r_{ij'})}{1+\sum_{k=1}^{i-1}\exp(r_{ij'})} = \frac{\exp(x_i'\beta_j)}{1+\sum_{k=1}^{i-1}\exp(x_i'\beta_k)}$$

The Results: it is probably to expect the occurrence of any changes to this classification on terms to calculate the probability of the occurrence for any of the previous standards through the calculation of exp) as shown in the table 6 as we can find the flows

1- The least probable occurrence is 0.390 we find it in the variable of x4&and the highest Probable occurrence is 4.494 we can find it in the variable y

2- This model is considered as standard model, as we found sig =0.00 asit appears in Table 5, the probability for all classes' j in similar data can be calculated,

Testing the flexibility (objectivity feedback value) by using the Bayesian network model

The Bayesian network model provides a brief way to describe the joint probably distribution for certain group of the random variables (x1 x2x3 x4 x5, y) Conditional probabilities for each variable show that the values of the data are divided into proportion to the origin of the node and the similar node , as an alternative for the analysis process , you can use them as an indicator of the evaluative figure to compare the accuracy of the model that was predicted based on the figure, and is used as an indicator of the process of prediction and the probability of testing the flexibility (feedback value and objectivity.)

a. As it is the case with the node of probability. The figure shows that each model produces similar results. but the model(the retrained model is used for comparison of the data for these standards) of flexibility and objectivity) it is little better because it contains a higher level than that of the confidence in its prospects ,

b. We can suppose that v forms a set of the classy random variables and that v=g e the figure will be continual ring so we can find the direction of the node v and a set of the directive edges

c. The model of Bayesian network consists of the figure G besides the table of THE conditional probability for each specific node of the original node value

d. Thus it will be possible to calculate the joint probable distribution of the random variables in the shape OF V AS producer for the conditional probabilities of all the nodes , due to the value of each node ,

e. A set of variables is given in the shape of V and a sample of the adjacent data as in the table 7 and figure 9.that show the presentation of the task for installing the Bayesian network model. It is called for identifying the edges of the figure the structural building of the e variables.

Table 7: The following notation is used throughout this algorithm description

G	A directed acyclic graph
D	A Dataset
$X_{1,}X_2, X_{3,}X_4, X_5, Y$	Target variable

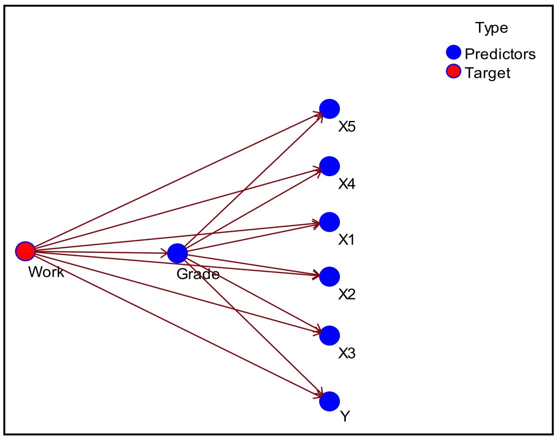
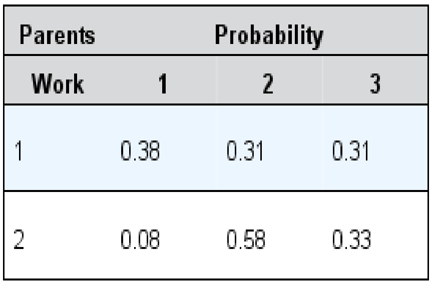


Figure 9: Bayesian Network

Table 8: conditional probabilities of grade



Parents	Probability		
Work	1	2	3
1	0.38	0.31	0.31
2	0.08	0.58	0.33

- The probability for the occurrence of the flexibility is none ,38% in the first case
- The probability of the occurrence of the non objectivity ,58% in the first case

