European Journal of Computer Science and Information Technology
Vol.2,No.4,pp.29-39, December 2014
Published by European Centre for Research Training and Development UK (www.eajournals.org)

NEURAL NETWORKS APPROACH FOR MONITORING AND SECURING THE E-GOVERNMENT INFORMATION SYSTEMS

Hanaa. M. Said, Mohamed Hamdy, Rania El Gohary and Abdelbadeeh M. Salem

Ain Shams University, Faculty of Computing, Information Science
El abasea squar, Cairo, Egypt

ABSTRACT: *Security must be addressed in the phase of planning and designing of e-government system. Management process is needed to assess security control, where management allows departments and agencies to maintain and measure the extent of data security depending on the mechanism of revealing the security weak points .Revealing the weak points is done by using a series of standards built on the application of machine learning methods specifically Using the Neural Networks Model, and intelligent data analysis. All these techniques are useful in monitoring and measuring the extent of the secured data and the provided services. The applied results on the data site of "Cairo cleanliness and beautification authority for cleaning" in Egypt showed that measurement qualifications were adequate, proper ,preaching, and can be generalized. The proposed approach of monitoring is very comprehensive where it limits the risk of information security that affect organizations' risk management decisions.*

KEYWORDS: *Government Cyber space, Data mining, the Neural Networks Model.*

INTRODUCTION

E-Government category includes interactions between governments and businesses that is (government is selling and providing the businesses with services while businesses is selling products and services to the government) [6]. E-Government can be defined as moving citizen services online, but in its broadest sense it refers to the technology-enabled transformation of the government. Governments must serve all members of the society irrespective of their physical capabilities where online services will have to be designed with appropriate interfaces [1].E-Governments have faced many challenges and obstacles such as: network crimes, lack of information technology rules security data which enables governments to create positive business climates by simplifying relationships with business and reducing the administrative steps needed to comply with the regulatory obligations.

The security Cyber space of e-government systems are monitored through: making information accessible, publishing government debates and minutes, budgets and expenditure statements, outcomes and rationales for key decisions, so as to get good services from the government as this helps to increase the transparency of the decision-making processes. The model is used to monitor and predict the measurement and data status, so that the ability to predict this outcome is the central goal of the predictive analysis [10]**.** It can be expressed for this real cyber space as if it is the series of the minor cyberspaces. Our objective is to analyze, study and secure one of the minor cyber space's which is the cyberspace for the authority of cleaning and beautifying Cairo in the Arab Republic of Egypt (www.ccba.gov.eg) so as to analyze the extent of the sufficiency for the suggested reasoning to measure the extent of securing data for the cyberspace. It is one of the important cyberspaces in the frame of mechanism for the e-

government services, and its effect on both the citizens, the investors and on the government. This cyberspace is connected with several electronic sites. Illegal entry or data manipulation as well as un-authorized access to the architecture underlying this data, can be limited by a Cyber space which hinders providing good and useful service. Identifying the difficulties is one of the major issues that face the government officials. The accuracy of the data and the extent of safety related requirements as well as monitoring the levels would be secured through data collection and analysis and identifying the most important obstacles that lead to reducing the level of government performance [3].

In this paper, we monitor Cyber space security of e-government systems by Using the Neural Networks and analysis of the data and making sure they are penetrated to provide strategic information to the various provided services which enables the province to find important points for effective management [2]. The Neural Networks Model is very useful in enabling the decision-maker to interact with the characteristics features of value which will be adapted with data mining tools. Data Mining is applied to the process of an actual sample of the government to prove the feasibility of this goal and in time shows a global view of the solution [5]. The design of the combined tool or generalized data search and selection algorithms is necessary to search for information in the context of scientific discoveries which is why the application of a set of module is to be monitored and measured.

We have a vision to decide whether these data, which we use through those sites are safe enough, or whether we must protect them against any probable attacks. Therefore, this research gives a general perspective to measure the level of security of websites, as we depend on a model to measure how much the data is valid and secure. That might require knowing other things than the data itself.

This Paper mainly discusses the issue of Neural Networks Approach for Monitoring and Securing the e-Government Information Systems and it is divided into the following sections: Section (1) The complete introduction: A General review on the security of cyberspace in e-governments. Section (2) presents the Related Work. Section (3) describes the Neural Networks Model which describes how to use a useful technology style in monitoring and measuring the extent of securing the data. Section (4) presents the discussions concerning the different results of this study. Finally, Section (5) ends this paper with a summary and conclusion.

