Cluster analysis server system using Self-Organizing Maps (SOM)

H.Sakaki, H.Yoshikawa

JAERI Tokai-mura, Naka-gun, Ibaraki, 319-1195

Abstract

We analyzed the instability behavior of the accelerator by using the Self-Organizing Maps (SOM) which Dr. T.Kohonen was advocating. The SOM is constructed by the neural net-work technique, and the best method on the pattern recognition.

In this report, the measurement data of the L-band LINAC at Osaka University was used for the SOM analysis. The result of this analysis and the different analysis technic, ARmodel, was extremely corresponding. Moreover, the SOM is able to be used from the World Wide Web browser on the Internet.

1 INTRODUCTION

The system analysis using the noise contribution rate has been shown to an effective way to analyze of the accelerator's system[1][2]. But, this is not recommended to beginners. Because, the analysis requires a lot of know-how to do.

Most the the noise analysis are applied by the power spectrum analysis using the Fast Fourier Transform(FFT). This analysis is simple to understand, so that we use one the first step. On the other hand, if only the similarities of the FFT spectrum are compared by visual observation, a wrong result is sometimes analysed by one's subjectivity. In addition, if many disturbance factors are analyzed using the FFT, clarifying the result becomes so difficult. The problem in the FFT analysis is solved by using Self-Organizing Map algorithm (SOM) which Dr. T.Kohonen advocates.

In this report, the data measured at Osaka University is analyzed using the SOM, as one of the applied case to the accelerator. And, to improve convenience, we had this analysis used by the Internet. If you use the Internet World Wide Web browser, you can obtain the result of the SOM from anywhere.

2 SELF-ORGANIZING MAPS AND APPLICATION TO TIME SERIESSIGNAL ANALYSIS

2.1 Self-Organizing Masp (SOM)

In general, the neural net-work is used to do the pattern recognition and system estimation. The SOM is a category which is called "unsupervised neural net-work", and some models need not be used. This technique is a suitable tool for automatically classifying and visualizing large-scale data. These method is one of the Data-mining. Especially, analyzing the similarity grouping is called Cluster analysis.

In this paper, the cluster analysis is done by using SOM algorithm which Dr. T.Kohonen advocates[3].

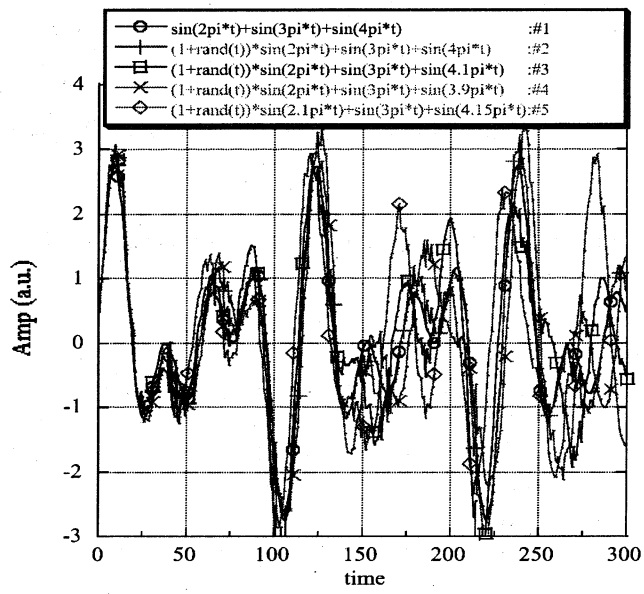

Figure 1: **Prepared time series data. Three sine waves are added each other.**

2.2 Procedure of time series signal analysis

This paragraph outlines the analysis of time-series signal procedure which uses the SOM.

The data preparation First of all, the signal data made by the computer is prepared (Fig.1). These signals were made from the addition of three sine waves, $sin(2\pi t), sin(3\pi t), and sin(4\pi t)$. '#1' is a base waveform. In '#2', the Gauss noise was added to the amplitude of the $sin(2\pi t)$. '#3' and '#4' adds the Gauss noise to the amplitude, and changes the frequency of $sin(4\pi t)$. In addition, '#5' changes both frequencies of $sin(4\pi t)$ and $sin(2\pi t)$.

Fast Fourier Transform(FFT) The FFT is done to these data. Each spectrum data has three

spectral-peaks, as shown in Fig.2. And, each spectrum pattern is different little by little.

