

ISSN 1913-8989

COMPUTER AND INFORMATION SCIENCE

Vol. 1, No. 3
August 2008



Canadian Center of Science and Education



Contents

Artificial Neural Networks for the Prediction of Thermo Physical Properties of Liquid Mixtures <i>R. Baskaran, S. Arunachalam, K. Manjunath & T.R Kubendran</i>	3
Computer Aided Investigation towards the Wind Power Generation Potentials of Guangzhou <i>Gang Yang, Yongxian Du & Ming Chen</i>	13
Autonomous Notification and Situation Reporting for Flood Disaster Management <i>Ku Ruhana Ku-Mahamud, Norita Md Norwawi, Norliza Katuk & Safaai Deris</i>	20
The Role of Session Border Controllers in the DMZ of Voice over IP(VoIP) Networks <i>Huihong Chen & Zhigang Chen</i>	27
Image Analysis for Online Dynamic Steganography Detection <i>Kevin Curran & Joanne Mc Devitt</i>	32
Model of On-line Room Reservation System Based on Web Service and XPMS <i>Jiali Wang, Jun Zheng & Wenxin Hu</i>	42
Construction of New Forest Roads in Malaysia Using a GIS-Based Decision Support System <i>Kamaruzaman Jusoff</i>	48
Revelations of the New Developments of Brain Science on Early Education <i>Tianbao Zhuang, Hong Zhao & Zheng Tang</i>	60
Similarity Matrix Based Session Clustering by Sequence Alignment Using Dynamic Programming <i>Dr.K.Duraiswamy & V. Valli Mayil</i>	66
Research of Information System of Monitoring Remote on Distributed Micro-turbine <i>Liangwei Zhong, Kangmin Chen, Jing Ni & Anyang He</i>	73
A Novel Sanitization Approach for Privacy Preserving Utility Itemset Mining <i>R.R.Rajalaxmi & A.M.Natarajan</i>	77
The Drive and Support of Web3.0 to Rich Media Advertisement <i>Xiaonan Jin</i>	83
Power Estimation in Mobile Communication Systems <i>Lenin Gopal, Ashutosh Kumar Singh & Veeramani Shanmugam</i>	88
Design of Calendar Clock Based on DS12C887 Chip <i>Xiao Chen</i>	95
Creating Websites: Adobe Go Live vs. LiveText <i>Donna F. Herring, Charles E. Notar & Mrs. Judy McEntyre</i>	100
The New Development of CAD in Clothing <i>Lingxia Zhang, Hongzhi Zhang & Yinglin Li</i>	104
A New Approach to Sensitive Rule Hiding <i>Dr.K. Duraiswamy, Dr.D. Manjula & N. Maheswari</i>	107
New Path Filling Method on Data Preprocessing in Web Mining <i>Chungsheng Zhang & Liyan Zhuang</i>	112
Hough Transform and Its Application in Vehicle License Plate Tilt Correction <i>Zhenhui Zhang & Shaohong Yin</i>	116



Contents

Study on the Scheme for RSA Iterative Encryption System	120
<i>Yulian Shang, Wuyuan Jia, Peng Li & Pengfei Zhu</i>	
The Design and Implementation of Precious Test and Quality Control System Based on Delphi	125
<i>Ruoping Wang & Liguo Lai</i>	
A New Vision Inspection Method for Wood Veneer Classification	129
<i>Mengxin Li & Chengdong Wu</i>	
Summarization of Program Development in Multi-core Environment	135
<i>Hongyu Li</i>	



Artificial Neural Networks for the Prediction of Thermo Physical Properties of Liquid Mixtures

R. Baskaran

Department of Chemical Engineering

St. Joseph's college of Engineering

Chennai-119, India

Tel: 91-99-4128-5866 E-mail: rbaskaran2000@yahoo.com

S. Arunachalam & K. Manjunath

Department of Electrical and Electronics Engineering

St. Joseph's college of Engineering

Chennai-119, India

T.R Kubendran

Department of Chemical Engineering

Alagappa College of Technology

Anna University

Chennai-600025, India

E-mail: trkubendran@yahoo.co.in

Abstract

A predictive method based on Artificial networks has been developed for the thermophysical properties of binary liquid mixtures at (303.15, 313.15 and 323.15) K. In method 1, a committee ANN was trained using 5 physical properties combined with absolute temperature as its input to predict thermo physical properties of liquid mixtures. Using these data, predicted values were determined for intermediate mole fraction of different systems without conducting experiments. In method 2, a committee ANN was trained using mole fraction and molecular weight as its input to predict the thermo physical properties of liquid mixtures. The five physical properties of five binary mixtures were taken for this study along with their molecular weights. ANN with back-propagation algorithm is proposed, for Multi-pass Turning Operation and developed in MATLAB. Compared to other prediction techniques, the proposed ANN approach is highly accurate and error is <1%.

Keywords: Artificial neural network, Neurons (nodes), Perceptron, Network training, Para anisaldehyde

1. Introduction

In many real world applications, we want our computer to perform complex recognition problems. Since our conventional computers are obviously not suited to this type of field, we therefore borrow features from the physiology has come to be known as Artificial Neural Systems (ANS) Technology or Neural Networks. Artificial neural network is a branch of artificial intelligence (AI) that attempts to achieve human brain like capability. Traditional approaches of solving chemical engineering problems frequently have their limitations, as for example in the modeling of highly complex and nonlinear systems. Artificial neural networks (ANN) have proved to be able to solve complex tasks in a number of practical applications. The utility of artificial neural network models lies in the fact that they can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical. Because ANN are nets of basis functions, they can provide good empirical models of complex nonlinear processes useful for a wide variety of purposes. The applications of ANN include detection of medical phenomena, stock market prediction, credit assignment, monitoring the condition of

machinery and engine management. There are only few reports of using ANN in the prediction of physiochemical properties, these reports have generally been restricted to equilibrium rather than transport properties (John.et.al 1999). It is believed that so far there has been no attempt to truly predict the properties of liquid mixtures across their wide temperature range using ANNs. This paper presents the findings of a programme of work devoted to the application of ANNs to thermo physical properties of binary mixtures. The thermo physical properties of binary mixtures are found by experimental work done by us. A predictive method based on Artificial Neural Networks (ANN) has been developed for ultrasonic velocity, density, kinematic viscosity, surface tension and refractive index for para anisaldehyde liquid mixtures over a wide range of mole fraction and temperatures. The study of various properties like viscosity, refractive index, density, surface tension and ultrasonic velocity for different systems are very important to understand the molecular interactions and purity of compounds. Using these experimental data, predicted data for intermediate mole fraction of different systems were generated, without conducting experiments.

2. Working procedure

A Neural Network is an interconnected assembly of simple processing elements, *units* or *nodes*, whose functionality is loosely based on the animal neuron. The processing ability of the network is stored in the inter-unit connection strengths, or *weights*, obtained by a process of adaptation to, or *learning* from, a set of training patterns. It has been shown that non linear feed forward neural networks are capable of universal functional approximation and that a single hidden layer is sufficient to uniformly approximate any continuous function Hornic.et.al.(1989). The neurons in a single hidden layer tend to interact globally but in complex functions this interaction makes it difficult to improve the approximation Heykin (1994), Maren .et.al (1990). The brain is principally composed of a very large number (circa 10,000,000,000) of *neurons*, massively interconnected (with an average of several thousand interconnects per neuron, although this varies enormously).

2.1 Artificial neurons

To capture the essence of biological neural systems, an artificial *neuron* is defined as follows:

- It receives a number of inputs (either from original data, or from the output of other neurons in the neural network). Each input comes via a connection that has a strength (or *weight*); these weights correspond to synaptic efficacy in a biological neuron. Each neuron also has a single threshold value. The weighted sum of the inputs is formed, and the threshold subtracted, to compose the *activation* of the neuron
- The activation signal is passed through an activation function (also known as a transfer function) to produce the output of the neuron.

In Figure 2.1, various inputs to the network are represented by the mathematical symbol, $x(n)$. Each of these inputs is multiplied by a connection weight. These weights are represented by $w(n)$. In the simplest case, these products are simply summed, fed through a transfer function to generate a result, and then output. This process lends itself to physical implementation on a large scale in a small package. The commonest type of artificial neural network consists of three groups or layers of units: input, hidden, and output. The layer of input neurons receives the data either from input files or directly from electronic sensors in real-time applications. The output layer sends information directly to the outside world, to a secondary computer process, or to other devices such as a mechanical control system Necat(2006). Between these two layers can be many hidden layers. These internal layers contain many of the neurons in various interconnected structures. The inputs and outputs of each of these hidden neurons simply go to other neurons.

2.1.1 Prediction Of Thermo physical Properties

Step 1: Collect data:

Things to be kept in mind while choosing the input data .

- The variables that are influential must be chosen
- Numeric and nominal variables can be handled. Convert other variables to one of these forms, or discard.
- Hundreds or thousands of cases are required; the more variables, the more cases.

Step 2: Define a network structure

An appropriate network topology is selected. Here for prediction problem, the neural network developed is fully connected feed forward multilayer perceptron. For this problem, the input variables are Temperature and mole fraction. So, the number of Input nodes is 2. Here, we have to predict the density, kinematic viscosity, ultrasonic velocity, surface tension and refractive index of the systems. Hence, the number of output nodes is 5. After a number of experiments, the transfer function of the hidden layer and the number of epochs are set, the details of which are given below. For example consider a feed forward net with 30 hidden nodes in a single layer.

Step 3: Random initialization of weights and biases for the defined network

Step 4: Back propagation learning Mandic (2001), Reilly (1982) algorithm is selected for training the network

Step 5: Training parameters are set

Number of iterations: 600

Performance goal: 0

Step 6: Start training

During training, the input and output data are presented to the network.

Step 7: Termination

The error for the network is calculated. The error calculation and weight updating process continues until the neural network reaches the performance goal or number of iteration.

3. Results and Discussion

Experimental thermo physical property values were extracted from the data base for all the mixtures studied. The list of binary liquid mixtures studied was displayed in Table 1. Input and output variables of method-1 and method -2 are listed in Table 2. The thermo physical properties were density, viscosity, refractive index, surface tension and ultrasonic velocity. The input to the ANN consisted of five thermo physical properties of five binary mixtures at (303.15,313.15 and 323.15) K for method - 1. The out put from the ANN consisted of five predicted thermo physical properties of five binary mixtures. In method -2 the input consisted of mole fraction, molecular weight and out put was thermo physical properties of five binary liquid mixtures. A BPN simulator is designed and the input data were fed in. BPN simulator is trained several times using selected data from the collected data, which consists of normal as well as abnormal data. During training, the simulator is presented with both input and output pairs and the error is generated which is the difference between actual and desired output. The error is minimized using the steepest descent technique. When the error obtained is of acceptable value, then the simulator is said to be trained. Then, the data for prediction is presented to the neural network after training. The mean square error is calculated using BP algorithm and the learning curve is plotted between the mean square error and the number of generations (iterations) i.e. epochs. Performance and error graph of method -1 and 2 of the trained committee ANN for the binary systems are presented in Figure 3.1- 3.10. While examining the results of these various neural networks, they were found to be in agreement with the desired results and within permissible error range. The algorithms traingdm or traingd are not producing satisfactory results for the 2 X 20 X 10 X 5 neural network for the maximum epochs of 800 and for a performance goal of 0. The trained algorithm with a single hidden layer with 30 nodes i.e 2 X 30 X 5 neural network. (Figure2.1.1) is producing the desired result for the prediction of these properties. The predicted values of other new mole fractions of the mixtures were taken and verified experimentally in Table 3. The percentage of error is <1.

4. Conclusion

Due to high speed of processing, low consumption of memory, great robustness, possibility of self learning and simple incorporation into chips the approach ensures prediction condition in real time. It provides robust representation clue to the fault-tolerant nature of neural networks. When comparing methods 1 and 2, the first method holds good which is indicated by low error. Our future proposal is Extending the proposed idea of ANN prediction to systems at extremely high or low temperatures where conducting of experiments are difficult.

References

- John, H., Sotos, C.G., John, H.R. (1999). Artificial neural networks for the prediction of liquid viscosity, density, heat of vaporization, boiling point and pitzers acentric factor. *Phys.Chem.Chem.Phys.* 1, 4075-4081.
- Hornik, K., Stinchcombe, M., White, H.(1989). Multilayer feed forward neural networks are Universal approximators. *Neuron Networks*, 2, 359-366.
- Heykin,S.(1994). *Neural network a comprehensive foundation*, New York, ch.6.12.
- Maren, A., Harston,C., Pap,R. (1990).*Handbook of neural computing applications*, New York, 238.
- Necat Altinkok. (2006). Use of Artificial Neural Network for Prediction of Mechanical Properties of -Al₂O₃ Particulate-reinforced Al-Si10Mg Alloy Composites Prepared by using Stir Casting Process. *Journal of Composite Materials*, 40, 9, 779-796.
- Mandic, D., Chambers, J. (2001). *Recurrent Neural Networks for Prediction: Architectures, Learning algorithms and Stability*. Wiley.
- Reilly, D.L., Cooper, L.N., Elbaum, C. (1982). A Neural Model for Category Learning. *Biological Cybernetics* 45, 35-41.

Table 1. Details of the systems studied

Sl.no	systems	Temperatures	Net work topology	Training algorithm
1	Anisaldehyde + chlorobenzene	303.15, 313.15 and 323.15 K	Feed forward multilayer perceptron	Back propagation learning algorithm
2	Anisaldehyde + bromobenzene			
3	Anisaldehyde + nitrobenzene			
4	Anisaldehyde + ethyl benzene			
5	Anisaldehyde + benzene			

Table 2. Input and output variables of method - 1 and method - 2

Method - 1		Method -2	
Input	Output	Input	Output
Mole fraction Experimental five Thermo physical properties	Predicted five thermo physical properties	Mole fraction Molecular weight	Predicted five thermo physical properties

Table 3. Random experimental verification of ANN predicted thermo physical properties (not included in the training data) of mixtures with percent standard deviation.

Systems	x_1	T/ K	P (pre)	S %	η (pre)	S %	n_D (pre)	S %	σ (pre)	S %	u (pre)	S %
Anisaldehyde and benzene	0.1500	303.15	0.890	0.28	0.825	0.48	1.509	0.37	37.21	1.72	621	0.20
	0.5000	313.15	1.000	0.98	1.214	0.42	1.524	1.40	64.21	0.58	998	0.47
	0.9000	323.15	1.100	0.99	2.610	0.91	1.544	0.81	77.61	0.85	1600	0.21
Anisaldehyde and chloro benzene	0.1500	303.15	1.108	0.85	0.945	0.95	1.558	0.89	47.91	0.65	1361	1.14
	0.5000	313.15	1.100	0.98	1.252	0.59	1.550	2.60	68.71	0.74	1121	1.27
	0.9000	323.15	1.110	1.14	1.999	0.98	1.546	1.96	88.21	0.12	1405	1.11
Anisaldehyde and bromobenzene	0.1500	303.15	1.421	2.18	0.988	2.48	1.558	3.04	44.21	2.04	1260	2.00
	0.5000	313.15	1.300	2.26	1.344	2.67	1.550	2.11	64.21	2.12	1364	2.54
	0.9000	323.15	1.155	2.18	2.125	2.98	1.547	2.25	88.21	2.79	1521	2.91
Anisaldehyde and nitrobenzene	0.1500	303.15	1.190	0.14	2.014	0.35	1.553	0.11	44.21	0.24	1461	1.02
	0.5000	313.15	1.154	0.62	2.125	0.49	1.548	3.28	61.21	0.18	1468	1.05
	0.9000	323.15	1.116	0.19	2.366	0.42	1.546	3.15	80.18	0.19	1547	1.00
Anisaldehyde and ethylbenzene	0.1500	303.15	0.922	0.85	1.320	0.11	1.510	1.11	42.21	0.00	1402	1.06
	0.5000	313.15	1.001	0.13	1.910	0.41	1.520	1.15	70.12	0.00	1457	1.07
	0.9000	323.15	1.077	0.61	2.301	0.20	1.540	0.01	92.95	0.04	1532	1.08

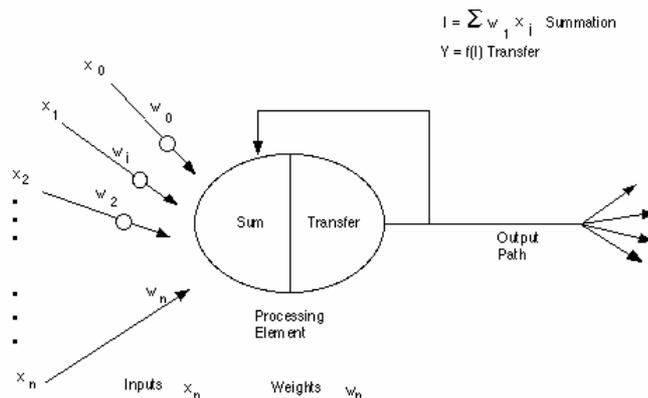


Figure 2.1 A Basic Artificial Neuron.

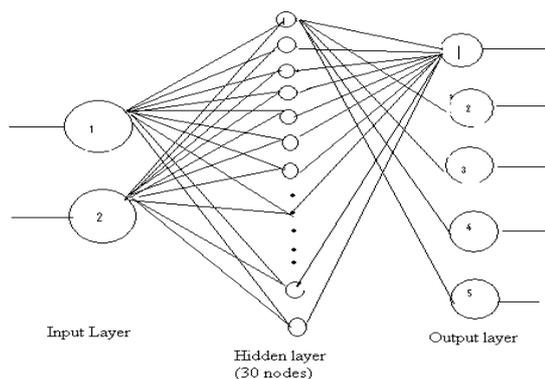


Figure 2.1.1 Feed forward net work

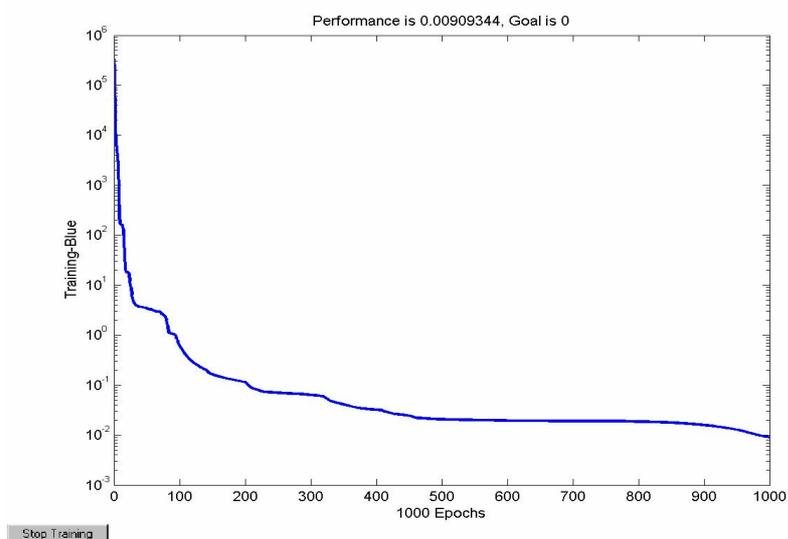


Figure 3.1 Plot of error and number of iterations - Error performance for anisaldehyde-chlorobenzene mixture

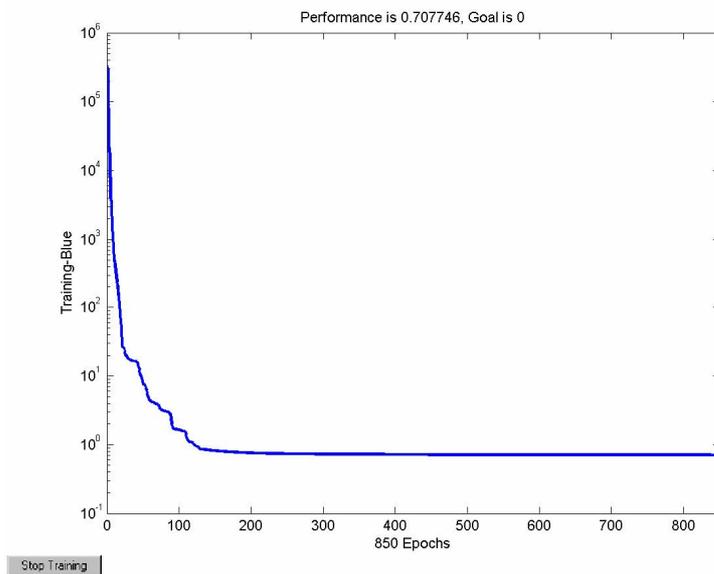


Figure 3.2 Plot of error and number of iterations - Error performance for anisaldehyde-chlorobenzene system with molecular weight.

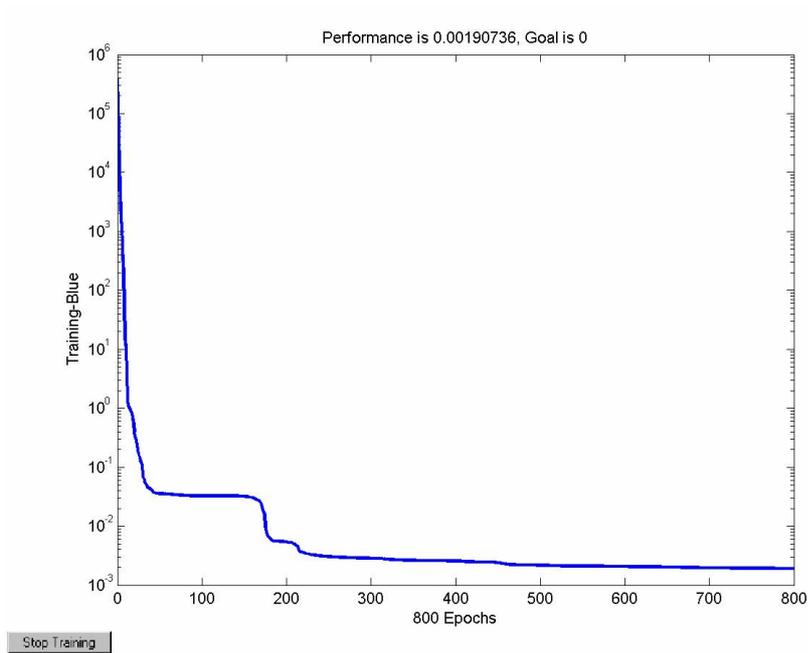


Figure 3.3 Plot of error and number of iterations - Error performance for anisaldehyde-bromobenzene system

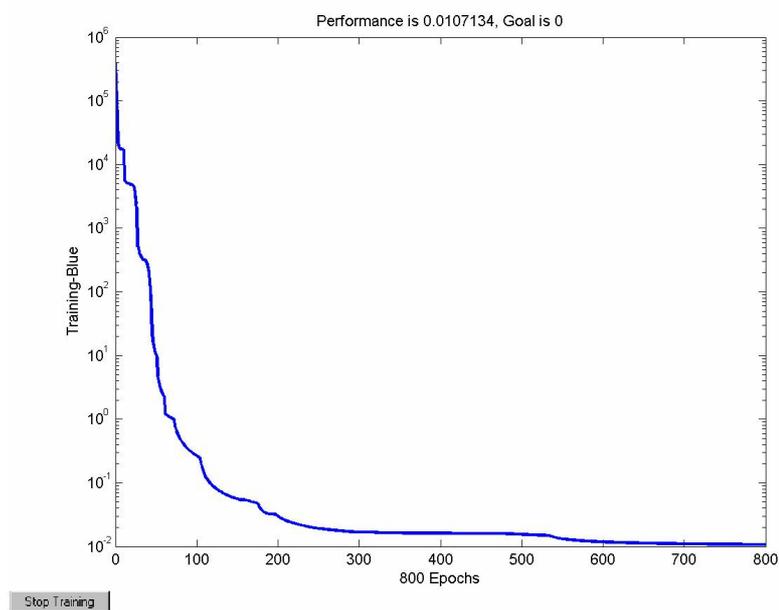


Figure 3.4 Plot of error and number of iterations - Error performance for anisaldehyde-bromobenzene system
with molecular weight

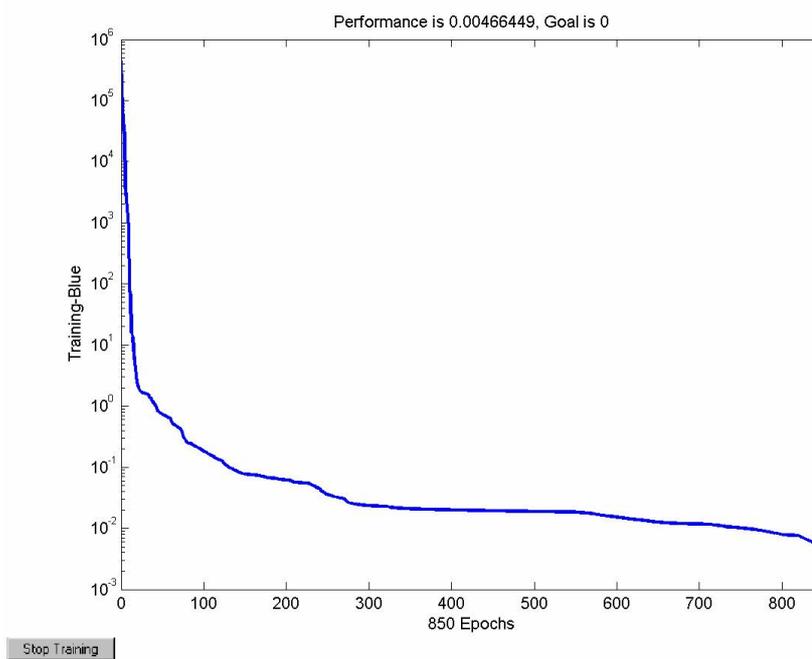


Figure 3.5 Plot of error and number of iterations - Error performance for anisaldehyde-nitrobenzene system

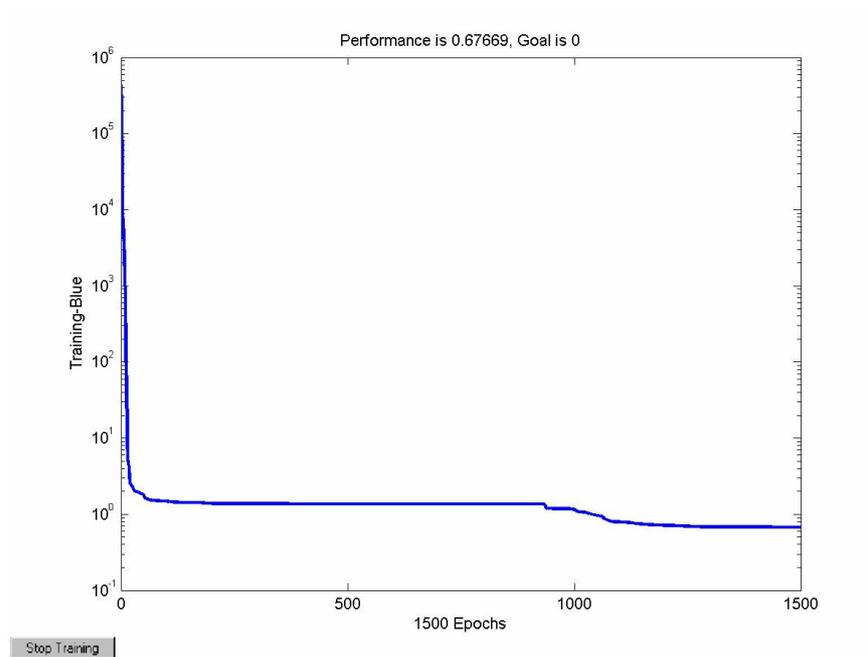


Figure 3.6 Plot of error and number of iterations - Error performance for anisaldehyde-nitrobenzene system with molecular weights

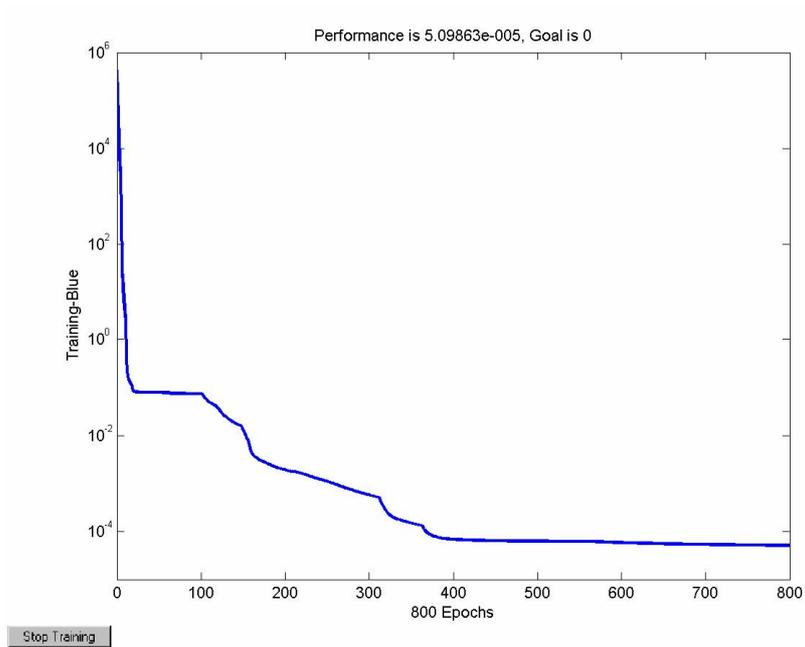


Figure 3.7 Plot of error and number of iterations - Error performance for anisaldehyde-ethylbenzene system

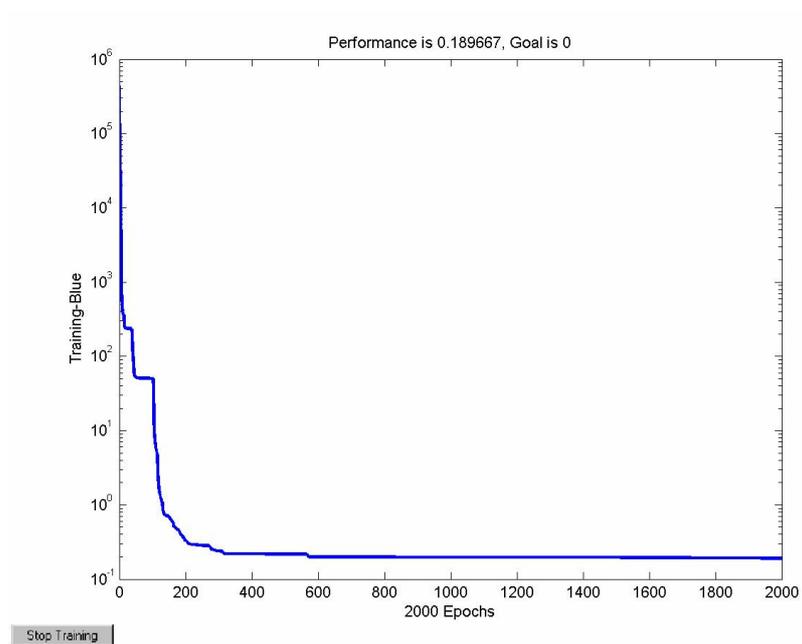


Figure 3.8 Plot of error and number of iterations - Error performance for anisaldehyde-ethylbenzene system with molecular weights

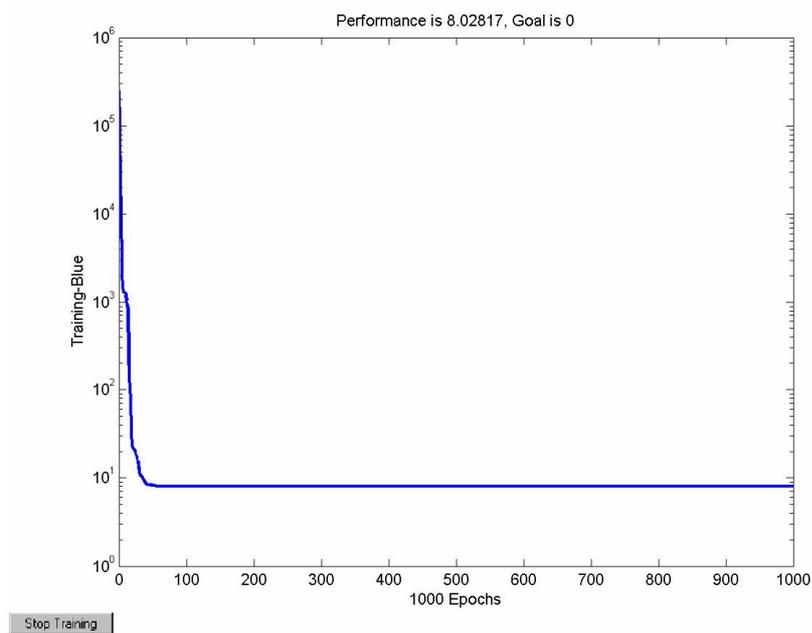


Figure 3.9 Plot of error and number of iterations - Error performance for anisaldehyde-benzene system

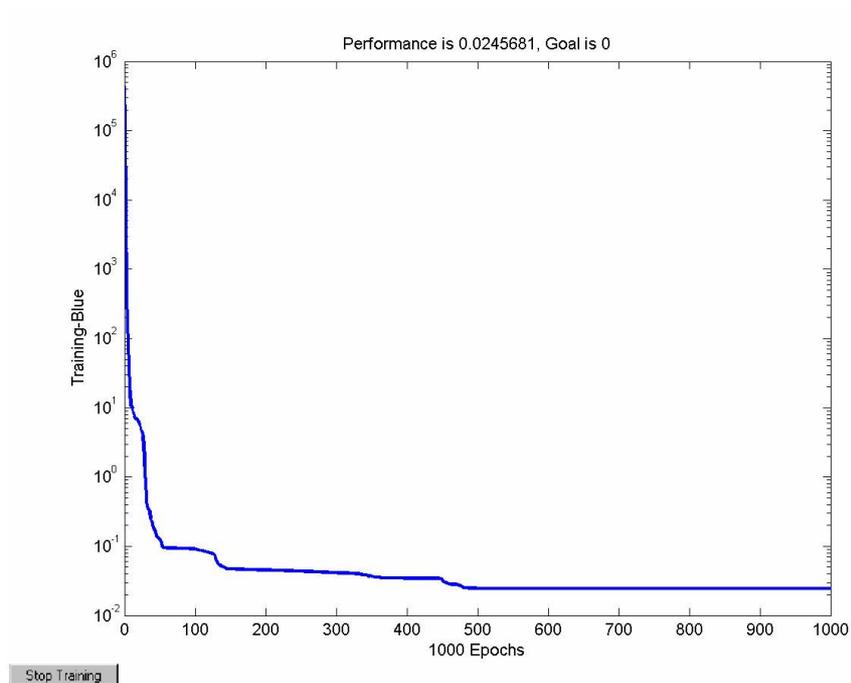


Figure 3.10 Plot of error and number of iterations - Error performance for anisaldehyde-benzene system with molecular weights



Computer Aided Investigation towards the Wind Power Generation Potentials of Guangzhou

Gang Yang, Yongxian Du & Ming Chen

Institute of Solar Energy System

Key Laboratory of Guangdong Educational Department

School of Physics and Engineering

Sun Yat-sen University

Mingde Garden 9-322, University City, Guangzhou 510006, China

Tel: 86-20-3933-0895 E-mail: ee04yg@sysu.edu.cn

The research is financed by the Chinese national 863 plan projects. No. 2006AA05Z409

Abstract

Wind power generation potential of Guangzhou has been analyzed based on the wind speed data collected in the recent 30 years. In order to make the results more accurate, the Weibull distribution function is introduced and the two parameters can be identified by CAD method together with local weather data available. By referring to the power characteristics of the wind turbines, the monthly and yearly operating probabilities and the output power under different heights are analyzed. Given by the above data and the power generation cost, it is proved that although the wind power resources in Guangzhou is not very ideal, it can also be utilized to generate power.

Keywords: Weibull distribution function, CAD, Height, Power characteristic, Power generation cost

1. Introduction

Guangzhou is one of China's leading economic cities and a massive manufacturing center. Along with the rapid development of economy, the demand of energy, especially electricity, is also growing at a fast speed. As the conventional energy generation will inevitably cause the emissions of poisonous gas and the greenhouse effect, it is necessary to ensure rational utilization of renewable energy to help protect the environment. Wind power generation, as an important part of the renewable energy, is put emphasis on in the recent years all over the world. However, the study towards the wind power generation potential of Guangzhou is very limited. This paper aims to analyze the wind power generation potentials of Guangzhou by using the collected weather data and wind turbine characteristics.

2. Models for wind power potential analysis

2.1 Weibull distribution function

The wind frequency distribution is used to determine the probability under which one wind speed occurs. Among the many different ways used to build such a distribution, the Weibull distribution function, as expressed by Eq. (1), is proved to be a simple but efficient way to fit the real wind speed distribution. The two adjustable parameters, k and c can be identified by using the average wind speed \bar{v} and the maximum wind speed v_{max} .

$$f(v) = \frac{k}{c} \left(\frac{v}{c} \right)^{k-1} \exp \left[- \left(\frac{v}{c} \right)^k \right] \quad (k > 0, v > 0, c > 1) \quad (1)$$

$$c = \frac{\bar{v}}{\Gamma(1+1/k)} \quad (2)$$

where $f(v)$ is the Weibull distribution probability of wind speed, c is the Weibull scale parameter and k is the dimensionless Weibull shape parameter.

The wind speed data is measured hourly and in a time segment t , the probability that the maximum wind speed v_{max} occurs is expressed by Eq. (3) and the result is expressed by Eq. (4).

$$P(v \geq v_{\max}) = \exp \left[- \left(\frac{v_{\max}}{c} \right)^k \right] = \frac{1}{T} \quad (3)$$

$$\frac{v_{\max}}{\bar{v}} \Gamma \left(1 + \frac{1}{k} \right) = (\ln T)^{\frac{1}{k}} \quad (4)$$

where $\Gamma(1+k) = \int_0^{\infty} x^k e^{-x} dx$; $T = 24dv_{\max}$; d is the number of days per month.

The parameter k can be calculated by the average wind speed \bar{v} and the maximum wind speed v_{\max} while the parameter c can be determined by the Eq. (2). The Eq. (4) is a transcendental equation which can be calculated by using the computer-aided method, shown in the Figure 1. Eq. (4) can also be calculated by using the iterative method which is very complex for the hand-calculation. It can be seen from the Figure 1 that the calculated parameter k is exact until the 3rd decimal place.

2.2 Wind speed adjustment

Wind speed near to the ground changes with height following the exponential law, thus the wind turbine hub height has a great influence on the amount of energy generated by the wind turbines. To calculate the output power of wind turbine, it is necessary to convert the measured wind speed to the wind speed at the wind turbine hub height. The most frequently used expression is the power law, expressed in Eq. (5).

$$v = v_0 \left(\frac{h}{h_0} \right)^{\alpha} \quad (5)$$

Where v is the wind speed at hub height h compared to the sea level, m/s; v_0 is the wind speed measured at the height h_0 compared to the sea level, m/s; α is the ground surface friction coefficient, while it is about 0.107 at the sea surface, 0.146 at the weald and 0.25 at the town.

2.3 Two meaningful wind speeds

As the shape and scale parameter of Weibull distribution can be calculated, two meaningful wind speeds, the most probable wind speed v_{MP} and the wind speed carrying maximum energy v_{MaxE} can be obtained by the Eq. (6) (7).

$$v_{MP} = c \left(\frac{k-1}{k} \right)^{\frac{1}{k}} \quad (6)$$

$$v_{MaxE} = c \left(\frac{k+2}{k} \right)^{\frac{1}{k}} \quad (7)$$

v_{MP} denotes the most frequent wind speed for a given wind distribution function and v_{MaxE} represents the wind speed that carries the maximum amount of wind energy.

2.4 Power output characteristics of wind turbines

According to the wind energy conversion theory and the Bates theory, the amount of power that one wind turbine actually obtains is summarized in the following Eq. (8):

$$P = \frac{1}{2} C_p \rho S v^3 \quad (8)$$

where P is the amount of power that one wind turbine actually generates, W; C_p is the wind power coefficient, with a limit value of 0.593; ρ is the air density, kg/m³; S is the sweeping area of the wind wheel, m²; v is the wind speed, m/s;

The Eq. (8) points out the relationship between wind speed and the power of wind turbines. The output power characteristic curve of wind turbines can be regarded as a 3rd power curve with constant parameters before reaching the maximum power point. This is also to say that $C_p \rho S$ can be regarded as a constant. After the wind turbine reaches its maximum output power, the sweeping area of the wind wheel changes as a result of the change of the blade angle. Thus the output power of wind turbine stops increasing and begin to remain steady or a slightly decreasing. The output characteristic of wind turbines can be expressed in the Eq. (9).

$$P(v) = \begin{cases} 0, (v \leq v_{ci}) \\ \varepsilon v^3, (v_{ci} \leq v \leq v_m) \\ P_m, (v_m \leq v \leq v_{co}) \\ 0, (v \geq v_{co}) \end{cases} \quad (9)$$

where $\varepsilon = \frac{P_R}{v_R^3}$, $v_m = \sqrt[3]{\frac{P_m}{\varepsilon}}$; v_{ci} is the cut-in wind speed, m/s; v_{co} is the cut-out wind speed, m/s; P_R is the rated

electrical power of the wind turbine, W; v_R is the rated wind speed, m/s; P_m is the maximum output electrical power, W; v_m is the wind speed that corresponds to the maximum output power, m/s;

With the output power characteristic of the wind turbine and the Weibull distribution probability of wind speed, the energy that a wind turbine produces can be obtained by using the Eq. (10)

$$\int_{v_{ci}}^{v_{co}} P(v)f(v)dv \quad (10)$$

3. Analysis of Guangzhou's wind power generation potential

3.1 wind speed data of Guangzhou

The wind speed data collected during the recent 30 years in each city can be found online at the ‘‘China Meteorological Data Sharing Service System’’. The weather station of Guangzhou locates at a height of 41 meters and it show that the average wind speed is 1.7m/s and the average maximum wind speed in the 30 years is 19.1 m/s. Figure 2 shows the above data collected from that weather station in recent 30 years.

As shown in the Figure 2, the average wind speed in every month is among 1.5-2.0 m/s which is a low value and there is no significant differences of the average wind speed among different months. The average wind speed is low in June, July and September while the maximum wind speed in the three months is larger as a result of the frequent typhoons in Guangzhou. That is also why there are significant fluctuations of the maximum wind speed each year. In winter, to the contrary, the average wind speed is larger and the wind resources are relatively more abundant.

3.2 Weibull probability density function

The parameters of Weibull distribution can be obtained by the wind speed data available. The yearly wind probability density function of Guangzhou can be calculated by using the Eq. (1) (2) (3) (4) and the result is shown in Figure 3.

As is shown in Figure 3, the most probable wind speed occurs at the wind speed less than 1m/s which represents that the wind energy resources in Guangzhou is not an ideal resource for power generation. Using the equation (6) (7), the monthly Weibull function parameters are shown in Table 1.

3.3 Operating probability of wind turbines

The Weibull distribution density function gives the probability of the wind speed exceeding the value u and between u_1 and u_2 , given in the Eq. (11) (12).

$$F(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right] \quad P(v \geq u) = \exp\left[-\left(\frac{u}{c}\right)^k\right] \quad (11)$$

$$P(u_1 \leq v \leq u_2) = \exp\left[-\left(\frac{u_1}{c}\right)^k\right] - \exp\left[-\left(\frac{u_2}{c}\right)^k\right] \quad (12)$$

Figure 4 shows the probabilities that the wind speed exceeding a certain value when the wind turbine hub height is 41m compared to the sea level in Guangzhou. For most wind turbines, the cut-in wind speed is about 3m/s which signify that the wind turbine will not generate power unless the wind speed exceeds 3m/s. It is usually that the cut-out wind speed is about 20-25m/s, and the probability of wind speed exceeding 20m/s can be ignored. If 3m/s and 20m/s are regarded as the cut-in wind speed and cut-out wind speed, the operating probability of wind turbines in Guangzhou is 16.96% which signifies that the wind turbine will work 1486 hours/ year. As wind speed is closely related to the wind turbine hub height, the operating probabilities of wind turbines under different heights can be obtained by the Eq. (5) and the $\alpha=0.146$. The operating probabilities of wind turbines under different heights in Guangzhou are shown in Figure 5.

As can be seen from Figure 5, the operating probabilities of wind turbines increase along with the increase of the wind turbine hub heights. In the wind turbine hub height of 200m, the operating probability rises from 0.1696 to 0.2490

which signifies that the operating hours rises to 2182 hours/year.

3.4 Analysis of energy generated by wind turbines

One frequently used wind turbine FD2.5-300, produced in the Inner Mongolia, is used as an example to illustrate the power generation situation. The cut-in wind speed, the cut-out wind speed and the rated wind speed of the wind turbine FD2.5-300 is 3m/s, 20m/s and 8m/s, respectively. The rated output power is 300W and the maximum output power is 500W. By using the Eq. (9) (10), the energy generated by such a wind turbine per month in Guangzhou at the wind turbine hub height of 41m is calculated, shown as in the Table 2.

From the Table 2, it is obvious that a yearly energy of 83.195 kW·h is generated by the wind turbine FD2.5-300 in Guangzhou; however, that number differs from the wind turbine hub heights. Figure 6 shows the energy generated by the wind turbine changes along with the wind turbine hub heights.

The wind turbine hub height has a great influence on the energy generated by the wind turbine, as shown in the Figure 6. As a result of the increase of the wind turbine hub height, the wind speed rises and the operating probability also tends to be larger. When the height reaches 200m, the energy that generated by the wind turbine will be 161.382 kW·h, nearly twice than that of at 41m.

3.5 The estimation of the generation cost

The cost of a wind turbine FD2.5-300 is about 2200 RMB and the typical service life of a wind turbine is 20 years, thus the wind power generation cost can be obtained by the following equation.

$$m = \frac{M}{yW} \quad (13)$$

Where M is the total investment of a wind power system; y is the lifetime of a wind turbine; W is the energy generated by the wind turbine per year.

If the cost for batteries and further maintenances are not counted into the investment of a wind power system, the wind generation cost under different heights are analyzed and shown in Table 3.

The data of Table 3 shows that the height of wind turbine hub is a closely related factor to the energy generated by the wind turbine. At the height of 41m, the generation cost is 1.322 RMB/kW·h. This figure falls to 0.68 RMB/kW·h at the height of 200 which is near to the cost of the commercial power.

4. Conclusions

This paper analyzed the wind power generation potential of Guangzhou. The Weibull distribution density function shows that the most probable wind speed of Guangzhou is lower than 3m/s and the wind energy resources in Guangzhou is not very ideal to generate power. The analysis of the operating probabilities proves that this figure remains between 0.17-0.25 which signifies that the operating hours per year is among 1500-2200 hours. The wind turbine hub height is proved to have a great influence on the energy generated by the wind turbine. The output power of wind turbines grows along with the rises of the height and the power generation cost decreases dramatically. The generation cost under a wind turbine hub height of 200m is close to the cost of commercial power. Thus although the wind resources of Guangzhou is not very ideal, it can also be utilized. As wind power is proved to be a kind of green and clean resource and has no pollution that will be produced during the conventional power generation process, moreover, the wind energy in the offshore areas is better than the inland, thus the Guangzhou's wind energy also has a non-negligible utilization value.

References

- Ai, Bin, Yang, Hongxing, Shen, Hui & Liao, Xianbo. (2003). Optimum sizing of PV/wind hybrid system (1) CAD method. *Acta Energetica Sinica*, 4, 540-547.
- E.Kavak Akpinar & S.Akpinar. (2005). An assessment on seasonal analysis of wind energy characteristics and wind turbine characteristics. *Energy Conversion and Management*, 46, 1848-1867.
- Lu, Lin & Yang, Hongxing. (2001). Wind data analysis and a case study of wind power generation in Hong Kong. *Wind Engineering*, 2, 115-123.
- Lu, Lin, Yang, Hongxing & John Burnett. (2002). Investigation on wind power potential on Hong Kong islands-an analysis of wind power and wind turbine characteristics. *Renewable Energy*, 1, 1-12.
- Wang, Changgui, Cui, Rongqiang & Zhou, Huang,. (2003). *Generating electricity technology by new energy*. Beijing: China electric power press. Pp. 177-189.
- Xu, Weimin, Zhang, Xinglin & Kong, Xinhong. (2007). Comparison for three methods to calculate Weibull distribution parameters of wind speed. *Jiangxi Electric Power*, 1, 1-3.

Zhou, Wei, Yang, Hongxing & Fang, Zhaozhong. (2006). Wind power potential and characteristic analysis of the Pearl River Delta region. *Renewable Energy*. 6: 739-753.

Table 1. The monthly Weibull function parameters of Guangzhou

Month	k	c	v_{MP} (m/s)	v_{MaxE} (m/s)
January	1.292	1.838	0.581	3.790
February	1.074	1.748	0.145	4.652
March	1.280	1.727	0.527	3.601
April	1.172	1.690	0.329	3.952
May	1.145	1.784	0.294	4.310
June	1.125	1.879	0.266	4.660
July	1.036	1.928	0.076	5.439
August	1.000	1.451	0.000	4.973
September	1.000	1.391	0.000	5.595
October	1.007	1.805	0.013	5.350
November	1.115	1.874	0.244	4.709
December	1.252	1.826	0.507	3.915

Table 2. Energy generated by wind turbine FD2.5-300 per month

Month	Energy generated per month (kW·h)	Month	Energy generated per month (kW·h)
January	4.743	July	11.113
February	6.679	August	6.965
March	3.863	September	8.065
April	4.719	October	10.019
May	6.368	November	7.890
June	7.742	December	5.102

Table 3. The energy generated under different wind turbine hub heights

Heights (m)	Energy generated per year (kW·h)	Generation cost (RMB/kW·h)
41	83.195	1.322
80	111.041	0.991
120	131.314	0.838
160	147.368	0.746
200	161.382	0.680

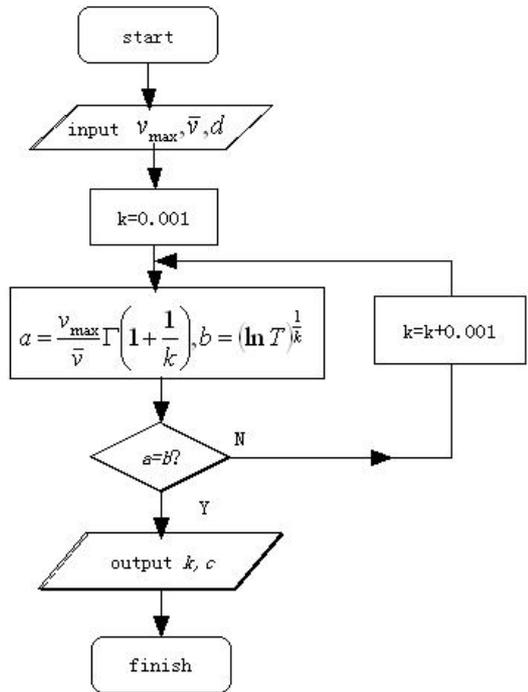


Figure 1. The flow chart of identifying the Weibull parameter k and c

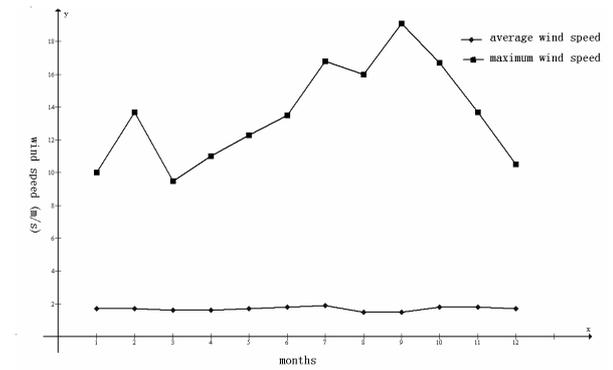


Figure 2. The average wind speed and the maximum wind speed in Guangzhou.

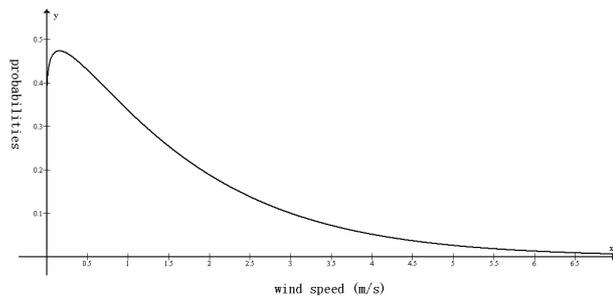


Figure 3. Yearly wind probability density function

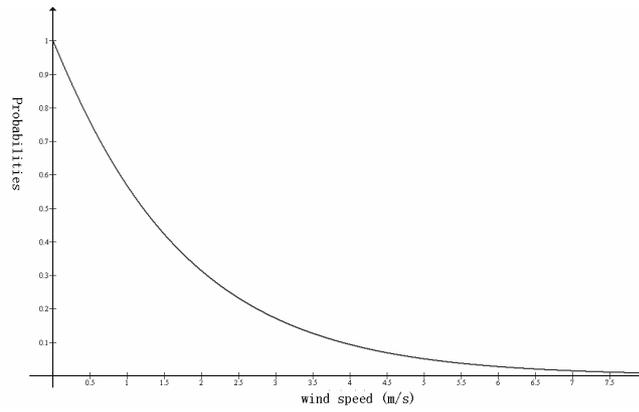


Figure 4. Probabilities of the wind speed exceeding a certain value.

