Prognosis of Electrical Faults in Permanent Magnet AC Machines using the Hidden Markov Model

IECON10

Syed Sajjad Haider Zaidi, Wesley G. Zanardelli, Selin Aviyente, Elias G. Strangas

IECON10

November 2010

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE
23 NOV 2010

2. REPORT TYPE
Briefing Charts

3. DATES COVERED
09-03-2010 to 21-10-2010

4. TITLE AND SUBTITLE
Prognosis of Electrical Faults in Permanent Magnet AC Machines using the Hidden Markov Model

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Wesley Zanardelli; Selin Aviyente; Elias Strangas; Syed Sajjad Haider Zaidi

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army TARDEC,6501 East Eleven Mile Rd,Warren,Mi,48397-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
#21410

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army TARDEC, 6501 East Eleven Mile Rd, Warren, Mi, 48397-5000

10. SPONSOR/MONITOR'S ACRONYM(S)
TARDEC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
#21410

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
For 36TH ANNUAL IEEE INDUCTRIAL ELECTRONICS CONFERENCE (IECON 2010)

14. ABSTRACT
-Failure prognosis and prediction of future state of operation is important to ensure continued operation and to exercise condition base maintenance. -Most of the work in machines is focused on the diagnosis instead of prognosis. But time to failure, or remaining useful life is important. -Generally prognosis needs large historic data sets to extract fault progression trends, which are not available in most of the cases.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Public Release**	**23**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Overview

Introduction

- Failure prognosis and prediction of future state of operation is important to ensure continued operation and to exercise condition base maintenance.
- Most of the work in machines is focused on the diagnosis instead of prognosis. But time to failure, or remaining useful life is important.
- Generally prognosis needs large historic data sets to extract fault progression trends, which are not available in most of the cases.

The problem under study

- A method for prognosis of electrical failures in a PM synchronous motor is presented.
- Transient increased contact resistance faults are investigated. Similar results for turn-turn and turn to frame faults.
- Objective: Determine the probability of failure at the next step.

Introduction

Methodology

- Fault prognosis is the next step of diagnosis and diagnosis information forms the basis of prognosis techniques
- the $q-$ axis current is used to extract information about faults
- Features of fault characteristics are extracted from Time Frequency distributions
- Diagnosis - Linear Discriminant Classifier (LDC)
  - Training of classifier for discrete fault states
  - Classification of test samples
- Prognosis - Hidden Markov Model (HMM)
  - Training/defining of HMM model parameters
    - Calculating State dependent observation probability (B)
    - Calculation of State Transition probabilities (A)
    - Defining the initial state probabilities (π)
  - Prediction of failure state probability

Type of fault and experimental setup

- Samples were time frequency features of the current i_q,
- Artificial faults were imposed. These are transient faults representing increased contact resistance, of fixed value and of fixed duration.
- A fault is identified by recognizing both inception and clearing, but only the inception transient is used to determine fault severity.

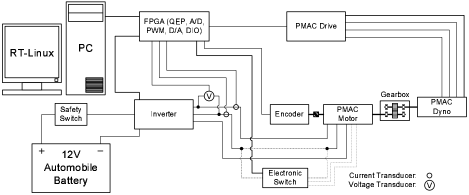


Figure: Experimental Setup

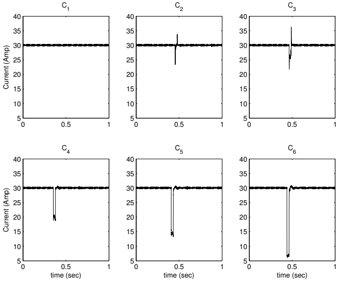


Figure: Sampled I_q Current

Machine and fault characteristics

- Specification of the test machine:

No of poles	6
Construction type	Surface mounted PMAC machine
Rated voltage	12V
Usage	Automotive application
Rated power	1HP
No load speed	3000rpm

- A custom FPGA-based I/O board was used as the drive controller. Its response was much slower than the fault dynamics
- The sampling frequency was 16.67kHz.

Application of Proposed Method

- The fault studied was an intermittently increased contact resistance, a transient electrical fault in the winding of the PMSM machine.
- The fault was created in the lab by inserting a resistor in series with a terminal in the winding to simulate a momentary breaking of the contact.
- The severity of the fault was primarily defined by the value of the inserted resistance.
- Resistances of $2.14pu$, $2.80pu$, $4.03pu$, $6.33pu$, and $15.84pu$ were inserted to mimic a progressively worsening fault.
- The transient faults have two stages, inception and removal. Both can be classified as separate events by the analysis of the time frequency features.
- In this work, the fault features were extracted from the inception event.