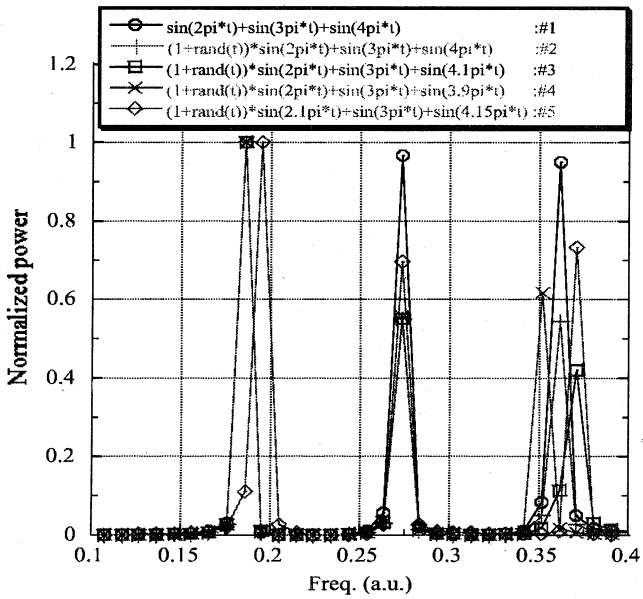

Figure 2: **Power spectrum from fast Fourier transform.**

These spectra are intentionally made from the waves which understand the frequency. Therefore, a slight difference of the pattern can be obviously confirmed from Fig.2. But, if these spectra were made from the actual measurement data, the analytical person might explain, "Considering the measurement noise, all spectra are the same characters."

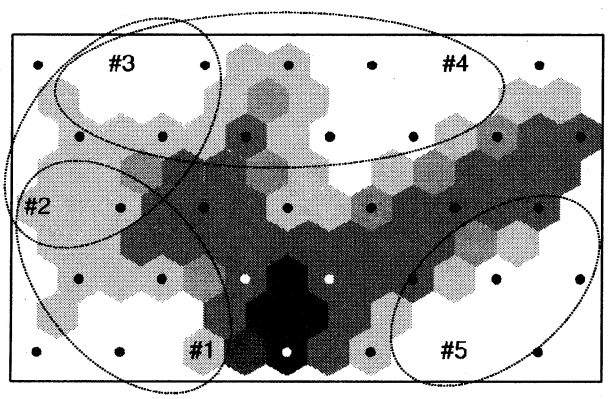

Figure 3: **Cluster analysis of pseudotime series data which use the SOM. It can be understood that '#5' is another group.**

Analysis using threSOM Finally, The FFT spectrum is analyzed with the SOM. The SOM map is shown in Fig.3. In the SOM, factors with high similarity are displayed in the same white region. In addition, if each distance is near it is a similar factor. From Fig.3, we can understand that '#1'~'#4' is a similar group.

However, '#5' is not in the same white area, so it is understood to belong to another group.

3 ANALYSIS OF UNSTABLE BEHAVIOR IN LINEAR ACCELERATOR

The SOM is used to analyze unstable behavior in the actual linear accelerator. The data collected from L-band LINAC at Osaka Universityis shown in Fig.4 [4]. The data of the RF phase is shown in the upper part, and the data of the disturbance factor measured simultaneously is shown in the lower side. Observing these data, we can't understand which factor contributes to the RF phase.

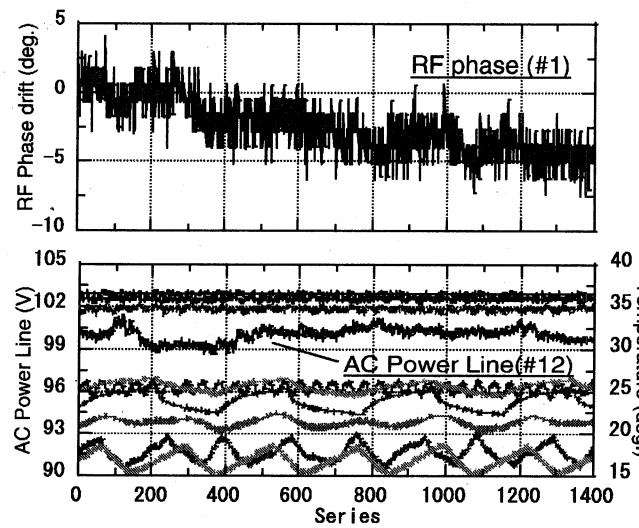

Figure 4: **Time series data collected in Osaka University L band-Linac. In the upper part, the RF phase data and the lower side are the disturbance factor data.**

To tell the truth, this data is being analyzed using the AR model by Konishi and Isaka[5]. And, they analyze that the AC Power Line contributes greatly in the RF phase.

Next, this data was analyzed by the SOM, Fig.5 was obtained. The RF phase and the AC Power Line are included in a similar group. Moreover, it is understood that other several kinds of fluctuation factor groups are formed. These results agree with the result of Konishi and Isaka.

