SYSTEM DESIGN AND BEHAVIOR OF PATTERN RECOGNITION OF CONTROL CHART USING NEURAL NETWORK

Justin Yulius Halim
Department of Industrial Engineering, Dongseo University, South Korea

ABSTRACT

Recognition of unnatural patterns in control charts is important to get more understanding about the process problem. Shewhart control chart can recognize shift pattern only, while the others cannot be detected by this chart. Neural networks were proposed by many researchers to achieve the better recognition of the patterns. A system of neural networks needs to be fitted to this problem by realizing the behaviors of control charts. Each behavior will affect the development of each part of the system. Some problem need to be pointed also when dealing with the patterns of the control chart.

Keywords: Control charts, pattern recognition, neural networks, control chart behaviors

1. INTRODUCTION

1.1 Control Chart and Behavior

Control chart as the tool of product quality control has been widely used in industry, whether the product is tangible or intangible. The usage of control chart is based on one or more parameters of the product. As the example to learn the behavior of control chart, take one most common control chart as study object, the x-bar control chart.

The X-bar control chart is a two dimensional graph with two axis as the representation of the value range of the dimension. The X axis generally represents time or sample number which is ordered by time. The Y axis is assigned for the parameter of the product. This behavior is usually called time series. The plotting of the values will takes unlimited number of plot in application. The X-bar chart usually show a number of the most recent values only. The cropping of the values is called as window. When a new value get into the chart then the oldest value plotted in the chart will be eliminated. The behavior of one in and one out is called moving window, because the window is looked like move forward time by time.

Another behavior of the chart is noise. Actually, noise is not the direct behavior of the chart but it belongs to the parameter. Control chart always face noise with very small value or big value of noise. Big value of noise makes the controlling become hard, because x-bar control chart is used to find any shift in the mean of the parameter value. Then noise could cause a false alarm of the shift.

These behaviors have a big impact in building a proper neural network for the pattern recognition of a control chart.

1.2 Why Use Neural Network

The first question can be retrieved from the use of neural network in control chart pattern recognition is why use neural network?

Neural networks have become a very popular method in many research field such as data mining, neurobiology, image processing, computer science, and physics. The ability of neural network is to tolerate noisy data highly as well as their ability to classify patterns on which they have not been trained (Han and Kamber, 2001).

Linear and nonlinear regression provides a smooth and robust curve fitting to training patterns. Linear and nonlinear regression is trained using a supervised learning procedure. Neural network is a method that considered successful if it can closely approximate the training values and make a good interpolation to the test set. This ability is called generalization (Kung, 1993) which differs to the regression.

The characteristics and advantages of neural networks (Kung, 1993) can be summarized to three points:

- Adaptiveness and self organization
  Neural networks have the adaptive learning types which lead to the flexibility of the usage of neural networks. The rules made by neural networks can be self organized.
- Nonlinear network processing
  The processing of input in neural network is done by nonlinear function which enhances the network's approximation, classification and noise-immunity capabilities.
- Parallel processing
  Neural networks is consists of many neuron which each of them process the inputs in parallel and enhanced by extensive interconnectivity.

The use of neural network has spread to many fields and one of that field is quality control. The number of the application of neural networks to recognize defects in automated manufactured products is large. But the research of the application of neural networks in pattern recognition on control chart still have young age. The research of pattern recognition by neural networks on control charts started from the recognition of a pattern and grew to a number of patterns. This paper start the study from the basic, neural network, on section two with brief review and then continued to the focused application, the pattern recognition on control charts, on section three. The last section discussed more about the usage of neural network in pattern recognition on control chart. There are many factors to be considered in the application and also many point-of-views given by many researchers.

2. NEURAL NETWORK

2.1 The Origin of Neural Network

Neural networks, precisely artificial neural networks, are the study of the work of human brain cells. The most effecting cells of human brain are the neurons which act as the storage of information and have analysis capabilities. The neurons itself are connected each other by connectors named axon. Each neuron received inputs from their neighbors (the other neurons) and sums the inputs and process it with step function. The result of the process is the output of the neuron and sent to next neurons as their input. The Figure 1 below gave better description of the process.