Related Work

In this section we shall see a real-life example of Networks Testing that involved the civilian government Agency.

Most of the work in the area of building operational IDSs using data mining depends on an offline analysis phase to build models of normal behaviour. Data mining algorithms applied to network traffic data include: Association Rules mining, classification, and clustering. Association Rules Mining [5-7] was used to identify interesting attributes that occur together with a high support in the traffic data, where these associations were used to classify the traffic data online. Classification Techniques such as Decision Trees [8] and Bayesian Classification [9] were used to build classifiers of normal network activity data and detect data that do not match the classifiers' rules.

Clustering Techniques work by grouping similar data instances into clusters regardless of the instances' class labels; they were used to tackle the problem of labelling data before training,

which is inherent in the association rule mining and in the classification techniques. Clustering is used in [10], where three clustering strategies are applied on data containing both attack scenarios and normal traffic, namely, the Self-Organizing Maps(SOM), K-means Clustering, and Expected Maximization Clustering. This work shows that SOM has the same complexity regardless of the data volume or clusters used. However, SOM can misclassify data inputs that will correspond to nodes which were not affected during training. The K-means clustering algorithm has a predictable performance and classification; however, it poorly manipulates highly dimensional data sets. The Expected Maximization Algorithm can tolerate missing and unlabeled data and can offer information about how close a data point is to each cluster since the data point has a varying membership to all clusters.

Stepien et al [17] presented an approach to Network testing for inherent to penetration testing of web application which 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 also described a message sequence diagram of a malicious bank website to show the XSS attacks. It generates the functional test cases.

Pietraszek et al [18] Presented an approach of Taint based Technique in which the author modified a 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.

Halfond et al [19] 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 where 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.

Neural Networks Model

The Neural Networks predict a continuous or categorical target based on one or more predictor's by finding unknown and possibly complex patterns in the data. We can build the network by using the calculation between each pair of the variables though using the Neural Networks Algorithms for building as the maximum level**.** It starts with the extended tree with the non-existence of the edges and the signs of the random variable as an approach. Then, we will find the variable of the non-controller where its weight with one of the observable variable is the maximum limit.

Applying to natural network log, the network consists of the six variables: (x1, x2, x3, x4, x5, y) where it constitutes a network of two heads of input and output changes which vary from (x1……x5, y) with different gradual colures (blue – yellow) the most coloured and thickness line represents the most effective and important.
These future variables have all been coded and compared with the dependent variable (z) where we can distinguish which of them is more or less important or need improvement or treatment. The network is formed from coloured lines. The blue colour represents the positive effect and the yellow colour is the negative effect. If the coloured line gets darker and thicker, it will be more positive and has a higher impact and interest. If the coloured lines become lighter and thinner, they will be less positive and have no impact or interest.

Figure (1) shows all variables from (x1 ..... x5, y) in different colours Gradient (Lama blue - yellow), we find that the line containing most of the colours is thicker and represents the most effective and is the most important. We find the network reflects the six variables: (X1, X2, X3, X4, X5, y). Using the hypothesis of the original premise that all of these variables in the future equal weights values, therefore they are all of equal importance where the sum of all these weights = 1, so that the value of each of them constitutes 17% of the original hypothesis value.

The following notation is used with multi-layers perceptions' unless otherwise stated:

Table 1: Notation Variables Description

Notation	**The following notation is used for multilayer perceptions**
$\boldsymbol{x^{(m)} = (x_i^{(m)} \ldots\ldots, x_p^{(m)})}$	Input Vector, Pattern m, where m=I …M
$\boldsymbol{Y^{(m)} = (Y_i^{(m)} \ldots\ldots, Y_p^{(m)})}$	Target Vector, Pattern (m).
I	Number of layers, discounting the input layer
$\boldsymbol{j_i}$	Number of units in layer $\boldsymbol{j_o = p_i j_i = R_i}$ discounting the bias unit
$\boldsymbol{r^c}$	Set of categorical outputs.
r	Set of continuous outputs.
$\boldsymbol{r_h}$	Set of sub vectors of $\boldsymbol{(Y)^m}$ cantoning 1 of 0 coded h the categorical Field
$\boldsymbol{a_{i:j}^m}$	Unit j of layer i, pattern m, j=0, $\boldsymbol{j^i; i = 0.., 1}$
$\boldsymbol{w_i, j, k}$	Weight leading from layer i-1,unit j to layer i, unit k, no weights
$\boldsymbol{w_i, j, 0}$ for any j	Connect $\boldsymbol{a_i^m - 1.j}$ and the bias $\boldsymbol{a_{i=0}^m}$ that is there is no
$\mathbf{c_{i:k}^m} \sum_{j=0}^{j_i} \boldsymbol{w_{i,k} a_{i-1,j;}^m i = 1, \ldots\ldots I}$ (1)	
W Weights$\mathbf{w_{1,0,1}, w_{1,0,2}, \ldots\ldots, w_{i;j_{i-1},} j_1}$	Weight of vectors containing all

