Paper GH-07

Categorizing the Degradation State of Aircraft Generators using Rank Order Statistics and SAS® CLUSTER Procedure

Tsung-hsun Tsai, Global Strategic Solutions, LLC, Sterling, VA

ABSTRACT

Maintaining an aircraft's electrical generator in operable condition is critical as failures in the power generator system can lead to catastrophic results. With a renewed emphasis on performing maintenance and repair before the problem arises, there is a strong push toward developing stringent health management technologies that can be realized in practice. This paper uses SAS® CLUSTER procedure and rank order statistics of symbolic sequence to assess the health state of Navy P-3 aircraft generators. By mapping the complex time series signal to binary sequence for encoding into abstract symbolic representation, the rank-frequency distribution is obtained by simple occurrence counting. The symbolic approach originated as a technique for analyzing hidden temporal structures in human cardiac dynamics. Two time series with similar patterns of fluctuations have similar probabilities and ranks of symbols. The objective of the study is to use the test data collected from Electrical Signature Analysis on five P-3 electrical power generators to estimate the differences of degradation state. Using a similarity measure and CLUSTER procedure, a hierarchical clustering tree can be constructed by using the TREE procedure from pair-wise measurements. The results uncover the difference of signal patterns of electrical power generators with quantitative information. Furthermore, the clustering tree not only discriminates patterns generated from different aging of the generators, but also reveals one of the generators that was used differently. A discussion of the potential for advancing health management techniques for aircraft generators is given.

INTRODUCTION

Electrical Power Generation System (EPGS), the power source for every subsystem, is a critical part on modern military aircrafts because premature failure of EPGS can lead to potentially serious consequences. Categorizing the degradation state of aircraft generators for early detection of EPGS degradation is an essential step toward the development of EPGS health management technology for preventing or recovering from impending in-flight failure. The effort may further improve operational efficiency by facilitating condition-based maintenance (CBM) and reducing unscheduled maintenance. It is worth noting that there are a number of integration issues and business factors to consider in applying any new technology to military platforms. Simple but effective methods are often sought for possible technology transfer in the future. This study, which is part of the research effort sponsored by the U.S. Naval Air Systems Command (NAVAIR), explores the use of rank order statistics to differentiate the status of system health among Navy P-3 aircraft generators. SAS® CLUSTER procedure is used for its simple visual interpretation in evaluating the degradation evidences of the state of health.

P-3 AIRCRAFT GENERATORS

The U.S. Navy P-3 Orion was introduced in the 1960s by Lockheed Martin as a land-based, long-range, anti-submarine warfare (ASW) patrol aircraft as a replacement for the aging P-2V Neptune (also by Lockheed). Many improvement programs were incorporated into its Updates during the past years to keep pace with increasing multi-mission requirements. Its mission in the early 21st century, such as the Operation Iraqi Freedom, has evolved to include surveillance of the battle space, either at sea or over land.

The power requirements are supplied by four interchangeable brushless air-cooled AC generators—three are engine driven and one is an auxiliary power unit (APU). The original generator that P-3 ASW patrol aircraft used was designed by Bendix. Allied Corporation (later named Allied Signal) acquired Bendix in 1983 and bought Honeywell Corporation in 1999. The combined company adopted the name “Honeywell” because of its superior brand recognition and Honeywell became the primary Original Equipment Manufacturer for Navy.

The original electric generator supplies 3-phase power with the rated voltage 115 VAC, 60 kVA (20kVA/ph) at 400Hz. The nominal frequency ranges from 380 to 420 Hz and the operating speed 5700 to 6300 rpm. The generator has a 12-pole AC exciter and a three-phase, half-wave diode rectifier rotating with the exciter armature. A single-phase permanent magnet generator provides control voltage and power for the voltage regulator. To ensure the P-3 remains a viable war fighter until Boeing P-8 Poseidon multi-mission maritime aircraft achieves full operational capability, many modernization programs were implemented during its production run and several improvements were made to satisfy Navy and joint requirements. Functional upgrade and replacement were made to the current 60 kVA generator for some aircrafts to provide additional power output for the modern subsystems.