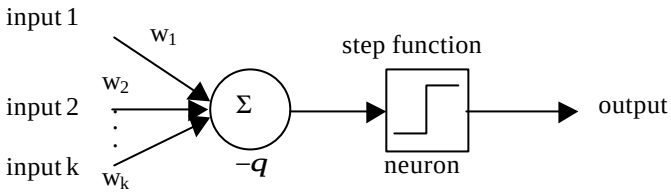

Figure 1. Neural Network

This main concept of neurons became the foundation of artificial neural networks (ANN), even not all of the concepts used at ANN. There are inputs from other neurons, which are product to their weights (w_i) and then subtracted by a threshold (q). This first function in artificial neural networks is called basis function or net function. Neurons have linear type of basis function to gain the information from other neurons. The networks value from basis function will be transformed by a non-linear function. The function type used here is step function. The value of the transformation became the output of the neuron.

2.2 Neural Network Type

The taxonomy of neural network could be started from the learning behavior of the neural networks. It classifies the neural networks to three category (Kung, 1993), Fixed, Unsupervised and Supervised learning. Actually, fixed learning was not learning because the weight (representation of the memory) is fixed. So the learning neural networks are supervised NN and unsupervised NN.

The learning system of supervised neural networks is based on the availability of a teacher. The teacher here is provided to give information about the correction of the weights needed. The correction of the weights is purposed to reduce the error that resulted from the difference between the target value of output and the real value of output. On the other hand, unsupervised neural network do not need teacher to learn. The ability of unsupervised neural networks to self-organize the networks made them a good algorithm to cluster classes.

The learning process time of supervised neural networks have a behavior that is turn to be a problem. As the size of the networks grow, the time needed to learn is grow exponentially. This behavior is called scaling problem. The development of unsupervised neural networks can overcome the scaling problem of supervised neural networks (Haykin, 1994). The learning time needed of unsupervised neural networks is grow linearly to the size of the neural networks.

Since the pattern recognition in control charts problem usually need relatively small size of networks as proved by many researchers, the scope of neural networks in this paper is based on the supervised neural networks only. The architecture of the neural networks used by many researchers is multilayer perceptron with back propagation algorithm. Multilayer perceptron is a neural network which had one input layer, one output layer and one or more hidden layers. Each layer had some units or nodes (in this paper these two term is used interchangeably) and each units is connected to all units at the next layer. An example of multilayer perceptron neural network is described at the Figure 2 below. The example is a neural network with an input layer, one hidden layer

and an output layer. There are two units in the input layer, three units in the hidden layer and one output unit in the output layer.

Back propagation algorithm name came from the fact that the network learn by propagating the error to the previous nodes.

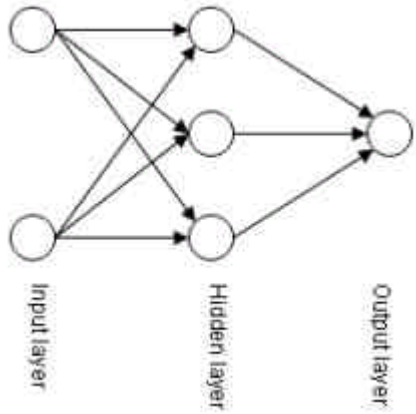

Figure 2. An Example of Multilayer Perceptron Neural Network

3. PATTERN RECOGNITION

3.1 Pattern Recognition in Shewhart Control Chart

Shewhart control chart or x bar and R control chart adopt the behavior of normal distribution to recognize the pattern made by some points. Generally, some points in the sequences of time are plotted. The plots then compared to the limits. When the plots exceed the upper limit of the chart then the process is defined as out of control and so does when the plots is less than the lower limit. This kind of method could be translated to a pattern recognition. The pattern that is recognized in this control chart is the shift pattern. When the process parameter have been shifted (whatever the value is), shewhart control chart will catch this shift. Shewhart control chart is purposely used to define the appearance of shift but this control chart cannot be used to define the size of the shift.

Any other pattern could not be recognized by shewhart control chart but shift pattern is actually the basic of all patterns. The need to control the process sometimes does not need any knowledge about the type of pattern so in some cases shewhart control chart is enough.

The ability of shewhart control chart in detecting shift at the process will be different if the shift is small. Small shift that is not exceed the control limit won't be detected by this control chart, thus another control chart will be needed.

3.2 Pattern Recognition in Cusum Control Chart

Small shift in the process usually can be detected using Cusum control chart. The ability to detect small shift in this control chart is derived from the method that use historical data. Cusum control chart make a cumulative of the difference between a plot and previous plot. The cumulative give the magnitude of the shift so that even a small shift can be detected.