Y(c) =c

This function of output layer when there are continuous targets.

Soft max

$$y(c_k) = \frac{\exp(c_k)}{\sum_{j \in Th} \exp(c_j)} \quad (2)$$

Error Functions- Sum-of-Squares

$$E_T(w) = \sum_{m=1}^{m} E_m(w) \quad (3)$$

The general architecture for MLP networks is:

Input layer: J0=P units,; with $a_{0;j} = x_j$.

It hidden layer: Ji units, $a_{0;1,\ldots\ldots a_{0;j}}$; with $a_{i;k} = y_i(c_{i;k})$ and $c_{i;k} = \sum_{j=0}^{j_i} wi,j,ka_{i-1,j}$ where $a_{i-1:0} = 1$.

Output layer: JI=R units, ; with Ji units ,$a_{0;1,\ldots\ldots a_{0;j}}$; with $a_{i;k} = y_i(c_{i;k})$ and$c_{i;k} = \sum_{j=0}^{j_i} wi,j,ka_{i-1,j}$

Where $a_{i-1:0} = 1$

Note that the pattern index and the bias term of each layer are not counted in the total number of units for that layer.

Activation Functions and Hyperbolic Tangent:

$$Y(c) = \tan(c)\frac{e^{c}-e^{-c}}{e^{c}+e^{-c}} \quad (4)$$

The function to the hidden layers Identity

Y (c) = tan (c)

Y(c) =c

The function of the output layer when there are continuous targets

Soft max

$$y(c_k) = \frac{\exp(c_k)}{\sum_{j\in Th}\exp(c_j)} \quad (5)$$

Cross-Entropy Where

$$E_M(W) = \frac{1}{2}\sum_{r=1}^{R}(y_r^{(m)} - a_{I:r}^{m})^2 \quad (6)$$

$$E_T(W) = \frac{1}{2}\sum_{m=1}^{M} 1\, E_m(w) \quad (7)$$

$$E_m(w) = -\sum_{r\in r^c}^{0} y_r^{(m)} \log\left(\frac{a_{I:r}^{m}}{y_r^{(m)}}\right) \quad (8)$$

Error Functions - Sum-of-Squares

$$E_T(w) = \sum_{m=1}^{m} E_m(w) \quad (9)$$

The function of Accuracy when all targets are categorical

$$\text{Accuracy}=\frac{1}{n}\sum_{M=1}^{M}\left(1 - \frac{\left|Y_r^{(m)} - y_r^{(m)}\right|}{\max_m\left(y_r^{(m)}\right) - \min_m\left(y_r^{(m)}\right)}\right) \quad (10)$$

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}}, \quad (11)$$

For each categorical target, this is the percentage of records for which the predicted value matches the observed value. On the Model Tab, the Predictor Importance Graph shows the relative effects of the various fields to the prediction. This shows us that Neural Networks Model has easily the greatest effect, while Unit Shape and Clump are also quite significant. Where x' i is the rescaled value of input field x for record i, (xi) is the original value of x for record i, x sub is the minimum value of x for all records, and (x max) is the maximum value of x for all records.

Re-code a symbolic field as a group of numeric fields with one numeric field for each category or value of the original field. For each record, the value of the derived field corresponding to the category of the record is set to 1.0, and all the other derived field values are set to 0.0. Such Derived fields are sometimes called (Indicator Fields), and this recoding is called (indicator Coding). For data set of Cairo Cleaning Authority, consider the following data, where x is a symbolic field.

In this data, the original set field x is recoded into three derived fields x1', x2', and x3'. X1' is an indicator for category A, (x2') is an indicator for category B, and x3' is an indicator for category C.