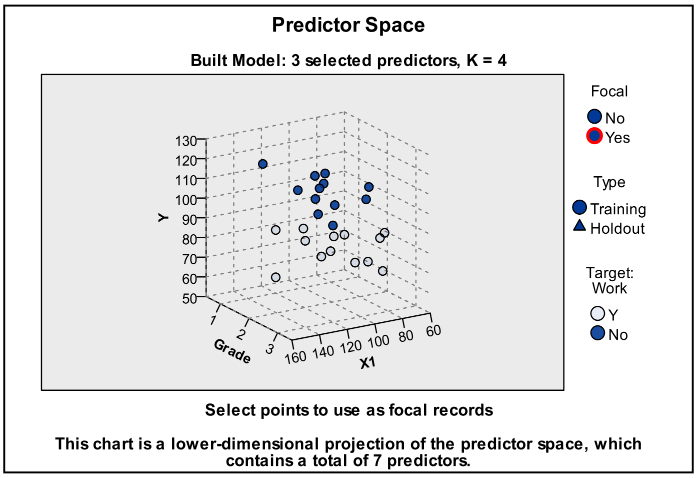


Figure 10: predictor space

Figure 10 shows the probability and expectation of the 3 dimensions and the distance that contains "v". random probabilities as shown in dark blue for non probability of the occurrence for which the measurement is run and the light blue color which achieves the probability of the occurrence for the event , table 10 shows the probabilities and the expectations of the 3 dimensions and the distances .

Table 10: possibilities and expectation triple dimensions and distances

Point Type	**X**	**Y**	**V4**
order set	0	0.7778	Grade
scale	0	6.2222	Bias
scale	0	5.4444	Y
scale	0	4.6667	X3
scale	0	3.8889	X5
scale	0	3.1111	X1
scale	0	2.3333	X2
scale	0	1.5556	X4
scale	1	5.25	Bias

scale	1	3.5	Hidden layer activation: Hyperbolic tangent Output layer activation: Soft max
scale	1	1.75	Hidden layer activation: Hyperbolic tangent Output layer activation: Soft max
set	2	3.5	Work

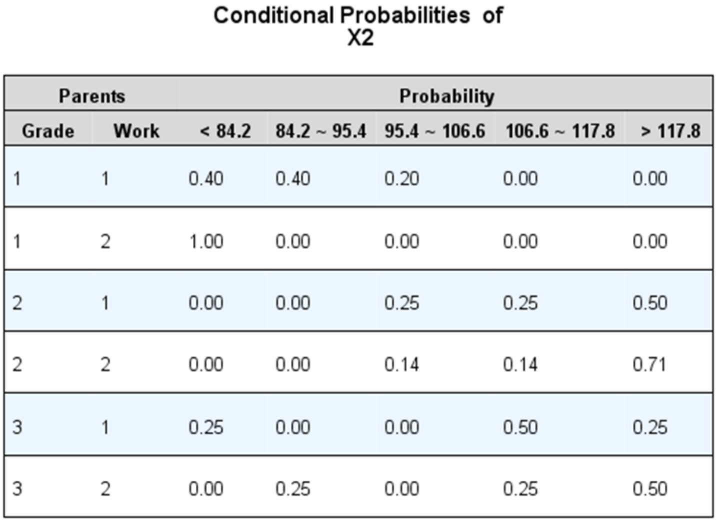
$$I(X_{i,}X_j) = \sum_{X_{i,}X_j} \mathrm{P_r}(X_{i,}X_j)\mathrm{Log}\left(\frac{\mathrm{P_r}(X_{i,}X_j)}{\mathrm{P_r}(X_i)\mathrm{P_r}(X_j)}\right)$$

We start by replacing the information exchanged between the two predictors with the information exchanged between the two conditional predictions given the goal.

$$I(X_{i,}X_j|y) = \sum_{X_{i,}X_j} \mathrm{P_r}(X_{i,}X_{j,},y_k)\mathrm{Log}\left(\frac{\mathrm{P_r}(X_{i,}X_j|y_k)}{\mathrm{P_r}(X_i|y_k)\mathrm{P_r}(X_j|y_k)}\right)$$

We can build the network by using the calculation between each pair of the variables by using the algorithms for building as the maximum level as it starts with the extended tree with the non existence of the edges and the signs of the random variable as an approach , then it will be found the variable of the non controller , whose weight with one of the observable is the maximum limit then this variable will coincide with this variable and adds its edge to the tree. this process is repeated till the sign is put on all the variables with setting standards for the measurement to the indicator of the assessment , so we can find the picture or the shape of the probabilities of the variable x2 which is shown clearly in table 11 as the below. We can find in the table 11 cases for the probabilities as the following " in the first case the highest rate of probability of the occurrence of the objectivity is.40% , the second case with 1,% , the third case is 50% , the fourth case is,71% , the fifth case is 50% and the sixth case is 50%:

Table 11: Photos of the possibilities of the variable



Conditional Probabilities of X2

Parents		Probability				
Grade	Work	< 84.2	84.2 ~ 95.4	95.4 ~ 106.6	106.6 ~ 117.8	> 117.8
1	1	0.40	0.40	0.20	0.00	0.00
1	2	1.00	0.00	0.00	0.00	0.00
2	1	0.00	0.00	0.25	0.25	0.50
2	2	0.00	0.00	0.14	0.14	0.71
3	1	0.25	0.00	0.00	0.50	0.25
3	2	0.00	0.25	0.00	0.25	0.50

The "figure" no 11 And no 12 shows the probabilities and the expectations that can build the net by using the non directive tree for the result into one of the outputs, then the node of the root is chosen and the direction from all the edges to be outside it for each variable of the random variables is identified, for the several predictions and probabilities that can be written or recorded to become indicator of the measurement,

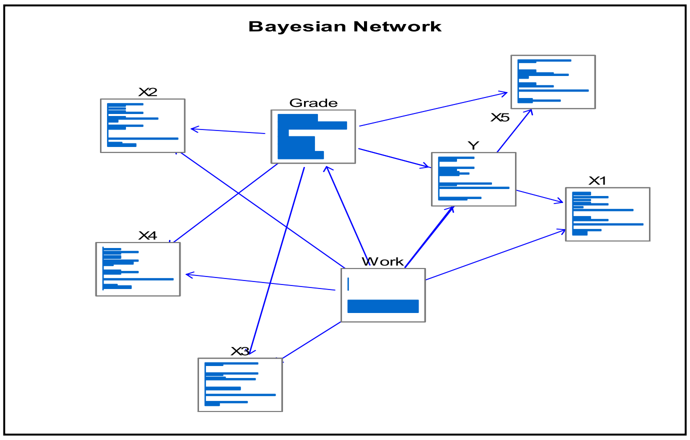
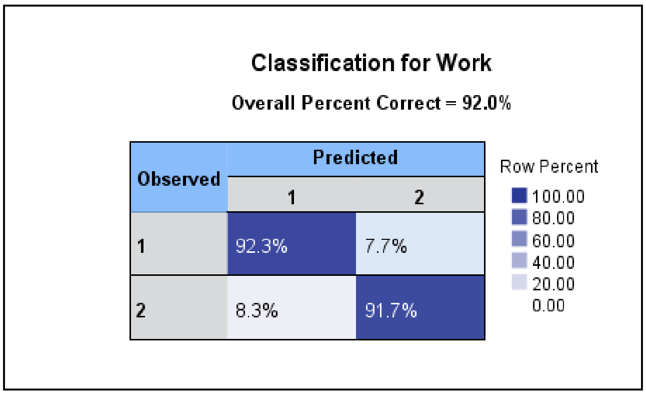


Figure 11 illustrates the possibilities and the network can be constructed expectation


Classification for Work

Overall Percent Correct = 92.0%

Observed	Predicted	
	1	2
1	92.3%	7.7%
2	8.3%	91.7%

Row Percent

- ■ 100.00
- ■ 80.00
- ■ 60.00
- ■ 40.00
- ■ 20.00
- 0.00

Figure 12 shows the variables and the average rating correct proportions

In this tactic we find several results for the occurrence that appear better largely. Several if the degree of decline is < 0.998, or better, it is an encouraging, to ensure the improvement in the model of analysis the classes then operating this technology to enable making the comparison between 2 or more models of the same type, the analytical output. this indicates that the function of the Bayesian network can properly predict the rate of 97,85% from the cases that are still very good .