The method that is used by Cusum control chart is a learning method also. The chart learned from past data about the magnitude. The magnitude of past data is recorded at the cumulative value.

3.3 Run Rules

Detection of small shift could be done also by applying run rules to the shewhart control chart. Run rules are applied by dividing the control chart to 3 zones. Zone A is between $2s$ and $3s$, Zone B is between $1s$ and $2s$, and Zone C is between center line and $1s$. There are some rules about the position plot in regard of the zone. For example, if there are 9 plots in zone C or beyond so there's shift in the parameter of the process. The shift is small so that the control limit cannot detect. This run rule's ability show that run rules is good also in detecting a small shift. But another research also shows that beside the detection of small shift, run rules also give contribution to false alarm (Cheng, 1995).

4. THE USAGE OF NEURAL NETWORK

4.1 Designing the Problems Definition for Neural Networks

There are many conditions and parameter should be set when designing the neural networks for control chart. Basically, neural networks need a definition for the input and output. Some researchers in this field have different approach to set the input and output. The inputs of neural networks for pattern recognition have many aspects that should be considered, for example: type of data that should be used, the number of input nodes, the generalization of the inputs and many others that will be explained at the next section. The outputs also have many aspects to be considered. The detection of the patterns must be designed at the output level. The number of the patterns desired to be detected could affecting the number of the output nodes. Another aspects of output is explained at the next section.

4.2 Input Aspects

The input data for neural networks could be divided into two kinds of inputs. First, (Shewhart) built a neural networks model which inputs were secondary data. Raw data are processed to statistical features such as: linear regression coefficients, the standard deviations of the first and second halves of the sequence, and the Box-Pierce Q-statistic. Another type of processing was the definition of Boolean type indicator such as whether the sequence has ten or more points and whether one of the regression coefficients is "low" in magnitude. The statistics features and Boolean type indicator itself built from a sequence of data which limited in a moving window. This approach had some merits besides some demerits also compared to raw data approach. Secondary data can process any size of raw data and provide the same length of inputs for the neural networks. In other words, if the window size was very big then the size of the networks will not be affected and still have the same size as the small window size. This can't be deal to the raw data approach, which the more the window size then the more the neural network size. In addition to the positive correlation of the size, the neural networks size grew exponentially as the window size or input data grew when raw data approach used. As a

result of the fixed size of neural networks, the secondary data approach can process data faster than the raw data approach when the size of the window is big. Another thing happened when the size of the window is relatively small. Secondary data approach obviously required extra computation to produce the secondary data first. Additional consideration should be given when the window size was almost the same or smaller than the number of secondary data used as the inputs. There was another demerit of this approach, which the features derivation was done outside the neural networks. Secondary data approach derived the features of the raw data and use it as inputs for neural networks. The derivation of the features itself was done with a limited knowledge about the right features of the raw data that enabled the networks to give the right output. Thus, the choosing of features used in this approach will determine the quality of the result of the neural networks. The comparison between raw data approach and secondary data approach is summarized in the Table 1.

Table 1. The Comparison Between Raw Data Approach and Secondary Data Approach

	Merit	Demerit
Raw data	Direct input gave a faster computation Brought all of the features to the neural networks to be learned	Size and speed of the neural networks grow exponentially as the window size grow
Secondary data	Window size did not affect the size and speed of the neural networks The computation speed of this method is faster than the raw data when the window size is big	Computation speed became slower than raw data method when the window size equal to input data Result quality highly depend on the features chosen

Raw data approach have been widely used by researchers, notably Cheng (1995) whose used 24 points of raw data as inputs and 1 additional input as a supplementary index to increase the sensitivity of the neural networks, Guh and Hsieh (1999) whose used 56 points of raw data as inputs, Guh et al. (1999) use 24 points of raw data as inputs, Reddy et al. (Reddy et al., 1998) used 16 points of raw data as inputs, Pham and Sagiroglu (2001) used 60 points of raw data as inputs. The number of points used were vary in every researches but all of the researchers agree that the number of points must be considered well regarding that it influence the speed of computation and false alarm. As the number of points decreasing from a certain point, the number of false alarm starts to increase. Spesifically, researchers must do experiments to decide the fittest window size to the neural networks. Another factor that should be decided beside the window size is the type of the problem solving. The neural networks can be designed as a general problem solver or specific problem solver. When the neural networks were designed to be a general problem solver then the raw data should be encoded. The encoding method could be vary such as change the data to standardized value (z) (Guh, 1999) or change it to an integer (Cheng, 1995).