Feature Extraction - Three methods

Un-decimated Wavelet Transform (UDWT)

- UDWT has greater flexibility compared to STFT
- Different base functions can be used
- Good time resolution and high frequency resolution
- Tiling is variable

Wigner Ville Distribution

Defined as

$$W(t,\omega) = \int s(t+\frac{\tau}{2})s^*(t-\frac{\tau}{2})e^{-j\omega\tau}d\tau.$$

where s is input signal, t is the time, ω is the frequency

- High time-frequency resolution
- No tradeoff problem between time and frequency
- The major shortcoming is of multi-component signals in terms of the cross-terms

Choi Williams Transform

Defined as

$$C(t,\omega) = \int\int\int \varphi(\theta,\tau)\, s(u+\frac{\tau}{2})\, s^*(u-\frac{\tau}{2})e^{j(\theta u-\theta t-\tau\omega)}du d\theta d\tau, \quad \varphi(\theta,\tau) = \exp\left(-\frac{(\theta\tau)^2}{\sigma}\right)$$

where s is input signal, t is the time, ω is the frequency, σ is the smoothing parameter

- Filtered/smoothed version of the Wigner distribution
- Amount of smoothing is controlled by σ
- Smoothing comes with a tradeoff of reduced resolution

Classification

- Linear Discriminant Classifier (LDC) is used
- The discriminant function is defined as:

$$D_k(x) = x_1\alpha_{1k} + x_2\alpha_{2k} + ... + x_N\alpha_{Nk} + x_{1+N}\alpha_{1+Nk}$$

  where α_{ik} are the trained coefficients for k_{th} class, x is the features vector used for training.
- A sample vector of coefficients belongs to a particular class if the discriminant function is greater for that class than for any other class i.e. x belongs to class C_j if

$$D_j(x) > D_k(x) \quad , for \quad k \neq j$$

Prognosticator - Hidden Markov Model

- A statistical modeling method which assumes states to be Markovian.
- States here correspond to discrete levels of fault severity.
- Finds the hidden variables from the observable parameters.
  - Observable Parameters - Features extracted form sampled signals.
  - Hidden Parameters - Machine States
- The model parameters
  - State transition probability matrix ($a_{ij} = p(x_{t+1} = j | x_t = i)$)
  - State-dependent observation density ($b_j(y_t) = p(y_t | x_t = j)$)
  - Initial state probability ($\pi_i = p(x_1 = i)$)
- Problem to be solved
  - Given the observation sequence $y = (y_1, \cdots, y_k)$ and set of model parameters, choose corresponding state sequence $x = (x_1, \cdots, x_k)$ which is optimal to have generated the observation sequence.

Prediction Algorithm

- The normalized forward probability, $\delta_t(i)$, at time t for each state S_i, and the state transition probabilities, a_{ij} are used to predict the probabilities of states at time $t+1$. The transition probability to state S_j at the time instance t+1 is given by

$$\mathrm{P[q_{t+1}= S_j|\lambda] = \sum_{i=1}^{j} P[q_t= S_i|\lambda]a_{ij} = \sum_{i=1}^{j} \delta_t(i)a_{ij}}$$

  where λ is the set of model parameters.

- The most probable state at time $t+1$ is one which has the highest probability
- The predicted state probabilities are updated using state dependent observation $b_t(j)$ at each time step. The algorithm works as follows:

Prediction Algorithm

- Initialize

$$\delta_1(i) = \pi_i b_i(O_1) \;\; 1 \leq i \leq N$$
$$q_1(i) = 0 \;\; 1 \leq i \leq N$$

- Recursion

$$q_t(j) = \arg\max \sum_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] \quad 2 \leq t \leq T, \quad 1 \leq j \leq N$$

$$\delta_t(j) = \sum_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t) \quad 2 \leq t \leq T, \quad 1 \leq j \leq N$$

State Dependent Observation Distribution **B**

Experimental method to obtain the distributions:

- The distributions of the projections of the observations are assumed to be Gaussian

$$P(O/S_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(O-\mu_{O/S_i})^2}{2\sigma_{O/S_i}^2}\right)$$

- Training phase gives a set of LDC coefficients corresponding to each class (called LDC plane)
- Training samples are projected on all LDC planes
- Mean and variances of projections are used to define the state dependent observation distribution

UNCLAS: Dist A. Approved for public release

State Transition ($\mathbf{A}$) and Initial State (π) Probabilities

Both should result from extensive testing and aging/fatigue models.