EVALUATION (ANALYSIS OF THE RESULTS)

Finally we can find in this paper that the cyberspace needs to be improved, and its sufficiency needs to be enhanced and necessary arrangements be taken to raise the efficiency of security. As the data is exposed to the occurrence of violations at the rate of 92,308& we can find that (timeliness& integrity &objectivity availability were achieved by high rate of7%91, 72, 6%92, 3% and we find that the (feedback value was achieved at the rate of medium. We conclude that the maximum number of the frequencies identified to set the accuracy of data the less probability (verifiability) of occurrence ,309, we find it in the variable x4& and the highest probability for * verifiability) occurrence is 4.494 as indicated clear in table 12.

Table 12: Report of the proposed procedure modeling

The factors	**Achieve percent**	**Percent did not achieve**
Availability	96 %	Limited
Timeliness	94 %	Limited
Security(Representational Faithfulness)	Limited and Poor	92.308 ٪
Feedback Value	48%	52 ٪
Objectivity	91.7 %	0.58 ٪
Verifiability	72.6 %	Limited
Integrity	92.3 %	Limited
puck up & Recovery & Agree	• We did not use these means and measure the Achieved and tested. •a future and we will supplement and follow-up research in this area	
Observation	We find that the Security needs to be improved and upgrade	

However, the above mentioned table which shows the outputs of the previous table as the function of assessment was of several values. Each case can be diagnosed correctly according to the standardized shape. However, in practice, it was not preferred to see100% accuracy, but you can use the assistant analysis in identification if the model of accurate and acceptable application of the cyberspace actually, or that there is no other type of function or of sins nor linear that can apply , however with the set of different data , it is possible for the results to be easily different. Thus, it is always worthy of trial, with full set of choices,

We also find that it was not possible for us to use means of measurement for the extent if achieving some standards and testing them among those of the above mentioned in which is using the policy of the pickup related with maintaining the data by using the policy of the relief coping by using means of secured storing using the policy of (recovery) related to retrieval of data that were lost for any reason by using the means of secured storing. The policy of the acceptable using (agree) related to achievement of satisfaction and the completeness when operating the cyberspace,

Penetration test also can be used as an important and useful indicator in security measurement. E-government needs to new technology in order to enable us to follow the new challenges that may face the cyberspace, in addition to identification of security threats.

❖ **Finally there are many reasons that make cyberspace needs for penetration test, these reasons can be summarized as follows:**

1. Determination of the effectiveness of the security controls and adjusting their appropriate locations.
2. Determination the points of weakness and strength of cyberspace security system.
3. Determination the sufficiency of the current controls in the cyberspace security system.
4. Determination of the threats against the organization's information.

CONCLUSIONS

Security is an important issue for the future of the cyberspace. Attacks against the nation’s computer infrastructures are becoming an increasingly serious problem. Over the past few years, there has been tremendous increase in the cyber threats due to penetration of new technologies within the global economy as it involves heavy usage/dependency of the Internet to carry out businesses for personal/business/governmental sectors. Computer security is defined as the protection of computing systems against threats against confidentiality, integrity, and availability.

Penetration testing is one of the oldest methods for assessing the security of a computer system. In the early 1970's, the Department of Defense used this method to demonstrate the security weaknesses in computer systems and to initiate the development of programs to create more secure systems.

This paper discusses the key issues related to penetration test. Penetration test examines the information from the client and the government sides. Penetration test for the security control system can be achieved through data mining analysis.

Data mining techniques a new approach used in penetration test model for intrusion detection. The penetration test aims to discover points of weakness in security system and infrastructure of the cyberspace. Also, this paper investigates and evaluates the decision tree data mining techniques as an intrusion detection mechanism. It also suggests a number of emerging applications for cyberspace affiliated to General Authority for Cleanliness and beautification of Cairo.

Finally the penetration test technique is very useful technique for building strategies for measuring the extent of securing data in order to improve the management performance, through the filtration of data. Also the suggested technique could become an important tool for the government and intelligence agencies in the decision-making and monitoring potential international terrorist threats.

In the future we will complete and follow up the research in this field through using search in data to be an active way in decision making. It is expected that there will be several challenges related to operation and development of cyberspace system. The penetration test will be an effective tool that will help in testing the security of the data.