The method to standardize the value is by normal distribution formula.

$$z(t) = \frac{x(t) - m}{s} \quad (1)$$

while $x(t)$ is a raw data at time t, the true mean and standard deviation should be

estimated from the samples at time *t*. *z(t)* is the standardized value of *x(t)* with a zero mean and one standard deviation that follows normal distribution.

$$z(t) = \frac{x(t) - E(m)}{E(s)} \tag{2}$$

The encoding of the data using this method led to a reducement of noise in the raw data and standardized the value to a range [-7, +7]. Actually, the range could be more than it but generally the value will be around that value.

Another method was changing the value to an integer. This method could limit the value to the range only to the pre-specified range, so that any kind of data used, the value will have value in the pre-specified range. For example, Cheng (Cheng, 1995) encode the raw data by the means of standard deviation. The range of the value is divided to zones of width $0.5\ s$. Mathematically, the coding scheme could be expressed as:

$$y_i = \begin{cases} +15 & \text{if} \quad \bar{x}_i > \bar{\bar{x}} + 4s_{\bar{x}} \\ (b_H + b_L)/2 & \text{if} \quad \bar{\bar{x}} + b_L s_{\bar{x}} < \bar{x}_i \le \bar{\bar{x}} + b_H s_{\bar{x}} \\ -(b_H + b_L)/2 & \text{if} \quad \bar{\bar{x}} - b_H s_{\bar{x}} < \bar{x}_i \le \bar{\bar{x}} - b_L s_{\bar{x}} \\ -15 & \text{if} \quad \bar{x}_i < \bar{\bar{x}} - 4s_{\bar{x}} \end{cases}$$

b_L and b_H is the lower limit and upper limit for each zones so that b_H-b_L=*0.5*.

The number of zones and zone rules can be changed to fit the design of the neural networks. Besides two encoding methods above, the combination of both methods can be made purposely. The combination of both methods can derive both methods' advantages.

4.3 Output Aspects

The number of output nodes dealt with the number of patterns need to be detected. There were some pattern that is mentioned in the Western Electric Handbook (1958), they were upward shift, downward shift, upward trend, downward trend, cycle and systematic patterns, and natural pattern (no pattern). The neural networks can be designed to output the prediction of the patterns by one node for each pattern or one node for all patterns. Another kind of neural networks can be made also to predict the parameters of the patterns. All of the parameters needed for the patterns summarized at Table 2.

Table 2. Summary of Parameters for Each Pattern

Patterns	Parameters
Upward and downward shift	Offset from the process mean
Upward and downward trend	Slope
Cycle	Amplitude and period
Systematic	Magnitude of process fluctuation

As noted at Table 2, the parameters of cycle patterns were two parameters. The number of parameters involved was very influencing the number of training set that must be given to the neural networks. This problem made the recognition of cycle pattern became more complicated. Some researchers try to cope this problem by simply limit the training set to one value of the cycle period (Guh, 1999). Of course, the problem will not be handled 100% with this method. A new method for handling the problem is needed.

Besides those parameters, all of the patterns should be appeared at a certain point. That point should be determined also.

4.4 Design of the Neural Networks

Designing the neural networks means that the design of the whole system need to defined also. If the purpose of the system is for pattern type recognition only then the system would be simpler than if the the purpose is pattern parameters prediction. The pattern type recognition system usually composed of one neural networks only, but when the pattern parameter prediction system is taken then the system would be composed of more than one neural networks depends on the number of patterns need to be recognized. Guh et al. (Guh, 1999) proposed a system which firstly used a neural network to recognize the type of pattern, if there was a pattern detected then the data processed again with the next neural network that is proper to the pattern detected to find the parameters of the pattern. Ensemble neural networks can be used also to produce a more reliable result. Ensemble neural networks worked as a group of neural networks which make a vote to decide an answer. The answer was the one which chosen by most of the networks.

There were two activation functions that were widely used in this application, sigmoid function and hyperbolic tangent function. Sigmoid function which became the most popular function at back propagation algorithm was considered not proper for pattern recognition problem (Cheng, 1995). The reason was that process changes of different directions must be detected equally well. Considering that hyperbolic tangent function have a range [-1, 1] then this function can fit the condition needed.