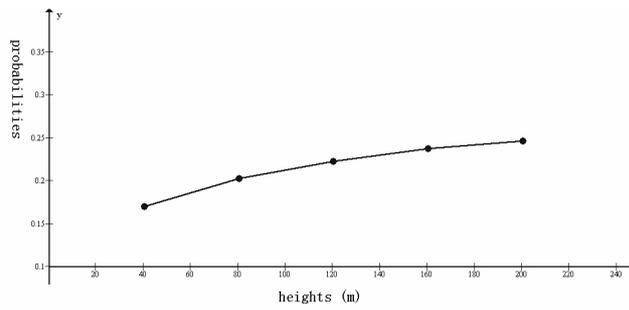


Figure 5. The operating probabilities of wind turbines in Guangzhou at different heights.

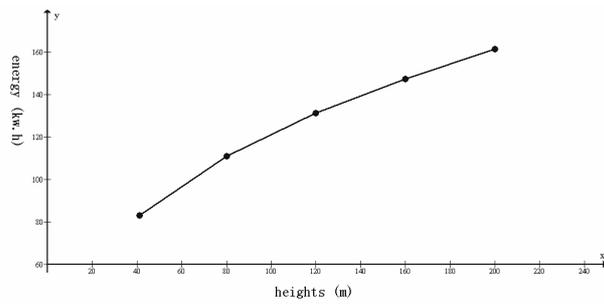


Figure 6. The yearly energy generated by the wind turbines in Guangzhou at different hub heights.



Autonomous Notification and Situation Reporting for Flood Disaster Management

Ku Ruhana Ku-Mahamud (Corresponding author)

College of Arts and Science

Universiti Utara Malaysia

06010 Sintok Kedah, Malaysia

Tel: 604-9284717 E-mail: ruhana@uum.edu.my

Norita Md Norwawi

College of Arts and Science

Universiti Utara Malaysia

06010 Sintok Kedah, Malaysia

Tel: 604-9284788 E-mail: nmn@uum.edu.my

Norliza Katuk

College of Arts and Science

Universiti Utara Malaysia

06010 Sintok Kedah, Malaysia

Tel: 604-9284660 E-mail: k.norliza@uum.edu.my

Safaai Deris

Faculty of Computer Science and Information Science

Universiti Teknologi Malaysia

83130 Skudai Johor, Malaysia

Tel: 607-5532343 E-mail: safaai@utm.my

The research is financed by MOSTI, IRPA grant No. 04-02-07-10040-EAR

Abstract

Failed management in emergency response effort due to inefficient and ineffective operation influences the adoption of information and communication technology specifically the Internet technology and World Wide Web. Timeliness, rapid response and efficient information dissemination are crucial in emergency management. This paper proposes a software design architecture taking advantage of the intelligent agent technology to facilitate autonomous notification and auto generates situation report. The adoption of agent technology has made it possible to produce scheduled report with standardized format that can reduce duplications and redundancies of information.

Keywords: Disaster management system, Intelligent agent, Emergency communication, Integrated reporting

1. Introduction

Emergency situations resulting from natural disasters such as flood, tsunami and earthquake represent complex and dynamic environments with high level of uncertainty. These are critical situations where human lives and properties are at stake. With the advancement of information and communication technology (ICT) especially widespread use of Internet and World Wide Web (WWW), improvement on the efficiency and effectiveness of the management of emergency is

possible. The emergency response operation can be enhanced through better coordination and communication among involved agencies by the adoption of ICT, in particular through the use of artificial intelligence which is suitable for dynamic, unpredictable and uncertain situations.

Emergency management in general consists of various phases such as preparation, mitigation, response, recovery and learning. In the preparation phase, a warning system is usually designed with mechanism for threat recognition, threat evaluation and alerting mechanism (Norwawi, 2004). However, there are reports on failure and inadequacy of warning systems which is self-defeating (Chan, 1997). Inefficient coordination among authorities and warnings which are issued on an ad-hoc basis are the main reasons for the warning failures (Chan, 1997).

Fedra (1999) developed a decision-support system (DSS) using high performance and networking technologies to improve the planning and operational decision-making process where the requirements are timeliness, accurate, directly understandable, usable and easily obtained. A similar network approach was also adopted by Bui and Sankaran (2001) taking advantage of the ICT specifically Internet, inter-platform connectivity software, group decision theoretical progress where large number of data can be processed, facilitating faster decision-making and high level of objectivity in decisions made. They proposed the concept of virtual information centre with a web centric management system.

Some examples of the application of the Internet technology specifically the WWW, is the South Asia Flood website which provides the platform in information sharing for the Hindu Kush Himalayan Region aimed at promoting regional cooperation (Shrestha, 2002). Hospitals were also reported using web based disaster management system that is accessible from desktop, laptop and mobile computers located at place designated as the command and control centre (McGee, 2005).

Taking into consideration that stress and pressures could influence and deteriorate human decision-making capability, an interactive assisted technology could offer significant benefit through its ability to give guidelines and recommendations to emergency managers. Iba and Gervasio (1999) claimed that an assisted technology has better advantage than an automation technology. Human expert is still a very important element in the decision system being a part of a socio-technical system linking people, organization, machine and computers (Norwawi, 2004). This is made possible through the application of software agents which is a computational entity that can be implemented in complex system due to its special characteristics such as autonomous, proactive, reactive, adaptive and sociable. A society of autonomous agents or better known as multiagent system can be developed where the agents can organize themselves in achieving a common goal. Norwawi (2004) discussed the matching characteristics between emergency environment and multiagent system which includes the multi organization and level of jurisdiction. These similar characteristics enable the multiagent system to be a natural representation of the roles in an emergency management.

Multiagent environment consists of multiple software agents working together to accomplish specific objective. An agent can be described as a computer program that performs some information gathering and holds very well-defined objectives of its creation. In computer science, agent has a very significant role especially in performing complex tasks in an application. Agents must be aware of their own capabilities and of changes to other agents and their environment. To remain effective, agents must be able to adapt their structures and knowledge while they execute. In order to achieve some objectives, systems employ a set of agents that are autonomous, proactive, adaptive, and reactive (Guessoum & Briot, 1999; Wooldridge & Jennings, 1995). Each agent has its own competencies and knowledge, but it needs to interact with other agents to solve complex problems, avoid conflicts, acquire and share information, and so on (Guessoum, 2004). The intelligent characteristics own by the multiagent will ensure that the complex tasks in flood management can be implemented in the most efficient way. The multiagent plays a very important role especially in predicting flood and notifying related people in flood management. During the emergency, fast and immediate response are highly needed to assist and evacuate victims. It also interacts and communicates with each other so that the functions can be performed faster.

Due to the above discussion, this study has develop a web-based flood management system that integrates intelligent features such as autonomous notification and auto-generate situation report that can be scheduled accordingly by the emergency management authority. This work is possible by adopting an intelligent agent technology, a software agent that can decide when to notify, to integrate reports from various agencies, auto-generate the report and sending it to all parties involved. The main aim of this paper is to present the software architecture for the notification and situation reporting embedded in a web based flood management system.

2. Web based flood management architecture

This study has adopted a multi methodological approach integrating the design research methodology (Hevner et al., 2004) with knowledge engineering. The first major step in KE specifically the knowledge acquisition activity consists of four major steps: planning of the knowledge acquisition sessions, extraction of knowledge, analyze the knowledge acquired and knowledge verification with the domain experts (Liou, 1992).

The state of Kedah in Malaysia is taken as a case study since flood is the most common natural disaster and frequently occurs during the wet season. In Malaysia a flood management committee (FMC) is formed at the district, state and national level depending on the magnitude of the flood. In this study, interviews with the agencies involved such as the National Security Division (BKN) and Social Welfare Department, were conducted to get an overview of the flood management operation. Among the problems highlighted were the difficulties in managing and distributing reports from various agencies involved to those interested parties. The cause of this difficulties were the use of facsimile machine, non-standardized report formatting, delay, duplication and even inconsistencies in the reported items. The second problem highlighted was that information dissemination among public is not well distributed.

The analysis and design for the web-based flood management system have been accomplished by following standard notation for object-oriented graphical modeling. This research has produced use case diagrams, activity diagrams, collaboration diagrams, sequence diagrams and class diagrams. All diagrams which have been mentioned representing different behavioral and structural of the system. The web-based flood management system consists of six use cases and three actors. Figure 1 illustrates the general use case diagram for web-based flood management system which has been design using Rational Rose 2000 software. The flooding phases are represented by use cases as depicted in Table 1.

Multiagent concept has been utilized in during pre-flood phase whereby two agents have been used for the purpose of assessing potential flooding and notifying the flood committee members. An agent has also been used for flood report generation during flooding and report generation after the flooding. For the purpose of assessing the potential flooding, the agent was given a name as Situation Assessment Agent and Notification Agent for agent responsible in sending notification.

3. The multiagent architecture

In this study, three agents have been initiated. The agents are for situation assessment, notification and report generating. Figure 2 shows the architecture which includes the interaction and communication between agents in the three phases of flooding. The process is initiated by the situation assessment agent that retrieves the hydrological data from the Drainage and Irrigation Department (DID) database. The agent will use current water level for each river in certain district and compare with the corresponding normal water level. Based on the water level and rainfall measurement, the agent can predict potential flooding. The agent monitors the situation for 24 hours a day. If there is a potential flooding, then a signal will be sent to the notification agent.

Figure 3 illustrates the situation assessment activity diagram. Upon receiving the message from situation assessment agent, the notification agent will create a message containing details regarding the location of potential flood. The message will be sent in the form of SMS and email. The two agents work together in the pre-flood phase. When a message is received by the Notification Agent, it will obtain information of corresponding flood committee members from the database. The phone numbers and the emails of the involved district flood committee members are used to notify the potential flooding. The Notification Agent will compose message in the form of SMS and email which contain the area of potential flooding and the nearest river water level reading. Next, the notification agent will also activate the virtual emergency operation centre EOC in order to allow the flood management committee to utilize the online flood management system as shown in Figure 4.

The Report Generation Agent is created for report generation. This agent will gather data from database and compile the report according to specific subject as required by users. Reports are presented in the form of Microsoft Office Document which later can be saved in hard disk and also can be directly printed. During flooding, the report generator agent will begin its activity once EOC is activated. This agent will gather information regarding the current flood occurrence and generate report in a printer-friendly form. The types of report generated in this phase will depend on the users' request. The report generated in this phase will be stored in the database and will be used to generate the formal post-mortem report. Figure 5 illustrates the interaction between agents with messages that need to be sent for the Situation Assessment Agent, Notification Agent and Report Generator Agent in this flood management system.

The autonomous notification and situation reporting functions have been integrated in web-based Flood Management System that has been developed using Active Server Page technology on the Microsoft platform. Figure 6 shows the screen shot of the system's main page.

4. Conclusion

Agent technology has demonstrated to be useful in facilitating faster notification and generating situational report through its autonomous, reactive and proactive characteristics. It helps in improving the emergency management operation through better communication and coordination via the Internet technology specifically the WWW. It will provide an anytime anywhere technology capability for the emergency managers and the public providing better access to information for the purpose of decision-making, planning and operating procedures. The agents used in this study have efficiently decreased the time in manual processes for notifying the related agencies of flooding situation.

References

Bui, T.X. & Sankaran S.R. (2001). Design considerations for a virtual information center for humanitarian assistance/disaster relief using workflow modeling. *Decision Support System*. 31(2): 165-179.

Chan, N.W. (1997). Increasing flood risk in Malaysia: causes & solutions. *MCB Disaster Prevention and Management*. 6(2): 72-86.

Fedra, K. (1999). Decision support and expert systems: implementation report. HITERM. Contract No. 22723, Technical Report. Environmental Software and Services (ESS).

Guessoum, Z. (2004). Adaptive agents and multiagent systems. *IEEE Distributed Systems Online*. 5(7): 1-4.

Guessoum, Z. & Briot J.P. (1999). From active object to autonomous agents. *IEEE Concurrency*. 7(3): 68-78.

Hevner, A.R., March S.T., Park J. and Ram S. (2004). Design science in information systems research. *MIS Quarterly*. 28(1): 75-105.

Iba, W. & Gervasio M.T. (1999). Adapting user preferences in crisis response. In *Proceedings of Intelligent User Interface*. 87-90.

Liou, Y.I. (1992). Knowledge acquisition: issues, techniques and methodology. *DATABASE*. 23(1): 59-64.

McGee M.K. (2005). New tools help hospitals handle terror attacks and other disasters. *Information Week*. April 14. Retrieved 8/10/2005 from <http://nformationweek.com>.

Norwawi, N.M. (2004). Computational recognition-primed decision model based on temporal data mining approach in a multiagent environment for reservoir flood control decision. PhD Thesis, Faculty of Information Technology, Universiti Utara Malaysia.

Shresta, M. (2004). Regional cooperation flood forecasting and information exchange in the hindu kush himalayan region. *Asian Disaster Management New*. 8(4). Oct-Dec.

Wooldridge, M.J. & Jennings N.R. (1995). Intelligent agents: theory and practice. *Knowledge Engineering Review*. 10(2): 115-152.

Table 1. Flooding Phase and Use Case

Phase	Use Case
Pre-Flooding	Manage Administration, Manage Preparedness Operation, Authenticate User
During Flooding	Manage Response Operation, Authenticate User
Post-Flooding	Manage Recovery Operation, Generate Flood Response Operation Report, Authenticate User

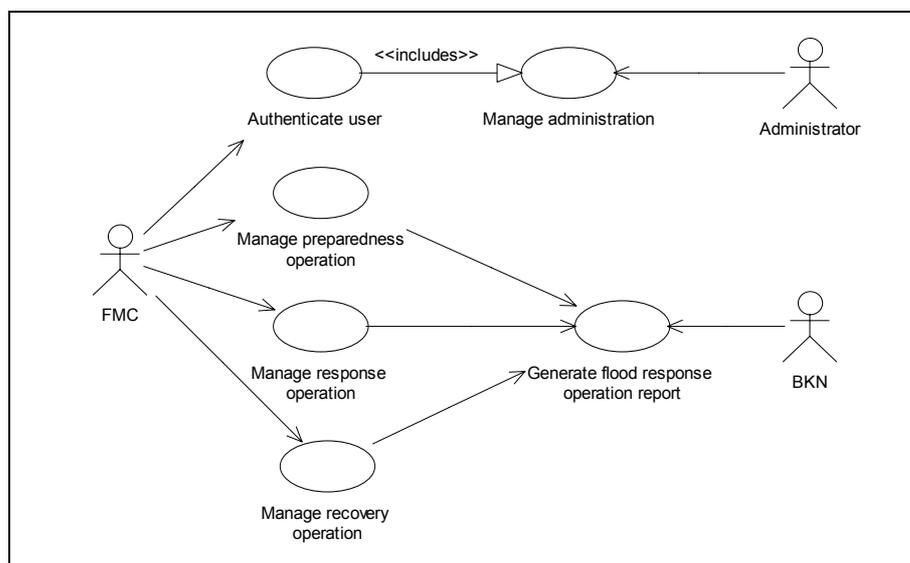


Figure 1. General Use Case Diagram for Web-based Flood Management System

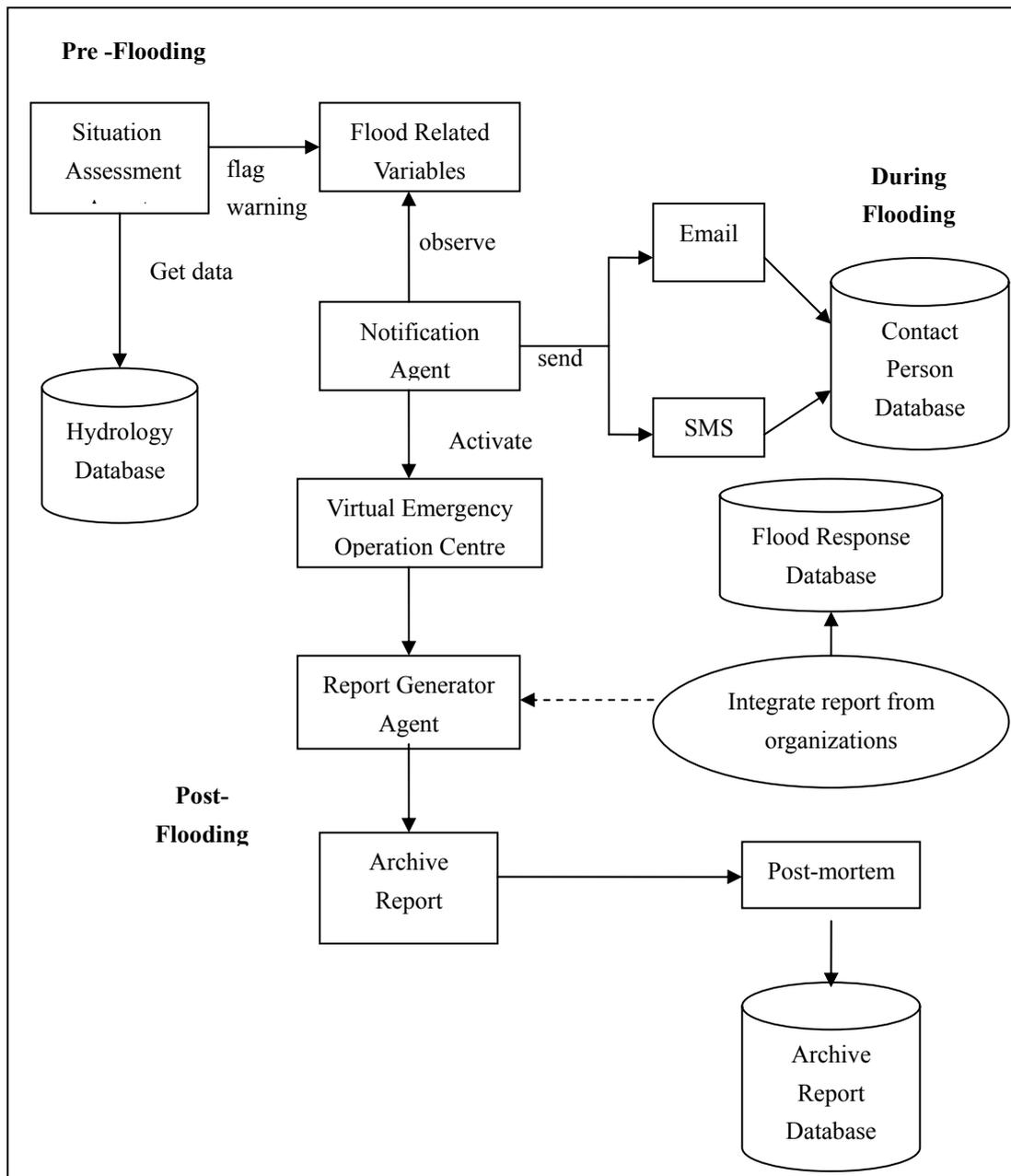


Figure 2. Multiagent Architecture for Flood Management System

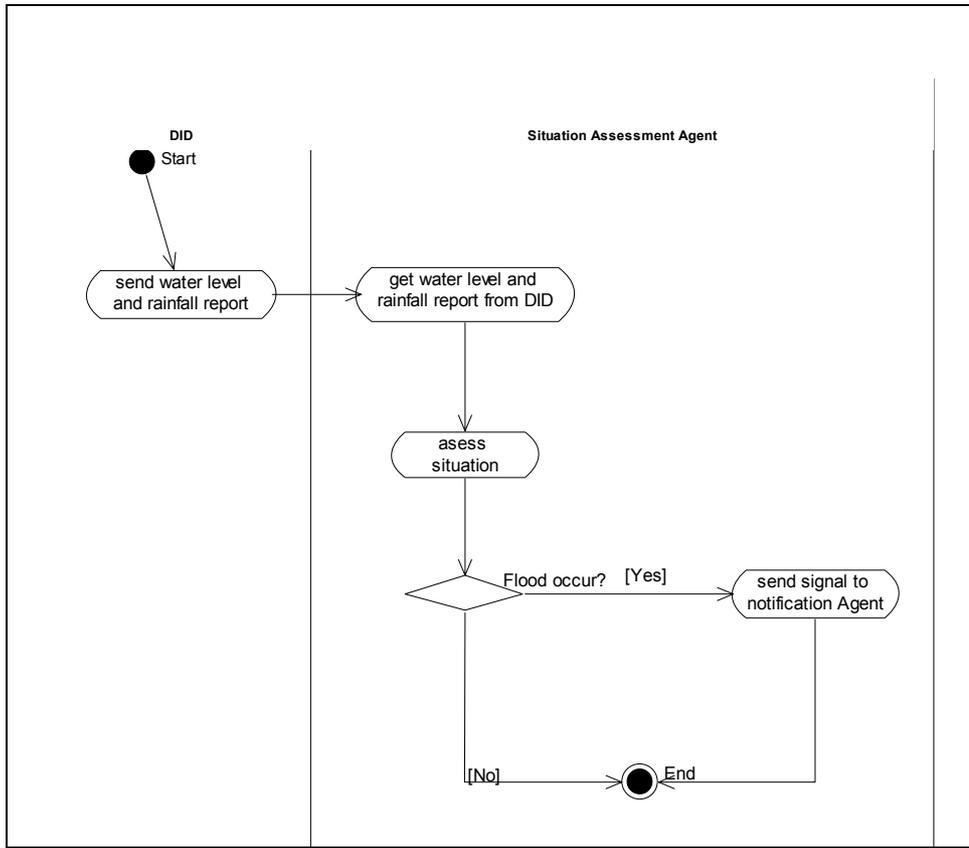


Figure 3. Activity Diagram for Situation Assessment Agent

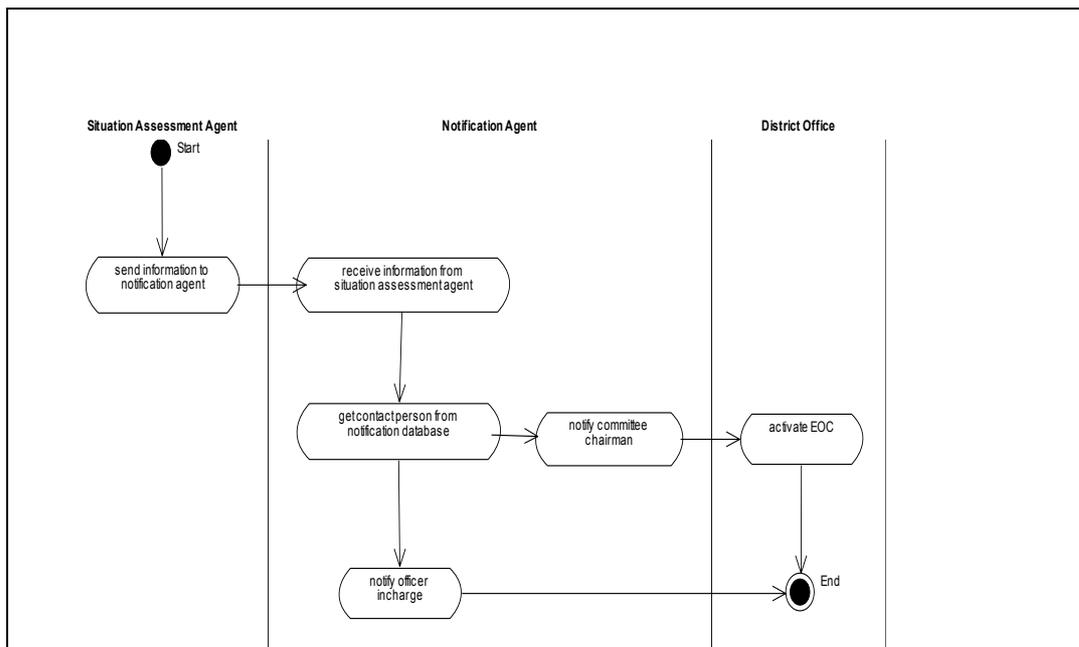


Figure 4. Activity Diagram for Notification Agent

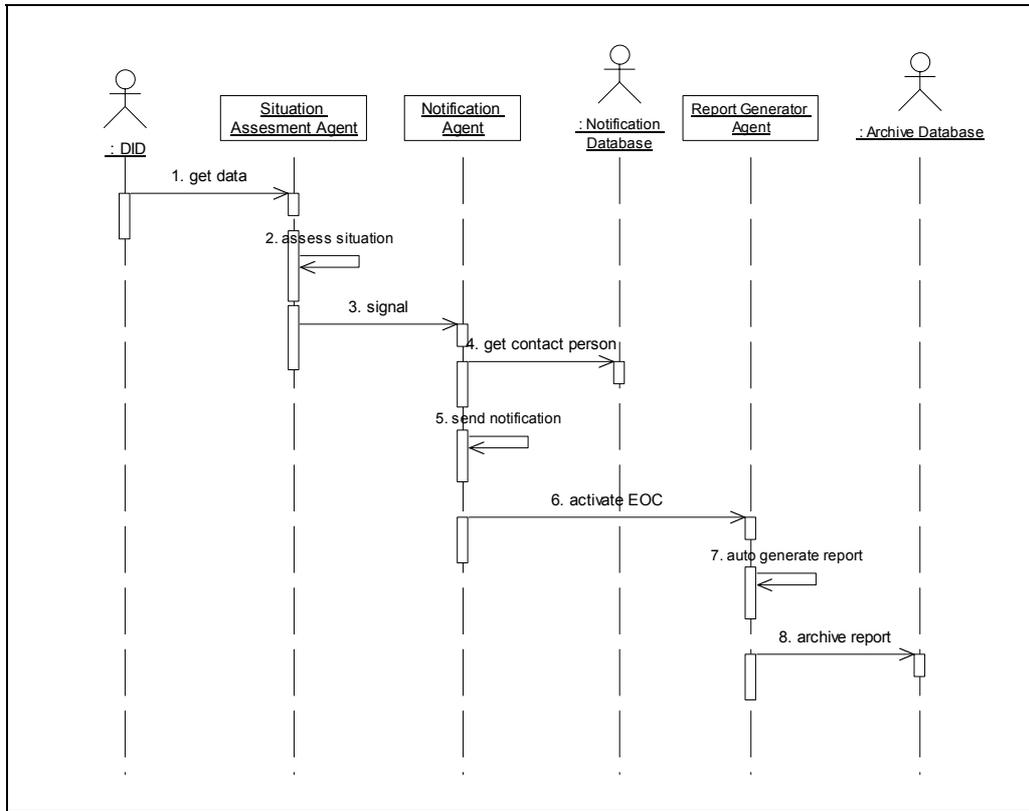


Figure 5. Sequence Diagram for Web-based Flood Management System



Figure 6. Screen shot for web-based flood response system



The Role of Session Border Controllers in the DMZ of Voice over IP(VoIP) Networks

Huihong Chen (Corresponding author)

School of Information Science Technology

Central South University

Changsha, Hunan 410083, China

E-mail: chenjin_2002@163.com

Zhigang Chen

School of Information Science Technology

Central South University

Changsha, Hunan 410083, China

Tel: 86-731-887-9636 E-mail: czg@csu.edu.cn

Abstract

SBCs usually sit between two service provider networks in a peering environment, or between an access network and a backbone network to provide service to residential and/or enterprise customers. They provide a variety of functions to enable or enhance session-based multi-media services (e.g., Voice over IP). This thesis analyzes the internet structure of SBC and shows single-box and dual-box SBCs in the DMZ and introduces firewalls and network address translation. And then explain how the SBC cooperates with the firewalls to ensure that VoIP signaling and media traverses the DMZ without compromising the security of the trusted network. And finally, describes topology hiding and bad protocol detection of other DMZ processing.

Keywords: VoIP, DMZ, SIP, SBC, NAT

Border Controllers (SBCs) have become an important element of modern Voice over IP(VoIP) networks, as service providers look to protect the integrity of their networks and business models while offering diverse services to their customers.

Most people would agree that an SBC is a kind of firewall for Voice over IP traffic. However, as soon as you start to look beyond this initial consensus, there is considerable disagreement as to what an SBC actually is, and what function it should offer! This is partly because SBC vendors are pushing out to cover a wide variety of niches in order to compete for market share, and partly due to the genuine range of scenarios where service providers are looking for solutions.

1. INTRODUCING SBCS AND DMZ

An SBC is a VoIP session-aware device that controls call admission to a network at the border of that network. Optionally (depending on the device), it can also perform a host of call-control functions to ease the load on the call agents within the network.

1.1 Internal structure of an SBC

An SBC device breaks down into two logically distinct pieces.

- The Signaling SBC function (SBC-SIG) controls access of VoIP signaling messages to the core of the network, and manipulates the contents of these messages. It does this by acting as a Back-to-Back User Agent (B2BUA).
- The Media SBC function (SBC-MEDIA) controls access of media packets to the network, provides differentiated services and QoS for different media streams, and prevents service theft. It does this by acting as an RTP proxy.

Some SBC devices offer both functions in a single box (referred to hereafter as single-box SBCs). Others take a distributed approach, and separate SBC-SIG and SBC-MEDIA onto separate machines (referred to hereafter as dual-box SBCs), using call control protocols such as H.248 and COPS-PR to link the two.

1.2 The Demilitarized Zone

The Demilitarized Zone (DMZ) is the conceptual term for a small subnetwork (or individual device) that sits between a trusted private network, such as a corporate private LAN, and an untrusted public network, such as the public Internet. Typically, the DMZ contains devices directly accessible to Internet traffic, such as web servers, FTP servers, or SBCs. The purpose of the DMZ is to prevent hostile or unwanted traffic from entering (or, in some cases, leaving) the private network.

2. DMZ PROCESSING

The DMZ is the “demilitarized zone” between two networks, as described in section 1.2, The Demilitarized Zone. This chapter provides more detail on the role played by SBCs in the DMZ.

2.1 Devices in the DMZ

All SBCs fall into one of the following two categories.

- Those that do not perform firewall processing in the DMZ, but instead rely on an external and internal firewall.
- Those that do perform firewall processing (i.e. that have a firewall on-board).

If an SBC does not perform firewall processing, then the DMZ looks like this

Figure 1. SBC does firewall processing

If the SBC is decomposed into the dual-box model, then the SBC-MEDIA device resides in the position indicated above, and the SBC-SIG device resides in the network core. In this case, SBC-SIG controls the firewall to allow VoIP signaling and media to pass through.

If the SBC does perform firewall processing, then the same DMZ looks like this:

Figure 2. SBC does firewall processing

If the SBC is decomposed into the dual-box model, then SBC-MEDIA implements a firewall that screens out unwanted signaling and media packets.

All SBCs typically incorporate NAT (Network Address Translator) function. The remainder of this section describes the role of firewall devices in the DMZ, and the role of the NAT component of the SBC.

2.1.1 The firewall

Firewalls prevent unwanted traffic from entering, or leaving, a network by performing basic packet filtering. Note that firewalls filter packets purely by examining packet headers, and do not parse or understand the payload of the packets. Therefore, they do not filter out all types of unwanted traffic. For example, firewalls do not perform Call Admission Control – SBCs do that. However, firewalls are valuable because they efficiently filter out large categories of unwanted traffic, leaving application-aware devices such as SBCs with much less work to do.

The external firewall in Figure 1 filters packets from the external network, but allows all packets from the internal network to pass through unfiltered. The internal firewall filters packets from the internal network, but allows all packets from the external network to pass through unfiltered (since they have already passed the external firewall).

Firewalls by default do not accept packets from the network, but are configured with rules that allow them to select and accept certain packets. Therefore, packets are admitted to (or from) the network based on *explicit configuration*, and not on default configuration. Firewalls are configured either

- by the network operator, using a human interface, *or*
- by trusted software, using an API.

There are no standards-defined APIs for configuring firewalls; however, the IETF’s MIDCOM working group is evaluating suitable protocols for this task. SNMP, RSIP, Megaco, Diameter, and COPS are all being considered. In addition, the MSF have made some steps towards defining their own protocol for firewall control (MSF2003.113.00 – Draft IA for RTP Proxy / FW Control Protocol).

2.1.2 The NAT

SBCs typically incorporate NAT function. NATs separate a network into distinct address spaces. In Figure 9, the NAT component of the SBC separates the internal network address space 10.1.0/24 from the external network address space 85.3/16. A few addresses from the 85.3/16 domain are used to represent all machines within the 10.1.0/24 domain, as described below.

The NAT maintains a table of mappings from {external address, port} to {internal address, port} and vice versa³. The table is a dual-index table, so a particular mapping can be looked up given either the internal or external addressing information. The NAT uses this table to rewrite the headers of the IP packets that it forwards.

- On receiving an IP packet from the external network, the NAT looks in its table for the destination address and port of the packet (which will be an address from the external address space). If a mapping is found, then the destination address header in the IP packet is changed to contain the corresponding internal address and port from the table, and the packet is forwarded towards the internal network. If no mapping is found, the packet is discarded.
- On receiving an IP packet from the internal network, the NAT looks in its table for the source address and port of the packet (which will be an address from the internal address space). If a mapping is found, then the source address header in the IP packet is changed to contain the corresponding external address and port from the table, and the packet is forwarded towards the external network. If no mapping is found, then a new mapping is created: the NAT dynamically allocates a new external address and port from the external address space for the packet (and all future packets from this source address and port tuple). Mappings in the table are created in one of two ways.
 - By packets traversing the NAT from the internal network towards the external network, as described in the second bullet point above.
 - By configuration, either from the network operator via a human interface, or programmatically from trusted software via an API.

2.2 How VoIP signaling packets traverse the DMZ

The NAT component of the SBC and the firewalls in the DMZ is configured (at start of day) as follows.

- The NAT is configured with a mapping that converts between the SBC's internal address (10.1.0.22 in Figure 9) and the port it uses for signaling, and some address and port taken from the external network's address space. This external address and port is used to identify the SBC in the public network. Packets sent from the external network and destined for the SBC are sent to this address and port.
- The external firewall is configured to permit IP packets whose destination address header contains the address and port that identify the SBC in the external network.
- The internal firewall is configured to permit IP packets whose destination address header contains the internal address of the SBC, and the port that it uses for signaling.

Note that this configuration could either come from human input, or from the SBC by programmatic API if the SBC and firewalls are collocated, or from the SBC by network protocol if the devices are separate.

This configuration allows all signaling packets addressed to the SBC to traverse the DMZ devices and reach the SBC, whether the packets originate from the internal or external network. In addition, it allows the SBC to send signaling messages towards either the internal or external networks.

This scheme relies on the fact that the external address and port that is used to identify the SBC on the public network for signaling is well-known to VoIP devices on the public network. Typically, this is achieved by using DNS records to associate this address with the SBC's hostname in the public network.

This scheme also relies on the SBC knowing its external IP address and port for signaling, because it must use these in the VoIP signaling headers that it sends in requests to the external network (as these fields are usually used to route the signaling response). This can be configured on the SBC.

2.3 How VoIP media packets traverse the DMZ

The situation with media packets is a little more complex than with signaling, because the media packets in a given call originate from, and are sent to, addresses and ports that are dynamically allocated by the RTP protocol when the call is established.

The SBC acts as a signaling Back-to-Back User Agent (B2BUA) and a media bridge, and so it terminates the media of a call on both the internal and external network sides. The ports that it uses to send and receive media on each side are allocated dynamically when the call is established.

Figure 3. VoIP media packets traversing the DMZ

This causes a couple of problems.

- The internal firewall needs to be configured to permit IP traffic sent to port 9900 on the SBC. Since 9900 is a dynamically allocated number, this must be done automatically during call set-up (i.e. without human intervention).
- The external firewall needs to be configured to permit IP traffic sent to the SBC's external address and port 12745. Again, this must happen automatically during call set-up.

If the firewall and SBC are on the same device, then these problems are easily overcome by implementing a programmatic interface that allows the SBC to dynamically configure the firewall software.

If the firewalls are separate from the SBC, then either

- the SBC dynamically configures the firewall over the network (although, as noted above, there are no standards for this at present), *or*
- the firewalls must be configured to permit all traffic sent to any port on the SBC (or at least, any port in the range used by the RTP protocol).

2.4 Other DMZ processing

SBCs also perform other DMZ-related processing, as described in the following sections.

2.4.1 Topology hiding

VoIP signaling messages convey information that can allow the recipient to determine both the internal topology of a network, and the route taken by a call across that network (and possibly out the other side). For example, the Via headers in SIP signaling messages carry this kind of information.

It is often undesirable to expose this information to users outside a network. For example, if you are a service provider who uses a second service provider to act as a carrier for your calls, you do not want to expose the identity of the carrier SP to your customers in case they approach the carrier SP directly for a better price.

To solve this problem, SBCs can remove sensitive information by rewriting the VoIP headers in the signaling messages that they send across the network boundary. SBCs achieve this by acting as B2BUAs. They terminate the VoIP signaling that they receive from within the private network, and signal a new call towards the public network. Since this is a new call, it does not require any of the routing information from the previous call (for example, none of the SIP Via headers are carried over into the public network call leg).

2.4.2 Bad protocol detection

The SBC processes all signaling and media that enter or leave the network. It can therefore screen the network from bad protocol within signaling or media packets, discarding or sending negative responses to badly-formed packets. This has two advantages.

- It reduces the load on the VoIP servers within the network, which can be significant if someone is attempting to mount a DoS attack on the network by sending poorly-formed packets.
- It reduces the likelihood of the badly-formed messages causing a crash on a key piece of VoIP infrastructure within the network.

The amount of checking that an SBC does on signaling messages should be configurable. For example, it could be configured to check only those fields that it itself needs to process the message, or it could check all fields in the message, or anywhere in between.

References

Rosenberg, J, Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M. and E. Schooler. (2002). SIP: Session Initiation Protocol. RFC 3261, June, 2002.

Schulzrinne H, Rosenberg J. (2000). The Session Initiation Protocol: Internet-Centric Signaling. IEEE Communications Magazine, 2000, (3): 134-141.

A. Johnston, S. Donovan, R. Sparks, C. Cunningham, K. Summers. (2003). Session Initiation Protocol(SIP) Basic Call Flow Examples. RFC 3665, December, 2003. 82.

M. Handley, V. Jacobson. (1998). SDP: Session Description Protocol. RFC2327, April, 1998.

H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. (2003). RTP: A Transport Protocol for Real-Time Applications. RFC 3550, July, 2003.

P. Srisuresh, M. Holdrege. (1999). IP Network Address Translator(NAT). RFC2663. IETF. 1999.8.

J. Rosenberg, J. Weinberger, C. Huitema. (2003). STUN—Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs). RFC3489. IETF. 2003.4

Alan B. Johnston. SIP—Understanding the Session Initiation Protocol Second Edition Section 3-3, 2004: section 3.3.

J. Rosenberg. (2006). Traversal Using Relay NAT (TURN) draft-rosenberg-midcom-turn-08, May, 2006.

Rosenberg. Interactive connectivity establishment(ICE): A methodology for network address translator (NAT) traversal for the session initiation protocol (SIP) [EB/OL].

<http://www.jdrosen.net/papers/draft-rosenberg-sipping-ice-00.html>.

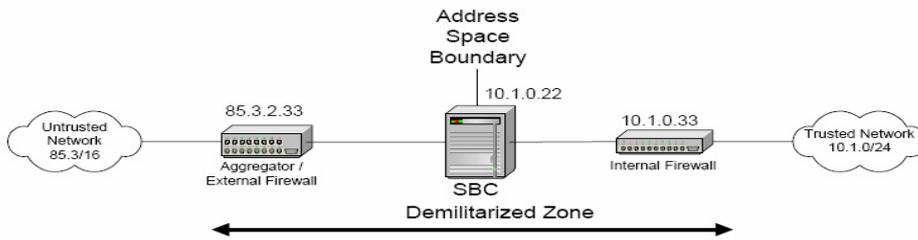


Figure 1. SBC does firewall processing

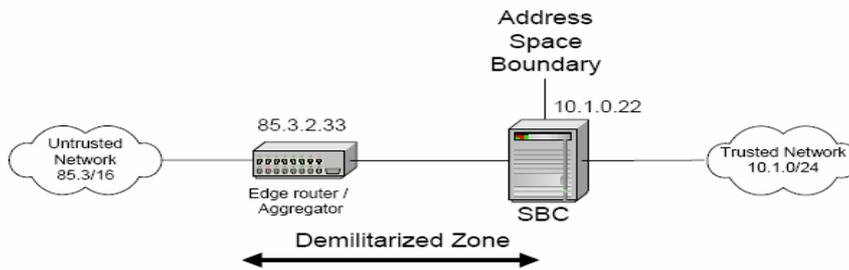


Figure 2. SBC does firewall processing

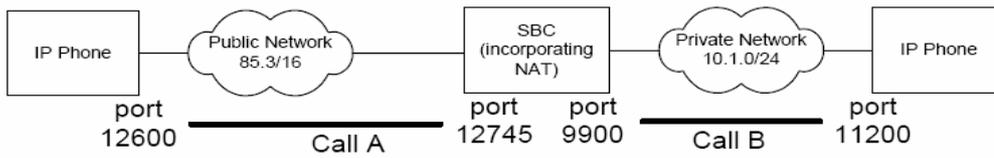


Figure 3. VoIP media packets traversing the DMZ



Image Analysis for Online Dynamic Steganography Detection

Kevin Curran & Joanne Mc Devitt
Intelligent Systems Research Centre
Faculty of Engineering
University of Ulster, Northern Ireland, UK
E-mail: kj.curran@ulster.ac.uk

Abstract

In recent years there has been a noticeable growth in the quantity of available Steganography tools on the World Wide Web. Steganography may be used to hide messages within images and it is widely believed that terrorist organizations may be communicating through the use of steganography. With this in mind there is a need to detect hidden data using Steganalysis – the art of detecting messages using Steganography. This paper presents a Steganalysis tool which scans images on the Web to test if they have been affected by Steganography.

Keywords: Steganography, Steganalysis, Watermarking, Security, Image processing

1. Introduction

Steganography derives from the Greek word Steganos meaning covered and Graphos meaning writing or drawing (Cole, 2005). Steganography is the art and science of writing hidden messages in such a way that no one apart from the intended recipient knows of the existence of the message. Steganography involves selecting an appropriate carrier file such as an image or audio file, removing the less important information from that file and injecting the hidden message in its place. When the cover message and the secret message are combined a stego image is created. Figure 1 illustrates a typical steganography system.

Steganographic messages may be encrypted before they are inserted into the cover image for increased security. Only those who know the technique used to encrypt the message can recover the message. The purpose of steganography is not to keep others from knowing the hidden information but to keep them from knowing the very existence of the information. If a steganographic method causes someone to suspect the carrier medium, the method has failed (Artz, 2001). Although messages embedded into an image are often imperceptible to the human eye, they often disturb the statistical nature of the image (Farid, 2006). The choice of the cover image in steganography is very important as it influences the security of the technique in a huge way. Some steganographic experts recommend greyscale images are the best to use as cover images (Fridrich & Goljan, 2006.). Images are the most popular medium for hiding data. The more detailed an image is, the fewer constraints there are on how much data it can hide before it becomes suspect (Artz, 2001). The internet is said to be a storehouse of steganographic material (Callinan & Kemick, 2006). There have been numerous rumours that terrorists are using steganography to exchange images over the internet. USA Today in 2001 stated that Al-Qaeda as well as other terrorist groups were using steganography to plan and implement terrorist attacks (Callinan & Kemick, 2006). Many people think steganography is a relatively new area in research but some examples of steganography trace far back to 440BC. Demeratus, a Greek at the Persian Court, warned Sparta of an invasion by Xerxes, King of Persia. Demeratus removed the wax from a writing tablet, wrote his message on the wood underneath and then covered the message with wax (Katenzenbeisser & Petitcolas, 2000). Another example of early steganography involved Histiaeus, who shaved the head of his most trusted slave and tattooed a message. Once the slave's hair had regrown, he was sent to transport the message (Katenzenbeisser & Petitcolas, 2000). This method was still used by German spies at the beginning of the 20th century. In 1499, Johannes Trithemius published *Steganographia*, one of the first books on steganography.

As technology evolved throughout the years so did steganography techniques. Invisible ink, created by juice or milk was discovered. When heat was applied to the document, the hidden writing became visible (Silman, 2001). In World War II, Germans used microdots to hide large amounts of data on printed documents, masquerading as dots of punctuation (Cole, 2003). Steganography has taken a giant leap forward which started in the 1990s when governments, industries and terrorist organizations began using software applications to embed messages and photos into various

types of media. Steganography has become a hot topic on the internet in terms of electronic privacy and copyright protection. In 1995, a search using steganography as a keyword on the World Wide Web produced fewer than a dozen responses. In 1996, the same search produced 500 hits. In 1998, such a search produced over a thousand hits (Cole, 2003). Today, a search in Google produces 2,200,000 hits. Computer technology has made it so much easier to hide messages and more difficult to discover that message.

There are important connections between steganography and cryptography but both have very different goals. The goal of steganography is to keep the existence of a message hidden or to hide the fact that communication is taking place. In contrast, the goal of cryptography is to obscure a message or communication so that it cannot be understood (Druid, 2006). With this in mind, the security ability of both techniques is also different. In cryptography, the message is not hidden therefore a hacker can try to intercept and decrypt the message but with steganography, the hacker must discover the medium before he/she can try to intercept it. Putting their differences aside, steganography and cryptography make great partners. It is very common to use these techniques together. As an additional security measure, a steganographer can encrypt the message before it is hidden using steganography (Grodzinsky et al, 2005). Steganography involves hiding data in such a way that it is difficult for an attacker to detect the existence of a secret message in the carrier file. Based on this, three principles can be used to measure the effectiveness of a steganography technique. The principles are amount of data, difficulty of detection and difficulty of removal (Cole, 2003).

- Amount of data suggests that the more data you can hide the better the technique.
- Difficulty of detection relates to how easy it is for someone to detect that a message has been hidden. Once you increase the amount of data hidden in a file, the risk that someone will be able to detect the message also becomes higher.
- Difficulty of removal suggests that someone intercepting your file should not be able to remove the data easily

Three steganography protocols exist to allow users more flexibility in their steganography systems and to prevent attacks on their systems. These protocols are Pure Steganography, Secret Key Steganography and Public Key Steganography. In Pure Steganography, the embedding and extraction algorithms should only be known by the message sender and the intended receiver (Anon, 2006). With Secret Key Steganography, it is assumed that a party other than the sender and intended receiver knows the embedding and extraction algorithms. The sender embeds a message in a cover object using a secret key known as a stego key. If a third party intercepts the stego object and extracts the information, the result will be scrambled. Only the intended receiver who possesses the same stego key can extract the original message. Public Key Steganography is based on the principles of Public Key Cryptography. In Public Key Steganography both a public key and a private key are used. The public key is used in the embedding process and the private key is used in the extraction process. This allows the sender and the receiver to avoid exchanging a secret message which might be compromised. However, this method is susceptible to a man-in-the-middle attack. There are two ways by which steganography techniques can be categorised; the file type and the method of hiding. File Type categorisation breaks down steganography based on the type of carrier file. Each file format has specific properties which control how data is hidden in that type of file. Based on this, knowing the host file type can give you an idea of where the data might be hidden (Cole, 2003). There are three general ways to hide data. They are injection, substitution and generation.

- Injection finds the areas of the file which will be ignored and injects the hidden message in that part of the file.
- Substitution finds the insignificant information in the file and replaces it with the hidden message
- Generation creates a new overt file based on the information contained in the covert message.

Today there are many uses of steganography on the World Wide Web. Most modern uses are based in the area of security. Governments use steganography for secure communication and to hide information from other governments. Steganography is also used by many businesses to maintain privacy and protect trade secrets. Technophiles use steganography to send secret messages just for fun. Another popular use for steganography today is watermarking. Watermarking involves injecting copyright marks and serial numbers into electronic mediums such as books, audio and video so that its source can be tracked or verified. There are many negative uses of steganography today. Many people believe that terrorists are using steganography to communicate by posting their secret messages in pictures of pornographic websites and sports chatrooms. Criminals use steganography to transmit child pornography. Steganography can also be used for corporate espionage where a person can get a job within a specific company with the intent of stealing valuable information. The stolen information can be transported to the recipient using steganography (Kn1ght10rd, 2005).

2. Steganography in Images

Digital images are the most widely used medium for steganography today. Digital images take advantage of our limited visual perception of colour. This field is expected to continually grow as computer graphics power also grows (Calpe,

2006). This report will focus on image steganography. In a computer, images are represented as arrays of values. These values represent the intensities of the three colours red, green and blue where the value of each of the three colours describes a pixel (Queirolo, 2006). These pixels are represented row by row. The number of bits in a colour scheme, called the bit depth, refers to the number of bits used for each pixel. The least number of bits in a current colour scheme is eight which means 8-bits are used to describe the colour of each pixel. Monochrome and greyscale images use 8-bits for each pixel and are able to display 256 different colours or shades of grey. Digital colour images are typically stored in 24-bit files and use the RGB colour cube (Morkel et al, 2005). The RGB Colour cube is illustrated in Figure 3.

Large images are the most desirable for steganography because they have more space to hide the data (Queirolo, 2006). The palette and composition of the image also contribute to how well the steganography tool does its job. Any images with gradual colour gradients or greyscale images are best for steganography as it is easier to insert "small errors" into them. The changes also appear more gradually and are less likely to be detected (Queirolo, 2006). Image compression is the application of data compression on digital images. The objective is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form. There are two types of image compression: lossy and lossless. A lossy data compression method is one whereby compressing data and then decompressing it retrieves data that may well be different from the original but is "close enough" to be useful in some way. Lossless data compression is a class of data compression algorithms that allows the exact original data to be reconstructed from the compressed data. Compression plays a key factor in choosing which steganography algorithm to use. Lossy compression techniques result in smaller file sizes but they increase the possibility that some of the hidden message may be lost due to the removal of excess image data. With lossless compression, the original digital image will remain the same without the chance of loss but the image will not be compressed to the extent it would be with lossy compression (Morkel et al, 2005). Different steganography techniques have been developed to handle both of these compression types.

Common methods for hiding information within images include least significant bit insertion, masking and filtering, and transformation. Least significant bit insertion is a very popular steganographical technique because it is the most basic. The least significant bit is the lowest bit in a series of binary numbers. The LSB is located at the far right of a string. E.g., in the binary number 1011101, the least significant bit is the far right '1'. In this method, the LSB of each byte in an image is used to store the secret data. The resulting changes are too small to be recognised by the human eye. A steganography tool which uses this method uses the RGB colour cube to make a copy of the image palette. The copy is rearranged so that the colours near each other in the RGB cube are near each other in the palette. The LSB of each pixel's binary number is replaced with one bit from the hidden message. A new RGB colour in the copied palette is found and a new binary number of the RGB colour in the original palette is found. The pixel is changed to the binary number of the new RGB colour (Si, 2004).

Masking and filtering techniques are usually restricted to 24-bit images or greyscale images (Krenn, 2006). These methods hide information in a manner similar to watermarks. While masking does not change the visible properties of an image, it can be performed in such a way so that it is not visible to the human eye. This method is much more robust than LSB modification with respect to compression since the information is hidden in the visible parts of the image. Transformation is a more complex way of hiding information. Various algorithms and transformations are applied to the image to hide information in it (Gupta, 2005). Discrete Cosine Transformations (DCT) is one such method. This method is used by the JPEG compression method to transform 8X8 pixel blocks of the image into 64 DCT Co-efficients each (Krenn, 2006). Steganography tools can use the LSB of the DCT Co-efficient to hide information. In addition to DCT, there are two other transformation steganographical techniques; fast fourier transformation and wavelet transformation (Gupta, 2005).

2.1 Current Steganography Software

S-Tools is a Steganographical tool that hides files in BMP, GIF and WAV files. It is a freeware program with a drag and drop interface that runs on most versions of Windows (from Windows 95 onwards). S-Tools offers the ability to hide multiple secret messages in one host file. It hides data in the three least significant bits of each byte of data. Hide and Seek is freeware software. It consists of a series of DOS programs which embed data in GIF files. Hide and Seek conceals data in the GIF files using the least significant bit of each data byte to encode characters. It then uses dispersion to spread the data throughout the GIF file in a pseudo-random fashion (Cole, 2003).

EZ Stego is a Java program which injects data in the least significant bit of GIF images. It sorts the colour palette so that closest colours appear next to one another. The message is then inserted into the least significant bits. The palette is then unsorted by renumbering all of the colours with their original values. The recipient can find the message by using the same algorithm to extract the message from the sorted palette (Si, 2004). Image Hide is a Steganography tools which can conceal data in a variety of formats. It alters the colour table of the pixels by replacing the least significant bits with bits of the hidden message, without increasing the size of the image. Digital Picture Envelope is a Windows 95/98/NT program which hides information in a bitmap file format. It can be used to conceal a large amount of data in an image without changing the size of the image file itself (Cole, 2003).