DATA DESCRIPTION

The data sets used in this study were provided by NAVAIR at Patuxent River, MD. NAVAIR recorded the data using Electrical Signature Analysis (ESA) test performed on five P-3 aircraft generators with three different running loads (0, 30, and 60 kVA) and four frequencies (395 Hz, 397.7 Hz, 400 Hz, and 405 Hz). Among the five generators, there is one brand new generator, one high hour generator, one APU generator, and two other low usage generators. The signals were acquired at high sampling rates at 100 kHz in the ESA test because one of NAVAIR's original goals is to examine the quality of signals. The five generators were all considered as "good" generators from NAVAIR's view point because they passed the ESA test. The purpose of this study is to categorize the in-service degradation among them.

Table 1: P-3 aircraft generators usage data.

1	73-A0255	Low usage generator	n/a
2	1182	APU generator	2131 hours
3	18700071	Brand new generator	0
4	700	High hour generator	2324 hours
5	765	Low usage generator	n/a

NAVAIR collected both vibration and electrical data. Only the electrical parameter data were used in this study. The electrical data sets consist of 14 variables. Each set contains about 10 seconds of signals. The data were stored into data files based on the combination of frequency and load condition. After consulting with some generator manufacturers, the exciter field current was selected for this analysis. We sliced 5-second sample from data sets of 60 kVA load and 400 Hz output frequency. It should be noted that NAVAIR provided the usage data as shown in table 1 after this study was finished. Besides, NAVAIR also provided some new test data from a recent enduring test performed on a similar P-3 generator. Likewise, the sampling rate is 100k Hz and the data size is 10 seconds, but the load varies and the details are unknown. NAVAIR performed the enduring test for certain purpose and shared the data with us. Even though the data sets were not collected for this research, the enduring test data were included in the analysis. Some data sets with 70 kVA load condition from the available data files were used in place of 60 kVA load condition. The data sets we used are not consecutive number (see table 2). We used a macro to retrieve the data files from a list.

Table 2: P-3 aircraft generators data sets.

1	73-A0255	3, 6, 9, 12
2	1182	15, 18, 21, 24
3	18700071	30, 33, 36, 39
4	700	48, 51, 54, 57
5	765	109, 112, 115, 118

```
*File IDs;
%let ifid=3 6 9 12 15 18 21 24 30 33 36 39 48 51 54 57 109 112 115 118 ;
%let nW=%eval(%sysfunc(count(%cmpres(&ifid),%str( )))+1);
%let ifname=C:\SESUG2011\gen\data\fdIe;  *path and file name;
%let ifext=txt;                          *file extension;
*Read frequency distribution files;
%macro rFd(fid=,fname=,fext=);
  %local i j k;
  %let j=1; %let k=1;
  %do %while(%scan(&fid,&j) ne);
    %let i=%scan(&fid,&j);
      data fd&k;
        %let inf=&fname&i..&fext;
        infile "&inf";
        input cnt;
      run;
    %let j=%eval(&j+1); %let k=%eval(&k+1);
  %end;
%mend rFd;
%rFd(fid=&ifid,fname=&ifname,fext=&ifext);
```

RANK ORDER STATISTICS

The rank order statistics method borrowed the modeling idea from symbolic dynamics that representing a dynamic system by sequences of abstract symbols. Originally, the method was proposed to study dynamical signature of physiologic signals. Consider a time series of length *n*

$$S_i = [x_1, x_2, \cdots, x_n].$$

The first step is discretization, which maps the complex time series signal to binary sequence of symbols 0 and 1, respectively:

$$I_n = \begin{cases} 0, & x_n \le x_{n-1} \\ 1, & x_n > x_{n-1} \end{cases}$$

The following step is a symbolic representation that maps successive binary sequence into *m*-bit word by shifting one point at a time. This data-driven technique does not demand huge amount of data storage and intensive data process resources because discretization and bit shifting can be implemented in hardware, symbolic representation and occurrence counting are simple operations, and only the occurrences of different words are stored for later analysis. Sorting the occurrences of different words, the rank-frequency distribution represents the characteristics of the original time series. A similarity measure between two signals is defined as follows based on the rank-frequency distribution:

$$D_m(S_1, S_2) = \frac{\sum_k \left| R_1(w_k) - R_2(w_k) \right| p_1(w_k) p_2(w_k)}{(2^m - 1) \sum_k p_1(w_k) p_2(w_k)}$$