4.5 Training of the Neural Networks

Neural networks need sets of training to produce a good result. The training sets of neural networks have to represent all of the conditions that will occur at the application time. The conditions mainly composed of the type of patterns, the value of parameters, and the first time pattern appear point. One of the problem arised was the value of parameters is a variable value. It would be impossible to build a training set that represent all of the value. In this problem neural networks advantage made a great help. Neural networks in a certain magnitude can interpolate the training set to be applied in the real condition. One consideration must be taken is the value that is used as the representation for training. The more the value the sharpest the prediction but the longer the time needed for training.

Neural networks also need some parameters for training. Mainly the learning rate and momentum give a big effect to the result of the training also. The bigger the learning rate then the bigger also the ability of the neural networks to jump out of the local minimum. But as the learning rate getting bigger, the ability of the neural networks to be converged become less. To avoid the danger of non-converged neural networks and increase the learning rate was by applying momentum (Haykin, 1994). Momentum make the neural networks learn also from the previous data so that it can control the learning process.

5. CONCLUDING REMARKS

It is clear that the need for more advanced control chart start to emerge. One of the advancement of control chart needed is the pattern recognition. Traditional control chart which is used in most industry can only tackle one pattern only, the shift pattern. The problem of the traditional control chart is the learning issue. Shewhart control chart did

not learn from past data except the mean and standard deviation. Cusum control chart learn past data but only the magnitude of the data from mean, so that this chart can tackle small shift better than shewhart control chart.

A new method of learning can be derived from neural networks. It learnt from training sets which is arranged to include all of the possible patterns and parameters. Some concerns should be made in designing the training set. Even neural network able to make an interpolation when some points only used in training sets but one should consider the best points that can represents the patterns and parameters. Another concern is about the cycle pattern. This pattern has two parameters which results in bigger training set. Some researchers try to cope with limiting the cycle period to a point. This method could result in poor recognition when another cycle period appears in the application. A new research would be needed to analyze the problem of the cycle pattern.

Besides the detail of the pattern, the system of pattern recognition should be designed to get the optimal usage of neural networks. There are many forms of system used by researchers depends on the aim of the recognition, whether to recognize patterns only or to recognize the patterns and the parameters. More researches would be needed to find the optimal system of pattern recognitions.

REFERENCES

Han, J. and M. Kamber, 2001. *Data Mining: Concepts and Techniques*, Simon Fraser University, Morgan Kaufmann Publishers.

Kung, S.Y., 1993. *Digital Neural Networks*, Princeton University. Prentice Hall.

Haykin, S., 1994. *Neural Networks, A Comprehensive Foundation*, Macmillan College Publishing Company.

Cheng, C.S., 1995. "A Multi-Layer Neural Network Model for Detecting Changes in the Process Mean", *Computers Ind. Engng*, Vol 28, 51-61.

Shewhart, M., *Interpreting SPC Charts Using Machine Learning And Expert System Techniques*, AFLC Center for Supportability and Technology Insertion (CSTI).

Guh, R-S and Y-C Hsieh, 1999. "A neural network based model for abnormal pattern recognition of control charts", *Computers and Industrial Engineering*, Vol. 36, 97-108.

Pham, D.T. and S. Sagiroglu, 2001. "Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithm", *Int. Journal of Machine Tools and Manufacture*, 41, 419-430.

Reddy, D.C., K. Ghosh, and V.A. Vardhan, 1998. "Identification and Interpretation of Manufacturing Process Patterns through Neural Networks", *Mathl.Comput. Modelling*, Vol. 27, 15-36.

Guh, R.-S. et al, 1999. "On-line control chart pattern detection and discrimination – a neural network approach", *Artificial Intelligence in Engineering*, Vol. 13, 413-425.

Guh, R.-S., J.D.T. Tannock, and C. O'Brien, 1999. "IntelliSPC: a hybrid intelligent tool for on-line economical statistical process control", *Expert Systems with Application* Vol. 17, 195-212.

Kopsco, D., L.Pipino, and W. Rybolt, 1993. *The Application of Artificial Neural Networks To Quality Control Charts*, IEEE.

Western Electric, 1958. *Statistical quality control handbook*, New York: Western Electric Company.