Outguess is a universal steganographic tool that allows the insertion of hidden information into the redundant bits of data sources. The program relies on data specific handlers that will extract redundant bits and write them back after modification. Currently only the PPM, PNM and JPEG image formats are supported for this method. Outguess allows for the hiding of two distinct messages in the data. GifShuffle is used to conceal messages in GIF image files by shuffling the colour map which leaves the image visibly unchanged. GifShuffle works with all GIF images including those with transparency and animation. It also provides for compression and encryption of the concealed message.

3. Steganalysis

Steganalysis is a relatively new research area with few articles appearing before the late 90s. Steganalysis is the process of detecting steganography by looking at variances between bit patterns and usually large file sizes. It is the art of discovering and rendering useless covert messages (Si, 2004). Steganalyst is a relatively new term that was developed to refer to someone who tries to break steganography techniques. Once a steganalyst has determined that a file has a message hidden in it and the message has been extracted, if it is encrypted, it then becomes the job of a cryptanalyst to try and break the cipher text and figure out what the plaintext message is (Cole, 2003). In steganalysis, comparisons are made between the cover-object, the stego-object and the possible portions of the message; the end result is the stego-object. If the message is encrypted, cryptanalysis techniques may be applied to further understand the message (Katzenbeisser & Petitcolas, 2000). The goal of steganalysis is to identify suspected information streams, determine whether or not they contain any hidden message and if possible, recover the hidden message (Si, 2004). The challenge of steganalysis is that:

- The suspect information stream may or may not contain hidden data
- The hidden data may have been encrypted before it was inserted into the carrier file
- Some of the suspect file may have noise or irrelevant data encoded into them which can make analysis very time consuming
- Unless it is possible to fully recover, decrypt and inspect the hidden data, often one has only a suspect information stream and cannot be sure if it is being used to transport secret data

In steganalysis, the steganalyst does not care which bits carry what information, he/she is trying to determine the existence of a hidden message (Zhang & Ping, 2006). If one can show that an image conceals a message, whether it can be read or not, then the stego system has failed (Zhang & Ping, 2006). Attacks and analysis on hidden information may take several forms: detecting, extracting and disabling or destroying hidden information (Chandramouli & Memon, 2006). A steganalysis attack represents the technique with which a steganalyst attempts to recover, modify or remove a stego message. An attack approach is dependent of what information is available to the steganalyst. Six steganalysis attacks exist which are incidentally derived from four cryptanalysis techniques: stego-only, known-cover, known-message, chosen-stego, chosen-message and known-stego. In the “stego-only attack” only the stego-object is available for analysis (Katzenbeisser & Petitcolas, 2000). This is the most difficult attack approach since there is no starting point from which to extract the hidden message. When the original cover data and the stego message are both available, this is known as the “known-cover attack”. The “known-message attack” assumes either a part of the entire hidden message is available to the steganalyst (Brainos II, 2006). Analyzing the stego-object for patterns that correspond to the hidden message may be beneficial for future attacks against that system. Even with the message available, this attack may be considered equivalent to the “stego-only attack” (Cole, 2003).

In the “chosen-stego attack” both the steganography tool (algorithm) and the stego-object are known. In this case the key, if the message is encrypted and the hidden message are unknown (Brainos II, 2006). “Chosen – message attack” occurs when the steganalyst generates a stego-object from some steganography tool or algorithm from a chosen message. The goal in this attack is to determine corresponding patterns in the stego-object that may point to the use of steganography tools and algorithms. In the “known-stego attack” the steganography tool is known and both the original and stego-objects are available (Cole, 2003).

3.1 Types of Steganalysis

Steganalysis can be classified into two categories, Passive Steganalysis and Active Steganalysis. Passive Steganalysis involves detection only. The steganalysis process ends when the following question has been answered, “Does the media contain steganographic data?” In Active Steganalysis, the process is complete only after the hidden data is removed, destroyed or strategically altered to render it useless.

Steganographic systems often leave detectable traces within a medium’s characteristics (Provos & Honeyman, 2006). This allows an eavesdropper to detect modified media, revealing that secret communication is taking place (Provos & Honeyman, 2006). With steganalysis, although the secret content is not exposed, its existence is revealed, which defeats the main purpose of steganography (Provos, Honeyman). Steganographic methods insert information and manipulate the images in ways to remain invisible (Katzenbeisser & Petitcolas, 2000). However, any manipulation to

the image introduces some amount of distortion and degradation of some aspect in the 'original' image's properties (Katzenbeisser & Petitcolas, 2000). In images that have colour palettes or indexes, colours are typically ordered from the most used to least used, to reduce table lookup time. The changes between colour values may change gradually but rarely. Greyscale image colour indexes do shift in 1-bit increments, but all the RGB values are the same. Applying a similar approach to monochromatic images other than greyscale, normally two of the RGB values are the same with the third generally being a much stronger saturation of colour. Having occurrences of single pixels outstanding may point to the existence of hidden information (Katzenbeisser & Petitcolas, 2000). Added content to some images may be recognisable as exaggerated noise. This is a common characteristic for many bit plane tools when applied to 8-bit images. Using the 8-bit images without manipulating the palette will cause colour shifts as the raster pointers are changed from one palette to another (Katzenbeisser & Petitcolas, 2000). If the adjacent palette colours are very similar there may be little change. However, if the adjacent palette entries are dissimilar, then noise due to the manipulation of the LSB is obvious (Katzenbeisser & Petitcolas, 2000). There are two main types of steganalysis techniques; Visual Analysis and Statistical Analysis.

The majority of steganographic algorithms embed messages by replacing carefully selected bits of an image with bits of the secret message. The human ability is used for visual analysis (Westfeld & Pfitzmann, 2006). Analysing repetitive patterns may reveal the identification of a steganography tool or hidden information. An approach used to identify such patterns is to compare the original cover images with the stego images and note visible differences (Katzenbeisser & Petitcolas, 2000). This is called a "known-cover attack". Another visual clue to the presence of hidden information is the padding or cropping of an image. With some stego tools, if an image does not fit into a fixed size, it is cropped or padded with black spaces. There may also be a difference in the file size between the stego image and the cover image which indicates steganography. Another indicator is a large increase or decrease in the number of unique colours in a palette which increase incrementally rather than randomly (except greyscale images) (Si, 2004). The idea of visual attacks is to remove all parts of the image covering the message. The human eye can then distinguish whether there is a potential message or still image content (Westfeld & Pfitzmann, 2006). The filtering process depends on the presumed steganographic utility. The process is shown in Figure 4.

Statistical analysis is frequently used in steganography detection. Statistical tests can reveal that an image has been modified by steganography by determining that an image's statistical properties deviate from the norm (Provos & Honeyman, 2006). This is done by measuring the colour pairs in a image. These two colours, which do not differ greatly, are known as a close colour pair. In a steganographic image, the individual colours become "less different" when a file is embedded into an image. The occurrences of these colours are measured using statistics (Marcus, 2006). One specific technique is the Raw Quick Pairs (RQP) method. In this method, the ratios between the close pairs and the colour pairs are measured in an image. Then, a message is deliberately embedded into the image and the measurements are taken again. If there is a large difference between the two measurements, then it is unlikely that a message has been embedded into the original image (Marcus, 2006). A more sophisticated technique is Ueli Maurer's Universal Statistical Test for Random Bit Generators (Provos, 2001). If we use a block size of 8 bits, the expected result from the maurer test for a truly random source is 7:184. We expect images with hidden data to have higher entropy than those without. Westfeld & Pfitzmann outline an interesting statistical attack in their paper "Attacks on Steganographic System" (Westfeld & Pfitzmann, 2006). They observe that for a given image, the embedding of encrypted data changes the histogram of colour frequencies in a particular way (Provos, 2001). When using the least significant bit insertion method to embed encrypted data into an image that contains colour two more often than colour three, colour two is changed more often to colour three than the other way around. As a result, the differences in colour frequency between two and three are reduced in the embedding (Provos & Honeyman, 2006).

The most popular steganalysis tool available is Stegdetect. Stegdetect, provided by Neils Provos, is an automated tool for detecting steganographic content in images. It is capable of detecting several different steganographic methods to embed information in JPEG images. Currently, the detectable schemes are Jsteg, Jphide (Unix and Windows), Invisible Secrets, Outguess 01.3b, F5 (header analysis) and Appendix and Camouflage. Stegdetect 0.6 supports linear discriminant analysis. Given a set of normal images and a set of images that contain hidden data by a steganographic application, Stegdetect can automatically determine a linear detection function that can be applied to unclassified images. Linear discriminant analysis computes a dividing hyperplane that separates the non-stego images from the stego images. The hyperplane is characterised as a linear function. The learned function can be saved for later use of new images. Stegdetect includes StegBreak which is a brute force attack tool for determining the pass phrase (if any), which has been assigned to the cover image embedded with a hidden message (Si, 2004). Stego Suite is another popular steganalysis tool. It is a software bundle available for the investigation, detection, analysis and recovery of digital steganography. This product is comprised of three products; StegWatch is WetStone's flagship steganography tool, Stego Analyst is a state of the art image and audio file analyzer and Stego Break is a password cracker.

It is important to detect hidden messages within images. On the steganography side, this is important in order to find methods to improve the algorithm implementing steganography. By exposing the flaws of the algorithm, the user can

further improve the algorithm in order to make it more difficult to detect whether or not data is hidden in the images (Blanco et al, 2005). Steganalysis is also extremely important with regards to security. In the age of Internet, images are sent via email or by posting them on websites. Detecting whether or not data is hidden in the images will allow us to monitor and further analyse the suspicious images in order to find what the hidden message is (Blanco et al, 2005).

4. Steganalysis Prototype

The Steganalysis system consists of two main components; the Steganalysis application and a web spider. Each component has an individual role within the system. The Steganalysis application and the web spider is installed on the user's computer and with the use of an internet connection, the system is connected to the World Wide Web. Figure 5 illustrates the Steganalysis system.

The Steganalysis application allows the user to load images into the system for steganography detection. The application scans the image based on the least significant bit scan method. Once the scan has completed the application will inform the user of the result of the scan. Web Spiders are specialist types of robots or automated programs that search the web. The web spider used within the Steganalysis system downloads images from specific website input by the user. The spider allows the user to download all images on the World Wide Web based on the image file type choice which is also chosen by the user. The user will setup a location where the downloaded images will be saved.

The user must enter a website URL which the spider can connect to. The user must also specify the location on the computer to which the downloaded images can be saved. The user can select the type of images which he/she wants to download. The downloaded images can be viewed within the web spider interface. Once the user clicks on the scan button, the image will be scanned using the least significant bit detection method. The system will then run the appropriate scan on the image.

The result of the scan will be displayed to the user. WebRipper is used to quickly download all the images contained in a specific website and store them to a specific folder on the computer. These will be the images used within the steganalysis system to detect steganography. The steganalysis system is based on the least significant bit steganography technique. The system will detect steganography in images by checking if the least significant bits of the image have been modified. To detect lsb steganography using the steganalysis system, the user must first select an image to scan. This is illustrated in Figure 6. When the user selects an image and clicks on the open button, the image will be displayed as shown in Figure 7.

The user then clicks on the detect button which will run the detection algorithm by calling the extract class. If the system detects the lsb of the image has been modified and the message contains hidden data, a message will be displayed to the user. If no hidden message has been detected, the system will inform the user of the result.

The steganography demonstration part of the system is an extra feature to allow the user to see steganography in action. This feature of the system demonstrates steganography to the user. To do this, the user must first select an image using the JFileChooser as discussed previously. Next, the user must click on the button named Steganography Demonstration which will call the steganographyImageDeconstructor class to convert the image pixel data into an array. This class then calls the steganography demonstration class which will hide a predefined message in a copy of the original image so that no permanent changes are made to the original image. The steganographyImageDeconstructor class will then build a JFrame which will display the original image and the new steganographic image which contains the hidden message.

The result of the scan will be displayed to the user. WebRipper is used to quickly download all the images contained in a specific website and store them to a specific folder on the computer. These will be the images used within the steganalysis system to detect steganography. The steganalysis system is based on the least significant bit steganography technique. The system will detect steganography in images by checking if the least significant bits of the image have been modified. To detect lsb steganography using the steganalysis system, the user must first select an image to scan. This is illustrated in Figure 6.

A number of images were included which we know to be steganographic images. As it is known that this image has been affected by steganography, the system responds with a message to say it has detected steganography. The result is shown in Figure 8. If the result is negative then the popup in Figure 9 is shown. Figure 10 illustrates the steganography demonstration.

5. Conclusion

At present, the system only reveals the existence of a message; it is not concerned with what the message is. This feature could be included within the system. Decryption could also be used within this feature as most steganography programs use encryption to make it more difficult to reveal the message in its original form. In conclusion, the project has been successful in designing and implementing a steganalysis program to detect lsb steganography. Although the

steganalysis system could be improved it effectively serves as a basic detection tool which can identify steganographic images.

References

- Artz, D. (2001). Digital Steganography: Hiding Data within Data.
http://www.cc.gatech.edu/classes/AY2003/cs6262_fall/digital_steganography.pdf
- Blanco, D. et al. (2005). Importance of Steganalysis.
<http://cnx.org/content/m13179/latest>
- Brainos II, A.C. (2006). A Study of Steganography and the Art of Hiding Information.
http://www.infosecwriters.com/text_resources/pdf/steganographyDTEC6823.pdf
- Callinan, J., Kemick, D. (2006). Detecting Steganographic Content in Images Found on the Internet.
<http://www.chromesplash.com/jcallinan.com/publications/steg.pdf>
- Calpe, A. (2006). Steganography in Images.
http://www.cs.ucf.edu/courses/cot4810/fall04/presentations/Steganography_in_Images.ppt
- Cole, E. (2005). Hiding in Plain Sight. Wiley, John & Sons, Incorporated.
- Chandramouli, R., Memon, N.D. (2006). Steganography Capacity: A Steganalysis Perspective.
<http://citeseer.ist.psu.edu/cache/papers/cs/27278/http:zSzzSzwww.ece.stevens-tech.eduzSz~moulizSzstegcap03.pdf/chandramouli03steganography.pdf>
- Esoteric Archives. (2006). Steganographia (Secret Writing), by Johannes Trithemius.
<http://www.esotericarchives.com/tritheim/stegano.htm>
- Druid. (2006). Steganography Primer: Introduction to Steganography.
<http://druid.caughq.org/presentations/Steganography-Primer/>
- Farid, H. (2006). Detecting Steganographic Messages in Digital Images.
<http://www.cs.dartmouth.edu/~farid/publications/tr01.pdf>
- Fridrich, J., Goljan, M. (2006). Practical Steganalysis of Digital Images – State of the Art.
<http://www.ws.binghamton.edu/fridrich/Research/Steganalysis01.pdf>
- Gupta, S. (2005). All about Steganography.
<http://palisade.plynt.com/issues/2005Apr/steganography>
- Katzenbeisser, S., Petitcolas, F.A.P. (2000). Information Hiding Techniques for Steganography and Digital Watermarking. Artech House Publishers.
- Knlgthl0rd. (2005). Steganography: The World of Data Hiding.
<http://www.infonomicon.org/text/Research.rtf>
- Krenn, R. (2006). Steganography and Steganalysis.
<http://www.krenn.nl/univ/cry/steg/article.pdf>
- Marcus, I. (2006). Steganography Detection.
<http://www.uri.edu/personal2/imarcus/stegdetect.htm>
- Morkel, T. et al. (2005). An overview of Image Steganography.
<http://mo.co.za/open/stegoverview.pdf>
- Provos, N., Honeyman, P. (2006). Detecting Steganographic Content on the Internet.
<http://niels.xtdnet.nl/papers/detecting.pdf>
- Roberts, M. (2006). The RGB Colour Model.
<http://www.cs.bham.ac.uk/~mer/colour/rgb.html>
- Si, B. (2004). Introduction to Steganography.
<http://www.infosyssec.com/infosyssec/Steganography/menu.htm>
- Silman, J. (2001). Steganography & Steganalysis: An Overview.
http://www.sans.org/reading_room/whitepapers/steganography/553.php

Queirolo, F. (2006). Steganography in Images.

<http://www.cse.buffalo.edu/~peter/cse741/Presentations/Refs/Queirolo.pdf>

Westfeld, A., Pfitzmann, A. (2006). Attacks on Steganographic Systems – Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools and Some Lessons Learned.

<http://www.ece.cmu.edu/~adrian/487-s06/westfeld-pfitzmann-ihw99.pdf>

Zhang, T., Ping, X. (2006). A Fast and Effective Steganalytic Technique Against JSteg Like Algorithms. http://citeseer.ist.psu.edu/cache/papers/cs/26891/http://zSzzSzwww.rbfm.comzSzTaoZhangzSzPaperzSzacm_sac_2003_zhang.pdf/zhang03fast.pdf

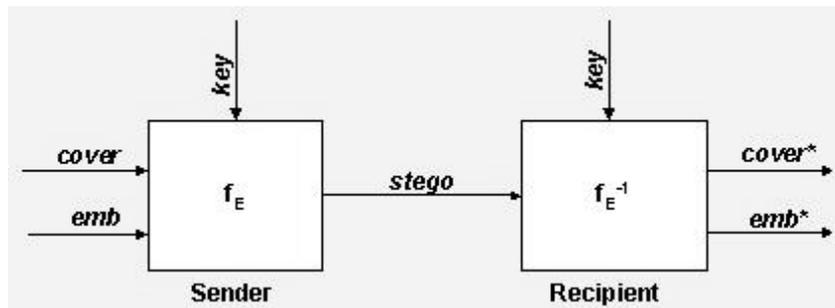


Figure 1. Steganography System (Si, 2004)

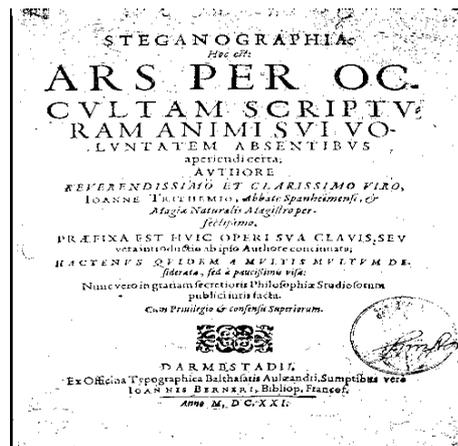


Figure 2. Steganographia by Trithemius (Esoteric Archives, 2006)

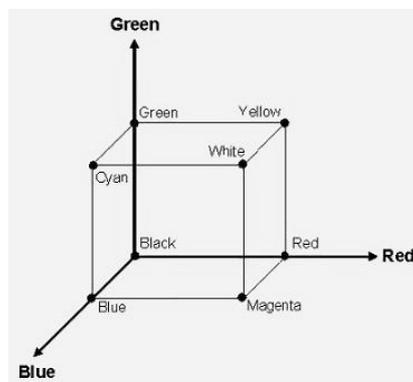


Figure 3. RGB Colour Cube (Roberts, 2006)

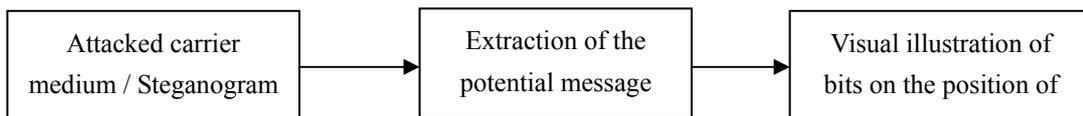


Figure 4. Filtering Process (Westfeld & Pfitzmann, 2006).

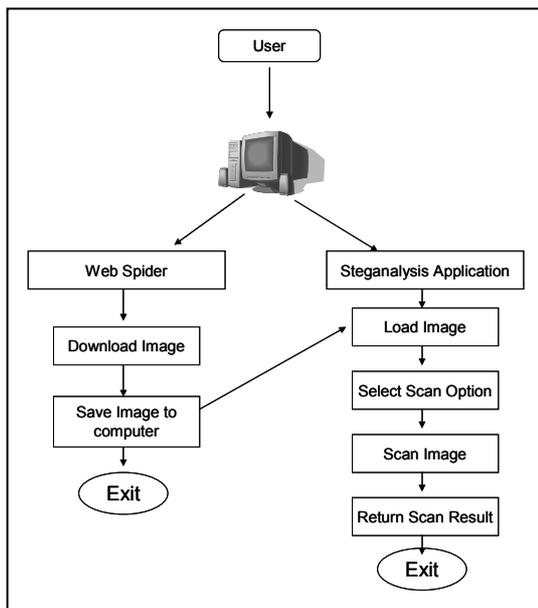


Figure 5. Architecture of Steganalysis System



Figure 6. JFileChooser with image preview.



Figure 7. Main GUI displaying selected image.



Figure 8. Detected Steganography Message

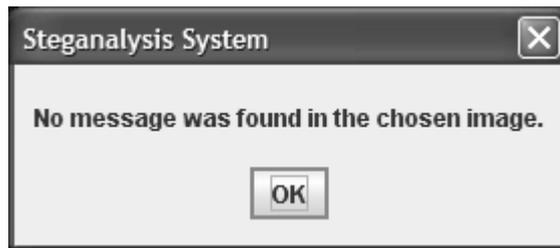


Figure 9. No steganography message detected



Figure 10. Steganography Demonstration.



Model of On-line Room Reservation System Based on Web Service and XPMS

Jiali Wang, Jun Zheng & Wenxin Hu

Computer Center

East China Normal University

Shanghai 200062, China

Tel: 86-21-6223-3797 E-mail:51061211008@ecnu.cn

Abstract

This article analyzed the on-line hotel room reservation systems which had being widely used. Aiming at the advantages and disadvantages of the systems, this article introduced the XPMS theory and put forward a sort of optimized model by using the Web Service technology. The new model can provide sorts of ports aiming at different commercial demands and exchange real-time data under the premise of semantic consistency, which can attract more product providers to join in this network platform and minimize the network cost in the effective range.

Keywords: Heterogeneous data integration, Semantic consistency, Web Service, XPMS, XML

1. Introduction

With the flourish development of network technology, E-business has been very important management mode and concept with the advantages of high efficiency and globalization. To establish the quick response between customers and enterprises, it requires that various systems with same functions and heterogeneous frames must be reintegrated to a new trans-regional network platform with loose coupling.

The on-line room reservation system is a typical representation of E-business information integration, which is different to general search engine. For example, Google computes the relativity among the contents through web link, and grasps data factors in them in time and returns the integrated and standardized result to customers. Because the hotel industry has its own characters such as numerous quantity and regionality, various hotels have different orientations according to actual situations, and many hotels may only construct their interior database, but not the exterior web sites, which can not depend on general search engine to implement.

About the on-line room reservation system, at present, there are some works and products such as www.elong.com and www.ctrip.com, which have been perfected in the aspect of customer service. However, because the network environments in various enterprises are different, so much data are under the distributed environment and have the diversity and complexity of semantic mode, and it is very difficult to realize high efficiently real-time inquiry to different database only through one on-line inquiry platform. Therefore, most present systems adopt the mode signing agreement with hotels. Even overseas hotels also adopt this mode. Customers put in order form first. Then the platform confirm and alternate information through telephone or fax. As viewed from users, it likes that the network platform and hotel databases are alternated real timely, but in actual operation, there are many limits in the on-line room reservation system, for example, the quantity of room reservation can not exceed the promissory quantity in the agreement.

Aiming at above situation, this article puts forward the on-line room reservation system model based on Web Service and XPMS, and except for dynamically implementing real-time inquiry, it also can solve the problem of semantic consistency, and it can classify hotels on different business needs to design and develop different ports and minimize the network spending in order to attract more possible hotels to participate in this platform.

2. Key technologies

2.1 Web Service

As a sort of new distributed computation system structure, Web Service provides a convenient, safe and high efficient platform for the commercial application. At present, there is not a strict definition for Web Service, and people generally think that Web Service is a sort of new Web application program, which has characters such as self-inclusion,

self-description and modularization, can implement issuance, search and binding through Web, can create the abstract definitions of the service and offer concrete implementation of the service, and realize the loose integration from simple asking to complex business logic processing. Web service can be a simple asking to respond the customer, and also can be to complete a complex flow. Once one Web Service is configured successfully, other application programs and other Web Services can directly find and transfer this service. It changes the commercial operation among enterprises and the applied design and development between business and business, and makes the “dynamic integration” among product provider, cooperators and users possible.

2.2 XPMS

XML allows the mutual communication of various programs compiled by different languages on different platforms with the standard form, so it has been the technical base of the trans-enterprise application based on Web Service. XPMS (Extensible Markup Language Product Map Schema) introduces XML Schema to describe the data, finds the mapped relationship among modes with data source through the definition of complete data mode, and offers theoretical base to realize the partnership of semantic consistency. For the data with different structures in Web, the complete mode is used as the uniform data view of complete user and application, and the developer defines corresponding exchange rules through the XML format, and shields the differences of application environment and data structure in the data source through the encapsulation of data sources such as exchange rules and mapped relationships.

3. System design and implementation

Considering that not all hotels are happy to accept the marketing strategy of on-line reservation, we divide hotels into two sorts including the mature type and the growing type according to the grown cycle. The mature hotels generally have centuries-old history and higher reputation, and they always possess stable customer flow even without some certain special on-line reservation platform, so they are not wild about joining in a certain on-line reservation platform. For growing hotels, they always possess modern management technology, but they have not established stable customer group and good public praise, so they need a network platform to develop the market and enhance their reputations, but this sort of hotels always will worry about higher maintenance cost of the network.

Aiming at different market orientations of hotels, we can divide these hotels into the common type and high-level type. Comparing with those common hotels facing the public, the hotels of high-level generally possess special sale channels, and they pay more attention to the band influence and better industrial state, so they will adopt the wait-and-see attitude to the network reservation platform facing the common public.

In a word, this article puts forward the strategy which designs different ports to different customers and controls the network cost in the reasonable range to attract different types of hotel.

For the growing hotels facing the common public, we should only offer a standard port supporting the port agreement and a set of standard stencil which they need, and the hotels put in data according to the fixed stencil and the service program automatically issues the accepted information to the platform, which can effectively control the network cost. For those hotels facing high-level customers, because they are absolutely necessary part in the network platform, so the network platform can attract their participations through offering more safe ports even assuming a part of network spending.

In the implementation, based on the XML standard frame, we can define the network ports supporting the visiting from different database systems and different programming languages, and describe these ports by WSDL (Web Services Description Language). The hotel server can issue a service offering inquiry by means of the document of WSDL, and the source data of the service is still stored in the local database, so the server can quickly complete the dynamic use of Web Services to the service through analyzing the inquiry demand. The network integration platform offers necessary control function of the service work, and the user can participate in the whole service through the mutual operation with Web Services.

The model offers the integrated market of Web Services, and many hotel servers coexist in it, and there are the relationship of competition and cooperation among them, so the security of the communication process must be ensured. To implement real-time data exchange among hotel servers and network integration platforms and ensure the semantic consistence in the network transfer, we design the system which is seen in Figure 1.

The session level is mainly to offer the operation interface. The presentation level (Servlet or EJB) encapsulates the information of SOAP, offers a logic view, treats the inquiry input from the session port, transfers the customers' demands to the business logic level, and transfers concrete implementation and complete the users' response. SOAP is defined on the base of XML and takes XML as the data transfer format, which completely inherits the opening and expansibility of XML, and can better solve the problem of passing the firewall and eliminate the differences among component platforms through defining basic rules and trans-platform information mechanisms.

The business logic level is responsible for the implementation of platform operation logic, describes and produces

SOAP asking information, binds the supplier of hotel service, and judge whether to use one or several hotel servers' remote services according to customers' asking. If needed, the business logic level will integrate corresponding source data based on XPMS and implements the remote use to the service by Web Services, and return the result to this level, and the business logic level implements logic judgment whether the remote use of the next service is needed, and if it is not needed, this level will put in the result to the presentation level, and the presentation level will finally obtain the SOAP information including the exertion result. Servlet asks EJB to treat SOAP information and use the JSP display result, and send HTML to the customer port.

The right part in Figure 1 is the analysis to the system data flow.

The transport layer on the bottom is to send and transport concrete data documents encapsulated by SOAP.

The structure layer designs and implements the semantic consistency arithmetic of data integration base on XPMS, and it is the key to integrate heterogeneous data.

The concept layer includes the descriptions of relative glossary, and it is the premise to implement the XPMS integration arithmetic. It is similar to a compendium which enumerates keywords in the document.

In the simulation experiment, hotel A and hotel B respectively adopt two different database, ACCESS and Sql server, to implement modeling and use different database fields. The integration process of semantic consistency based on XPMS is seen in Figure 2.

In the implementation, define the XML factor as the unit group <id, value>, where id is the only factor identifier and value is the string constant or the complex value including objectives, which is seen in following example.

< Concept >

```
<xpms : id="hotel_name"  xpms : value="concrete name of the hotel">
<xpms : id="hotel_address"  xpms : value="address of the hotel">
<xpms : id=" hotel_address.1 "  xpms : value="the country that the hotel is in">
<xpms : id=" hotel_address.2 "  xpms : value="the city that the hotel is in">
<xpms : id=" hotel_address.3 "  xpms : value="the region that the hotel is in">
<xpms : id=" hotel_address.4 "  xpms : value="the post number of the hotel">
<xpms : id="hotel_rating"  xpms : value="the class of the hotel">
<xpms : id="hotel_type"  xpms : value="the room type of the hotel">
<xpms : id="hotel_type.1"  xpms : value="the room type 1 of the hotel">  //(such as standard room, seascape room
and flatlet)
<xpms : id=" hotel_type.1.1"  xpms : value="the quantity of the room type 1">
<xpms : id=" hotel_type.1.2"  xpms : value="the price of the room type 1">
<xpms : id=" hotel_type.1.3"  xpms : value="the discount of the room type 1">
.....
<xpms : id="hotel_special_features"  xpms : value="the special establishment of the hotel">
<xpms : id="hotel_special_features.1"  xpms : value="the special establishment 1">
<xpms : id="hotel_special_features.2"  xpms : value="the special establishment 2">
<xpms : id="hotel_introduction"  xpms : value="the introduction of the hotel">
<xpms : id="customer_evaluation"  xpms : value="customers' evaluations">
.....
```

< /Concept >

Because the database system of every hotel server is independently designed, implemented and operated, so the data with same semantic contents may have completely different definitions or names, which will present non-structured character or half-structured character, and the data values changes continually. To keep the semantic consistency of data integration, this article puts forward the arithmetic based on XPMS, which can be described that the same semantic data are transformed into same ID number when they are transported to the network reservation platform, and this ID number is only confirmed by the network system platform. Therefore, even for different languages or overseas hotel servers, the expedite exchange of distributed heterogeneous data can be realized when the heterogeneous data are transported and integrated.

The concrete implementation includes following steps.

The total glossary catalog is uniformly regulated by the network reservation platform, and the same semantic keywords possess same ID. It concludes all key glossaries in the databases of all service suppliers. Therefore, various words with different forms always occur after one ID, and the occurrence frequency of some words is even very low, which may only occur in one supplier's database.

The glossary catalog of various hotels is formed according the definitions in the databases of various hotel servers, and it is a subset of the total glossary catalog, and the keyword in the database of creation hotel only occur after ID. This module can largely enhance the efficiency of the data integration. To those hotel servers facing the common public, they can use the uniform stencil, i.e. putting in the data according to the uniform format, so the network platform can adopt the optimized arithmetic of mode matching and reduce the repeated network consumption.

The collection engine module of various hotel glossaries can implement the function that abstracts the keywords in the inquiry consequence of database (the hotel server uses his own database language to complete the inquiry that the customer required and form the inquiry result document which is stored by the fixed XML form). The local glossary collection engine of the hotel server analyzes the keyword glossary in the result documents and put in these keyword glossaries to the upper glossary catalog module to match. Generally, it should be true inclusive relationship. If some new keywords occur, the sub-glossary catalog will inform the total glossary catalog and join the new keywords into the total glossary catalog and relative sub-glossary catalog. If the database of the hotel server changes, for example, some certain keywords are not be used any longer, so it should inform the network system platform and timely refresh relative data catalogs to prevent producing garbage data.

When all matching results return logic true, the data integration engine of various hotels will alternate various inquiry result documents and sub-glossary catalogs, and make same concept possess uniform ID, and put in it to the network reservation platform through the port, and the network platform integration engine module cleans up various documents and return them to clients.

4. Characters of the system

The on-line room reservation system based on Web Service and XPMS in this article realizes the real-time data exchange of trans-database platform and keeps the principle of semantic consistency when implementing data integration. One important problem faced by the integration is the incompatibility of expression and semantics. The incompatibility of expression rules can easily adopt the software to analyze and arrange, but the semantic incompatibility is a problem with large disputes. This article uses the uniform ID to identify through establishing various conceptive keywords, and enhances the matching efficiency through establish the subsets, and accordingly abstracts the design of commercial concept from the whole system design, which possesses good expansibility and flexibility and makes the semantic consistency more easily completed. Secondly, according to the market investigation, this article puts forward adopting different strategies to develop different ports and minimize the network cost aiming at different commercial demands of service suppliers, in order to attract more hotel suppliers' participations.

To the on-line room reservation system, there are many service platforms at present, which can be mainly divided into two sorts. One sort is the hotel chain such as Motel168 and the Star of Jinjiang, this sort of hotel has uniform management and can adopt uniform database and uniform document format, so it is easily to implement real-time data exchange. But this sort of platform has single information, which can not offer information to overseas consumers and other hotels, and the semantic integration is still in the blank stage. The other sort is the third party server such as Ctrip and Elong. This sort of web site can offer supports of software and hardware, and it only needs supporting corresponding information to join in the hotel. But the exchange maintenance of data largely depends on the manual telephone or fax.

Table 1 lists the comparisons among three systems.

5. Conclusions

This system establishes highly efficient enterprise application model by means of the technologies of XPMS and Web Service. It can better integrate the heterogeneous data of various hotel suppliers to one information insurance and inquiry platform and keep the semantic consistency, which is propitious to enhance the enterprise image, develop international market, bring many new opportunities for the enterprise, enhance the economic benefits of the enterprise and strengthen the competitiveness of the enterprise. With the further development and maturity of the XPMS technology, network security and Web Service, the structure design of this system will be realized and applied better.

References

- Cao, Jingxia. (2006). Research on Semantic Overlay P2P Info-sharing System. *Computer Engineering*. No.12.
- Guo, J. (2006). Inter-Enterprise Business Document Exchange. *Proc. of the 8th Int'l Conf. on Electronic Commerce, ACM International Conference Proceeding Series*.

Shegalow G, Gillmann M. (2001). XML-enabled Workflow Management for E-services across Heterogeneous Platforms. *The International Journal on Very Large Data Bases*.

Songhui & Zhangming. (2004). *A First Step Towards the Semantic Web*. Beijing: China Higher Education Press.

Tang, Zhenyu, Chen, Fengyan & Feng, Yuqiang. (2006). Research of Library Network Information Resources Integration Based on XML. *Journal of Harbin Institute of Technology*. No.7.

Zhang, Xiaodong. (2006). Project and Application Based on Web Service and Information Search Technology for Information Integration. *Application Research of Computers*. No.23(3).

Table 1. Comparisons among three sorts of system

Comparison of capability	First sort	Second sort	Model of this article
Data integration	Middle (can only integrate single brand)	High	High
Real time of data	Middle	Low	High
Semantic consistency	Null	Middle	High
Expansibility	Low	Middle	High
Private capability	Middle	Low	High
Friendly	Middle	Middle	High
Main technology	XML	J2EE	XPMS & Web Service

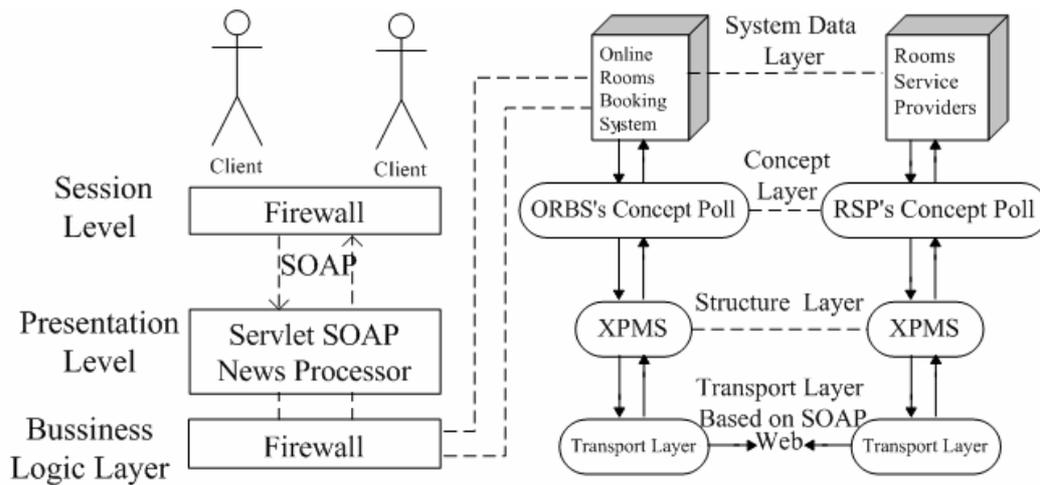


Figure 1. The Frame of On-line Room Reservation System Based on Web Service and XPMS

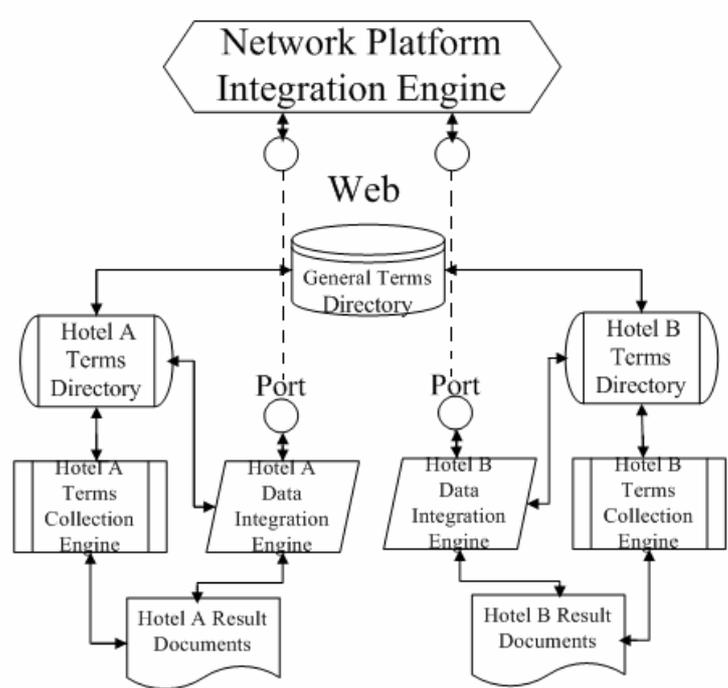


Figure 2. The Semantic Consistency Integration Process Based on XPMS in the Simulation Experiment



Construction of New Forest Roads in Malaysia Using a GIS-Based Decision Support System

Kamaruzaman Jusoff (Corresponding author)

Yale University

Centre for Earth Observation (CEO)

Environmental Science Centre

21 Sachem St, New Haven CT 06511, USA

Tel: 60-3-89467176 E-mail: kamaruz@putra.upm.edu.my

The research is financed by Malaysian MOSTI's Intensive Research Priority Area (Sponsoring information)

Abstract

This paper describes how a Geographical Information System (GIS)-based Decision Support System (DSS) was applied in selecting the most compatible block or compartment to construct a new forest road for a sustainable timber harvesting purposes. The study area was located in Gunong Stong Permanent Forest Reserve, Kelantan, Malaysia between latitude 5°00' N and 5°25' N and longitude 101°20' E and 102°05' E. Four criteria have been considered in this paper which is timber volume (m³), slope (degree), ground condition and distance from the primary and secondary forest road that previously exists (m). ArcView and ArcMap softwares were used to evaluate all the criteria using remote sensing and field data. Results indicated that the best block that fulfill all the criteria chosen for the new forest road construction is Block Nos. 9 and 11.

Keywords: Geogrpahic information system, Decision support system, Forest harvesting, Forest road, Road construction

1. Introduction

Virtually everyone makes hundreds of decisions each day. Good decision making means that the person are informed and have relevant and appropriate information on which to base the choice. Information has been referred as facts, numbers, historical data, graphics, pictures, and sounds (Sauter, 1997). Decision support system (DSS) technology and applications was built to support the decision making. DSS is being used in many fields including forestry. GIS is one of the decision support system tools that help the forest manager or silviculturist to do better decision making. GIS help the forester to do planning and preparation of a block or compartment based on their objectives or targets. GIS has the capability to map, record data, planning and involve in special projects.

Construction of new forest road is one of the main issues that arise in forestry management. Many criteria aspects have to be considered before a new road is constructed. In this study, four criteria have been identified which is timber volume (m³), slope (degree), ground condition and distance from the road that previously exists (m). Scoring model is being used as a weighting for each classes of criteria (Williams, 2000) in this study. ArcMap and ArcCatalog is the GIS software that has been selected to decide the best block that fulfills all criteria.

2. Methodology

2.1 Study area

State of Kelantan, Malaysia at the northern part of Peninsular Malaysia, in a northern-east direction, between latitude 5°00' N and 5°25' N and longitude 101°20' E and 102°05' E had been selected as the study area (Figure 1). Compartment 1 until compartment 16 has been selected as the demonstrative area (Figure 2). The terrain of the study area was classified as very steep slope (greater than 10°). The climate in the study area is a hot humid tropical climate. Temperature and rainfall has been categorized by three distinct features: uniform temperature, high humidity and heavy rainfall. The temperature rarely fell below 20.5°C nor rose beyond 36°C. The mean annual rainfall was seldom less than 2000 mm. The heaviest rainfall occurred during the November, December and January. Dry season was been marked from February until April.

<Figure 1.>

<Figure 2.>

2.2 Identification and evaluation of criteria

The determination of criteria is very important as guidance for collecting data that is needed to help the effectiveness of decision making process. Four criteria have been identified: timber volume (m³), slope (degree), ground condition and distance from the road that had been exist (m). All the selected criteria were classified. Table 1 through Table 4 shows the classification of the criteria.

<Table 1: The volume of the harvest block classes>

<Table 2: Criteria for evaluating slope>

<Table 3: Criteria for evaluating ground conditions>

<Table 4: Criteria for evaluating distance from existing secondary roads>

Next, each classification of criteria was given a score. The maximum score given was one hundred and the minimum was zero (Table 5). This evaluation was very important for the calculation process in ArcMap software to decide the best compartment to construct the new forest road.

<Table 5: Evaluation criteria for the harvest block selection>

2.3 Data collection

Data collection process is very important for data validation and integration. Data that was used in this study are described in Cabral, 2000. Table 6 shows the data for demonstration area (Block No.1-16).

<Table 6: Information of demonstration area>

2.4 System implementation

ArcMap and ArcCatalog software had been chosen as decision support tools to find the best block or compartment to construct new logging roads. Both ArcMap and ArcCatalog software are integrated. In ArcMap, information concerning the demonstration area will be stored in a database (Figures 3 and 4). Next, scores will be given for each data that has been stored (Figure 5).

<Figure 3.>

<Figure 4.>

<Figure 5.>

The calculation of score value was done to get the result of this study using Raster Calculator function that exists in ArcMap. The formula used for calculation is:

$$\text{Calculation} = [\text{distance2} - \text{distance2}] + [\text{ground_con2} - \text{ground_con2}] + [\text{slope5} - \text{slope5}] + [\text{volume2} - \text{volume2}]$$

ArcCatalog has been used to save all the layer files from ArcMap. Combination of all layers will also be saved in ArcCatalog (Figure 5).

3. Results and discussion

Figure 6 shows the classes of volume of harvest for the blocks. The highest score, 60 shows the highest volume that can be harvested in a block. Meanwhile, 40 the lowest score showed the lowest volume that can be harvested. The highest volume that can be harvested in the demonstration area ranges from 40 m³/ha to 50 m³/ha whereas 20 m³/ha to 30 m³/ha is the lowest volume. Table 7 showed results of a selected demonstration area for the estimation of timber volume.

<Figure 6.>

<Table 7: Attributes of volume for harvest block classes>

The second criteria that had been considered in this study are slope. The demonstration area has all classes of slope from 0° until more than 45°. From the illustration in Figure 7, Block Nos. 8, 9 and 10 has the lowest slope. Block Nos. 3, 4 and 16 have the highest slope. If the slope is high, the location is considered not suitable to construct new forest roads. Table 8 shows the attributes of slope for each block.

<Figure 7.>

<Table 8: Attributes of slope>

The next criterion is the ground condition. The best ground condition is in block number 3, 4 and 16 (score = 100). The soil texture of these three blocks is sandy loam and fresh for soil moisture. The worst ground condition is in Block Nos.

8, 9 and 10 (score = 10) in which the soil texture is silty clay loam and the soil moisture is very wet. A good ground condition will make the construction of the road much better. Figure 8 shows the score of ground condition and Table 9 shows the attributes.

<Figure 8>

<Table 9: Attributes of ground condition>

The last criterion taken into consideration is the distance of the block from an existing forest road. Distance is considered because it affects the cost for construction of the road. Block No. 13, 14, 15 and 16 has the highest score because those blocks have existing forest roads. Block No. 7 has the highest potential to construct new forest roads. This is followed by Block No. 8, 10 and 12 (score = 80). The lowest score is in Block Nos. 1, 3, 6 and 9 (score = 20) where the distance is very high from the existing forest road. Scores of distance from existing forest roads are shown in Figure 9 and the attributes shown in Table 10.

<Figure 9>

<Table 10: Attributes of distance from existing forest road>

Figure 10 shows the result of this study which is the combination of all criteria (timber volume, slope, ground condition and distance from the road that previously existed) that had been considered. Block Nos. 13, 14, 15 and 16 have existing forest road. The block that has the highest total score is block number 9 and 11. Block Nos. 7 and 8 have total score equal to 250. This is followed by Block No. 5 (total score = 240), 10 (total score = 230), 6 (total score = 220), 4 (total score = 200), 3 (total score = 180) and 1 (total score = 160). The worst block to construct a new forest road is Block No. 2 (total score = 150).

<Figure 10>

4. Conclusion

GIS has a lot advantages in supporting the decision making process in sustainable forestry management especially for construction of new forest roads in minimal impact timber harvesting. The forester can achieve their objectives by using any combination layers in ArcMap and save the result using ArcCatalog for future reference and implementation.

References

- Cabral, J. C. (2000). *GIS-Based Support System for Tactical Timber Harvest Planning: Design and Development*. Ph.D. Thesis, Universiti Putra Malaysia.
- Sauter, V. (1997). *Decision Support Systems*. John Wiley & Sons, Inc., United States of America.
- William, A. S. (2000). *An Introduction to Management Science, Quantitative Approaches to Decision Making*. South-Western College Publishing / ITP.