where p and R represent probability and rank of a specific word, w_k, using *m*-bit word in time series S. Probability and rank can be computed from a number of approach. Below is a macro that uses PROC SQL with inner sub-query to compute probabilities and uses PROC RANK to obtain ranks.

```
* Compute probability and rank of words;
%macro cPR(fd=,pr=);
  proc sql;
  create table &pr as
    select
      int(cnt) as cnt,
      calculated cnt/scnt as P
      from &fd,
      (select sum(cnt) as scnt from &fd);
  quit;
  proc rank data=&pr out=&pr ties=high; var cnt; ranks R; run;
  data &pr; set &pr(drop=cnt); run;
%mend cPR;
%macro caPR; %do i=1 %to &nW; %cPR(fd=fd&i,pr=pr&i); %end; %mend caPR;
%caPR;
```

HIERARCHICAL CLUSTERING

PROC CLUSTER and PROC TREE were used in conjunction to handle the similarity measure of the signals. By default, PROC CLUSTER interprets the input data set as coordinates in Euclidean space and computes Euclidean distance. To replace a customized distance measure computed by ourselves instead, we need to create a special SAS® data set. The input data set for PROC CLUSTER to use directly can be created by arranging the similarity measures for each pair of time series in a matrix form and specifying the `type=distance` in parentheses after the data set name in the DATA statement. We used PROC IML to compute the similarity measure and arrange distance matrix.

```
*Compute similarity measures;
%macro cSim(dM=,dsID=,nDs=);
proc iml; t=j(&nDs,&nDs,0); store t; quit;
%do i=1 %to &nDs-1;
  %do j=&i+1 %to &nDs;
    data pri; set pr&i; run;
    data prj; set pr&j; run;
    proc iml;
      load t;
      use pri; read all var _num_ into s1;
      use prj; read all var _num_ into s2;
```

```
        p1=s1[,1]; r1=s1[,2];
        p2=s2[,1]; r2=s2[,2];
        Dd=p1#p2;
        Dn=abs(r1-r2)#Dd;
        D8=Dn[+]/Dd[+];
        t[&j,&i]=D8;
        store t;
      quit;
    %end;
  %end;
  proc iml; load t; create dM from t;  append from t; quit;
  data &dM(type=distance); set dM; &dsID=scan("&ifid", _n_); run;
  %mend cSim;
  %cSim(dM=p31,dsID=dsName,nDs=&nW);
```

PROC CLUSTER will interpret the input data set to be clustered as a distance matrix. The output data set from is then rendered by the following TREE procedure to generate a dendrogram. In PROC CLUSTER, the analyst can select from a number of different mathematical methods to hierarchically cluster observations in the data set and provide related statistics to indicate the number of clusters. The code segment below uses single linkage method. We found that all of the methods suggested similar division of the clusters for 60 kVA cases.

```
  goptions htext=0.15in htitle=0.15in;
  proc cluster data=p31 outtree=tree method=sin; id dsName; run;
  title2 'METHOD: Single Linkage' ;
  proc tree; id dsName; run;
```