Table 2: The following Codec variables

Record #	**X**	$\mathbf{x_1'}$	$\mathbf{x_2'}$	$\mathbf{x_3'}$
1	B	0	1	0
2	A	1	0	0
3	c	0	0	1

It possible values A, B, and C. Feed-forward Calculations Information flows through the network as follows:

Input neurons have their activations set to the values of the encoded input fields. The activation of each neuron in a hidden or output layer is calculated as follows:

$$a_i = \sigma(\left(\sum j w_i o_j\right), \quad (12)$$

$$\sigma(x) = \frac{1}{1+e^{-x'}} \quad (13)$$

The weight of the connection between neuron (i) and neuron (j), is the output of neuron j, and is the sigmoid or logistic transfer function Back-propagation Calculation**s** at the beginning where all weights in the network are set to random values in the interval. Records are presented in cycles (also called epochs), where each cycle involves presenting randomly selected records to the network, where (n) is the number of records in the data set. Because of the random selection process in any particular cycle, some records may be presented more than once and others may not be presented at all.

For each record, information flows through the network to generate a prediction, as described above. The prediction is compared to the target value found in the data for the current record, and that difference is propagated back through the network to update the weights. To be more precise, the change value Δ w for updating the weights is calculated as where is the rate parameter, is the propagated error (described below), is the output of neuron i for record p, is the momentum parameter, and is the change value for at the previous cycle.

The value of Input is fixed during cycles, but the value of output varies across cycles of network. It starts at the user-specified initial σ(x), decreases logarithmically to the value of low σ(x), reverts to the high σ(x) value, and then decreases again to low value. The value of σ(x), is calculated as indicator where d is the user-specified number of σ(x) decay cycles. If, then is set to. And continues to cycle thusly until is complete. The back-propagated error value is calculated based on where the connection lies in the network for connections to output neurons.

The target values which will be recorded are calculated. The disconnected neurons represent propagated errors through the network.

In figure (1) we see that the colours are used as indicators of the level of security, In order to figure out which of the flowing factors is affected i.e. (Availability, Timeliness, security, Objectivity and Integrity). We focus on these characteristics which are important to site the government cyber space. We need to detect and monitor which of these measurements with more weight when decoding the random variables (x1, x2, x3, x4, x5, y) and giving them one of the two values (0 or 1) at three levels for the measurement to the indicator of the assessment.

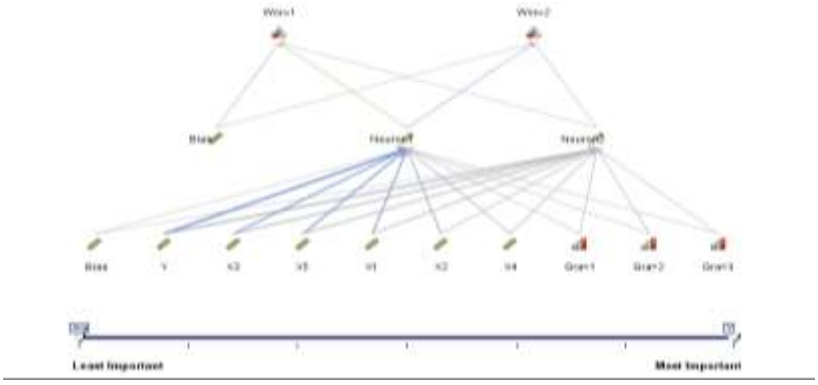

Figure 1: Neural Networks Model

- Probability of occurrence (92.0 %)
- Failure probability of occurrence (0.8%)

Figure 2 show as reduce these variables work to hide some of the variables and identify the hidden three variables by the inputs and outputs and determine the weight, we find that security is the most important property.

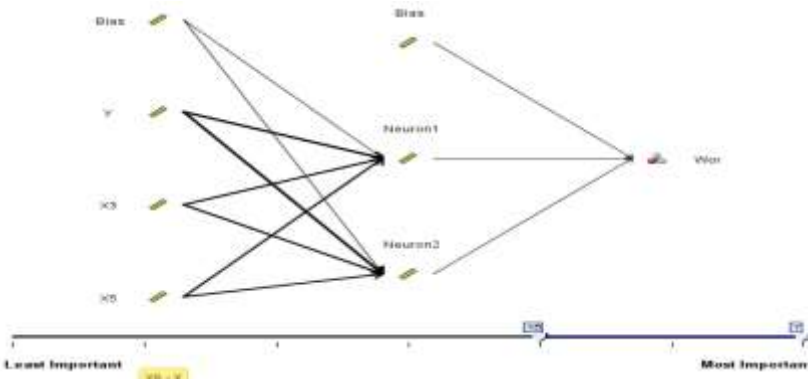

Figure 2: Predict the value of each parameter

• Probability of occurrence (95.0 %)