Table 1. The volume of the harvest block classes

Class	Criteria
I	>50m ³ /ha
II	40m ³ /ha – 50m ³ /ha
III	30m ³ /ha – 40m ³ /ha
IV	20m ³ /ha -30 m ³ /ha
V	<20m ³ /ha

Table 2. Criteria for evaluating slope

Slope Class	Range (degree)	Description
1	0° – 15°	Level
2	15° – 25°	Gentle
3	25° – 35°	Moderate
4	25° – 45°	Steep
5	>45°	Very Steep

Table 3. Criteria for evaluating ground conditions

	1 Good	2 Moderate	3 Poor	4 Very Poor
Sandy Moisture	Fresh (10-20%)	Moist (20-30%)	Wet (30-40%)	Very Wet (>40%)
Soil Texture	Sandy Loam	Sandy Clay Loam	Clay	Silty Clay Loam

Table 4. Criteria for evaluating distance from existing secondary roads

Class	Distance from secondary roads (m)
1	0 – 750
2	75 – 1500
3	1500 – 2250
4	2250 – 3000
5	>3000

Table 5. Evaluation criteria for the harvest block selection

Criteria	Classification	Score
Timber Volume	I	100
	II	80
	III	60
	IV	40
	V	20
Slope	0° – 15°	100
	15° – 25°	70
	25° – 35°	40
	25° – 45°	10
	>45°	0
Ground Conditions	Good	100
	Moderate	70
	Poor	40
	Very poor	10
Distance from the existing road (m)	0 – 750	100
	75 – 1500	80
	1500 – 2250	60
	2250 – 3000	40
	>3000	20

Table 6. Information of demonstration area

Block	Hectare	Volume (m ³)	Slope (degree)	Ground Condition	Access to Existing Road (m)
1	303	9136	35 – 45	moderate	3160
2	280	10104	35 – 45	poor	2500
3	378	12572	>45	good	3350
4	390	11722	>45	good	2760
5	344	11356	15 – 25	moderate	3090
6	416	15563	15 – 25	moderate	3340
7	374	12692	25 – 35	moderate	2360
8	390	12345	0 – 15	very poor	1000
9	385	13157	0 – 15	very poor	1100
10	426	16057	0 – 15	very poor	2425
11	330	9274	25 – 35	moderate	1160
12	349	9651	15 – 25	moderate	50
13	340	9262	15 – 25	moderate	100
14	365	13265	25 – 35	moderate	180
15	411	15613	35 – 45	poor	190
16	342	14322	>45	good	50

Table 7. Attributes of volume for harvest block classes

Harvest block	Volume/hectare (m ³ /ha)	Class of volume	Score of volume
1	30.15	3	60
2	36.09	3	60
3	33.26	3	60
4	30.06	3	60
5	33.01	3	60
6	33.01	3	60
7	26.54	4	40
8	33.74	3	60
9	30.51	3	60
10	29.45	4	40
11	21.77	4	40
12	41.7	2	80
13	39.01	3	60
14	42.78	2	80
15	37.99	3	60
16	41.88	2	80

Table 8. Attributes of slope

Harvest block	Slope (°)	Class of slope	Description of slope	Score of slope
1	35-45	4	Steep	10
2	35-45	4	Steep	10
3	>45	5	Very steep	0
4	>45	5	Very steep	0
5	15-25	2	Gentle	70
6	15-25	2	Gentle	70
7	25-35	3	Moderate	40
8	0-15	1	Level	100
9	0-15	1	Level	100
10	0-15	1	Level	100
11	25-35	3	Moderate	40
12	15-25	2	Gentle	70
13	15-25	2	Gentle	70
14	25-35	3	Moderate	40
15	35-45	4	Steep	10
16	>45	5	Very steep	0

Table 9. Attributes of ground condition

Harvest block	Soil moisture	Soil texture	Description of ground condition	Class of ground condition	Score of ground condition
1	Moist	Sandy clay loam	Moderate	2	70
2	Wet	Clay	Poor	3	40
3	Fresh	Sandy loam	Good	1	100
4	Fresh	Sandy loam	Good	1	100
5	Moist	Sandy clay loam	Moderate	2	70
6	Moist	Sandy clay loam	Moderate	2	70
7	Moist	Sandy clay loam	Moderate	2	70
8	Very wet	Silty clay loam	Very poor	4	10
9	Very wet	Silty clay loam	Very poor	4	10
10	Very wet	Silty clay loam	Very poor	4	10
11	Moist	Sandy clay loam	Moderate	2	70
12	Moist	Sandy clay loam	Moderate	2	70
13	Moist	Sandy clay loam	Moderate	2	70
14	Moist	Sandy clay loam	Moderate	2	70
15	Wet	Clay	Poor	3	40
16	fresh	Sandy loam	Good	1	100

Table 10. Attributes of distance from existing forest road

Harvest block	Access from existing forest road (m)	Class of access	Score of access
1	3160	5	20
2	2500	4	40
3	3350	5	20
4	2760	4	40
5	2360	4	40
6	3090	5	20
7	50	1	100
8	1000	2	80
9	3340	5	20
10	1160	2	80
11	2425	4	40
12	1100	2	80
13	100	1	100
14	180	1	100
15	190	1	100
16	50	1	100

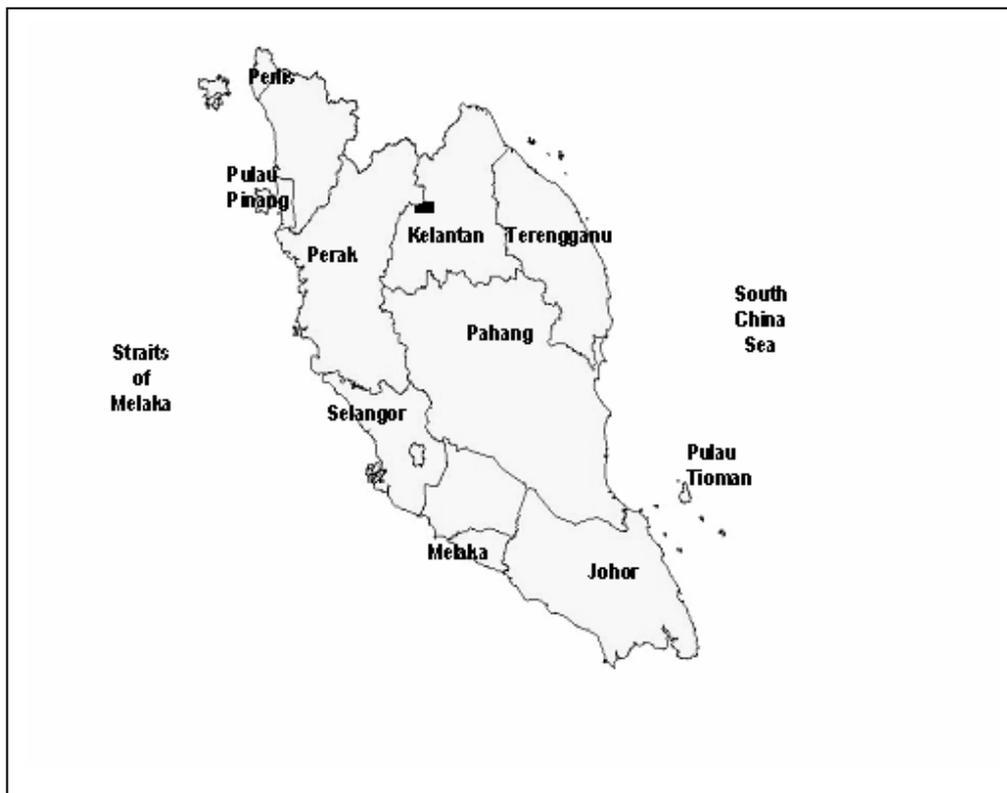


Figure 1. Study area and location

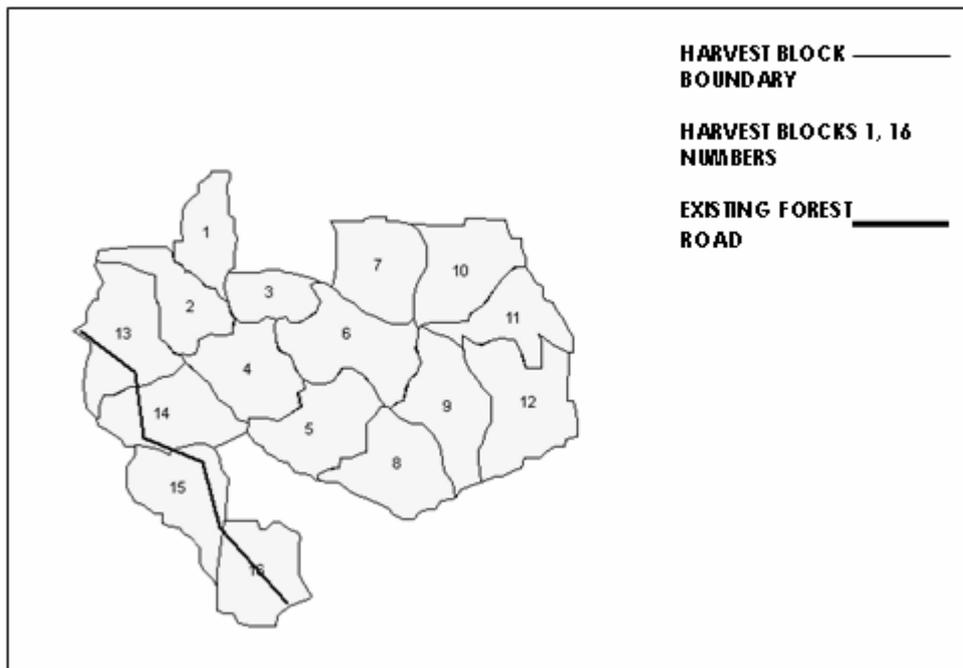


Figure 2. Study and demonstration area

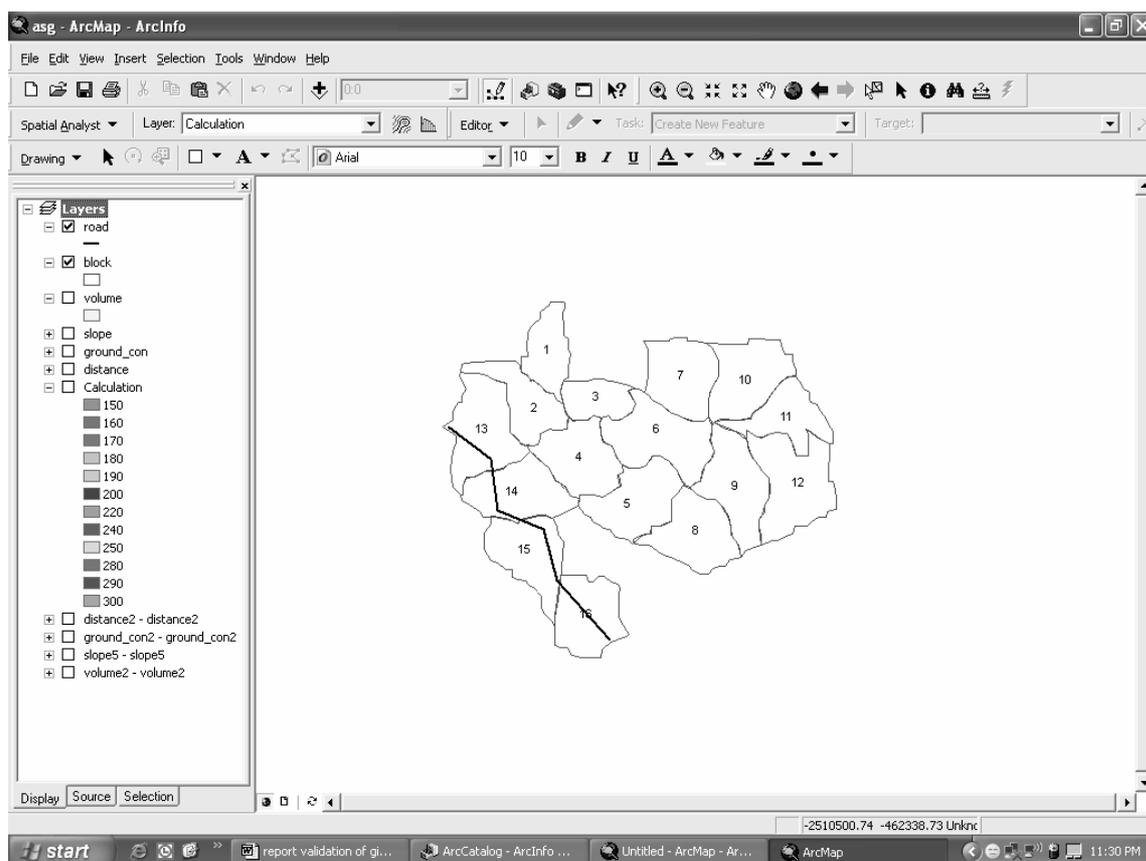


Figure 3. Interface of ArcMap

Id	FID	Shape	slope	Slp_class	Value_slp	Des_slp
1	0	Polygon	35-45	4	10	steep
2	1	Polygon	35-45	4	10	steep
3	2	Polygon	>45	5	0	very steep
4	3	Polygon	>45	5	0	very steep
6	4	Polygon	15-25	2	70	gentle
9	5	Polygon	0-15	1	100	level
5	6	Polygon	15-25	2	70	gentle
8	7	Polygon	0-15	1	100	level
12	8	Polygon	15-25	2	70	gentle
11	9	Polygon	25-35	3	40	moderate
10	10	Polygon	0-15	1	100	level
7	11	Polygon	25-35	3	40	moderate
13	12	Polygon	15-25	2	70	gentle
14	13	Polygon	25-35	3	40	moderate
15	14	Polygon	35-45	4	10	steep
16	15	Polygon	>45	5	0	very steep

Figure 4. Example of database for slope

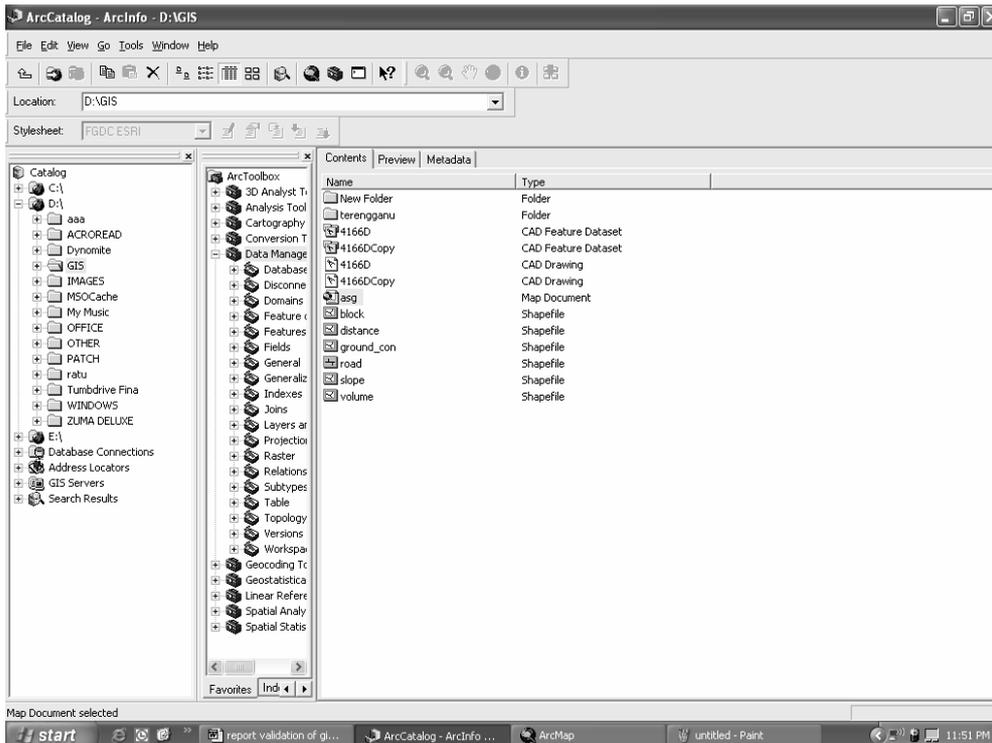


Figure 5. Interface of ArcCatalog

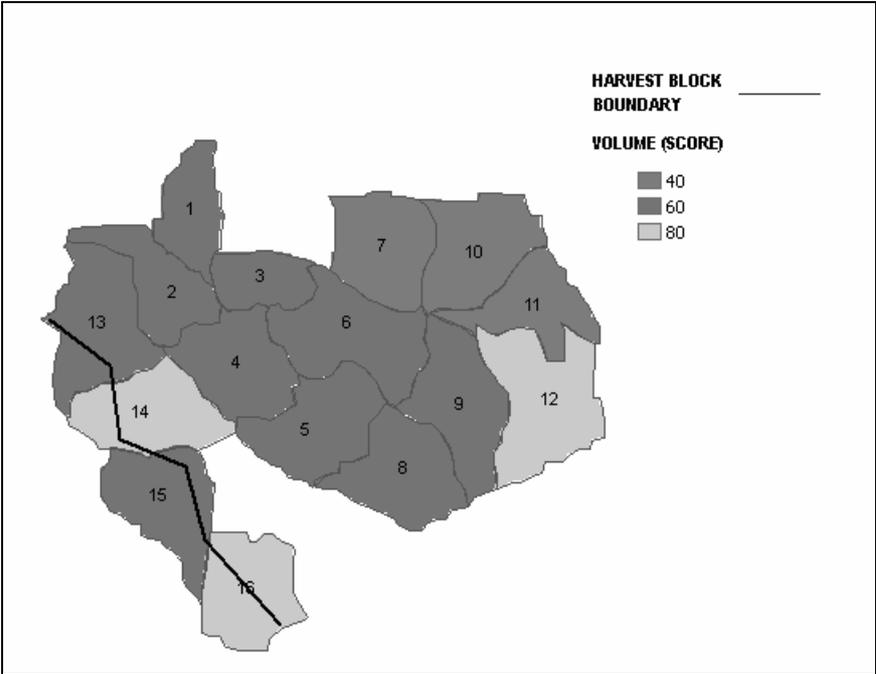


Figure 6. Score of volume for harvest block classes

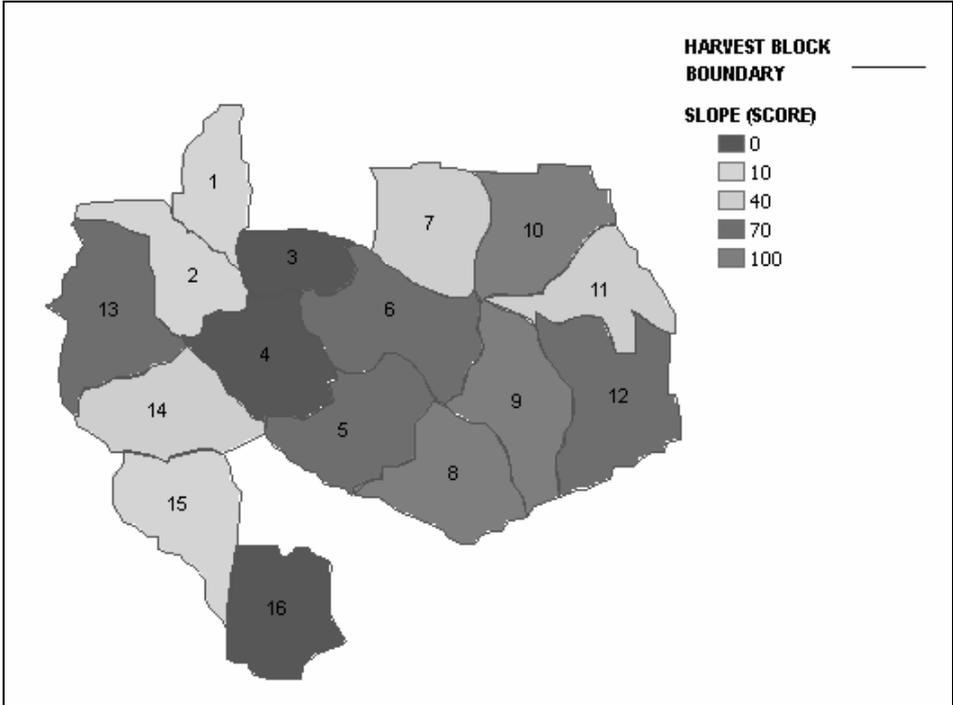


Figure 7. Score of slope

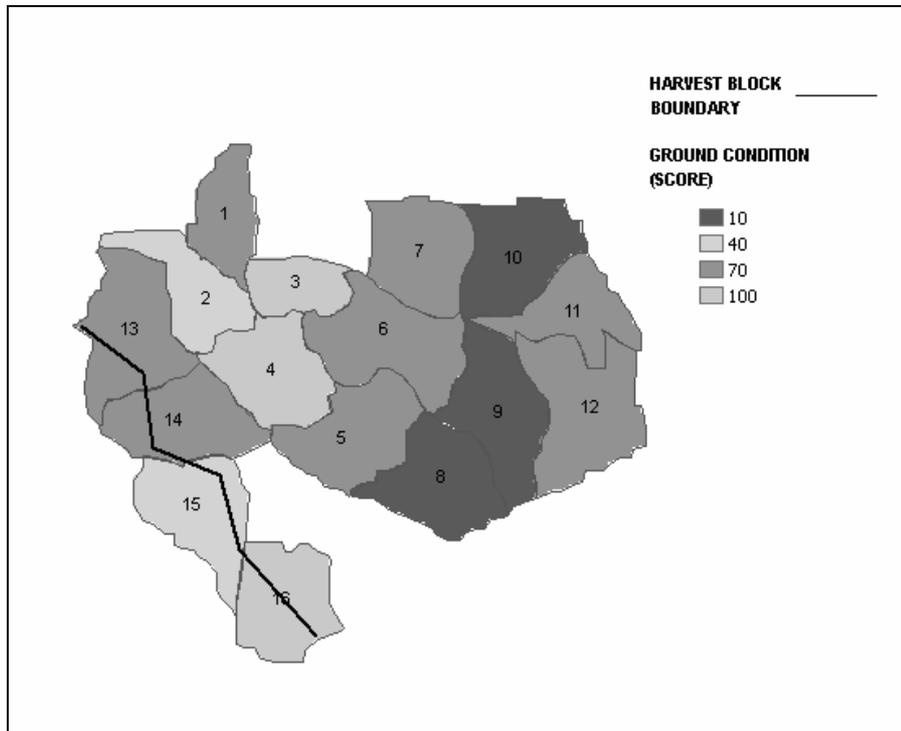


Figure 8. Score of ground condition

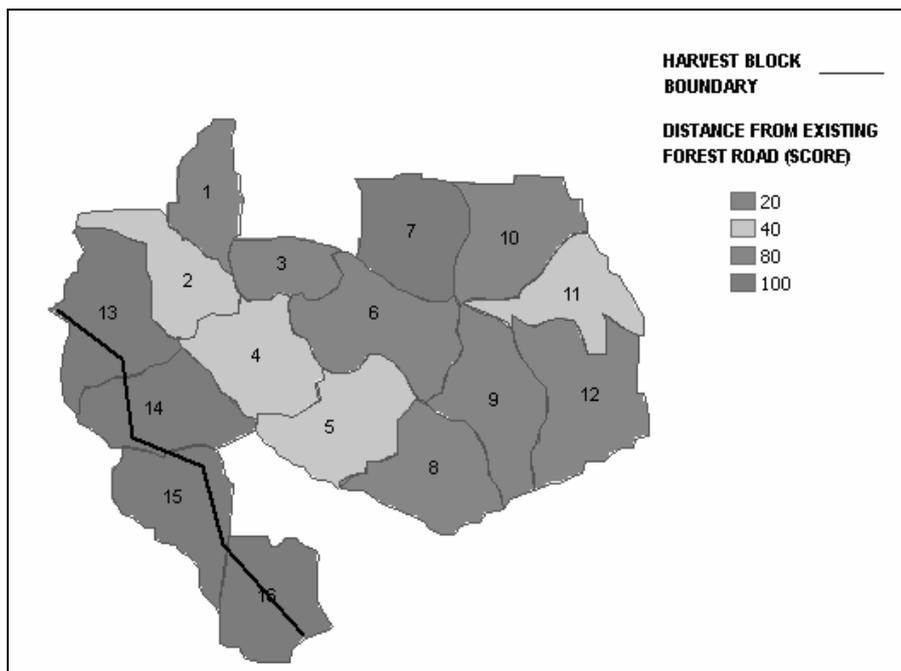


Figure 9. Score of distance from existing forest road

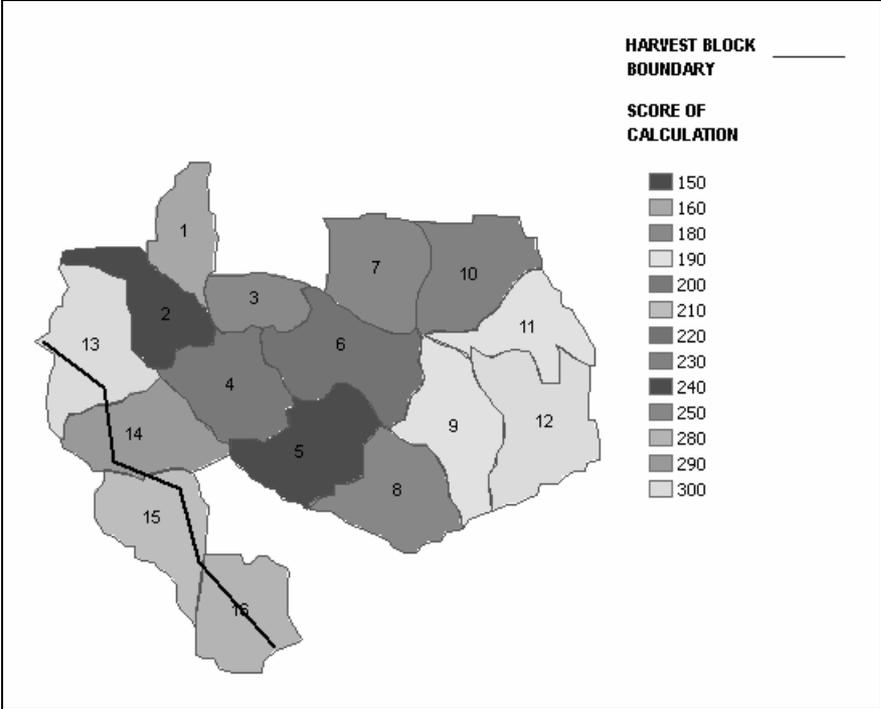


Figure 10. Calculation of results using ArcMap



Revelations of the New Developments of Brain Science on Early Education

Tianbao Zhuang & Hong Zhao

Graduate School of Innovative Life Science

University of Toyama

Toyama, Japan

Tel: 81-090-3886-9168 E-mail: bfztb@sina.com

Zheng Tang

University of Toyama

Toyama, Japan

E-mail: ztang@eng.u-toyama.ac.jp

Abstract

The development of information technology has drove the large revolution of brain science and made it go into the new development stage. The new development of brain science must deeply influence the education, especially for the early education, which makes our attention focus on how to develop children brain again. From the world developments of the research of “brain science and education”, in this article, we will research the revelations of the concepts such as “learning” and “education” in the domain of the cognitive neuroscience and the development of brain science on the early education.

Keywords: Brain science, Information technology, Early education, Neural circuit

1. Introduction

The competition among countries in 21st century is the competition of talent in the final analysis, and the cultivation of talent can not leave the education. Therefore, the good education is extremely important for national security and flourish. The education is also the important topic for discussion in the human society, and to cultivate good moral character, excellent intelligence and healthy physique is the material base to strengthen any one nation.

The continual development of information technology brings revolutions for the development of brain science and fluctuate the traditional concept of brain science and makes it obtain new scientific meanings. Taking the neurobiology as the center, the brain science is the science on the based of animal experiment and brain injury researches. In recent years, with the mature of brain imaging and scanning technology, this science has directly and safely observed super functions in the brain including human mental activities, which is the result that many subjects including molecular biology, medicine, psychology, neurobiology and engineer intercross each other and research together, and it makes the researches of brain science go into the new stage.

The new development of brain science deeply influences the domain of education, especially for the early education. Therefore, it is more and more necessary to research relative problems about early education from the view of brain science. The environment and education in the period of infant is very important to form one’s personality and develop one’s ability, and many research projects about how to develop infants’ brain and how to implement more effective early education to children are more and more emphasized. So it is a very important topic to develop early education and do well infants’ brain developments and even the future of the whole mankind.

In this article, we will discuss many aspects such as the background, feasibility and implementation methods of early education and infant brain development.

2. The world developments of the research of “brain science and education”

2.1 The research developments of OECD

In 1999, the CERI of OECD put forward the first research plan (1999-2001) of “learning science and brain science”. That research topics included “brain structure and infant learning”, “brain structure and youth learning” and “brain structure and aged learning”, which are convened international forum respectively in New York, Granada and Tokyo. The New York forum mainly discussed the brain plasticity and the child special sensitivity period, reported some problems about early education and brain science, such as language study, early cognitive ability, reading structure, mathematic thinking and feeling growth. There were two topics in Granada forum, one was to apply up to date results of cognitive neuroscience to the educational domain by the simple and pellucid form, and another one was to apply brain function imaging technology to find that the brain in the youth period (20-30 years old) was still in the growth. The topic center of Tokyo forum was to open out the brain essential in the aged period and how to enhance the cognitive enginery in the aged period.

In May of 2002, OECD begun the second plan (2002-2005) of “the application research of brain science in the domain of education”, some scientists in US, Europe and Japan had formed three research divisions and cooperative organization. US mainly researched the relationship between language and brain, Europe mainly researched the relationship between arithmetic and math with brain, and Japan mainly assumed the research about the combinations of brain and education such as lifelong learning and human observation ability, memory and “desire”. These researches had produce large influence in many countries and been changing the educational policy and educational practice of many countries.

In July of 2007, one up to date report issued by OECD showed that the brain would never lose its learning ability, not only the learning process could not be stopped, but the brain could be changed physiologically by learning. The emotion could rebuild human nerve organization, and one of the strongest learning motivations came from the understanding to the new concepts. The report also advised that the early education should obtain the experience of “elicitation” for children to ensure they could feel the happiness of learning. The report also called on the education science should cooperate with brain science to establish a new subject, i.e. the learning science, and finally made the education decision and education practice obtain stable scientific proofs.

2.2 The research developments in US

In 1990s, US defined the period from 1990 to the end of the 20th century as “ten years of brain”, and passed the topic to push the topic of brain science research, and the extensive researches about brain science, such as the pathogeny explanation and cure method of spirit disease. In virtue of this chance, various main universities in US established the special research institutions of brain neuroscience, and researches of brain science were continually coming forth, and when they took part in the mutual research of OECD, they developed their own independent researches.

The US National Institutes of Health (NIH) is pushing the research of brain science, and especially the National Institute of Biomedical Imaging and Bioengineering (NIBIB) established by NIH took the development of “Imaging technology” as the intention, and attracted attentions in the measurement technology of brain enginery. Otherwise, US National Science Foundation (NSF) and Finance Ministry also put forward the research project to develop and comprehensively improve human ability and technology in Dec of 2001, and they integrated multiple subject researches to promote the strategic development of brain science and education research project.

2.3 The research developments in Japan

After Japan instituted and implemented the “human proceeding science plan” which put the brain research on the important position in 1986, it also pushed the brain science plan program of “brain science times” with 20 years in 1996. Scientists predict that the brain science would occupy especially important status in 21st century. Based on that, fully considering the actuality and tendency of social development and combining with the cognitions of deficiencies about illegibility and experience existing in the research of education science, scientists should closely combine the brain science and education, and develop application research. In May of 1997, the Brain Science Committee of Life Science Branch in the Scientific Technology Meeting put forward the “long-term plan about the development of brain research”, and established the strategic objective of “understanding brain (the explanation of brain function)”, “protecting brain (the overcoming of brain disease)” and “creating brain (the development of brain computer)”. Japan established the brain science comprehensive research center of physical and chemical graduate school in Oct of the same year, which became the core institution of the research of brain science to implement the explanation of brain function. Above various researches have obtained many developments, and the non-invasive brain scanning and imaging technology has got quick development, which has developed the technology with more exact and safe, and applied the research results into the child education, school education, social life and aged nursing. Based on that, the evaluation branch of scientific technology and learning consideration research plan of Japanese Ministry of Education put forward the new strategic objective of brain science research, i.e. “brain cultivation”, in 2002.

Japanese Ministry of Education convened the meeting of “brain science and education” in March of 2002, and in this meeting, Japanese brain scientists definitely pointed out that the research combining brain science with education is an important tendency in the domain of education research in present Euro and American developed countries. After that, Japanese Ministry of Education convened meeting many times to further definite the research details, and formally started large research project of “brain science and education” on the new year’s day of 2003.

3. To explain the concepts of “learning” and “education” from the view of cognitive neuroscience

3.1 New explanations of learning and education

Owing to the traditional opinions, the “education” is the practical domain on the base of psychology, pedagogy, philosophy, sociology and cognitive science, and the “learning” is the process of repeat practices under the instruction from teacher. But as viewed from the view of the cognitive neuroscience, both learning and education are closely relative with the development of brain, because the brain is an information processor which can continually adapt the stimulus from the exterior environment.

The so-called “learning” is the process constructing the nerve centre circuit because of exterior environmental stimuli (except for oneself). The “education” is the process to control and complement exterior stimulus (H. Koizumi, 1999, p.5-24).

The learning is the process that the brain reacts to the environmental stimulations through establishing the nerve link, and this nerve link in the process is used in the information disposal circuit and the information memory. The education is to lead and inspire the construction of brain information disposal structure through offering, controlling and inputting stimulus for learners. The learning process includes many postnatal factors such as the formation of basic nerve link in the infant period, the formation of long term memory, and the reconstructions of some functional areas.

According to this concept, in the process of education, pedagogues should add effective stimulus to help learners construct corresponding nerve circuit and enhance their learning efficiencies. However, many problems such as how to discriminate “effective stimulations” to fulfill different learners’ demands, whether the “learning” from exterior stimulations breaches the “education” of exterior stimulus controlled and offered by the school and so on also need us to further study.

3.2 To treat the relationship between brain science and education from the dimension of “life-long education”

As viewed from systematic opinion, the “learning and education” can be treated as a whole concept from the fetus period, the infant period, the youth period, until to the aged period and death (seen in Figure 1).

As seen in Figure 1, the above “learning” and “education” are the lifetime concepts from birth to death. As viewed from the view of the research of “brain science and education”, the “learning and education” has been enlarged from the past narrow concept to the concept of “lifelong education”. Through experiments, we have found that human could learn from natural environment, cognize space through vision, take love through touch, be familiar with tone and rhythm through music, and understand phoneme and rules to produce expression through language from birth to death, which was induced by the feeler combination in the brain. So we can utilize language, music and natural environment to cure the group with obstacles. Otherwise, from the water environment in the fetus period to child education, elementary, middle and higher education, language education, obstacle child education, healing, occupational education, brain disease prevention, violence prevention, education economics, we can re-research these phenomena about learning and education from the view of biology. From the Figure 1, the infant period is the important period to construct the basic nerve circuits such as thinking and feeling.

In the infant period, even the fetus period, we should offer the environmental stimulus according with the growth of children possibly, which will influence the lifelong development for children. Most parts of this work are completed by the children’s parents, so to enhance the parents’ or quasi-parents’ qualities, means enhancing the quality of the whole nation, and further enhancing the strength of the country to some extents.

4. The revelations of the developments of brain science on early education

Generally speaking, the early education is the education from zero to six years old. In this period, many modes such as game can cultivate good abilities including language, painting, space, computation, music, sports, feeling and communication for children and the educational activities to promote their developments are called early education.

According to the up to date developments of world countries to the research of brain science and the new concept of “learning and education” put forward by Japanese brain scientist H. Koizumi, we thought that the early education should pay more attention to the infants’ potentials and implement education as early as possible according to the natural rule of child development, that is to say, offer them relative stimuli environment as early as possible to establish the nerve centre circuits according to their individualities, and grasp the “intellect dawn” occurred in the development process of infant and timely supply appropriate stimulus to fulfill their demands to the development.

4.1 Emphasizing the “potential ability” infants’ developments and instruct them early

Japanese early education expert Kimura thought that the children potential ability followed the rule of descending. Supposed that the child with 100 degree potential ability when he was born, and if since his born, we offer ideal education for him, so he may become an adult with 100 degree ability. If we teach him since he is 5 years old, even if the education is very excellent, he can only become the adult with 80 degree ability. And if we teach him since he is 10 years old, no matter how good the education is, he can only achieve the adult with 60 degree ability. In another words, the education begins later, the chance that the child can realize his ability is less. That is the famous child potential ability descending rule (Kimura. Kyuichi, 2005).

Some researches showed that though the inheritance is important, but in the growth process of every child, the early education also exerts large function and build the character from inheritance, and the influence of early education is much stronger than the inheritance. The early education exerts sustainable influences to one’s lifetime, and when the infant is just born, his brain is in the status of extreme vacuum and possesses hundreds of millions brain cells and nerve cells which are the physiological base for their lifelong thinking, communication and learning. However, these nerve cells are not linked to be nerve circuits with complex functions. Obviously, the stimulation exerts large function to form and develop the nerve circuit, and it also can change and adjust the nerve circuit system in the development. The learning starts from the initial several hours in life and with the advent of new stimulation, the brain of the infant form and strengthen millions nerve knots in nerve cells to construct the individual nerve centre circuit system.

The following experiment indicated the neonate possesses special sensitivity to the language. According to the result that researchers measured neonatal hearing, we can clearly see that the neonate has obvious reaction to the mother language. In the experiment, researchers used the “optical tester” to measure the neonate in Italy when they were born in five days (seen in Figure 2).

According to Figure 2, when the neonate hear the tape of mother language (Italian) conversation, because the red area represent the activity of brain cell in the brain, so we can think the reactions in the brain are very obvious, after that, when the tape is over and rewind, because the voice is not the language, we can see the red area in the brain becomes very weak, which means the reaction in the brain is very small, and when the tape stops, the red area basically disappears, which means there is not any reactions in the neonatal brain, and thus it can be seen the neonate possesses special sensitivity to the mother language.

We thought that immeasurable potential abilities exist in the children, and we should offer them appropriate environmental stimulus as early as possible, i.e. construct the corresponding nerve centre circuits in the children’ brain as soon as possible, and make their various potential exert, which is the premise to develop the early education.

4.2 Flowing the “intellect dawn” for infants’ developments and instruct them timely

The Germany famous educationist Karl Weter emphasized that the education to the children must begin with children’s “intellect dawn”. In the process of early education, the children are been found to possess “intellect dawns” in different stages, and if we can grasp these “intellect dawns” written in water and timely teach children, we will get twice the result with half the effort. The similar opinion is another concept of “critical period”. The Nobel Prize of medicine and physiology winners Hubel and Wiesel had put forward the concept of “critical period”, and they thought no matter what children do, when they miss the special period, they will be difficult to be cultivated in certain aspect, that means, there is a time limitation. The term is the important shed whether the children can grasp certain ability, and it is called as “critical period”. About the experiments of “critical period”, the example which was cited most is the vision experiment of a little cat. Put one just born cat in the environment that the walls in all sides are painted with transverse lines, after two weeks, when this cat enters into usual environment, the obstacles will appear in the vision of the cat, and the cat doesn’t know vertical lines, and it only know the transverse lines on the horizontal direction. That means two weeks since the cat is born is the “critical period” to obtain normal vision.

When we investigate the reason which induces the obstacle of cat vision, we find that the body of the cat is normal. The cat originally possesses the visual function, only in the growth period of cat vision, we don’t offer it appropriate environment, so the cat doesn’t establish its normal vision.

In the same way, in the growth process of children, there are many periods of “intellect dawn”, and when every “intellect dawn” comes, children will be very interested to certain one of special abilities such as painting, music and language and grasp it well. When the period is missed, to one child, he will be very difficult to grasp this sort of ability. Some researches indicated that children’s “intellect dawn” of language was before their 8 years old, and they investigated the relationship between age and present English ability to Chinese and Koreans when they migrated, and the result showed that the residents they migrated before they were 7 years old, they possesses same English ability with American, and when they exceeded 8 years old, their English learning abilities would toboggan, and once they exceeded 15 years old, their English learning abilities were not relative with the age of migration. That is to say, the plasticity of language intellect will achieve flood tide when children are in about 8 years old, and it toboggans when

children's ages exceed 15.

The unfortunate girl Gini of US also shows that in the "critical period" of language, if children can not obtain enough language stimulation, they will lose their abilities to obtain language. When Gini was 20 months old, she was locked in a dark room by her father, and she was forced to separate with her mother and brother of 6 years old, and in 12 years' underground living without any freedom, what Gini heard and felt were only father's abuses and maltreatments. When she was rescued in 12 years old, she could only said several simple words such as "stop" and "no". Through the research to Gini, experts knew that the development of language needed exterior stimulations, and when children are separated with language environment, they would lose the ability to grasp the language.

We thought whether "critical period" or "intellect dawn", all indicated children would come forth the learning tendency to certain ability in the appropriate period, we should timely grasp this sort of tendency and quickly offer corresponding environment to fulfill children's curiosity and learning desire, which is the process to strengthen children's corresponding nerve centre circuit in their brains, and that is the key to develop early education.

4.3 Offering the "right soil" for infants' developments and instruct them naturally

"Right soil" is the good environment to suit for children's developments. To offer excellent learning environment for children is the guarantee for the success of early education. The "Suzuki Method" put forward by Japan famous educationist Shinichi Suzuki also emphasized "human is the son of environment". Suzuki thought if parents could pay attention to children's instincts and offer them the ideal environment, so all children will obtain a series of very outstanding abilities, and this theory was the same with children all over the world, and there was no exception (S. Suzuki, 1981). He gave the examples that the children born in the tropic are sent to the cold area to grow up, they will also produce anti-cold abilities, and the children born in China are sent to US to grow up, the English will become into their mother language. That indicates many human abilities are cultivated by the environment, not naturally.

The basic concept of "Suzuki Method" is that the music is just like a sort of language, and to grasp one sort of language, we must begin learning it from childhood, the begin is earlier, we can grasp the language like mother language. The music leaning also follows same principle, only children have the ability to grasp certain one musical instrument a little, parents should let them to study. When we teach children, at the very start, we should not teach them music score or any music theories, because that breaches natural rule just as teaching lettering and language grammar to a normal child when he could not speak.

The "Suzuki Method" emphasized letting children listen music first and following the natural rule. Before children learn a new music, however the rhythm is simple, they must repeat to listen the record of this music. And the method of listening should be also natural, they don't need absorbedly sit there and listen facing the recorder, they only listen the music in the most relaxed estates such as dining, gambol, various rest and half rest estates before falling asleep. Though exerting a subtle influence on children to be familiar with the music unconsciously, teacher will let them play the music on the musical instrument under his instruction, far from looking at music book or learning music theory, and the whole process is so natural like children learn to speak when they hear something.

Suzuki had oppugned that "have you seen the child without physiological deficiency could not speak?", and the learning of language was so complex, but there were no failing examples, and other learning could not fail of course, and it is obvious that there were only failed educations or teachers, there were no failed students. Whether teachers or parents follow natural rule, and don't compel to create an appropriate learning environment for children is extremely important.

We thought in the implementation of early education, the soul of Suzuki Method should be extended, and parents and teachers should pay more attention to offer appropriate learning environment for children, and adopt the rule to teach children according to the nature, which is the guarantee to successfully develop early education.

5. Conclusions

Children possess infinite potentials, the premise to develop early education is to offer appropriate stimulus for children and construct corresponding nerve centre circuits in children's brains. The "critical period" or "intellect dawn" also indicate that children would come forth the learning tendency to certain ability in the appropriate period, we should timely grasp this sort of tendency and quickly offer corresponding environment to fulfill children's curiosity and learning desire, which is the process to strengthen children's corresponding nerve centre circuits in their brains, and that is the key to develop early education. In the process of early education, we should adopt the natural rules to teach children, which will certainly fully enhance the efficiency of early education, and offer guarantees to successfully develop early education.

References

- H. Koizumi. (1999). *A Practical Approach to Trans-disciplinary Studies for 21st Century*. Seison Life Sci. No.9. p.5-24.
 Kimura Kyuichi. (2005). *Early Education and Genius*. Beijing: China Literature and History Press.

Konishi Yukuo. (2003). *Baby and Brain Science*. SHUEISHA.

Liu, Shanxun. (2001). How the Brain Potential is Wasted? *Digest of Science and Technology*. No.8.

Sarah-Jayne Blakemore & Uta Frith (2005). *The Learning Brain: Lessons for Education*. Published by Blackwell Publishing Ltd., Oxford.

S. Suzuki. (1981). *Ability Development from Age Zero*. Tokyo: Senzay.

S. Suzuki. (1969). *Nurtured by Love*. OH: A Senzay Publications.

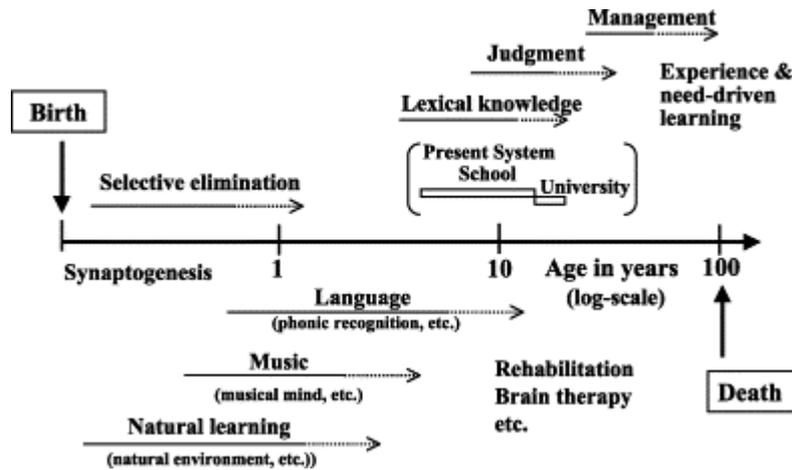


Figure 1. Life-long Learning and Education (A Natural Curriculum) (H. Koizumi, 2004)

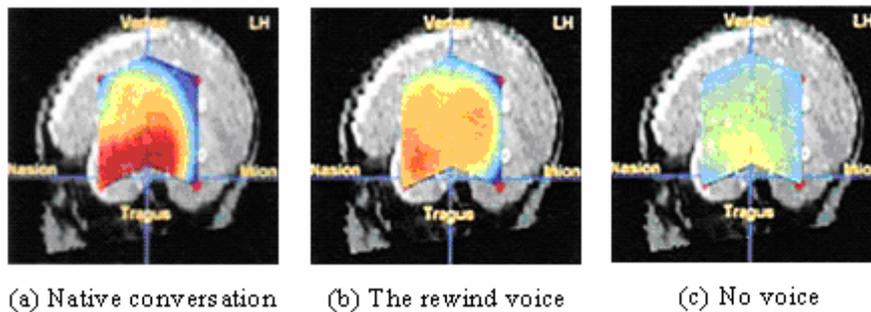


Figure 2. The Neonatal Brain Activity when Heard Native Language (Neonate: In Five Days after Birth & Language: Italian) (J. Mehler's Group, 2003)



Similarity Matrix Based Session Clustering by Sequence Alignment Using Dynamic Programming

Dr.K.Duraiswamy

Dean, Academic

K.S.Rangasamy College of Technology

Tiruchengode, India

V. Valli Mayil (Corresponding author)

P.G.Department of Computer Science

Kongu Arts and Science College, Erode, India

Tel: 91-0424-233-9149 E-mail: vallimayilv@gmail.com

Abstract

With the rapid increasing popularity of the WWW, Websites are playing a crucial role to convey knowledge to the end users. Every request of Web site or a transaction on the server is stored in a file called server log file. Providing Web administrator with meaningful information about user access behavior (also called click stream data) has become a necessity to improve the quality of Web information and service performance. As such, the hidden knowledge obtained from mining, web server traffic data and user access patterns (called Web Usage Mining), could be directly used for marketing and management of E-business, E-services, E-searching , E-education and so on.

Categorizing visitors or users based on their interaction with a web site is a key problem in web usage mining. The click stream generated by various users often follows distinct patterns, clustering of the access pattern will provide the knowledge, which may help in recommender system of finding learning pattern of user in E-learning system , finding group of visitors with similar interest , providing customized content in site manager, categorizing customers in E-shopping etc.

Given session information, this paper focuses a method to find session similarity by sequence alignment using dynamic programming, and proposes a model such as similarity matrix for representing session similarity measures. The work presented in this paper follows Agglomerative Hierarchical Clustering method to cluster the similarity matrix in order to group similar sessions and the clustering process is depicted in dendrogram diagram.

Keywords: Clustering, Sequence alignment, Dynamic Programming, Agglomerative Clustering, Similarity matrix, Preprocessing, Session

1. Introduction

The World Wide Web is continuously growing with the large volume of transaction information and access request of web server. The task of obtaining hidden knowledge from web log file is called Web Usage Mining. As such, the enormous information in log file, it can not be used directly to obtain the knowledge. Data mining from web access log is a process consisting of three consecutive steps: (1) data gathering and pre-processing for filtering and formatting the log entries,(2) pattern discovery which consists of the use of a variety of algorithms such as association rule mining, sequential pattern analysis, clustering and classification on the transformed data in order to discover relevant and potentially useful patterns, and finally,(3) pattern analysis during which the user retrieves and interprets the patterns discovered.

The work in this paper follows sequence of processes as follows. Identification of session information from server log file is one of the important tasks in preprocessing of log file. Session is nothing but a sequence of web pages accessed in a period of time. Continuous click of web pages is identified as a session. Clustering of user session starts with finding similarity between all pairs of user sessions. Here we use the sequence alignment using dynamic

programming technique to find session similarities. A matrix is constructed for maintaining session similarity values and it is clustered based on Agglomerative technique. The knowledge obtained after grouping similar web session's is used to group the visitors of the system who has same access behavior in a system.

1.1 Related work

Most of the studies in the area of web usage mining are very new, and the topic of clustering web sessions has recently become popular in the field of real application of clustering techniques. Shahabi et al. [5] introduced the idea of Path Feature Space to represent all the navigation paths. Similarity between each two paths in the Path Feature Space is measured by the definition of Path Angle which is actually based on the Cosine similarity between two vectors. In this work, k-means cluster method is utilized to cluster user navigation patterns. Fu et al. [4] cluster users based on clustering web sessions. Their work employed attribute oriented induction to transfer the web session data into a space of generalized sessions, then apply the clustering algorithm BIRCH [6] to this generalized session space. Most of the previous related works apply either Euclidean distance for vector or set similarity measures, Cosine or Jaccard Coefficient. Shortcomings for doing this are obvious. Another method of grouping the session is finding similarity measure between sessions and clusters them accordingly. Fig 1.1 discusses the sequence of procedures of grouping the session

2. Session Similarity Measure by sequence alignment

As the Web log contains data of every request of the server, it is to be preprocessed to obtain the relevant data, consists of sequence of sessions. The first and foremost question needed to be answered in clustering web sessions is how to measure the similarity between two web sessions. A web session is naturally a stream of hyper link clicks. Here we consider the original session data as a set of sequences, and apply sequence alignment method to measure similarity between sessions.

FastLSA [8] is a dynamic programming algorithm produces the optimal alignment for a given scoring function. We use sequence alignment techniques to analyze the sequence of user requests that appear in user sessions. For two strings of length m and n , optimal sequence alignment has zero or more gaps inserted into the sequence to maximize the number of positions in the aligned strings that match.

2.1 Session Similarity Measure by Dynamic Programming

Following data cleaning and preprocessing steps, the session similarity measure by dynamic programming is used in calculating the similarities between all pairs of session. Since user sessions are ordered URL requests, the session is referred as sequences of Web pages. The problem of finding the optimal sequence alignment is solved using dynamic programming approach [7,3,2]. Briefly, the algorithm consists of three steps. The first step is initialization where a dynamic programming matrix is created with $K+1$ columns and $N+1$ rows where K and N correspond to the sizes of the sequences to be aligned. One sequence is placed along the top of the matrix (sequence#1) and the other one along the left-hand-side of the matrix (sequence#2). A gap is added to the start of each sequence which indicates the starting point of calculation of similarity score.

To translate a matrix path to an alignment, follow the path from the top left to the bottom right. Every diagonal move corresponds to aligning two letters as either a match or a mismatch. A right move corresponds to the insertion of a gap in the vertical sequence. Down move corresponds to the insertion of a gap in the horizontal sequence. Any path can be translated to an alignment, Consider the session sequences as in table 2.1.

Considering first two sequences from table 2.1

S1=P1,P5,P7,P3,P6

S2= P1,P5,P3,P6

The alignment process, such as gap may be inserted between sequences S1 & S2 in order to make it match

S1 = P1 P5 P7 P3 P6
S2 = P1 P5 - P3 P6

The alignment path matrix is given in the following table 2.2

Two sequences are scanning from left to right and the corresponding arrow is placed in the table.

Diagonal arrow - Match or mismatch of pair of web pages from sequences

Right arrow gap on top

Down arrow - gap on left

In order to obtain the optimal alignment for a given scoring function, we need to identify the corresponding *optimal path*. To derive the optimal path in the matrix, the algorithms can be divided into two phases, which we call *FindScore*

and *FindPath*. Table 2.3 shows the DPM (Dynamic Programming Matrix) scores for the example sequences that are computed during the *FindScore* phase. The entries with numerical subscripts form the optimal path, which is computed in the *FindPath* phase.

In the *FindScore* phase, a 0 is placed in the upper-left corner of the matrix. Each algorithm propagates scores from the upper-left corner of the matrix to the lower-right corner of the matrix. The score that ends up in the lower-right corner is the optimal alignment score.

The score of any entry is the maximum of the three scores that can be propagated from the entry on its left, the entry above it and the entry above-left + Scoring function. For every pair of identical pages a positive score of value 20 is given as scoring function and for mismatch or gap negative value -10 is given as scoring function. For Example the score of 20₅ in (P1/P1) entry is the maximum of the score from its left entry (-10 + -10), above entry (-10 + -10), and above left entry (0 + 20) is assigned. A path was read from the top left to the bottom right. The entry in lower right corner is the value of optimal score. In this example the optimal score value is 70. Following algorithm discusses the methods to obtain DPM and optimal score.

Algorithm: Optimal similarity score value for 2 sessions

Input: Pair of session sequences S1,S2

Output: Optimal Similarity score Value

- 1 Matrix DPM is created with K+1 columns and N+1 rows where K and N correspond to the sizes of S1 & S2 sequences respectively
- 2 Align the sequence by providing gap in between the sequence so two sequences can be matched as much as possible
- 3 Place sequence S1 in top of Matrix and sequence S2 on left side of Matrix
- 4 Assign top left corner of matrix = 0;
- 5 Find optimal path

a) Find score

Find DPM with every cell entry = max (left entry, right entry, above left

Entry + scorevalue)

Where score value = 20 for pair of matching pages

Score value = -10 for mismatch pages.

b)Find path

For every 2 sequences the pair of web pages is read and arrow is placed as per the following rule.

Construct the arrow as follows

Diagonal arrow - Match or mismatch

Right arrow gap on top

Down arrow - gap on left

- 6 Traversing from top left to lower right corner of DPM , the optimal score value is available at lower right corner

2.2 Session Similarity measure

Session similarity measure [3] has two components, such as *alignment score* component and *local similarity component*. The alignment score component computes how similar the two sessions are in the region of their overlap. If the highest value of the score matrix of two sessions, s1 & s2, is v, and the number of matching pages is M, matching score value is s (m), then the alignment score component Sa is:

$$Sa(s1,s2) = v / (s(m) * M)$$

This value is normalized by the matching score and the number of matching pages. The local similarity component computes how important the overlap region is. If the length of the aligned sequences is L, the local similarity component is

$$Sb = M/L$$

Then the overall similarity between two sessions is given by

$$\text{sim}(S1,s2) = S_a * S_b.$$

For the above example the similarity measure of two session S1 & S2 is 0.7

$$\text{Sim}(S1,S2) = 70/(20*5) = 0.7$$

In this way, Similarity value of all pairs of session sequences of table 2.1 are found and depicted in table 2.4 & table 2.5. The highest value of two pair of session defines the more similarity between them and lowest value shows the dissimilarity between the pages.

3. Agglomerative hierarchical Clustering

Clustering is the process of grouping objects into clusters such that the objects from the same cluster are similar and objects from different clusters are dissimilar. Similarity matrix obtained in previous section with largest value has high similarity whereas the less number have low similarity.

The objective of clustering in similarity matrix is to group the data with similar characteristics. In this work the sessions are arranged in similarity matrix and extracting knowledge from them, groups the similar session together. In this work Agglomerative hierarchical clustering process is used to cluster the similarity matrix. The process starts by finding the clusters with the high similarity value and putting those clusters into one cluster. The hierarchy of clusters formed can be represented by a dendrogram. Single linkage nearest neighbor technique is used across cluster to determine the similarity between the clusters.

Consider the array of similarity as given in table 2.4. The highest similarity value in the matrix is 70, available for S1 & S2, and at first stage sessions S1, S2 are combined to form second stage of matrix as follows.

The next highest similarity value is 50 for the sessions S4, S5. Now sessions S4, S5 are combined and clusters (S1,S2) & (S4,S5) are formed. The session S1, S2, S3 are combined, with the session similarity value 40, and (S1, S2, S3), (S4, S5) clusters are available; finally these two clusters are formed to produce resultant clusters. The dendrogram of clustering process is shown in fig 3.1

Algorithm : Agglomerative Hierarchical clustering

Input : Similarity matrix

Output : Dendrogram cluster

1 Select the largest similarity value from matrix and its session is S_i, S_j and combine and form its composition $S_{i,j}$

2 Form a matrix with $S_{i,j}$

3 Find the cell value of matrix as

$$\text{Similarity}(S_{i,j}, S_k) = \min \{ \text{similarity}(S_i, S_k), \text{Similarity}(S_j, S_k) \}$$

4 Repeat step 2 until single cluster in matrix cell.

4. Conclusion

Categorizing visitors or users based on their interaction with a web site is a key problem in web usage mining. This paper focuses on clustering the session details obtained from logfile. Session is a sequence of web pages accessed within a period. Session details are obtained after the preprocessing of log file. The work presented in this paper finds the similarity between every pair of session by sequence alignment using dynamic programming and similarity matrix is clustered using Agglomerative hierarchical technique in order to group the same similarity session. The objective of this process is to create a recommender system which will be used in finding group of visitors of session with the same nature in E-learning system, finding group of visitors with similar interest, categorizing customers in E-shopping etc or group of customers' session in marketing environment.

References

- A. Banerjee and J. Ghosh. (2001). Clickstream clustering using weighted longest common subsequences. In Proc. of Workshop on Web Mining in First International SIAM Conference on Data Mining, pages 33–40, Chicago, April 2001.
- Aaron Davidson A Fast Pruning Algorithm for Optimal Sequence Alignment.
- S., Gunduz and M. T. Ozsu. (2003). A web page prediction model based on click-stream tree representation of user behavior. In Proceedings of Ninth ACM International Conference on Knowledge Discovery and Data Mining (KDD), August 2003.
- Y. Fu, K. Sandhu, and M.-Y. Shih. (1999). Clustering of web users based on access patterns. *WEBKDD* workshop, 1999.

C. Shahabi, A. Zarkesh, J. Adibi, and V. Shah. (1997). Knowledge discovery from users web-page navigation. *In workshop on Research Issues in Data Engineering*, England.