4 CLUSTER ANALYSIS SERVER SYSTEM

If you spend many time to install the SOM analysis software or a lot of preprocessings are neces-

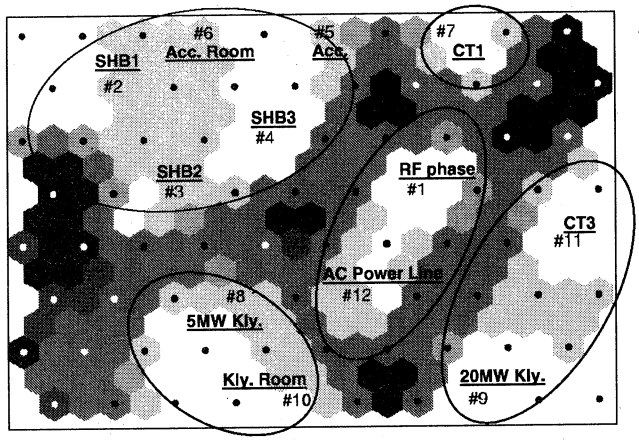

Figure 5: **SOM map of Osaka University's LINAC. The AC Power Line was included in the same group as the RF phase. This is corresponding with the result of obtaining because of the noise contribution analysis.**

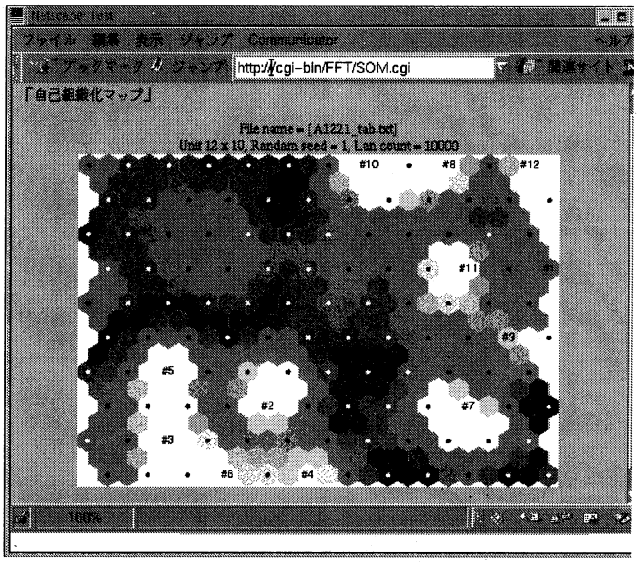

Figure 7: **The SOM map sent back from the cluster analysis server.**

sary for data analysis, the impression of this analysis might worsen. Then, we enabled you to use this analysis handily. In a word, the measurement data is transmitted to our Sever using World Wide Web browser, you can get the SOM map easily.

This usage is sequentially explained as follows. The top page shown with Fig.6 opens, when URL of the server is selected on the browser. You selects the data file which wants to be analyzed. SOM map, Fig.7, is sent back from our server only by pushing the transmission button. These operations are very simple.

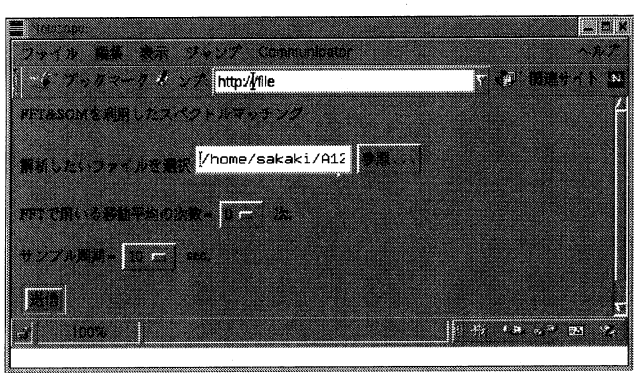

Figure 6: **Top page of cluster analysis server. To access, World Wide Web browser, Netscape, is used.**

5 CONCLUSION

The instability behavior cause in the accelerator was analyzed by the cluster analysis which used the SOM. The effectiveness was confirmed by comparing this and other analysis technics ; the AR model analysis.

Moreover, everyone from anywhere can easily do "Cluster analysis" by SOM using the World Wide Web browser for an analytical interface.

6 ACKNOWLEDGEMENT

In the SOM analysis in this work, we used the data which Mr. Konishi and Mr.Isaka at Osaka University had collected. We wish to express our gratitude.

7 REFERENCES

[1] H.Sakaki et al: "Statistical Analysis and Control of the Electron Beam Energy Fluctuation at Linear Accelerator", SICE, 35-10, 1283/1291 (1999).

[2] H.Sakaki et al.: Proc. of THE 25th LINEAR ACCELERATOR MEETING IN JAPAN, 261/263 (2000).

[3] T.Kohonen et al: "SELF-ORGANIZING MAPS", Springer-Verlag Japan (1996).

[4] T.Konishi et al: Osaka-univ. master's thesis (2001).

[5] S.Isaka et al.: Proc. of THE 26th LINEAR ACCELERATOR MEETING IN JAPAN, 104/106 (2001).