• Failure probability of occurrence (0.5%)

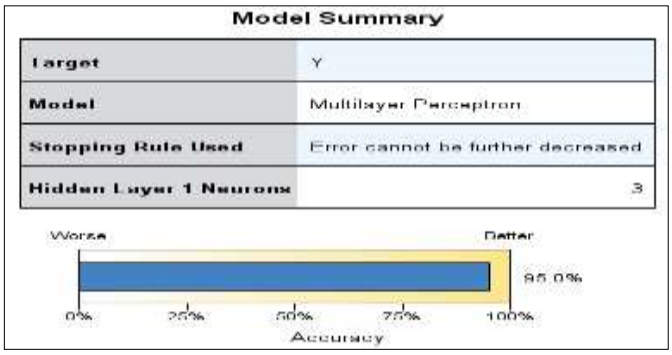

Figure 3: Sample of Predict Accuracy

The experiment achieved its targets in the image of indicators for both the input and output which clearly shows the high level of security we achieved. This model is used to predict and monitor the status of data, and decides to use it as a quality indicator, to improve the performance as this show in Figure (3).

Intentionally or not there should be bias to data accuracy (morally in the real data) especially the significant and critical data and these are the values which constitute the importance of the variable which appear**s** in the (I &OP)**.**

Comparing the values on the rectangles diagram with hypothesis values 17%; we can find that (Y) represents for security form the value 28% and this parameter is very important and has its effect. We must take care with it and improve its performance. Breaking them will be more negative, has no effect and can be cancelled.

X3 = 21% has its effect and importance but less importance than Y,

X5= 18% has also its effect than the hypothesis value 17%, it has its effect and importance but less importance than parameter (Y).

X1 = 16% it is less bias and importance than X5, X3 and Y.

Both X2 and X4 equal to 7 % and can be neglected due to their low effect, so we can avoid and pay no attention to it.

The advantage of this experiment, can be summarized in determination of the important of the effective dependent and independent variables also through this technique we can determine the relative importance of each variable and determination of the priority of the variables. In this work, security is most important variable.

Neural network was the guidance for the form among different relationships.

Determine the actual grouping attacks in a set of unlabeled data. For algorithms on enhancing model accuracy or stability or working with very large datasets, a Neural Network is a structure of many such neurons connected in a systematic way.

RESULTS AND DISCUSSION

The Results for Polynomial look much better. Many of the propensity scores are 0.995 or better, which is very encouraging. To confirm the improvement in the model, an Analysis Node was attached to the class-poly model nugget. Open the Analysis Node and click run this technique to enable you compare two or more model nuggets of the same type. The output from the Analysis Node shows that the RBF function correctly predicts 97.85% of the cases, which is still quite good. However, the output shows that the Polynomial Function has correctly predicted the diagnosis in every single case. In practice, you are unlikely to see 100% accuracy, but you can use the Analysis node to help determine whether the model is acceptably accurate for your particular application. In fact, neither the Sigmoid nor the linear function performs as well as the Polynomial function on this particular dataset. However, with a different dataset, the results could easily be different, so it is always worth trying the full range of options.
It is taken for granted that preparing or how to use the means of research measurements and testing the security of the data submitted recently for E-government and effective for the cyberspace may help in identifying the important and useful indicators from the practical side. In addition to that, I see that the government or any other system needs a new technology or modern one to include the follow up and knowing the challenges that face the cyberspace, and automatically identify the directions of the used electronic security threats.