T. Zhang, R. Ramakrishnan, and M. Livny. (1996). BIRCH: an efficient data clustering method for very large databases. *In ACM SIGMOD*, pages 103–114, June 1996.

Weinan Wang Osmar R. Zaiane Clustering Web Sessions by Sequence Alignment.

K. Charter, J. Schaeffer, and D. Szafron. (2000). Sequence alignment using fastlsa. In Proceedings of the 2000 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS'2000), pages 239–245.

Karypis, E.-H. Han, and V. Kumar. (1999). Chameleon: A hierarchical clustering algorithm using dynamic modeling. *IEEE Computer*, 32(8):68–75, August 1999.

Table 2.1 Session sequences

Session number	Sequences
S1	p1,p5,p7,p3,p6
S2	p1,p5,p3,p6
S3	p5,p7,p6
S4	p2,p8,p6,p5
S5	p2,p6,p5

Table 2.2 alignment path of session S1 & S2

	-	P1	P5	P7	P3	P6
-						
P1						
P5						
P3						
P6						

Table 2.3 DPM - optimal score value

	-	P1	P5	P7	P3	P6
-	0	-10	-20	-30	-40	-50
P1	-10	20 ₅	10	0	-10	-20
P5	-20	10	40 ₄	30 ₃	20	10
P3	-30	0	30	30	50 ₂	40
P6	-40	-10	20	20	40	70 ₁

Table 2.4 optimal score value of pair of sessions

	S1	S2	S3	S4	S5
S1	---	70	40	-50	10
S2	70	----	20	-20	-20
S3	40	20	----	-10	-10
S4	-50	-20	-10	---	50
S5	10	-20	-10	50	----

Table 2.5 Similarity measure Matrix

	S1	S2	S3	S4	S5
S1	1	0.7	0.4	-0.5	0.1
S2		1	0.25	-0.25	-0.25
S3			1	-0.25	-0.125
S4				1	0.625
S5					1

Table 2.6 Similarity matrices on applying single linkage rule

	S1,S2	S3	S4	S5
S1,S2	1			
S3	40	1		
S4	-20	-10	1	
S5	-20	-10	50	1

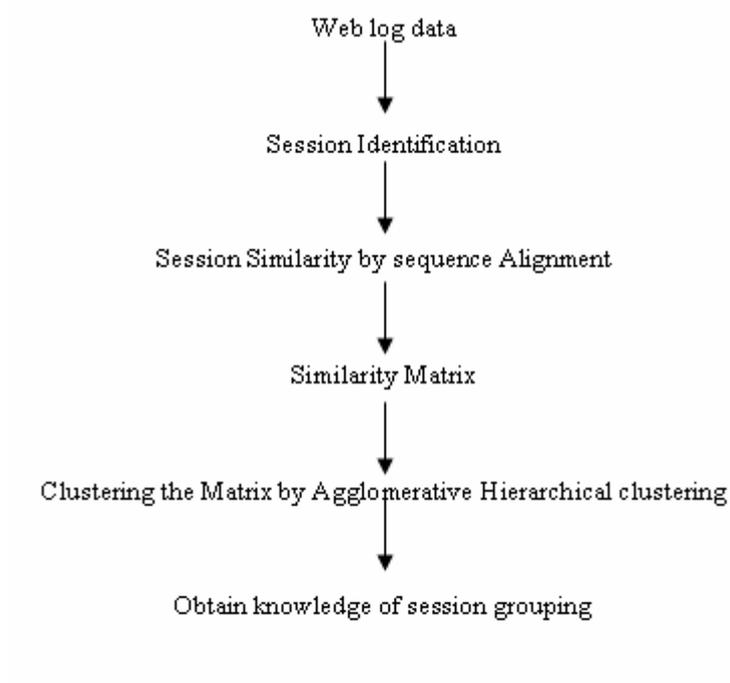


Figure 1.1 procedures for grouping the session

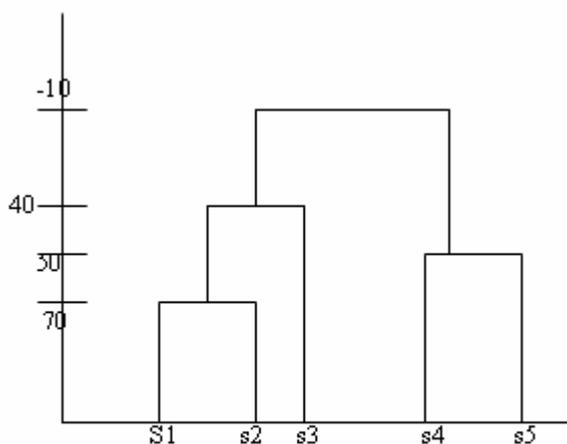


Figure 3.1 dendrogram for session clustering



Research of Information System of Monitoring Remote on Distributed Micro-turbine

Liangwei Zhong

CAD center

University of Shanghai for Science and Technology

Shanghai 200093, China

Tel:86-21-5527-6487 E-mail: zlvcad@126.com

Kangmin Chen

College of Power Engineering

University of Shanghai for Science and Technology

Shanghai 200093, China

Tel:86-21-5527-4627 E-mail: chenkm@usst.edu.cn

Jing Ni

Business School

University of Shanghai for Science and Technology

Shanghai 200093, China

Tel:86-21-5527-6487 E-mail: nijing501@126.com

Anyang He

CAD center

University of Shanghai for Science and Technology

Shanghai 200093, China

Tel:86-21-5527-6487 E-mail: heay@yahoo.com

Supported by Shanghai Leading Academic Discipline Project (T0502) and Shanghai Education Committee (05EZ30)

Abstract

Aiming at the problems existing in fault diagnosis of Distributed micro gas turbine, a kind of new information system frame of fault Diagnosis on Distributed micro gas turbine is proposed in this paper. The consumers and the manufactures contact with each other through Internet, share the corresponding data of the Distributed Micro-turbine, and diagnose and solve the faults together. Through the study of collecting method of the micro gas turbine generating sets' vibration signals, the technique of gathering reliable vibration signals is designed. The traditional method is revised by this new system, and this can greatly improve the efficiency of Fault Diagnosis, reduce the cost of equipment maintenance, and be in favor of the accumulation and updating of techniques in this field.

Keywords: Fault Diagnose, Distributed Micro-turbine, Monitoring Remote

Compared with the traditional Power station, Distributed Generation refers to the small-scale power generation system locating near the consumers' houses or moving with the consumers. Distributed Generation system is equipped with the advantages like enhancing energy utilization, improving security, avoiding environment pollution and so on. However, when there are faults appearing in the system, the operating crew must detect the state and parameters of the Distributed Micro-turbine on site, and at the same time, the closing down of the Distributed Micro-turbine will bring economic loses to the consumers. Fortunately, with the rapid development of remote telecommunication technology, the increasing abroad application of Internet and the tendency of perfection of CSCW technology, the remote fault diagnose is becoming possible.

1. Remote monitoring information system architecture

Figure 1 shows the system global architecture. Because the location of the Distributed Micro-turbine is scattering, spindle speeding, the natural frequency of rotor, the temperature of axial bearing, the frequency of vibration signals and other data of the Distributed Micro-turbine are collected through intelligent collecting terminal. All these data are sent to the consumer monitoring center by WEB or GPRS while the Distributed Micro-turbine is being monitored. By using the system information management software, we can monitor the Distributed Micro-turbine online and practice the remote fault Distributed Micro-turbine. The manufacturer can contact the consumer monitoring center through Internet, exchange and update the data, and help the consumer to diagnose the faults and repair the system. Both the consumer monitoring center and the manufacturer maintenance center are equipped with three databases. All kinds of data concerning faults are saved in database 1, the dynamic refreshing data of working condition are saved in database 2, and the engineering data of CAD/CAPP applicable to the local Distributed Micro-turbine are saved in database 3.

2. Function of system

By the usage of system information management software through GPRS from long distance, the Distributed Micro-turbine's working condition can be supervised. Usually, only the information of the Distributed Micro-turbine's working condition in the latest three months is stored in the consumer monitoring center, and the information is continuously updated. Once there is something wrong with one of the Distributed Micro-turbines, the working condition within one week before and after the fault clearance will be stored in the database of the consumer monitoring center. The manufacturer maintenance center and the consumer monitoring center exchange the fault information and update the database periodically. If the engineers for the consumers can not solve the faults, they can ask help from the engineers of the manufacturers and clear the faults jointly. The record of the fault clearance will be added into the database. By using this system, not only the Distributed Micro-turbine can be diagnosed and monitored from a long distance, but the consumers and manufacturers can contact with each other closely and share the data. Thus, the efficiency of fault Distributed Micro-turbine and maintenance can be greatly improved, and the cost of equipment maintenance can also be reduced a lot.

3. Characteristic of system

Because the Distributed Micro-turbines are located separately and their location might be changed continuously, we can supervise them effectively through GSM/GPRS. By using GSM/GPRS, we can diagnose and locate the Distributed Micro-turbines with something wrong so as to clear the faults and maintain the equipment conveniently. In addition to this, through the application of control equipment, the Distributed Micro-turbine can be supervised online, and this can not only cut the cost but guarantee the personal security. The consumers and manufacturers can contact with each other through Internet, exchange the corresponding data of the Distributed Micro-turbine, and update and enrich the databases of faults and working condition. And this can be detailed data support for the future fault Distributed Micro-turbine and replacement of the older generation by new ones of the Distributed Micro-turbine. This system can greatly improve the efficiency of equipment Distributed Micro-turbine and maintenance, reduce the cost of equipment maintenance, and be in favor of the accumulation and updating of techniques in this field. It is a kind of excellent management pattern which can benefit both the consumers and the manufacturers.

4. Consumer monitoring center, manufacturer maintenance center & software and hardware configuration

The configuration of the consumer monitoring center and the manufacturer maintenance center is similar. They both need the network environment of Internet, by which the consumer monitoring center and the manufacturer maintenance center can share the audio and video frequencies, examine the files, extract the data and modify the parameters. The Distributed Micro-turbine can be monitored through video frequencies by the CCD cameras, and this is convenient for the remote controlling engineers and those in the manufacturer maintenance centers to observe and diagnose the faults of the Distributed Micro-turbine.

Because the fault Distributed Micro-turbine bears the multi-media characteristics including all kinds of filed data, acoustic data, pictures and so on, databases supporting the multi-media function are required. The databases should also be equipped with the function of unfailing visiting because of the dynamic characteristics of them. In addition to that, the databases are required to have the open interconnect function as the consumers' needs are different. Oracle database can satisfy all the requirements stated above, and is a kind of suitable application database for the consumer monitoring center and the manufacturer maintenance center.

5. Operating mode of the online fault diagnose system

The intelligent collecting terminal collects the working condition of the Distributed Micro-turbine periodically, and all this information is stored in the working condition database of the Distributed Micro-turbine's consumer monitoring center. The working condition within one week before and after the fault clearance will be stored in the database chronically so as to be used as reference when there is anything wrong. The Distributed Micro-turbine methods applicable to this system include the method based on statistics, the method based on expert system and the method

based on the artificial neural network. The manufacturer maintenance center of the Distributed Micro-turbine can transfer the data in the consumer monitoring center under the following two kinds of conditions:

- (1) When the local engineers can not solve the faults of the local Distributed Micro-turbine and ask for help to the manufacturer maintenance center, the engineers in the manufacturer maintenance center can have the right to open the consumer's Website and transfer the data of fault and working condition within a certain period of time before and after the occasion.
- (2) The manufacturer maintenance center can transfer the new information, successful solving methods and unsuccessful experience in the fault database of the consumers all over the world periodically through the Internet, so as to enrich its own database and provide better service to the consumers.

In the remote service PCs of the local consumer monitoring center, a maintenance expert system is built up and its basic mode is provided by the manufacturer maintenance center. Every consumer can have his own faults memorial and store it in the fault database eternally. After the occurrence of some fault (eg. Iron sick), the frequency of the occurrence of the same kind of fault in the same place or different places will be enhanced a lot. So the memorial will benefit the future maintenance of the local Distributed Micro-turbine, and at the same time, the manufacturer maintenance center will surely encourage the behavior of building the faults memorial because they can get more fault cases from it and the expert system can provide different maintenance scheme aiming at different kinds of faults.

The faults which can be removed quickly by the operators are called as the first kind of faults. Towards this kind of faults, the local expert system will give hints to the operators like closing down the machine, examining the corresponding part and taking necessary measures, eg. tightening the screws, alternating unit and so on.

The faults which are already recorded in the local memorial are called as the second kind of faults. All necessary data, steps of fault clearance and notice points are provide by the expert system, and the local engineers can solve the fault easily with all these abundance.

The faults which can not be solved by the local engineers and already recorded in the manufacturer maintenance center are called as the third kind of faults. The consumer can download the record from the manufacturer maintenance center by inputting his own user name, code and account number (if it is necessary to pay). During the process of fault clearance, the local engineers can seek help from the consumer providing this record as he has the experience of solving this kind of fault. And there are several ways seeking help, like E-mail, telephone, fax, video conference call and so on.

The faults which can not be solved by the local engineers and never recorded in the manufacturer maintenance center are called as the fourth kind of faults. When meeting with this kind of faults, the professional maintenance engineers in the manufacturer maintenance center can transfer the data of working condition and faults, and the data of the local engineers' trying repairing. Through the Internet conference system, the remote engineers and the local engineers can discuss the problems by means of shared whiteboard, discourse, expression, gesticulation and other means and set down the maintenance scheme. If it is necessary, the professors in the corresponding field and the consumers ever experiencing the similar faults can be invited to join the discussion, and the consumers all over the world can take part in the discussion through Internet. When solving the fourth kind of faults, we can take the number of the changed parameters, the time cost in repairing and the time the Distributed Micro-turbine ever operating into account comprehensively to make sure the maintenance cost. Solving a new kind of faults can perfect the database in the manufacturer maintenance center, so the manufacturer should encourage this kind of behavior by low charge.

In addition to that, the manufacturer maintenance center can also provide a hyperlink by which the consumers can visit the manufacturer's engineering database in which the data of CAD/CAPP suitable for the machine tools produced by the manufacturer are stored and these data can be used by the consumers with or without charges. The consumers can get the engineering data either through the PCs of local monitoring center or provided by the engineers of the manufacturer. The manufacturer will update and publicize the engineering data for free periodically, and the long-term friendly relationship between the consumers and the manufacturer will be maintained.

6. Conclusion

In this paper, an information system used for remote fault Distributed Micro-turbine of the Distributed Micro-turbine is provided by combining the telecommunication technique, multi-media technique, the collaborative work technique supported by the computers and Internet together. The traditional method is revised by this new system, and this can greatly improve the efficiency of Distributed Micro-turbine and reduce the cost of equipment maintenance. At the same time, this system can tighten the communication between the consumers and the manufacturer, strengthen the capability of fault diagnosis of Distributed Micro-turbine and clearance for both parts, and be in favor of the accumulation and updating of techniques in this field. By the application of this system, the online monitoring of the Distributed Micro-turbine through GPRS and control module is becoming available.

References

Hu, Wenbin. etc. (1997). The Design of an Integrated Fault diagnoses system for Modern Manufacturing Equipments. *JOURNAL OF HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY*. 1997. 01.

Xia, Feng etc. (1997). Design and Implementation of Distributed and Open Telemedicine System. *JOURNAL OF SHANGHAI JIAOTONG UNIVERSITY*. p115~118. 1997. 06

Zhang, Hao and Wu, qidi. (2002). *Teleservice and remote engineering support system of Nanyang's manufacture corporation*. TongJi University Press. 2002.

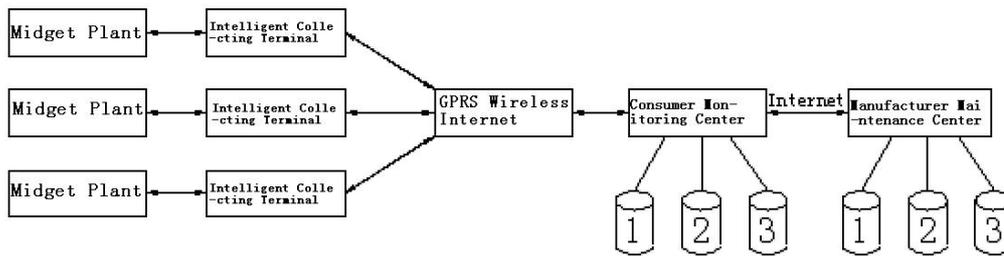


Figure 1. system global architecture



A Novel Sanitization Approach for Privacy Preserving Utility Itemset Mining

R.R.Rajalaxmi (Corresponding author)

Computer Science and Engineering

Kongu Engineering College

Erode – 638 052, Perundurai, TamilNadu, India

E-mail: rrr_kec@yahoo.co.in

A.M.Natarajan

Bannari Amman Institute of Technology

Sathy, TamilNadu, India

Abstract

Data mining plays a vital role in today's information world wherein it has been widely applied in various business organizations. The current trend in business collaboration demands the need to share data or mined results to gain mutual benefit. However it has also raised a potential threat of revealing sensitive information when releasing data. Data sanitization is the process to conceal the sensitive itemsets present in the source database with appropriate modifications and release the modified database. The problem of finding an optimal solution for the sanitization process which minimizes the non-sensitive patterns lost is NP-hard. Recent researches in data sanitization approaches hide the sensitive itemsets by reducing the support of the itemsets which considers only the presence or absence of itemsets. However in real world scenario the transactions contain the purchased quantities of the items with their unit price. Hence it is essential to consider the utility of itemsets in the source database. In order to address this utility mining model was introduced to find high utility itemsets. In this paper, we focus primarily on protecting privacy in utility mining. Here we consider the utility of the itemsets and propose a novel approach for sanitization such that minimal changes are made to the database with minimum number of non-sensitive itemsets removed from the database.

Keywords: Privacy Preserving Data mining, Frequent Itemset mining, Association Rule Mining, Utility mining, Data sanitization

1. Introduction

Large volumes of detailed personal data are regularly collected. Such data include shopping habits, criminal records, medical history, and credit records, among others (Brankovic L and Estivill-Castro V. 1999). These data can be analyzed by applications which make use of data mining techniques. Hence such data is an important asset to business organizations and governments for decision making processes and also to offer social benefits, such as medical research, crime reduction, national security, etc. (Jefferies P. 2000). On the other hand, analyzing such data opens new threats to privacy and autonomy of the individual if not done properly (Culnan M.J, 1993, Vassilios S. Verykios, Ahmed K. Elmagarmid, Elisa Bertino, Yucel Saygin and Elena Dasseni. 2004).

With the conventional data analysis methods there is a limited threat to privacy. Also these techniques mainly present the results based on the mathematical characteristics associated with the data. Making use of such techniques may not reveal some interesting patterns which are hidden in the data. By using appropriate data mining techniques it is possible to explore the hidden patterns. But the threat to privacy becomes real since data mining techniques are able to derive highly sensitive knowledge from unclassified data which is not even known to database holders (Elisa Bertino, Ravi Sandhu, 2005). In order to overcome this issue the data owners may decide not to share or release such data for analysis provided they should make a compromise for exploring hidden knowledge (Estivill-Castro V and Brankovic L, 1999, Atallah M., Bertino E., Elmagarmid A., Ibrahim M., and Verykios V. 1999).

Privacy preservation in data mining is the emerging research area which addresses the different ways to protect the sensitive knowledge discovery (Agrawal R and Srikant.2000, Verykios V.S., Bertino E., Fovino I.N., Provenza L.P.,

Saygin Y., and Theodoridis Y. 2004). It is mandatory to gain the benefit of data mining and as well as maintaining privacy. This paper focuses on the privacy preservation in utility mining. The rest of the paper is organized as follows: Section 2 discusses the background and related work in privacy preserving data mining. Section 3 gives an overview of the utility mining. In section 4 we describe the problem to be solved. Section 5 explains the proposed approach for privacy preservation in utility mining. Section 6 discusses the various experimental results carried and the detailed discussion on the analysis of results. Finally we conclude the paper in Section 7.

2. Background and Related Work

2.1 Frequent pattern Mining

Let $I = \{i_1, i_2, i_3, \dots, i_n\}$ be a set of items. Let D , the task-relevant data, be a set of database transactions where each transaction T is a set of items such that $T \subseteq I$. Each transaction is associated with an identifier, called TID. Let A be a set of items. A transaction T is said to contain A if and only if $A \subseteq T$. A set of items is referred to as an itemset. An itemset that contains k items is a k -itemset. The occurrence frequency or support of an itemset is the number of transactions that contain the itemset (Han J and Kamber M. 2006). If the relative support of an itemset I satisfies a prespecified minimum support threshold, then I is a frequent itemset.

2.2 Privacy preservation of frequent itemsets

Since frequent itemset mining is a preliminary step in the association rule mining, most of the researches have addressed the privacy preservation of frequent itemsets with respect to association rule mining. A heuristic approach is presented in (Charu C. Aggarwal, Jan Pei, Bo Zhang, 2006) to hide the sensitive association rules against adversarial data mining. Randomization is an approach used to protect the discovery of association rules (Brankovic L and Estivill-Castro V. 1999). Additive Perturbation (Muralidhar K., Parsa R. and Sarathy R. 1999) is used to provide security to the databases which does not reveal sensitive information. To hide restrictive patterns which are generated after the mining process a simple and effective way is to decrease their support in a given database (Agrawal D and Aggarwal C.C. 2001, Alexandre Evfimievski, Ramakrishnan Srikant, Rakesh Agrawal, Johannes Gehrke. 2002, Evfimievski A, Srikant R, Agrawal R, and Gehrke J. 2002, Rizvi S. J. and Haritsa J. R. 1994). Data sanitization is the process of altering the transactions and was introduced in (Atallah M., Bertino E., Elmagarmid A., Ibrahim M., and Verykios V. 1999). To do so, a small number of transactions have to be modified by deleting one or more items from them or even changing items in transactions, i.e., adding noise to the data. However this work relies on boolean association rules. The authors also proved that the optimal sanitization problem is NP-hard. On the other hand, the approach must hold the following restrictions: (1) the impact on the data in the database should be minimal and (2) an appropriate balance between privacy and knowledge discovery must be guaranteed. A set of new strategies and algorithms were proposed (Vassilios S. Verykios, Ahmed K. Elmagarmid, Elisa Bertino, Yucel Saygin and Elena Dasseni. 2004) for hiding sensitive knowledge from data by reducing the support and confidence of rules which specify how significant they are. The sensitive transactions are modified by removing some items, or inserting new items depending on the hiding strategy. However the proposed strategies do not consider the undesired side effects of hiding the sensitive rules. To overcome the limitations the authors in (Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen. January 2007) proposed a new approach which classifies the valid modifications for hiding sensitive rules and represents each class of the modifications by three attributes. The work in (Oliveira S. R. M. and Zañane O. R. 2002, Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen. 2007),” concentrates on hiding the frequent itemsets which specifies a new framework for the hiding process. But the authors in (Zhihui Wang, Wei Wang, Baile Shi, S. H. Boey. 2006), proposed a novel approach by incorporating the hiding process in the frequent itemset mining algorithm itself so that it does not generate sensitive frequent itemsets.

The earlier work of association rule mining or frequent itemset mining considers either the presence or absence of items in the transaction. Hence privacy preservation is performed by either decreasing the confidence or support. However in real world the transactions made by the customer consist of quantities of items purchased with unit price of the items. In order to address privacy preservation in such databases a new set of strategies are required. Our work differs from the earlier work of frequent itemset hiding wherein we consider the utility of the items in the transactions.

3. Utility Mining

Earlier studies in frequent itemset mining focus in generating frequent itemsets by considering the presence or absence of items in the transactions. But in practical applications the transactions contain the purchased quantities of the items. In order to address this utility mining model was introduced in (Hong Yao, Howard J. Hamilton, and Cory J. Butz. 2004). Utility of an item specifies how useful an item is. The main goal in utility mining is to find the itemsets which yield high utility. In traditional association rule mining the utility of an item is either 0 or 1 (Rakesh Agrawal Ramakrishnan Srikant. 1994).

Let $I = \{i_1, i_2, \dots, i_m\}$ be a set of items, D be a transaction database, and $UT \langle I, U \rangle$ be a utility table, where U is a subset of the real numbers that reflect the utilities of the items. The *utility mining problem* is to discover all itemsets in a

transaction database D with utility values higher than the *minimum utility threshold*, given a utility table UT . We use the definition of a set of terms that leads to the formal definition of utility mining problem (Hong Yao, Howard J. Hamilton, and Cory J. Butz. 2004). Utility mining is to find all the itemsets whose utility values are beyond a user specified threshold called MUT . An itemset X is a *high utility itemset* if $u(X) \geq MUT$, where $X \subseteq I$ and MUT is the minimum utility threshold, otherwise, it is a *low utility itemset*.

4. Problem Formulation

In this work, our goal is to hide a group of interesting patterns which contains highly sensitive knowledge. We refer to these interesting patterns as sensitive patterns, and we define them as follows:

Definition: Let D is a transactional database, P be a set of all interesting patterns that can be mined from D , and P_s be a set of sensitive patterns that need to be hidden according to some security policies. A set of patterns, denoted by P_s , is said to be sensitive if $P_s \subset P$

The specific problem addressed in this paper can be stated as follows:

Given a transaction database, utility database, MUT , and a set of sensitive high utility itemsets (P_s), how can we modify the database such that using the same MUT , the set of non-sensitive high utility itemsets in the modified database can still be mined?

5. Conflict based Sanitization Approach

To solve this problem, we propose a novel approach called Conflict based Utility Itemset Sanitization (CUIS) that strategically modifies the database to decrease the utility of the sensitive itemsets. The approach is iteratively applied in a greedy fashion until the utility of each sensitive itemset falls below the threshold MUT while the number of non-sensitive itemsets with utility smaller than MUT is minimized. For each sensitive transaction the conflict degree is computed as the number of sensitive items that the transaction is supporting. It selects the transactions with maximum conflict degree for modification. The approach CUIS is described as follows:

Input: the collection P of high utility itemsets, the set P_s of sensitive high utility itemsets., Minimum Utility Threshold MUT .

Output: safe and sharable Database D'

Steps:

$D' = D$;

For each sensitive itemset p in P_s

1. Identify the set of transactions T_s supported by P_s
2. Find the difference of the sensitive itemset p as

$$Diff = U(p) - MUT$$

While ($Diff > 0$) do

3. Choose transaction $T \subset T_s$ such that the conflict degree is maximal

Find the item $i_{max} \in T_s$ which has maximum transaction utility

4. if ($U(i_{max}, T) < Diff$)

- a. $O(i_{max}, T) = 0$

- b. $Diff = Diff - U(i_{max}, T)$

else

- c. $O(i_{max}, T) = O(i_{max}, T) - ((diff / S(i_{max}))$

- d. $Diff = 0$

The set of transactions supported by the sensitive itemsets are identified (Step 1). For each of the private itemset the utility difference is computed (Step 2). Selecting the transaction with the maximum conflict degree can reduce the number of non-sensitive itemsets deleted in the sanitization process and the victim item which has maximum utility is selected for modification (Step 3). The process is repeated until the utility of the sensitive itemset falls below the MUT (Step 4).

6. Experimental Analysis

To measure the effectiveness, we adopt the set of metrics proposed in (Oliveira S. R. M. and Zaniane O. R. 2002) in terms of information loss and non-sensitive patterns removed as a side effect of the transformation process. The performance measures are specified as follows:

Misses Cost: It denotes the percentage of legitimate patterns that are not discovered from D'

$$MC = \frac{\# \sim Ps(D) - \# \sim Ps(D')}{\# \sim Ps(D)} \quad (1)$$

where $\# \sim P_s(X)$ denotes the number of non-sensitive patterns in the database X .

Original and Sanitized database Difference: It denotes the difference between the original (D) and sanitized database (D')

$$\text{diff}(D, D') = \frac{|D - D'|}{|D|} \quad (2)$$

To measure the effectiveness of the algorithm, experiments were conducted on the synthetic datasets generated using IBM synthetic data generator (<http://www.almaden.ibm.com/software/quest/Resources/index.shtml>, 2003). It is highly advanced and considers typical properties of real transactional databases such as high frequencies of some itemsets, mean length of transactions, etc. This generator can produce only a binary form of transactional databases. Therefore, internal and external utilities were generated separately from the log-normal distribution in range $[1, \dots, 10]$ (internal) and $[1, \dots, 20]$ (external). Table 1 lists the characteristics of the datasets used in the experiment.

The experiments are conducted on a PC with Pentium IV processor (2 GHz) having 256MB main memory, running Windows XP. The effectiveness of the algorithm is studied based on the following condition: we varied the number of sensitive itemsets to hide and the disclosure threshold is zero. Based on this condition, no sensitive itemsets is disclosed from the sanitized database. Given the minimum utility threshold and an original dataset, the high utility itemsets are generated using the two-phase algorithm proposed in (Ying Liu, Wei-keng Liao, and Alok Choudhary. 2005). A certain number of sensitive itemsets were randomly selected from a set of high utility itemsets.

A standard method of frequent itemset hiding approach reduces the support of the itemsets either by deleting the items or the transactions that support it (Evfimievski A, Srikant R, Agrawal R, and Gehrke J. 2002). In order to compare our results we decrease the utility of the itemsets in High Utility Itemset Sanitization (HUIS) approach by choosing the victim item for modification such that it yields maximum profit in that transaction. However the proposed approach (CUIS) selects transactions based on the number of sensitive itemsets that are supported by it and shows promising results in terms of misses cost and the difference between the original and sanitized database.

Fig. 1-6 shows the summary of the sanitizing algorithm based on varying size of the database with the performance measurements. We fixed the minimum utility threshold as 0.1% and $|Ps|=60$. Fig.1-6 shows that CUIS outperforms HUIS and attains lowest miss cost and difference in almost all the datasets.

7. Conclusion

Privacy becomes an important factor in data mining so that sensitive information is not revealed after mining. However data quality is important such that no false information is released provided privacy is not jeopardized. In this paper we proposed a novel approach for preserving privacy in sensitive high utility itemset mining. The various experimental results show that the approach applies minimum number of changes to the database and minimal amount of non-sensitive itemsets are missed which is the ultimate aim of data sanitization. Future work has to be carried over to develop optimal algorithms for data sanitization.

References

- Agrawal D and Aggarwal C.C (2001). "On the Design and Quantification of Privacy Preserving Data Mining Algorithms", In Proc. of ACM SIGMOD/PODS, pages 247–255, Santa Barbara, CA.
- Agrawal R and Srikant (2000). R, "Privacy Preserving Data Mining", Proc. ACM SIGMOD Conf. on Management of Data, pp. 439-450.
- Alexandre Evfimievski, Ramakrishnan Srikant, Rakesh Agrawal, Johannes Gehrke. (2002). "Privacy preserving mining of association rules", Proceedings of SIGKDD.
- Atallah M., Bertino E., Elmagarmid A., Ibrahim M., and Verykios V.. (1999). *Disclosure Limitation of Sensitive Rules*. In Proc. of IEEE Knowledge and Data Engineering Workshop, pages 45–52, Chicago, Illinois.
- Brankovic L and Estivill-Castro V. (1999). *Privacy Issues in Knowledge Discovery and Data Mining*. In Proc. of Australian Institute of Computer Ethics Conference (AICEC99), Melbourne, Victoria, Australia,.
- Charu C. Aggarwal, Jan Pei, Bo Zhang, (2006). "On Privacy Preservation against Adversarial Data Mining", Proceedings of Knowledge Discovery in Databases
- Culnan M.J, (1993), "How Did They Get My Name?: An Exploratory Investigation of Consumer Attitudes Toward Secondary Information", MIS Quarterly, 17(3): 341–363.

- Elisa Bertino, Ravi Sandhu, (2005), “*Database Security—Concepts, Approaches, and Challenges*”, IEEE transactions on dependable and secure computing, vol.2, no. 1.
- Estivill-Castro Vand Brankovic L, (1999). “*Data Swapping: Balancing Privacy Against Precision in Mining for Logic Rules*”, In Proc. of Data Warehousing and Knowledge Discovery (DaWaK-99), pages 389–398, Florence, Italy.
- Evfimievski A, Srikant R, Agrawal R, and Gehrke J. (2002). “*Privacy Preserving Mining of Association Rules*”, In Proc. of the 8th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, pages 217–228, Edmonton, AB, Canada.
- Han J and Kamber M. (2006). “*Data Mining: Concepts and Techniques*”, Morgan Kaufmann Publishers, San Francisco, CA.
- Hong Yao, Howard J. Hamilton, and Cory J. Butz. (2004). “*A Foundational Approach to Mining Itemset Utilities from Databases*”, In Proceedings of the 4th SIAM International Conference on Data Mining.
- Jefferies P. (2000). “*Multimedia, Cyberspace & Ethics*”, In Proc. of International Conference on Information Visualization, pages 99–104, London, England.
- Muralidhar K., Parsa R. and Sarathy R. (1999). “*A General Additive data Perturbation method for Database Security*”, Management Science, 45(10):1399-1415.
- Oliveira S. R. M. and Zaiane O. R. (2002). “*Privacy Preserving Frequent Itemset Mining*.” In Proc. of the IEEE ICDM Workshop on Privacy, Security, and Data Mining, pages 43–54, Japan.
- Rakesh Agrawal Ramakrishnan Srikant. (1994). “*Fast algorithms for mining association rules*”, Proceedings of the 20th VLDB Conference.
- Rizvi S. J. and Haritsa J. R. (1994). “*Privacy-Preserving Association Rule Mining*”. In Proc. of the 28th International Conference on Very Large Data Bases, x(2002), Hong Kong, China.
- Vassilios S. Verykios, Ahmed K. Elmagarmid, Elisa Bertino, Yucel Saygin and Elena Dasseni. (April 2004). “*Association rule hiding*”, IEEE Transactions on Knowledge and Data Engineering, Vol. 16, No. 4.
- Verykios V.S., Bertino E., Fovino I.N., Provenza L.P., Saygin Y., and Theodoridis Y. (2004). “*State-of-the-Art in Privacy Preserving Data Mining*” ACM SIGMOD Record, vol. 3, no. 1, pp. 50-57.
- Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen. (January 2007). “*Hiding sensitive association rules with limited side effects*”, IEEE Transactions on Knowledge and Data Engineering, Vol.19, No.1.
- ZhihuiWang, Wei Wang, Baile Shi, S. H. Boey. (2006). “*Preserving Private Knowledge in Frequent Pattern Mining*”, Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06).
- Ying Liu, Wei-keng Liao, and Alok Choudhary. (2005). “*A Two-Phase Algorithm for Fast Discovery of High Utility Itemsets*”, PAKDD 2005, LNAI 3518, pp. 689 – 695.
- URL <http://www.almaden.ibm.com/software/quest/Resources/index.shtml>, 2003.

Table 1. Summary of the datasets used in the experiment.

Database name	D	I	L	T
T10I10L5D1K	1K	10	5	10
T10I10L5D2K	2K	10	5	10
T20I10L5D4K	4K	10	5	20

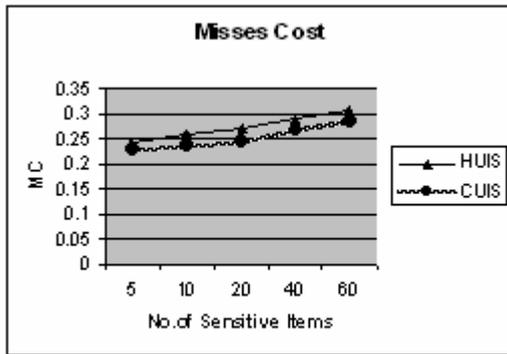


Figure 1. Misses cost for $|Ps|=60$ and T10I10L5D1K

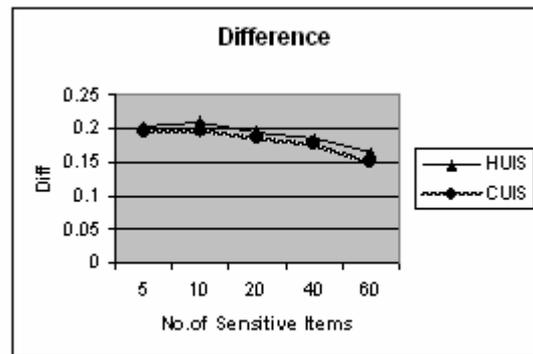


Figure 2. Difference for $|Ps|=60$ and T10I10L5D1K

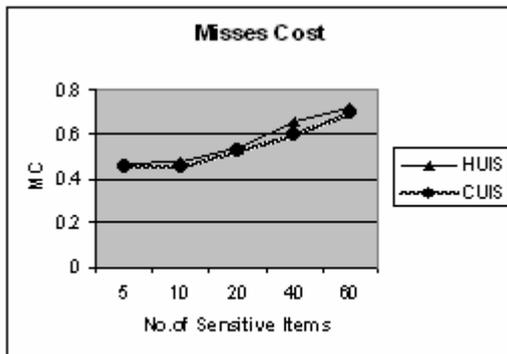


Figure 3. Misses cost for $|Ps|=60$ and T10I10L5D2K

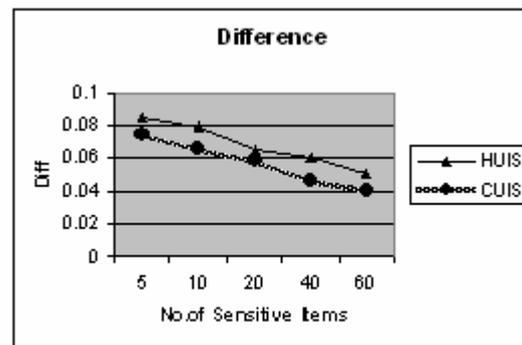


Figure 4. Difference for $|Ps|=60$ and T10I10L5D2K

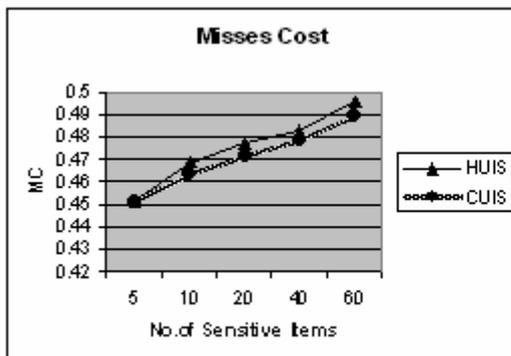


Figure 5. Misses cost for $|Ps|=60$ and T20I10L5D4K

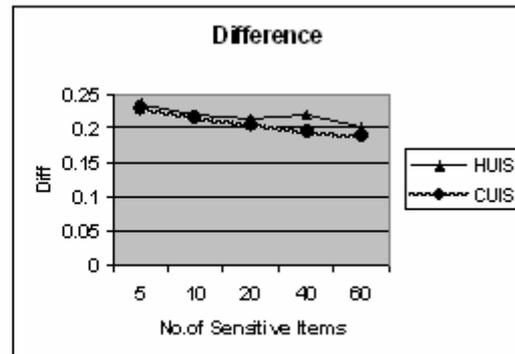


Figure 6. Difference for $|Ps|=60$ and T20I10L5D4K



The Drive and Support of Web3.0 to Rich Media Advertisement

Xiaonan Jin

Tianjin Polytechnic University

Tianjin 300387, China

Tel: 86-22-8395-1147 E-mail: perfectfeeling@163.com

1. Introduction

When the new thing, internet, formally entered into the commercial operation in 1990s, few people could think that the internet, which came from military technology, would drastically change the modes of obtaining information and communication by unheard-of speed. Undoubtedly, the developmental course less than 50 years from the advent of internet to the establishment of information society have made media theoreticians believe that internet was the most powerful media form beyond the memory of men, and it changed the way of traditional marketing, the medium and form of communication.

There is medium, there is advertisement. The development of network advertisement form is continually evolves with the development of application mode of internet, and every sort of application mode will go with a sort of network advertisement form. Through the comparison and analysis of the development course of network advertisement, this article will analyze the development foreground of Web3.0 and its drive and support to the rich media advertisements.

2. Web1.0 time: the network extension of traditional public communication mode

The earliest application mode of internet was the mode of Web1.0, which belonged to some static applications, and the typical application is to obtain the page payout of HTML, or implement simple data exchanges with the server, such as user log on, inquiring database and putting in data. Various web portals take the obtainment, issuance and disposal as the core, and try to collect various sort of news to make users browse, and the information communication mode is from internet to users, and few users' feedbacks can possess the statistical and analytical meanings, and this mode undoubtedly is the network extension of plan medium communication mode (seen in Figure 1). At the same time, in this term, the application of internet is extremely limited by the bandwidth, so it is called narrowband application.

In the term of narrowband application, the main bodies which compose the contents of website are the text and few GIF, JPG, PNG pictures with low quality and simple Flashes. The internet advertisement is thought to have great future, but under the limitation of bandwidth, it is still in the time of narrowband and is limited extremely for the form. For example, some page advertisement could be only composed by some simple word information or pictures about 30~40k, and comparing with TV advertisement and plane advertisement with beauties and art effects, this sort of narrowband advertisement is whiter and meager in the expressive force.

3. Web2.0 time: the time of "You"

Because of large-sized popularization of the broadband technology, the problem of bandwidth which had limited the development of internet has not been the bottleneck. When the information gap in the world range is gradually filled by the widespread broadband technology, more and more enterprises and users begin to go into the tide of internet, which makes information intersperse among every place, and the search engine and individual portal produce charms and the tide of Web2.0 emerges as the times require.

Under the support of broadband technology, Web2.0 makes internet possess richer contents such as search, picture, music, video, mash-ups, Wiki, blog and community. The internet in the time of Web2.0 takes every network people as the information node and every node is taken as one information source, even various web portals implement collection, disposal, filtration and republication, so search engine, blog, video, P2P and chat tool would assume main roles, and the information transfer mode is changed into from internet to users and to internet, or from users to users, or from users to internet and other multilateral communication modes, and people can search for what they need on different nodes (seen in Figure 2).

The time of Web2.0 really embodies the spirit of network, which is different to past inventions such as print and TV, and the network is not a sort of superincumbent spread, and it is a sort of spread form with self-organizational network, and

a sort of spread from bottom to top, which makes everyone has the public platform to speak, makes everyone possess the talking right in deed, and the grass-root estate, users begin becoming into the masters of the internet, so US Time magazine awarded the annual character of 2006 to “You”, which is the acute and appropriate grasp for this time.

The advertisements of Web2.0 are the advertisements of broadband. When the bandwidth bottleneck disappeared, the network advertisements which can realize 2D and 3D cartoons, video, audio with rich visual effects and communication function effects began occurring in the time of Web2.0 and get rapid developments, and these advertisements are just rich media advertisements with most potentials.

The rich media advertisement is not a single technology or form, and it is not a sort of advertisement platform or tool, and it is the network advertisement which takes the network as the carrier, combines with complex visual effects and communication function effects such as 2D and 3D videos, audio, JAVA, Flash and dynamic scripts, and its core rests with alternation and accuracy, and it can offer the advertisement experience most according with objective customers' demands. The rich media advertisement fully enriches the contents and effects of banner advertisement, pop-up advertisement, interstitial advertisement and other former narrowband network advertisements, and because it offers bran-new and superior use experiences for the network customers, so since it was born, it is popular and obtains quick development.

In Feb 2007, according to the data analysis of iResearch market consultation, the proportion that US online practitioners thought that the rich media advertisement was the best network advertisement strategy ascended from 28% in 2005 to 38% in 2006. And according to the data from eMarketer, in 2005, the increase rate of US rich media advertisement was 27.6%. eMarketer predicted that US rich media advertisement would increase quickly from 2004 to 2010, and the increase rate in 2006 would achieve 46.1%, and after that the increase speed would slow down, but the increase speed would still keep above 25%.

However, the mode of Web2.0 had not solved the problems of accuracy, efficiency and information integration, and it has not offered services according to every customer's individual behavior and demand, and this sort of network mode without enough subdivision and accuracy can not fully exert the advantages of communication and trace of customers' behaviors for rich media advertisement, and can not fulfill the higher requirement of issuance accuracy, so the time of Web2.0 is not the time that rich media shows its skills completely.

4. Web3.0 time: the time that the forthcoming rich media advertisement will be extraordinary splendor

Other than Web2.0, the time of Web3.0 is not only to offer comprehensive service and establish comprehensive service platform according to users' demands, and the most important is that it offers the individual integration service based on users' favors. The individualization and intelligence of search engine make users only list what they want, and without analyzing and testing how to combine keywords, the search engine will offer proper data and rapidly search information and solve problems. And according to one user's habit, the internet can automatically integrate his information, which is similar with individual portal, and everyone's browser pages are combined according to his noticed information favor and behavior habits, even the internet can offer not only information service, but individual consultation service, for example, based on human internet behavior traces, professional website will be turned into a artificial intelligent system which can give out reasonable and complete answers aiming at simple questions.

The most values of Web3.0 time are that the most important thing is not to offer information, but offers intelligent filter and comprehensive platform which can fulfill demands based on different demands, and every sort of filter is based on one user's demand, and the comprehensive platform is to fulfill users' comprehensive demands. The internet service is the aggregation which is set up according to everyone's noticed information type, individual demand and favor, and internet doesn't belong to one group of people any longer, but belongs to “You” (seen in Figure 3).

The new characters of Web3.0 bring large transformation for the internet marketing, and these characters can just fully fulfill the demands of rich media advertisement and exert its advantages, and both sides have many close combining points.

4.1 The character of Web3.0 accords with the demand of rich media advertisement in the aspect of marketing

As from the situation of present development, the click cost of rich media advertisement keeps the trend of stable increase, and the increase speed is obvious, and the average click cost in 2005 achieved 2.21 Yuan, but the average click cost of other network advertisement forms at the same term was only 1.56 Yuan. In order to enhance the utilization efficiency of capital and attract more advertisement businesses, the rich media advertisement has urgent requests for the accuracy of advertisement issuance.

On the contrary, for the mode of Web3.0, if the time of Web1.0 belongs to the time of public marketing, and the time of Web2.0 belongs to the time of focus marketing, so the time of Web3.0 is the time of accurate marketing. In the time of Web3.0, the information is highly integrated and users are also deeply fractionized and integrated, so internet offers higher accuracy for enterprises, and enterprises can clearly know what kind of users the information transfer to and they

can also exactly lock objective customers.

4.2 The accuracy of Web3.0 offers powerful supports for the exact investment of rich media advertisement

To realize the exact issuance of rich media advertisement, it needs fulfilling following three conditions which can be completely fulfilled by born advantages of Web3.0 technology.

(1) Enough samples. If the quantity of sample is too small, it is hard to find enough exact groups when the samples are filtered. The internet in the time of Web3.0 is every network user's internet, large and quickly increasing user quantity offers enough sample quantity for orientating objective customer group.

(2) A good semantic analysis technology. When users see the web page, the server can analyze the contents that the web page displays to users, i.e. similar with the technology of search engine, the server should know what the user notices at present. And this sort of technology is just the technical base of Web3.0 mode, and the core of Web3.0 marketing is to intelligently analyze users' behaviors, exactly know every user's network behavior track, and develop marketing through concluding different segment groups' behavior tracks.

(3) Strong database. The database is the database of user behavior. If there is not a large database as the support to analyze network people's behaviors, the exact issuance can not be performed. In the time of Web3.0 which emphasizes individualization, every website should have its own user database, and every website should grasp users' basic characters and nuclear values, and the use database of the website will limit its own group, and the website should periodically dig these databases to predict future trend, which will be the powerful base for the accurate issuance of rich media advertisement.

4.3 Web3.0 possesses the interactively integrated character with rich media advertisement in the aspect of communication

In the times of Web1.0 and Web2.0, the internet is in the time that people search for information, and though communication occurs in the time of Web2.0, but the internet users are disperse in various corners, the deficiency of subdivision and integration of users induces the users' values can not be exerted, and the values of internet advertisement can not be evaluated.

In the time of Web3.0, the internet integrates human demands and information, and the interaction between human and human or between human and internet establishes a new and interaction social space people depend on, and the rich media advertisement can realize interactive game, interactive research, interactive display and various interactive advertisement demands, and this sort of powerful interaction can perfectly infiltrate and integrate various applications of internet and become into the necessary part, and this sort of objective lock technology can ensure the exact arrival of advertisement to really interested customers, and repeat playing for these customers by its powerful display function, and accordingly influence the formation of users' brand images and exert the maximum benefits of advertisement.

5. Conclusions

The medium and advertisement are an indivisible integration supplemented each other, and both mutual dependences and promotion can maximally exert their own functions. From above analysis, we can see that the close combination of rich media advertisement and Web3.0 mode in the exact marketing and interaction can largely enhance the spread effect of advertisement, more effectively utilize limited advertisement space resource, reduce customers' dissatisfaction to the advertisement, and finally make the advertisement investment of enterprise exceed its values. Therefore, the fittest mode exerting function for rich media advertisement is the Web3.0 application mode of internet, and the time of rich media advertisement is the Web3.0 time of advertisement.

References

- Liu, Jia. (2007). Study on the Actuality and Development of Rich Media Advertisement. [Online] <http://media.people.com.cn/GB/40628/5726827.html>. (May 14, 2007).
- Xiao, Mingchao. (2007). Internet Web3.0: New Integration Time. [Online] Available: *Network Magazine of Market Research*. No.38.
- Yang, Weiqing. (2006). iResearch: Annual Report of Rich Media Advertisement in 2006. [Online] Available: <http://www.iresearch.com.cn>. (2006).
- Yang, Weiqing. (2006). iResearch: Data Report of Global Advertisement Market in 2006. [Online] Available: <http://www.iresearch.com.cn>. (2006).
- Yang, Weiqing. (2006). iResearch: Report of China Advertisement Market Share in 2006. [Online] Available: <http://www.iresearch.com.cn>. (2006).
- Yang, Weiqing. (2006). iResearch: Research Report of China Advertisement Market in 2006. [Online] Available: <http://www.iresearch.com.cn>. (2006).

Zhaojuan. (2007). Can Rich Media Advertisement Exceed Traditional Advertisement? [Online] <http://tech.sina.com.cn/i/2007-02-14/15351385130.shtml>. (Feb 14, 2007).