REFERENCES

[1] J.Long, “Google Hacking for Penetration Testers”, e-book.
[2] “Six hours to hack the FBI (and other pen-testing adventures)”
http://www.computerworld.com/action/ article.do?
[3] command=viewArticleBasic&articleId=9 087441”, 2009-04-25.
Kenneth R. van Wyk, “Adapting Penetration Testing for Software Development
[4] Purposes”, 2007, Carnegie Mellon University.
Jaak AAVIKSOO, Minister of Education and Research, Estonia; “Cyber attacks
Against Estonia Raised Awareness of Cyberthreats; Defence Against Terrorism
[5] Review Vol.3,No. 2 F all 2010,pp. 13-22 Copyright © COE-DAT ISSN:1307-9190
Halfond WGJ, Shauvik Roy Choudhary and Alessandro Orso" Improving penetration
testing through static and dynamic analysis "SOFTWARE TESTING,
VERIFICATION AND RELIABILITY *Softw. Test. Verif. Reliab.* (2011) Published
[6] online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/stvr.450
Halfond WGJ, Viegas J, Orso A, A classification of SQ Linjection attacks and counter
measures, Proceedings of the International Symposium on Secure Software
[7] Engineering, Washington, DC, U.S.A., March 2006.
Halfond WGJ, Anand S, Orso A. Precise interface identification to improve testing
and analysis of web
[8] applications. *Proceedings of the International Symposium on Software Testing and*
Analysis, Chicago, IL, U.S.A., 2009.
[9] Halfond WGJ, Orso A, Manolios P. WASP: Protecting web, applications using
positive tainting and syntax-aware evaluation,Transactions on Software Engineering
2008; 34(1):6581.
[10 Xiong Deng, Moustafa M. Ghanem, YikeGuo, “Real-Time Data Mining Methodology
and a Supporting Framework”Conference: International Conference on Network and
[11 System Security - NSS, pp. 522-527, 2009.
Pulei Xiong, Liam Peyton, A Model-Driven Penetration Test Framework for Web
Applications, 2010 Eighth Annual International Conference on Privacy, Security and
Trust.
[12 books.google.com.eg/books? Is b n=0521771455 Julien Bogousslavsky ,Louis R.
[13 Caplan - 2001 - Medical This may lead to platelet activation via 5-HT2 receptors (
Pietraszek et al., 1993). Clinical features TAO is characterized by claudicating or
[14 ischemia of both
T. Pietraszek and C. V. Berghe , Defending Against Injection Attacks through
Context-Sensitive String Evaluation, In Proceedings of Recent Advances in Intrusion
[1] Detection (RAID2005), 2005
Arkin, B., Stender, S., McGraw, G. (2005). “Software Penetration Testing”, IEEE
[16 Security and Privacy, Volume 3, Issue 1.
[17 Pierce, J., Jones, A., and Warren, M. (2007). “Penetration Testing Professional Ethics:
[18 a conceptual model and taxonomy”, Australasian Journal of Information Systems,
13(2). Available at: http://dl.acs.org.au/index.php/ajis/article/view/52 [Accessed 25
[19 July 2010]
[20 McRue, A. (2006). “University opens school for hackers”. URL:
http://news.cnet.com/University-opens-schoolfor-hackers/2100-7355_3-
[21 6085375.html [Accessed 8 August 2010]
Matt Bishop, "Introduction to Computer Security", Addison-Wesley.
[22 Matt Bishop, “About Penetration Testing”, Security & Privacy, IEEE.