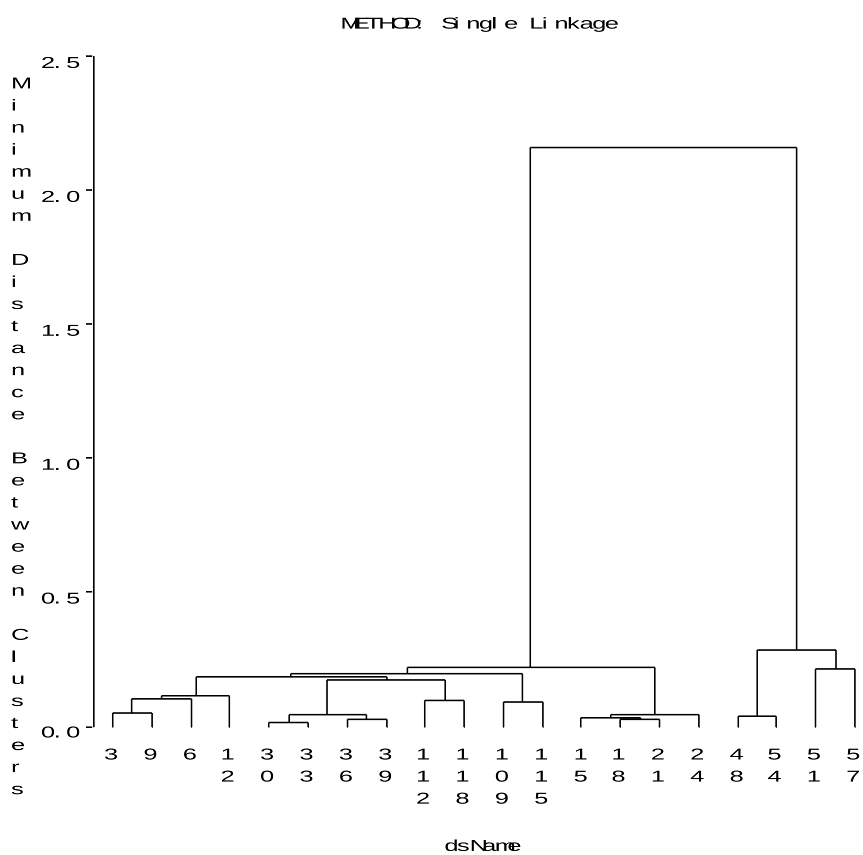

Figure 1: Dendrogram of the data sets of five generators

KEY OBSERVATIONS AND REMARKS

Figure 1 shows the arrangement of clusters using the data sets of the five P-3 generators. It is obvious that the high hour generator data sets form a distinct cluster. When using the signal without load, we were not able to separate the generators with the same approach. We concluded that early degradation signature can be seen only during excitations (i.e., applied load).

Visual inspection and related statistics suggest there seems to be reasonable separation of APU and the rest generators even though the generator is an identical product as the other four engine driven generators. NAVAIR provided the usage hours of the brand new, high hour, and APU generators later. In retrospect, the usage hours of the high hour and APU generator are comparable. One explanation might be that when a generator is used differently, it may have a different degradation signature. The APU is a critical support component for aircraft systems that made up of a turbine compressor driving an APU generator. It is used in an aircraft to start the engines and provides electrical power while the aircraft is on the ground, such as ground air conditioning and other pre-flight checks. Generally it will not be used with full power. Figure 2 is the dendrogram with the enduring test data added.