- The state transition probabilities are computed using a heuristic method
- The probabilities are computed from matching pursuit decomposition of the sampled data

- The initial state probabilities should be
  - From the manufacturer
  - Repair facilities
  - Large scale sample analysis
- Assumed values are used for the demonstration
- Prognosis algorithm was tested with an assumption that the machine already has traces of fault
- The initial probability of class 2 is set high as compared to others.

State Dependent Observation Density Statistics - UDWT

- Means of the projection on each plane

	P_1	P_2	P_3	P_4	P_5	P_6
1	**0.6263**	0.8329	1.6662	1.9190	3.0788	4.43410
2	-0.0668	**14.2365**	17.0266	20.4563	28.99.79	43.0349
3	-0.0177	14.0670	**17.3373**	20.5170	29.3376	43.7851
4	-0.0268	13.2008	16.0605	**22.1399**	32.5496	48.0109
5	-0.0034	12.5658	15.4205	21.8770	**32.9591**	48.6149
6	-0.0110	12.5841	15.5384	21.8047	32.8978	**48.6836**

- Variances of the projections on LDC planes

	P_1	P_2	P_3	P_4	P_5	P_6
1	0.2188	0.3748	0.6602	0.2251	0.1785	0.2058
2	0.0031	0.0241	0.0176	0.1725	0.0470	0.0425
3	0.0072	0.0812	0.0228	0.4906	0.0600	0.0500
4	0.0066	0.0299	0.0251	0.0873	0.0269	0.0583
5	0.0094	0.0168	0.0328	0.2754	0.0238	0.0626
6	0.0086	0.0302	0.0278	0.4548	0.0201	0.0584

  - C is the class, P is the LDC plane

A State Dependent Observation Density Statistics- WVD

- Means of the projection on each plane

	P_1	P_2	P_3	P_4	P_5	P_6
1	**6.8544**	5.8337	5.0972	3.6628	2.3158	1.1945
2	6.8033	**5.8776**	5.1430	3.7109	2.3900	1.3013
3	6.7765	5.8629	**5.1559**	3.7687	2.4843	1.4219
4	6.5098	5.6555	5.0384	**3.8567**	2.7262	1.7653
5	5.5336	4.8971	4.4652	3.6653	**2.8685**	2.1619
6	3.4635	3.2352	3.1011	2.8799	2.6232	**2.3641**

- Variances of the projections on LDC planes

	P_1	P_2	P_3	P_4	P_5	P_6
1	0.0286	0.0458	0.2820	4.7961	0.2259	0.0182
2	0.0284	0.0521	0.3000	4.6780	0.2371	0.0162
3	0.0268	0.0515	0.2849	4.5926	0.2294	0.0152
4	0.0233	0.1134	0.2013	4.7297	0.1701	0.0120
5	0.0137	0.2317	0.0988	3.5576	0.0948	0.0069
6	0.0033	0.2408	0.0153	1.2946	0.0203	0.0039

  - C is the class, P is the LDC plane

State Dependent Observation Density Statistics- CWD

- Means of the projection on each plane

	P_1	P_2	P_3	P_4	P_5	P_6
1	**6.8545**	5.9820	5.2338	3.8745	2.6148	1.5756
2	6.8426	**5.9924**	5.2388	3.8752	2.6176	1.5820
3	6.8364	5.9823	**5.2477**	3.9147	2.6789	1.6580
4	6.7027	5.8607	5.1847	**3.9623**	2.8096	1.8467
5	6.2210	5.4443	4.8794	3.8640	**2.8811**	2.0450
6	5.0197	4.4060	4.0439	3.4009	2.7384	**2.1516**

- Variances of the projections on LDC planes

	P_1	P_2	P_3	P_4	P_5	P_6
1	0.0218	0.0554	0.2132	3.5857	0.1953	0.0143
2	0.0215	0.0584	0.2179	3.5532	0.1981	0.0140
3	0.0212	0.0522	0.2066	3.5301	0.1898	0.0136
4	0.0197	0.0427	0.1680	3.4105	0.1592	0.0120
5	0.0160	0.0502	0.1116	2.9207	0.1113	0.0089
6	0.0096	0.0891	0.0426	1.9052	0.0464	0.0046