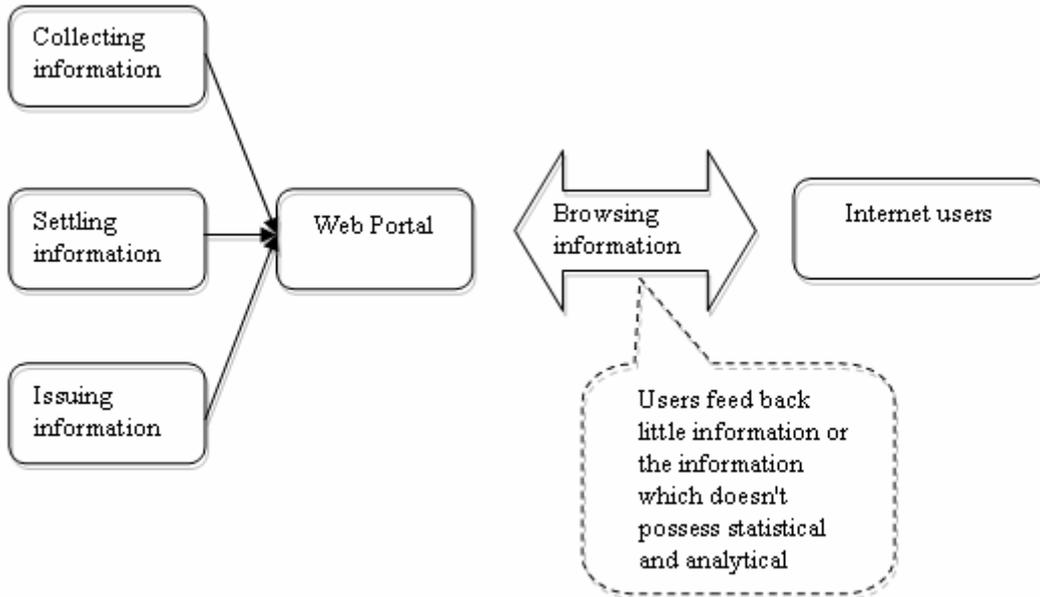


Figure 1. The Mode of Web1.0

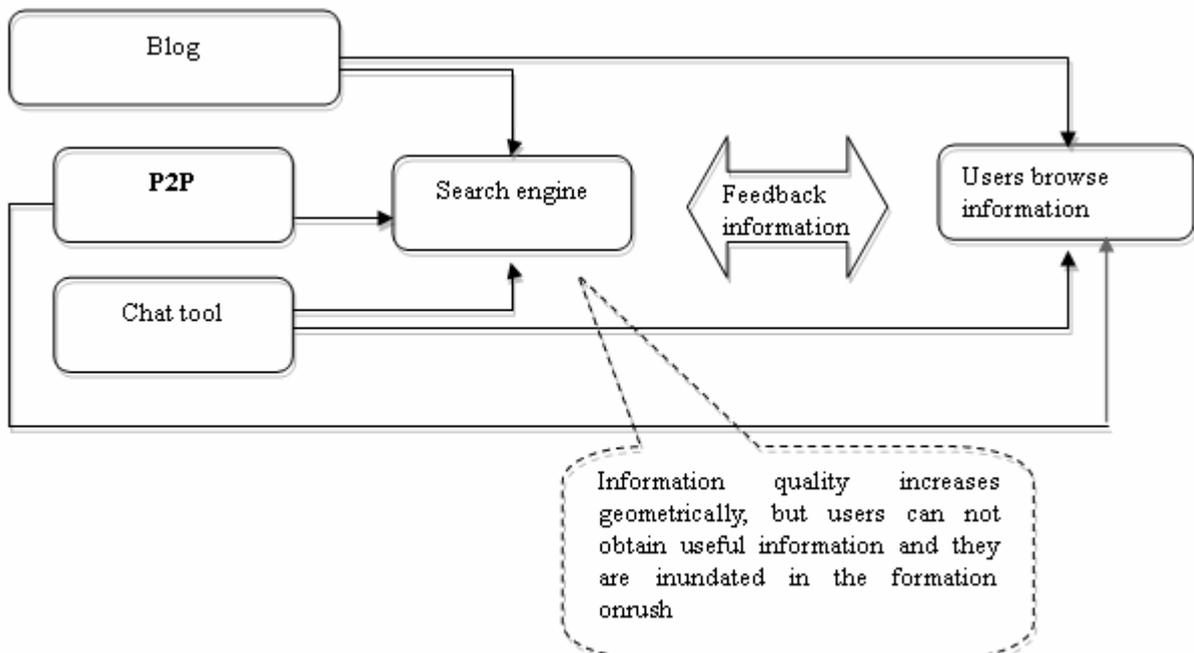


Figure 2. The Mode of Web2.0

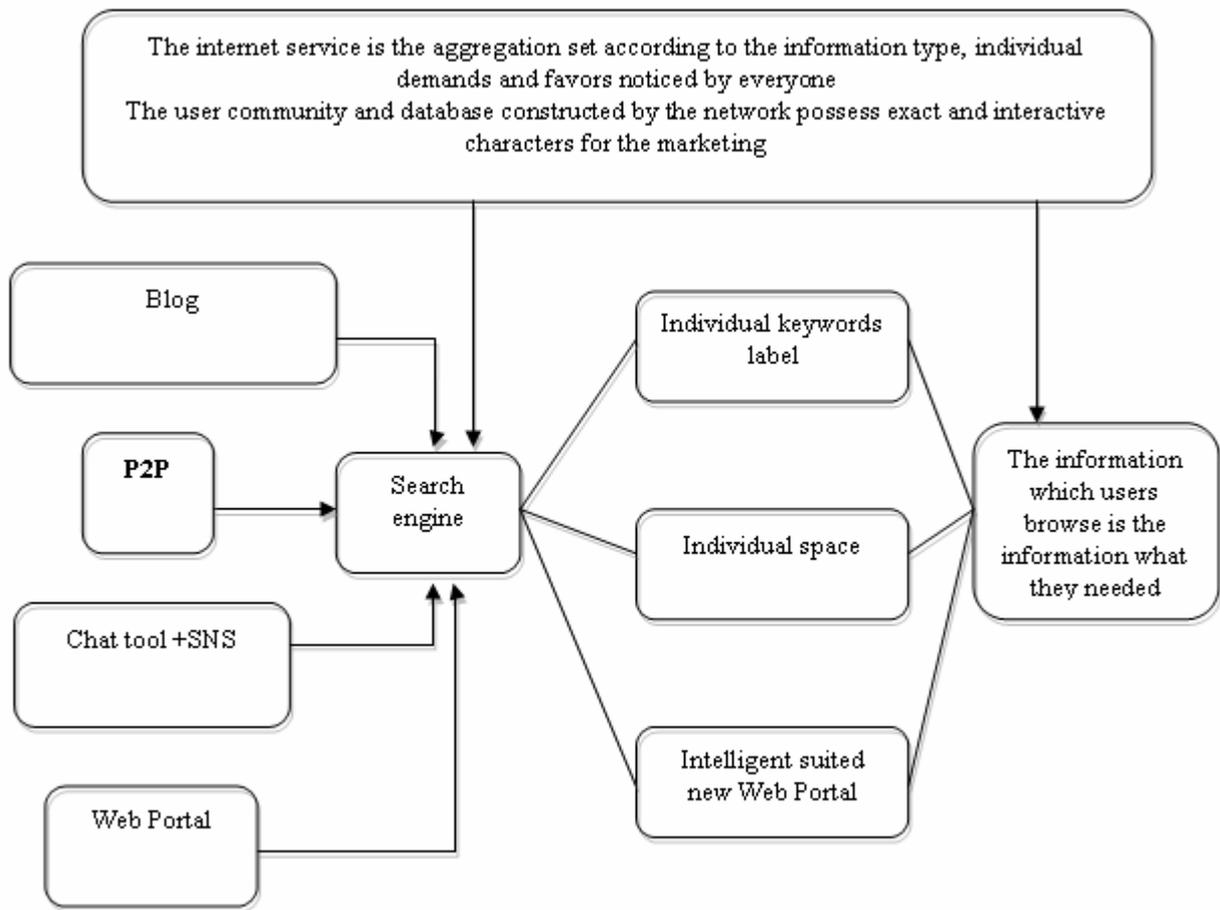


Figure 3. The Mode of Web3.0



Power Estimation in Mobile Communication Systems

Lenin Gopal, Ashutosh Kumar Singh & Veeramani Shanmugam

School of Engineering and Science

Curtin University of Technology

CDT 250, Miri, Sarawak, Malaysia

Tel: 60-85-443-932 E-mail: lenin@curtin.edu.my

Abstract

Accurate power estimation has an important role for power control and handoff decisions in mobile communications. Window based weighed sample average power estimators are commonly used due to their simplicity. In practice, the performances of these estimators degrade severely when the estimators are used in the presence of correlated samples. In this paper performances of the three local mean power estimators namely, sample average, optimum unbiased and maximum likelihood estimators, are analysed in the presence of correlated samples. The variance of the estimators is used as performance measures. Finally, the simulation results show that the performances of the optimum unbiased and maximum likelihood estimators are very good as compared to the performance of the sample average estimator.

Keywords: Power estimation, Local mean power, Rayleigh fading, Mobile communications

1. Introduction

In mobile communication systems, local mean power estimation is very important and it must be an accurate estimate of the received signal power. Some system function like channel access, handoff, and power control use local mean signal level to indicate the wireless communication link quality.

Mockford, Turkmani and Parsons (1990) analysed local mean signal variability in rural areas. The local mean of the signal envelope was described by a lognormal distribution. The power spectral density of the local mean signal was estimated along rural routes. Goldsmith, Greenstein and Foschini (1994) studied the error statistics of real time power measurements in cellular channels with multipath and shadowing. The authors considered two measurement methods (filtering the squared envelop, and filtering the logarithm of the squared envelop) and two filter types (integrate-and-dump and RC). Accurate measurements were obtained by filtering either the logarithm of the detected power or the power itself. Valenzuela, Landon and Jacobs (1997) explored techniques for the measurement of local mean signal strength at 900 MHz and 2 GHz. Linear averaging technique was used to estimate the local mean and ray tracing propagation model was used to evaluate different methods of calculating the local mean signal strength for indoor environment. Manohar Das and Coopriider (1997) introduced new techniques for detection of changes in the local mean of a signal. The first one utilized a discrete cosine transform (DCT) based data compression principle, and the second one was based on a robust piecewise linear approximation (PLA) of the given signal. The results of experimental studies were compared with the performance of the two methods with the existing filtered derivative method. Wong and Cox (1999) derived the optimal local mean signal level estimator for the Rayleigh fading environment and compared with the sample average estimator. Variance of the two estimators were estimated and compared with the Cramer-Rao Lower bound. The 5th and 95th percentiles of the estimators were obtained by computer simulation. Antilog Rayleigh (ALR) distribution was used to estimate the signal variation in a Rayleigh fading environment.

Young-Chai Ko and Mohamed-Slim Alouini (2001) presented two local mean power estimation techniques over Nakagami fading channels. Chai Ko and Mohamed-Slim Alouini presented maximum likelihood as well as minimum variance unbiased estimators for the local mean signal power estimation. De Jong and Herben (2001) presented new method for the computation of local mean power from individual multipath signals predicted by two-dimensional ray tracing based on an expression for the spatial average of the received power, which takes into account the spatial correlation between signals. The presented method is based on the spatial average of the received power over each pixel area. Avidor and Mukerjee (2001) investigated the possibility of obtaining better estimates or prediction of the path loss between a mobile and the surrounding base station by processing more measurements, including older measurements of received power that are not used by current deployed algorithm. The current or near future value of the local mean

received power including the shadow loss was estimated by using the algorithm. Tepedelenlioglu, Sidiropoulos and Giannakis (2001) derived the maximum likelihood and median filtering for power estimation in mobile communication system. Linear filtering techniques for power estimation was compared with the maximum likelihood (ML) estimator and the median filtering techniques for power estimation was compared with the linear filtering, maximum likelihood (ML) estimator and uniformly minimum variance unbiased (UMUV) estimator. Jiang, Sidiropoulos and Giannakis (2003) proposed a scalar Kalman-Filter-based approach for improved local mean power estimation. The performance of the Kalman filter (KF) was compared with the window based estimators, like the sample average estimator of (Goldsmith, Greenstein & Foschini, 1994) the uniformly minimum variance unbiased (UMVU) estimator of (Wong & Cox, 1999) and the maximum likelihood (ML) estimator of (Tepedelenlioglu, Sidiropoulos and Giannakis, 2001).

In this paper, the performances of the sample average estimator, the optimum unbiased estimator and maximum likelihood estimator are studied with the use of uncorrelated and correlated samples (Rayleigh fading signal). For correlated samples, the variances of the three estimators are obtained while varying the correlation between the samples.

The first part of this paper introduces the system model for the local mean power estimation using uncorrelated and correlated samples. Following that, the three estimators details are included. The performances of the three local mean power estimators and conclusion are included in the subsequent sections.

2. System Model

2.1 Uncorrelated Fading Samples

The model for the uncorrelated fading samples at the log amplifier output of the receiver is shown in Fig .1. The uncorrelated samples can be expressed as

$$Y = r_i^2 + r_q^2 \quad (1)$$

The output of the amplifier is expressed as

$$X = 20 \log Y \quad (2)$$

X has a probability density function called the Antilog Rayleigh (ALR). The ALR probability density function is given by

$$f_x(X) = (\ln 10) \frac{10^{\frac{x}{10}}}{20p} \exp \left(-\frac{10^{\frac{x}{10}}}{2p} \right) \quad (3)$$

where $p = E(Y^2)/2$ is the power of each phase component.

2.2 Correlated Fading Samples

Fig.2 shows the simulator model for correlated samples (Rayleigh fading signal). A block of $N_f = 1024$ uncorrelated samples are generated and the maximum Doppler frequency shift $f_m = 50$ Hz is specified to produce proper Doppler shift. The (Smith, 1975) steps are used to generate series of simulated Rayleigh fading signal with proper Doppler spread and correlation between the samples. The correlated output samples are fed into the logarithmic amplifier. The output of the amplifier can be expressed as in eq.2.

3. Local Mean Power Estimators

In this section, the conventional local mean power estimators for the Rayleigh fading channel are briefly reviewed.

3.1 Sample Average Estimator

Due to the rapid changing characteristics of the Rayleigh fading process, an estimate of local mean power is obtained by averaging the samples X . The average of the received samples as follows,

$$E_{SA} = \frac{1}{N} \sum_{j=1}^N X_j \quad (4)$$

where N is the window size, X_j is the received power measurement.

3.2 Optimum Unbiased Estimator

The optimum unbiased estimator was derived in (Wong, 1999) as follows.

$$E_{OU} = 10 \left[\log T - \frac{H_{N-1}}{\ln 10} \right] \quad (5)$$

$$\text{Where } T = \sum_{j=1}^N 10^{X_j/10} \quad (6)$$

and
$$H_N = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{N} \tag{7}$$

for $N= 1,2,3, \dots$ H_N is known as the N^{th} harmonic number.

3.3. Maximum Likelihood Estimator

In (Tepedelenlioglu, 2001), the maximum likelihood estimator for local mean power estimation was derived as follows.

$$E_{ML} = 10[\log(T) - \log(N)] \tag{8}$$

where
$$T = \sum_{j=1}^N 10^{X_j / 10} \tag{9}$$

4. Simulation Results and Comparisons

In this section, the performances of the three estimators, sample average estimator, optimum unbiased estimator and maximum likelihood estimator are compared. Simulations are carried out in uncorrelated and correlated samples environment. The channel parameters are obtained by the multipath fading generator as in Fig. 1 and 2.

4.1 Effect of Number of Samples, N

Fig.3 and 4 show the performance of the three estimators for uncorrelated and correlated samples. All the three estimators' variance gradually decreases for increasing number of samples. The variance obtained with optimum unbiased estimator and maximum likelihood estimators are consistently lower than that obtained with sample average estimator. From Fig. 3, the sample average estimator requires 50 samples to achieve 0.65 dB² variance, but optimum unbiased and maximum likelihood estimators need 30 samples to achieve the 0.65 dB² variance for uncorrelated samples.

Fig.4 shows the variance of the three estimators for correlated samples. From the Fig.4, the sample average estimator requires 200 samples to achieve 0.65 dB² variance but the optimum unbiased and maximum likelihood estimators require only 100 samples to achieve the same performance. Optimum unbiased and maximum likelihood estimators consistently perform better than the sample average estimator over the entire range of number of samples tested for correlated samples.

Fig.5 and 6 compare the performance of the 5th and 95th percentiles of the three estimator output for uncorrelated and correlated samples. In Fig.5, 50 samples are needed for the difference between the two percentile lines to be 2 dB for optimum unbiased and maximum likelihood estimators. About 85 samples are needed for sample average estimator to obtain the 2 dB differences between the two percentile lines. In Fig.6 approximately 150 samples are required to achieve the 2 dB differences between the two percentile lines for optimum unbiased and maximum likelihood estimators. About 400 samples are required to achieve the similar performance with the sample average estimator for correlated samples.

4.2 Effect of Doppler Frequency, f_m

Fig.7 and 8 show the relation between the variance of the estimators with normalised Doppler frequency. Doppler frequency of the Rayleigh fading signal is varied from 15 Hz to 300 Hz. $M=2$ sample distances is used in the simulation. Sample average estimator has 17.6 dB² variance at low normalised frequency (0.01). The performance of the sample average estimator is improved at high Doppler frequency. The variance of the optimum unbiased and maximum likelihood estimators is 13 dB² at low normalised frequency (0.01). The optimum unbiased and maximum likelihood estimators have better performance as compared to the sample average estimator for the tested range of Doppler frequency.

In Fig.8, the variance of the three estimators is compared over various normalized Doppler frequencies. Sample distance used in the simulation is $M=20$ and the Doppler frequency is varied from 15 Hz to 300 Hz. Variance of the sample average estimator is 3.2 dB² at 0.01 normalized Doppler frequency. Sample average estimator achieves the same performance at 0.08 normalized Doppler frequency for 2 sample distance. Optimum unbiased and maximum likelihood estimators require 20 sample distance to obtain 2.4 dB² variance for 0.01 normalized Doppler frequency. Both estimators achieve the similar performance for 2 sample distance at 0.09 normalized Doppler frequency. The optimum unbiased and maximum likelihood estimators have very low variance as compared to the sample average estimator for 20 sample distance. The variance of the estimators is very high at low Doppler frequency and decreases at high Doppler frequency.

4.3 Effect of Correlation Between Samples

Fig.9 shows the relation between the variance of the estimators with the autocorrelation between the samples. 2 sample distances is used in the simulation. Optimum unbiased and maximum likelihood estimators have very low variance as compared to sample average estimators. The variance of all the estimators is very high at 0.9 to 1 auto-covariance. The variances of the estimators reduce for decreasing value of the autocorrelation between the samples.

In Fig.10, the performance of the three estimators is compared with the autocorrelation between the samples. 20 sample distance is used in the simulation and the Doppler frequency is varied from 15 Hz to 300 Hz. The variance of the

estimators is reduced by six times at 0.9 to 1 autocorrelation of the samples as compared to 2 sample distance simulation. Optimum unbiased and maximum likelihood estimators have very low variance as compared to the sample average estimators for the tested range of the autocorrelation of the samples. Performance of the estimators improves for increasing value of the sample distance used in the simulation. The variances of the three estimators reduce for increasing value of the sample distance.

5. Conclusions

In this paper, the performance of the local mean power estimators is analysed for uncorrelated and correlated samples (Rayleigh fading channel). The comparisons are made between the sample average estimator with optimum unbiased and maximum likelihood estimators. Optimum unbiased and maximum likelihood estimator requires much less samples to achieve the desired result as compared to the sample average estimator using uncorrelated samples. All the three estimators require more test samples to achieve the desired result using correlated samples as compared to the performance of the estimators using uncorrelated samples. The performances of the three estimators improve for the increased number of tested samples.

Finally, for the correlated samples, the required sample distance is investigated to achieve a similar performance of the uncorrelated samples. Large numbers of samples distances are required for the correlated samples to achieve similar performance of the uncorrelated samples. It is also noted that the higher the normalized Doppler frequency, the better is the estimator performance. Similarly, the higher the correlation between two samples, the poorer is the estimator performance.

References

- Tao Jiang., Sidiropoulos, N. D., & Giannakis B. Giannakis. (2003). Kalman Filtering for Power Estimation in Mobile Communications. *IEEE Trans. On Wireless Communications*, Vol.2, No-1, pp.151-161.
- Avidor, D., & Mukherjee, S. (2001). Estimation and Prediction of the Local Mean Signal Power in Mobile System. *Veh. Technol. Conference, VTC 2001 Spring. IEEE VTS 53rd*, Vol.4, 6-9 pp.2751-2755.
- De Jong, Y. L. C., & Herben, M. H. A. J. (2001). Prediction of Local Mean Power using 2-D Ray-Tracing-Based Propagation Models. *IEE Trans.on. Vehicular Tech.*_Vol.50, No.1, pp.325-337.
- Tepedelenlioglu, C., Sidiropoulos, N.D., Giannakis B. Giannakis. (2001). Median Filtering for Power Estimation in Mobile Communication systems. *in Proc. 3rd IEEE Signal Processing workshop on Signal Processing Advances in Wireless Communications*, Taiwan, pp.229-231.
- Young Chai Ko., & Mohamed Slim Alouini. (2001). Estimation of the Local Mean Power over Nakagami Fading Channels. *12th IEEE International Symposium on personal, Indoor and Mobile Radio Communications*, Vol.1, pp c-107-c-112.
- Wong, D., & Cox, D. C. (1999). Estimating Local Mean Signal Power Level in a Rayleigh Fading Environment. *IEEE Trans. on Veh. Technol.*, Vol-48, No-3, pp.956-959.
- Manohar Das., & Coopriider, A. (1997). New Technologies for Detection of Changes in the Local Mean of a Signal. *in proc.of the 40th Midwest Symposium on circuit and systems*_ Vol.2, 3-6, pp.849-852.
- Valenzuela, R. A., Landon, O., & Jacobs, D. L. (1997). Estimating Local Mean Signal Strength of Indoor Multipath Propagation. *IEEE Trans. On Veh. Technol.*, Vol.46, No 1.
- Goldsmith, A. J., Greenstein, L. J., & Foschini, G. J. (1994). Error Statistics of Real-Time Power Measurements in Cellular Channels with Multipath and Shadowing. *IEEE Trans on Veh.Technol.*_ Vol.43, No.3, pp.439-446.
- Mockford, S., Turkmani, A. M. D., & Parsons, J. D. (1990). Local Mean Signal Variability in Rural Areas at 900 MHz. *in Proc. 40th vehicular Technology Conf.*_ pp. 610-615.
- Smith, J. I. (1975). A Computer Generated Multipath Fading Simulation for Mobile Radio. *IEEE Transactions on Vehicular Technology*, Vol. VT-24, No.3, pp. 39-40.

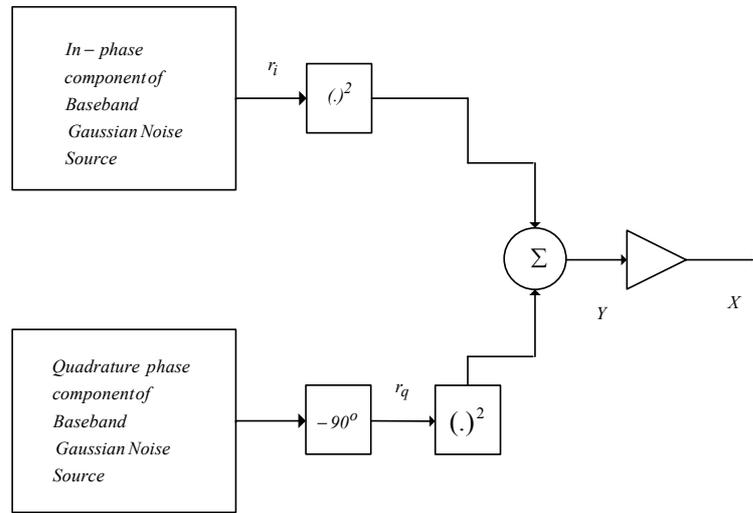


Figure 1. Simulator of uncorrelated fading samples.

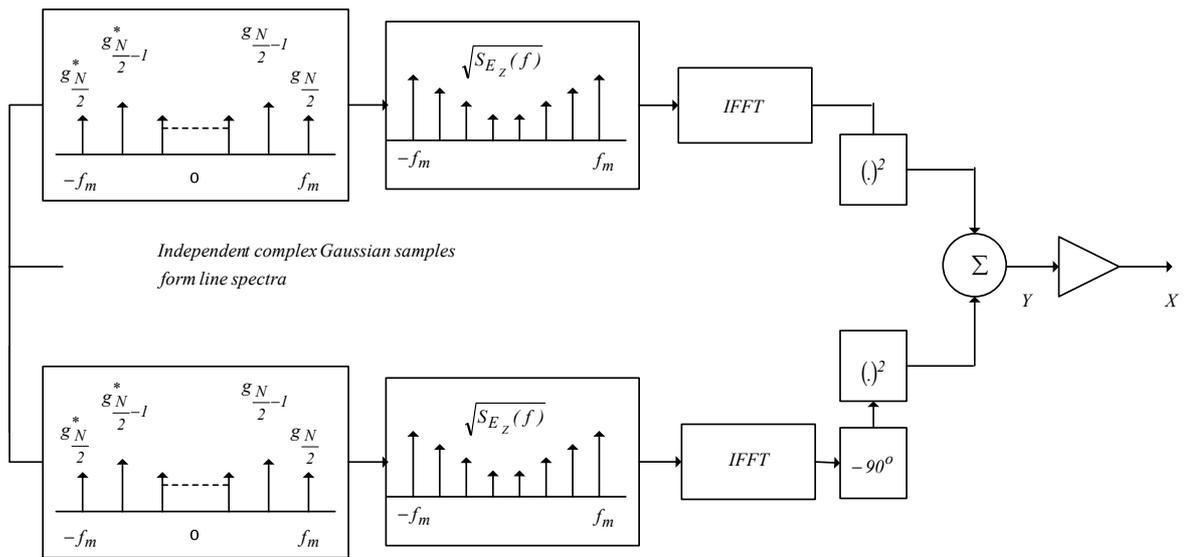


Figure 2. Simulator of correlated fading samples.

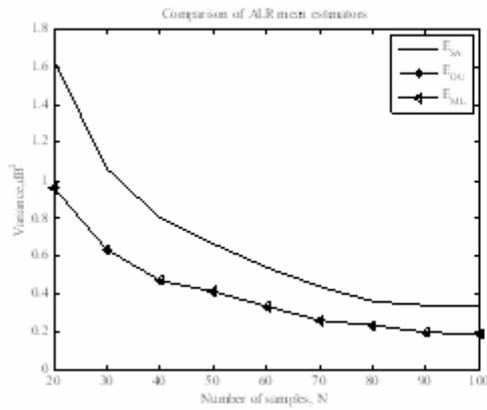


Fig. 3. Variance of the estimators for uncorrelated samples.

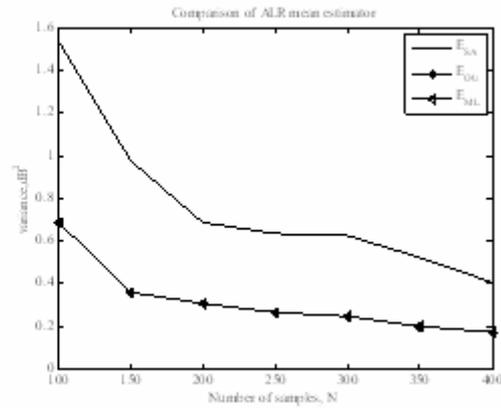


Fig. 4. Variance of the estimators for correlated samples.

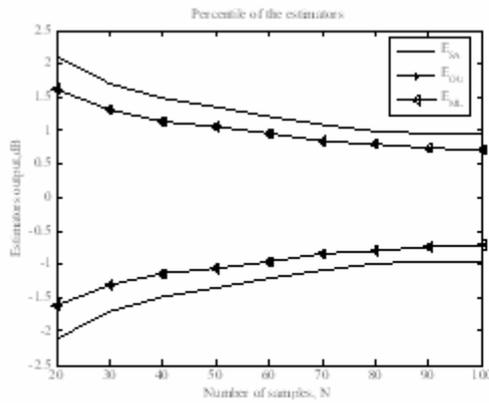


Fig. 5. The 5th and 95th percentiles of the estimators for uncorrelated samples.

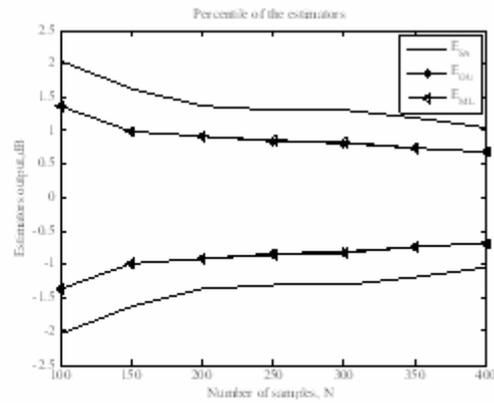


Fig. 6. The 5th and 95th percentiles of the estimators for correlated samples.

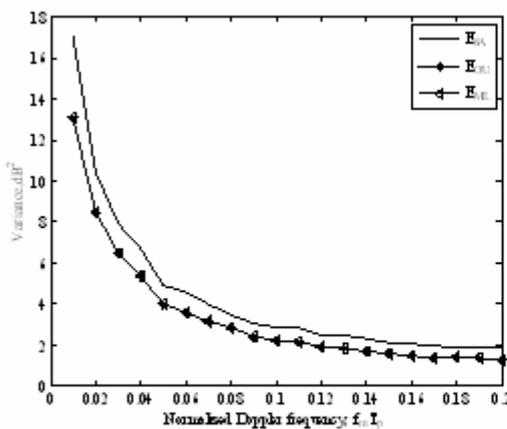


Fig. 7. Variance of the three estimators [$f_m=15$ Hz, 30 Hz... 300 Hz, Number of samples per second $N_f=1/T_p=1500$, $N=20$, $M=2$]

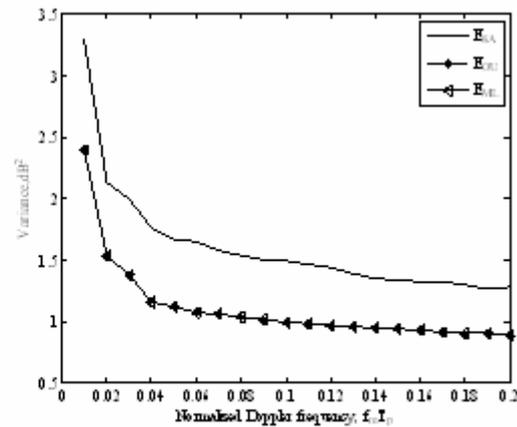


Fig. 8. Variance of the three estimators [$f_m=15$ Hz, 30 Hz, ..., 300 Hz, Number of samples per second $N_f=1/T_p=1500$, $N=20$, $M=20$]

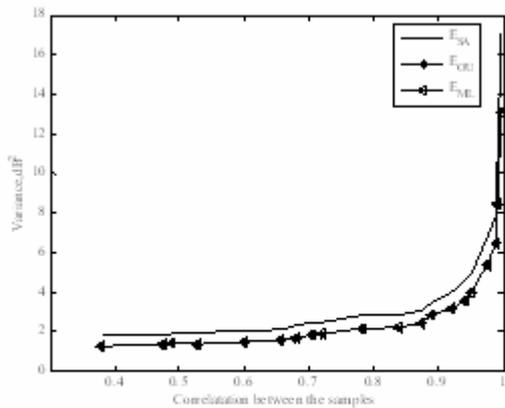


Fig. 9. Variance of the three estimators [$f_m=15$ Hz, 30 Hz... 300 Hz, Number of samples per second $N_f=1/T_f=1500$, $N=20$, $M=2$]

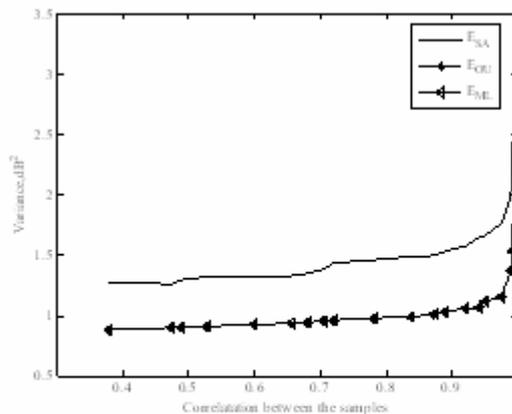


Fig. 10. Variance of the three estimators [$f_m=15$ Hz, 30 Hz... 300 Hz, Number of samples per second $N_f=1/T_f=1500$, $N=20$, $M=20$]



Design of Calendar Clock Based on DS12C887 Chip

Xiao Chen

Department of Electronic Information Engineering
Nanjing University of Information Science and Technology
Nanjing 210044, China
Tel: 86-25-5873-1196 E-mail: rainofsun@netease.com

The research is supported by the Basic Research Program (Natural Science Foundation) of Jiangsu Province, China (No. BK2007601) and the Natural Science Foundation of Jiangsu Higher Education Institutions of China (No. 06KJB510048). (Sponsoring information)

Abstract

The digitization of clock has brought us enormous conveniences and the development of the special-purpose clock chip would extend the traditional functions of clock. In this article, we design a new calendar clock based on DS12C887 chip and AT89C52 SCM. The system includes 5 parts: the AT89C52 SCM module, the calendar clock module, the LED display module, serial communication module and the C51 program. The SCM can get real-time information through the correct operations of addressing and the register. Users can set up initial time through serial communication between the PC and SCM. The LED unit is used to display real-time clock information.

Keywords: DS12C887, SCM, Clock, LED display

1. Introduction

The clock has had hundreds years' history, and the digitization of clock has brought us enormous conveniences and extended the traditional function of clock such as telling time. Many new functions such as automatic timing alarming, automatic bell on schedule, automatic control of time program, timing broadcasting, automatic street lamp, timing switch oven, automatic drive equipment break, and even automatic start of various timing electric appliances are based on the digitization of clock. Therefore, it is very significant to study the digital clock and extend its applications.

With the developments of computer and SLSI technology, the technology of professional digital clock chip is more and more mature, and the technology of SCM has gradually entered into the design of electric clock (Tian, 2005, p.61-63 & Xiehui, 2005, p.34-38). The function of digital clock has been extended significantly, for example, when the data is collecting, for some important information, not only the contents but also the concrete time that this event happens should be recorded, and for another example, the display screen of interest rate or exchange rate used in the bank hall should display information of calendar clock such as information of year, month, day, week, hour, minute and second, except displaying interest rate and exchange rate.

In this article, we design a real-time calendar clock display system based on DS12C887 chip, which can display second, minute, hour, week, day, month and year, and possesses the function of leap year rehab, and for the time display, it can select 24 hours mode or 12 hours mode with "AM" and "PM" indications, and the time, calendar and clock have two sorts of form such as binary code and BCD code, and it can work for a long time when the exterior power supply conks out.

2. Design ideas

The real-time display can be realized through software programming, but this method needs complex program, and the codes are numerous, and the spending of SCM software is high, and the time information is difficult to be stored for a long term. The special real-time clock chip can avoid these problems, and it can save time information chronically nonvolatile. Therefore, we can use special clock chip DS12C887 to implement the real-time calendar clock system. This chip can offer detailed information including year, month, day, week, hour and minute for the system. The interior of DS12C887 has information memorizers of clock, week and date, and the real-time time information is in these nonvolatile information registers. In the same way as 51SCM, DS12C887 also adopts 8 bits address/data reuse bus, which also has interior control register and the interior time information register reading DS12C887. Various registers of DS12C887 have corresponding fixed addresses in the interior space, so SCM can get time information needed through correct addressing and register operations.

According to the partition of function module, the hardware structure of system includes 5 following parts which are seen in Figure 1.

- (1) The module of 51 SCM. Its function is to communicate with exterior clock chip, control the data transfer process, collect and dispose time information.
- (2) The module of calendar clock. It is composed by special real-time clock chips. And it is the core module of the system, and it can offer real-time calendar clock information.
- (3) The module of LED display. It is used to display calendar clock. Its function is to display output data and state. So, as typical exterior parts, the LED display unit is the effective part to reflect system output and operation input.
- (4) The module of serial communication. The user can set initial time information through the series communication between PC and SCM.
- (5) The C51 program. It includes the interface program (it can realize the data transfer process between SCM and clock chip) and display program that SCM controls clock chip.

3. Design of hardware

The hardware of the system includes 4 parts. They are the module circuit of calendar clock chip DS12C887, module circuit of 51 SCM AT89C52, the module circuit of serial communication, and the module circuit of LED display output. Figure 2 is the structure of circuit principle of the system. The user can set initial time information through the serial communication between PC and SCM, and after the initial time information is set up, the system begins to work, and the calendar clock chip DS12C887 would offer real-time calendar clock information which will implement data transfer with SCM AT89C52. The SCM AT89C52 communicates with the calendar clock chip DS12C887, and controls the data transfer process, and collects and dispose time information. Finally, AT89C52 will transfer time information to the display system, and the real-time time information is displayed through 4 bits dynamic display circuit.

U1 is the SCM chip AT89C52 of Atmel Company, and it works in the clock of 11.0592MHz (Horenstein, 1996, p.24-57). P0 port is used as the address/data reuse bus AD [0, 1 ...7] which is connected with the AD [0, 1 ...7] of calendar clock chip DS12C887. The P1 port is used as the interface of LED display, and because the display in the article has not decimal, we only use a, b, c, d, e, f and g, and don't use the segment of dp (decimal). The P2.1/P2.4 of P2 port are used as the digit code interfaces of LED, and they need to control the digit code selection of 13 bits LED through 4-16 decoding circuit and drive circuit. The P2.0 of P2 port offers patch selection signals for the calendar clock chip DS12C887 after reverse phase, because the patch selection signals are effective on the low level. The series lead feet of SCM P3.0 (RxD) and P3.1 (TxD) will realize serial communication with PC through the serial level switch circuit. The SCM lead feet P3.7 (/RD) and P3.6 (/WR) are directly connected with the write and read pins of calendar clock chip DS12C887, and both them are effective on the low level. The ALE lead foot of SCM is directly connected with the memory latch pin of real-time clock chip, and it is used as the memory address latches which can realize the time reuse of data and address.

U2 is 6 bits phase inverter 74LS04, which implements reverse phase to P2.0 of SCM and transfer it to the patch selection pin of DS12C887.

U3 is the calendar clock chip DS12C887. In the design, its MOT lead foot needs to be grounding, and the Intel bus time sequence mode is selected. When U3 works in the mode of Intel bus time sequence, it completely allows the interface of 51 SCM, so its address/data reuse lines AD0-AD7, latch input ALE, read input DS and write input R/\overline{W} are completely connected with corresponding lead feet of the interface of 51 SCM. The square wave output SQW of DS12C887 and interrupt application /IQW are needed in the design, so its lead feet can be hang in the air.

The serial communication adopts the serial port of SCM to communicate by the mode of RS232. The hardware of data communication adopts 3 lines system which is to respectively connect three lead feet (RxD, TxD, GND) between SCM and PC, i.e. across connecting the TxD line and RxD line between PC and SCM, and directly connecting both GND lines, and using no other signal lines, and adopting the mode of software handclasp. In this way, it can not only realize scheduled rash, but simplify the circuit design. Because the TTL logic level of SCM has different electric character with RS232, so the MAX232 chip must be used to switch the level.

Because the system needs display year, month, day, weak, hour, minute and second, i.e. 13 bits effective information, so 4 LED are needed. The 4 bits 7 segments common-cathode LED is selected and the LED are connected with SCM in parallel and dynamically display the time information. Following aspects should be noticed, the confirmation of word digit port address and word type port address, how to realize the data display of LED through word bit code and word type code and realize dynamical display through timing interrupt.

The connections between SCM and various module pins are shown as follows.

TxD: The lead foot11 and the TTL level input lead foot of MAX 232 respectively connect with series input signals TXD and TTL of SCM.

RxD: The lead foot 12 and the TTL level output lead foot of MAX 232 respectively connect with series input signals

RxD and TTL of SCM.

SER-IN: The lead foot 14 and the RS-232 level output lead foot of MAX 232 respectively connect with series output signals RxD and RS232 of RS-232.

SER-OUT: The lead foot 13 and the RS-232 level input lead foot of MAX 232 respectively connect with series output signals RxD and RS232 of RS-232.

4. Design of software

The emphasis of software design is how SCM control calendar clock chip and read time information through the operation setup of register. The program flow of the real-time calendar clock display system designed in this article is shown in Figure 3.

5. Conclusions

The calendar clock chip DS12C887 can offer complete time information from year, month, day and weak to hour, minute and second, and it is broad applied in the SCM system which needs getting or displaying real-time calendar clock information. In this article, we design a sort of real-time calendar clock based on DS12C887 chip and AT89C52 SCM. The whole clock display system sets up the initial time information through the series communication of PC and SCM, and at the same time SCM exchanges data with the calendar clock chip DS12C887, and finally the real-time calendar clock information implement digital display through the digital display module.

References

Horenstein M N. (1996). *Microelectronic Circuits and Devices*. Prentice-Hall Inc, New Jersey. p.24-57.

Tian, Fuyang. (2005). The Application of the DS12C887 in the SCM System. *Electrical Measurement & Instrumentation*. No.42(3). p.61-63.

Xie, Hui. (2005). The Application of Real Time Clock Chip DS12C887 in Vehicle Data Recorder of Electrical Vehicle. *Machine Tool & Hydraulics*. No.12. p.34-38.

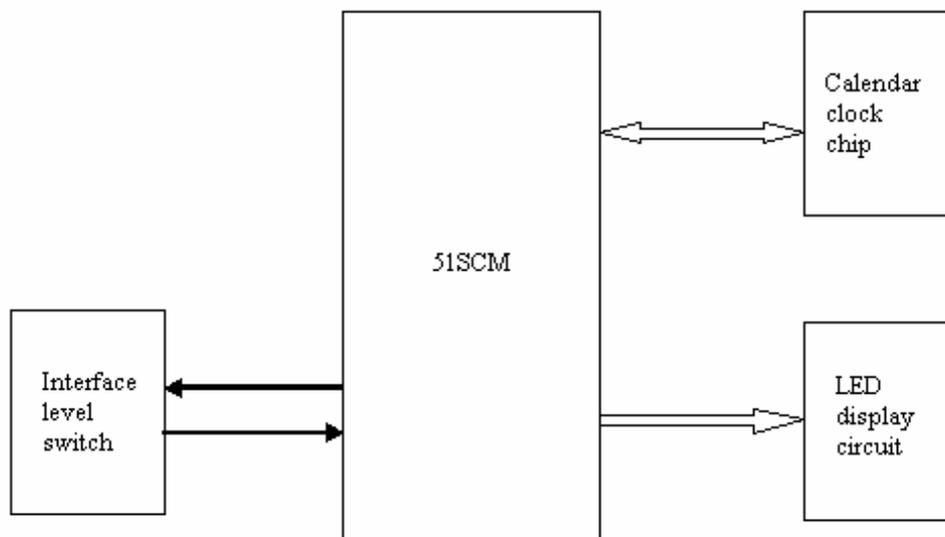


Figure 1. The Structure of System Hardware

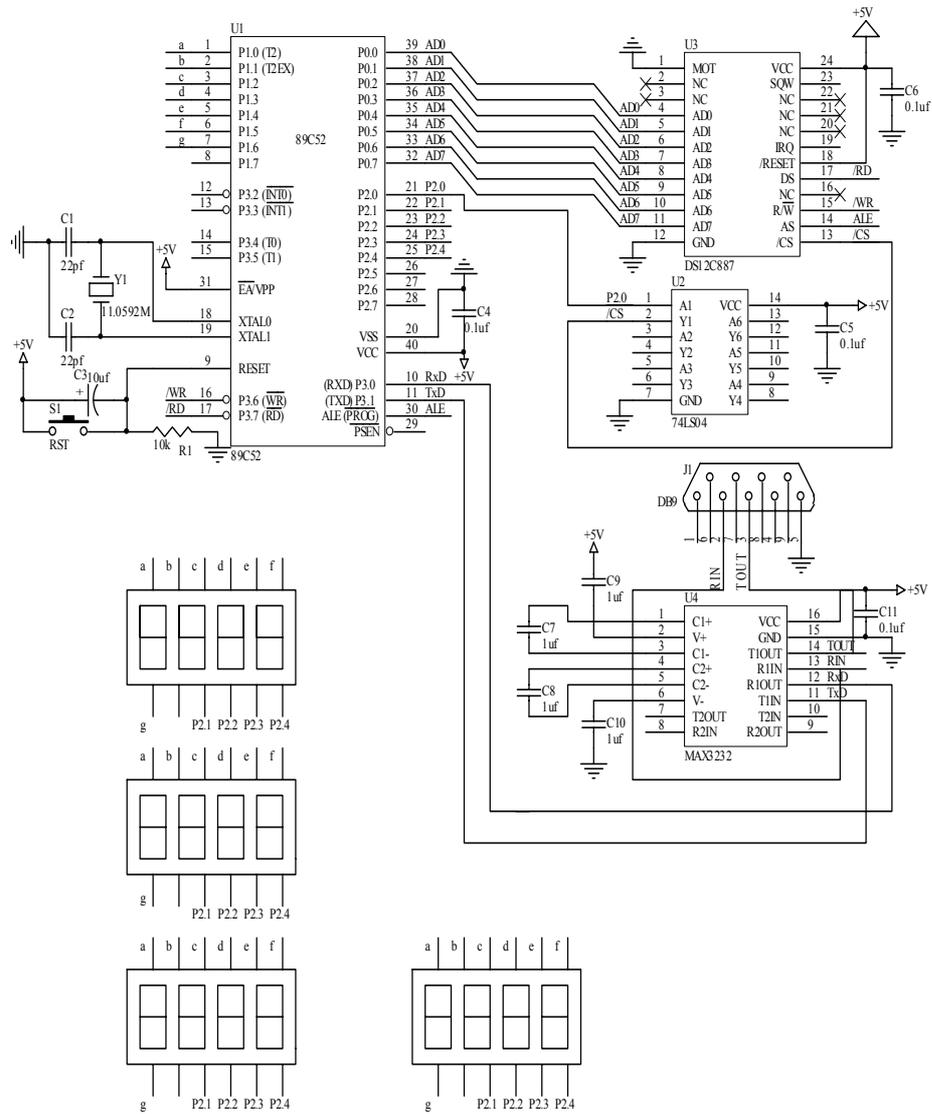


Figure 2. The Structure of Circuit Principle

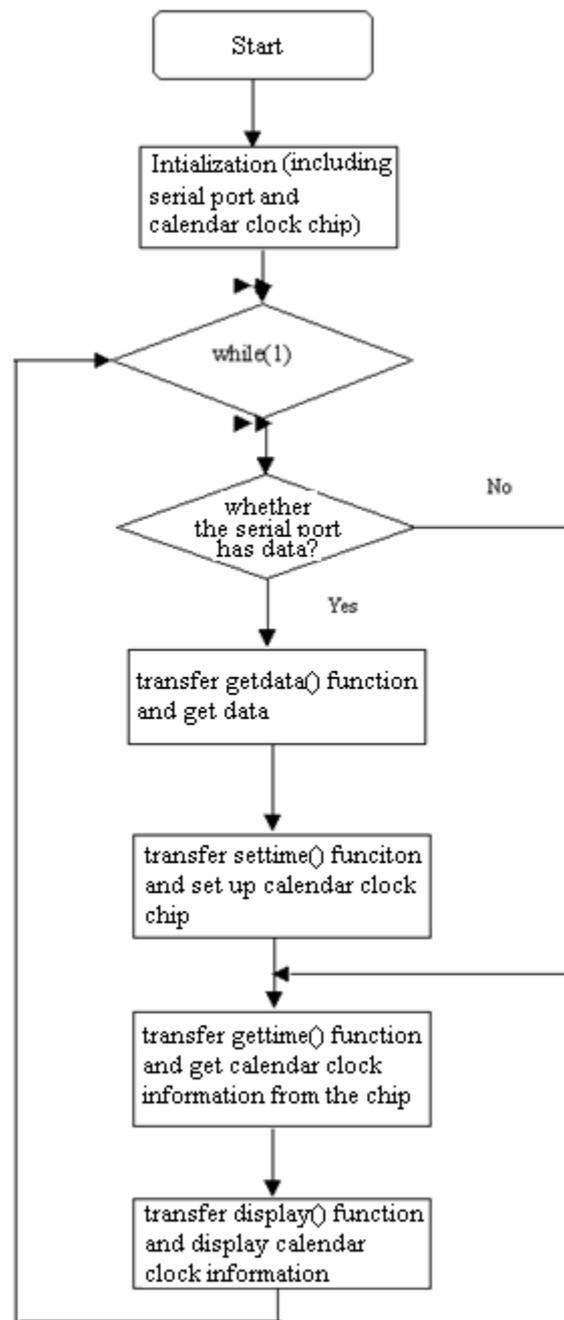


Figure 3. Program Flow



Creating Websites: Adobe Go Live vs. LiveText

Donna F. Herring

Department of Educational Resources
College of Education and Professional Studies
Jacksonville State University
E-mail: dherring@jsu.edu

Charles E. Notar (Corresponding Author)

Department of Secondary Education
College of Education and Professional Studies
Jacksonville State University
Jacksonville, Alabama 36265-1602
Tel: 1-256-782-5832 E-mail: cnotar@jsu.edu

Mrs. Judy McEntyre

Dalton State College
Educational Technology Training Center
Dalton, Georgia 30720
E-mail: jmcentyre@dscctc.org

Abstract

This article reports student comments regarding their experiences in developing a website as part of a federally funded educational technology grant. The primary goal of the grant was to create a program where pre-service teachers developed competency in the use of technology for the academic success of their students. The students during the three year grant constructed websites for two years using Adobe Go Live and one year using LiveText. The two vehicles were compared on Ease of Use, Appearance, and Content with LiveText being preferred.

Keywords: Websites, LiveText, Adobe Go Live

Creating Websites: Adobe Go Live vs. LiveText

Objectives of a federally funded educational technology grant were the pre-service teachers will be able to:

- (1) Create learning environments where advanced technology is used to teach content standards,
- (2) Use a variety of classroom management techniques necessary for successful technology integration,
- (3) Create technology connected lessons that include new designs for learning and enhanced pedagogy,
- (4) Analyze student achievement/assessment data to make decisions for structuring technology connected lessons, and
- (5) Create an educational website to enhance classroom learning.

Objective 5 was achieved through instruction the pre-service teachers received based on a state-developed educational technology program and the development of a website, integrating technology for the grade level of their clinical experiences. Pre-service teachers used the "Adobe-Go-Live" software program during the first two years of training, and LiveText during the third.

Evaluation of the grant activities included the following:

- (LOTI) pretest/post-test process;
- university supervisor observation of the teachers' use of technology in classrooms during student teaching;
- a survey of pre-service teachers' students' computer abilities;
- intern pretest/post-test self evaluation of technological abilities;
- reflective journals;
- evaluator observation of pre-service teachers in their intern classroom;
- evaluator end-of-instruction interviews; and
- website evaluations.

Information for the article was gathered from pre-service teachers' reflective journals; evaluator observation of pre-service teachers in their intern classroom; evaluator end-of-instruction interviews; and website evaluations.

The information was compiled under three headings: Ease of Use, Appearance, and Content. Ease of use is defined as ability to use the software without assistance from the instructor.

Most students experienced problems when uploading or downloading GoLive files to the web. Many could not perform this activity on their own and needed assistance every time the site was uploaded or downloaded. The LiveText site was updated by simply saving. Students did not have a problem performing this task.

Adobe GoLive required an organized file structure for the site to work while LiveText did not. Students often deleted part of the structure or saved files into the wrong folder. By moving files or folders, it was easy to cause a broken link in GoLive. Again, this was not an issue in LiveText.

GoLive was driven by tools or palettes where LiveText was very similar to MicroSoft Word. LiveText tools were very familiar and easy to use and did not require a technical background. Online, 24-7, technical help was available via LiveText but was not for GoLive.

LiveText is a web-based software package and did not require access to a school server. GoLive did require access to a local web server. The ETTC hosted the GoLive websites but expected students to gain access to a server during their first year teaching. This did not happen in many cases.

Appearance is based on the appropriateness of the site for a particular grade level.

Students were able to add very creative backgrounds to GoLive. This was not an option in LiveText. However, the LiveText websites had a more consistent look that did the GoLive websites. Students spent more time working on content when using LiveText since they were limited in ability to use backgrounds and JAVA scripts with their site.

Content is based on the degree to which the site content supports and enhances the learning of the content standards.

The conclusion that can be drawn from the observations of the students is that LiveText provided the students with an easier to use, content-rich, standards-based website. Because of the ease of use, the LiveText websites were kept current and correlated well with the curriculum. The LiveText Rubric tool and the use of United Streaming Video within the software were not available as tools inside the GoLive software and added several dimensions to the students' lesson plans.

References

- Bannister, S., Vannatta, R. A., & Ross, C. (2006). Testing electronic portfolio systems in a teacher education: Finding the right fit. *Action in Teacher Education*, 27(4), 81-90.
- Carlson, J. (2001). Easy web-site plans. *Macworld*, 18(2), 78,-80.
- Carnevale, D. (2004). Files not found: Web-hosting company loses data of students and professors. *Chronicle of Higher Education*, 50(32), A34-A34.
- Dysart, J. (2007). The beauty of web templates. *American School Board Journal*, 194(11), 50-51.
- Fleishman, G. (2002). Go dynamic with GoLive. *Macworld*, 19(11), 86-89.
- Goldsborough, R. (2003). Sprucing up a website. *Reading Today*, 20(6), 16.
- Gordon, R. S. (2003). Computer media. *Library Journal*, 128(6), 122.
- Holzberg, C. S. (2005). No code? No problem! *Technology & Learning*, 25(8), 18-21.
- Myung-Ah, L. & Hare, M. K. (2007). A web-based learning portfolio system. *The Journal of Physical Education, Recreation & Dance*, 78(9), 45-48.
- O'Reilly, D. (2002) Top-notch tools for web development. *PC World*, 20(8), 56.

Go Live	LiveText
Uploading Difficulties	Save and Make Public
Downloading Difficulties	Available with Internet access
Structure Required	Not Required
Broken Links	Less of an issue
Palette Driven	Very similar to MS Word
More technical knowledge	No technical knowledge necessary
Must have server access	Hosted by Livetext
No 24-7 support	Provides 24-7
Hugh learning curve	No learning curve

Figure 1. Ease of Use

Go Live	LiveText
Backgrounds	Option not available
Inconsistency of page look	Very consistent look from page to page

Figure 2. Appearance

Go Live	LiveText
None	Standards Tool
None	Rubric Generator Tool
None	United Streaming Videos
Update in class with help	Update on their own
Less content rich	More content rich
No K-12 student use	Easy for K-12 student projects

Figure 3. Content



The New Development of CAD in Clothing

Lingxia Zhang, Hongzhi Zhang & Yinglin Li

Tianjin Polytechnic University

Tianjin 300160, China

E-mail: zlingxia2006@gmail.com

Abstract

This text gives introduction to the history and the status quo of CAD/CAM in clothing and puts forward some questions. It also proposes a neural network model of artificial intelligence, which indicates the direction of intellectualized and three-dimensional development of clothing by use of neural network.

Keywords: Digital clothing, Artificial intelligence, Visible technology

In 1959, CAD (computer aided design) as a conception was proposed. But since it was introduced into the clothing industry in 1970s, it has accelerated to change the traditional manual way of production. At the same time, CAD has been developing at a surprising speed. At present, the application of the computer has involved in various clothing fields of design, manufacture, sales, management, and education. It has included the whole process of clothing production.

1. The development of digital clothing technology.

In 1970s, the mature CAD in clothing proposed on personal computers has laid a foundation for a wide range of application.

In 1980s, CAM (computer Aided Manufacturing) and FMS (flexible manufacturing system) further improved the key technologies relevant to CAD in clothing.

In 1990s, the development of computer information management technology promoted the integration of design, manufacture and management. And clothing CIMS (computer integrated manufacturing system) has come into being.