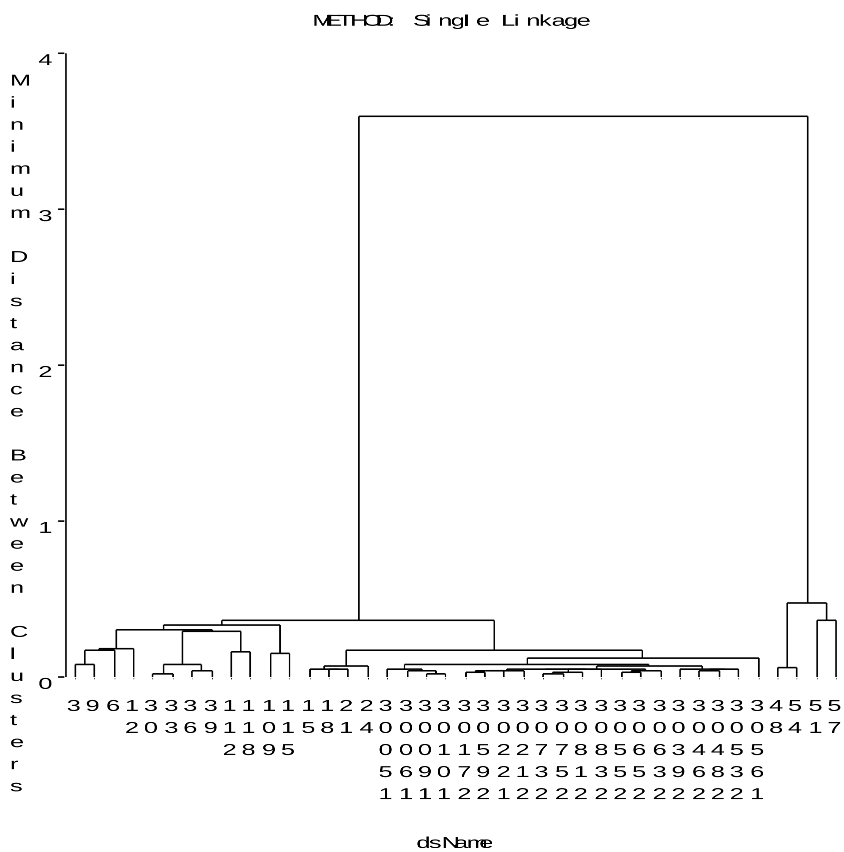

Figure 2: Dendrogram of the data sets of five generators and data sets from enduring test

CONCLUSION

P-3 long range ASW patrol aircraft is one of the key Navy aircrafts. The availability and performance are critical. During the past years, numerous modifications and improvement have been made to the P-3 updates to ensure provisional acceptance for service. Extra electric power is needed in order to operate the new subsystems effectively. With the increasing multi-mission requirements, the original reliability data are no longer valid because generators were undergoing more frequent emergency operations and intermitted system overload. The decline in the performance of the aircraft's power system is primarily due to continuing degradation in availability and reliability of the generators. Because the health of P-3 aircrafts is affected by their severe use profiles over a wide range of application environment, the CBM concept has become a paradigm shift in how the Department of Defense executes aircraft EPGS sustainment.

CBM, in essence, is a broad-based maintenance concept, which is intended to reduce maintenance down time by taking corrective actions based on the actual condition of the equipment, as opposed to other traditional (run-to-failure or time-based) maintenance concepts. Quantifiable degradation evidence is at the forefront of CBM before time reference degradation data can be collected. In this study, we attempted to categorize different aging and degradation state of the generators using exciter field current data of Navy P-3 ASW patrol aircraft collected from ESA testing. The proposed data-driven approach revealed underlying dynamics of the signal using modeling method inspired from nonlinear dynamics. We introduced clustering techniques and dendrogram in SAS/STAT® to reveal the structure of hidden health state, which reflects on the rank order statistics of a signal. The method is mathematically simple and tractable, easy to implement, and only requires very minimum computing and storage resources, which address one of the critical concerns in implementing health management solutions in aircraft EPGS. Data used in this approach are currently observable data—no need to re-certify any additional sensor. The technique may be used to examine the quality of refurbished or overhauled generators. Another attractive feature is that the data can easily be analyzed algebraically or graphically. The method for constructing dendrogram is useful in presenting the concept because it can easily be understood by the end users which is an important factor for possible technology transfer in the future.

REFERENCES

Kraskov, A., Nadler, W., Stogbauer, H, and Grassberger, P. (2004). Comment. Phys, Rev. Lett. 92(10), 109801.

SAS Institute Inc., *SAS®/Stat 9.3 User's Guide.* http://support.sas.com/documentation/onlinedoc/stat/930/statug.pdf

Smith, T., Lyle, D., Oliver, S., & Millett, b. (2003). USAF condition-based maintenance plus (CBM+) initative (AFLMA Report LM 200301800). Gunter Annex, AL: Air Force Logistics Management Agency.

Yang, A.C.-C., Hseu, S.-S., Yien, H.-W., Goldberger, A.L., and Peng, C.-K. (2003). Linguistic analysis of the human heartbeat using frequency and rank order statistics. *Phys. Rev. Lett.*, 90(10), 108103.

Yang, A.C.-C., Hseu, S.-S., Yien, H.-W., Goldberger, A.L., and Peng, C.-K. (2004). Comment. *Phys. Rev. Lett.*, 92(10), 109802.

ACKNOWLEDGMENTS

The author would like to thank Sean Field and Nathan Kumbar for their sponsorship of this research under NAVAIR contract number N68335-08-C-0112 and Charles Singer for his technical assistance.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Tsung-hsun Tsai
Global Strategic Solutions, LLC
22375 Broderick Drive
Sterling, VA 20166
Work Phone: 703-466-0500
E-mail: ttsai@gssllc.net
Web: www.gssllc.net

The views expressed in this publication do not necessarily reflect the views of NAVAIR or Global Strategic Solutions, LLC.