  - C is the class, P is the LDC plane

Plots of the State Dependent Observation Density Satistics

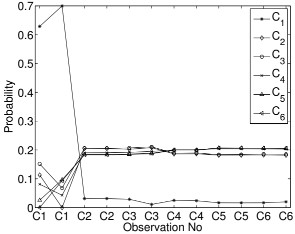


Figure: UDWT

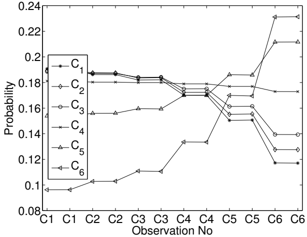


Figure: Wigner

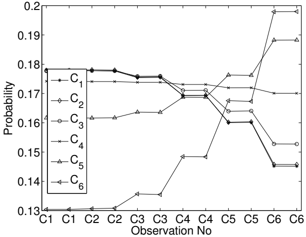


Figure: Choi-Williams

UNCLAS: Dist A. Approved for public release

HMM Parameter Values - State Transition Probability Matrix

- Sample mean of each class was calculated
- MP decomposition was performed
  - A greedy adaptive algorithm
  - Chooses the atom of the dictionary that best represents the signal
  - Decomposition was performed by Gabor dictionary
  - Uses Gaussian window for atom generation $g(t) = e^{-\frac{1}{2}t^2}$
  - 3905 normalized atoms were generated by time shifting, scaling, and modulation

$$g_\gamma(t) = (k_\gamma/\sqrt{s})g(t - \tau/s)cos(\xi t + \phi)$$

where s is scaling constant, τ is time shift, k is normalizing coefficient

- From the decomposed samples the state transition probabilities are calculated
- Only forward path probabilities were allowed as the fault analyzed is non reversible

Class	1	2	3	4	5	6
1	0.5063	0.2435	0.1481	0.0800	0.0200	0.0021
2	0	0.4935	0.2127	0.1651	0.0967	0.0320
3	0	0	0.4542	0.2709	0.1535	0.1214
4	0	0	0	0.4529	0.2900	0.2571
5	0	0	0	0	0.5800	0.4200
6	0	0	0	0	0	1.0000

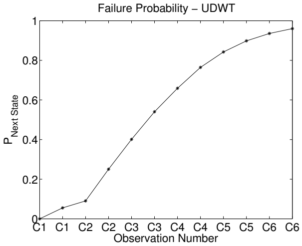
HMM Parameter Values - Assumed Initial State Probabilities (π)

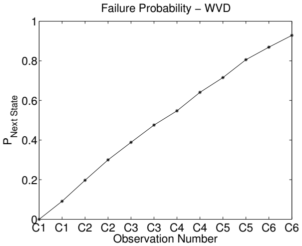
- Historic data, or manufacturer input was not available
- Assumed values were used to demonstrate the developed method

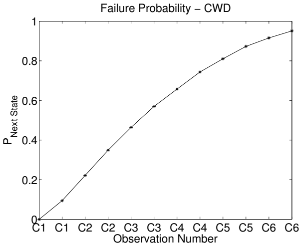
Class	1	2	3	4	5	6
Probability	0.07	0.60	0.15	0.1	0.08	0

Probability of the Failure State (C_6)

An artificial sequence of observations was constructed to test the method. These observations were actual sampled signals arranged in order of severity to mimic a natural fault progression.







UNCLAS: Dist A. Approved for public release

Results

- State Dependent Observation probabilities are higher if the projection are made on the corresponding plane
- The values of b_i were computed using UDWT have less discriminative value, except for the healthy class (C_1), in comparison with the the b_i obtained using the Wigner or Choi-Williams distributions
- Although for Wigner and Choi-Williams distributions the probabilities are close for the early fault severities, these are adequately discriminative for the high fault severities.
- The values of b_i computed using features extracted from Choi-Williams distribution are slightly more discriminative then the Wigner distribution, due to the smoothness provided by Choi-Williams distributions.
- The probability of failure, class 6, increases as the current state of fault severity increases, which is in accordance with the expected results.

Conclusions

- A fault prognosis method was developed and demonstrated based on the Hidden Markov Model
- Parameters of the HMM were calculated through a mixture of experiments and heuristic methods
- The HMM is used to estimate the most probable next state at every time step, as well as the probability of failure
- The prognosis algorithm produces similar failure probabilities for all three distributions used