Nowadays, the new technologies are undergoing a revolutionary change. Made-to-measure, three-dimensional computer simulation and visualization have become the important study programs of the clothing digital technology.

Now a variety of CAD systems have been widely used in the field of clothing. Lectra system gives a comparative concentration on the integration of CAD software. Since 1998, it has adopted new strategies to expand its technology and provided a system of simulation for the entire textile industry from yarn and textile design to the manufacture of virtual goods and to the ready-to-wear. It represents the direction of development of a new generation of textile and clothing CAD.

Assyst system can simulate three-dimensional effect and give the design of structure and fabric. It has more than 400 kinds of databases for optional format-making, picturizing and modifying. It can make format automatically or according to the actual size and change the size freely. The system also has intelligent database, which can make management of styles quick and efficient.

Spanish Investronica system has made effective efforts in the application of artificial intelligence and other technologies. Its CAD / CAM system is featured with a sound cost management system, production-oriental system, and warehouse management system. After a customer selects the style and the material, it measures the customer, then makes an automatic sampling design, changes the size, arranges the material, makes autonomous single cutting out, and finishes the final step of the ready-to-wear. It is really a highly automated customer-oriented system.

2. The limitations of clothing CAD technology.

In recent years, there have existed some problems in the rapid development of the CAD technology. For example, the development of cell technology is overstressed; too much emphasis is on the general system while the design of product life cycle in the process, the development of net application and database, and feature bank are neglected.

Second, CAD is lack of software features and effects of humanity, intelligence, and visualization. Although some software can be used for the three-dimensional garment paper pattern design at present and can simulate a stronger real wearing effect, if the designer feels dissatisfied, it must return to the plain plate for further modification. Such repetition is bound to lose the advantages of CAD technology in the users' eyes.

3. The new development of apparel CAD Technology.

To make apparel CAD system more humane, more convenient and more efficient to be accepted by the designers, we tries to propose a more humane process-oriented intelligentized artificial neural network garment CAD technology. Its principle is to, on the basis of the 3-dimensional body measurement, design the clothing pieces according to the data measured, and turn the two-dimensional pieces into three-dimensional ready-to-wear, and making it simulate the motion of human beings, as shown in Figure 1.

Then through artificial neural networks, each human act will be put into the computer. The size degree of each body part will be decided and then a change will be made of the ready-to-wear model. Finally the model will be seen as two-dimensional clothing pieces.

Artificial neural network function can adopt the simple RBFNN(radial basis function neural network), which has such advantages as the strong ability of nonlinear approximation, the simple network structure, the linear net weights and output, fast learning speed, and good visual effect. It has been widely used in pattern recognition, function approximation, signal processing, system modeling, intelligent control, and other fields. RBFNN is shown in Figure 2, including an input layer, a hidden layer and an output layer. There is only one output node, indicating the degrees of fitness in GA. Its value is from 0 to 1.

For a given training set, RBF neural network can approximate the function relationship in it. By using gradient descent training, RBF neural network parameters can be obtained, including RBFNN's centre, the width and the output weight. At the initial stage when the system begins to run, the movement of every part is grouped into the random design programs. Each program means the corresponding state of the parts when body moves and clothing fitness degrees. Professional designers will judge the various programs and give the evaluation. Such evaluation value is the fitness degrees of the apparel. A set made up of groups of such relations can be used for training RBF neural network. The parameters of RBF network is the fitness degree function of different parts. With this function every part of the clothing fitness degree can be calculated.

4. Conclusions.

This text has introduced the development process of CAD technology and its future direction of growth. It has also proposed an intelligent artificial neural network garment CAD system, which emphasizes the humanity, automation, and integration of the design. Its intelligence supports self-teaching, self-organization, self-adapting, self-correcting, and parallel searching, associative memory, pattern recognition, automatic acquisition of knowledge, and other technologies. The methods used have been also studied in the medical and many other fields.

At present, the system still exist such problems as numerous calculation, poor simulation effect of the apparel with complex structure. In the clothing area, the development and application of three-dimensional CAD still lag behind other fields. To achieve the conversion from two-dimensional to three-dimensional clothing CAD, a lot of problems should be solved such as the fabric texture and dynamic performance, three-dimensional reconstruction, vivid and flexible surface modeling, and the conversion from three-dimensional garment design model to two-dimensional plain clothing pieces, and many other technical issues, which has resulted in the long cycle of developing three-dimensional garment CAD and more technical difficulty. But it represents the future direction of development of apparel CAD and has a good prospect for further growth.

References

- Gray, Stephen. (1998). *CAD / CAM in clothing and textile*. Gower publishing Limited.
- Hagan, Martin T. Demuth, Howard B. and Beale, Mark H. (2002). Trans. by Dai Kui. *Neural Network Design*. China Machine Press.
- Jiang, Zongli. (2001). *Introduction to Artificial Neural Network*. Higher Education Press.
- Qiu, Peiqiong and Chen, Jihong. (2005). Talking about the development of domestic garment CAD/CAM system. *Shandong Textile Science and Technology*. Vol.1, pp. 40-42.
- Wen, Lisheng. (1998). The use and growth of garment CAD/CAM in China. *Textile Herald*. Vol. 1, pp.45-47.
- Yang, Qin-pu., Sun, Bingnan. And Jin, Hu. (2006). The application of artificial neural network in civil engineering. *Hypothermia Construction Techniques*. Vol. 6 (114), pp. 64-65.
- Zhao, Nuoping, Zhang, Weiyuan. and Zhang, Hongzhi. (2000). The development and prospect of apparel CAD technology. *Tianjin Textile Engineering Journal*. vol.19 (5), pp.70-73.

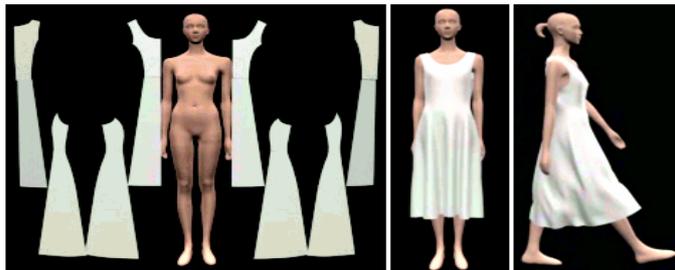


Figure 1. The virtual design effect of apparel

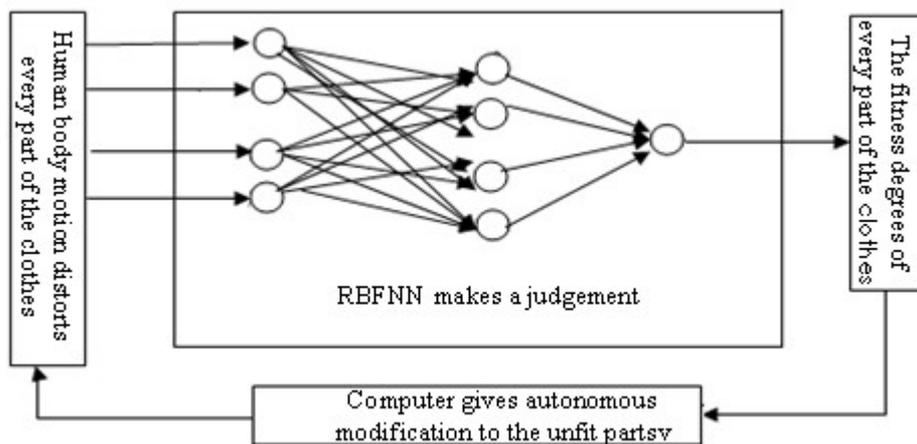


Figure 2. The application theory of neural network system



A New Approach to Sensitive Rule Hiding

Dr.K. Duraiswamy

K.S. R. College of Technology

Tiruchengode- 637 209, Tamil Nadu, India

E-mail: kduraiswamy@yahoo.co.in

Dr.D. Manjula

Department of Computer Science and Engineering

Anna University

Chennai, Tamil Nadu, India

E-mail: manju@annauniv.edu

N. Maheswari (Corresponding Author)

P.G. Department of Computer Science

Kongu Arts and Science College

Erode-638 107, Tamil Nadu, India

E-mail: mahii_14@yahoo.com

Abstract

Privacy preserving data mining is a novel research direction in data mining and statistical databases, which has recently been proposed in response to the concerns of preserving personal or sensible information derived from data mining algorithms. There have been two types of privacy proposed concerning data mining. The first type of privacy, called output privacy, is that the data is altered so that the mining result will preserve certain privacy. The second type of privacy, called input privacy, is that the data is manipulated so that the mining result is not affected or minimally affected. For output privacy in hiding association rules, current approaches require hidden rules or patterns to be given in advance. However, to specify hidden rules, entire data mining process needs to be executed. For some applications, only certain sensitive rules that contain sensitive items are required to hide. In this work, an algorithm SRH (Sensitive Rule Hiding) is proposed, to hide the sensitive rules that contain sensitive items, so that sensitive rules containing specified sensitive items on the right hand side of the rule cannot be inferred through association rule mining. Example illustrating the proposed approach is given. The characteristics of the algorithm are discussed.

Keywords: Data Mining, Privacy Preserving, Association Rules, Sensitive Rules, Clustering, Minimum Support, Minimum confidence

1. Introduction

The concept of Privacy-Preserving has recently been proposed in response to the concerns of preserving personal or sensible information derived from data mining algorithms. Successful applications of data mining have been demonstrated in marketing, business, medical analysis, product control, engineering design, bioinformatics and scientific exploration, among others. The current status in data mining research reveals that one of the current technical challenges is the development of techniques that incorporate security and privacy issues. The main reason is that the increasingly popular use of data mining tools has triggered great opportunities in several application areas, which also requires special attention regarding privacy protection. There have been two types of privacy concerning data mining. The first type of privacy, called output privacy, is that the data is minimally altered so that the mining result will preserve certain privacy (Evfimievski, 2002, Oliveira, Zaiane, 2003 a, Oliveira, Zaiane, 2003 b). The second type of privacy, input privacy, is

that the data is manipulated so that the mining result is not affected or minimally affected (Dasseni, Verykios, Elmagarmid, Bertino, 2001).

For example, through data mining, one is able to infer sensitive information, including personal information, or even patterns from non-sensitive information or unclassified data. As a motivating example of privacy issue in data mining discussed in (Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen, 2007). Consider a supermarket and two breads suppliers A and B. If the transaction database of the supermarket is released, A (or B) can mine the association rules related to his/her breads and apply the rules to the sales promotion and the goods supply. As a result, a supplier is willing to exchange a lower price of goods for the database with the supermarket. From this aspect, it is good for the supermarket to release the database. However, the conclusion can be opposite if a supplier uses the mining methods in a different way. For instance, if A finds the association rules related to B's breads, saying that most customers who buy cheese also buy B's breads, he/she can run a coupon that gives a 10 percent discount when buying A's breads together with cheese. Gradually, the amount of sales on B's breads is down and B cannot give a low price to the supermarket as before. Finally, A monopolizes the bread market and is unwilling to give a low price to the supermarket as before. From this aspect, releasing the database is bad for the supermarket. Therefore, for the supermarket, an effective way to release the database with sensitive rules hidden is required. This leads to the research of sensitive rule hiding.

In this work, the sensitive rules are given and the algorithm SRH is proposed to modify data in database so that sensitive rules containing specified sensitive items on the right hand side of rule cannot be inferred through association rule mining. The proposed algorithm is based on modifying or perturbing the database transactions so that the confidence of the association rules can be reduced.

The rest of the paper is organized as follows. Section 2 gives the view of the previous works. Section 3 presents the statement of the problem. Section 4 presents the proposed algorithm for sensitive rule hiding. Section 5 shows the example of the proposed algorithm. Section 6 analyzes the characteristics of the algorithm. Concluding remarks and future work are described in Section 7.

2. Related Work

In output privacy, given specific rules or patterns to be hidden, many data altering techniques for hiding association, classification and clustering rules have been proposed. For association rules hiding, two basic approaches have been proposed. The first approach (Saygin Y, Verykios V, Clifton C, 2001, Verykios V, Elmagarmid A, Bertino E, Saygin Y, Dasseni E, 2004) hides one rule at a time. It first selects transactions that contain the items in a give rule. It then tries to modify transaction by transaction until the confidence or support of the rule fall below minimum confidence or minimum support. Either removing items from the transaction or inserting new items to the transactions does the modification of transaction. The second approach (Oliveira, Zaiane, 2002 a, Oliveira, Zaiane, 2002 b, Oliveira, Zaiane, 2003 a, Oliveira, Zaiane, 2003 b) deals with groups of restricted patterns or association rules at a time. It first selects the transactions that contain the intersecting patterns of a group of restricted patterns. Depending on the disclosure threshold given by users, it sanitizes a percentage of the selected transactions in order to hide the restricted patterns. However, both the above approaches require hidden rules or patterns been given in advance.

The work presented here differs from the related work in some aspects are as follows: First, correlations among the sensitive rules are considered. Second, avoids the modification in transactions unnecessarily, if the confidence of the sensitive rule gets reduced. Third, alter the transactions in the cluster and finally changes can be updated in the database, which reduces the time period of database updating.

3. Problem Statement

The problem of mining association rules was introduced in (Agrawal, Imielinski, Swami, 1993). Let $I = \{i_1, i_2, \dots, i_m\}$ be a set of literals, called items. Given a set of transactions D , where each transaction T in D is a set of items such that $T \subseteq I$, an association rule is an expression $X \Rightarrow Y$ where $X \subseteq I$, $Y \subseteq I$, and $X \cap Y = \Phi$. The X and Y are called respectively the body (left hand side) and head (right hand side) of the rule. An example of such a rule is that 90% of customers buy hamburgers also buy Coke. The 90% here is called the confidence of the rule, which means that 90% of transaction that contains X (hamburgers) also contains Y (Coke). The confidence is calculated as $|X \cup Y| / |X|$, where $|X|$ is the number of transactions containing X and $|X \cup Y|$ is the number of transactions containing both X and Y . The notation \cup here is not the set union operator. The support of the rule is the percentage of transactions that contain both X and Y , which is calculated as $|X \cup Y| / N$, where N is the number of transactions in D . In other words, the confidence of a rule measures the degree of the correlation between item sets, while the support of a rule measures the significance of the item sets. A typical association rule-mining algorithm first finds all the sets of items that appear frequently enough to be considered significant and then it derives from them the association rules that are strong enough to be considered interesting. The problem of mining association rules is to find all rules that are greater than the user-specified minimum support and minimum confidence.

The objective of data mining is to extract hidden or potentially unknown but interesting rules or patterns from databases. However, the objective of privacy preserving data mining is to hide certain sensitive information so that they cannot be

discovered through data mining techniques (Agrawal, Imielinski, Swami, 1993, Evfimievski A, Gehrke J, Srikant R, 2003). In this work, an algorithm SRH (Sensitive Rule Hiding) is proposed, to hide the sensitive rules that contain sensitive items, so that sensitive rules containing specified sensitive items on the right hand side of rule cannot be inferred through association rule mining. More specifically, given a transaction database D , a minimum support, a minimum confidence and a set of sensitive items Y , the objective is to minimally modify the database D such that no sensitive rules containing sensitive items Y on the right hand side of the rule will be discovered.

4. Proposed Algorithm

In order to hide an sensitive rule, $X \Rightarrow Y$, it can be either decrease its supports, $(|X|/N$ or $|X \cup Y|/N)$, to be smaller than pre-specified minimum support or its confidence $(|X \cup Y|/|X|)$ to be smaller than pre-specified minimum confidence. In the transactions containing both X and Y , to decrease the support of Y only, the right hand side of the rule, it would reduce the confidence of the rule. In order to hide sensitive rules, when considering hiding sensitive rules with 2 items, $x \Rightarrow z$, where z is a sensitive item and x is a single large one item. In theory, association rules may have more specific rules that contain more items, e.g., $xY \Rightarrow z$, where Y is a large item set. However, for such rule to exist, its confidence must be greater than the confidence of $x \Rightarrow z$, i.e., $\text{conf}(xY \Rightarrow z) > \text{conf}(x \Rightarrow z)$ or $|xYz| > \text{conf}(x \Rightarrow z) * |xY|$. For higher confidence rules, such as $\text{conf}(x \Rightarrow z) = 1$, there will be no more specific rules. In addition, once the more general rule is hidden, the more specific rule might be hidden as well.

The algorithm tries to decrease the support of the right hand side of the rule.

Algorithm SHR

Input:

1. source database D
2. min support
3. min conf
4. sensitive items Y

Output:

a transformed database D' , where rules containing Y on RHS will be hidden.

Step 1:

Generate the association rule.

Selecting the Sensitive rules with single antecedent and consequent with the sensitive item in the consequent.

Step 2:

Clustering the rules with the right hand side has the common item and indexing the rules

Step 3:

Check all the rules in the cluster.

for cluster $i=1$ to n

{

do

{

- 1 select the rule with high confidence
- 2 find $K = \text{mincount}$
- 3 $|\text{Tr}| = \text{no of sensitive rules that are correlated}$

Such that $k < |\text{Tr}|$

- 4 $T_k = \text{find first } k \text{ transactions that contains the items in the correlated rules}$
- 5 sort T_k in ascending order by the number of items.
- 6 For $i=1$ to k

{

- 7 Choose the first transaction t from T_k .
- 8 Modify t so that y is not supported

```

9      Find the confidence of every rule
10     If confidence of all the rules < min conf
      {
11     Update the transactions in the clusters
12     Exit
13     Else
14     Update the transactions in the clusters
      }
    }
15     unchecked the rules< minconf
16     }while(rules in cluster[i]==checked)
17     }

```

Step 4: update database D as transformed database D’.

The algorithm tries to generate the association rule using (Agrawal, Imielinski, Swami, 1993). Then it selects the sensitive rules with the sensitive items in the right hand side. Cluster the rules with the common item in the right hand side of the rule and index the rule using (Oliveira, Zaiane, 2003 a).

The mincount=minimum (MSC, MCC) (Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen, 2007)

$$MSC\ x \rightarrow y = \text{count}(xUy) - [|D| * \text{min support}] + 1$$

$$MCCC\ x \rightarrow y = \text{count}(xUy) - [\text{count}(x) * \text{min conf}] + 1$$

$$MPCC\ x \rightarrow y = [(\text{count}(xUy) - \text{count}(x) * \text{min conf}) / (1 - \text{min conf})] + 1,$$

MCC=minimum (MCCC, MPCC).

Therefore it would take maximum of k no of executions to hide the rule. The rules in every cluster will be hidden.

5. Example

This section shows the example to demonstrate the proposed algorithm to hide the sensitive rules.

The items in database can be represented as a bit vector 1 and 0.(Saygin Y, Verykios V, Clifton C, 2001)

TID	Items	Items
T1	ABC	111
T2	ABC	111
T3	ABC	111
T4	AB	110
T5	A	100
T6	AC	101

Frequent item sets are generated with minimum support 0.50. The following are the frequent item sets and its support value A(1.0), B(0.6),C(0.6), AB(0.6), AC(0.6), BC(0.5), ABC(0.5). Association rules with minimum confidence 0.75 are generated. The rules and the corresponding confidence values are as follows:

C->B0.75, B->C0.75, C->A1.0, B->A1.0, BC->A1.0, AC->B0.75, AB->C0.75, C->AB0.75, and B -> AC0.75.

The item A is considered as sensitive item. The sensitive rule with single antecedent and consequent is C->A, B->A. The rules are clustered and to hide such rules the no. of transactions required is 2. In the first and second transaction the item A will be removed and placed as 0.

Then the database will be updated as

TID	Items	Items
T1	BC	011
T2	BC	011
T3	ABC	111
T4	AB	110

T5	A	100
T6	AC	101

After updating the database the rules $C \rightarrow A$, $B \rightarrow A$ will have the confidence as 0.5, which is less than minimum confidence and are hidden. While hiding the two rules the rules $BC \rightarrow A$, $C \rightarrow AB$, $B \rightarrow AC$ are also hidden and the two rules $AB \rightarrow C$ and $AC \rightarrow B$ are lost as side effects.

6. Analysis

This section analyses some of the characteristics of the proposed algorithm. The first characteristic is the time effect. The time taken to scan the database to search the sensitive rules in the database is reduced because of clustering the sensitive rules. The second characteristic is the database effect. The minimum numbers of transactions are modified because of correlation among the sensitive rules. The third characteristic is the efficiency of the algorithm. The database is updated after all the rules are hidden that saves the updating time. The fourth characteristic is the transaction effect. The alteration in the transactions are stopped when the confidence of the sensitive rules are reduced than the minimum confidence.

7. Conclusion

In this work, the database privacy problems caused by data mining technology are discussed and the algorithm for hiding sensitive rules is presented. The proposed algorithm here can automatically hide sensitive rule sets. The previous works does not consider the characteristics that are discussed here. Example illustrating the proposed algorithm is given and the characteristics of the algorithm are analyzed. Further the efficiency of the algorithm will be analyzed and improved by reducing the side effects.

References

- Agrawal R, Imielinski T, Swami A (1993). Mining association rules between sets of items in large databases. *In: Proceedings of ACM SIGMOD International Conference on Management of Data*, Washington DC.
- Dasseni E, Verykios V, Elmagarmid A, Bertino E (2001). Hiding association rules by using confidence and support. *In: Proceedings of 4th Information Hiding Workshop*, Pittsburgh, PA, pp. 369–383.
- Evfimievski A (2002). Randomization in privacy preserving data mining. *SIGKDD Explorations* 4(2), Issue 2:43–48.
- Evfimievski A, Gehrke J, Srikant R (2003). Limiting Privacy Breaches in Privacy Preserving Data Mining. *PODS 2003*, San Diego, CA.
- Evfimievski A, Srikant R, Agrawal R, Gehrke J (2002). Privacy preserving mining of association rules. *In: Proceedings of the 8th ACM SIGKDD Int'l Conference on Knowledge Discovery and Data Mining*, Edmonton, Canada.
- Oliveira S, Zaiane O (2002 a). Privacy preserving frequent item set mining. *In: Proceedings of IEEE International Conference on Data Mining*, November 2002, pp. 43–54.
- Oliveira S, Zaiane O (2002 b) A framework for enforcing privacy in mining frequent patterns. *Technical report, TR02-13*, Computer Science Department, University of Alberta, Canada.
- Oliveira S, Zaiane O (2003 a). Algorithms for balancing privacy and knowledge discovery in association rule mining. *In: Proceedings of 7th International Database Engineering and Applications Symposium (IDEAS03)*, Hong Kong.
- Oliveira S, Zaiane O (2003 b). Protecting sensitive knowledge by data sanitization. *In: Proceedings of IEEE International Conference on Data Mining*.
- Saygin Y, Verykios V, Clifton C (2001). Using unknowns to prevent discovery of association rules. *SIGMOD Record* 30(4): 45–54.
- Verykios V, Elmagarmid A, Bertino E, Saygin Y, Dasseni E (2004). Association Rules Hiding. *IEEE Trans Knowledge and Data Eng* 16(4): 434–447.
- Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen, (2007). Hiding Sensitive Association Rules with Limited Side Effects, *IEEE Trans Knowledge and Data Eng*, Vol. 19, No. 1.



New Path Filling Method on Data Preprocessing in Web Mining

Chungsheng Zhang & Liyan Zhuang

School of Mathematics and Computer Science

Inner Mongolia University for the Nationalities

536 Huo Lin He Street, Tongliao, NeiMonggol 028043, China

E-mail: zhangcs_817@www.yahoo.com.cn

Fund Project: Inner Mongolia Educational Research Projects Subsidize(8)

Inner Mongolia talents Fund Projects Granted

Abstract

The article discusses the importance of data preprocessing in web mining and gives the topology structure for the website in the view of actual condition, analyzes the limitation of reference [3] and proposes a data structure based on adjacency list. The proposed method satisfies the actual condition of topology structure for the existed website. The special data structure and path filling algorithm based on adjacency list are given. The data structure satisfies the commonness of topology structure for the existed website and the time complexity is lower.

Keywords: Web mining, Data preprocessing, Path filling

1. Introduction

With the rapid development of Internet, and the gradual increase the amount of information, it is estimated that there has 350 million web pages in 1999, and is increasing the speed of one million per day. Google has recently declared it has indexed 3000 million web pages. The World Wide Web is the largest database at present, and it is a challenging task how to access effectively these data ^[1].

An effective approach to solving these problems is web mining. Web mining is that data mining technique is applied to web data to the discovery of the interesting usage patterns and implicit information ^[2].

However in fact, data mining has the strict quality requirements to these data that are deal with. A key step of data mining is the establishment of appropriate data sets ^[7], so it seems very important to carry out data preprocessing before data mining. According to statistics, two-thirds data mining analysts consider a complete data preprocessing spends about sixty percent of the whole mining time ^[8].

Web mining is classified into three categories: web content mining, web log mining, and web structure mining ^[6]. To web content mining and web structure mining, it seems be not critical of users identity, but when users are browsing web pages, because of existing the local cache and the proxy server cache, the web page got by users pressing “backward” button on browser, hasn’t corresponding records in server log to web log mining ^[2], so we must carry out path filling, otherwise, it will seriously affect mining results.

The article [3] proposed an algorithm of STT, which a topology structure for web transforms into a binary tree, the method made some innovation indeed, but I think that there still exists some limitations in algorithm.

2. STT principle and limitation

The topology for website described in article [3] shows as figure 2-1, which is a tree structure, and in comparison with the actual condition exists in the following three problems:

- (1) The current real topology for website should be graph structure, shows as figure 2-2, and can completely exist path $E \rightarrow F$;
- (2) Generally, the depth is much larger than the width in the topology of website, tree transforms into binary tree in the article [3], which certainly will cause its search depth increase and algorithm efficiency decrease.
- (3) When general references solving path, all start from root node, while the actual condition is not true, solving path can start from any node.

3. Data structure based on adjacency list and path filling algorithm

3.1 Data structure of topology of web

After analysis and selection, data structure of topology for website uses the adjacent list data structure, constructed as follows.

Definition 3.1 Website nodes set can be described as a sequence L , L equals to $\{i \mid 1 \leq i \leq n, \text{ in which } n \text{ is the total node numbers}\}$, numbering of i starts from root node, the first is the internal layers from left to right, and the last is external layer from up to down, that is the width priority method sorts nodes.

For example, in figure2-2 symbolic node sequence is $\{A, B, C, D, E, F, G, H, I, J, K, L, M\}$, its corresponding digital sequence is L and L equals to $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$.

Theorem 3.1 Any filling path must be a node in searched path.

For example, suppose that there exists user path ABEIF in figure2-2, and filling path is ABEIBF, B is obviously a node in searched path.

Prove: The reasons of causing discontinuous point is that users come true by “backward” button on browser, and the backward node is certain in searched path, so theorem holds.

Theorem 3.2 If the sub-node numbers of a node in a searched path are less than or equal to 1, the node can not be filling node.

Prove: If the sub-node numbers of a node in a searched path equal to 0, the node is leaf node and can not be accessed by “backward” button, so the node can not be taken as filling node.

If the sub-node numbers of a node in a searched path equal to 1, and if the node can be arrived by “backward” button, the node is the direct parent node of its only sub-node, because it do not exist other sub-nodes, the node can not be taken as filling node.

Definition 3.2 Data structure of the topology for website is defined in language C as follows:

```
// defining adjacency sub-node
typedef struct CTNode{
    int child;                // sub-node number
    struct CTNode *next      // pointer of next node
}*ChildPtr;
// defining webpage node list
typedef struct{
    int child_count;         // the number of sub-node
    TElem Type data;        // node marking
    ChildPtr firstchild;    // pointer of the first node
}CTBox;
// defining website structure
typedef struct{
    CTBox nodes[MAX_TREE_SIZE]; // node list
int n, r;                   // the total number of node and the root of node
    }Ctree;
    int s_path[ ];           // original path
    int d_path[ ];           // path after filling
```

3.2 Path filling algorithm

Suppose i is the original path scan variable, j is the object path scan variable, and pre_i is the original path backtracking scan variable.

```
Void main ( )                // the main function of algorithm
{
int i =0;                    // initialize variables
```

```

int j=0;
int pre_i=0;
d_path[0]=s_path[0];           // deal with the first node of the original sequence
i=1;
for (i=1; len(s_path); i++)    // start from the second node and scan the original
    // sequence
{
    if (nodes[i-1].first=s_path[i]) { // is the first sub-node of next node, and assign value
// directly
        d_path[j]=s_path[i];
    else // otherwise, scan other sub-nodes
        { if (!find (i, j, nodes[i-1].first. next))
            for (pre_i = i-1; 0 ; i -- ) { // is not sub-node, and start backtracking
                if (nodes[pre_i].child_count <= 1)
                    continue; // if the sub-node numbers of the current node are less
else //than or equal to 1, by theorem 3.2, then continues
// backtracking, because the node is not filling node
                { if (!find ( pre_i, j, nodes[pre_i].first )
                    continue; // otherwise, scan sub-chain of the current node
                else
                    break; }; // found then return
                };
            };
        };
    i++;
    j++;
}

int function find ( int i, j, CTNode q ) // scan sub-chain function
{
int tag=0; // setting found mark: tag equals to 0, and means no
// found; tag equals to, and means found.

while (q) {
    if (q. child= s_path[i]) // found, then ends circulation
        { d_path[j]= s_path[i];
            tag =1;
            break; };
    else // otherwise, continues to scan
        { q=q. next;
            tag=0; };
    };
return tag;
}

```

4. Conclusion

Data structure based on adjacency list is proposed in the article, the proposed method satisfies the actual condition of topology structure for the existed website. Meanwhile, path filling algorithm based on adjacency list data structure is given. The algorithm overcomes the shortcomings that existed website transforms into binary tree to cause larger search depth, generally the width is more than the depth in existed website. And the algorithm uses only array, structure and simple circulation, so it is simple and low complexity.

References

Doro Tanasa, Brigitte Trousse. Advanced data preprocessing for intersites web usage mining[J]. IEEE Intelligent Systems, 2004, (3/4): 59-65.

Han, Jiawei, Kamber M. (2001). Conception and Technique of Data Mining. Beijing: Mechanical Industry Publishing House.

Jaideep Srivastava, Robert Cooley, Mukund Deshpande, et al. Web usage mining: Discovery and applications of usage patterns from web data [J]. Proc ACM SIGKDD, 2000, 1(2): 12-23.

Li, Chaofeng. (2007). Design and realization of data preprocessing algorithm in web usage mining. *Journal of South-Central University for Nationalities*, 26, 56-60.

Li, Liebiao, Zhang, Haipeng, Zhou, Yafeng. (2007). Research on data preprocessing methods in web log mining. *Computer Technology and Development*, 17,45-52.

Liu, Lijun, Zhou, Jun, Mei, Hongyan. (2007). Data preprocessing of web usage mining. *Computer Science*, 34, 200-204.

Ma, Ruiming, Li, Xiangyun. (2007). Research on data preprocessing technology in web log mining. *Computer Engineering and Design*, 28, 2358-2360.

Margaret H.Dunham. (2005). Data Mining Course. Beijing: Tsinghua University Press.

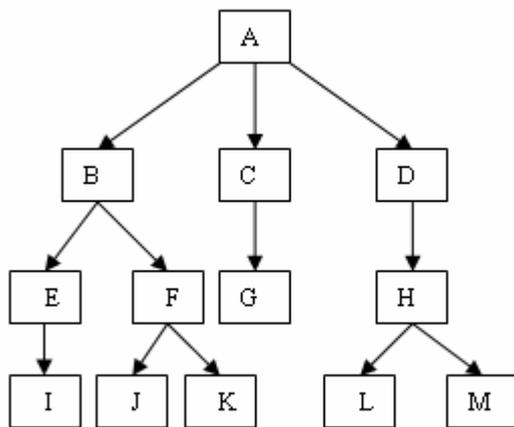


Figure 2-1 topology of web in article [3]

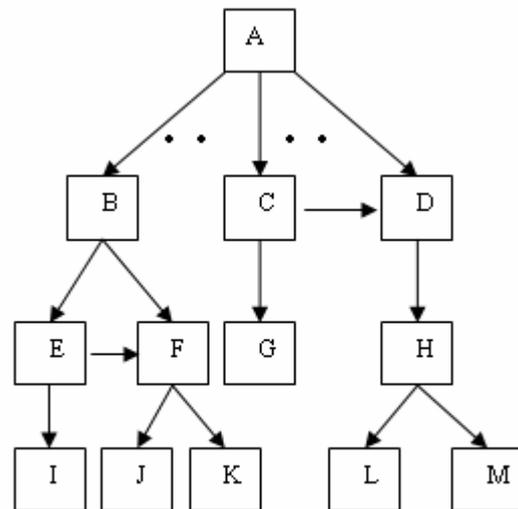


Figure 2-2 the actual topology for existed web



Hough Transform and Its Application in Vehicle License Plate Tilt Correction

Zhenhui Zhang & Shaohong Yin

School of Computer Technology and Automation

Tianjin Polytechnic University

Tianjin 300160, China

E-mail: zzh56888@163.com

Abstract

In a vehicle license plate recognition system, tilt vehicle license plate has a bad effect on its character segmentation and recognition. In this paper, tilt models of a plate are analyzed and a approach for number plate tilt correction is presented. Hough Transformation is an effective method to obtain vertical or horizontal angle. Though rotating a correct angle, Tilt vehicle license will be rectified using interpolation a rotation method. Experimental result shows that the method can be implemented easily when dealing with dirty number plates and license plate in variant lighting conditions.

Keywords: Vehicle license plate, Hough transformation, License plate recognition

1. Introduction

The Hough Transform is a global method for finding straight lines hidden in larger amounts of other data. It is an important technique in image processing. For detecting lines in images of vehicle plate, the image is first binarised using some form of threshold and then the positive instances catalogued in an examples dataset. The images of vehicle plate have obvious line which is edge of plate. By using binarised, the edge will be a rectangle. Because each point (d, t) in Hough space corresponds to a line at angle T and distance D from the origin in the original data space. The value of a function in Hough space gives the point density along a line in the data space. For each point in the original space consider all the line which go through that point at a particular discrete set of angle, chose a priori. For each angle T , calculate the distance to the line through the point at that angle and discretise that distance using a priori chosen discretisation, giving value d . Make a corresponding discretisation of the Hough space –this will result in a set of boxes in Hough space. These boxes are called the Hough accumulators. For each line we consider above, we increment a count (initialized at zero) in the Hough accumulator at point (d, t) . After considering all the lines through all the points, a Hough accumulator with a high value will probably correspond to a line of point. So the angle of line will be obtained. At last, rotating a certain angle will rectify the vehicle license plate.

2. Pre-processing of image

In order to improve the accuracy and efficiency in the process of Hough transformation, the image of the vehicle plate must be processed by the Global Threshold Binarization Method. The simplest way is choose a threshold value, and classify all pixel with value above this threshold as white, and all other pixels as black. For the vehicle plate the pixel value is easy to find the right the threshold by calculating average value of all pixel value in the image. Experiment shows that this method can get the best binarization threshold distinguish the edge and background perfectly. Meantime, it can also avoid disconnection. Figure.1, Figure.2 show the result of binarizing processing.

3. Hough transformation and straight line detection

The Hough transformation is a standard tool in image analysis that allows recognition of global patterns in an image space by recognition of local pattern in a transformed parameter space. It is particularly useful when the patterns one is looking for are sparsely digitized have “holes” and the picture are noisy. Especially in detected straight line in the licenses plate.

The basic idea of this technique is to find curves that can be parameterized like straight lines in a suitable parameter space.

As show in figure 3, we assume them parameterized in the form: $\rho = x \cos(\theta) + \sin(\theta)$, where ρ is the perpendicular

distance from the origin and θ the angle with the normal. Collinear point (x, y) with $i=1, \dots, N$, are transformed into N sinusoidal curves $\rho = x_i \cos(\theta) + y_i \sin(\theta)$ in the (θ, ρ) plane, which intersect in the point (θ, ρ) .

Care has to taken when one quantizes the parameter space (θ, ρ) , when the bins of the (θ, ρ) space (it is easy to visualize the transform as a two-dimensional histogram) are chosen too fine, each intersection of two sinusoidal curves can be in a different bin. When the quantization is not fine enough, on the other hand, nearly parallel lines which are close together will lie in the same bin.

For a certain range of quantized values of parameters ρ and θ , each (x, y) is mapping into the (θ, ρ) space and the points that map into the locations are accumulated in the two-dimensional histogram THIST (ρ_m, ρ_m) , i.e. $\text{THIST}(\rho_m, \rho_m) = \text{IHIST}(\rho_m, \rho_m) + 1$. if a grey level image $g(x, y)$ is given, and g is the grey value at the point (x_i, y_i) the grey values are accumulated: $\text{IHIST}(\rho_m, \rho_m) = \text{IHIST}(\rho_m, \rho_m) + g_i$.

In this form, the Hough transform is not basically different from the discrete Radon Transform typically used for reconstruction of three-dimensional images from two-dimensional projections.

Local maxima of the pixel intensity $\text{IHIST}(\rho_m, \rho_m)$ identify straight line segments in the original image space. Ideally, the Hough domain has to be searched for a maximum only once. In situation where a picture contains many patterns of different size, it may, however, be necessary to take out first those patterns in the original image space that correspond to clearly identifiable peaks in the Hough domain and to repeat the process.

We can see the process that The Hough Transformation used in the license plate from Figure3 and Figure4. The Figure3 is original image space, the pixel value is accumulated in The Hough Transformation space. In the Figure5, there are two peak points which point into the long parallel lines which are the edge of The License Plate. The value θ is the angle of The Vehicle Plate Tilt. At last, the angle will be obtained.

4. The types of vlp tilt

The image of the vehicle plate is got by camera from different direction. So the pictures are different. On the whole, they can be divided into three different types such as the horizontal tilt, the vertical tilt, the horizontal and vertical tilt. The models of their tilts are shown in Figure.6, Figure.7, Figure.8.

5. Image rotation

The angle of the vehicle plate obtained, the last step is how to rotation certain angle. Image rotation is performed by computing the inverse transformation for every destination pixel. Output pixels are computed using bilinear interpolation. RGB image are computed by evaluating one color plane at a time. The follow formulation will show the principle of the rotation.

Before rotating

$$\begin{cases} x_0 = r \cos(\alpha) \\ y_0 = r \sin(\alpha) \end{cases}$$

After rotating

$$\begin{cases} x_1 = r \cos(\alpha - \theta) = x_0 \cos(\theta) + y_0 \sin(\theta) \\ y_1 = r \sin(\alpha - \theta) = -x_0 \sin(\theta) + y_0 \cos(\theta) \end{cases}$$

Matrix form

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix}$$

Inverse operation

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

6. Result and conclusion

The Hough transformation is a standard tool in image analysis that allows recognition of global patterns in an image space by recognition of local pattern in a transformed parameter space.

A new method of Hough transformation is presented in this paper, experiments show that this method can accurately detected the line and obtain the angle. It is proved to be effective.

In the course of experiment, the methods of Hough transformation can accurately detected the line and obtain the angle.

This paper used by the algorithm can effectively in the car license plate image positioning, and can calculate the tilt angle plates, using interpolation method of rotating the position of cutting a correction. The results from simulation, the algorithm can tilt plates in a control error correction. About the results of this system to meet the requirements for the future lay the foundation for the segmentation of the characters.

Experiment results show that the method can be implemented easily and offers robustness when dealing with dirty number plates and license plates in variant lighting conditions.

References

Bao Ming, Lu Xiaobo. (2004). Distorting correction of license plate based on Hough transform [J]. *Computer and Communications*, 2004, 22 (2): 57-60.

Bernsem J, (1986). Dynamic thresholding of gray-level image [A], Pro of 8th Intel Conf on Patt Recog [C], Paris, France: IEEE computer Society Press, 1986:1251~1255.

Lu, Y I. (1995). Machine printed character segmentation-An overview [J]. *Pattern Recognition*, 1995, 28, 10: 67~80.

Codaro E N, Nakazato R Z,Horovistiz A, et al. (2002). An Image Processing Method for Morphology Characterization and Pitting Corrosion Evaluation [J]. *Materials Science and Engineering*, 2002,334A(1-2): 298-306.

Lee E.R., Kim P.k., Kim H.J. (1994). Automatic recognition of car license plate using color image processing. In: Proceedings of IEEE International Conference on Image Processing, Austin, Texas, 1994, 2:301~305.

Li, Wenju, Liang, Dequn, Cui, Lianyan. (2004). A novel approach for VLP tilt correction [J]. *Information and Control*, 2004, 33: 104-108.

Wen C-Y,Yu C-C,Hun Z-D. (2002). A 3-D transformation to improve the legibility of license plate numbers [J]. *Journal of Fotrnsic Sciences*, 2002, 47(3):578-585.

Xu, Hui. (2002). Practical Projects of Digital Image Implemented by VC++ [m]. The People's Posts and Telecommunications Publishing House, Beijing, 2002.

Zhang, Yujin. (2000). *Image process and analysis* [M]. Beijing: Tsinghua University Press, 1999, 68-70.

Bian, Zhaoqi, Zhang, Xuegong. (2000). Pattern recognition. 2nded [M]. Beijing: Tsinghua University Press, 2000.223-228.



Figure 1. primitive image



Figure 2. Adjusted image

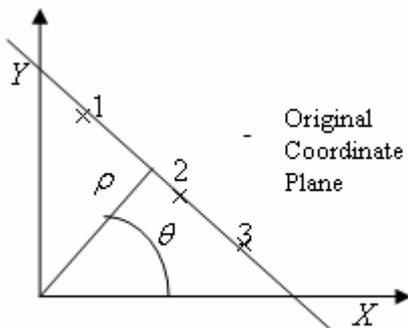


Figure 3. Define of line parameter

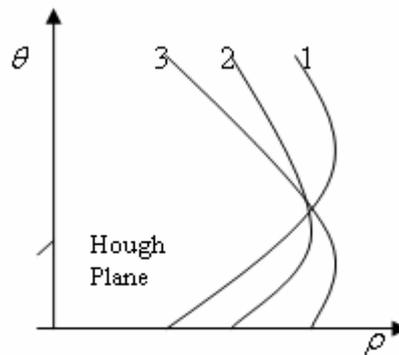


Figure 4. Hough transformation space

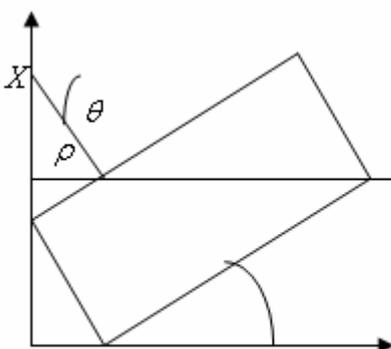


Figure 5. Original image space

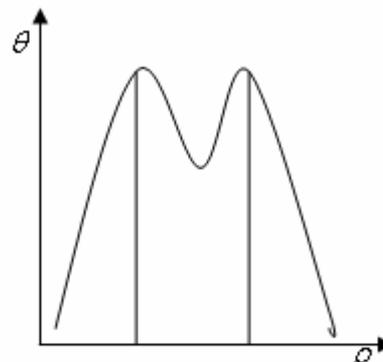


Figure 6. Hough transformation space

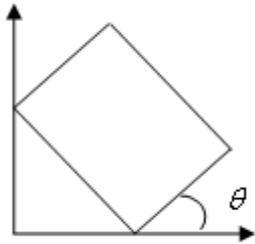


Figure 7. Horizontal tilt

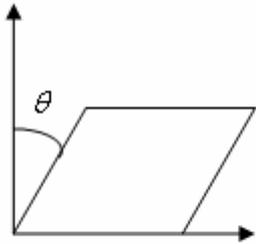


Figure 8. Vertical tilt

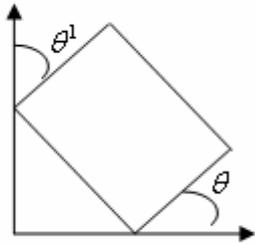


Figure 9. Horizontal and vertical tilt

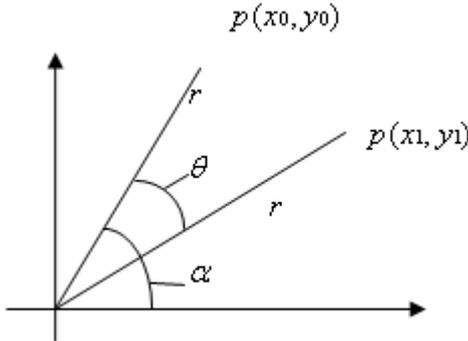


Figure 10. Rotating coordinates changes of point



Study on the Scheme for RSA Iterative Encryption System

Yulian Shang

School of Information and Engineering

Taishan Medical University

Taian 271016, China

Wuyuan Jia

School of Chemistry and Chemical Engineering

Taishan Medical University

Taian 271016, China

Peng Li

School of Information and Engineering

Taishan Medical University

Taian 271016, China

Pengfei Zhu

National Laboratory of Pattern Recognition

Automation Institute of Chinese Sciences Academy

Beijing 100080, China

Abstract

Based on the RSA public key cryptosystem, in this article, we put forward one sort of iterative encryption scheme based on RSA. The multiple keys can make attackers' attacks to the system more difficult and further enhance the security of the key. Through the simulation of software, we realize the encryption transmission of binary image and validate the improvement of the RSA iterative encryption effect.

Keywords: Public key cryptosystem, Iterative encryption system, Key security

1. Introduction

In 1978, Rivest, Shamir and Adleman in MIT put forward the asymmetric key (public key) system which was called the RSA cryptosystem (R. L. Rivest, A. Shamir & L. M. Adleman, 1978, p.20-26). The RSA cryptosystem is one milestone in the developmental history of cryptography, and it is the representative arithmetic of the public key cryptosystem based on big integer decomposability. Because of its simple arithmetic, it is very easily realized in actual application, and it is the most successful cryptography in theory and one of public key system which is most extensively applied at present (Yang, 1999, p.63-66 & Shang, 2004, p.238-242). In the traditional RSA encryption system, we require that the key length has 2048 bits in order to ensure the security of the encryption system, and Lu Changjin and Shi Kaiquan adopted the iterative encryption-decryption arithmetic and obtained better encryption effect (Lu, 2003, p.546-549 & Lu, 2003, p.681-684). In this article, we also put forward the iterative encryption-decryption arithmetic based on the RSA public key cryptography. Because the operations used in this scheme are power arithmetic and modular arithmetic, the complexity of the operation is same to the RSA cryptography. In practice, we can select iterative times properly according to concrete situation. The result of the experiment indicates that the security of the iterative encryption is higher than the security of the single encryption, and the fatalness that the illegal attacker attacks the encryption system is lower and the encryption system is safer.

2. The iterative public key encryption system based on RSA

Based on the RSA public key encryption system, in this article, we put forward an iterative encryption scheme based on RSA to enhance the security of the key. We will concretely introduce the iterative public key encryption system based on RSA and give the structure of this iterative arithmetic (because of the limited length of the article, we only give the

iterative RSA arithmetic with multiple key pairs and the iterative RSA signature arithmetic).

2.1 The RSA iterative encryption arithmetic with multiple key pairs

Supposing: A and B are the two parties of encryption communication, A is the transmitter of information, and B is the receiver of information. Starting from the universality, we suppose that A and B respectively select random t public keys and private keys, the public keys are public, and the private keys are secret, and they are respectively independent each other, and one can not be deducted from the other one. The public information is $m \in M$, and the secret information is c .

For A: Select t pairs of big prime number p, q randomly and compute $n^A = \{n_1^A, n_2^A, \Lambda, n_t^A\}$. The corresponding public key set is $e^A = \{e_1^A, e_2^A, \Lambda, e_t^A\}$, e_i^A is one random public key of A. The corresponding private key set is $d^A = \{d_1^A, d_2^A, \Lambda, d_t^A\}$, d_i^A is one random secret key of A. $i = 1, 2, \Lambda, t$.

For B: Select t pairs of big prime number p, q randomly and compute $n^B = \{n_1^B, n_2^B, \Lambda, n_t^B\}$. The corresponding public key set is $e^B = \{e_1^B, e_2^B, \Lambda, e_t^B\}$, e_j^B is one random public key of B. The corresponding private key set is $d^B = \{d_1^B, d_2^B, \Lambda, d_t^B\}$, d_j^B is one random secret key of A. $j = 1, 2, \Lambda, t$.

Definition 2.1: Define the public information m , $c_1 = (m^{e_1^B}) \bmod n_1^B$ is the first order encryption of m , $c_i = (c_{i-1}^{e_i^B}) \bmod n_i^B$ is i times iterative encryption of m , where $i \in \{1, 2, \Lambda, t\}$.

Definition 2.2: Define $c_t = (c_{t-1}^{e_t^B}) \bmod n_t^B$ is t times iterative encryption of m , and it is also called as the secret information c corresponding with m in the encryption communication.

Definition 2.3: Define the corresponding information c with t times iterative encryption, and $m_1 = (c^{d_1^B}) \bmod n_1^B$ is once decryption of c , $m_j = (m_{j-1}^{d_{t-j+1}^B}) \bmod n_{t-j+1}^B$ is j times decryption of c , where $j \in \{1, 2, \Lambda, t\}$.

Definition 2.4: Define $m_t = (m_{t-1}^{d_1^B}) \bmod n_1^B$ is t times decryption of c with t times iterative encryption, i.e. the m produced through decryption.

The process of iterative encryption communication between A and B is described as follows.

A gets m and completes the iterative encryption, and transmit c to B.

Suppose that m is one random public information sect in the public information M which is digitized and grouped, and according to the appointed order, A gets m and encrypt it as c by the public key $e^B = \{e_1^B, e_2^B, \Lambda, e_t^B\}$ of B.

The process of t 'th order encryption is:

$$c_1 = (m^{e_1^B}) \bmod n_1^B$$

$$c_2 = (c_1^{e_2^B}) \bmod n_2^B$$

...

$$c = c_t = (c_{t-1}^{e_t^B}) \bmod n_t^B$$

Where, the expression of the i 'th time encryption is $c_i = (c_{i-1}^{e_i^B}) \bmod n_i^B$, c_i is the secret information of m through i times encryption. A transmits c to B. $i \in \{1, 2, \Lambda, t\}$.

B receives c and completes the iterative decryption and obtains m .

When B receives c , according the corresponding order, B implements iterative decryption to the secrete information by his own private key, and finally obtains the public information m .

The process of t 'th order decryption is:

$$m_1 = (c^{d_1^B}) \bmod n_1^B$$

$$m_2 = (m_1^{d_{t-1}^B}) \bmod n_{t-1}^B$$

...

$$m = m_t = (m_{t-1}^{d_1^B}) \bmod n_1^B$$

Where, the expression of the j 'th time encryption is $m_j = (m_{j-1}^{d_{t-j+1}^B}) \bmod n_{t-j+1}^B$. B decrypts c , obtains the public information m and completes the whole encryption communication. $j \in \{1, 2, \Lambda, t\}$.

Obviously, when $t=1$, the public key of A is e^A , the corresponding private key is d^A , the public key of B is e^B , the corresponding private key is d^B , so the above encryption communication process can be simplified as follows:

A encrypts the public information m to the secret information c and transmits it to B.

$$c = (m^{e^B}) \bmod n^B$$

B decrypts the secret information c to the public information m .

$$m = (c^{d^B}) \bmod n^B$$

The process of encryption communication can be reverted the encryption form in the initial RSA system.

2.2 Arithmetic discussion

(1) As viewed from the encryption-decryption arithmetic, the iterative encryption system based on RSA is still based on the RSA public key cryptography, so its security is still based on the problem of big integer decomposability. Though the arithmetic still uses the iterative arithmetic, but it is the modular arithmetic and power arithmetic, so the complexity of the operation is not be added comparing with RSA cryptography.

(2) In this iterative encryption system based on RSA, because the transmitter A implements iterative encryption to the public information by t public keys $e^B = \{e_1^B, e_2^B, \Lambda, e_t^B\}$, the key attacker can decrypt the secrete information when he obtains all t secrete keys $d^A = \{d_1^A, d_2^A, \Lambda, d_t^A\}$ which is very difficult to be completed, so when B loses one or $t-1$ keys (i.e. the number of the key which is lost is less than t), the information can not be lost and juggled. Therefore, the key space in the iterative arithmetic increase the randomness of the key selection, and the order that the iterative encryption selects the key is random, which all can enhance the security of the key.

3. Example analysis and software simulation

In this article, we use the improved encryption scheme to encrypt the image (image element value), select the binary image lean.bmp (256×256) to implement the simulation experiment, and we can see the improvement of encryption effect obviously.

3.1 Software simulation

The experiment of software simulation mainly uses Matlab6.0 to compete the encryption transmission of binary image (the encryption of image element value in fact), and the experiment result includes following aspects which take the image lena.bme as the examples.

The initial image is Figure a, the once encryption-decryption of image are shown in Figure a1 and Figure a2, and the image iterative encryption arithmetic with multiple key pairs are shown in Figure a3 and Figure a4.

3.2 Analysis of experiment result

From above the results of simulation experiment, we can obtain following conclusions.