[23 B. Duan, Y. Zhang, D. Gu, "An Easy-to-deploy Penetration Testing Platform", The
] 9th International
[24 Conference for Young Computer Scientists, 2008.ICYCS 2008.
] Bruce Schneier Blog http://www.schneier.com/blog/archives/2007/05/is_
[25 penetration.html, 2009-05-01.
] James F. Kurose, Keith W. Ross, "Computer Networking – A top Down Approach",
[26 4th edition, Addison
] Wesley Computing.
27] Dr. Daniel Geer and John Harthorne, "Penetration Testing: A Duet", Proceedings of
[the 18th Annual Computer Security Applications Conference (ACSAC '02).
28] "Nmap - Free Security Scanner For Network Exploration & Security Audits"
[http://nmap.org/, 2009-05-
[29 "The Metasploit Project" http://www.metasploit.com/, 2009-05-08.
] "Tenable Network Security" http://www.nessus.org/nessus/, 2009-05-08
[30 http://searchnetworking.techtarget.com/generic/0,295582,sid7_gci1083683,00.html
] http://www.weblaw.co.uk/templates_agreements/ethical_hacking_penetration_testin
[31 g/
] http://www.infosecinstitute.com/blog/ethicalhacking_computer_forensics.html
http://www.oissg.org/wiki/index.php/PENETRATION_TESTING_METHODOLOG
[32 Y
] http://searchnetworking.techtarget.com/generic/0,295582,sid7_gci1083715,00.html
International Journal of of Grid and Distributed ComputingVol.2, No.2, June 2009
Oona A. Hathaway, Rebecca Crootof, Philip Levitz, Haley Nix,AileenNowlan,
William Perdue, Julia Spiegel; "The Law of Cyber Attack" ; Forthcoming in the
California Law Review, 2012
[33 Mindy Chidester ;"The Exploitation of Social Media by Clandestine Groups, How
] Law Enforcement & Intelligent Can Better Utilize Social Media, and Legal Concerns
to Ensure its Appropriate Use by Government Entities " ; A Thesis Presented to the
[34 Faculty of San Diego State University In Partial Fulfillment of the Requirements for
] the Degree Master of Science in Homeland Security by Mindy Chidester Summer
2012
Julian Charvat;"Radicalization on the Internet";Defence Against Terrorism Review
[35 Vol.3, No2,F all 2010,pp. 75 85 Copyright © COE-DAT ISSN:1307- 9190
] Jaak AAVIKSOO, Minister of Education and Research,Estonia; "Cyberattacks
Against Estonia Raised Awareness of Cyberthreats; Defence Against Terrorism
Review Vol.3,No. 2 F all 2010,pp. 13-22 Copyright © COE-DAT ISSN:1307-9190
[36 Maslin Masrom, NikHasnaaNikMahmood, Othman Zainon, Hooi Lai Wan, Nadia
] Jamal ;"Information and Communication Technology Issues: A Case of Malaysian
Primary School" ; VOL. 2, NO. 5, June 2012 ISSN 2225- 7217 ARPN Journal of
Science and Technology ©2011-2012. All rights reserved.
[37 JeffyMwakalinga and Stewart Kowalski; "ICT Crime Cases Autopsy: Using the
] Adaptive Information Security Systems Model to Improve ICT Security" ; IJCSNS
International Journal of Computer 114 Science and Network Security, VOL.11 No.3,
March 2011
[38 PetterSvenhard& Amir Radaslic"A Penetration Test of an Internet Service Provider"
] ;© Copyright PetterSvenhard, Amir Radaslic, 2012. All rights reserved Bachelor
Thesis Report, IDE1256 School of Information Science, Computer and Electrical
Engineering Halmstad University

[39] Rizik M.H Al-Sayyedatel ;“ Search Engines in Website Security Leak” ; World Applied Sciences Journal 20 (5): 753-759, 2012 ISSN 1818-4952 © IDOSI Publications, 2012 DOI: 10.5829/idosi.wasj.2012.20.05.261212

Mrs.Yogini A. Kulkarni Mr. Rajendra.G. Kaduskar Department Of Computer Engg., Department Of E &TC Engg. PVG’s COET, PVG’s COET, Pune, India Pune, India,;”

[40 Security against Malicious Code in Web Based Applications”; 978-0-7695-4246-1/10 © 2010 IEEE DOI 10.1109/ICETET.2010.53

Eric KeWang,Yunming Ye, XiaofeiXu Department of Computer Science Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China, S.M.Yiu, L.C.K.Hui, K.P.Chow Department of Computer ScienceThe University of Hong Kong Pokfulam, Hong Kong;” Security Issues and Challenges for Cyber Physical System; 2010 IEEE/ACM International Conference on Green Computing and Communications & 2010 IEEE/ACM International Conference on Cyber, Physical and Social Computing 978-0-7695-4331-4/10 © 2010 IEEE DOI 10.1109/GreenCom-CPSCom.2010.36 733

“IBM – Features” http://www- 935.ibm.com/services/us/index.wss/detail/iss/a1027213?cntxt=a1027208, 2009-05-08.

Kenneth R. van Wyk, Software Engineering Institute, “Penetration Testing Tools”, 2007, Carnegie Mellon University.