CONCLUSION

E-Governments have faced many challenges and obstacles such as: network crimes, lack of information technology rules security data which enables governments to create positive business climates by simplifying relationships with business and reducing the administrative steps needed to comply with regulatory obligations. Monitoring security Cyber space of e-government systems is through: making information accessible, publishing government debates and minutes, budgets and expenditure statements, outcomes and rationales for key decision set good services from the government; this helps to increase the transparency of decision-making processes. We use the neural networks model to Manage process is needed to assess security control, this management allows departments and agencies to maintain and measure the extent of data security depending on the mechanism of revealing the security weak points, Analyze the extent of the sufficiency for the suggested reasoning to measure the extent of securing data for the cyberspace. It is one of the important cyberspaces in the frame of the mechanism for the e-government services, and its effect on both the citizens, the investors and on the government, this cyberspace is related with several electronic sites. The applied results on the data site of the ”Cairo cleanliness and beautification authority for cleaning” in Egypt refer to that measurement qualifications were adequate, proper, preaching and can be generalized .The proposed approach of monitoring is very comprehensive and limit the risk of information security that affect organizations’ risk management decisions. Results of this research are very useful to build a strategy for measuring the extent of securing data in order to improve the management of servants effective government, any type of data to be used, any type of data was transferred in a proper way, Could be this study remarkable as one of the first studies on the use of data mining tools in Cyberspace. Finally, this paper could become an important tool for the government and intelligence agencies in the decision-making and monitoring potential international terrorist threats in real time present at the talks and research and blogging.

Reference

[1] Aggarwal, C. C., P. S. Yu. (1998). Online generation of association rules. In: Proceedings of the 14th International Conference on Data Engineering, Los Alamitos, Calif: IEEE Computer Society Press, 402–411.
[2] Agrawal, R., R. Srikant. (1994). Fast Algorithms for Mining Association Rules. In: Proceedings of the 20th International Conference on Very Large Databases, J. B. Bocca,M. Jarke, and C. Zaniolo, eds. San Francisco: Morgan Kaufmann, 487–499.
[3] Agrawal, R., R. Srikant. (1995). Mining Sequential Patterns. In: Proceedings of the Eleventh International Conference on Data Engineering, Los Alamitos, Calif.: IEEE Computer Society Press, 3–14.
[4] Aitkin, M., D. Anderson, B. Francis, and J. Hinde. (1989). Statistical Modelling in GLIM. Oxford:Oxford Science Publications.Albert, A., and J. A. Anderson. 1984. On the Existence of Maximum Likelihood Estimates in Logistic Regression Models. Biometrika, 71, 1–10.
[5] Anderson, T. W. (1958). Introduction to multivariate statistical analysis. New York: John Wiley &Sons, Inc... Arya,S., and D. M. Mount. (1993). Algorithms for fast vector quantization. In: Proceedings of the Data Compression Conference (1993), , 381–390. collinearity. New York: John Wiley and Sons.
[6] Belsley, D. A., E. Kuh, and R. E. Welsch. (1980). Regression diagnostics: Identifying influentialdata and sources of Biggs, D., B. de Ville, and E. Suen. (1991). A method of choosing multiway partitions for classification and decision trees. Journal of Applied Statistics, 18, 49–62.
[7] Bishop, C. M.(1995). Neural Networks for Pattern Recognition, 3rd ed. Oxford: Oxford UniversityPress.
[8] Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. Journal of the Royal Statistical Society, Series B, 26, 211–246.
[9] Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. (1994). Time series analysis: Forecasting andcontrol, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.
[10] Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone.(1984). Classification and RegressionTrees. New York:Chapman & Hall/CRC.
[11] Breslow, N. E. (1974). Covariance analysis of censored survival data. Biometrics, 30, 89–99.
[12] Brockwell, P. J., and R. A. Davis. (1991). Time Series: Theory and Methods, 2 ed. :Springer-Verlag.
[13] Cain, K. C., and N. T. Lange. (1984). Approximate case influence for the proportional hazardsregression model With censored data. Biometrics, 40, 493–499.
[14] Cameron, A. C., and P. K. Trivedi.(1998). Regression Analysis of Count Data. Cambridge:Cambridge University Press.
[15] Chang, C. C., and C. J. Lin. 2003. LIBSVM: A library for support vector machines. Technical Report. Taipei, Taiwan: Department of Computer Science, National Taiwan University.
[16] Chow, C. K., and C. N. Liu. (1968). Approximating discrete probability distributions with Dependence trees. IEEE Transactions on Information Theory, 14, 462–467.
[17] Kenneth R. van Wyk, Software Engineering Institute, “Penetration Testing Tools”, 2007, Carnegie Mellon University.
[18] Matt Bishop, "Introduction to Computer Security", Addison-Wesley.

[19] Matt Bishop, “About Penetration Testing”, Security & Privacy, IEEE.