(1) When the image is implemented encryption-decryption operation every time, it can be resumed to the image before encryption without distortion. We use the data encryption arithmetic to encrypt and transmit the image element values, and its principle is based on the data encryption, but not the direct encryption transmission of the image, which also proves the validity of various encryption-decryption algorithms based on RSA when the image element values are taken as the public information and implemented image encryption arithmetic.

(2) As viewed from the visual effect of image encryption, the iterative encryption (two times) (Figure a3) has better effect than the once encryption (Figure a1), and the image is more disorder which indicated the change of the image element values are larger. Therefore, the system security of the iterative encryption is higher than the system security of the once encryption, and the key of the encryption system is safer.

(3) The improvement of encryption arithmetic based on RSA encryption can be used in the encryption transmission of image, which can implement encryption transmission when the image is not be distorted, and complete the image encryption transmission from another view, and it has very extensive applications in practice.

4. Conclusions

In this article, we put forward the iterative public key encryption system based on RSA, and the practice proves that the RSA iterative encryption system can obtain better encryption effect. However, in this iterative encryption system, large of encryption key set and decryption key set are used, and one important problem that we face is how to manage these enormous key sets and make illegal attackers can not purloin the keys of the encryption communication system, otherwise, there are not effective method to produce randomly big random number to produce the RSA keys.

References

- Huang, Yuanfei & Chenlin. (2001). *Information Security and the Core Technology of Encryption and Decryption*. Shanghai: Pudong Electronic Press. p.134-154.
- Lu, Changjin & Shi, Kaiquan. (2003). One Dimension Iteration Encryption-recognition of Image. *Journal of Shandong University (Engineering Science)*. No. 33(5). p.546-549.
- Lu, Changjin & Shi, Kaiquan. (2003). Two Dimension Mixed Iteration Encryption-recognition of Image. *Journal of Shandong University (Engineering Science)*. No. 33(6). p.681-684.
- R. L. Rivest, A. Shamir & L. M. Adleman. (1978). A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. *Communications of the ACM*. No.21(2). p.20-26.
- Shang, Yulian, Fu, Haiyan & Shi, Kaiquan. (2004). Improved Authentication and Key Distribution for Internet Environment. *Journal of Hainan Normal University (Natural Science)*. No. 17(3). p.238-242.
- Yang, Zhimin. (1999). Making the Way of RSA Reusing the Public Key to Truth in Computer Information Network. *Journal of Anhui Institute of Architecture & Industry (Natural Science)*. No. 7(4). p.63-66.



Figure a. lena.bmp (Initial Image)

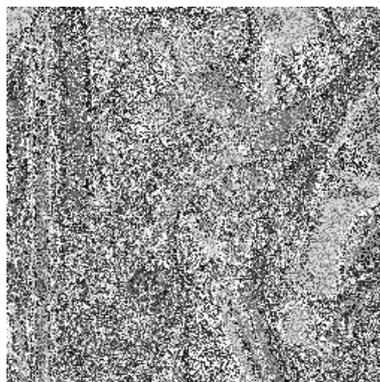


Figure a1. Once Encryption Image



Figure a2. Once Decryption Image

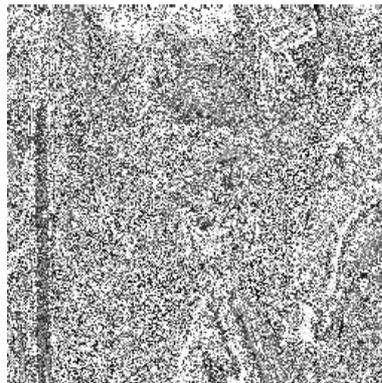


Figure a3. Iterative Encryption Image (Twice)



Figure a4. Iterative Decryption Image



The Design and Implementation of Precious Test and Quality Control System Based on Delphi

Ruoping Wang & Liguo Lai

Changchun Institute of Equipment and Technics

Changchun 130012, China

Tel: 86-431-8550-3161 E-mail: cc.ruoping@163.com

Abstract

In this article, utilizing Delphi 7.0 as the development platform, we design the system to collect or record measurement data through electronic digital caliper and digital table, compute control parameters of quality control chart, draw the process control chart of controlled objectives and read, analyze and manage the measurement data.

Keywords: Delphi 7.0, Electronic digital caliper, Precious test, Quality control

1. Introduction

The electric digital caliper is a sort of apparatus to measure the distance of relative removal between two measured faces, and it is one product of mechanical and electrical integration. It can transfer the collected and tested accessory data to the computer through the communication interface. Because the electric digital caliper has many advantages such as clear reading, clear display, convenient use and high effectiveness, it is broad applied in the domain of manufacturing in China.

2. Objective of system

The objective of the system is to utilize the electric digital caliper and digital table to collect or manually record measurement data, compute the control parameters of quality control chart, draw the process control chart of controlled objective, effectively analyze and judge the stability of working procedure quality in the production process, find the abnormalities occurred in the production process in time, exactly the trend of quality, and through translating the analysis control chart into the production control chart, prevent and reduce the occurrence of over-tolerance products, make the making process in the technical stable state, and make the quality of working procedure divivable, and quickly measure the influence degree of system error and adopt corresponding measures to control the quality, and follow, transfer and manage measurement data to offer quality guarantee for the successive productions.

3. System composing and operation environment

3.1 Composing of system

The precious test and quality control analysis system mainly includes seven function modules such as system management, test database, test repository, quality control, basic information query and print, and system service. The basic design concept and flow of the system is shown in Figure 1.

3.2 Operation environment

3.2.1 Hardware

The lowest collocation of hardware that the software operates includes Pentium 3 CPU (500M and over 500M main frequency), EMS memory of 128M or over 128M and hard disk of 6G or over 6G.

3.2.2 Development environment

This system is designed, developed and transferred under Microsoft Windows XP, and the front-end client program uses the Delphi 7.0 as the software development program, and the background database uses the Microsoft Access, and the documents in the program are made through Microsoft Office FrontPage 2003.

3.2.3 Operation environment

The operating system can use Windows NT, Windows 98, Windows 2000 or Windows XP.

4. System functions and software structure design

4.1 System functions

- (1) Realizing the supervision of machining quality for the structured components connected with the electric digital caliper.
- (2) Data read. Reading the quality control curve or histogram drew through comparing the test data on the electric digital caliper with the drawing size.
- (3) Realizing historical data query and displaying historical data curve or histogram.
- (4) Realizing exports of report forms and data.

4.2 Structure design of software

The implementation of precious test and quality control includes three parts. The first part realizes the record of basic data. The second part realizes the transfer, reading, display, treatment, memory and curve (histogram) drawing of data through the measurement of electric digital caliper. The third part realizes the historical data query, and curve (histogram) query, printing of report forms. The works of various parts include following aspects.

4.2.1 Data transfer

The data is transmitted by ASCII code (and text formatting), and every measurement data is ended by the enter symbol and is transmitted continually. The data has no frame head, and the ASCII code of measurement has 10 bytes, i.e. the front is the positive sign or minus, then 7 figures and one decimal, and the transmitted byte is fixed.

4.2.2 Data reading and graph protracting

- (1) The core codes of data readin.

try case Msg of

```

Apw_triggerdata: {got 'login', send response} ;
Apw_triggeravail: {extract and display/process the data};
begin
S:="";
for W:= 1 to Data do
begin
C := ComPort1.Char;
if CB_hexGet.Checked then
begin
S:=S+inttohex(byte(C),2)+' ';
end else
S:=S+C;
inc(I1);
end;

```

- (2) Drawing the curve. Use the function in Delphi 7.0 to draw the curve which is seen in Figure 2.

```

Canvas.MoveTo(0, 0);
Canvas.LineTo(X, Y);

```

4.2.3 Data memory and history data query

- (1) The core codes in the data memory process.

```

table.FieldName('drawing size').AsString:=bsskinedit9.Text;
table.FieldName('actual measurement result').AsString:=bsskinedit10.Text;
.....
table.Post;

```

- (2) Historical data query (seen in Figure 3). According to the query condition inputted, we can confirm the query result. The core codes:

```

query.SQL.Add('select * from testing records');

```

```

stemp:='select * from testing records ';
query.Open;
begin
stemp:='select * from testing report where number=
'+ '' + trim(edit5.text)+ ''';
end;
    
```

5. Conclusions

Through the process control chart of controlled objective, the precious test and quality control system designed in this article can effectively analyze and judge the stability of working procedure quality in the production process, find the abnormalities occurred in the production process in time, exactly the trend of quality, and through translating the analysis control chart into the production control chart, prevent and reduce the occurrence of over-tolerance products, make the making process in the technical stable state, and make the quality of working procedure divivable, and quickly measure the influence degree of system error and adopt corresponding measures to control the quality, and follow, transfer and manage measurement data to offer quality guarantee for the successive productions. This system makes the software interface more beautiful and the data read more convenient and better fulfills users' demands through the exertion of the third party's controls such as Businessinform and Spcomm.

References

Eric Harmon. (2000). *Delphi COM Programming*. Beijing: Electronic Industry Press. October, 2000.
 Huangjun, Xiongyong & Liuyan. (2002). *Manual of Oracle9iPL/SQL for Developing Personnel*. Beijing: People's Post and Telecommunication Publishing House. January, 2002.
 Liwei. (2005). *Practical Object-oriented Development with Delphi*. Beijing: China Machine Press. April, 2005.

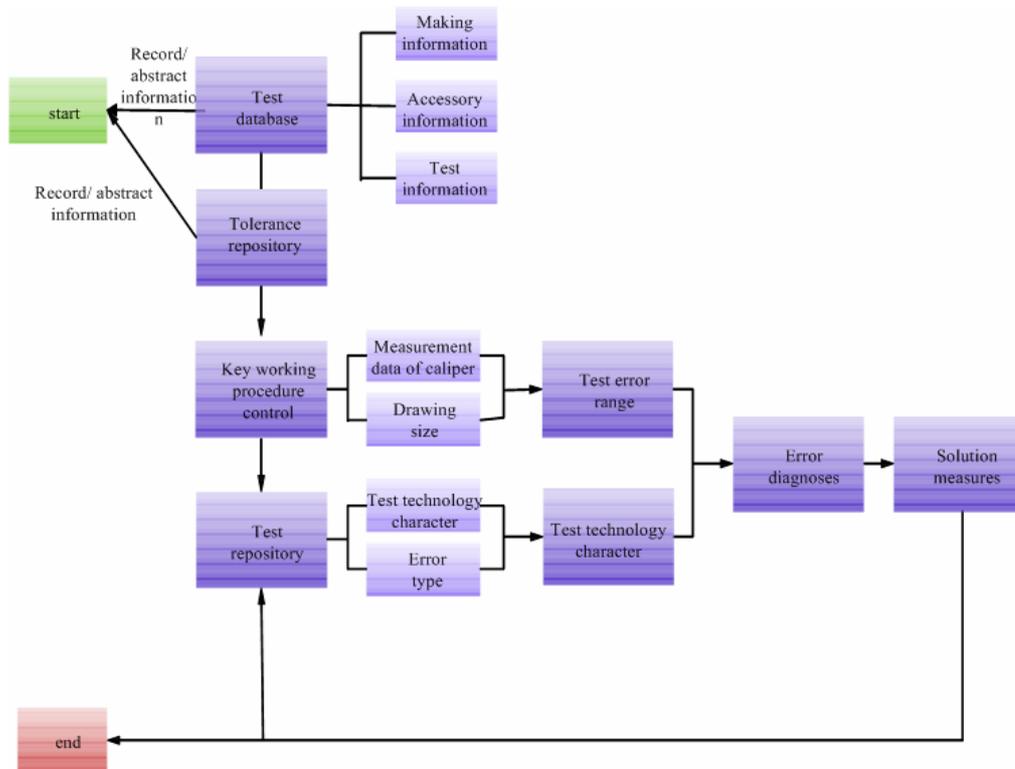


Figure 1. Main Function Design Flow of the Precious Test and Quality Control System

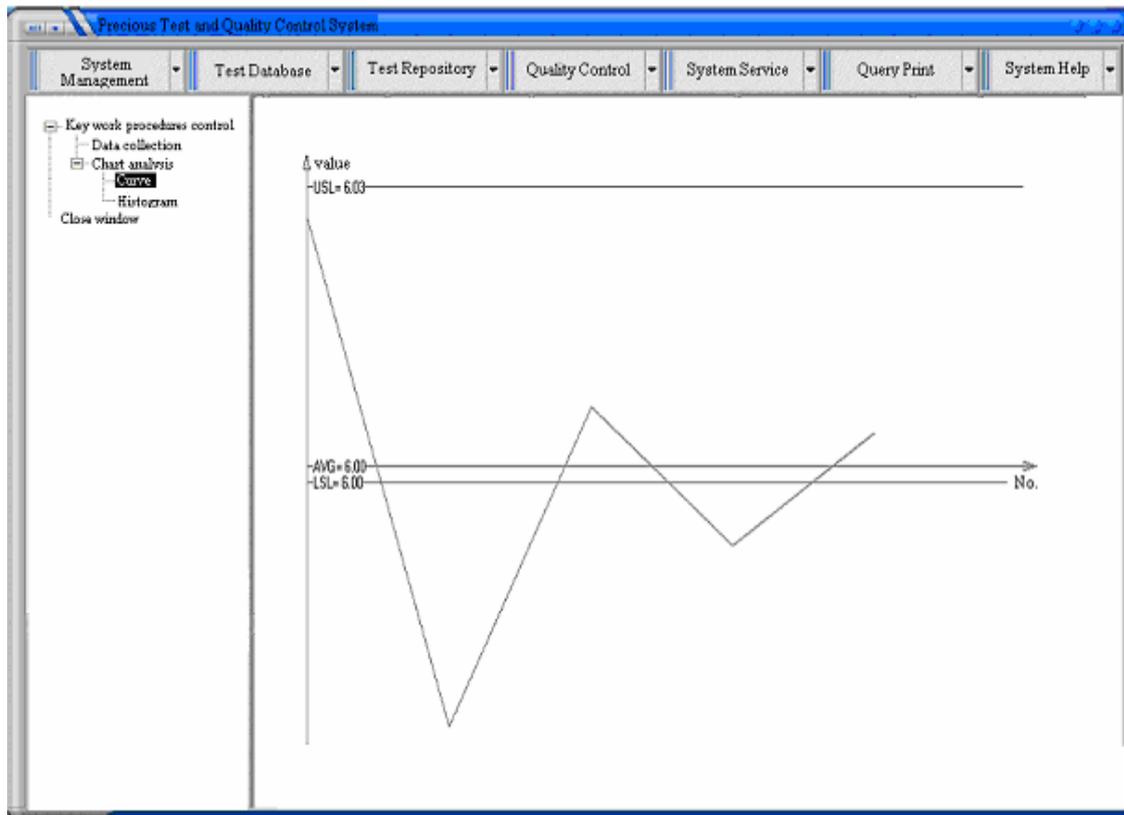


Figure 2. Quality Control Curve

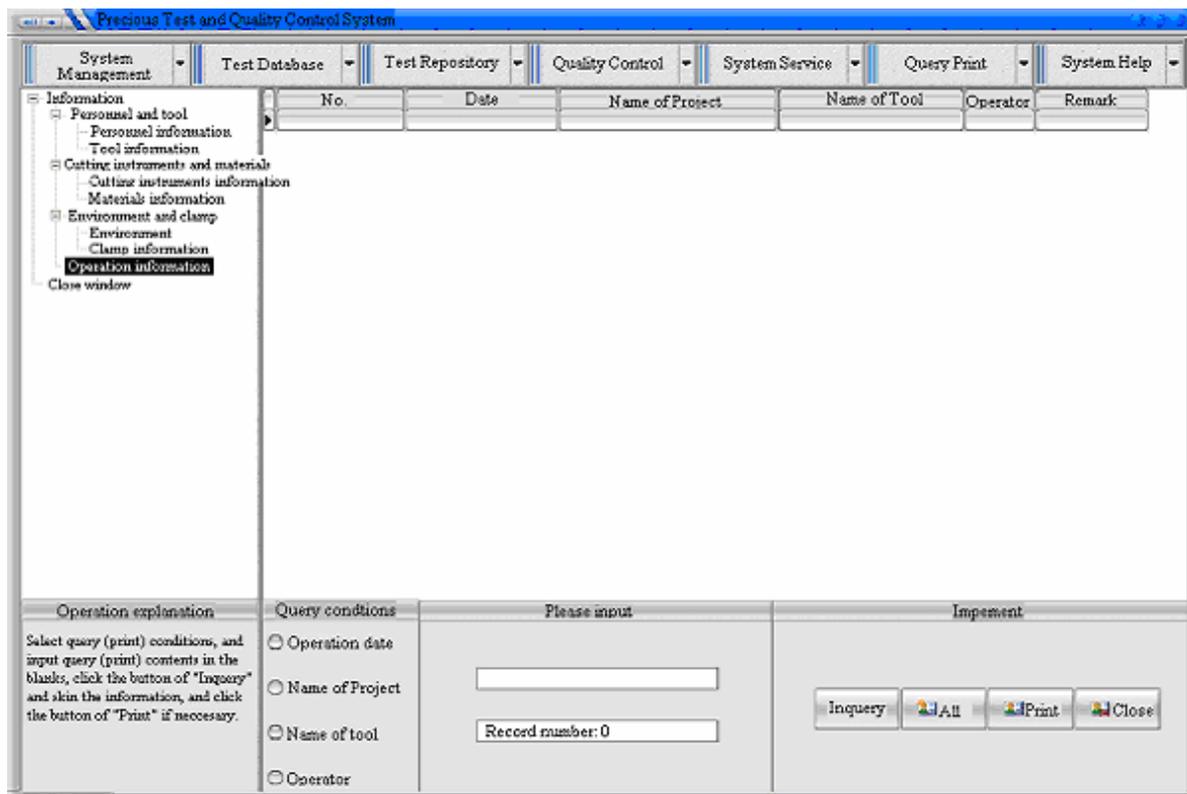


Figure 3. Historical Data Query



A New Vision Inspection Method for Wood Veneer Classification

Mengxin Li

School of Information and Control Engineering

Shenyang Jianzhu University

Hunnan New District, Shenyang 110168, China

Tel: 86-24-2469-0042 E-mail: limengxinf1972@yahoo.com.cn

Chengdong Wu

School of Information Science and Engineering

Northeastern University

Shenyang 110004, China

Tel: 86-24-2424-3208 E-mail: wuchengdong@mail.neu.edu.cn

Abstract

A vision-based inspection method based on rough set theory, fuzzy set and BP algorithm is presented. The rough set method is used to remove redundant features for its data analysis and procession ability. The reduced data is fuzzified to represent the feature data in a more suitable form as input data of a BP network classifier. By the experimental research, the hybrid method shows good classification accuracy and short running time, which are better than the results using BP network and neural network with fuzzy input.

Keywords: Vision-Based Inspection, Classification, Fuzzy input, Rough Set, Neural Network

1. Introduction

In the wood factory, the veneer sheets are placed on a conveyor which runs at a speed of 2.2m/s and the sheets appear at one to tow second intervals for human inspection. This task is extremely stressful and a little disturbance or loss of attention will result in a misclassification. Huber et al (1985, p.79) made a series of experiments and found an accuracy of 68% with human inspection of boards. Similar experiments carried out by Polzleitner and Schwingshagl (1992, p.2283) indicated an accuracy of 55%. It is imperative to develop an automatic visual inspection system to relieve the human inspector and improve the classification accuracy, thus improving the productivity and profitability of the plywood factory.

Various pattern recognition methods have been researched during the last four decades, such as statistical decision theory, syntax and neural networks. Hybrid methods are of great interest due to their proven adaptability and advantages. This research presents a method using fuzzy rough neural network method. The rough set method is proposed to remove redundant features for its data analysis and processing. The reduced data is fuzzified to represent the feature data in a more suitable form for input to a BP network classifier. The BP neural classifier is considered the most popular, effective and easy-to-learn model for complex, multi-layered network.

2. Rough set for feature selection of wood veneer defect

2.1 Basic concepts

The following terms or concepts are introduced in order to facilitate and understand the proposed algorithm.

Information System: It is assumed that the given set of training samples represents the knowledge about the domain. In the approach described here, the training set is described by a classification system. The objects in a universe U are described by a set of attribute values.

Formally, an information system S is a quadruple $\langle U, A, V, f \rangle$, where $U = \{x_1, x_2, \dots, x_N\}$ is a finite set of objects, which in this case are states of the environment; A is a finite set of attributes and the attributes in A are further classified into two disjoint subsets, condition attributes C and decision attributes D , such that $A = C \cup D$ and $C \cap D = \emptyset$; $V = \bigcup_{a \in A} V_a$ is a set of

attribute values and V_a is the domain of attribute a (the set of values of attribute a); and $f: U \times A \rightarrow V$ is an information function which assigns particular values from domains of attributes to objects such that $f(x_i, a) \in V_a$ for all $x_i \in U$ and $a \in A$.

Decision Table: An information system can be designed as a decision table if the attribute set is divided into two disjoint sets - condition attribute set C and decision attribute set D , and $C, D \subset A$.

Indiscernibility Relation: For every set of attributes $B \subset A$, an indiscernibility relation $IND(B)$ is defined in the following way: two objects, x_i and x_j , are indiscernible by the set of attributes B in A , if $b(x_i) = b(x_j)$ for every $b \in B$. The equivalence class of $IND(B)$ is called the elementary set in B because it represents the smallest discernible group of objects. For any element x of U , the equivalence class of x_i in relation $IND(B)$ is represented as $[x_i]_{IND(B)}$. The equivalent class of relation $IND(C)$ and relation $IND(D)$ are called condition class and decision class respectively for condition attribute set C and decision attribute set D .

Lower and Upper Approximation: Let X denote a subset of elements of the universe U ($X \subset U$). The lower approximation of X in B ($B \subset A$), denoted as \underline{BX} , is defined as the union of all these elementary sets which are contained in X . More formally:

$$\underline{BX} = \{x_i \in U \mid [x_i]_{IND(B)} \subset X \neq \emptyset\} \quad (1)$$

The above statement is to be read as: the lower approximation of the set X is a set of objects x_i , which belongs to the elementary sets contained in X . The upper approximation of the set X , denoted as BX , is the union of these elementary sets, which have a non-empty intersection with X :

$$BX = \{x_i \in U \mid [x_i]_{IND(B)} \cap X \neq \emptyset\} \quad (2)$$

For any object x_i of the lower approximation of X , it is certain that it belongs to X . For any object x_i of the upper approximation of X , it may belong to X .

$BNX = BX - \underline{BX}$ is called a boundary of X in U .

CORE: The set of all attributes indispensable in C is denoted by $CORE(C)$.

$$CORE(C) = \bigcap RED(C) \quad (3)$$

where $RED(C)$ is the set of all reducts of C .

Accuracy of Approximation: An accuracy measure of the set X in $B \subset A$ is defined as

$$\mu_B(X) = \text{card}(\underline{BX}) / \text{card}(BX) \quad (4)$$

The cardinality of a set is the number of objects contained in the lower (upper) approximation of the set X , $0 \leq \mu_B(X) \leq 1$.

- A rough sets feature selection method

For an information system for wood veneer classification, there are 17 condition attributes including mean grey level, mode grey level, median grey level, standard deviation, skewness, kurtosis, lower number of pixels, higher number of pixels, lower grey level, higher grey level, dark grey level, bright grey level, number of edge pixels (threshold= μ), number of edge pixels (threshold= $\mu-2\delta$), number of edge pixel for feature 14, number of edge pixels (threshold= $\mu+2\delta$), number of edge pixel for feature 16, and 1 decision attribute expressed by 1-13 that means 13 defects including holes, pin knots, rotten knots, roughness, splits, strips, discoloration, coloured strips, barks, worms holes, curly grain, clear wood and sound knots. The task is to find out the optimal attributes and acquire decision rules.

- Data discretization

It is necessary to process the attribute values with a discretization algorithm to express and simplify the decision table. Clustering is a useful tool for analysing the structure of attribute spaces, and deducing similarity and dissimilarity among the observations. In terms of its high dimensionality, discovering clusters of arbitrary shapes, and dealing with different types of attributes, a hierarchical clustering method is presented (Li et al, 2004, p.3650) and adopted for data discretization before attribute reduction. This method has advantages such as embedded flexibility regarding the level of granularity and ease of handling any forms of similarity or distances. It can be divided into agglomerative and divisive algorithms. The agglomerative algorithm usually produces a sequence of clustering schemes of a decreasing number of clusters at each step, which results from merging the two closest clusters. The agglomeration schedule can be visualised by a dendrogram that shows which samples are combined. Nevertheless, it is not certain how many clusters are in the data for a group can be merged into different clusters. So the number of clusters further comes from the idea about statistics such as PST2, PSF, CCC which can judge how many classes should be suitable. Chosen clusters should make the number of clusters as small as it could be. All the optimal values constitute the information system.

The 17 defect features of wood veneers, used as condition attributes $C = \{X_1, X_2, \dots, X_{17}\}$, are discretized using hierarchical clustering and the values of decision attributes are expressed by 1-13.

● Attribute reduction

The aim of attribute reduction is to find a minimal subset of related attributes that preserves the classification accuracy of the original attributes of the decision table. It is therefore necessary to identify important attributes. There are many reducts, but in most cases it is not necessary to find all the reducts. The reduct with the least number of combinations of its attributes is selected [4]. In this research, 17 condition attributes are reduced according to the consistency principle in the following steps.

Step 1: The repetitive samples in the decision table are merged.

Step 2: The data in the decision table is further processed through attribute reduction based on the consistency principle. Important attribute sets thus remain.

The principle of rough sets for reducing redundant attributes can be expressed as the following:

Supposing $C=\{X_1, X_2, \dots, X_n\}$ is an attribute set, if $POS_C(D) = POS_{(C-\{X_i\})}(D)$, then X_i in C is omissible or superfluous; otherwise attribute X_i in C is indispensable.

This has been implemented, as shown in Table. 1. The reduct of condition attributes determines whether there are different decision values when the attribute values are the same in the decision table. Based on consistency principle, if an attribute set is removed and harmony still remains unchanged, the attribute set is removable. Among the 17 attributes of wood veneers, {c}, {g}, {i}, {l}, {m}, {q} are omissible after the attribute reduction, and the remaining attributes are sufficient. Mode grey level {c} and Kutosis {g} are insensitive to defining the changes of pixels. Because only splits and holes belong to white defects among the 13 defects, {i}, {m} and {q} are considered omissible. Dark grey level {l} contributes little to classification decision. After the attribute reduction, the optimum attributes are combined to obtain the decision rules for classification.

3. A NEURAL NETWORK WITH FUZZY INPUT FOR WOOD VENEER INSPECTION

3.1 Fuzzifier

A fuzzy set (Zadeg,1965, p.338) can be represented as membership function μ_A that associates with each element x of the universe of disclosure X , a number $\mu_A(x)$, i.e. membership grade, in the interval [0, 1]. In particular, $\mu_A: A \rightarrow [0, 1]$, where set A can also be treated as a subset of X . The main function of the fuzzifier maps a crisp input point $x \in X$ into a fuzzified value in $A \in U$ (the universe). There are two types of fuzzifier:

- Singleton fuzzifier: fuzzy set A with support x_i , where $\mu_A(x_i)=1$ for $x= x_i$ and $\mu_A(x_i)=0$ for $x \neq x_i$, for which the input measurement x is perfect crisp, $i=1,2,\dots,n$.
- Non-singleton fuzzifier: $\mu_A(x_i)$ reaches maximum value 1 at $x= x_i$ and decreases from 1 to 0 while moving away from $x= x_i$.

The determination of the fuzzy membership function is the most important issue in applying a fuzzy approach. No common approach is available for determining such a function. In some cases, the fuzzy membership function is attained subjectively as a model.

3.2 A NEURAL NETWORK ALGORITHM WITH FUZZY INPUT

Taking advantage of data processing using fuzzy sets, fuzzy theory is considered to combine with a BP algorithm. Suppose that we have a recognition problem of m classes, which has m nodes in the output layer. The weighted distance is first defined between the i th class and the j th class:

$$z_{ji} = \sqrt{\sum_{k=1}^n \left(\frac{x_{jk} - m_{ik}}{\sigma_{ik}} \right)^2}, \quad i = 1, 2, \dots, m \tag{5}$$

where x_{jk} is the k th vector of the j th pattern vector; $1/\sigma_{ik}$ is a normalisation factor, which results in small class weights for high variance, and m_i and σ_i are the mean value and standard dispersion respectively. The ambiguity of the j th pattern belonging to the i th class is then defined as follows:

$$\mu_{ij} = \frac{1}{1 + (Z_{ji} / \alpha)^\beta}, \quad i = 1, 2, \dots, m \tag{6}$$

where α and β are parameters used for controlling the fuzzy degree and $\alpha, \beta > 0$. According to equation 6.5, there is a low attributive degree if there is a large distance between a pattern and a class. If all elements satisfy $\mu_{ij} \neq 0$, a high fuzziness exists. If only one element satisfies $\mu_{ij} \neq 0$, no fuzziness exists. Under the condition of a high fuzziness, there is a need to modify the ambiguity factor in order to enlarge the difference of membership function.

$$\mu_{ij,INT} = \begin{cases} 2(\mu_{ij})^2 & 0 \leq \mu_{ij} \leq 0.5 \\ 1 - 2(1 - \mu_{ij})^2 & \text{others} \end{cases} \tag{7}$$

For the j th pattern, x_j , the i th subvector of the desired output, y_j is defined as:

$$y_{ij,INT} = \begin{cases} \mu_{ij,INT} & \text{high fuzziness} \\ \mu_{ij} & \text{others} \end{cases} \quad (8)$$

where $0 \leq y_{ij} \leq 1$. All the input and desired output vectors (x_j, y_j) can be used for training with the improved BP neural network.

4. A CLASSIFIER USING ROUGH SETS BASED NEURAL NETWORK WITH FUZZY INPUT

The input data is dealt with using rough sets and fuzzy sets for their powerful function of disposing infinite and incomplete information. This will decrease the number of input nodes and complexity of neural network. A classifier using rough sets based neural network with fuzzy input is proposed (Figure 1).

The system includes input layer, data reduction layer, neural network layer with fuzzy input and output layer.

- Input data

This raw data is preprocessed to suppress noise and normalise the input, and the processed data is input to the feature reduction layer. The normalisation formula is

$$Z = \frac{x - \mu}{3\delta} \quad (9)$$

where x is raw data, μ is the sample mean, δ is the sample deviation and Z is the normalised data which is restricted to $[-1, 1]$ [Kjell et al., 1995,p.1222].

- Feature reduction layer

The normalised data is processed with a hierarchical clustering discretization method. The data reduction is performed using rough sets, and important features for classification remain in the data.

- Neural network layer with fuzzy input

This layer includes a fuzzifier for the reduced data and the improved BP neural network. The crisp input data is converted into fuzzy data through fuzzification, and the fuzzified data is input to the improved BP neural network for defect classification. It is expected that a high classification accuracy and rapid running speed will be achieved through fast data processing.

- Output layer

Output results are obtained from this layer. The maximum coding method is used for the classification decision, which sets the highest output value to 1 and the others to 0. In other words, the defect class corresponds to the output neuron with the largest value.

Sample data from wood veneer defects is used for testing, of which 80% is used as training data and 20% as testing data. The experiments are carried out in the 3 groups of data. The results of feature reduction are used by the neural network classifier with fuzzy input.

The fuzzified data is then input to the BP network for training and testing. It should be noted it is not necessary to change the learning rate by large amounts. Especially, the closer the training to convergence is, the smaller the learning rate change should be. Therefore a learning coefficient that is very close to 1, but less than 1, 0.999, is chosen for adjusting the weight slowly. The learning rate η is expressed below:

$$\eta(k+1) = 0.999\eta(k) \quad (10)$$

where k is the number of epochs and maximum of k is 6000, and initial learning rate is set to 0.5. The experimental result is shown in Figure 2, which takes Group 2 of sample data as an example.

For comparison, the BP method with fuzzy input and BP network are adopted and trained with the original sample data, and comparison results are provided in Table 2.

The results are further improved with each method proposed in that the average accuracy of classification increases from 86.01, to 95.88% and the average running time drops from 324.7, 14.79s to 12.67s. Obviously, the rough fuzzy neural network method has the best classification performance.

5. CONCLUSIONS

To summarise, the classifier has several obvious characteristics which listed below:

- The neural network acquires knowledge through the process of training. The BP algorithm is a mature and effective one.
- An important idea of rough sets is to remove redundant attributes where there is a large amount of sample data. Elimination of redundant attributes can help identify strong, non-redundant classification rules. The rough set method for feature selection is particularly useful for dealing with imprecise, inconsistent and overlapping information.
- Fuzzy set theory is an expressive tool for coding. It entails a low computational load compared to neural networks and makes decisions in a simple and robust way. Data fuzzification applied to the neural network classifier can also reduce the system running time while improving the classification accuracy.
- The hybrid method incorporating a neural network, fuzzy sets and rough sets can achieve a high classification accuracy and rapid speed taking advantage of the complementary characteristics of these techniques even in situations where data is imprecise, noisy, inconsistent and huge. There is no need for extra hardware to deal with uncertainties.

This paper has presented the rough sets based neural network with fuzzy input for pattern recognition as a more effective hybrid approach. In the feature selection process, redundant features are reduced significantly without losing essential information using rough set. In the feature classification, data fuzzification is used to deal with imprecise data and shorten the running time, and the improved BP neural network tackles the local minimum problem to achieve a good accuracy. The hybrid method has taken the advantages of all the techniques incorporated.

Experimental results have shown that the rough fuzzy neural network classifier has a high classification accuracy of wood veneer defects and a short running time. The method is considered general and can be applied to inspection of other products such as ceramic tiles.

References

Huber H.A. , Mcmilin C.W. & Mckinney J.P., (1985), Lumber defect detection abilities of furniture rough mill employees, *Forest Products Journal*, 35 (11/12), pp.79-82.

Polzleitner W. & Schwingshagl G., (1992), Real-time surface grading of profiled wooden boards, *Industrial Metrology*, pp.2283-298.

Li M.X., Wu C.D. & Y.Yue, (2004), A hierarchical clustering method for attribute discretization in rough set theory. *Proceedings of International Conference on Machine Learning and Cybernetics (ICMLC)*, Shanghai, China, 26-29 August, pp.3650-3654.

Zadeg L.A., (1965), Fuzzy sets, *Information and Control*, 8, pp.338-352.

Kjell B., Woods W.A. & O.Freider, (1995), Information retrieval using letter tuples with neural network and nearest neighbour classifiers, *IEEE International Conference on Systems, Man and Cybernetics*, Vancouver, Canada, pp.1222-1226.

Table 1. The representation of discretized knowledge

U	b	c	d	e	f	g	h	...	r	Y
1	2	1	2	1	1	1	3	...	1	1
2	1	1	2	1	2	2	1	...	1	1
3	2	1	1	1	2	2	2	...	1	1
4	1	1	2	1	1	1	1	...	1	1
5	2	1	2	1	2	2	3	...	2	1
...
232	2	2	1	1	1	1	2	...	1	13

Table 2. Comparison of overall accuracy and running time

	Accuracy	Average times	Average epochs
BP network	86.01%	324.7s	44323
A fuzzy rough neural network	95.88%	12.67s	4322

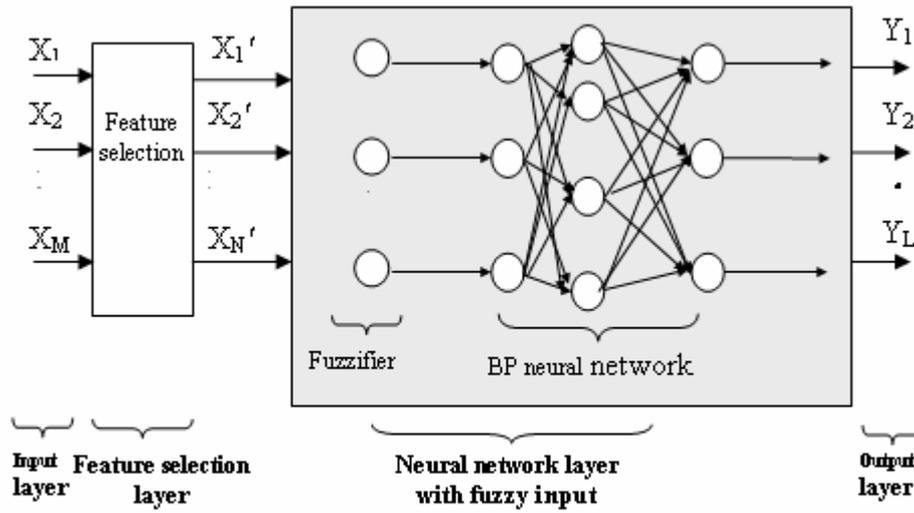


Figure 1. A classifier using rough sets based neural network with fuzzy input

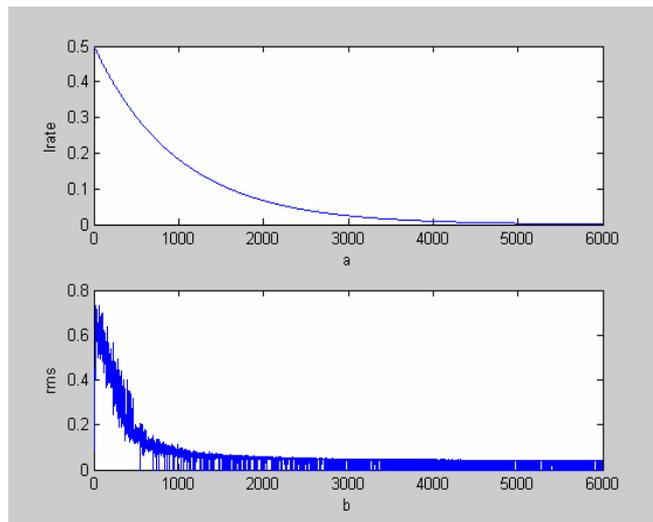


Figure 2. Training results of the RNNFI classifier with Group 2



Summarization of Program Development in Multi-core Environment

Hongyu Li

College of Computer and Automatization

Tianjin Polytechnic University

Tianjin 300160, China

E-mail: milanbaggio18@sina.com

Abstract

It is not transparent from single core time to multi-core time for programmer not like the enhancement of clock frequency of processor, and if we don't design the programs what we compile aiming at multi-core characters, we can not obtain the enhancement of performance from multi cores. In this alternative time of old and new concepts, we should fully use former development experiences for references.

Keywords: Multi-core programming, Multithreading, Parallel computer, OpenMP

1. Introduction

The computer system adopting the multi-core processor is a parallel computer system with multi processors in the chip. This sort of parallel computer system only occurred in the super computer center and the cluster computer system in the past. At now, with that the processor chip has entered into the multi-core time, millions computer systems in our homes have been the parallel computer system with multi processors. This sort of parallel computer system is the system with parallel course and parallel thread, and many courses and threads can really run at parallel, and it is not like the single processor system that the course and thread can only alternatively run in CPU. But in the multi-core system, the running speed of the former single program can not be enhanced. To enhance the running performance of single program, we need redesign the former program, decompose the single computation task into many parallel subtasks, and let these subtasks respectively run in different processor cores. Thus, we need adopt different programming methods. With the advent of multi-core time, the parallel programming will be the necessary knowledge and skill in the software industry.

2. The advent of multi-core time revolution

In 2001, IBM pushed the Power4 processor based on dual-core, after that, both Sun and HP successively pushed the UltraSPARC 4 processor and PA-RISC8800 processor based on dual-core frame. But these RISC processors facing high-end application were too high to be popular, and they didn't obtain attentions by people. Until the second quarter in 2005, Intel issued the desktop dual-core processor based on X86, so the multi-core processor begun to go into common families.

Today, the multi-core processor has occupied more and more market shares, and the programming personnel also must face the collision brought by the revolution of multi-core processor. The multi-core programming is not only the opportunity but also the challenge, and it is the urgent task for us that how to grasp the direction and advance with time in the industry revolution. Because if we don't design the programs what we compile aiming at multi-core characters, we can not obtain the enhancement of performance from multi cores. In this alternative time of old and new concepts, what we can choose and whether we can use former development experiences for references? We should use former experiences for references and actively study the skill of parallel programming. The multi-core, especially the dual-core is very similar with the frame of dual-route SMP (symmetric multiprocessing).

From Figure 1, we can see that though there are differences in the dual-core technology between Intel and AMD, but the so-called dual-core processor is to centralize two computation cores in one processor. That is quite similar with the dual-route SMP system which centralizes two processors on one mainboard, and the difference is only in that the data exchange between two computation cores of the dual-core computation system doesn't need the front system bus (FSB), but the two processors of the dual-route processor system is to exchange data through FSB, which is one tiny detail that we should notice when we compile programs.

Similar with SMP programming, we must use the form of multithread or multi-course to compile the application program to obtain the performance improvement brought by multi-core processor aiming at the multi-core processor programming. So most experiences what we accumulate in the SMP parallel programming can be applied in the multi-core programming.

3. The revolution of programming

The advent of multi-core time brings large collision to our programming thinking. In order to fully utilize the multi-core property, we must learn to design the program with the partial thinking and compile the program with form of multi-course or multithread. Whether we should use the multi-course form or the multithread form to compile the program is one of confused problems for programmers, the author thought we should decide it according to concrete application, but in usual conditions, the multi-core programming using multithread has more advantages than multi-course.

Except for programming form, the motivation that we use multithread programming has been changed. In the past, for the Windows programmer, one of main reasons using multithread was to enhance the user experience, for example, to enhancing the response speeds of UI, I/O or network in the long-time computation. But in the multi-core time, the reason that we compile the application program is to fully utilize many computation cores, reduce the computation time or compute more tasks in the same period. For example, in the game programming, the form of multithread can detract the computation of collision inspection into many CPU cores to largely reduce the computation time, and it also can use the multi-core to make more meticulous inspection computation and can simulate more real collision.

In the multi-core time, we should be more careful to choose the programming language. Relative to compiled languages such as C/C++/Fortran, the script languages such as C#/java/Python may be the better choice. The reason is that the script language is super, and it usually offer the original support to the multithread, such as the System Threading.Thread of C#, the java.lang.Thread of java, and the Threading.Thread of Python. Comparably, the compiled language usually supports the multithread through the relative base of the platform, such as Win32 SDK and POSIX threads. There is not uniform standard, so more details are should consider when we compile multithread program through C/C++, which increase the project costs. Now, through the users of C/C++ are numerous, but the script language in the multi-core time will be more popular, because the script language usually has not ISO standard, so it can be changes conveniently, and the explainer and compiler will be occur quickly aiming at multi-core. However, the script languages such as PHP, Ruby and Lua are difficult to obtain programmers' favors, because they don't offer the interior core thread support, and their multithreads are for users, and they even can not support the thread, so the multithread program complied by them can not still fully utilize the multi-core advantages.

Through C/C++ losses part advantages in the multithread programming because of the lack of support of language, but because most present main operation systems offer API to create the thread by the form of C language interface, so C/C++ has quite abundant program base, which can compensate the deficiency of language to some extent. We can not only use the Win32 SDK, but can use POSIX threads, MFC and boost.thread to compile the multithread program. Through these bases offer certain encapsulations and reduce programmers' multithread burdens to certain extent, but for the multi-core programmers whose aims are to enhance the performance of condensed program, these methods are still very complex, because the use of these bases almost needs increase double key codes, and the corresponding transfer and testing costs will be increased largely. The better choice is to use the compiler such as OpenMP to support and strengthen the fundamental base of multithread. The OpenMP appoints the parallel code segment through using #pragma compiler order, and it changes the program little, and it can appoint the serial edition to compile for debugging conveniently, and it can also coexist with the compilers that can not support OpenMP.

It is obvious that through the script language offers the original support of the multithread programming on the language layer, but it doesn't lead the C/C++ far. The essential reason is the base of the script language, the fundamental base and CRT/STL of data structure and algorithm. The fundamental base of C/C++ is designed and developed by the serial form. To modify the fundamental base aiming at the multi-core programming, which is the extremely urgency faced by all compiled languages, is the war of life and death to separate the leading advantage of two camps, but the property is centralized in the script languages such as C#/java/Python in certain company or organization, which will decide the key battle. That is the reason why we commend choosing the script language to compile program.

4. Multi-core programming

With the advance of time, we will finally face the multi-core system to design program. For the multi-core programming, the author thought it is equal to the parallel programming sharing memory, and the multi-core programming can use former experiences of past parallel programming for references, such as the partial design thinking, the parallel design method and multifold parallel support methods.

First, we discuss the partial design thinking. Because the thread is the minimum unit to distribute CPU resource for the operation system, so if we want to design the multi-core parallel program, we must form the designing thinking to

divide the program into parts. The Mr. Hua Luogeng's overall planning method can be referred to discuss how to divide the program into parts.

For example, we want to steep the tea to drink. The situation is that there is not boiled water, the water jug, the teakettle and the teacups should be washed, the fire burns, and there is tea leaf. How to do?

Method A: wash the water jug, fill the cool water, put the teakettle on the fire, when we wait the boiled water, we can wash teakettle and teacups, take the tea leaf, and when the water is boiled, and we can steep the tea to drink.

Method B: first do some preparation works, wash water jug, teakettle and teacups, and take the tea leaf, and when every thing is ready, we fill and boil the water, and when the water is boiled, we steep the tea to drink.

Method C: clean the water jug, fill the cool water, put the water jug on the fire, wait that the water is boiled, and when the water is boiled, we find the tea leaf in hurry, wash the teakettle and teacups, and steep the tea to drink.

Which method can save time? Of course, the first method is the best one, the others waste time. Suppose Mr. Hua Luogeng has two robots to steep tea, the best method is to distribute the work according to the "Method A", i.e. robot A go to boil the water, and robot B wash the tea set, and when the water is boiled, the tea can be steeped. Unconsciously, we use the partial thinking, i.e. distributing irrelevant works to different processor to implement.

The above partial thinking is simple and direct, but for complex task, it is not so easy to find the partial plan, so we need the method of parallel design to guide us. Through tens years' researches of parallel program, former researchers have summarized many effective methods of parallel design, here, we introduce one classic method, i.e. the data correlation graph. Still taking the classic steeping tea as the example, we can draw the following data correlation graph seen in Figure 2.

From Figure 2, we can see that the data correlation graph is a directional graph, and every acme in the figure represents a task which should be completed, and the arrowhead denotes the task pointed by the arrowhead depends on the task educed by the arrowhead, and if there is not the route from one task to another task, so two tasks are not irrelevant, so we can implement parallel treatment. If Mr. Hua want to steep the tea to drink himself, so the smaller oblong in the data correlation graph of Method A is can be parallel, and if Mr. hua has two robots to help him steep tea and there are two water faucets at least for robots, so the bigger oblong area can be parallel and obtain higher efficiency. So the resources that can be reasonably used are more, the parallel acceleration ratio is higher.

Now that the partial thinking and method of parallel program design can be referred, we only lack in how to develop the parallel program. There are three popular ideas.

(1) Extended compiler. The parallel compiler can find and express the parallel character in the present serial language program, for example, the Intel C++ Compiler has the functions to automatic parallel circulation and vector data operation. This method leaving parallel works to compiler reduces the costs to compile the program, but because of the complex combination of circulation and branch languages, the compiler can not identify quite more parallel codes and falsely compile serial edition.

(2) Extended serial compiled language. It is the most popular method, and it can denote lower layer language to obtain parallel program through increasing function transfer or compiling orders. The users can create and end the parallel course or thread, and offer function of synchronization and communication. The outstanding bases include MPI and OpenMP, and in the explained script camp, the ParallelPython also win partial market.

(3) Creating a parallel language. Through it is a crazy idea, but in fact, in tens of years, there are many people to do the work all along, for example, HPF (High Performance Fortran) is the extension of Fortran90, and it supports the parallel data program through various modes.

In the future multi-core programming, we will find that there is more knowledge in the computation domain to deserve us to study, and there is more extensive space to realize out ideas.

References

- Akett Roberts, interpreted by Li, Baofeng et al. (2007). *Multi-core Programming Technology*. Beijing: Electronic Industry Press. Mar. 2007.
- Chen, Guoliang. (2002). *Parallel Computer System Structure*. Beijing: Higher Education Press. Sep. 2002.
- David R. Butenhof, interpreted by Yulei et al. (2003). *POSIX Multithread Programming*. Beijing: China Power Press. Apr. 2003.
- Huangkai. (2000). *Extended Parallel Computation Technology, Structure and Programming*. Beijing: Machine Industry Press. May 2002.
- Jim Beveridge, interpreted by Houjie et al. (2002). *Win32 Multithread Programming*. Wuhan: Huazhong University of Science and Technology Press. Jan. 2002.

Kalla R. (2004). IBM Power 5 Chip: A Dual-core Multithreaded Processor. *IEEE Micro*.

Kai Hwang, interpreted by Zheng, Weimin et al. (1995). *Higher Computer System Structure*. Beijing: Tsinghua University Press.

Micheal J. Quinn, interpreted by Chen, Wenguang et al. (2004). *MPI and OpenMP Parallel Programming for C Language*. Beijing: Tsinghua University Press. Oct. 2004.

Zheng, Weimin. (1998). *Computer System Structure*. Beijing: Tsinghua University Press.

Table 1. Comparison of various languages to support multithread

	Compiled languages such as C/C++	Scripts such as C#/java/Python	Scripts such as PHP/Ruby/Lua
Whether the language support the multithread	No	Yes	No
Whether the base support the multithread	Yes	Yes	No
Whether supporting interior core thread	Yes	Yes	No
Whether support user thread	Can be simulated	Can be simulated	Yes
The complexity of thread programming	Commonly/Easy	Easy	N/A
Recommendation degree	★★★☆☆	★★★★★	★☆☆

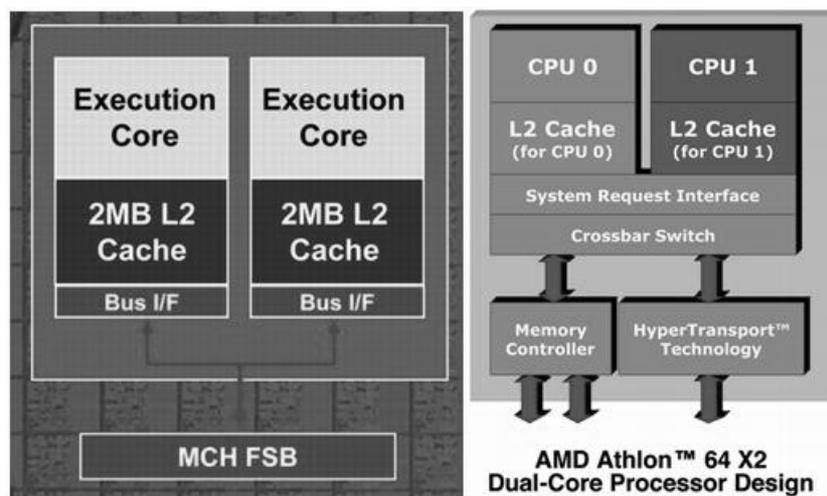


Figure 1. Dual-core CPU Structures of Intel and AMD

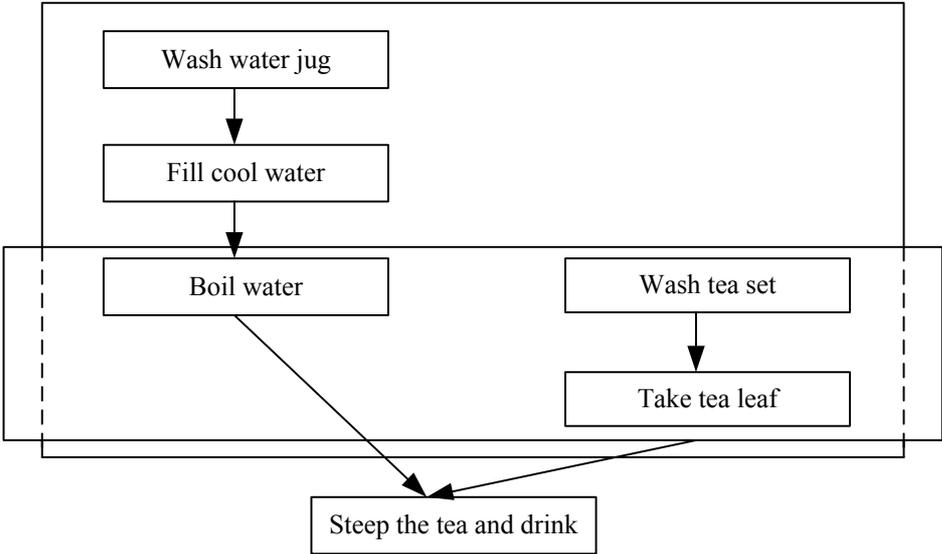


Figure 2. Data Correlation of Method A in “Overall Planning Method”

A journal archived in Library and Archives Canada
A journal indexed in CANADIANA (The National Bibliography)
A journal indexed in AMICUS
A leading journal in computer and information science

Computer and Information Science

Quarterly

Publisher Canadian Center of Science and Education

Address 4915 Bathurst St. Unit # 209-309, Toronto, ON. M2R 1X9

Telephone 1-416-208-4027

Fax 1-416-208-4028

E-mail CIS@ccsenet.org

Website www.ccsenet.org

Printer William Printing Inc.

Price CAD.\$ 20.00

